THE PROPROTEIN CONVERTASES OF *Aedes aegypti:* IMPLICATIONS DURING FLAVIVIRUS AND ALPHAVIRUS INFECTION

by

Carlos Andres Brito Sierra

A Dissertation

Submitted to the Faculty of Purdue University

In Partial Fulfillment of the Requirements for the degree of

Doctor of Philosophy

Department of Biological Sciences

West Lafayette, Indiana

December 2021

THE PURDUE UNIVERSITY GRADUATE SCHOOL
STATEMENT OF COMMITTEE APPROVAL

Dr. Richard J. Kuhn, Chair

Department of Biological Sciences

Dr. Andrew D. Mesecar

Department of Biological Sciences

Dr. Douglas J. LaCount

Department of Medicinal Chemistry and Molecular Pharmacology

Dr. Robert V. Stahelin

Department of Medicinal Chemistry and Molecular Pharmacology

Approved by:

Dr. Janice Evans

and

Dr. Jason R. Canon

To my mother (Martha C. Brito Sierra) and my grandfather (Azael Brito Aristizabal)

All that I am is the result of your hard work and dedication to make me be a good world citizen.
I love you so much

Todo lo que soy es el resultado de su arduo trabajo y dedicación para hacerme un buen
ciudadano del mundo
Los amo mucho

ACKNOWLEDGMENTS

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS

AaFUR1	*Aedes aegypti* furin1
AaFUR2	*Aedes aegypti* furin2
Aag2	Cell line derived from *Aedes aegypti*
ACE2	Angiotensin Convertin Enzyme-2
Ad5	Adenovirus
ADE	Antibody dependent enhancement
AgFUR	Anopheles gambiae furin
AMC	7-Amino-4-Methylcoumarin
AmFUR	Apis mellifera furin
BEI	Biodefense and Emerging Infections resources
BHK	Baby hamster kidney cells
BiP	Binding immunoglobulin protein
BME	ß-mercapthoethanol
bp	Base pair
BSA	Bovine serum albumin
BSL3	Biosafety level 3
C	Capsid
C/EBPβ	CCAAT-enhancer-binding proteins - transcription factor
C6/36	Cell line derived from Aedes albopictus
CAD	Coronary artery disease
CHIKV	Chikungunya virus
CMV	Cytomegalovirus
CRISPr	Clustered Regularly Interspaced Short Palindromic Repeat
CRR	Cysteine Rich Region
Da	Daltons
DENV	Dengue virus
Dfurin1	*Drosophila* furin1
Dfurin2	*Drosophila* furin2

DMEM	Dulbecco's Modified Eagle Medium
DMSO	Dimethylsulfoxide
DNA	Deoxyribonucleic acid
DPP4	Dipeptidyl Peptidase 4
DsiRNA	siRNA substrate of Dicer
dsRED	Red fluorescent protein derived from Discosoma
E	Envelope
EDTA	Ethylenediaminetetraacetic acid
EEEV	Eastern equine encephalitis virus
EGTA	Ethylene glycol-bis(β-aminoethyl ether)-N,N,N′,N′-tetraacetic acid
FBS	Fetal Bovine Serum
FCA	Complete Freund's Adjuvant
FES	Human Proto-Oncogene homolog of feline sarcoma retrovirus
GFP	Green Fluorescent Protein
GMO	Genetically modified organism
HDPG	Heparan sulfate proteoglycans
HEK	Human embryonic kidney cells
HiBiT	Small tag for luminescense-based detection of proteins
HmFURIN	Human furin
HPAI	Higlhy Pathogenic Avian Influenza
IBS	Illustrator of biological Sequences)
IDT	Integrated DNA Technologies
IL	Interleukin
IRS	Indoor residual spraying
ITN	Insecticide treated nets
LVP	Liverpool strain of *Aedes aegypti*
MBD	Mosquito-borne disease
MEM	Minimum Essential Media
MEROPS	Online database for proteases and their inhibitors
MERS	Middle East Respiratory Syndrome
mL	Milliliter

Mmfurin	*Mus musculus* furin
MOI	Multiplicity of infection
Mpro	Main protease of ß-coronavirus
mRNA	Messenger RNA
Mt	Metallothionein
NC2	Neuroendocrine convertase 2
NCBI	National Center for Biotechnology Information
NEB	New England Biolabs
Ni-NTA	Nickel-Nitriloacetic acid
nsps	Non-structural proteins
ORF	Open reading frame
PACE	Paired amino acid converting enzyme
PC	Proprotein Convertase
PCR	Polymerase chain reaction
PCSK	Proprotein Convertases Subtilisin/Kexin
PE2	Precursor of E3 and E2 in alphaviruses
Pfam	Database that provides alignments and hidden Markov models for protein domains
PFU	Plaque forming units
pH	Power of hydrogen
prM	Pre-membrane protein
Puro	Puromycin
qRT-PCR	Quantitative Real Time PCR
RFU	Relative Fluorescent Units
RISC	RNA-induced silencing complex
RLU	Relative Light Units
RNA	Ribonucleic acid
RnRp	RNA-dependent RNA polymerase
RRV	Ross river virus
S2 cells	Cell line derived from *Drosophila melanogaster*
S8A	Subtilisin Carlsberg (*Bacillus licheniformis*)

S8B	Kexin (*Saccharomyces cerevisiae*)
SARS	Severe Acute Respiratory Syndrome
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel
SEC	Size Exclusion Chromatography
Sf9	Cells derived from *Spodoptera frugiperda*
SG	Salivary glands
SINV	Sindbis virus
siRNA	Small interfering RNA
SKI-1	Subtilisin Kexin Isozyme-1
SOSUI	Online tool that predicts a part of the secondary structure of proteins from a given amino acid sequence
SV40	Simian vacuolating virus 40
TGF	Tumor growth factor
TGN	Trans-Golgi Network
TM	Transmembrane region
TMHMM	Membrane protein topology prediction tool
UB	Ubiquitin
UTR	Untranslated region
VEEV	Venezuelan equine encephalitis virus
WEEV	Western equine encephalitis virus
WHO	World Health Organization
WNV	West Nile virus
ZIKV	Zika virus

ABSTRACT

In 1741, a British fleet of around 124 ships attacked Cartagena. The purpose of the siege was to gain control of the Spanish port and eventually use it to invade inner colonial lands. The siege involved invasion by sea and land and was at first successful. There was one remaining obstacle to win the victory, the San Felipe de Barajas Castle. This attack resulted more difficult as it involved mobilization of British troops into the jungle, where soldiers were exposed to mosquitoes. As the battle progressed, the British army was forced to retreat as they had lost thousands of men, the majority from yellow fever virus, leading to a Spanish victory. This is just an example of how mosquitoes can influence the outcome of history. Even in our days, we see how the mosquitoes can affect the way we live and the tremor they can cause with outbreaks like Zika, yellow fever, chikungunya, or dengue viruses. As such, it is important to be prepared and develop strategies that would harm the tight mosquito-virus relationships. For that reason, understanding the life cycle of these viruses in the mosquito would provide targets for disease control. One of the major steps in the life cycle of the virus is the maturation process, which heavily relies on the host proteases. The objective of this dissertation was to identify the mosquito proteases that are necessary for the maturation of flaviviruses and alphaviruses. Given that multiple viral families utilize these proteases, disrupting their function would prove harmful for different viruses at once. In the mosquito, *Aedes aegypti,* there are three proprotein convertases, named furin1, furin2 and NC2. These proteases retain high similarity with the human and *drosophila* homologs and were initially hypothesized to be involved in the viral maturation process. Recombinant expression and enzymatic activity assays of these proteins showed that furin1 exhibited activity comparable to human furin but significantly higher than furin2, whereas NC2 was not active under the experimental conditions. Further, CRISPr knockouts in mosquito Aag2 cells revealed that furin1 but not furin2 is required for efficient maturation flaviviruses and alphaviruses, as reduced proteolytic cleavage is also reflected in reduced viral titer. Combined, these experiments suggest that furin1 has a role in the proteolytic cleavage of different families of arboviruses in mosquitoes and is associated with their maturation and infectivity. Attacking the maturation process of these viruses during the infection of the mosquito will change the history of vector-borne diseases control.

CHAPTER 1.   THE PROPROTEIN CONVERTASES, PROCESSING OF VIRAL PROTEINS AND MOSQUITOES AS VECTORIAL AGENTS

1.1   Chapter Summary

Proteins undergo different types of post translational modifications that can be either reversible or irreversible. The irreversible modifications are performed by proteases that can process different types of substrates to generate active forms. The serine proteases are an extensive group of enzymes that cleave residues using a catalytic triad composed of serine, histidine and aspartic acid. Among the serine proteases are the proprotein convertases, which are evolutionary related to bacterial subtilisin and yeast kexin. These proteases perform multiple functions in the cells, such as processing of pro-hormones, growth factors and cellular receptors. However, the proprotein convertases can also be exploited by invading pathogens such as viruses. Furin is a member of the proprotein convertase family, and it has been extensively studied. Furin has the typical three domains of proprotein convertases: pro-segment, catalytic domain and P-domain, which work together during the regulated autoactivation and processing of substrates. Furin has been involved in the processing of viral substrates such as flaviviruses and alphaviruses. These viruses cause medically relevant illnesses and are primarily transmitted by mosquitoes. The efficiency at which the mosquitoes can transmit viruses depends on multiple aspects such as geographical distribution, vectorial competence and mosquito-virus interactions. However, little to no studies have addressed the furin-mediated processing of viruses in mosquitoes. Understanding the proprotein convertases of important insect vectors such as *Aedes aegypti* would provide targets for development of strategies to control and prevent transmission.

1.2   The serine protease superfamily

Post translational modifications constitute an important source of diversity among proteins. There are two types of modifications: reversible and irreversible. Glycosylation, sulphation, phosphorylation and palmitoylation are among the reversible modifications, which help in regulating the proper function of the cell by activation or inactivation of proteins upon specific signals. In contrast, proteolysis is an irreversible post-translational modification that results in products of different sizes with diverse biological functions (Figure 1.1).

19
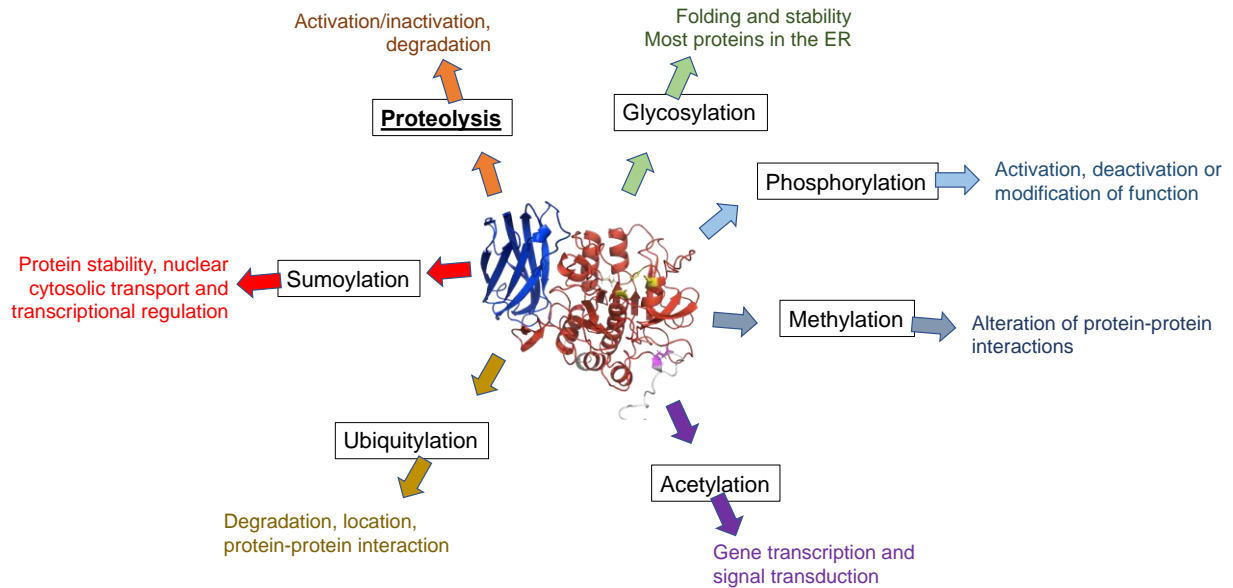
Figure 1.1 Post-translational modifications. Examples of common post-translational modifications. The proteolysis is an irreversible modification that can activate/inactivate substrates. Glycosylation, phosphorylation, methylation and acetylation coordinate the stability, activation/deactivation and protein-protein interactions. Ubiquitylation regulates the lifespan of protein as well as their interactions. Finally, sumoylation can regulate protein stability, transport and regulation of expression.

Proteolysis is carried out by proteases, also named peptidases, which comprise 641 and 677 coding genes in humans and mice, respectively [1]. There are different types of proteases which share similar mechanism of action but use different active amino acid residues to perform the proteolytic cleavage. According to the MEROPS Peptidase Database, the proteases can be divided into six major groups: Aspartic, Cysteine, Glutamic, Metallo, Asparagine, Threonine and Serine proteases [2], [3].
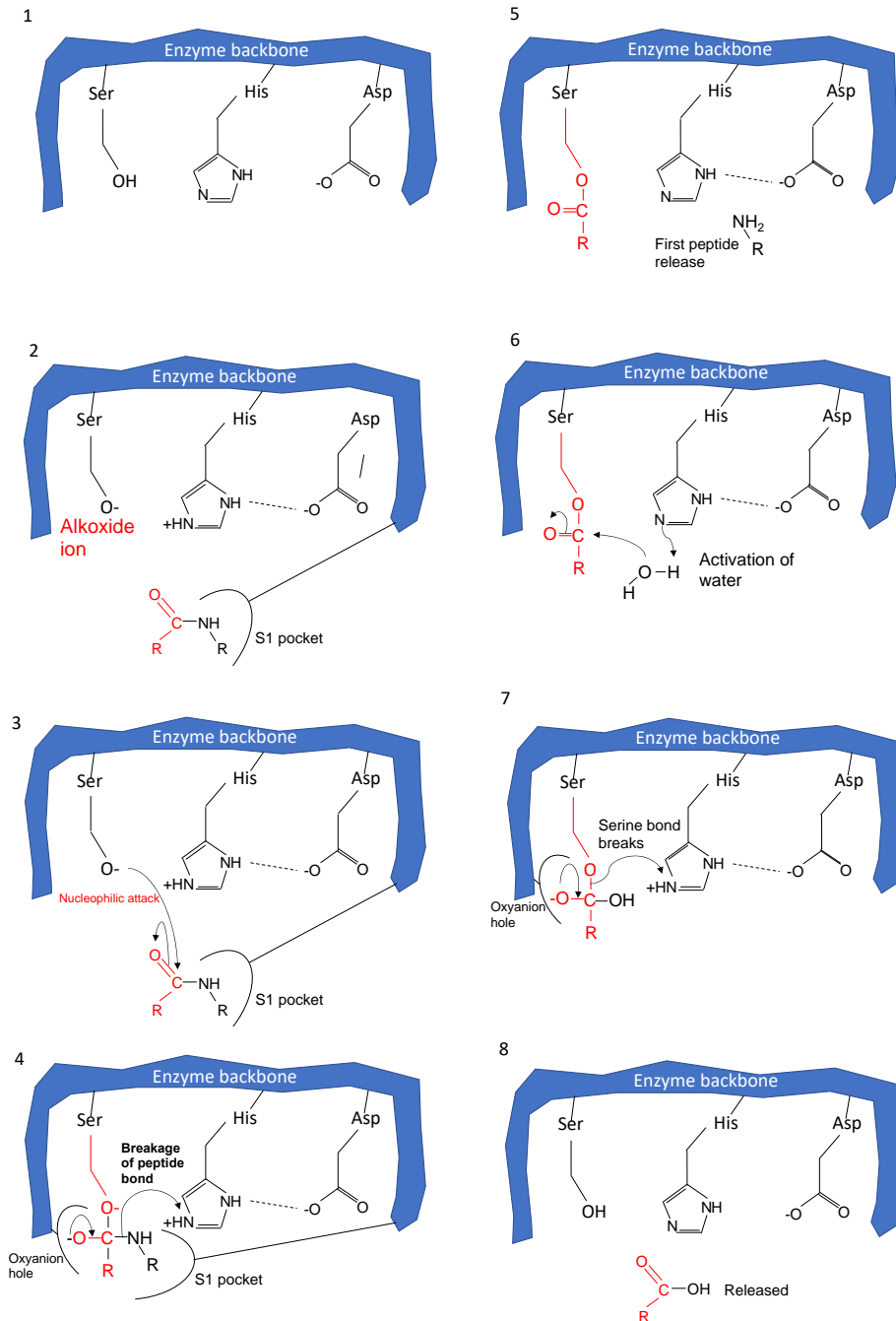
Figure 1.2 Mechanism of action of the Serine Proteases.
1) Catalytic triad. 2) Binding of substrate induces conformational change that results in the generation of alkoxide ion. 3) Alkoxide ion performs a nucleophilic attack on the carboxyl side of the cleavage site. 4) The oxyanion hole stabilizes the intermediate generated during the breakage of the peptide bond. 5) The first peptide gets released but the N-terminal peptide remains bound to the serine. 6) Histidine activates water. 7) Another nucleophilic attack creates an intermediate that results in the breakage of the serine bond. 8) Second peptide is released and the active site is restored.
Figure was adapted from
https://bio.libretexts.org/Bookshelves/Biochemistry/Book%3A_Biochemistry_Free_For_All_(Ahern_Rajagopal_and_Tan)/04%3A_Catalysis/4.03%3A_Mechanisms_of_Catalysis Accessed on July 13th, 2021

After metalloproteases, the serine proteases are the second most abundant group of proteases in humans, mice and *drosophila*. All serine proteases have in common a catalytic triad, composed of Aspartic acid, Histidine and Serine as well as a pocket called the oxyanion hole, which stabilizes the intermediate state during the catalytic cleavage [4]. However, the difference among the serine proteases resides in the binding pocket, which confers specificity for the different proteases. For example, the binding pocket of chymotrypsin is highly hydrophobic, whereas the pocket of trypsin is negatively charged [4], [5].

The catalytic mechanism of serine proteases is very conserved and occurs upon binding of the substrate to the binding pocket. This induces a conformational change of the enzyme that brings together the catalytic triad. The close proximity of the negatively charged aspartic acid to the protonated histidine, induces the abstraction of the proton present in the side chain of serine. This creates an alkoxide ion that attacks the carbonyl group of the substrate, breaking the peptide bond. This results in the release of part of the peptide, but retention of the peptide that is in the same chain as the carbonyl group that is now covalently attached to the serine. The second part occurs slower: a molecule of water is attacked by the histidine, creating a hydroxyl group that performs a nucleophilic attack on the carbonyl-serine bond. This breaks the bond, releasing the peptide and recovering the enzyme back to the original state [6], [7] (Figure 1.2).

The serine proteases are classified in 13 clans and 40 families, where each clan has a common catalytic mechanism, and each family has a common ancestor [8]. The clans can be grouped into chymotrypsin/trypsin-like and subtilases [9]. The subtilases are enzymes evolutionary related to the bacterial subtilisin, which share the same mechanism of action as chymotrypsin/trypsin but with an unrelated protein fold [10]. The subtilases are divided into two subfamilies, the S8A and S8B, which differ in the specificity of the substrate [11]. The S8B subtilases are related to the yeast kexin and have specificity for cleaving after dibasic residues [12]. Kexin, Kex2, is the type of the subfamily S8B, which also contains furin and the rest of the proprotein convertases. (Figure 1.3)

1.3 The repertoire of proprotein convertases of humans

Furin and the rest of the proprotein convertases are calcium dependent and contain three minimal defining features: 1) the pro-segment, which acts as an intramolecular chaperone and covalent inhibitor, 2) the catalytic domain, which contains the catalytic triad and has all the characteristics of the serine protease subtilisin, and 3) the P-domain which is required for stability and activity of the catalytic domain [9], [13]–[15]. These proteases are typically referred as PCSKs (Proprotein Convertases Subtilisin/Kexin) and there are total nine proprotein convertases in humans.

The first seven members are kexin-like subtilases, PC1, PC2, furin, PC4, PC5, PACE4 and PC7. They process substrates at basic residues in the motif $(R/K)X_n(R/K)\downarrow$ and are located in different parts of the secretory pathway, including the Golgi apparatus, secretory granules, plasma membrane, extracellular space and even in the endoplasmic reticulum (ER) [9], [14]. They are necessary for processing of multiple precursor proteins such as growth factors, hormones, adhesion molecules and other enzymes.
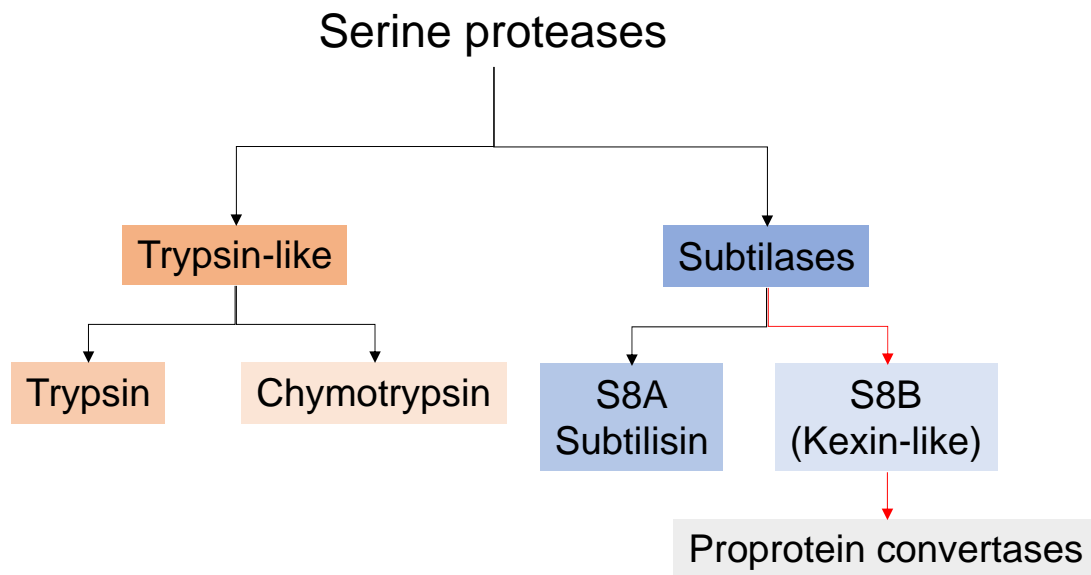
Figure 1.3 Classification of the Serine Proteases. The serine proteases are divided into two major groups: the trypsin-like and the subtilases. The trypsin-like has the trypsin and chymotrypsin. Subtilases have the S8A which is originated from *Bacillus licheniformis* and S8B, which contain all the kexin-like enzymes and the proprotein convertases.

The eighth proprotein convertase, termed as SKI-1, is a pyrolysin-like subtilisin that cleaves after non-basic residues in the motif RX(L/V/I)X↓ and is mostly accumulated in the Golgi apparatus. SKI-1 is involved in the processing of transcription factors and regulation cholesterol synthesis [14], [16], [17], as well several homeostatic processes such as bone mineralization and cellular stress response [9], [18]. The ninth proprotein convertase, PCSK9, is the only proprotein convertase that remains bound to the pro-segment, rendering it without enzymatic activity other than cleaving itself at the site $VFAQ_{152}$↓. Instead, PCSK9 retains its binding efficiency for its substrate, LDLR, targeting it for degradation in the lysosomes [14], [19].

1.3.1   Tissue distribution and biological roles

The differential distribution, both in tissues and subcellular locations, partially determines the substrates that each PCSK processes. PC1 and PC2 are mostly present in secretory granules of neural and endocrine cells, where they process pro-hormones under the acidic conditions of the secretory pathway [20], [21]. Furin is ubiquitously expressed but accumulates in the trans-Golgi network (TGN), but can also travel to the plasma membrane, be recycled or shed to the extracellular space [22], [23]. PC4 is exclusively expressed in testes, ovary and placenta [24]. PC5 is widely expressed and has two alternatively spliced products that can be membrane bound or secreted. PACE4 is also ubiquitously expressed and accumulates in the TGN [9], [25]. Both PACE4 and PC5 are also active at the plasma membranes where they bind heparan sulfate proteoglycans (HSPG). PC7 can transit directly from the ER to the plasma membrane, but like furin or PC5B, can be recycled to the TGN through endosomes. SKI-1 is also widely expressed and can be present mostly in the early compartments of the Golgi apparatus. PCSK9 is mostly expressed in the liver, kidney and small intestine, and can be secreted out of the cell with the attached pro-segment. When PCSK9 interacts with its substrate, it can enter the cell through endosomes and promote degradation of its target in the lysosomes [26].

1.4   Furin

Furin is the most widely studied proprotein convertase. To date, there are 2265 research articles on PubMed that have addressed furin, which is staggering when considering that it took nearly 20 years to discover furin for the first time [27]. Furin is ubiquitously expressed and more

than 100 substrates have been identified, including growth factors, cytokines, hormones, adhesion molecules, receptors, coagulation factors, metalloproteinases and albumin [28]. The wide spectrum of furin substrates is also explained by its localization in different parts of the secretory pathway. Furin accumulates mostly in the trans-Golgi network (TGN) [29], [30], but can be trafficked to the secretory granules, plasma membrane and be recycle through the endocytic pathway [31].

The importance of furin is evident by knock-out of mice that die at 11 days during the embryogenesis stage, associated with severe ventral closure defects and failure of the heart during morphogenesis [32]. Also, having mutations in the furin cleavage site of the substrates result in genetic disorders such as X-linked hypohidrotic ectodermal dysplasia [33]. In addition, changes in the expression of furin also have correlation with disease. For example, during coronary artery disease (CAD), there is an increase in furin expression that promotes migration of macrophages and prevents apoptosis, which results in increased risk of atherosclerosis [34].

1.4.1   The life cycle of furin

Furin, also known as PCSK3 or PACE, is expressed by the FUR gene (FES upstream region) in chromosome 15 and can be ubiquitously present, but with variation among different tissues [25]. There are at least three promoters (P1, P1A and P1B) that produce different transcripts of furin with an identical coding sequence[35]. P1A and P1B are constitutively promoters, whereas P1 can be activated by the transcription factor C/EBPβ, meaning that is sensitive to cytokine regulation, such as with TGFβ1 [36], [37] and IL-12 [38], [39]. Ironically, TGFβ1 activates the expression of its own processing enzyme.

Once the promoters are activated, and the furin mRNA is transcribed, the protein is initially synthesized as a proenzyme that enters the secretory pathway. Furin is inserted in the membrane of the endoplasmic reticulum through its C-terminal transmembrane region while the N-terminus signal peptide is removed co-translationally (Figure 1.4, step 1). Then, the pro-segment, which is required for correct folding of the catalytic domain, gets cleaved in a two cleavage events, utilizing the enzymatic rules of furin. First, the pro-segment is cleaved at Arg107 in a fast reaction (t/2=10 min) in the neutral pH of the ER in a localized Ca^{++} dependent manner (Figure 1.4, step 2). The

peptide remains bound to the protein and acts as an inhibitor. Then, once the protein reaches the acidic environment of the trans-Golgi network (TGN), the second cleavage occurs at Arg75, dependent on both Ca^{++} and pH. This results in disassociation of the pro-segment and activation of furin (Figure 1.4, step 3) [27], [40]. Simultaneously, during the transit of furin through the ER, furin gets glycosylated and subsequently sialylated in the Golgi apparatus [41]. Furin accumulates primarily in the TGN but can also move to the plasma membrane and back via the endosomal pathway (Figure 1.4, step 4). In addition, furin can also be shed out of the cell by removal of the transmembrane region (Figure 1.4), step 5) [23].

5- Shedding

4- Accumulation in the TGN

3- Removal of the **prosegment**

2- Cleavage of the **prosegment**

1- Removal of **signal peptide**
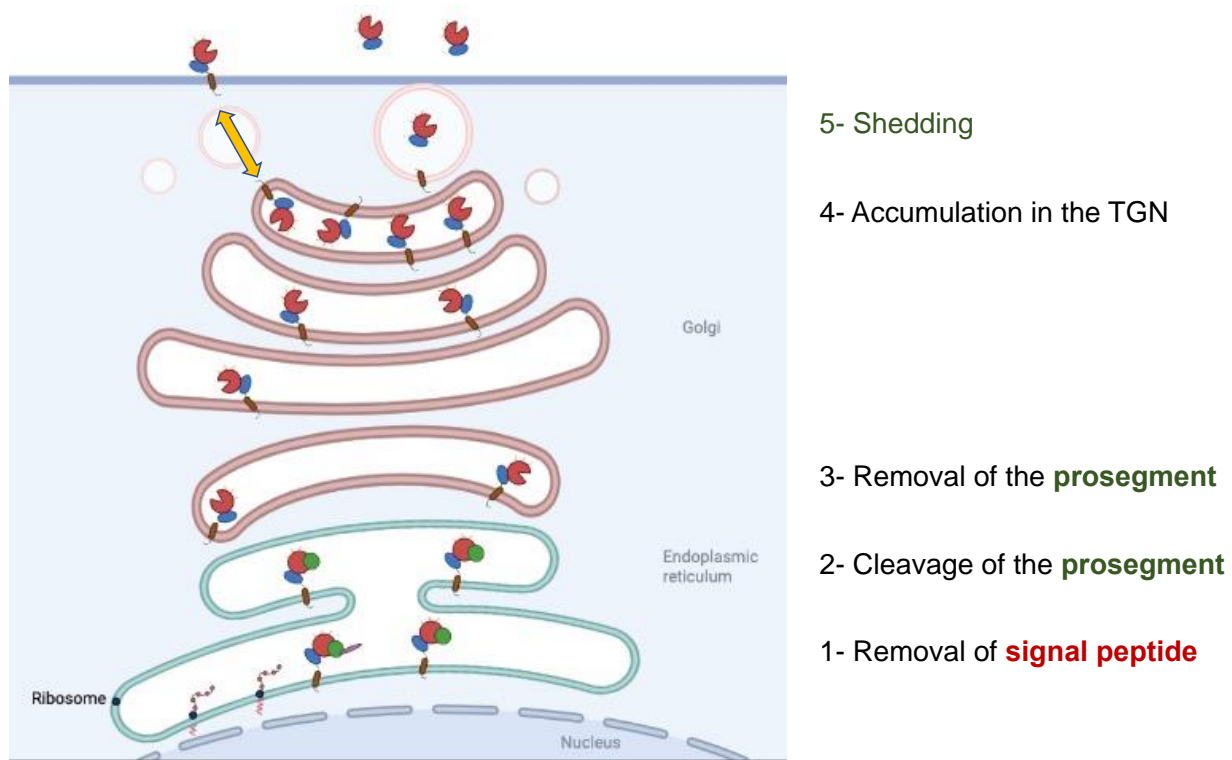
Figure 1.4 Life cycle of furin. Furin is initially translated and glycosylated in the cytoplasm and the signal peptide directs the protein into the secretory pathway. Once in there, the signal peptide is removed by host signalases. Then, an autocatalytic cleavage of the prosegment occurs at Arg107, but the peptide remains bound as an inhibitor. Once it reaches the acidic pH of the Golgi apparatus, a second cleavage occurs at Arg75. This renders the enzyme active and it accumulates in the trans-Golgi network (TGN). From there, the enzyme can be transported to the plasma membrane and be recycled or shed out of the cell.

1.4.2 Role of furin in cancer progression

Furin is known as the major switch towards cancer progression. Different types of cancer are associated with increased expression of furin, as furin can also promote cell proliferation and

migration. Increased levels of furin have been reported for lung cancers, squamous cell carcinomas, breast and cervical cancer, gastrointestinal track, sarcomas, and cancers of the brain and the skin [42]. Furin has been associated with processing of cancer-related proteins such as growth factors and their receptors (e.g. IGF1 and IGFR), matrix metalloproteases (e.g. MMP14), cell adhesion molecules (e.g. integrin α-subunits), as well as angiogenic and lymphangiogenic factors (e.g. VEGF). In addition, furin has been used as a target to anti-cancer treatments such as inhibition of furin with camelid antibodies [43].

1.5 The proprotein convertases and processing of viral substrates

Furin can be involved in the processing of viral glycoproteins in at least four different forms. First, furin can cleave viruses after they are assembled and transit through the secretory pathway. Second, furin can also process individual viral proteins before being assembled at the plasma membrane. Third, viral proteins can also be cleaved during entry in endosomal vesicles. And fourth, furin can process viral proteins at the plasma membrane, before entry. Furin is involved in the processing of evolutionary diverse viruses, including families such Flavi-, Toga-, Retro-, Papiloma-, Corona-, Herpes-, Orthomyxo-, Paramyxo- and Filoviridae [23], [44]–[50]. In addition, furin can also influence the outcome of the infection with the virus. For example, highly pathogenic avian influenza viruses (HPAI) display a polybasic cleavage site that exploit furin activity and allow the viruses to spread systemically [51]. In the following sections, the life cycle of three viral families are addressed with emphasis on furin-mediated processing.

1.5.1 Coronaviruses

Until 2002 the coronaviruses were not recognized as a threat, but only associated with mild colds. However, in 2003 the Severe Acute Respiratory Syndrome (SARS-CoV) alarmed the world as it was the first highly pathogenic coronavirus with a mortality rate close to 10%. Then, in 2014 another coronavirus, Middle East Respiratory Syndrome, MERS, emerged with an even higher mortality rate, 30%. These two viruses were however not efficiently spread as the onset of symptoms happened early in the infection, which allowed the contention and prevention of world-wide spread. However, in 2019 a new coronavirus, this time named SARS-CoV-2 emerged as a

highly transmissible virus, which ended up in a pandemic currently ongoing which has claimed the lives of more than 4 million people as of August 2021.

The coronaviruses are enveloped single stranded RNA viruses with a genome of about 30 Kb, being the largest of the RNA viruses. Two-thirds of the genome encode for non-structural proteins (nsps) required for transcription and viral replication. This part of the genome is translated as a long polyprotein that is cleaved by viral and host proteases. The first part of the genome, named *Rep1a*, is translated into 11 proteins (nsp1 to nsp11), but 25% of the time, the ribosome slips into the -1 position of the stop codon of *Rep1a*, reading a different coding frame [52]. When this happens, the ribosome translates a longer polyprotein named *Rep1b*, which encodes five additional proteins (nsp12 to nsp16), among them the RNA-dependent RNA polymerase (RnRp). The last third of the genome undergoes a different strategy. In this case, seven to ten additional open reading frames (ORFs) are transcribed in a discontinuous manner, generating sub-genomic RNAs that encode for individual proteins including the Spike (S), nucleocapsid (N), Matrix (M) and Envelope (E).

The spike proteins (S) protrude out of the viral membrane giving the characteristic feature of the family (corona=crown in Latin). Among other domains, the spike protein contains the receptor binding domain and the fusion peptide, playing a key role both during attachment and entry, respectively. The receptors of human coronaviruses have been identified, including Angiotensin Converting Enzyme-2 (ACE2) and dipeptidyl peptidase 4 (DPP4) [53]. Upon binding to the receptor, there is a large-scale conformational change of the spike protein that exposes the hydrophobic fusion peptide. The exposure of the fusion peptide is critical for viral entry and there is extensive literature suggesting that it can be dependent on catalytic cleavage. There are three sites that have been suggested to be involved in the proteolytic cleave of the Spike protein: S1/S2, ECP and S2'. The S1/S2 site is the least conserved but most studied of the proteolytic cleavages of human coronaviruses. SARS-CoV does not have this site, but is present in MERS-CoV and the current SARS-CoV-2. Previous work done on the S1/S2 site of MERS-CoV suggest that the cleavage might happen during secretion, but whether or not furin is the main player, remains a topic of controversy [54]–[56]. This topic will be addressed in more detail in chapter 6.

1.5.2 Flaviviruses

The flaviviruses are a group of medically relevant blood-borne viruses. The Flaviviridae family has three major groups: the flaviviruses, the hepaciviruses and the pestiviruses. The genus Flavivirus include dengue, Zika and yellow fever viruses virus alone causes more than 390 million infections with more than 4000 deaths every year. The complications with dengue can vary from fever, hemorrhagic fever to shock syndrome. Yellow fever is the most lethal mosquito-borne Flavivirus, it can cause high fever, bleeding into the skin and damage of the liver leading to jaundice. In contrast, ZIKV was for a long time a relatively unknown Flavivirus given that it was sporadically appearing in isolated regions of the world. However, since 2007 a few epidemics occurred and in 2015 the virus spread in the Americas. The occurrence of ZIKV in the Americas was associated with increased levels of microcephaly in newborn, as well as the Guillain-Barré Syndrome. All these viruses are transmitted by mosquitoes, with *Aedes aegypti* as the preferred vector given its high anthropophilic lifestyle.

The flaviviruses have a complex life cycle. First, the virus first attaches to a receptor at the plasma membrane and gets internalized through the endocytic pathway. The low pH of the endosomes triggers exposure of the fusion loop present in the E protein, which brings together host and viral membranes. This results in release of the viral genome into the cytoplasm where it undergoes an initial translation into a polyprotein, which are processed into individual proteins by viral and host proteases. Then, the replication of the viral genome occurs in the membranes of the endoplasmic reticulum using the non-structural proteins of the virus (NS1, NS2A/B, NS3, NS4A/B and NS5) [57], [58]. Viruses are assembled into immature particles containing the structural proteins Capsid (C), prM (pre-membrane) and Envelope (E). At the surface of the immature virus, 60 heterotrimers of prM-E are found with a spiky appearance. The pr peptide covers the fusion loop of the E protein and prevents premature fusion with internal membranes in the cell. The assembled immature virus enters the secretory pathway and when it reaches the acidic pH of the Golgi apparatus, the virus undergoes a conformational change where the prM and E proteins are rearranged into 90 homodimers [50], [59]–[61]. At this point, the virus does not have the spiky appearance as before, instead it has a flat surface which make it look similar to a golf ball. This conformational change also exposes the furin cleavage site between pr and M (Figure 1.5), allowing the proteolytic cleavage performed most likely in the TGN, where furin is highly

abundant. The pr peptide remains bound to the viral particle after the catalytic cleavage, until it reaches neutral pH when the viral maturation is complete and the virion is budded out of the cell.

The furin-mediated cleavage of flaviviruses has been widely studied, as it is another factor in pathogenicity. The immature viral particle is presumed to lack infectivity, therefore inefficient cleavage results in lower viral titer. However, dengue virus is an example where the furin-mediated maturation affects the outcome of the disease. Dengue viruses contain an inefficient furin cleavage site (R-E/D-K-R), where the negatively charged E/D is believed to interfere with the negatively charged binding pocket of furin. This results in inefficient viral maturation and generation of mosaic particles, which can contain a mixture of both prM and M in the same virion [61]–[63]. As a consequence, when these virions are secreted and present in the serum, they can trigger a different immune response. There is extensive literature suggesting that antibodies against pr are not neutralizing but instead can cause what is called Antibody-Dependent Enhancement (ADE), where the viral particle can use the non-neutralizing antibody as a receptor and gain entry into the cells.

1.5.3   Togaviruses: alphaviruses

alphaviruses are a group of medically relevant arboviruses that are primarily transmitted by arthropods. These viruses are traditionally classified in two groups: New and Old-World alphaviruses. The New World alphaviruses, such as the Venezuelan equine encephalitis virus (VEEV), Eastern equine encephalitis virus (EEEV) and Western equine encephalitis virus (WEEV) typically cause encephalitis in humans and animals. In contrast, the Old World alphaviruses, such as Chikungunya (CHIKV), Ross River Virus (RRV) and Sindbis Virus (SINV) typically cause fever, rash or arthralgia, leading in rare cases to death. SINV is the most studied alphavirus given its high titer and versatility, however the outbreaks of chikungunya in different parts of the world, including Italy, have caused alarm for the control of these viruses.

The alphaviruses have a single stranded RNA genome with positive polarity. The life cycle of these viruses is complex. The virus attaches a receptor through interactions with the E2 protein. Then, the virus is internalized through clathrin-mediated endocytosis and in mature endosomes where the pH is low, the virus undergoes a conformational change that results in a homotrimeric E1 protein exposing its fusion peptide and inserting it into the endosomal membrane. This forms

a pore that releases the nucleocapsid into the cytoplasm, which gets disassembled rapidly. The viral genome RNA is translated initially into the non-structural proteins into P123 and in some cases using by leaking stop coding into P1234. P1234 is initially cleaved into P123 and nsp4 making an unstable replication complex capable of producing only negative strand RNA. Then, P123 is cleaved into nsp1 and P23, and in complex with nsp4, are able to produce both negative and positive strand, but no subgenomic RNA. Later, the P23 is cleaved into nsp2 and nsp3, which produces positive and subgenomic RNA. The subgenomic RNA is produced from negative sense RNA and contains the structural proteins which is initially translated into a polyprotein containing the capsid, PE2 (precursor of E3 and E2), 6k and E1 [64]. The capsid is first cleaved off and utilized to assemble the viral nucleocapsid with newly synthetized genomes. The E3 present in PE2 works as a signal peptide that targets the polyprotein into the secretory pathway. The PE2 and E1 are initially assembled in the ER as heterodimers. The PE2 regulates the fusogenic activity of E1, but when the PE2/E1 heterodimers reaches the low pH of the Golgi, PE2 is processed into E3 and E2 by furin or other proprotein convertases [23], [47], [64] (Figure 1.5). This cleavage induces a conformational change that weakens the E2/E1 interaction, priming E1 for fusion at lower pH. The mature E2/E1 proteins are transported to the plasma membrane where they assemble along with the nucleocapsid [65], [66].
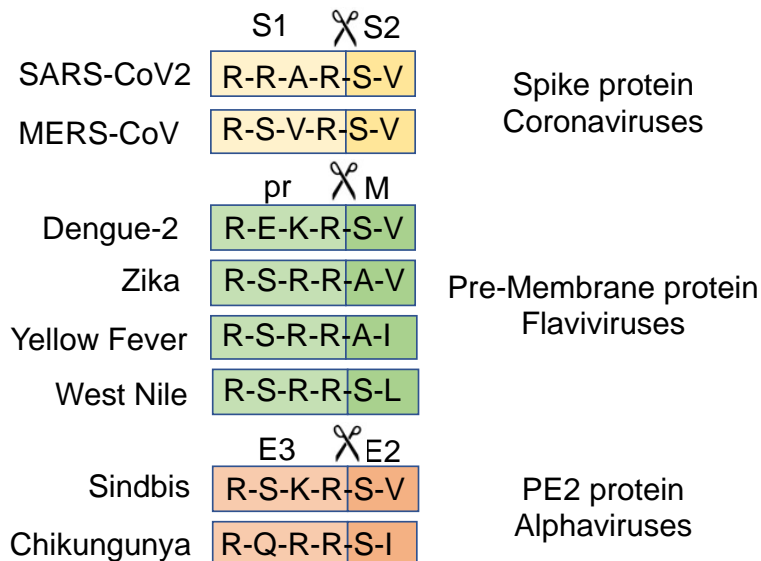
Figure 1.5 furin cleavage site among different viruses. The scissors represent the site where furin cleaves the substrate. The amino acid before the cleavage site is known as P1, and as is shown in this figure, it is the positively charged amino acid Arg. Positions P1 and P4 are always required to be a positively charged amino acid. The residues in between can vary, but negatively charged amino acids are less preferred, as is the case of DENV2.

1.6   Arthropod proprotein convertases

1.6.1   Mosquitoes as major bridges of viral infections

Mosquitos have been cataloged as the deadliest animal on earth as they can transmit multiple diseases of public health relevance. Every year, more than 800 thousand people die because of mosquito transmitted diseases, mostly malaria. However, mosquitoes also transmit viruses that cause significant burden, leading to increased levels of morbidity and mortality. For example, more than 40% of the world population is at risk of getting infected with at least one of the serotypes of dengue. Despite having a low mortality associated with DENV, the disease can cause significant health burden.

The risk of getting infected with a mosquito-borne disease is primarily determined by the distribution of the insect. The two major mosquito vectors, *Aedes aegypti* (the yellow fever mosquito) and *Aedes albopictus* (the Asian tiger mosquito), are widely distributed around the globe. However, *Aedes albopictus* is less resistant to extreme heat, making its distribution markedly expanded in temperate areas such as North America and Europe. However, global warming poses a threat to change in the dynamics of the distribution, and it has been suggested that overtime the impact of *Aedes albopictus* will be reduced in tropical areas where temperatures might increase, but this mosquito could eventually move to higher latitudes. In contrast, the more heat resistant *Aedes aegypti* mosquito is likely to stay in tropical areas, but also spread to temperate regions that foresee an increase of temperatures [67].

Not only the distribution of the mosquitoes affects the risk of infection, but also the lifestyle of the mosquito. For example, *Aedes aegypti* is a highly anthropophilic mosquito that prefers to stay indoors, allowing it to have contact with the humans in a more routine way, hence transmitting the virus to susceptible people more efficiently. In contrast, *Aedes albopictus* is also anthropophilic, but also an outside mosquito, which reduces the potential interactions with humans and eventually diluting the transmission of disease. *Culex quinquefasciatus,* which is the vector of West Nile virus (WNV), is more of an ornithophilic mosquito, suggesting that in very few cases the mosquito will bite a human but it prefers to feed on birds. This reduces further the chances of transmitting the virus among humans, and on top of that is the fact that WNV is a zoonotic disease

with the humans being the dead-end host, meaning that the mosquito cannot pick up sufficient virus from an infected human and transmit it to a new host.

It is important to point out that mosquitoes are not syringes that simply pass blood from one infected human to another. Instead, when a mosquito bites an infected human, the virus must undergo through several steps/barrier before it can be transmitted to the next host. First, the virus needs to infect the cells of the midgut and pass through the basal lamina. After this step, the virus must infect other cells of the mosquito, including the fat body and hemocytes. When this happens, the virus is spread all around the mosquito and able to reach the next barrier, which involves the infection of the salivary glands (SGs). After the virus infects this tissue, it can be secreted into the lumen of the SGs along with saliva. At this point, the mosquito is believed to have completed the extrinsic incubation period and be infectious for the rest of its life.

The vectorial competence is a way of measuring how efficient a mosquito is at transmitting a disease, and it is mostly determined by intrinsic factors involving the mosquito-virus interaction. As it was stated above, the virus must infect different types of cells before it can be transmitted to the next host, suggesting that both mosquito and viral factors must come into play. Several studies have suggested that the viruses can go through a few bottlenecks in the mosquito, determined primarily by evasion of the immune response of the mosquito (Figure 1.6). From these bottlenecks, the virus that comes out in the saliva might have incorporated mutations that were not present in the incoming virus [68]. In addition, specific point mutations have been shown to enhance the infection of the mosquito.

Mosquito abundance and distribution
Vectorial Competence
Mosquito-virus interactions*
Mutations in the virus
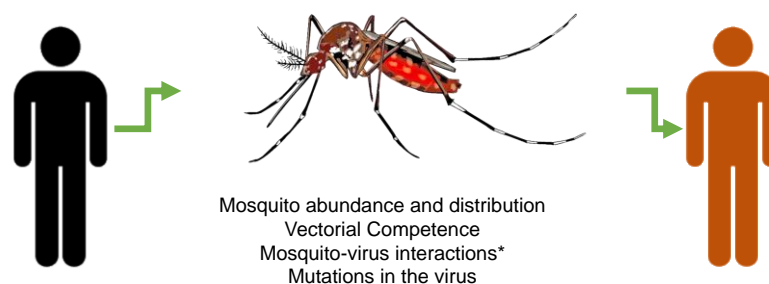
Figure 1.6 Determining factors of mosquito viral transmission. Transmission of a mosquito-borne virus relies on several factors such as the extrinsic incubation period, meal preferences of the mosquito, interactions of the virus with the mosquito and mutagenesis in the virus.

From the virus perspective, it must undergo multiple infection and replication cycles within the mosquito, suggesting that not only the challenge of the mosquito immune system is important. The virus relies on several host factors during its life cycle including proteases, glycosylation and other posttranslational modifications. Therefore, the maturation step in the life cycle of the virus would rely on the mosquito equivalents of the proprotein convertases and furin.

1.6.2   Control of infectious diseases at the vector level

The World Health Organization (WHO) proposes several approaches to control Vector-borne Diseases such as insecticide treated nets (ITNs), indoor residual spraying (IRS), outdoor spraying, addition of chemicals to water, biological and genetic control and integrated mosquito management. ITNs and IRS have been the major contributors of reducing malaria in Africa between 2000 and 2015 [69] and aerial spraying is necessary when an outbreak is imminent. The most common insecticide that is being used in the United States is Naled, an organophosphate implemented when people are getting sick or when the number of mosquitoes is large. However, it has been reported that Naled can also harm bees and other pollinators [70]. Another approach is Integrated Mosquito Management (IMM), which involes conducting mosquito surveillance, removal of breeding places where mosquitoes lay eggs or control of larvae/pupae and adults by using larvicides and adulticides, respectively.

Numerous genetic approaches have been done to alter mosquito populations by introducing Genetically Modified (GMO) Mosquitoes: i) modified males carrying a dominant lethal allele that produces offspring without females, ii) modified mosquitoes where females transform into males, and iii) introduction of sterility genes through drive systems (transposons, *Wolbachia*, *Medea*) [71]. There are many problems associated with GMO mosquitoes and some of them are around the implementation of the techniques themselves: how to keep these genes in the field and make them competitive against the natural population genes? However, this is not the biggest problem. Even if technical dilemmas of GMO mosquitoes can be successfully addressed, the next concern would be around the moral, ethical, natural and ecological consequences of eliminating mosquitoes from the world. Mosquitoes may play pivotal roles in ecosystems such as pollinators and components of the food chain. There could be poor evidence about the actual ecological function

of mosquitoes but there is absolutely no evidence of the consequences of a world without mosquitoes, "something even worse could take over" [72].

1.6.3 Current understanding of arthropod proprotein convertases

Most of the research on insect PCs has been done in *Drosophila melanogaster*, where unlike humans, it has three PCs: Dfurin1, Dfurin2 and Amontillado. Dfurin1 is involved in antimicrobial response [73] and fibroblast growth factor cleavage [74]. It has been shown that Dfurin1 efficiently cleaves the PE2 protein in alphaviruses [75]. Dfurin2 is involved in cleavage of the growth factor TRK as well as processing bone morphogenic proteins during embryogenesis [76]. Amontillado is involved in neuropeptide processing and hatching behavior in the pupal stages [77].

In mosquitoes, there are only two studies addressing the proprotein convertases. First, Chen and Raikhel described a subtilisin-like convertase named ´vitellogenin convertase´ present in the fat body and involved in the processing of pro-vitellogenin [78]. Second, Cano-Monreal suggested that the genome of *Aedes aegypti* encoded two furin-like proteases with high sequence similarity to Dfurin1 [75], but no functional studies were made on those. Therefore, there is a critical need to describe and functionally characterize the proprotein convertases expressed in mosquitoes. This project is significant because these proteases could be used as targets for inhibitory molecules. If mosquito proteases get inactivated, then flaviviruses would not be able to complete the maturation process and mosquitoes would produce immature non-infectious particles. The reasoning is that furin and other proprotein convertases are host factors common to multiple viruses. Therefore, understanding the proprotein convertases of mosquitoes will provide new targets for strategies to control arbovirus infection and transmission.

CHAPTER 2.   BIOINFORMATIC, TRANSCRIPTOMIC AND EXPRESSION ANALYSIS OF MOSQUITO PROPROTEIN CONVERTASES

2.1   Chapter Summary

Proprotein convertases are a family of structurally and catalytically related proteases that cleave at polybasic amino acid residues and render active proteins and peptides. The most studied proprotein convertase, furin, plays major roles in homeostasis, embryogenesis and neurodegenerative diseases, but is also exploited by pathogens such as viruses. Different virus families have been reported to use furin at some point of their life cycle. Alphaviruses and flaviviruses are both arthropod-borne viruses that also exploit furin to produce mature infectious particles. The mosquitoes are the primary route of transmission for medically relevant viruses like dengue, Zika or chikungunya, but despite their importance, little is known about the repertoire of proprotein convertases of *Aedes aegypti*. In this chapter, using bioinformatic and annotation tools, we have identified three proprotein convertases in the genome of *Aedes aegypti*: *Aa*FUR1, *Aa*FUR2 and *Aa*NC2. All three have the canonical domains of proprotein convertases: the prosegment, the catalytic domain and the P-domain. However, the proteases show some differences in the N-terminal and C-terminal regions. For example, both *Aa*FUR1 and *Aa*FUR2 do not have a predicted signal peptide bond, instead have a transmembrane region further down stream in the sequence, making a potential signal peptide with a longer extension than the one found in *Aa*NC2. In addition, we found that these genes are being expressed both in Aag2 cells and in adult female mosquitoes from the LVP strain, and their expression is differential among tissues. Finally, we attempted to produce antibodies for studying expression of these proteases. The antibodies were created in rabbits using mono-specific peptides. Together, these *in silico* analyses and *in vivo* experiments produced the first detailed description of the proprotein convertases of *Aedes aegypti* as well as an insightful expression profile.

2.2   Introduction

The proprotein convertases comprise an evolutionarily diverse group of serine proteases that are catalytically and structurally similar to bacterial subtilisin and yeast kexin [14], [79],

leading to the abbreviation of the family as PCSKs (Proprotein convertase subtilisin/kexin). In humans, this family comprises 9 proteases, out of which 7 show the kexin-like feature. They are characterized for carrying out the endoproteolytic cleavage of pre-proteins at conserved residues, mostly polybasic, producing mature and biologically active polypeptides and proteins.

Human furin is the most studied proprotein convertase with over 2000 research articles reported in Pubmed. This is because furin is ubiquitously expressed in different tissues, enabling furin to have key roles among multiple physiological aspects [23], [45], [46]. For example, furin can be involved in embryogenesis, homeostasis, tumor metastasis and neurodegenerative diseases [20], [27], [80]. Moreover, furin can be exploited by multiple bacterial and viral pathogens. Among the viruses that can use furin at some point of their life cycle are the coronaviruses, papillomaviruses, retroviruses, alphaviruses and flaviviruses [23].

Most alphaviruses and flaviviruses are considered arboviruses because their primary route of transmission is through an arthropod, which in most cases are mosquitoes and ticks. Yellow fever, dengue, Zika, and chikungunya are medically relevant viruses that affect morbidity mostly in tropical countries [68], [81], [82]. These viruses are mainly transmitted by *Aedes aegypti*, an insect of the *Culicidae* family and order *Diptera*. However, despite the importance of this mosquito in the dynamics of several arboviruses, little is known about the repertoire of proprotein convertases that might enable virus maturation and infectivity.

A few studies have addressed the role of proprotein convertases of other insects, including *Drosophila melanogaster* and *Locusta migratoria.* In *Drosophila melanogaster,* it has been shown that there are at least three proprotein convertases, including Dfurin1, Dfurin2 and Amontillado [77], [83]–[85]. Dfurin1 is involved in antimicrobial response, embryogenesis and neuromuscular junction [73], [74], [76], [86] , Dfurin2 is involved in regulation of the secretory pathway and embryonic patterning [76], [87], [88], whereas Amontillado is required for processing precursor proteins involved in muscle elongation,[89], [90]. Similarly, another study suggested that *Locusta migratoria* expresses an insect furin that is involved in the post-translational activation of the lipoprotein apoLp-II/I [91].

Regardless of the species and even the proprotein convertase, they all share three common domains that characterize the family: a pro-segment, a catalytic domain and a P-domain. The pro-segment (or pro-domain) acts as an intramolecular chaperone that enables the correct folding of the catalytic domain as well as inhibitory properties to enable localized activation [13], [92], [93]. The catalytic domain contains the catalytic triad D (aspartic acid), H (Histidine) and S (Serine), as well as N (Asparagine) which forms the oxyanion hole [27], [79]. Finally, the P-domain is necessary for providing stability to the catalytic domain by hydrophobic interactions and by stabilizing the pH [94].

In the present study, we performed bioinformatic analyses using the AaegL5.0 genome of *Aedes aegypti*. We have identified three proprotein convertases named *Aa*FUR1, *Aa*FUR2 and *Aa*NC2. These proteases contain all the canonical domains of a proprotein convertase, and their phylogenetic organization showed strong associations with other closely related species. However, *Aa*NC2 aligns better with human PC2. In addition, all three genes are producing RNA among different tissues of adult female mosquitoes of *Aedes aegypti* and in Aag2 cells, which is a cell line derived from the same mosquito species. Further, we attempted to produce antibodies against *Aa*FUR1 and *Aa*FUR2 for immunoblotting assays using specific epitopes in the catalytic and P-domains. Combined, these analyses and experiments suggest that mosquitoes produce three proprotein convertases with conserved amino acid sequence and diverse expression.

2.3 Materials and Methods

2.3.1 Gene annotation

The amino acid sequence of human furin was retrieved from NCBI (FURIN Gene ID: 5045) and used to identify orthologue proprotein convertases (PCs) in the AegL5.0 genome of *Aedes aegypti* using BLASTp. Mosquito PCs were subsequently annotated. Briefly, exonic sequences were retrieved from GenBank and manually curated using the coordinates in the genome. Gene models were constructed using IBS (Illustrator of Biological Sequences [95]). Protein structure models were constructed by predicting motifs (Pfam [96], Prosite [97], ClustalOmega [98]), signal peptide (SignalP5.0 [99]) and transmembrane regions (SOSUI [100]

and TMHMM [101]). Genes were named as members of the proprotein convertase family based on the BLASTp score and presence of determining domains: prosegment, catalytic and P-domain.

2.3.2 Phylogenetic analysis

A Bayesian inference of phylogeny was performed using the amino acid sequence of PCs of *Aedes aegypti, Anopheles gambiae, Drosophila melanogaster, Apis mellifera, Mus musculus and Homo sapiens.* Yeast Kexin was used as an outgroup. A sequence alignment with ClustalW was performed prior to the tree construction in phylogeny.fr [102]. The substitution model used for the Bayesian inference was Blosum62 and the Markov Chain Monte Carlo parameters included 100000 generations with sampling every 10 generations, discarding the first 250 trees. The resulting tree was annotated and curated in iTOL [103].

2.3.3 Structure prediction and analysis

Putative sequences of furin1, furin2 and NC2 were used to predict the structure using the iTasser server. The predicted structures were annotated using pymol and different domains were curated using the same software. The catalytic triad as well was highlighted and used for subsequent analysis and antigen development for antibody production.

2.3.4 mRNA extraction

Adult *Aedes aegypti* female mosquitoes from the Liverpool strain were dissected following the protocol of Schmid *et al*. 2017 [104]. The tissues extracted were salivary glands, malphigian tubes, ovaries, midgut and head. Tissues from >70 mosquitoes were homogenized using mortar and pestle in liquid nitrogen and then were lysed in RLT buffer with ß-mercaptoethanol. The rest of the RNA extraction was performed using the RNeasy mini-kit (Qiagen ID 74104) following manufacturer instructions.

RNA extraction from Aag2 and C6/36 cells was carried out from confluent 6-well plates. The cells were scraped off and lysed using RLT buffer plus ß-mercaptoethanol. The suspended cells were homogenized by passing through pipet tips multiple times. The rest of the RNA extraction was carried out following manufacturer instructions.

2.3.5 Quantitative Real Time PCR (qRT-PCR)

To determine the relative transcription of proprotein convertases in both insects and cells, specific primers targeting *Aa*FUR1, *Aa*FUR2 and *Aa*NC2 were designed to amplify a region of 200-250 bp. The PCR reaction was performed using the SuperScript™ III Platinum™ One-Step qRT-PCR Kit (Cat. 11732020). Data was represented as relative to the housekeeping gene Rps17.

2.3.6 Antibody production against AaFUR1 and AaFUR2

Antigen prediction of immunogenic amino acid epitopes was using the online prediction software SVMTrip [105]. Epitopes of 10-25 amino acids were selected based on: 1) potential off-targets in the genome of either mosquitoes or rabbits were detected by BlastP, sequences that shared >70% identity with other proteins were discarded; 2) accessibility of the peptide based on the predicted iTasser structure on Pymol [106]; and 3) ease of peptide synthesis based on its hydrophobicity and charged amino acid residues.

Selected peptides were synthetized by GenScript with a C-terminal cysteine for conjugation with KHL (Keyhole Limpet Hemocyanin). The peptides were used by Poccono farms laboratories for immunization of rabbits as follows: on day 0, rabbits were immunized with 200 µg of conjugated peptide in Complete Freund's Adjuvant (FCA). On days 14 and 28, rabbits were injected with 100 µg of conjugated peptide in Incomplete Freund's Adjuvant (IFA). Finally, on days 56 and 77, rabbits were injected with 50 µg of conjugated peptide in Incomplete Freund's Adjuvant (IFA). All injections were subcutaneous. Pre-immune, and partial immune bleeds were collected on days 0, 42, 63, 70, 84, 91 and 124.

2.3.7 Western Bloting

Proteins were run under reducing and denaturing conditions in a 10% SDS-PAGE for 1 hour 10 min at 110 volts. Then, the proteins were transferred onto a 0.45 µm nitrocellulose membrane (Biorad 1620115) for 1 hour at 100 V at 4 ºC. Subsequently, the membrane was blocked with 5% fat free milk in 1X TBS (0.05% Tween). After that, the polyclonal serum from rabbit was diluted 1:500 in blocking buffer and left overnight with the membrane at 4 ºC. The next day, the membrane was washed three times with 1X TBS-T before adding the secondary antibody 1:10000

diluted in blocking buffer. Membranes were imaged in an Odyssey machine (LI-COR Biosciences).

2.4   Results

2.4.1   Gene models

Using genetic and transcriptomic data from the AaegL5 genome assembly, we identified three proprotein convertases (Table 2.1) and predicted the gene models and coding regions of *Aa*FUR1, *Aa*FUR2 and *Aa*NC2 (Figure 2.1a). *Aa*FUR1 and *Aa*FUR2 have five and two transcript variants, respectively. For both *Aa*FUR1 and *Aa*FUR2, the transcript variants differ only on the first to second exons (Figure 2.1a, Table 2.2), but the same coding sequence is maintained. *Aa*NC2 only has one transcript variant.

2.4.2   Protein motifs and structure

The deduced protein sequences of *Aa*FUR1, *Aa*FUR2 and *Aa*NC2 were used to identify conserved domains/motifs of proprotein convertases (PCs) and predict the protein structure models. All three proteins displayed the minimum characteristics of a PC: pro-segment, catalytic domain, the catalytic triad and a P-domain (Figure 2.1b). However, there are some variations in the N- and C-termini. For example, *Aa*FUR1 and *Aa*FUR2 do not have a typical signal peptide. Instead, they display a transmembrane region proximal to the N-terminus. Another variation is at the C-terminus, where *Aa*FUR1 and *Aa*FUR2 display two extra features: a cysteine-rich region and a tyrosine growth receptor-like region. Finally, unlike *Aa*FUR1 and *Aa*FUR2, *Aa*NC2 does not display a C-terminal transmembrane region.

Table 2.1 Proprotein Convertases: annotation summary

NCBI LOC Accession	NCBI XM Accession	NCBI XP Accession	Chromosome	Location	Length (bp)	Length (amino acids)	Max number of exons	No. Transcript variants	No. Isoforms	% Ident with human furin
AaFUR1										
AaFUR1 LOC5578803	XM_021854000.1	XP_021709692.1	3	65770170 to 66596795, complement	5566	1135	13	5	1	47.65
	XM_021854002.1	XP_021709694.1			4725					
	XM_021854003.1	XP_021709695.1			5018					
	XM_021854004.1	XP_021709696.1			4594					
	XM_021854005.1	XP_021709697.1			4306					
AaFUR2										
AaFUR2 LOC5573800	XM_021848508.1	XP_021704200.1	3	130090025 - 131184447, complement	8535	1535	16	2	1	47.8
	XM_021848509.1	XP_021704201.1			6730					
AaNC2										
AaNC2 LOC5564601	XM_001648899.2	XP_001648949.2	1	150140028 - 150170019, complement	3564	640	12	1	1	47.43
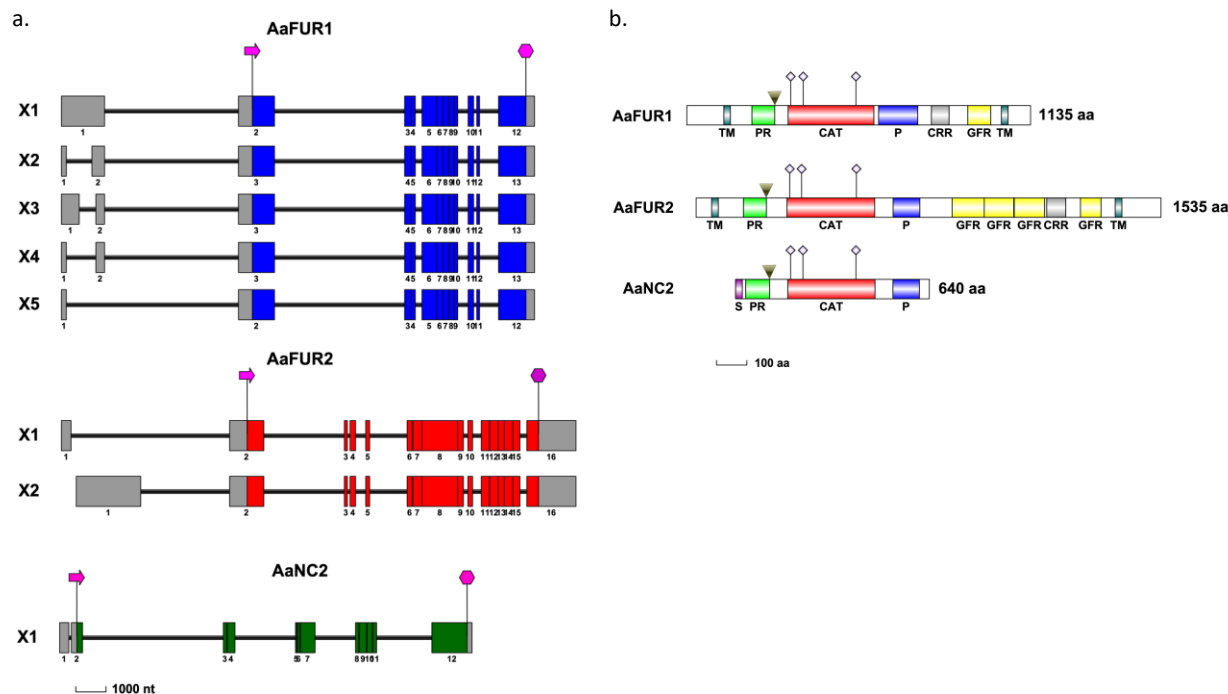
Figure 2.1 Gene and protein models of *Aedes aegypti Aa*FUR1, *Aa*FUR2 and *Aa*NC2
(*a*) Schematic of the gene models of the mosquito proprotein convertases. Exons are drawn as boxes and introns as black interrupted line. Only exons are drawn to scale. The X*n* represent the transcript variants. Pink arrow and hexagon represent start and stop codons, respectively. Grey boxes represent 5' and 3' UTR sequences, whereas colored blue, red and green boxes represent the coding sequence of *Aa*FUR1, *Aa*FUR2 and *Aa*NC2, respectively. (*b*) Protein structure models. Amino acid sequences are drawn to scale. Specific domains/motifs are differently colored and labeled as follows: TM Transmembrane region, S Signal Peptide, PR Pro-segment, CAT Catalytic Domain, P P-domain, CRR Cysteine Rich Region, GFR Growth Factor Receptor-like. In addition, the location of the catalytic triad DHS is represented by diamonds in the CAT domain. Gold/black inverted arrowhead indicates the autocatalytic cleavage site.

2.4.3 Phylogenetic tree

To determine conservation of *Aa*FUR1, *Aa*FUR2 and *Aa*NC2 with different orthologs, we performed a percent identity analysis of individual domains (Figure 2.2, Tables 2.3 and 2.4). As expected, the catalytic domain is the most conserved across species. In addition, we performed a Bayesian phylogenetic analysis to discern relationships of *Aa*FUR1, *Aa*FUR2 and *Aa*NC2 with PCs of other organisms (Figure 2.3). *Aa*FUR1 clustered in a clade with furin-like protease-1 of *Anopheles gambiae* (*Ag*FUR1)*, Drosophila melanogaster* (*Dm*FUR1) and *Apis mellifera* (*Am*FUR1). It did not show any direct relationship with any of the mammalian PCs. Similarly, *Aa*FUR2 was included with the furin-like protease-2 of *Anopheles gambiae* (*Ag*FUR2)*, Drosophila melanogaster* (*Dm*FUR2) and *Apis mellifera* (*Am*FUR2). Finally, *Aa*NC2 clustered

43

Table 2.2 Variation of exon length in base pairs (bp) among transcript variants

Exon	AaFUR1					AaFUR2		AaNC2
	Transcript variant 1: XM_021854000.1	Transcript variant 2: XM_021854002.1	Transcript variant 3: XM_021854003.1	Transcript variant 4: XM_021854004.1	Transcript variant 5: XM_021854005.1	Transcript variant 1: XM_021848508.1	Transcript variant 2: XM_021848509.1	Transcript: XM_001648899.2
1	1419	159	583	159	159	2120	314	302
2	1186	419	288	288	1186	1128	1128	369
3	216	1186	1186	1186	216	105	105	108
4	129	216	216	216	129	183	183	256
5	482	129	129	129	482	135	135	38
6	206	482	482	482	206	160	160	77
7	187	206	206	206	187	304	304	479
8	174	187	187	187	174	1173	1173	101
9	119	174	174	174	119	188	188	234
10	179	119	119	119	179	157	157	165
11	91	179	179	179	91	275	275	118
12	1178	91	91	91	1178	277	277	1317
13	-	1178	1178	1178	-	200	200	-
14	-	-	-	-	-	279	279	-
15	-	-	-	-	-	236	236	-
16	-	-	-	-	-	1616	1616	-
Total	*5566*	*4725*	*5018*	*4594*	*4306*	*8536*	*6730*	*3564*

Table 2.3 Percent identity of domains of Proprotein Convertases among different organisms

	Domain	*Ag*FUR1*	*Dm*FUR1*	*Am*FUR1*	*Mm*FURIN*	*Hm*FURIN*
AaFUR1	Prosegment	68.42	65.33	52.00	53.33	52.00
	Catalytic	94.29	87.62	81.69	71.11	71.76
	P-domain	89.23	80.77	70.59	43.31	43.31
AaFUR2	Prosegment	81.82	64.94	61.04	50.67	50.67
	Catalytic	99.69	91.64	91.38	70.65	72.00
	P-domain	93.53	85.61	82.02	41.35	40.60
AaNC2	Prosegment	72.50	65.00	60.76	29.33	28.00
	Catalytic	96.63	98.16	95.04	57.01	57.09
	P-domain	86.86	73.72	80.46	41.35	40.60

* Ag. Anopheles gambiae, Dm. Drosophila melanogaster, Am. Apis mellifera, Mm. Mus musculus, Hm. Homo sapiens.

not only with the insect NC2/Amontillado PCs (*Ag*NC2, *Dm*AMON and *Am*NC2), but also with the putative mammalian orthologues of *Mus musculus* (*Mm*PC2) and *Homo sapiens* (*Hm*PC2).

*Aa*FUR1

*Aa*FUR2

% identity = 59%

% identity = 55%
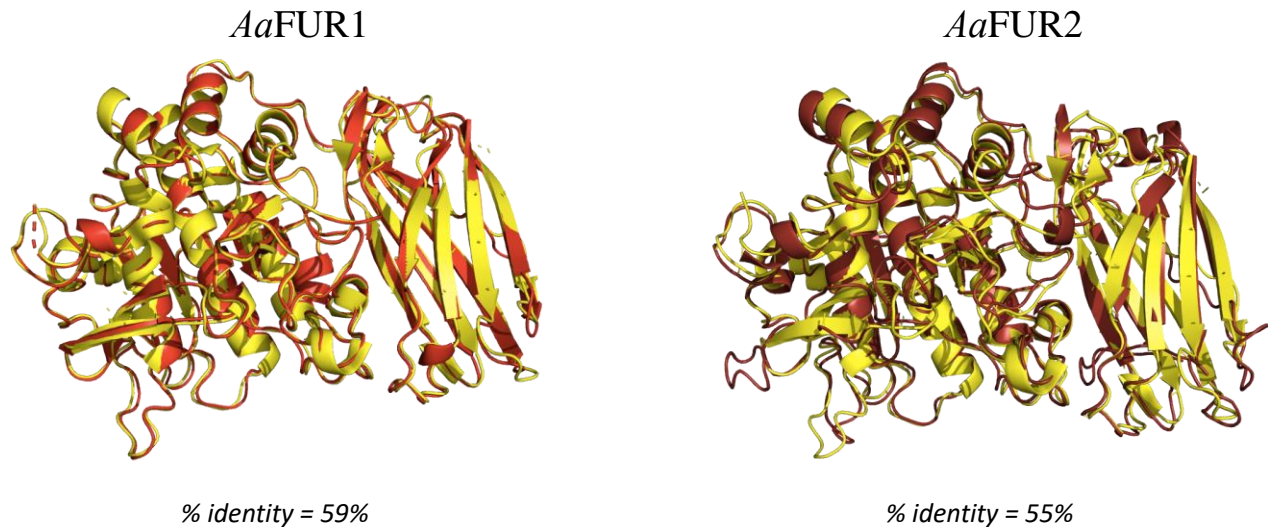
Figure 2.2 Superposition of mosquito furin-like proteases with human furin.
The crystal structure of the catalytic and P-domains of human furin was retrieved from PDB (1P8J) and was aligned in with the i-Tasser predicted structures of *Aa*FUR1 or *Aa*FUR2. Human furin and mosquito furin-like are shown in yellow and red, respectively. The overall identity percentage is displayed under each structure.

ScKEXIN

MmPC7
HsPC7

AmNC1

MmPC1
HsPC1

MmPC6
HsPC6
MmPC5
HsPC5B
HsPC5A

AmFUR2
DmFUR2
AgFUR2
AaFUR2

MmPC4
HsPC4

MmFURIN
HmFURIN

AmFUR1
DmFUR1
AgFUR1
AaFUR1

MmPC2
HsPC2

AmNC2
DmAMON
AgNC2
AaNC2

Tree scale: 0.1
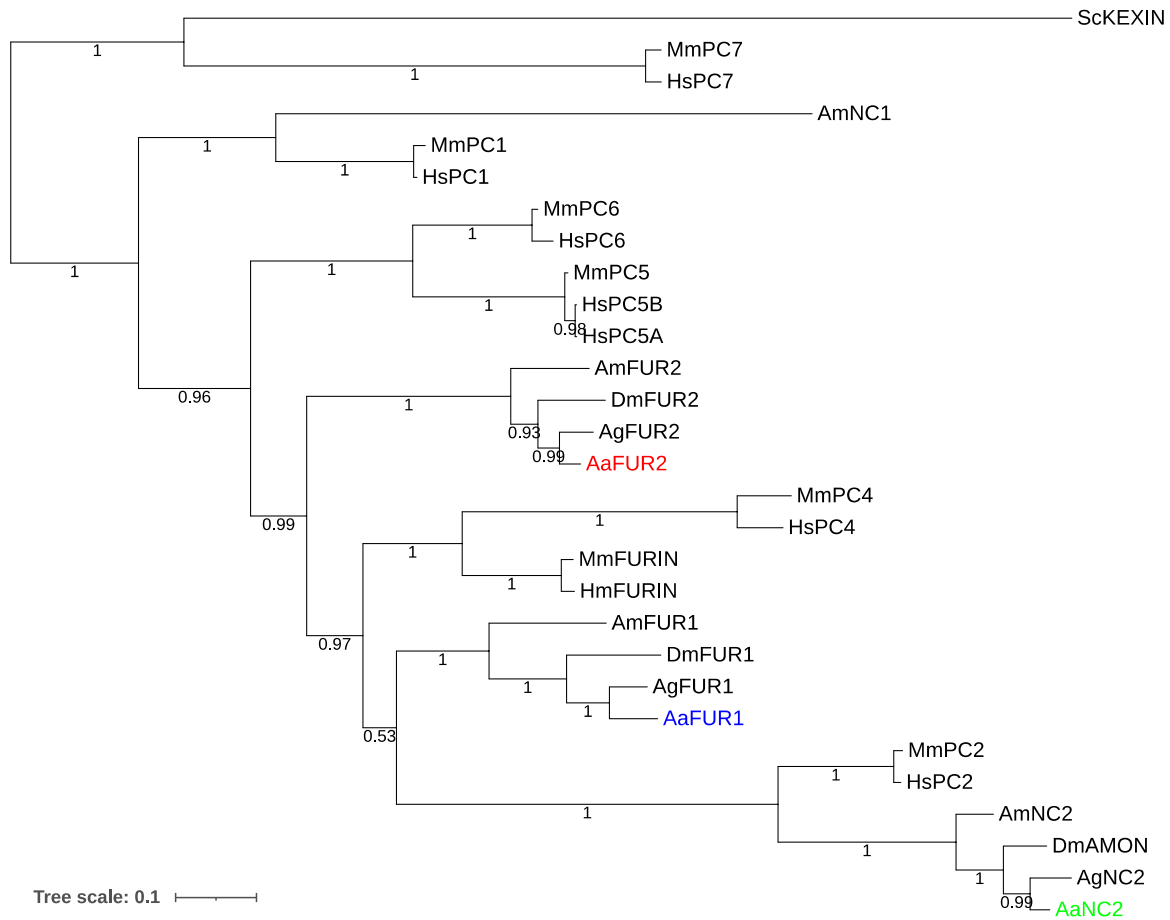
Figure 2.3 Bayesian phylogeny of proprotein convertases
Unrooted tree showing the main relationships of mosquito PCs to its counterparts in Ag: *Anopheles gambiae,* Dm: *Drosophila melanogaster,* Am: *Apis mellifera,* Mm: *Mus musculus*, Hm: *Homo sapiens* and Sc: *Saccharomyces cerevisiae*. The branches are supported by posterior probability values. The tree is drawn to scale: amino acid changes per site. Mosquito AaFUR1, AaFUR2 and AaNC2 are highlighted in blue, red and green, respectively.

2.4.4   Expression in tissues of *Aedes aegypti* and Aag2 cells

To determine whether these genes were being expressed, we extracted RNA from adult female mosquitoes of *Aedes aegypti* of the Liverpool strain. The RT-PCR indicates that these genes are producing RNA transcripts (Figure 2.4a) and the expression is different among tissues. *Aa*FUR1 and *Aa*FUR2 have the highest expression in the salivary glands, whereas AaNC2 has the highest expression in the head (Figure 2.4b). In addition, we also checked for the relative expression in Aag2 cells (Figure 2.4c). AaFUR2 is the proprotein convertase with significantly minor expression in Aag2 cells.

Table 2.4 Percent identity among domains of Proprotein Convertases of *Aedes aegypti*

		AaFUR1	*AaFUR2*	*AaNC2*
Pro-segment	*AaFUR1*	100	46.67	31.58
	AaFUR2	46.67	100	31.17
	AaNC2	31.58	31.17	100
		AaFUR1	AaFUR2	AaNC2
Catalytic domain	*AaFUR1*	100	69.01	53.18
	AaFUR2	69.01	100	54.43
	AaNC2	53.18	54.43	100
		AaFUR1	AaFUR2	AaNC2
P domain	*AaFUR1*	100	48.06	44.19
	AaFUR2	48.06	100	37.04
	AaNC2	44.19	37.04	100

		AaFUR1	AaFUR2	AaNC2
Full length	AaFUR1	100	41.99	44.33
	AaFUR2	41.99	100	43.52
	AaNC2	44.33	43.52	100

2.4.5 Antibody production

With the aim to produce antibodies against *Aa*FUR1 and *Aa*FUR2, we first predicted antigenicity of epitopes in the catalytic and P-domains (Figure 2.5 a,b). The epitopes selected showed a percent amino acid identity below 70% with other proteins in the mosquito (Figure 2.5c) and were relatively accessible in the predicted protein structure. The synthetized peptides were used to immunize the rabbits and the development of specific antibodies was detected from different bleeds for a period of 124 days. As evidenced by western blot, none of the bleeds collected showed a strong specific band for the expected sizes of each protein in any of the immunized rabbits (Figure 2.6). In addition, it is difficult to conclude which band corresponds to the corresponding furin because there is a lot of background in the pre-immunized sera, especially in the range of the expected molecular sizes. This is particularly true for Rabbits A and B, where the expected size of *Aa*FUR1 should be between 92 and 123 kDa. In other words, none of the bands that are present in the post-immunization bleeds are new because they all are also present in

the pre-immunized bleed. In contrast, rabbit D showed a set of bands in the range of *Aa*FUR2 that were not present in the pre-immunized bleeds. These include a band present that is in the range of *Aa*FUR2 at around 167 kDa.
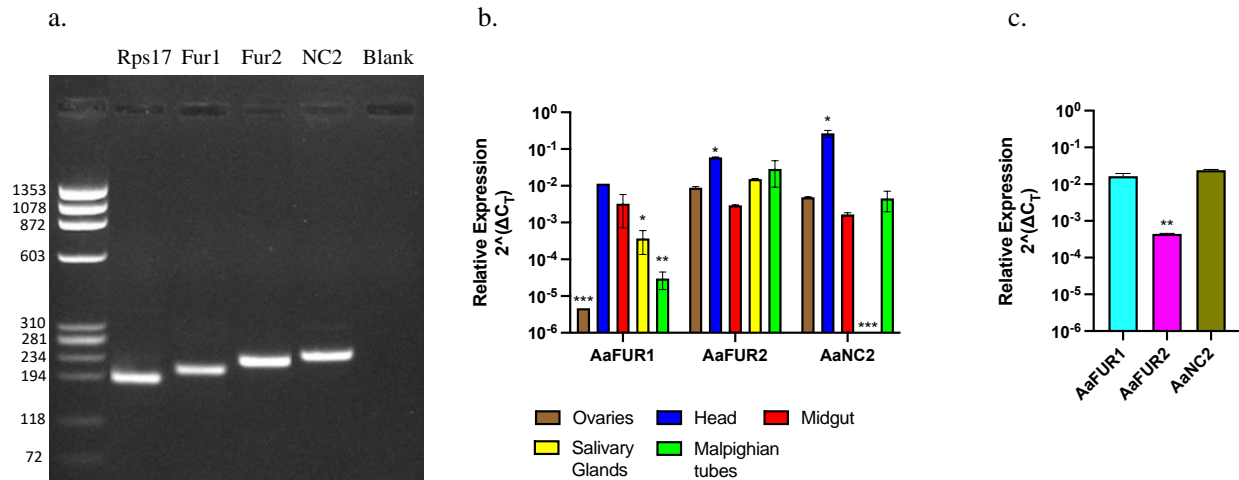
Figure 2.4 Expression of mosquito Proprotein Convertases
(*a*) RT-PCR from whole adult female mosquitoes. Rps17 195bp, furin1 216 bp, furin2 235 bp and NC2 246 bp. (*b*) Relative expression of mosquito proprotein convertases among different tissues of adult female mosquitoes. (*c*) Relative expression in Aag2 cells to the housekeeping gene Rps17. N=2. Significance was calculated with a two-tailed t-test, P<0.05.

	Amino acid sequence	Length	%Identity to AaFUR1	% Identity to AaFUR2	% Identity to Hfurin
AaFUR1	SDHPDLEHNYDPKAC	15	100%	69.23%	72.73%
AaFUR2	EINEDRPIEGSVGC	14	50%	100%	35.70%
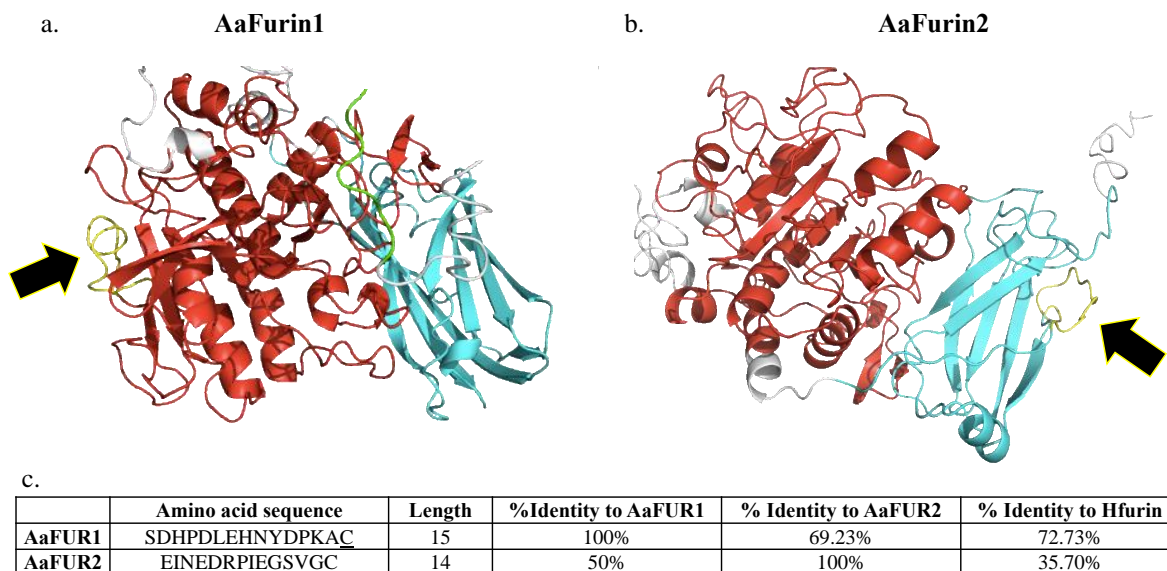
Figure 2.5 I-Tasser predicted structures with antigens colored in yellow
The black arrows point the location of the antigen in the protein. Prosegment, catalytic domain and P-domain are labeled in green, red and cyan, respectively. (*a*) AaFUR1 and (*b*) AaFUR2. (*c*) Percent identity table comparing the identity of epitopes between the AaFUR1 and AaFUR2 homologue sequences.
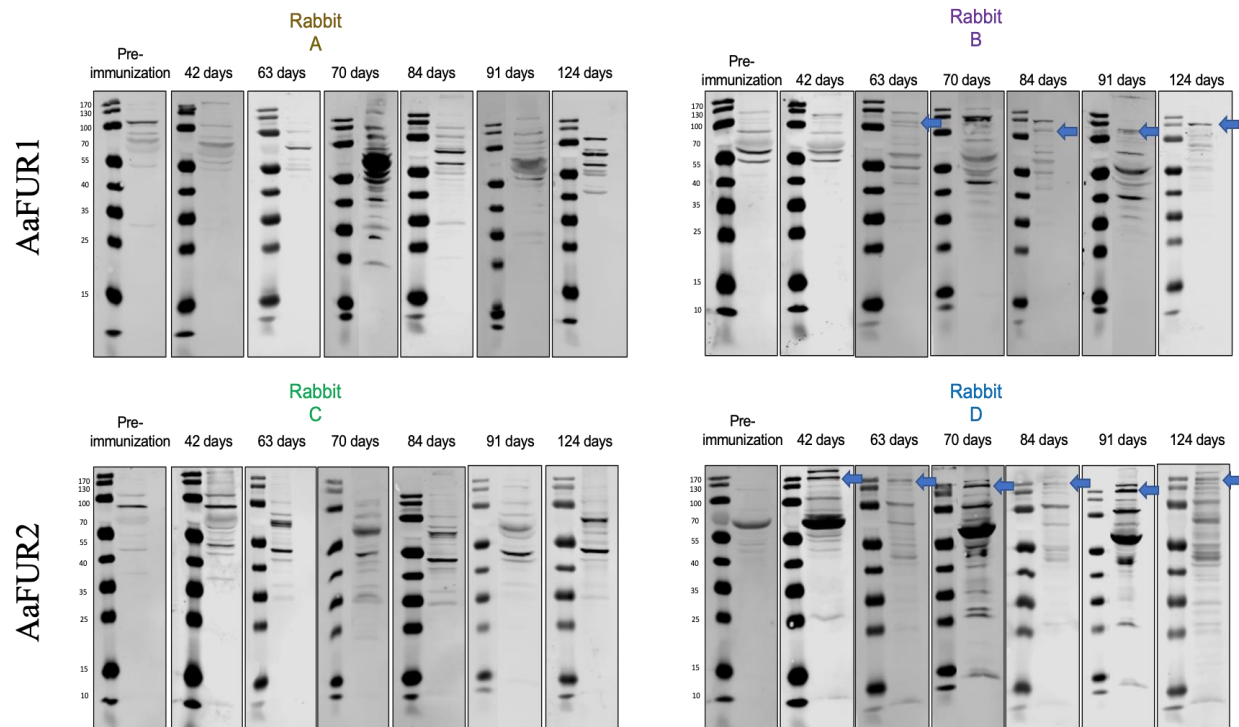
Figure 2.6 Western Blot using bleeds of immunized rabbits against *Aa*FUR1 and *Aa*FUR2
Bleeds collected at different times pre- and post-immunization, were tested against cell lysate of Aag2 cells. The bleeds were 1:500 diluted in 5% fat-free milk/TBS-T. Immunized rabbits against AaFUR1 are shown as A and B. Immunized rabbits against AaFUR2 are shown as C and D. The blue arrows point to bands that correspond to the expected size of the full-length protein: AaFUR1= 123 kDa ; AaFUR2 = 167 kDa. Note that these expected sizes correspond to the pre-mature protein as it should be before transiting in the trans-Golgi network. The size of the post-maturation proteases would be 92 kDa and 142 kDa, respectively.

2.5   Discussion

In this chapter, we described for the first time the proprotein convertases of *Aedes aegypti* and provided evidence of RNA expression. Our annotations highlight the diversity in transcript variants of *Aa*FUR1 and *Aa*FUR2, particularly in exons 1 and 2 (Figure 2.1). This variation in splicing does not have any effect in the amino acid sequence because the alternation happens at the 5' UTR. However, this difference might have an impact in localized expression among different tissues. This would require further studies to determine where each of these transcript variants are prevalent.

Notwithstanding, the protein models showed the minimal requirements of a proprotein convertase: the Prosegment, the Catalytic domain and the P-domain. However, there are variations among other motifs such as the signal peptide. Neither AaFUR1 nor AaFUR2 showed a canonical

signal peptide. Instead, they have a transmembrane region located near the N-terminal region which might act as a signal peptide. This aspect will be further addressed in Chapter 3. In addition, the Growth Factor Receptor-like (GFR) domain was present in AaFUR1 and repeated four times in AaFUR2. The function of GFR is not clearly understood but it displays overlapping amino acid sequence similarities with the furin Cysteine Rich Region (CRR) domains. The CRRs, present both in AaFUR1 and AaFUR2, are probably implicated in protein aggregation and signal transduction of tyrosine kinases as it has been reported for this particular type of motif [107].

The phylogenetic analyses of this chapter emphasize the homology of AaNC2 with amontillado and human PC2, which correltate with the fact that it was the only proprotein convertase without a transmembrane region. Instead, AaFUR2 appeared at the base of the clade containing human furin, AaFUR1 and AaNC2. This is in agreement with other studies that have suggested that furin2, is a protease specific to protostomes (represented here by the honey bee, fruit fly and mosquitoes) and was lost on chordates (mice and humans) [108]. In contrast, other studies have concluded that furin1 was the original furin and was eventually duplicated into furin and PC4. This makes sense in our phylogenetic tree because AaFUR1 (along with the other insect orthologs) appear at the base of the clade that contains furin along with the more divergent NC2/PC2/Amontillado group. In addition, it is important to point out that the prosegment was the least conserved domain among the species analyzed in this study. Given that the prosegment is implicated in ensuring the proper folding of the catalytic domain, this might account for the differences in the overall amino acid sequence of the protein in conjunction with cellular factors associated with posttranslational modifications. Interestingly, the prosegments of *Aa*FUR2 and *Aa*NC2 share higher amino acid sequence identity with the prosegments of other furin1 from other organisms than *Aa*FUR1.

The expression of RNA suggests that these genes are not pseudogenes but are transcribed and expressed. In addition, the differential expression among different tissues might account for other biological questions, in particular the protease that is exploited by viruses when they infect different organs. For example, the midgut, which is the initial tissue that viruses infect in mosquitoes, shows equal expression levels of all three proteases, whereas the salivary glands, where the virus needs to be produced before being transmitted to the next host, does not have any

expression of *Aa*NC2, but both *Aa*FUR1 and *Aa*FUR2 are present. Further functional and biological assays that highlight the importance of these proteins in the context of viral infection are addressed in chapters 3, 4 and 5. Finally, the production of antibodies against host factors is important to establish expression levels of the proteins, which ultimately represents the final step in expression. We attempted to produce antibodies against *Aa*FUR1 and *Aa*FUR2 using specific peptides. However we were not able to detect protein from Aag2 cells. This experiment requires further investigation because the expression levels in Aag2 cells might be extremely low.

In conclusion, in this study we provided extensive annotation information about the proprotein convertases of *Aedes aegypti* as well as expression analyses. We conclude that mosquitoes in fact have a repertoire of furin-like proteases that have homologs in humans, mice, flies and bees. Therefore, given the relevant role of mosquitoes in the transmission of viruses and the viral need of furin during its life cycle, these newly discovered mosquito PCs need to be further studied, which is the main objective in this thesis.

CHAPTER 3.   HETEROLOGOUS EXPRESSION STUDIES OF PROPROTEIN CONVERTASES OF MOSQUITOES IN MAMMALIAN AND INSECT CELLS

3.1   Chapter Summary

Expression of recombinant protein is a powerful tool for understanding the functions of a particular biologic of interest. In virology, the expression of host proteins utilized during viral infection is helpful to understand viral host interactions. Protein expression systems exist in different formats, including bacterial, yeast, insect and mammalian cells. However, the selection of one versus the other depends on the protein of interest. Here, the proprotein convertases of *Aedes albopictus* and *Aedes aegypti* were expressed in S2, C6/36, Aag2 and HEK293T cells. Expression of individual domains bound to the BiP signal peptide of *Drosophila* did not result in detectable protein by western blot. Similarly, the expression of full length soluble furin1 and furin2 of *Aedes albopictus* did not produce secreted protein as the N-terminal region corresponded to an inaccurate annotated sequence. Furthermore, secretion was not recovered even after replacing with human furin signal peptide, emphasizing the importance of the native N-terminal region. However, expression of well genome annotated furin1, furin2 and NC2 from *Aedes aegypti* did result in protein expression and secretion, and the co-expression of NC2 with the peptide 7B2 increased the secretion but not the activity. However the amount of protein was very low and was difficult to purify. Finally, codon optimization for expression in *Drosophila* S2 cells and under the inducible metallothionein (*Mt*) promoter resulted in higher amounts of protein detectable by SDS-PAGE. The purification of furin1 was performed in three different strategies: i) Ni-NTA affinity alone, ii) Ni-NTA plus FLAG-Immunoprecipitation and iii) Ni-NTA plus gel filtration. The first strategy resulted in samples with contaminants with a lesser intensity. The second strategy resulted in purified protein with almost no contaminants detected by silver staining. And the third strategy produced larger amounts of protein but with some contaminants. The expression and optimization experiments addressed in this chapter set the stage for producing protein and characterize the enzymatic activity in further chapters of this thesis.

3.2   Introduction

Protein expression and purification is an important aspect in many as areas of virology research, including structural biology and biochemistry. Production of recombinant protein is necessary to elucidate important steps in the life cycle of a virus, and therefore identifying potential targets for antiviral development. The most common expression systems include *E. coli,* yeast, insect, mammalian cells and cell-free. The *E. coli* system is the preferred one in many instances given its versality and ease, but it has limitations when certain post-translational modifications such as disulfide bonds and glycosylation, necessary for some protein of interest [109]–[111]. Similarly, the cell-free expression systems are good for proteins that are difficult to express, such as toxins [112]. However, complex proteins produced in cell-free systems do not get properly folded and the post-translational modifications do not occur. Therefore, the eucaryotic systems are more suitable for proteins that rely on post translation modification events [113].

Furin is a proprotein convertase with wide variety of cellular functions [27]. Similar to viruses, furin also has a life cycle in which the protein is initially translated in the endoplasmic reticulum with a signal peptide which targets the protein into the secretory pathway [23]. The signal peptide is recognized and cleaved by signalases. Then, in the endoplasmic reticulum, the protein undergoes glycosylation events in at Asn387 and Asn440. Once the protein progresses towards the Golgi, it undergoes autocatalytically activation and accumulates mostly in the trans-Golgi network (TGN). Furin can be recycled between the TGN and the plasma membrane, but can also be truncated and shed out of the cell [22], [107]. Therefore, it is evident that furin relies on multiple cellular factors for post-translational modification. Lack of glycosylation in Asn387 and Asn440 results in protein deficient in autoactivation [15]. As a result, recombinant expression of furin and furin-like proteases cannot be done in either cell-free or in *E. coli* systems. Instead, an eucaryotic system is more suitable.

Mammalian expression systems offer the advantage of providing the environment needed to express most eukaryotic proteins. CHO and 293 cells are the most popular mammalian expression systems [114]. These cells can be either stable or transiently transfected. Transient transfections offer high yields within one to two weeks [115]. The 293 cells, also known as HEK293 cells were derived from a human embryonic kidney in 1977 [116] and transformed with

sheared adenovirus 5 (Ad5), making them more efficient at producing high amount of protein under the CMV promoter [116]. In addition, the HEK293 cells were further modified with the incorporation of the SV40 large antigen T, which makes the cells to produce large amounts of protein from plasmids with the SV40 origin of replication [117], [118]. This daughter cell line is called HEK293T and is highly transfectable and routinely used in gene expression and protein production.

In terms of posttranslational modifications, insect expression systems offer the same advantages as mammalian cells. However, insect expression systems can be scaled up to produce larger amounts of proteins and even express multicomplex proteins. The most common insect expression system is the baculovirus in Sf9 cells [119]–[121]. However, the *drosophila* S2 cells have been widely used for protein production recently. This is because the S2 cells offer the same advantages of the baculovirus Sf9 system with additional flexibilities around the post-translational modification and production of active proteins. The S2 cells were originated from early embryo of *drosophila* and are believed to be derived from macrophage-like cells [122], [123]. S2 cells can be stable and transiently transfected with plasmids carrying the gene of interest. Different expression plasmids have been used in S2 cells, including constitutively expressed actin 5C promoter (*Ac5*) or inducible metallothionein (*Mt*) [124]. The metallothionein promoter offer the advantage of expressing proteins that under constitutive conditions would result toxic for the cells [125], [126], once again, showing the versatility of the system. In addition, stable transfected S2 cells can be scaled up to several liters and produce enormous amount of protein for further analytical and function steps.

In addition to expression, different detection and purification strategies exist. The most common is SDS-PAGE and staining with Coomassie or another similar reagent. However, in some cases, the protein expression is very low, and needs to be detected by more sensitive methods such as western blot, fluorescent or luminescent tags. HiBiT is a recent technology developed by Promega in which proteins are tagged with a sequence that when incubated with a larger protein portion (Largebit), it will emit light [127]. This technology is able to detect proteins at very low levels without the need to make time consuming western blots. In addition, the HiBiT works well for detection of intracellular and extracellular protein and can be applied to different expression

systems with ease. On the other hand, purification based on affinity can be performed using tags such as poly-histidine for Ni-NTA purification and FLAG for immunoprecipitation. The Ni-NTA purification has the advantage of being versatile and having a bigger binding capacity but can result in samples without complete purity. FLAG immunoprecipitation offers the opposite, highly pure samples but lower protein yield.

In this chapter, attempts to express proprotein convertases of *Aedes albopictus* and *Aedes aegypti* were done in mammalian and insect cells. Different systems were sought, including the HEK293T cells, S2 cells, C6/36 and Aag2 cells. In addition, three plasmid vectors were used: pcDNA3.0, carrying the CMV promoter, pMT and pMT puro, carrying the inducible metallothionein and pUB, carrying the promoter of polyubiquitin of *Aedes aegypti.* In addition, multiple detection platforms were tested, including western blot, HiBiT signal, FLAG peptide and furin activity assays. Finally, different purification strategies were implemented to achieve high protein yield, such as Ni-NTA purification, FLAG immunoprecipitation and size exclusion chromatography.

3.3   Methods

3.3.1   Cells

C6/36 cells were maintained in Minimum Essential Media (MEM) supplemented with 10% FBS at 30 C with 5% CO_2. Aag2 and S2 cells were maintained in *Drosophila* Schneider's Medium, supplemented with 10% Fetal Bovine Serum (FBS) at 27 C and no CO_2. In addition, Aag2 cells were also supplemented with non-essential amino acids and PenStrep. HEK293T cells were maintained in Dulbecco's Modified Eagle Medium (DMEM) with 10% FBS at 37 C and 5% CO_2.

3.3.2   RNA extraction and first strand of cDNA

RNA from C6/36 and Aag2 cells was extracted using the Invitrogen Purelink RNA Mini Kit. Briefly, sub confluent cells in a 6-well plate were scrapped off with pipette tips and treated with lysis buffer containing ß-mercaptoethanol. Further purification steps were performed using manufacturer instructions. For the cDNA synthesis, approximately 1 µg of RNA was used with the iScript Select cDNA Synthesis Kit with oligo dT primers.

3.3.3   Clones construction

Catalytic and soluble clones of *Aedes albopictus* were done by traditional cloning. *Catalytic portions*: furin1 was amplified using the primers F-CAGAATccatggGGGAAGGGCGTCGTGGTGACGATTCTGGAC and R-TCGCCGaccggtACCGTAACCGAACGAGTGCGACACTCGCCG. Furin2 was amplified using primers F- CAGAATccatggGGGAAAGGCGTCGTCGTGTCGATCCT and R-TCGCCGaccggtACCATAACCGAACTTGTGGCTGACTTTCCT. Soluble clones: furin1: F-CAGAATccatggATGCCTTGTCCGGAGGCTAGCGGGAGCGGT and R-ACACCGacgcgtGAGGGCTCCGGAGTATTTGGCGCACACCGA. Furin2: F-CAGAATccatggATGTCCTACTCGGCTGCCTCGGTCCA and R-TCGCCGaccggtGTCGTGCAGTACGACCAGCTTGTTGTTCGA. The PCRs were performed using Phusion™ High-Fidelity DNA Polymerase (Thermo Scientific. F-530XL) and cellular derived cDNA as template. The PCR products were gel purified and digested with *NcoI* and *AgeI*-HF to generate sticky ends. The pMT/BiP/V5-His vector was digested with the same restriction enzymes. Digested products were subsequently purified and incubated overnight at 16 C with T4 DNA Ligase (NEB M0202). Dh5alpha cells were transformed with the ligated products and growing colonies were screened by colony PCR. Plasmids were purified with the QIAprep Spin Miniprep Kit (Qiagen 27104) and analyzed with restriction digestion. Positive clones were sent for sequencing to confirm accuracy in the sequence.

For the addition of the N-terminal region of furin1 and furin2 of *Aedes albopictus,* the predicted coding sequence was synthesized as gBlock by IDT. Furin1 = 369 bp and furin2 = 249 bp. Then, these sequences were added to the rest of the gene and cloned into the pcDNA3.0 vector and the FLAG, HiBiT and 6x-Histidine tags using the Gibson Assembly® Master Mix (NEB E2611). Transformation and screening were performed as stated above.

The addition and exchange of signal peptide of furin1 with human furin signal peptide (Hfsp) was performed by site directed mutagenesis. First, specific primers were designed using the NEBaseChanger server [128]. These primers had overhangs containing the Hfsp (=78 bp). Then, PCRs were performed using Phusion™ High-Fidelity DNA Polymerase using the original plasmid (i.e., full-length N-terminal furin1 or furin2 clones) as template. Then, the samples were treated

with *DpnI* to get rid of plasmid DNA. The remaining PCR product was phosphorylated using T4 Polynucleotide Kinase (NEB M0201) and ligated with T4 DNA ligase. Colonies were screened only by sequencing. In addition, when switching systems (i.e., from pcDNA3.0 to pMT/BiP/V5-His or PSL1180polyUB), the full-length recombinant gene: from the N-terminus until the FLAG-HiBiT-6x-Histidine tags was transferred using Gibson assembly.

Similarly, the cloning of the proprotein convertases from *Aedes aegypti* into the pcDNA3.0 (mammalian expression) and PSL1180polyUB (insect expression), were performed by Gibson assembly. However, this time the full-length genes were amplified from cDNA produced from RNA of Aag2 cells. The furin1 and furin2 clones were truncated before the transmembrane region, whereas NC2 and 7B2 were full-length.

3.3.4 Transfections, protein expression, HiBiT signal and Western Blots

The correct plasmids obtained after screening and sequencing were expanded with the HiSpeed® Plasmid Midi Kit (Qiagen 12643). Then, low passage HEK293T, S2 cells, C6/36 or Aag2 cells were transfected with lipofectamine 3000 following the manufacturer's instructions. Transfection efficiency was determined with fluorescent controls: pEGFP-puro (HEK293T), pUB-dsRED (S2, C6/36, Aag2) and inducible pMT-GFP (S2). The cells transfected with constitutive over expression promoters such as CMV (pcDNA3.0) or polyUB (PSL1180polyUB) were checked for expression 48 hours post transfection. For the S2 cells transfected with inducible promoter (pMT/BiP/V5-His or pMT/BiP/V5-His-PURO), addition of 600 µM $CuSO_4$ occurred 24 hours post transfection and protein expression was assesed 48 hours after induction.

Depending on the experiment, protein expression was determined in different formats: Western Blot and/or HiBiT activity assay. HiBiT signal was detected either intracellularly or extracellularly of transfected cells. For intracellular expression, the cells were treated with Nano-Glo® HiBiT Lytic Detection System (Promega N3030), whereas extracellular expression was determined with Nano-Glo® HiBiT Extracellular Detection System (Promega N2420). For the western blots, the samples were run on an SDS-PAGE (10% acrylamide), transferred to a nitrocellulose membrane, 0.45 µm, blocked with 5% fat-free milk in TBS-T and incubated overnight with primary antibody. For detection of 6x-Histidine, the primary antibody was mouse

monoclonal anti-polyHistidine−Alkaline Phosphatase antibody (Sigma A5588) and for detection of FLAG-tagged proteins, the primary antibody was mouse monoclonal anti-FLAG® M2 antibody (Sigma F3165). After that, three washes with TBS-T (0.05%) were performed and the membrane was then incubated for one hour with the secondary antibody: anti-mouse 680/800. The membranes were washed three more times with TBS-T and imaged on the LI-COR Odyssey instrument.
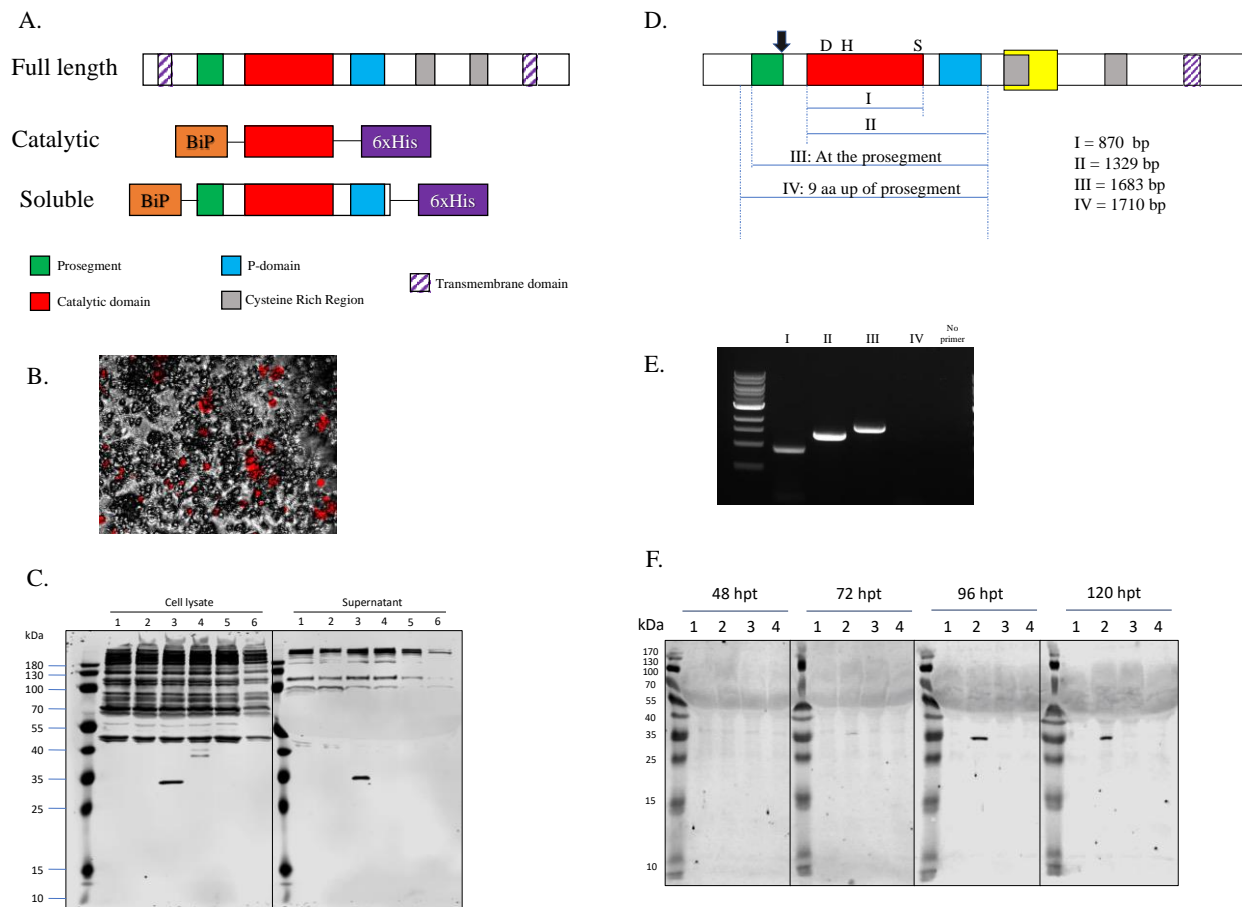
Figure 3.1 Cloning and expression of catalytic and soluble portions of furin1 and furin2 of *Aedes albopictus* in S2 cells
A. Representation of the catalytic and soluble clones in comparison with the full-length protein. Both recombinant clones had a BiP signal and a 6X-Histidine tag. B. Image showing the transfection efficiency of S2 cells with dsRed, using calcium phosphate. C. Western blot of lysis and supernatant from S2 cells transfected using calcium phosphate and induced with $CuSO_4$. (1) Blank, (2) dsRed, (3) GFP, (4) Catalytic domain of furin1, (5) Catalytic domain of furin2, (6) Soluble furin1. Only 3-6 should have 6x-Histidine tag. Expected sizes: (3) GFP = 33.6 kDa, (4) Catalytic domain of furin1 = 33.3 kDa, (5) Catalytic domain of furin2 = 31.9 kDa and (6) Soluble furin1 = 61.3 kDa. D. Representation of furin2, indicating the different set of primers used to amplify sections of the gene. E. 0.8 % Agarose gel showing the PCR products as in D. F. Western blot on supernatant at of S2 cells at different time points after transfection (1) empty plasmid, (2) GFP, (3) Soluble furin1 and (4) Soluble furin2 = 62.1 kDa.

58

3.3.5 Kinetic activity assays

To detect furin activity, the samples were incubated in 100 mM HEPES (pH 7.5 @ 25°C), 0.5% Triton X-100, 1 mM CaCl2, 1 mM ß-mercaptoethanol, 100 µM BoC-RVRR-AMC in a 100 µL volume at 30°C. For the experiments that used unpurified protein, 60 µL of sample was used and mixed with the reagents mentioned above, whereas for the purified sample experiments, 10 µL were used. In addition, the experiments that required different conditions, such as calcium and pH, were performed as needed: first, lower pHs (5.5, 6.0 and 6.5) required the presence of sodium acetate instead of HEPES.

3.3.6 Size exclusion chromatography

The Superdex 200 Increase 10/300 G column was used for the gel filtration. Briefly, samples were concentrated to a final volume of 500 µL. The sample was injected into the pre-washed and pre-equilibrated column with 20 mM MES, 150 mM NaCl, 2 mM CaCl$_2$. The pressure limit was 2.7 pa, and the flow rate was at about 0.5 mL/min.
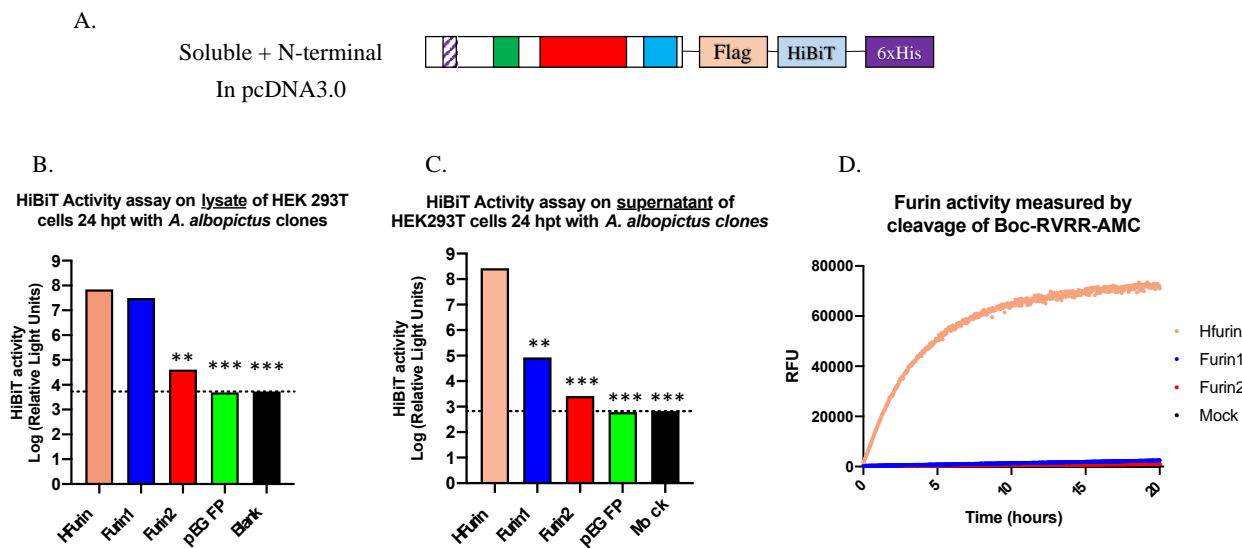
Figure 3.2 Cloning and expression of soluble portions of furin1 and furin2 of *Aedes albopictus* in HEK293T cells
A. Representation of the soluble portion plus the synthetized N-terminal region and the three tags included in this occasion: FLAG, HiBiT and 6xHis. B and C. HiBiT signal from lysate (B) and supernatant (C) of HEK293T cells 24 hours post transfection. D. Furin activity measured from supernatant of transfected cells. Significance was calculated with a two-tailed t-test compared to the corresponding levels of human furin, P<0.05. N=1

3.4　Results

3.4.1　Part 1. Attempts to express and purify proprotein convertases of *Aedes albopictus.*

C6/36 cells were initially used as the source of RNA for the production of cDNA for the different plasmid clones. However, multiple difficulties were encountered and this led to different approaches, including three expression systems, synthesis of the N-terminal region, exchange of signal peptide sequences and inclusion of additional detection and purification tags.

Expression of domains/regions of furin1 and furin2 of Aedes albopictus

C6/36 cells were selected because they are a robust cell line, are accessible in the lab and are routinely used for multiple virology assays including preparation of viral samples for structural biology [129]. The initial clones created were: 1) the catalytic portion, which included the predicted catalytic domain and 2) the soluble portion which included the pro-segment, the catalytic domain and the P-domain (Figure 3.1a). Both of these constructs were cloned into the pMT/BiP/V5-His plasmid in frame with the BiP signal sequence at the N-terminal and the 6x-Histidine tag at the C-terminal region. At this point, the soluble clone of furin2 was not produced given the difficulties amplifying the N-terminal region which will be discussed below.

Drosophila S2 cells were initially chosen for two reasons: 1) when stable S2 cells are produced, they can be scaled up and produce large amount of protein and, 2) because *Drosophila melanogaster* is closely related to *Aedes albopictus* and this can potentially account for post-translational modifications such as glycosylation. The transfection of S2 cells was initially attempted using the calcium phosphate method, and the transfection efficiency was measured using the fluorescent plasmid dsRed, which is about 30% (Figure 3.1b). Then, an initial transfection with the catalytic domain of furin1 and furin2, and the soluble construct of furin1 was performed and induced with $CuSO_4$ 48 hours post transfection. Following 72 hours post induction, the supernatant of transfected cells was collected and analyzed for expression on a western blot against 6x-His tag. Only the inducible GFP control was detected under the experimental conditions (Figure 3.1c).

The full-length N-terminal region was never amplified for either furin1 or for furin2. However, furin2 was more challenging because none of the primers would work at any region of

the N-terminal section., but the positive control, which was the catalytic domain, was always a successful amplification. Therefore, different PCR lengths were attempted using combination of primers (Figures 3.1.d and 3.1.e). The conserved regions were easily amplified: catalytic domain, P-domain and pro-segment. However, when the forward primer was placed 27 nucleotides upstream of the predicted pro-segment, the PCR failed. Further, a soluble version of furin2 was created using the cDNA at the beginning of the pro-segment and in frame with the BiP sequence and the 6x-His tag. Then, the new soluble furin2 construct was tested along with GFP and soluble furin1 for expression at different points after transfection (Figure 3.1f). Once again, none of the constructs were expressed at any time point, but 96 hours post transfection (i.e., 48 hours post induction) was identified as the highest expression time point for the GFP control.

Table 3.1 Signal Peptide Prediction

Software	Hs_furin	Aa_furin1	Aa_furin2
SignalP 4.1	Yes (26-27)	No	No
SignalP 3.0	Yes (26-27)	Yes (55-56)	Yes (22-23)
PrediSi	Yes (21)	Yes (57)	No
Phobius	Yes (24)	No (two transmembrane)	Yes (22)

Addition of synthesized N-terminal region, FLAG-tag and HiBiT signal for expression in HEK293T cells

Several studies have suggested that signal peptides have functions other than the directionality of proteins through the secretory pathway, including chaperone functions [130], [131], and in some cases they are not interchangeable [132]. However, none of the mosquito furins have a canonical signal peptide predicted by online software. For this reason, we thought that the

61

signal peptide of mosquitoes might have specific roles to the protein and was best to keep it during expression rather than replacing it with a BiP sequence. Nevertheless, as mentioned above, PCRs including primers in the N-terminal region always failed. To address this problem, the region was synthetized as gBlocks by IDT (furin1 = 369 bp and furin2 = 249 bp) and added to the construct (Figure 3.2a). In addition, the transfection efficiency in S2 cells was extremely low, therefore the constructs were transferred to the pcDNA3.0 plasmid under the CMV promoter for expression in HEK293T cells where the transfection efficiency is more than 80%. Finally, these clones also included the FLAG-tag and HiBiT, in addition to 6xHis sequence for detection purposes and further purification purposes.
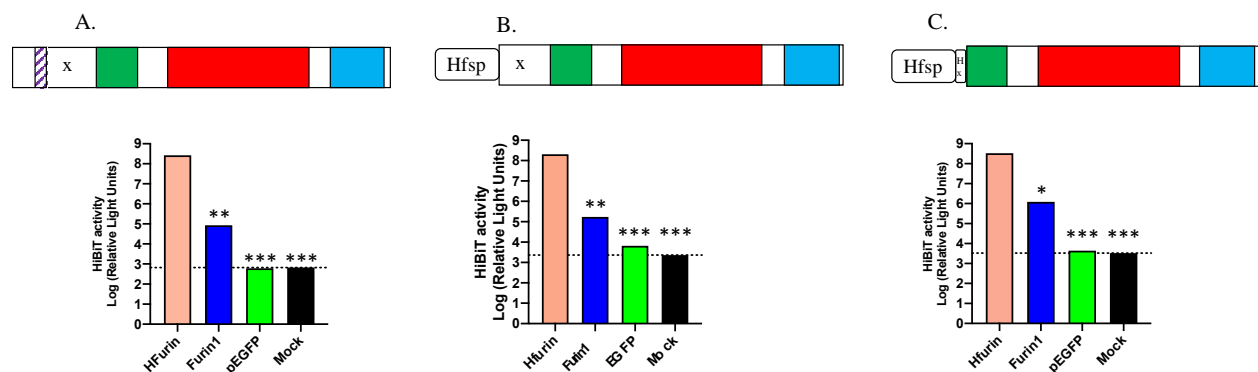
Figure 3.3 Humanization of the N-terminal of furin1 of Aedes albopictus
A. Original clone with the synthetized N-terminal portion. B. Replacement of the predicted signal peptide of furin1 with the human furin signal peptide. N-terminal region completely humanized: with both signal peptide and the "X" region of human furin. Each experiment shows the diagram and the corresponding HiBiT signal obtained in supernatant of HEK293T cells 24 hours post transfection using Lipofectamine 2000. Hfurin was used as a positive control for secretion. pEGFP was used as a transfection control and a negative control for secretion. Hfsp: Human furin signal peptide. Hx: is the X region in the human furin protease. Significance was calculated with a two-tailed t-test compared to the corresponding levels of human furin, $P<0.05$. N=1

At this point, the human furin construct was also generated to serve as a control. The human furin construct included the secreted version, truncated at Ala574. The human furin construct also contained the three tags: FLAG, HiBiT and 6xHis. The addition of the HiBiT signal was definitely successful in the sense that for the first time it provided a high-throughput method to detect expression without relying on western blot. Human furin and furin1 were both detected intracellularly (Figure 3.2b), but only human furin was also detected in the extracellular space (Figure 3.2c). Furin2 did not even get expressed intracellularly. Then, the supernatant was tested for activity using the BoC-RVRR-AMC substrate. As expected, human furin showed the most activity over a period of 20 hours, however, little or no activity was detected in the supernatant of

the other samples (Figure 3.2d), confirming that there was no active protein present. These results suggest that furin1 was getting expressed but not secreted. This set the stage for the next set of experiments attempting to improve the secretion.

Humanizing the N-terminal region of furin1 of Aedes albopictus

It is important to mention that neither furin1 nor furin2 have canonical signal peptides, instead they have a hydrophobic sequence in the N-terminal that might function as a signal peptide in a unique way (Table 3.1). To address the secretion issue of furin1 in HEK293T cells, we thought that changing the sequence would make a difference in mammalian cells. And given the previous results obtained with human furin, we decided to remove the predicted signal peptide of furin1 and replace it with the signal peptide of human furin (Figure 3.3b). This change did not improve secretion and did not reduce it when compared to the original furin1 (Figure 3.3a). Then, we analyzed the sequence in further detail and identified a long non-conserved region between the predicted signal peptide and the prosegment of furin1. This region was named as the "X" region and was exchanged with the human furin x region "X" which is much shorter. Once again, the change did not help with the secretion of furin1 (Figure 3.3c). These experiments suggest that the replacement with the signal peptide of human furin does not improve the secretion of furin1.
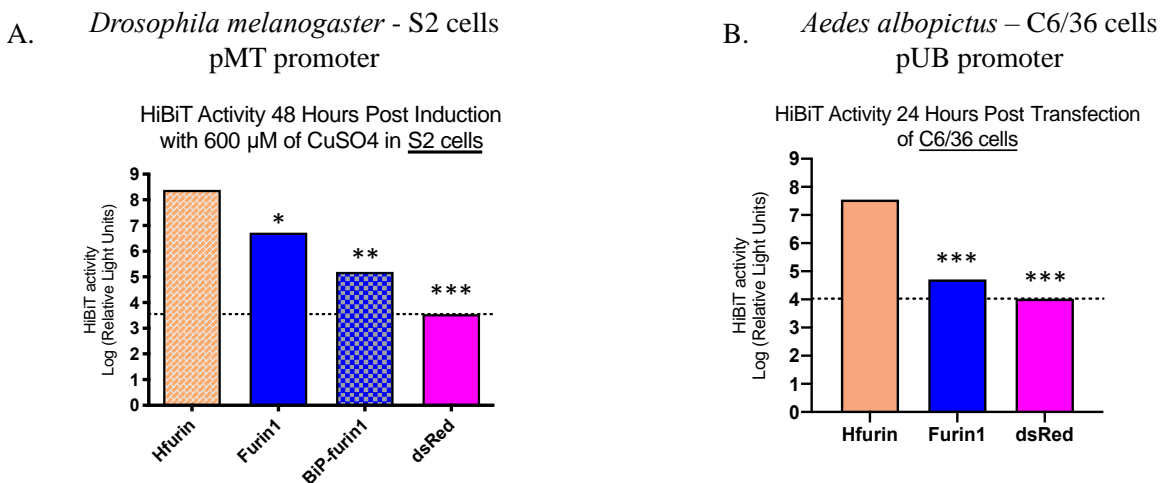
Figure 3.4 Expression and secretion of full length N-terminal furin1 of Aedes albopictus in S2 cells and C6/36 cells

A. HiBiT signal from supernatant of S2 cells 48 hours post induction with CuSO4. This figure compares the secretion levels of the original furin1 against the furin1 plus the BiP signal. B. HiBiT signal from supernatant of C6/36 cells 24 hours post transfection. Both experiments use human furin as positive secretion control and dsRED as transfection control and negative control for HiBiT signal (i.e., dsRed does not have HiBiT tag). Significance was calculated with a two-tailed t-test, $P<0.05$. N=1

Expression of furin1 with the full-length N-terminal region in S2 cells and C6/36 cells

Post translational modifications can be necessary for ensuring the average life of a protein and can be host and even cell specific. For that reason, we attempted to express furin1 with its full-length N-terminal region in the insect systems: Drosophila S2 cells and C6/36 (Figure 3.4). Human furin got secreted efficiently in both systems, but furin1 did not reach the same levels as human furin. Interestingly, when testing removing the original signal peptide of furin1 and replacing it with BiP, it decreased the secretion further (Figure 3.4a). Also, furin1 did not get secreted at all in C6/36 cells, opposite to what was expected because C6/36 are the original source of the RNA for making the clones (Figure 3.4b).
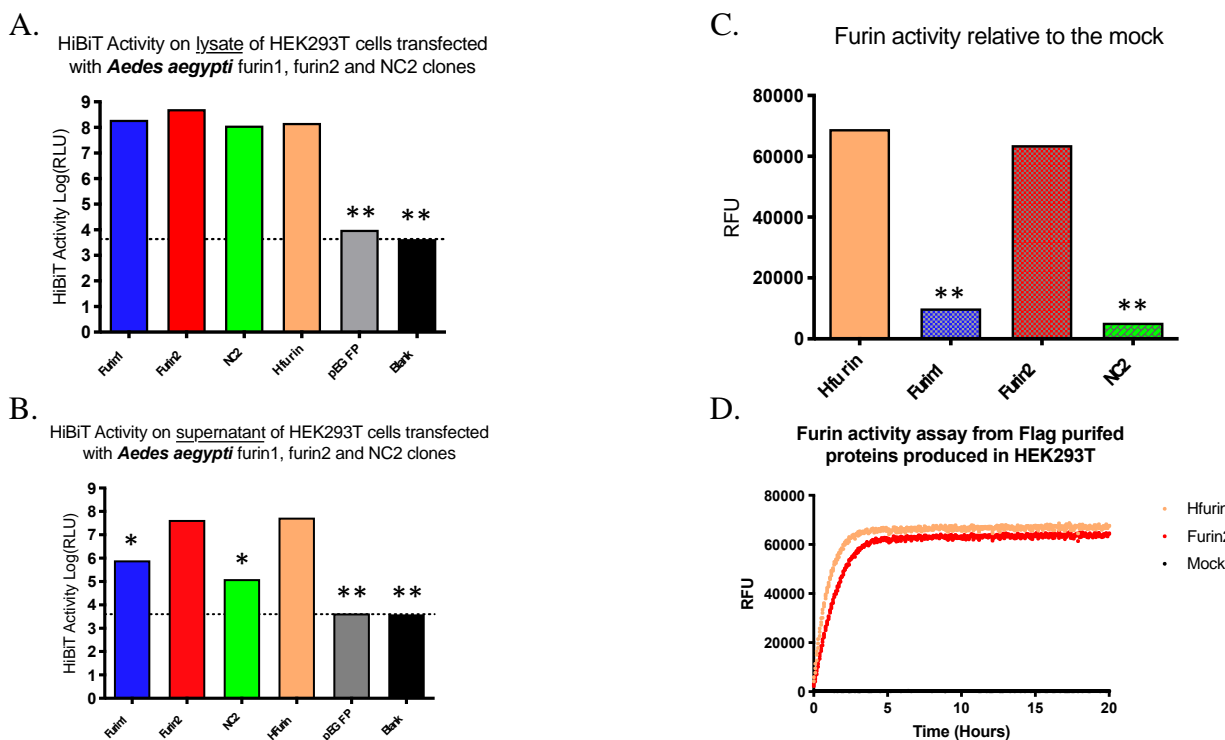
Figure 3.5 Expression and activity of the PCSKs of *Aedes aegypti* in HEK293T cells
A and B. HiBiT signal from lysate (A) and supernatant (B) of HEK293T cells 24 hours post transfection. C. Furin activity measured in relative fluorescent units from immunoprecipitated protein from supernatant of cells. D. Kinetic activity assay of human furin and furin2. Significance was calculated with a two-tailed t-test compared to the corresponding levels of human furin, P<0.05. N=1

At this point, none of the attempts to express and secrete furin1 of furin2 of *Aedes albopictus* were successful. However, there was an uncertainty about the N-terminal region. First, we learned that it is important and should not be replaced (Figures 3.3 and 3.4a). Second, the N-

terminal region was never amplified by PCR and needed to be synthetized. This last statement is concerning because the reported sequence in the genome of *Aedes albopictus* (canu_80X_arrow2.2) might be erroneous given that this genome is assembled in contigs, rather than scaffolds or chromosomes. In addition, we found out that these sequences had been *DISCONTINUED from NCBI because the model on which they were based was not predicted in a later annotation.* For this reason, we decided to move to the proprotein convertases of *Aedes aegypti,* which have a better annotated genome, the AaegL5.0.
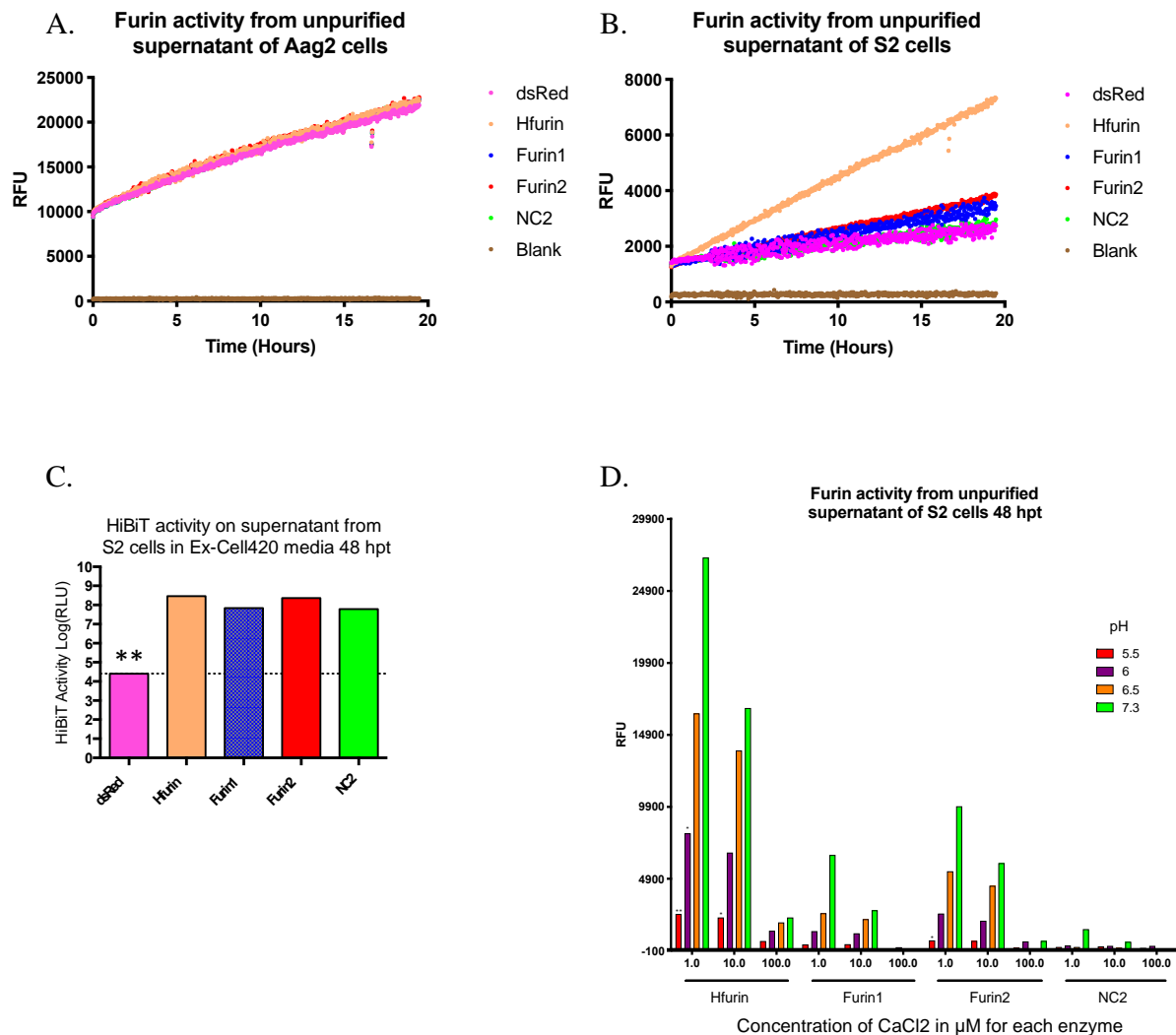
Figure 3.6 Expression and activity of proprotein convertases of Aedes aegypti in transient transfected S2 cells under the constitutively expressed poly-ubiquitin promoter (pUB)
A and B. Kinetic activity assay from supernatant of S2 cells 48 hours post transfection in 10% FBS (A) or ExCell420 serum-free media (B). C. HiBiT signal from supernatant of S2 cells 48 hours post transfection in ExCell420 media. D. Activity of unpurified proprotein convertases of Aedes aegypti at different pHs and concentrations of calcium. dsRed represents a flurescent transfection control but with no HiBiT tag. Significance was calculated with a two-tailed t-test compared to the corresponding levels of human furin (C) or to the neutral pH 7.5 (D), P<0.05. N=1

3.4.2 Part 2. Expression and purifications attempt of PCs of *Aedes aegypti* – non-codon optimized.

In this section, we used Aag2 cells as the source of genetic material of *Aedes aegypti* for cloning purposes. Once again, we tested expression in HEK293T cells and also in S2 cells. However, in both cases the expression attempts were made with the constitutively expressed pUB promoter. In addition, we also tested for the secretion of NC2 when co-expressed with the peptide 7B2, which is suspected to be required for adequate folding and improvement of secretion.

Expression and activity of proprotein convertases of Aedes aegypti in HEK293T cells.

The first attempt of expressing proprotein convertases of *Aedes aegypti* was performed in HEK293T cells because this is a robust cell line with a high transfection efficiency. The coding sequence for furin1, furin2, NC2 and the human furin control were cloned under the control of the CMV promoter for constitutive over expression in mammalian cells. The clones were truncated right before the transmembrane region, furin1: Phe1037, furin2: Gly1025 and Hfurin: Ala574, or NC2 included the entire coding sequence: Met640. In addition, all the clones included FLAG-HiBiT-6XHistidine tags at the C-terminal region for detection and purification.
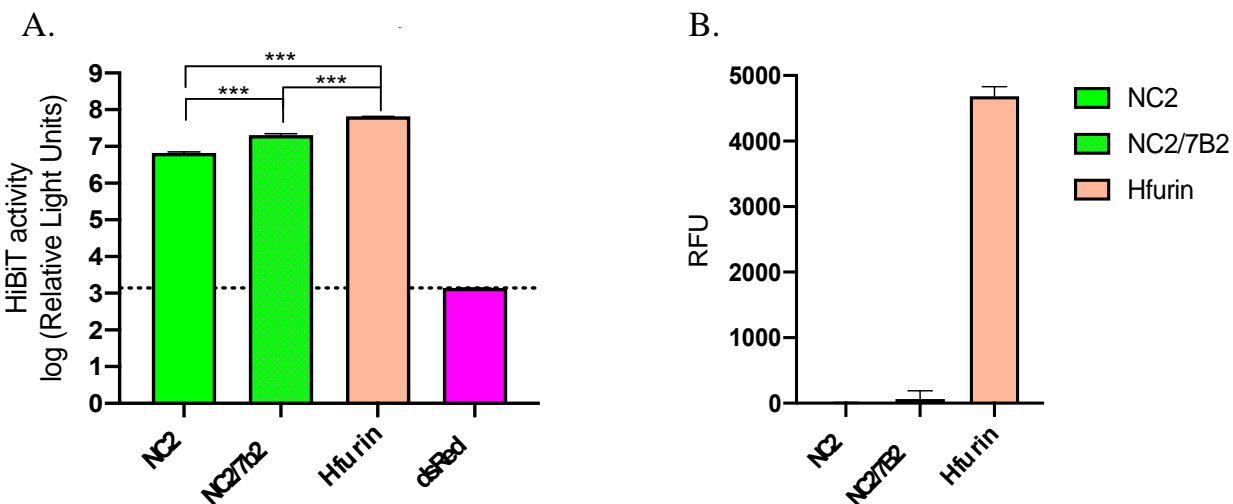
Figure 3.7 Co-expression of NC2 with the 7B2
A. HiBiT signal from supernatant of S2 cells 48 hours post transfection. B. Furin activity from the supernatant of transfected S2 cells. Significance was calculated with a two-tailed t-test compared to the corresponding levels of human furin, P<0.05. N=2

Intracellular expression was similar among the different clones (Figure 3.5a). However, protein detected in the extracellular media, also referred as the secreted protein, was identified at high levels for human furin and furin2 (Figure 3.5b). Furin1 and NC2 were secreted too, but the HiBiT signal was almost 100-fold reduced. In addition, when tested for furin activity in supernatant of transfected cells, only human furin and furin2 showed the most activity (Figure 3.5c), which correlates with the amount of protein detected by HiBiT signal. Next, we proceeded to purify protein by FLAG-immunoprecipitation of both human furin and furin2, and performed activity assay (Figure 3.5d). Activity was shown to be similar to human furin under these experimental conditions.
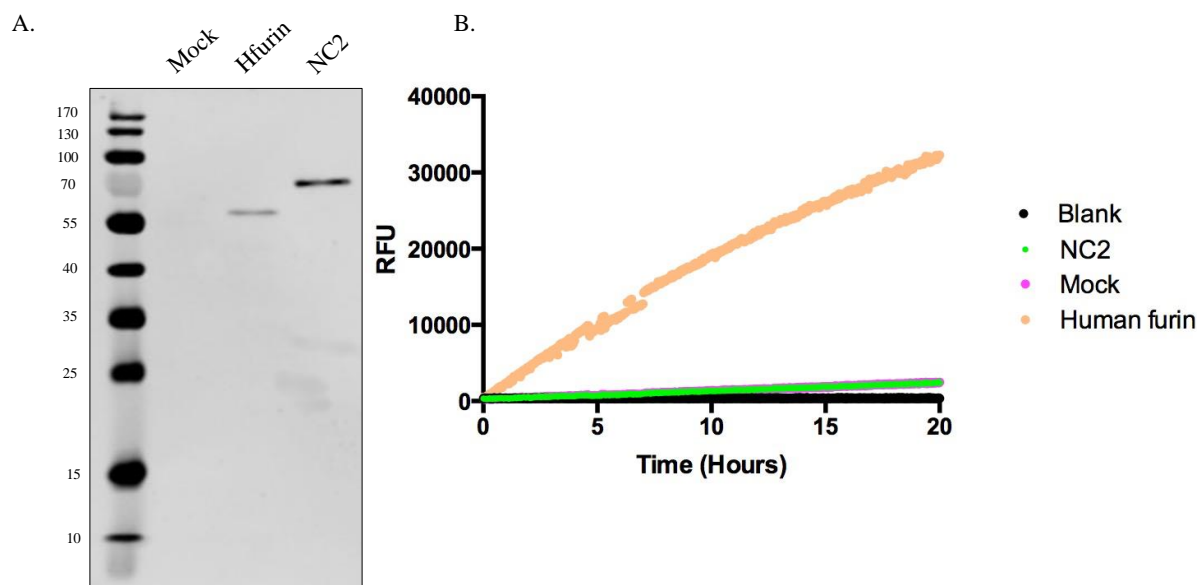
Figure 3.8 Purified NC2 from Aag2 cells is an inactive protease under these experimental conditions
A. Western-blot of FLAG-immunoprecipitated protein. Primary antibody: mouse anti-FLAG. Secondary antibody: goat-anti mouse wavelength at 680. Expected sizes: Human furin = 63.26 kDa, NC2 = 70.9 kDa. B. Activity assay using the purified samples and the BoC-RVRR-AMC substrate.

Expression and activity of proprotein convertases of Aedes aegypti in transient transfected S2 cells under the constitutively expressed poly-ubiquitin promoter (pUB).

To address the secretion issue of furin1 and NC2, we attempted to express proprotein convertases of *Aedes aegypti* back in the insect system, *Drosophila* S2 cells. In addition, for rapid screening of expression and activity, the genes were cloned under the polyubiquitin promoter for rapid constitutive overexpression. First, we aimed to check for activity from supernatant of S2 cells supplemented with 10 % FBS. However, in this experiment there was too much background

to identify the activity coming from recombinant protein (Figure 3.6a). Therefore, we attempted another experiment in which the media was exchanged with ExCell420 serum-free 24 hours after transfection. In this experiment, it was easy to identify the activity coming exclusively from Hfurin, and to some extent from furin1 and furin2 (Figure 3.6b). Unlike HEK293T cells, the HiBiT signal was detectable for all the proteases at similar levels (Figure 3.6c). And, using unpurified supernatant of S2 cells in serum-free, we were able to test different biochemical conditions for activity. All proteases, including human furin, showed highest activity at neutral pH (=7.3), and 1 mM of $CaCl_2$ (Figure 3.6d). These results were significant, because for the first time they showed that similar to other calcium-dependent proteases, the experimental conditions here showed that mosquito proprotein convertases were affected by the concentration of calcium and their activity was different among pH gradients.

The next question that we were addressing was whether NC2 needed a co-factor for being expressed and secreted. Previous studies [133], [134] have suggested that PC2 and amontillado require to be co-expressed with the peptide 7B2. Thus, we cloned the 7B2 sequence of *Aedes aegypti* and co-expressed with NC2. As expected, when co-expressed, there is an increase in the HiBiT signal in the extracellular space of almost a log; however, it was still lower than human furin (Figure 3.7a). The supernatant of these transfected cells was used for activity assays, and the co-expression with 7B2 did not improve the activity (Figure 3.7b). Then, we questioned whether Drosophila S2 cells lack additional cofactors that were necessary for folding/activation of NC2. Therefore, we overexpressed NC2 and the human furin control in Aag2 cells. These cells do not survive in serum-free medium, therefore the proteins were FLAG-immunoprecipitated. Since the expression in these cells is extremely low, the purified proteins had to be checked on a western blot (Figure 3.8a). However, when tested for activity, NC2 did not show activity (Figure 3.8). Combined, these results suggest that NC2 is not with this substrate.

Expression and purification of proprotein convertases of Aedes aegypti from stable transfected S2 cells with the pUB promoter and the selection marker pCo-PURO.

Next, I attempted to produce stable transfected S2 cells to produce larger amounts of furin1 and furin2. For that, S2 cells were co-transfected with the selection plasmid pCo-PURO and treated with puromycin to remove un-transfected cells. The ratio of recombinant protein plasmid: pCo-

PURO was 20:1. Cells were treated with or without puromycin, and the expression was determined by HiBiT signal 6 days after incubation in a shaker at 120 rpm. The background signal was determined in the mock control (i.e. cells transfected with dsRed) and two different media: old and new. Conditioned medium, refers to medium in which untransfected cells were growing for 6 days. In both, instances, cells with and without puromycin, showed the highest expression for furin2 (Figure 3.9a). Cell density was also determined at the time of sample collection and was used to estimate the HiBiT signal coming out per cell with and without puromycin treatment (Figure 3.9c). As expected, the cell density was higher in cells without puromycin treatment (Figure 3.9b). The HiBiT signal was higher for furin2 in cells without puromycin treatment, but with puromycin, the signal for furin1 increased by almost three-fold (Figures 3.9c and 3.9d). HiBiT signal from human furin and furin2 did not change in cells treated with puromycin suggesting that S2-furin1 cells were expressing both the puromycin selection marker and the recombinant furin1. Then, we attempted to do a FLAG-immunoprecipitation assay and found that human furin was easily purified and detectable by SDS-PAGE and instant blue staining (Figure 3.9e), however, furin1 and furin2 were hardly, if at all, detected. For this reason, we performed a western blot and found that only human furin and furin2 were present in the purified fraction. Combined, these results suggest that the expression of proprotein convertases in S2 cells is not optimal and might require a better platform for increasing the amount of protein produced. Therefore, the next step was to produce codon optimized clones under the regulation of a more robust promoter in *drosophila*, the inducible metallothionein promoter.

3.4.3   Part 3. Expression and purifications attempt of PCs of *Aedes aegypti* – codon optimized.

The last section of this chapter included the generation of codon-optimized clones for expression in Drosophila S2 cells in the pMT-Puro plasmid, which contains both the metallothionein promoter and the puromycin selection marker, removing the need of co-transfection with other plasmids. In this section, the purification attempts are discussed as part of an optimization strategy as well as an activity assay to demonstrate serine-protease activity.

Expression and purification of codon-optimized proprotein convertases of Aedes aegypti
in S2 cells under the inducible promoter pMT.

The codon optimized sequences of furin1, furin2 and NC2 were synthetized and cloned into the pMT-PURO plasmid by Biobasic Inc. The plasmids were transfected into S2 cells. 48 hours after transfection, the S2 cells were treated with puromycin and selected for two weeks. After that, the cells were expanded. Once cells were expanded, they were induced with $CuSO_4$. The supernatant was collected and processed for further Ni-NTA purification (Figure 3.10a). Before the purification, the supernatant was checked for detection of HiBiT signal. This was the first time that HiBiT signal was observed to be equivalent among the different proteases (Figure 3.10b), which indicated that the codon optimization contributed to more expression. The induced supernatant was incubated overnight with Ni-NTA beads and next morning the beads were washed and eluted. The eluted samples were then checked on a SDS-PAGE and stained with instant blue (Figure 3.10c). Once again, a new milestone had been achieved at this point: this was the first time that recombinant furin1 and furin2 were detectable in a gel without doing a western blot. However, as seen in the gel, the samples were not pure and there were likely other contaminants in the sample. A kinetic activity assay was performed and only human furin and furin1 showed activity (Figure 3.10d).
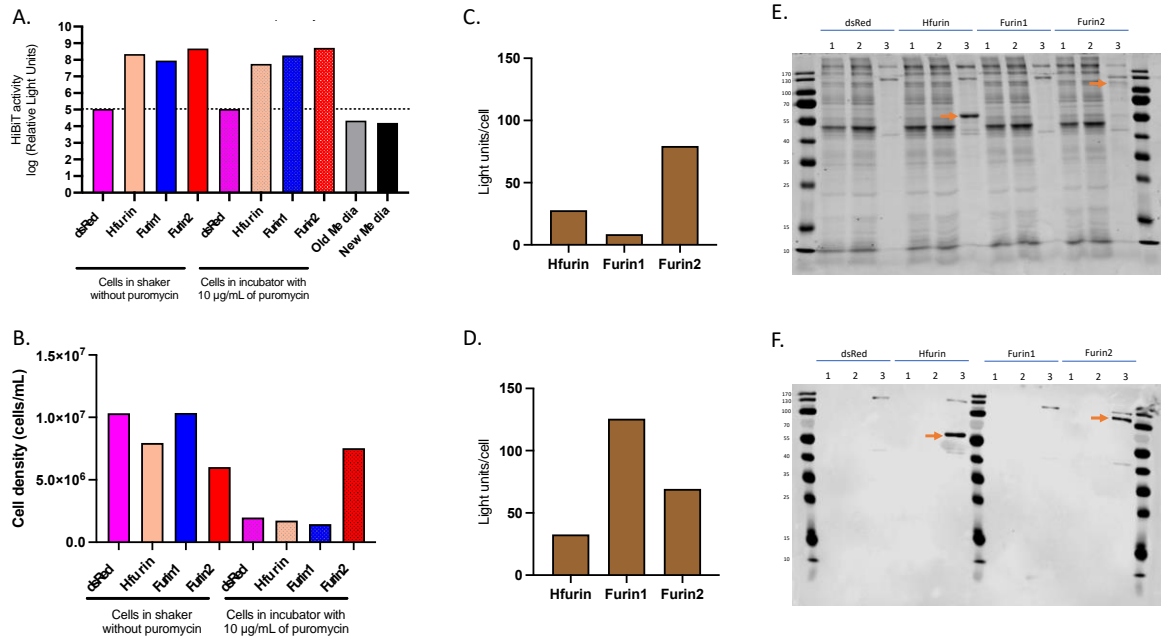
Figure 3.9 Expression and purification of proprotein convertases of Aedes aegypti from stable transfected S2 cells with the pUB promoter and the selection marker pCo-PURO

A.HiBiT signal from stable transfected S2 cells. S2 cells were co-transfected with the plasmid containing the proprotein convertase under the constitutively expressed polyubiquitin promoter (pUB) and the selection marker pCo-PURO. The signal was measured at 6 days after constant rotation at 120 rpm at 27 C, either with or without puromycin (10 µg/mL). B. Cell density at 6 days after constant rotation of stable transfected S2 cells. C and D. HiBiT signal relative to the amount of cells measured as light units per cell with (C) and without (D) puromycin. E and F. SDS-PAGE (E) and Western Blot (F) of immunoprecipitation steps with FLAG agarose beads of human furin, furin1 and furin2. (1) Unprocessed supernatant of S2 cells, (2) flow-through after overnight binding to the FLAG agarose beads and (3) samples eluted with 150 µg/mL of the FLAG-peptide. The orange arrows point at bands in the expected size. N=1

Given that the furin1 histidine purified sample was not completely pure, there was uncertainty whether the activity corresponded only to recombinant protein. Therefore, an additional purification step was included. Basically, the imidazole-eluted sample was immunoprecipitated with FLAG-agarose beads. The protein was eluted with flag peptide and then assessed for purity on an SDS-PAGE and silver staining, which is much more sensitive than instant blue (Figure 3.11a). Purified furin1 showed to be pure. However the amount of protein recovered was extremely low. Nevertheless, this purified sample of furin1 was used assess the dependency of calcium for activity. For that, a kinetic activity assay was performed in different conditions: with and without calcium and with chelators EDTA and EGTA (Figure 3.11b). This experiment demonstrated for the first time that furin1 is dependent on calcium, and potentially other metal ions. As observed in this activity assay, the control with furin displays high activity which eventually gets saturated and is not readable after 15 hours. The treatment without calcium still

showed some activity but was dramatically reduced when compared to the control. And finally, the treatments with either EDTA or EGTA makes the protease completely inactive.
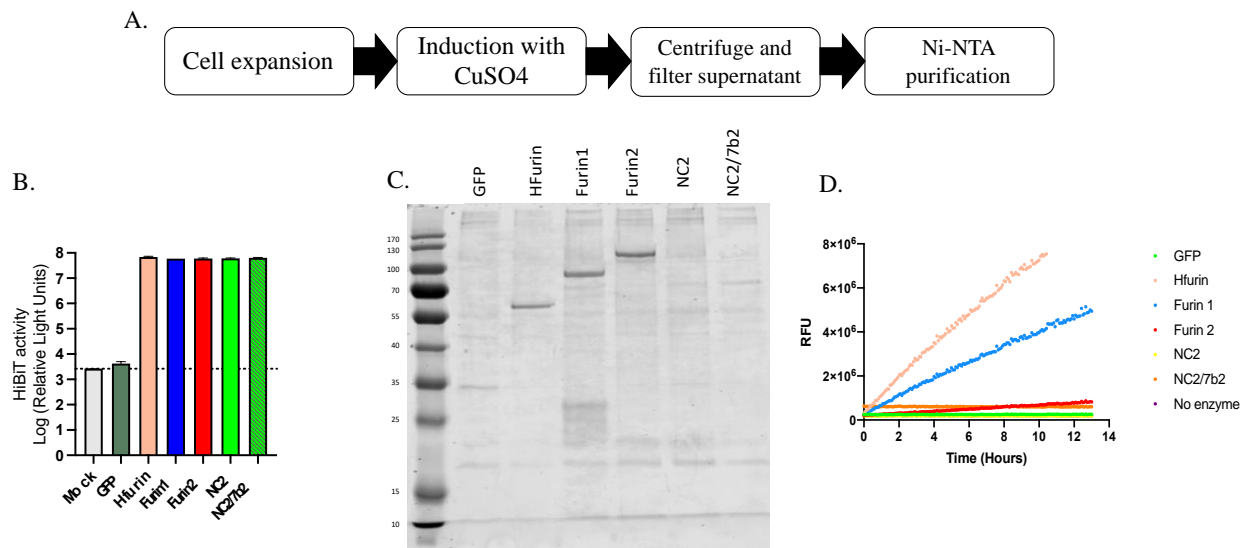
Figure 3.10 Expression and purification of codon-optimized proprotein convertases of Aedes aegypti in S2 cells under the inducible promoter pMT
A. Diagram depicting the process for expression and purification. B. HiBiT signal from stable transfected S2 cells. The S2 cells were transfected with the codon optimized genes in the pMT-PURO plasmid, which contains the inducible promoter and the selection marker simultaneously. Cells were selected with 10 µg/mL of puromycin for 2 weeks and then expanded and induced with 600 µM of CuSO4. C. SDS-PAGE of purified proteins by Ni-NTA batch purification. D. Furin activity of the Ni-NTA purified samples. Expected sizes: Hfurin= 53.71 kDa, furin1= 84.2 kDa, furin2: 89.5 kDa, NC2: 70.9 kDa. N=2

Batch Ni-NTA purification and size exclusion chromatography of furin1.

The two-step purification - Ni-NTA plus FLAG-immunoprecipitation - produced very low amount of protein as observed in the silver stain (Figure 3.11a). Therefore, another purification attempts were tested. Given that the binding capacity of the FLAG-immunoprecipitation beads is only 0.6 mg/mL, then this step was avoided. Instead, an additional purification step was tested: Size Exclusion Chromatography. Different buffers were tested to optimize the separation of furin1 in the Superdex 200 column. The best one was 20 mM MES, 150 mM NaCl, 2 mM CaCl2. After the Ni-NTA batch purification, the sample was concentrated and injected into the column. The chromatograph shows that there are two peaks (Figure 3.12a), suggesting a potential contaminant in the second peak. According to the SEC standard curve (Figure 3.12b), furin1 should be eluted from the column at 13-14 mL, which corresponds to the first peak (Figure 3.12a). However, other

72

fractions were also collected and analyzed. The silver stain indicated that the protein was not as pure, but the purest sample with the strongest band at the expected size was fraction B11. Additional bands were also detected, with the strongest potential contaminant at fraction B9 (Figure 3.12c). Then, a western blot was performed to determine whether the suspected band corresponded to the purified protein. Surprisingly, all bands reacted to the western blot, including the contaminants (Figure 3.12d), suggesting that the protein might have gone further cleavage or degradation events. Finally, an activity assay was performed with all the fractions, and contrary to what was expected, the fraction with the most activity was B9 (Figure 3.12e), which indicates that there is more furin1 present in this sample and the additional band observed at a lower molecular weight is the same protein.
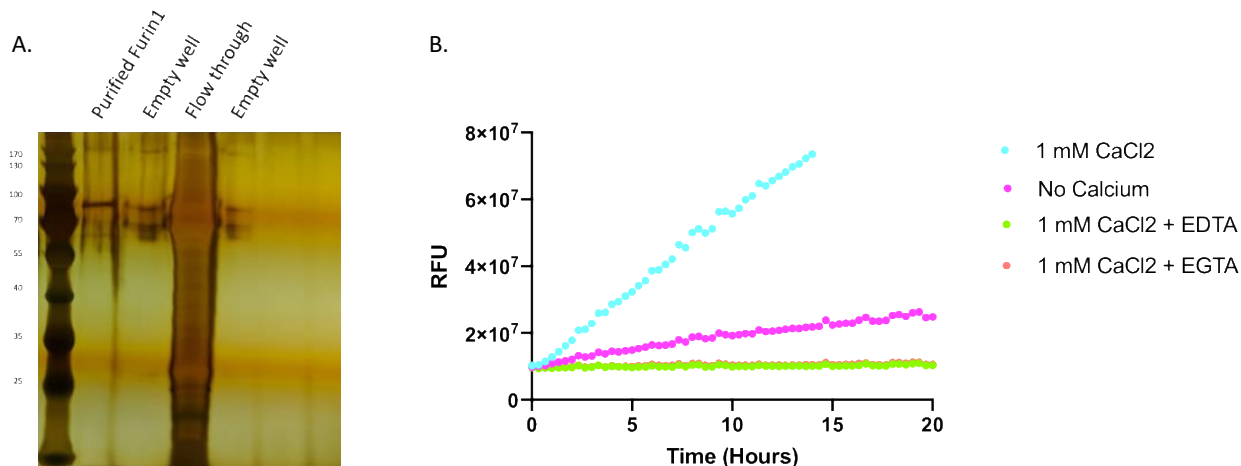
Figure 3.11 Furin1 shows typical serine protease activity and dependance on CaCl$_2$
A. Silver staining of double purified furin1. After batch purification with Ni-NTA, furin1 was immunoprecipitated with the FLAG-agarose beads. Purified furin1 is the eluted protein with 150 µg/mL of FLAG peptide. The flow through represents the sample unbound to the FLAG agarose beads. Empty wells represent background signal of the silver stain. B. Kinetic assay of double purified furin1 in the presence of CaCl2, EDTA, EGTA or no calcium. N=2

3.5 Discussion

Expression and purification of recombinant proteins is an important step in virology research. However, in most instances, protein expression needs optimization, especially true for unknown proteins. The main aspects that need to be considered when producing an unknown protein are the region in which truncations are made, the expression system and the purification conditions. Expression of individual domains of a protein is often used for antigenic production and development of antibodies, however in several instances the proteins can collapse because

domains might need other parts of the protein for proper folding [135]. On the other hand, many proteins can be expressed in *E. coli* because it provides the most robust system: it can be scaled up and is inexpensive. However, many proteins require specific post-translational modifications that *E. coli* does not provide.
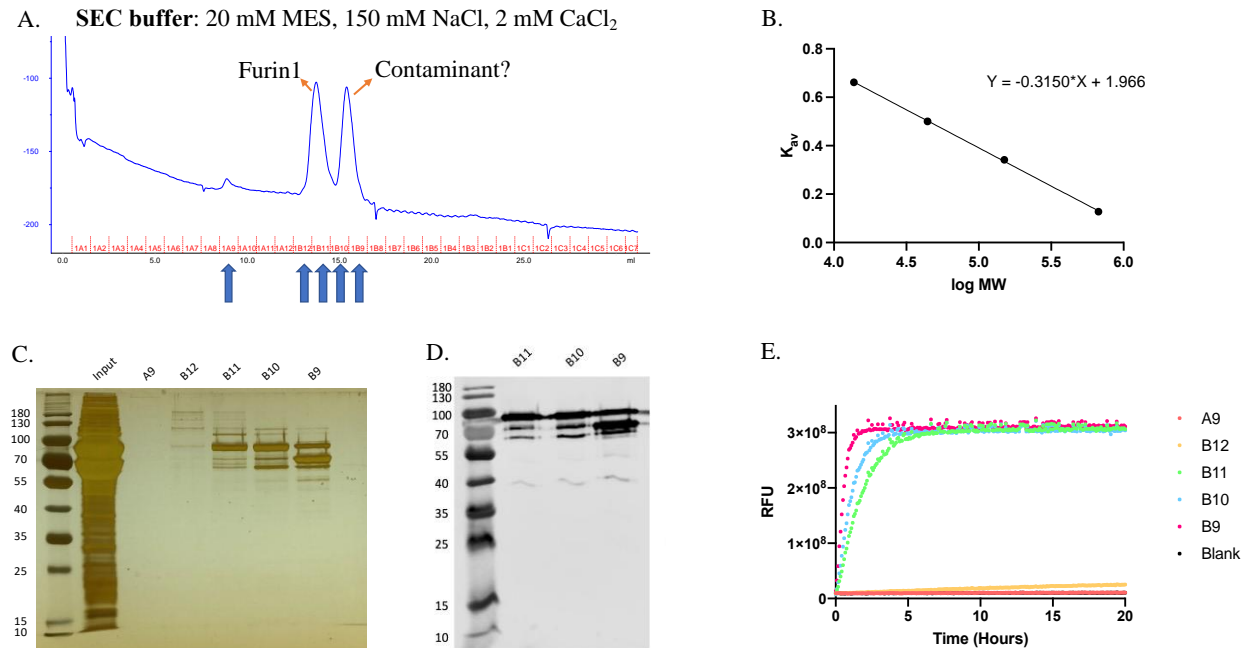
Figure 3.12 Batch Ni-NTA purification and SEC of furin1. 300 mL of supernatant of stable transfected S2 cells was used for batch purification with the Ni-NTA beads
Then, the sample was concentrated and used for further analytical steps. A. Size exclusion chromatography: the blue arrows indicate the samples analyzed on silver stain (C), Western Blot (D) and kinetic assay (E). B. Size exclusion chromatography standard curve. C. Silver stain. The input corresponds to the histidine purified sample before size exclusion chromatography. Expected size: furin1= 84.2 kDa. D. Western blot. The membrane was blotted with anti-Flag antibody and then ant-Mouse 680 as secondary antibody. E. Kinetic activity assay. The fractions of the SEC were tested on a furin activity assay using the substrate BoC-RVRR-AMC. N=2

In the present chapter, expression of mosquito proteases was attempted in different formats. The aspects that were taken into account for expression were 1) the solubility of the protein, which meant to produce globular portion(s) of the protein without transmembrane sections, 2) the expression system and 3) the secretion, for purification purposes. The first attempt was to express furin1 and furin2 of *Aedes albopictus.* Initially, the catalytic domain of furin1 or furin2 alone was cloned into the pMT vector for expression in S2 cells. The S2 cells were selected because *Drosophila* is evolutionary close to *Aedes*, they both belong to the order *Diptera*. However, this attempt did not result in detectable protein by western blot (Figure 3.1). The reason might be that

the catalytic domain by itself was unstable and did not get properly folded and was degraded upon translation. Nevertheless, the expression of the soluble truncated protein with the prosegment and the P-domain, but without the N-terminal did not produce protein either, inferring that this portion of the protein is important for expression. The next step was to add the full-length N-terminal section, however PCR amplification of this region from cell derived cDNA was not possible and required the synthesis, which improved the detection of intracellular furin1, but no of furin2 (Figure 3.2). No secretion of protein was detectable.

The fact that the N-terminal region was not amplifiable by any of the PCR troubleshooting, suggested that the primers were not annealing properly, which might suggest that the sequence was incorrect. The ends of gene models are known for having errors in annotation, and it is especially true for genomes that were assembled from short reads as contigs, which is the case of the genome of *Aedes albopictus*. This explains why the protein with the full-length synthesized N-terminal was not expressed nor secreted in any of the systems: S2, HEK293T or C6/36. In support of this idea, NCBI removed the sequences of furin1, furin2 and NC2 from GenBank explaining that the model on which they were based was not predicted in a later annotation.

In contrast, the newest genome of *Aedes aegypti*, AaegL5.0 has longer reads and it is assembled into three chromosomes rather than thousands of contigs. This makes the genome of *Aedes aegypti* to be the best annotated. As a consequence, the proprotein convertases of *Aedes aegypti* instead of *Aedes albopictus* were further pursued for expression purposes. At a first glance, the expression attempt in HEK293T cells, all proprotein convertases were expressed intracellularly, but only furin2 was successfully secreted and active. However, when expressed in S2 cells, all proteases were secreted and both furin1 and furin2 showed activity, but not NC2. This suggests, that HEK293T cells lack element(s) that influence the behavior of furin1.

Despite success, it is important to mention that the expression levels of PCs of *Aedes aegypti* in S2 cells were still very low, and further challenges needed to be addressed before large amounts of protein production and purification could be accomplished. Two possible reasons might explain these low levels of expression: 1) the clones were in the PSL1180polyUB vector, which contains the polyubiquitin promoter of *Aedes aegypti* and might not be efficiently

recognized by the transcription machinery of *drosophila* S2 cells, and 2) the polyubiquitin promoter of *Aedes aegypti* is constitutively overexpressed which might result in problematic instances when the protein being expressed is toxic for cells. Therefore, to further increase the expression amounts of these proteins, the sequences of PCs of *Aedes aegypti* were codon optimized and cloned into a highly inducible promoter, the metallothionein (*Mt*), which is known for producing large amounts of protein when S2 cells are incubated in the presence of metals like copper or cobalt [125], [126], [136]. This was the step that improved the amount of protein produced.

Several questions remain open, such as why only furin2 showed strong activity when expressed in HEK293T cells, or why after codon optimization furin1 gained more activity than furin2 when expressed in S2 cells. However, the optimization of expression and protein production that was performed in this chapter, set the stage for the more functional approach performed in chapter 4, where different aspects of the biochemistry of these proteases are addressed.

In summary, this chapter helped understand important aspects of expression and purification of proprotein convertases of mosquitoes. First, the production of individual domains, such as the catalytic domain is not possible, and it likely requires other sections of the protein for stability. Second, the unusual long N-terminal region of furin1 and furin2 is absolutely necessary for expression and secretion and cannot be exchanged with the N-terminal region of human furin. Third, the protein expression is enhanced in codon optimized clones that harbor the pMT promoter. Fourth, FLAG immunoprecipitation produced purer protein but in lesser amounts. Finally, furin1 shows activity comparable to human furin.

CHAPTER 4.   ENZYMATIC CHARACTERIZATION OF PROPROTEIN CONVERTASES OF *Aedes aegypti*

4.1   Chapter Summary

Viruses can utilize host factors during their life cycle. The flaviviruses are a good example of viruses that use host proteins, especially during assembly, secretion and maturation processes. The prM protein of dengue or Zika viruses is cleaved by the host protease furin in a pH-dependent manner, serving as a key and lock strategy that prevents premature fusion of the virus with intracellular membranes. However, these viruses are transmitted by mosquitoes and the invertebrate furin has not been studied yet. Furin is a member of the proprotein convertases, which depend on calcium for activity and can be present in different locations within the cell. The mosquitoes have three proprotein convertases named furin1, furin2 and NC2. In this study, the proteases were expressed in *Drosophila* S2 cells and purified in a two-step affinity purification process, including Ni-NTA and FLAG immunoprecipitation. The purified samples were enzymatically characterized. Furin1 showed activity comparable to human furin, whereas furin2 showed significantly reduced activity. In contrast, NC2 showed no activity under the experimental conditions. The enzyme inhibition assays with the classic furin inhibitor-I (decanoyl-rvkr-cmk) demonstrated that furin1 is more sensitive than human furin. Furthermore, plaque inhibition assays in Aag2 cells showed that the titer of both, DENV2 and ZIKV, can be reduced in a concentration dependent manner. However, SINV was not inhibited when cells were infected at high MOI, but at lower MOI the titer was reduced with the inhibitor, suggesting that the lack of maturation of SINV affects its spread in cells deficient in furin activity. Finally, an *in vitro* maturation assay demonstrated that DENV2 can be primed to be infectious at low pHs with both furin1 and furin2. However, ZIKV was primed only with human furin. Combined, these results suggest the mosquito proprotein convertases have a role in viral infection and can be used as potential targets for antiviral development and control of vector-borne infections.

4.2   Introduction

Multiple viruses utilize host machinery during different steps in their life cycle. This means that host proteins are a good target to disturb the viral infection [137]. One of the steps in the life

cycle of viruses are the post-translational modifications. Similar to host proteins, the viral proteins need to undergo a several changes, including folding, glycosylation, and proteolytic cleavages[138]. Proteins can be expressed as precursors that are inactive, but for activity require the cleavage at specific residues[139]. This is usually done by endoproteases, and viruses rely on them during maturation. Furin, a proprotein convertase, is implicated in the cleavage of glycoproteins from different families of viruses, hence the importance of this protein in virology[27], [44], [48], [140].

Therefore, it is not surprising that furin, a protease involved not only in the viral life cycle but also in processing of multiple precursor proteins in the host, has been extensively studied. To date, there are more than 3000 research articles on Pubmed that have examined furin. Structure to biochemistry and genetic knockouts have been investigated for furin. However, all these studies have focused on the mammalian furin, mostly human and to some extent mouse [23], [79]. But, from the viral life cycle perspective, other hosts, and reservoirs should be taken into account.

The flaviviruses comprise a genus of medically relevant viruses that affect human health globally. This genus includes DENV2 and ZIKV, which utilize the furin cleavage during secretion [50], [57], [59], [62]. This cleavage is performed in a pH-dependent manner and results in presumably mature virions [61], [62]. One of the main aspects of the Flavivirus is that they can be transmitted by an invertebrate host like mosquitoes. However, before being transmitted to the next susceptible host, the virus needs to infect different organs and cells in the mosquito, meaning that the virus must undergo multiple maturation steps [68]. This maturation process has not been addressed in mosquitoes, but it is presumed that a similar protease, like human furin, is implicated.

Furin is a member of the proprotein convertase family, which are metabolically conserved serine proteases that cleave at poly basic residues. The diversity of proprotein convertases implies different locations in the cell, which can be different in terms of environmental conditions such as pH and ionic content [79]. As mentioned above, furin has been extensively studied, and its life cycle has been documented to involve multiple layers of activity tightly associated with pH and metal ions. Calcium is particularly important for activity, and it has been documented to work as a cofactor that regulates the stability of the catalytic domain [94], [141]; there are at least three

calcium binding sites in human furin [142]. This is an important aspect in the life cycle of the protease involving multiple locations within the cells. For example, the accumulation of calcium happens in the trans-Golgi network, which has deep implications in spatial-temporal interaction with viral pathogens that use the secretory pathway. With lower concentrations of calcium, the protease activity is reduced and with addition of chelators such as EDTA or EGTA, the activity is completely depleted [141].

The furin activity can be measured by cleavage of a fluorogenic substrate called BoC-RVRR-AMC. This substrate has a canonical furin recognition site that once cleaved by furin emits fluorescence that can be measured at a different wavelength. In addition, BoC-RVRR-AMC has been used for testing activity not only of furin, but also other members of the proprotein convertase family. This cleavage site is similar to the site in DENV2 (REKR), but better with ZIKV (RSRR) [47], therefore giving some insights on the functional characterization in terms of infection with these viruses.

One of the major enzymatic characterizations of human furin occurred in the 90s, when the first full length clone of furin was obtained. From these studies, human furin was shown to display a typical Michaelis-Menten behavior when incubated at different concentrations of substrate. This lead to identifying the enzyme constants such as Vmax and Km, which tell the efficiency of the protease [141], [143].

Similarly, one of the additional enzymatic assays that have been pursued with human furin, is the use of the furin inhibitor-I, which is a peptide that mimics the recognition cleavage site and renders the protein inactive upon binding. This inhibitor has been extensively used in multiple virology studies where processing of viral glycoprotein is the interest of the research[56], [144]. As expected, all studies have been focused on mammalian cells, with only one study addressing the effect of this inhibitor in insect cells[145]. Despite controversy about the specificity of this compound, it is still a good starting inhibitory drug to address viral infection in the context of furin cleavage.

Furthermore, functional assays involving *in vitro* maturation of viral particles have provided insights into the optimal pH, calcium concentration and amount of protein necessary to produce cleaved glycoprotein and generate infectious viruses [62], [63]. These studies involved the production of purified virus and incubation with the corresponding protease at different pHs during a stablished period of time that renders the virus infectious. These experiments work as an additional layer of information that demonstrate the role of a protein of interest in the processing of viral proteases, mimicking the environmental conditions in the cell.

In the present study, the recombinant proprotein convertases of *Aedes* aegypti, the main vector of flaviviruses like dengue and Zika, were expressed and purified. The purified proteins were used to characterize their enzymatic activity using the fluorogenic substrate BoC-RVRR-AMC and determine their Vmax and Km constants. The Neuroendocrine convertase NC2 did not show activity under the experimental conditions, but furin1 and furin2 did show activity. However, furin1 has a more robust activity comparable to human furin. Further, the furin inhibitor1 was used to determine the IC_{50} and perform plaque reduction assays against DENV2 and ZIKV. SINV was also used as a way of comparison with alphaviruses and displayed a unique behavior in which the inhibition is MOI-dependent. Finally, an *in vitro* maturation experiment was attempted using purified furin1 and furin2 with partially purified DENV2 and ZIKV. The results suggest that at lower pH the activation of DENV2 can be carried out by human furin and furin1, but not furin2. Combined, these experiments demonstrate that furin1 and furin2 are active proteases that are implicated in the maturation of flaviviruses and alphaviruses.

4.3   Methods

4.3.1   Cells and viruses

Aag2 and S2 cells were maintained in *Drosophila* Schneider's Medium, supplemented with 10% Fetal Bovine Serum (FBS) at 27 ºC and no CO_2. In addition, Aag2 cells were also supplemented with non-essential amino acids and PenStrep. C6/36 cells were maintained in minimum essential media (MEM) supplemented with 10% FBS at 30 ºC in an atmosphere of 5% CO_2. BHK and Vero cells were maintained at 37 ºC in an atmosphere of 5% CO_2. BHK were grown in minimum essential media (MEM) supplemented with 10% FBS, whereas Vero cells in

Dulbecco's Modified Eagle Medium (DMEM) with 10% FBS. Low passage DENV2 strain 16681, and ZIKV strain H/PF/2013, were grown in C6/36 cells in the presence of MEM and 2% FBS. SINV-mCherry virus (toto64)[64] was produced from RNA transfection into BHK cells.

4.3.2 Protein production, purification, and quantification

The coding sequences of the proprotein convertases of *Aedes aegypti* were codon optimized and cloned into the pMT-Puro plasmid (Addgene #17923) using the services of Biobasic Inc. The clones also contained three tags at the C-terminus: FLAG, HiBiT and 6x-Histidine. The codon optimized plasmids were expanded using a Midi-Prep Purification Kit (Qiagen #12643). Purified plasmid was transfected into S2 cells using lipofectamine 3000. Then, 48 hours post transfection, cells were centrifuged, and media exchanged for media containing 10 µg/mL of puromycin. Cells were selected for 2 weeks, in which media were changed every fourth day to remove dead cells. After that, the cells were expanded to a final volume of 100 mL in puromycin-media in five 20 mm dishes. After 3 days of growth, the cells were centrifuged and media were exchanged for serum-free media Ex-Cell 420 (Sigma #14420C) without puromycin but in the presence of 600 µM $CuSO_4$ to induce expression of the metallothionein promoter. 48 hours after, the cells were centrifuged and the media containing secreted protein was clarified using a 0.22 µm filter.

Taking advantage of the two affinity tags that the proteins had in the C-terminal region, a two-step purification process was performed. All steps were done on ice or at 4 ℃. First, 1 mL of Ni-NTA affinity resin was centrifuged at 8000 xg for 30 seconds, then washed with ultrapure water and equilibrated twice with Native Purification Buffer (50 mM NaH2PO4, pH 8.0 & 0.5 M NaCl). The cleared media containing secreted protein were incubated with the equilibrated Ni-NTA resin overnight at 4 ℃ under constant rotation. Next day, samples were centrifuged and the resin was washed three times with Native Purification Buffer containing 20mM of imidazole. Then, the protein was eluted using 250 mM Imidazole.

The imidazole-eluted sample was buffer exchanged for 20 mM MES, 150 mM NaCl, 2 mM $CaCl_2$ using an Amicon ultracentrifuge filter 30k MWCO (Millipore sigma #UFC9030). Simultaneously, 200 µL of FLAG-M2 agarose beads (Sigma #A2220) were centrifuged at 8000 x g for 30 seconds, washed three times with TBS, treated once with 0.1 M Glycine HCl pH=3.5 to

81

remove unbound antibody and then washed three additional times with TBS. The buffer-exchanged sample was incubated with the washed/equilibrated beads for 2 hours at 4 ℃. After that, the beads were washed three times with TBS and eluted with 150 µM Flag peptide dissolved in TBS. Once again, The eluted sample was buffer-exchanged as described before. The purified protein was mixed with 50% glycerol, flash-frozen in liquid nitrogen and stored at -80 ℃. For protein quantification, a standard was performed using dilutions between 78 and 4.9 µg/mL of Bovine Serum Albumin (BSA). The band intensity was determined using the LI-COR Image Studio software and a linear regression was produced and used to determine the concentration of purified protein and virus.

4.3.3 Enzymatic activity assays

For enzyme progression curves, 5µL of purified protein were incubated with 100 mM HEPES (pH 7.5 at 25°C), 0.5% Triton X-100, 1 mM $CaCl_2$, 1 mM ß-mercaptoethanol and 50 µM BoC-RVRR-AMC substrate in a 100 µL volume at 30°C. The measurements were done in iD5 instrument (Molecular Devices) every 5 minutes for 20 hours, with an excitation of 370 nm and emission of 470 nm. The integration time was 100 ms, OD=2 attenuated, and linear shake for 30 seconds in between reads. The reads were made from the bottom of a 96 well black clear bottom plate (Corning #3881). Each experiment was performed at least in duplicates.

For calculation of enzyme velocity, the values of the linear phase, corresponding to the 5-15% of the plateau, were used to determine the slope in RFU/min. These values corresponded to the time frame between 15 and 60 min. The experiments with change in calcium, substrate or temperature were a variation from the original reaction setup as described above. However, the experiment with different pHs required an additional change. At lower pHs (i.e. 5 and 6), the addition of 100 mM Sodium Acetate was used to buffer the solution.

For enzyme inhibition activity assays, the furin-inhibitor-I compound (Millipore Sigma) was dissolved in DMSO at a concentration of 2 mM for the stock. All the compound dilutions were made in DMSO to maintain a constant concentration of 5% among the samples.

4.3.4 Plaque reduction assays

The furin inhibitor-I was dissolved in DMSO at a stock concentration of 20 mM. Serial dilutions were made to achieve a final DMSO concentration of 0.5% to prevent cell toxicity, meaning the compound's highest concentration measured was 100 µM. First, Aag2 cells were assessed for cytotoxicity with the inhibitor using the WST-1 compound (Biovision K304). Briefly, cells were seeded in a 96-well plate at a density of 40000 cells/well. 24 hours later, the media were replaced with media-containing inhibitor at different concentrations and cytotoxicity was measured 48 hours later following manufacturer instructions. For the plaque reduction assays, Aag2 cells were treated and infected with DENV2, ZIKV or SINV for 2 hours. After that, the virus inoculum was removed and cells were washed two times with PBS and fresh media were added. Cells were incubated at 27 °C for 1 hour before changing media and replacing with the fresh media containing the furin inhibitor-I. 48 hours later the virus was collected and tittered on BHK (DENV2 and SINV) or Vero (ZIKV).
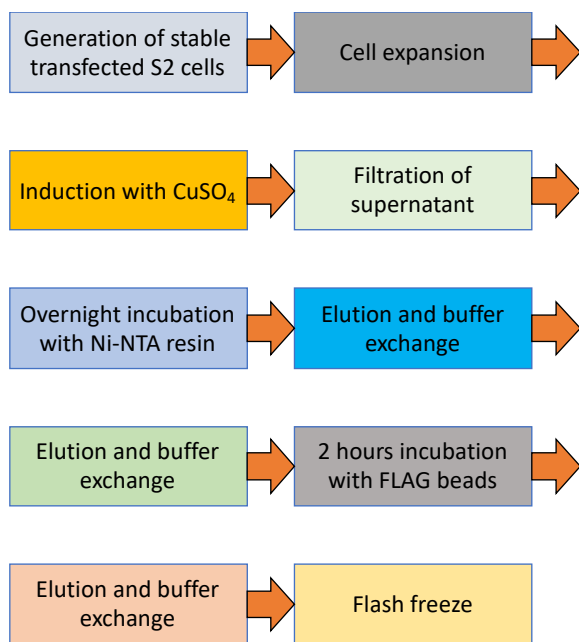
Figure 4.1 Procedure for protein purification

4.3.5   Partial purification of virus and in vitro maturation

C6/36 cells were infected with DENV2 or ZIKV at an MOI=1. Then, 96 hours post infection, the supernatant was removed and exchanged for media containing 2% FBS and 30 mM ammonium chloride (NH_4Cl). Cells were incubated in this media for two hours. The process was repeated two more times for a total of three incubations with NH_4Cl. After that, the cells were incubated overnight with no FBS and 30 mM NH_4Cl. Next morning, the supernatant was collected and virus was purified on a 20% sucrose cushion for a 2 hour centrifugation at 32000 rpm (Beckmand Coulter rotor 50.2 Ti). Tubes were incubated at 4 ºC overnight to slowly dissolve the virus pellet.

	Signal	Concentration (µg/mL)
Furin1	271	21.46
Furin2	401	27.00
NC2	168	17.07
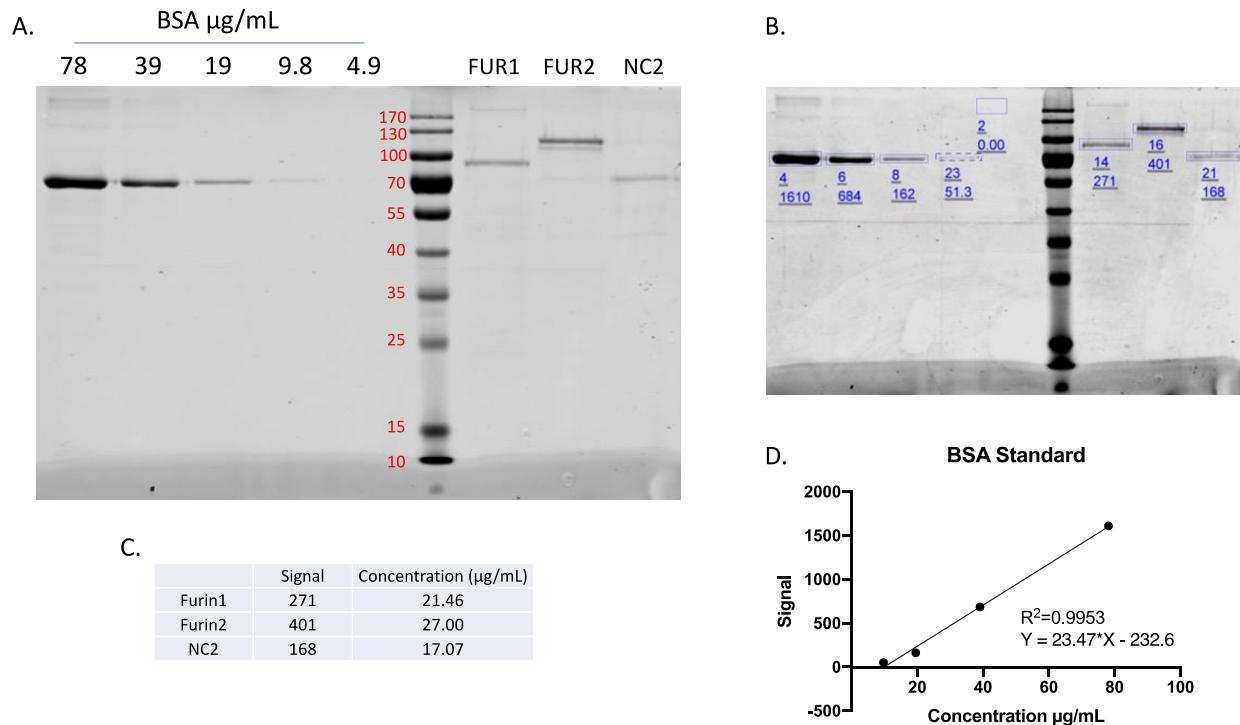
Figure 4.2 Protein quantification for enzymatic activity assays.
A. SDS-PAGE gel with the BSA standard and the purified proteins furin1, furin2 and NC2. B. Signal intensity of each band as in A. C. Concentration of protein calculated using the BSA standard. D. Linear regression of BSA standard. Expected sizes: BSA 66.5 kDa, furin1= 84.2 kDa, furin2: 89.5 kDa, NC2: 70.9 kDa.

The partially purified virus, along with purified proteases, were quantified using a BSA standard on an SDS-PAGE. For the *in vitro* maturation experiment, equivalent amounts of protease were used in nmol. The virus:protease ratio was 1:3. The samples were incubated at 30 ºC in the presence of 3 mM $CaCl_2$, for 16 hours in different buffers (citrate phosphate or Tris-HCl)

depending on the pH as described by others[62]. Next morning, the samples were collected and analyzed by plaque assay.

4.3.6 Structural analysis of the active site and calcium binding sites

The crystal structure of human furin in complex with the competitive inhibitor Arg-Arg-Arg-Val-Arg-Amba was retrieved from PBD (6EQX). The structure was analysed on pymol, where the amino acid residues within 4 Å were highlighted to identify the binding pockets and the catalytic triad along with the oxyanion hole. A similar approach was made to identify the glutamic and aspartic acid residues of the calcium binding sites. Then, the iTasser predicted structure of furin1 and furin2 was aligned with the human furin structure, and the amino acid residues of the subside pockets and the calcium sites were compared.
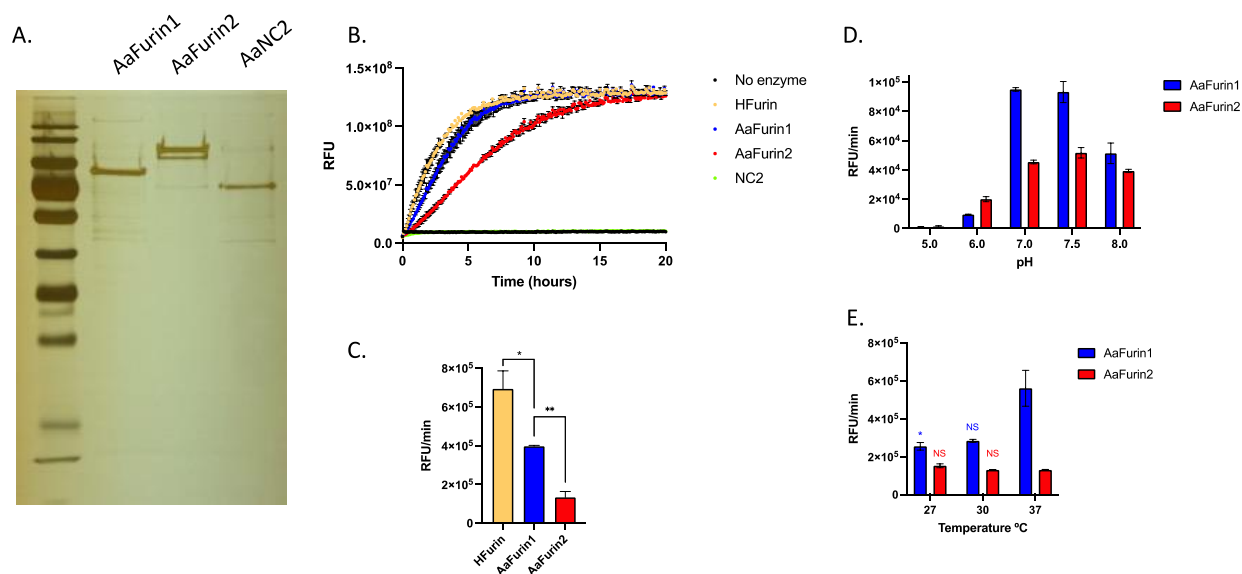
Figure 4.3 Enzymatic activity assays of purified AaFUR1 and AaFUR2
A. Silver stain of purified proteins. B. Kinetic activity assay using purified proteins and 25 µM of substrate, BoC-RVRR-AMC. C. Enzyme velocity calculated from the slope of the linear phase of "B". D. Enzyme velocity at different pHs. The lower pHs (5 and 6) were done in 50 mM sodium acetate, whereas the neutral-basic (7, 7.5 and 8) were done in 100 mM Hepes. E. Enzyme velocity at different temperatures using neutral pH. Asteriscs indicate statistical significance. Significantly different $P<0.05$. NS Not significant. N=2 independent replicates

4.4   Results

4.4.1   Protein production, purification and quantification

Drosophila S2 cells, stable transfected with the pMT-Puro clones, were induced with 600 µM of $CuSO_4$ and supernatant was collected 48 hours after. The supernatant was collected, filtered and incubated overnight with pre-equilibrated Ni-NTA affinity resin. The next day, the resin was washed and eluted with imidazole. The eluted sample was buffer-exchanged by ultracentrifugation. Then, the sample was incubated for 2 hours with pre-equilibrated FLAG beads and eluted with FLAG peptide, buffer exchanged and flash-frozen (Figure 4.1). This two-step purification process provided an advantage of purer sample while still maintaining enough amounts of protein.

One of the first analyses performed was to detect the protein on an SDS-PAGE gel stained with *instant-blue*. As discussed in chapter 3, the purification of these proteins was a great challenge, and multiple expression systems were attempted before producing enough protein to be detected by a standard protein gel, without relying on western blot. However, the purification strategy shown in this chapter provided sufficient pure protein that was able to be quantified using a Bovine Serum Albumin (BSA) standard (Figure 4.2). It is important to mention that the expected sizes (furin1= 84.2 kDa, furin2: 89.5 kDa, NC2: 70.9 kDa) correspond to the sizes after cleavage of the pro-segment. However, purified furin2 showed a higher molecular weight (>100 kDa), suggesting that the pro-segment was not cleaved (Figure 4.2 a and b). In addition, the furin2 lane showed two bands, indicating that there might be two species of the same protein, where the lower band might indicate the matured furin2. In addition, the protein amounts were determined using the BSA regression linear curve as shown in Figure 4.2 c and d. All proteins were in the range of 17 to 22 µg/mL.
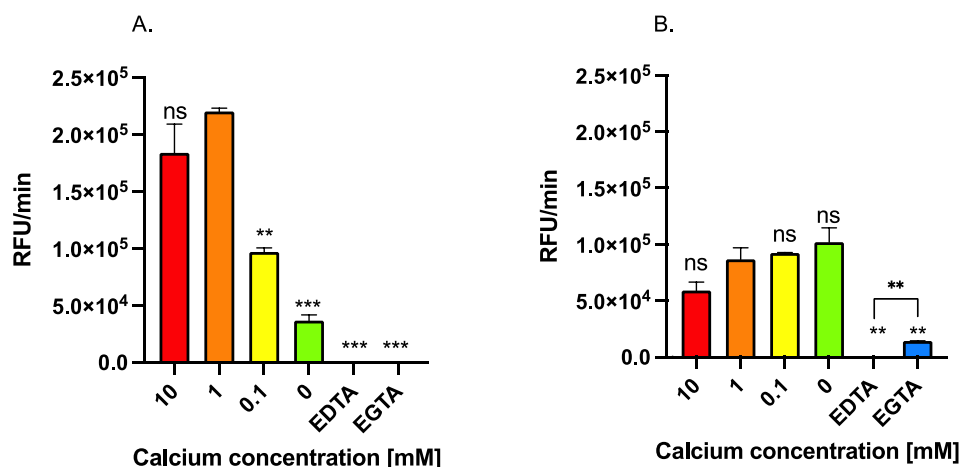
Figure 4.4 Furin1 is dependent on calcium concentration
Purified enzymes were incubated with different concentrations of calcium A. Furin1, B. Furin2. EDTA and EGTA were added at a concentration of 2 mM in the presence of 1 mM CaCl2. Significance was determined with a two-tailed t-test, where $P<0.05$ is significant. NS Not significant. N=2 independent replicates.

4.4.2 Enzymatic characterization

To perform the enzymatic activity assays, first the sample had to be assessed for any impurities. As shown in the silver stain, which has a detection limit of 0.25 ng (compared to 5 ng with instant blue), no contaminant bands were detected (Figure 4.3a). Therefore, the activity assays performed under the experimental conditions, demonstrate that the activity comes only from the protein of interest. As shown in figure 4.3b, both furin1 and furin2 have activity measured by cleavage of the BoC-RVRR-AMC substrate. On the contrary, NC2 did not show any activity and appeared as the no-enzyme control, suggesting that NC2 is either an inactive protease or the preferred substrate is different to BoC-RVRR-AMC.
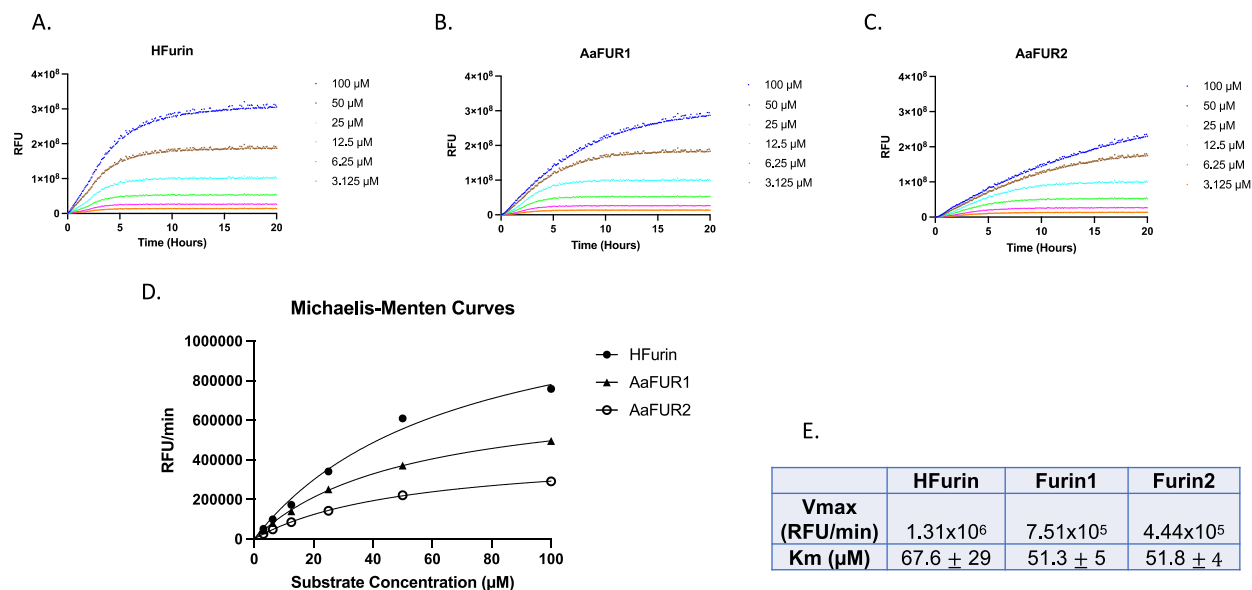
Figure 4.5 Enzymatic characterization of furin1 and furin2.
A,B and C. Kinetic activity assays with human furin, furin1 and. Furin2 at different concentrations of substrate. D. Michaelis-Menten curves using the slope of the linear stages from a-c. E. Vmax and enzyme constants Km. N=2

The next step was to compare the enzyme velocity using the slope during the linear stage of the enzyme progress curves. Commercially available human furin was used as a positive control, and it displayed the highest activity. Furin1 has activity that is as low as half the speed of human furin, whereas furin2's activity is at the lowest level, with a speed that is nearly one third of furin1 (Figure 4.3c). Then, different pHs were tested using pH7.5 as the control (Figure 4.3d). Furin1 had a significant reduced activity at lower pH (= 5 and 6), showed the highest activity in neutral pH (=7 and 7.5) but activity is reduced by half at pH=8. In contrast, furin2 did not show much variation among neutral and slightly alkaline pHs, but showed twice the activity of furin1 at acidic pH=6. In addition, the activity of furin1 and furin2 was also tested at different temperatures (Figure 4.3e). Furin1 showed the highest activity at 37 ℃ and lowest at 27 ℃, whereas furin2 did not have a significant change among temperatures.

Previous studies have shown that human furin depends on calcium for activity, and the addition of chelators to the enzymatic reaction, such as EDTA or EGTA, completely render the protein inactive. For this reason, furin1 and furin2 were incubated with different concentrations of calcium (Figure 4.4). As expected, furin1 showed the most activity at 1 mM calcium, but the activity was significantly reduced at lower concentrations of calcium and was reduced by one

fourth at 10 mM. When incubated with EDTA or EGTA, the activity was reduced by ten-fold (Figure 4.4a). In contrast, furin2 did not show any significant change at different concentrations of calcium, but was severely compromised in the presence of EDTA. EGTA also reduced the activity, but with less harmful effect than EDTA (Figure 4.4b).
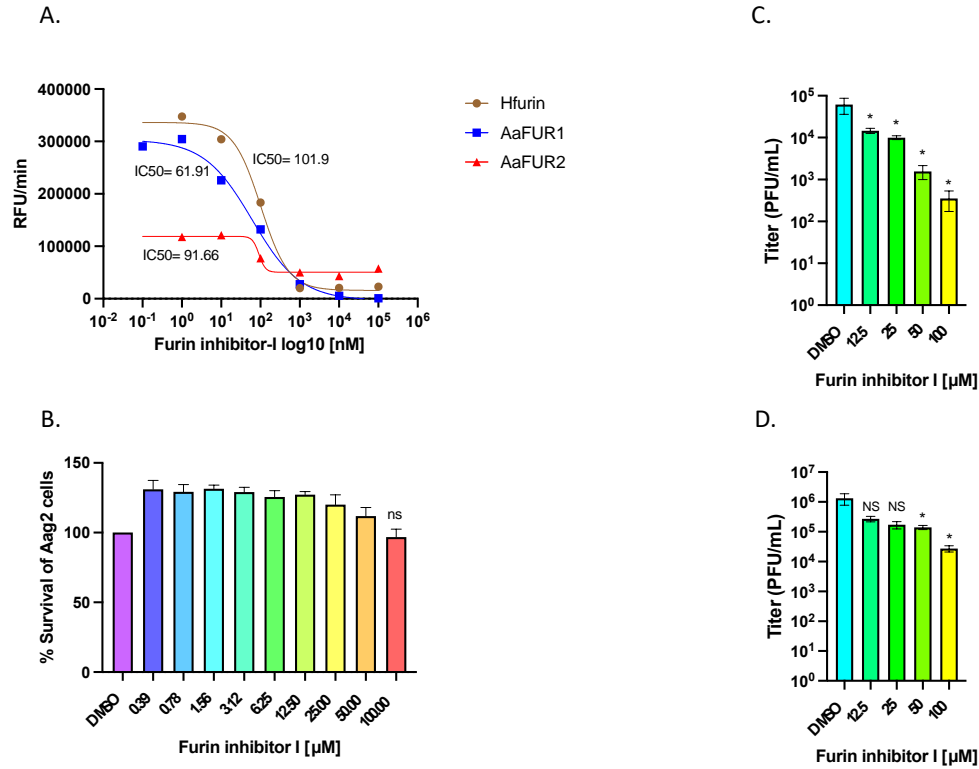
Figure 4.6 Effect of furin inhibitor-I in the DENV2 and ZIKV infection
A. In vitro inhibition of purified protein in the presence of the inhibitor. B. Cytotoxicity assay in Aag2 cells treated with furin inhibitor-I. C and D. Plaque reduction assays in the presence of furin inhibitor-1 for DENV2 and ZIKV, respectively. Significance was determined with a two-tailed t-test, where P<0.05 is significant. NS Not significant. N=2

Finally, to determine the enzyme kinetics constants, furin1 and furin2 were incubated with different concentrations of BoC-RVRR-AMC substrate -3.125 µM to 100 µM- at pH=7.5 and 1 µM $CaCl_2$ at 30 ºC. Human furin was used as a comparison. The enzyme progress curves showed that furin2 did not reach a plateau at the highest concentrations of 50 and 100 mM. Similarly, furin1 did not plateau at 100 µM (Figures 4.5a, b and c). Then, the kinetic points of the linear phase (which corresponds to values between 5-15% of the maximum RFU) were used to calculate the slope and determine the enzyme velocity at different concentrations of substrate and plot a Michaelis-Menten curve (Figure 4.5d). The maximum enzyme velocity (Vmax) for human furin

was almost twice the value of furin1, and three times the Vmax of furin2. In contrast, the substrate concentration that yield a half-maximal velocity, or Michaelis-Menten constant, Km, was similar between furin1 and furin2, but 17 µg lower than human furin. Together these enzyme kinetics assays suggest that, furin1 and furin2 get saturated at lower concentrations of substrate than human furin.
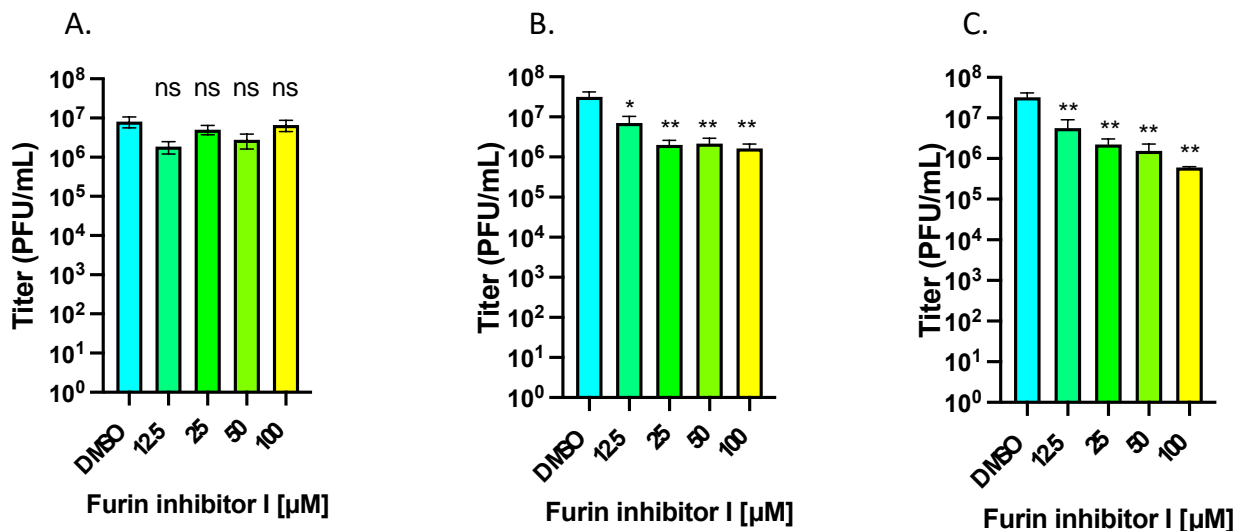
Figure 4.7 SINV infection is reduced with furin inhibitor-I in a MOI-dependent manner
A. MOI=10, B. MOI=1, C. MOI=0.1. Significance was determined with a two-tailed t-test, where P<0.05 is significant. NS Not significant. N=2.

4.4.3   Reduction of viral infectivity in Aag2 cells using human furin inhibitor-I

The furin inhibitor-I, decanoyl-Arg-Val-Lys-Arg-CMK, is a widely used compound used in virology studies addressing the furin cleavage site. This inhibitor is a peptide that mimics the canonical recognition site of furin and upon interaction with the enzyme, remains covalently bound to the catalytic site. In this study, the inhibition of mosquito proprotein convertases was addressed using this compound. First, the inhibition of purified protein was tested *in vitro* to calculate the IC_{50} (Figure 4.6a). Furin1 showed an IC_{50} of almost half of human furin (61.9 ± 1.8 vs 101.9 ± 2.0 nM) whereas furin2 showed an IC_{50} of 91.7 ± 2.0 nM. This suggests that under the experimental conditions presented here, furin1 is more sensitive to inhibition and it could potentially be targeted for *in vivo* assays with cells.

Aag2 cells were initially assessed for cytotoxicity against the furin inhibitor-I using the WST-I compound. The cells showed no toxicity at the highest concentration tested of 100 µM (Figure 4.6b). Therefore, Aag2 cells were treated with the inhibitor at concentrations ranging 100-12.5 µM, and then infected with DENV2 or ZIKV at a MOI=10. Virus was collected and quantified 48 hours after the infection. For DENV2, there was a significant reduction in viral titer starting at the lowest concentration tested of12.5 µM (Figure 4.6c). In contrast, ZIKV was inhibited only at the highest concentrations (Figure 4.6d).
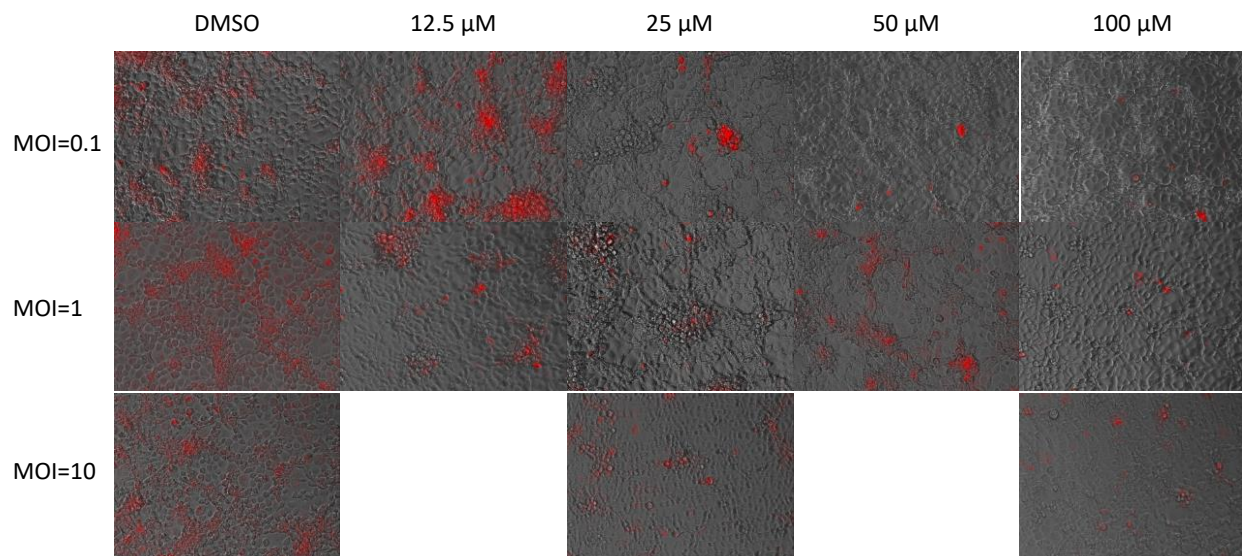
Figure 4.8 SINV-mCherry inhibition of spread in lower MOIs
Aag2 cells were infected at different MOIs and treated with the furin inhibitor-I. Viral spread was qualitatively determined using a fluorescent microscope 24 hours post infection.

Next, Aag2 cells were treated again with the furin inhibitor-I but this time the cells were infected with an alphavirus, SINV, which has a different maturation process, but also requires furin. At a MOI=10, there was no difference in the titer among different concentrations of the inhibitor (Figure 4.7a). However, at lower MOIs, the viral titer was significantly affected by the concentration of inhibitor, particularly at MOI=0.01 (Figure 4.7 b and c). To determine whether the reduction in titer was due to the reduction in viral spread among other inhibitor-treated cells, photos were taken of Aag2 cells infected with SINV-mCherry (Figure 4.8). As expected, viral spread was reduced at the lowest viral infection (MOI=0.1) and higher inhibitor concentrations. At an MOI=1 and MOI=10 there is not difference in the fluorescence among inhibitor concentrations. However, there is more signal in the DMSO control.
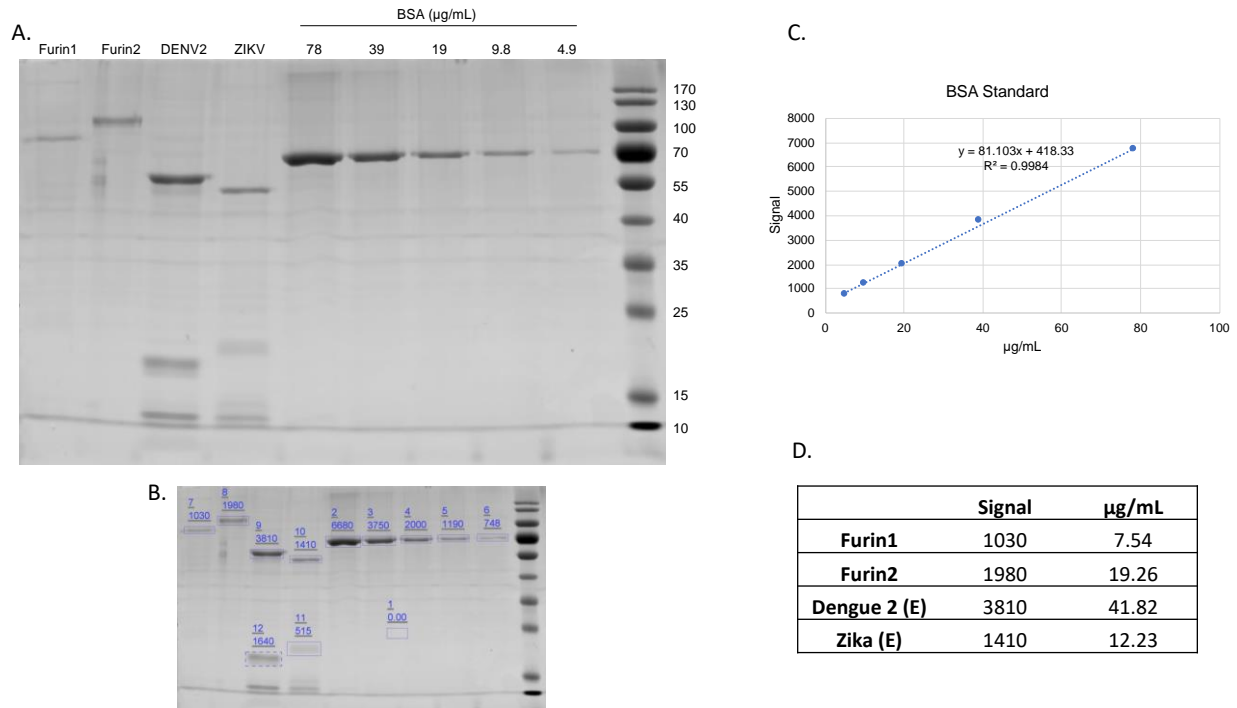
Figure 4.9 Quantification of viral protein E and proteases furin1 and furin2
A. SDS-PAGE gel with purified proteins, partially purified DENV2 and ZIKV, and BSA standard. B. Signal intensity of each band as in A. C. Linear regression of the BSA standard. D. Protein concentration based on the BSA linear regression. N=1.

4.4.4 In vitro maturation assays of DENV2 and ZIKV

Immature viruses were produced in C6/36 cells and partially purified using sucrose cushion. Then virus along with proteases were quantified using SDS-PAGE gel and BSA standard (Figure 4.9). The concentration of E protein was used to add an equimolar amount in a ratio of 16:1 protease:E. The samples were incubated overnight at 30 ℃ at different pHs and the virus was detected. Titer was detected the next day using BHK or Vero cells. DENV2 showed infectivity among all pHs. However, at the lowest pH tested (=5.5), only treatment with human furin and furin1 produced infectious particles. At pH 6.5, human furin-treated DENV2 had higher titer than furin1 or furin2, and at neutral pH all samples showed infectivity, including the no-enzyme control. In contrast, ZIKV did not show any infectious particles treated at pH=5.5, and at pH=6.5 only the samples treated with human furin showed some infectivity. Similar to DENV2, at neutral pH all the samples, including the no enzyme control showed infectivity.

4.4.5 Structural analysis of the active site, subside pockets and the calcium sites of human furin, furin1 and furin2

Using the crystal structure of human furin in complex with the competitive inhibitor (PBD 6EQX), the subside binding pockets and calcium binding sites were highlighted (Figure 4.11a). As shown in the superimposed structures of mosquito furin1 and furin2 with human furin, there is conservation of folding in most of the residues. Interestingly, the histidine of the catalytic triad (H194) shows variation in folding between furin1 and human furin (Figure 4.11b). In addition, two amino acid residues are different between human furin and furin1: D191S and E257D (Figure 4.11b, .1). In contrast, the alignment between human furin and furin2 showed more conservation in the folding, with more similar spatial organization of H194. However, there are three amino acids in the subsite S4 that show completely different organization (V231, T232N and D233) (Figure 4.11c, .1). This might be associated with the change in the amino acid T232N, which might affect the spatial organization in the site.

The calcium binding sites were also analyzed. There are three calcium sites that have been previously identified in the catalytic domain of furin [15], [142] (Figure 4.12). Alignments of human furin with mosquito furin1 or furin2 showed that these sites are conserved and the aspartic/glutamic acid residues are within 4 Å, which is a good distance for ionic bonds. However, the calcium site 2 of furin2 (Figure 4.12e) does not have an aspartic acid within 4 Å, and only one at 5.5 Å, suggesting that this protease might have less dependency on calcium given that only has two sites.

4.5   Discussion

Proprotein convertases are a group of serine proteases that are necessary for the activation of a variety of cellular and viral glycoproteins. Furin, the most studied proprotein convertase is implicated in the processing of prM in dengue and Zika viruses. Importantly, these viruses utilize the host machinery in multiple stages of their life cycle and targeting those steps during antiviral development is an alternative strategy to treat infections. In addition, both dengue and Zika viruses are transmitted by mosquitoes, where the viruses must infect and undergo processing at different steps. However, the maturation of viruses has not been addressed in mosquitoes.
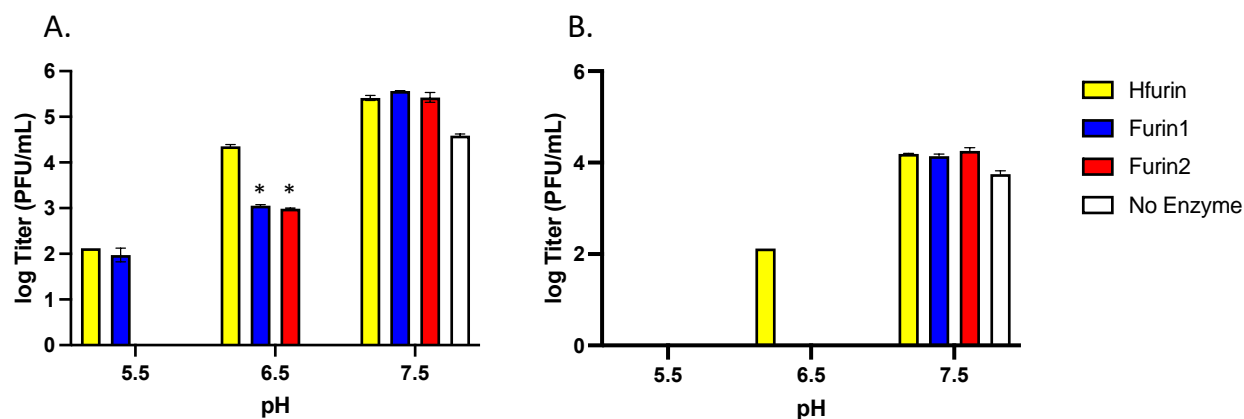
Figure 4.10 Titer of DENV2 and ZIKV after *in vitro* maturation
Partially purified viruses were incubated overnight with the corresponding protease at different pHs. A. DENV2. B. ZIKV. Significance was determined with a two-tailed t-test, where $P<0.05$ is significant. NS Not significant. N=1.

Expression of novel proprotein convertases and their subsequent enzymatic characterization are the first steps in understanding their function and potential use for drug discovery. The S2 cells and the inducible metallothionein promoter provided a robust expression of the proteases. Despite obtaining low levels of protein, the amounts were sufficient to perform activity assays, inhibition and *in vitro* maturation. In addition, the expression platform produced active protein, which in the case of furin1 was comparable to human furin. However, it is important to point out that purified furin2 showed a higher-than-expected size, but instead shows the size of the pre-protein, 112.7 kDa. As it is well known, these proteases undergo a maturation process where the pro-segment is autocatalytically cleaved [23], [27]. This suggests that furin2 is not mature and the reduction in activity compared to furin1 may be the result of immature protein. The explanation for this requires further testing, but it can be that either the S2 cells do not provide the ideal environment for the post-translational modifications of furin2, or that furin2 is indeed less active and is deficient at processing itself.

Furthermore, furin1 but no furin2 showed a change of activity at different concentrations of calcium and at different temperatures. Furin2 showed a slight increase in the absence of calcium and in addition, in the presence of EGTA furin2 has less reduction in activity. These two experiments suggest that furin2 is not dependent on calcium in the same way that furin1 is. Furin1 was significantly inhibited in the presence of both EDTA and EGTA, which are, but EGTA has a higher affinity for calcium. Similarly, furin1 was sensitive to temperature, but furin2 remained the

same. This is an interesting fact because most enzymatic reactions increase their rate with temperature, however it can be reduced when the protein reaches a denaturing stage. However, more temperatures need to be tested, because it appears that at lower temperature, 27 ℃, the activity is slightly higher for furin2. In addition, the enzyme constants Vmax and Km reported here for human furin are similar to those observed using the same substrate [143]. Furin1 and furin2 only varied in the Vmax, but the Km remained the same.
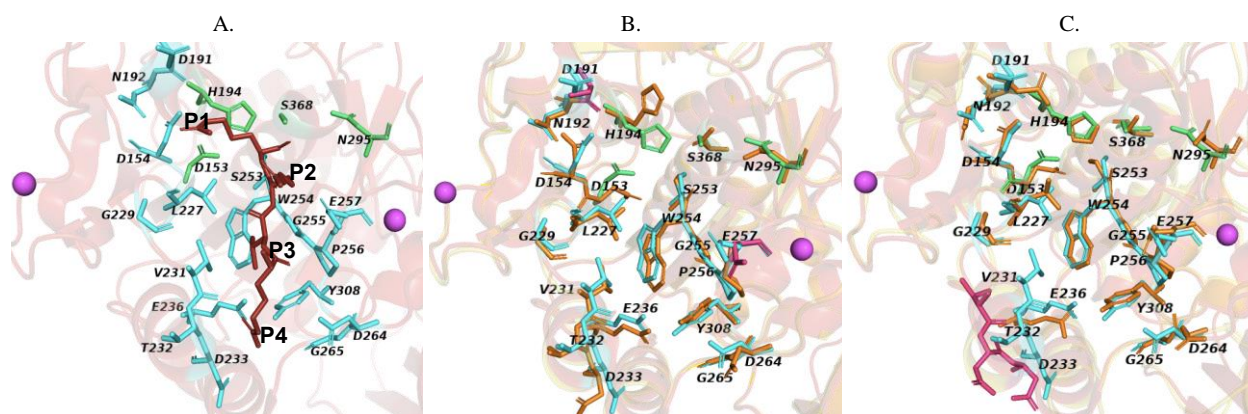
Figure 4.11 Comparison of the active site of human furin, furin1 and furin2. A. Active site of human furin in complex with the substrate RVRR (PBD 6EQX). The catalytic triad (D153, H194 and S368) is highlighted in green, as well as the oxyanion hole asparagine (N295). The substrate RVRR is in red and the P1, P2, P3 and P4 sites are shown accordingly. The amino acid residues of the binding pocket are indicated in cyan. B. Furin1 superimposed to human furin. The residues of mosquito furin1 are shown in orange. Magenta residues indicate amino acids that are different between the two proteases. C. Furin2 superimposed to human furin with similar coloring as in B. The magenta spheres represent calcium ions.

In contrast, the inhibition assays using furin inhibitor-I suggested that furin1 was more sensitive than human furin. For DENV2, the reduction in viral titer is noticeable even at the lowest concentrations, whereas for ZIKV is only observed at 50 µM. DENV2 is known for having a deficiency in the furin cleavage, rendering mosaic particles [61] that retain some uncleaved prM proteins. Therefore, it would be expected that additional perturbations to the cleavage of prM would have higher impacts on DENV2 than on a virus with an optimized furin site.

In contrast, alphaviruses have a different maturation process, in which the E3-E2 protein is cleaved by furin and the virus is subsequently assembled at the plasma membrane[66],. However, in the case of alphaviruses, the furin cleavage is not dependent on pH. This suggests that the cleavage can happen not only in the trans-Golgi, but also in other locations of the cells where

furin is present (the secretory granules, the plasma membrane, extracellular space and in the endosomes). This means that if an alphavirus such as SINV is produced in a cell that lacks furin activity, the E3-E2 proteins will not be processed and the virus will be immature. However, if the same virus is used to infect a normal cell, the cleavage of E3-E2 can occur during the attachment/entry process. This explains the results obtained for the SINV inhibition assays. First, at MOI=10, where the system is saturated, and equal amount of virus is produced from furin-inhibited Aag2 cells versus wild type. The virus coming out from the furin-inhibited cells is therefore immature but utilizes the furin that is present in the BHK cells that are infected during the plaque assays. In contrast, the infection at lower MOIs like 1 and 0.1, the virus would normally spread to other cells, but in the system, all the cells have deficiency in furin activity, suggesting that the virus would not be able to get the furin cleavage and over time it translates into less virus being used for the plaque assays.

Table 4.1 Amino acid residues of the subsite pockets of human furin, furin1 and furin2

Subsite pockets	Location in human furin	Hfurin	AaFUR1	AaFUR2
S1	154	D	D	D
	191	D	S	D
	192	N	N	N
	194	H	H	H
	227	L	L	L
	253	S	S	S
S2	254	W	W	W
	255	G	G	G
	257	E	D	E
S3	236	E	E	E
	256	P	P	P
	264	P	P	P
	265	G	G	G
	308	Y	Y	Y
S4	231	V	V	V*
	232	T	T	N*
	233	D	D	D*

*Amino acids that show strong variation in the folding of the predicted structure as seen in Figure 4.11. Purple residues indicate different amino acid at the corresponding position

The *in vitro* maturation assays demonstrated that DENV2 was able to mature and become infectious in the presence of furin1, furin2 and human furin. At neutral pH, even in the absence of protease, all samples produced titer. Interestingly, at the lowest pH, furin2 was not able to produce

infectious DENV2, suggesting that this protease could have a smaller range of pH activity, as was also shown in the activity assay. On the other hand, ZIKV did not show any maturation in the presence of furin1 or furin2 at low pHs, but only human furin was able to recover infectious particles.
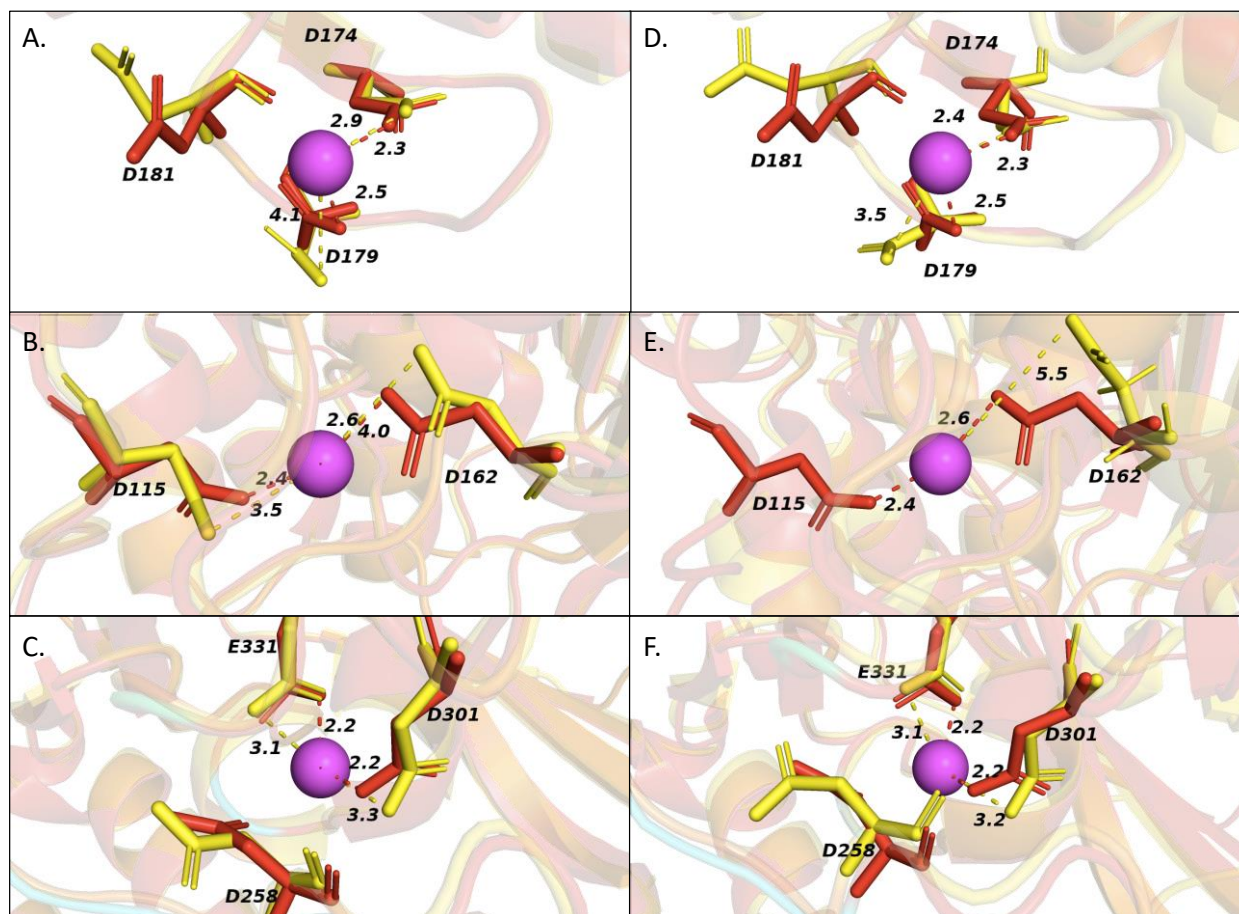
Figure 4.12 Calcium binding sites. Alignment of the calcium binding sites between human furin and furin1 (A-C) or human furin and furin2 (D-F). Calcium site 1 (A & D), site 2 (B & E) and site 3 (C & F). Human furin residues are shown in red and mosquito furin residues are shown in yellow. Magenta spheres represent the calcium ions. Distances between the calcium ion and the carbonil group of the aspartic/glutamic acids are shown with red/yellow dashed lines and represented in Å.

Finally, the structure analysis of the active site and the calcium binding pockets showed that there is high conservation in the folding pattern of the proteases. The active site shows slight changes between human furin and furin1, which might explain the difference in activity, where mosquito furin1 has more affinity for its substrate as shown in Figures 4.3 and 4.5. In addition, the calcium site 2 of furin2 does not have any aspartic/glutamic acid residue within 4 Å, which explains the less dependency on calcium as was see in Figure 4.4. However, this analysis is based

on predicted structure, which would require further validation with experimentally resolved structures.

In conclusion, this chapter addressed for the first time the enzymatic activity of the proprotein convertases of *Aedes aegypti*. This led to characterizing the activity at different pHs, temperatures and concentrations of calcium. Furin1 showed robust activity comparable to human furin, whereas furin2 had a reduced activity probably associated with the lack of autocatalytic processing. In addition, the experiments performed here demonstrate that a classic inhibitor such as furin inhibitor-I can be used for reduction in enzyme activity and reduce viral titer of both Flavivirus and alphavirus. Finally, the in vitro maturation assay demonstrated that furin1 and furin2 can make DENV2 infectious from a sample of immature viruses. Combined, these results suggest that the proprotein convertases of mosquitoes can play a major role in viral infection and can be used as targets for development of antivirals.

CHAPTER 5.   CRISPR-MEDIATED KNOCKOUT OF FURIN1 AND FURIN2 IN Aag2 CELLS

5.1   Chapter Summary

Viruses can utilize different vectors as means of transmission and spread. The flaviviruses are an example, in which they take advantage of the lifestyle of mosquitoes. The virus must infect the mosquito before it transmits to the next susceptible host. In addition, flaviviruses also rely on a host protease, named furin, for their maturation and generation of infectious particles. The furin of the mosquitoes has not been identified nor characterized. In the present study, the maturation process of flaviviruses was addressed in mosquito cells: C6/36 and Aag2. Mosquitoes have two furin-like proteases: furin1 and furin2. Using two gene silencing strategies, siRNA and CRISPr-Cas9, the furin1 protease was found to be necessary for viral infection. CRISPr-Cas9 mediated mutagenesis produce a cell line with a deletion of E311 of furin1. As demonstrated with recombinant protein expression, this single amino acid is required for activity. This is the first mosquito cell line deficient in furin activity. In Aag2 mutant cells, the trans-complementation with recombinant furin1 recovered the titer of DENV2, ZIKV and SINV. Despite having reduced titer, the prM content of DENV2 remained the same between parental and mutant cells, but in ZIKV the prM content increased significantly. In the case of SINV, the reduction in titer was evidenced by deficiency in spread to other cells that lack furin activity, which is observed at lower MOIs. In contrast, knockout of furin2 in Aag2 cells did not have an effect neither on ZIKV or DENV2, however in parental Aag2 cells, the expression levels of this protein are significantly lower, suggesting that a knockout would not affect the life cycle of viruses as these can rely on furin1. Combined, these experiments suggest that furin1 is necessary for maturation of Flavivirus and alphavirus in Aag2 cells and provide a potential target for development of strategies to control disease in the mosquito vector before it can be transmitted to susceptible human hosts.

5.2   Introduction

Mosquitoes have been attributed the title of being the single deadliest animal that have shaped human history. Despite being relatively small creatures, mosquitoes can carry dangerous diseases that lead to increased levels of morbidity and mortality. Such danger is enhanced by the

fact that about half of the world population is at risk of getting infected with a Mosquito-Borne Disease (MBD). For example, *Aedes aegypti* is a highly anthropophilic mosquito and is the bona fide vector of viral pathogens that can affect human health, including but not limited to dengue and Zika viruses (DENV and ZIKV). DENV alone infects more than 390 million people every year [81] and its complications can vary from subclinical to severe dengue shock syndrome, which in some cases can lead to death [146]. ZIKV caused the 2014-2015 outbreak in the Americas, which resulted in a total of 86 countries reporting local transmission [147]. The outbreak of ZIKV was associated with increased cases of microcephaly among infants as well as Guillain-Barré Syndrome in adults [147].

One of the most successful methods to combat the burden caused by MBDs is by controlling the mosquito populations. However, recent controversial opinions regarding the role of mosquitoes in the ecosystem have called for alternative methods to tamp down diseases without affecting the insect [72], [148]. Gene-drive systems and genetically modified mosquitoes (GMOs) offer an option to introduce self-limiting genes in the populations. One extension of such approach would be to produce mosquitoes that are not permissive to infection. To achieve that, it is necessary to identify host factors that are required during the life cycle of viruses in the mosquito.

The flaviviruses, which includes both DENV and ZIKV, are characterized for having a coordinated and complex life cycle: the virus enters cells via clathrin-mediated endocytosis [149], the low pH of the endosome triggers a conformational change of the viral glycoprotein E which allows for fusion of viral and host membranes, then the viral RNA is released into the cytoplasm and is translated into a polyprotein containing structural and non-structural proteins [57]. The genome is subsequently replicated and assembled into the endoplasmic reticulum forming immature viral particles containing uncleaved prM [57]–[60]. Then, the virus travels through the secretory pathway and when it reaches the low pH of the Golgi apparatus, Flavivirus maturation occurs: the immature viral particle undergoes a conformational change that exposes a furin cleavage site in the prM junction [62]. The cleavage of furin is irreversible and causes the virion to release the pr peptide at neutral pH in the extracellular space, and this results in a mature infectious virion [23], [59], [61]. Whether maturation is required or not for infectivity remains a topic of debate.

Furin is a serine protease, a member of the Proprotein Convertase (PC) family. In humans, the PC family contains 9 members, 7 of which cleave the substrate after basic residues in the consensus sequence K/R-Xn-K/R↓ [14], [79]. These proteases play key roles in the processing of several precursor proteins such as growth factors, polypeptide hormones, adhesion molecules and receptors [79]. Furin is ubiquitously expressed and localizes mostly in the trans-Golgi network [23], [27]. This property allows furin to coexist with the virus temporally and spatially as the virus passes through the secretory pathway. As a consequence, furin alone processes glycoproteins from several enveloped virus families: Corona-, Herpes-, Toga-, Retro-, Flavi-, Paramyxo-, Orthomyxo- and Filoviridae [23].

The study of PCs has been limited to mammalian systems, being humans and mice the main models. However, a number of studies have addressed the orthologues of PCs in *Drosophila*. Three PCs have been described in *Drosophila melanogaster*: Dfurin1, Dfurin2 and Amontillado. Dfurin1 and Dfurin2 were named after furin given that both share highly resemblance with human furin [73], [84]. However, Dfurin1 and Dfurin2 differ from human furin by having an extended N-terminal region which might serve as a transmembrane anchor[83] and a multiple repeat of the cytein-rich region [85], respectively. The functions of Dfurin1 have been highly addressed, including antimicrobial response [73], embryogenesis [150], processing of viral glycoproteins [75], as well as synergistic activity with Dfurin2 when processing growth factors [88]. Amontillado shares strong similarity with human PC2 which is entirely soluble and requires co-expression with the peptide 7B2 for being properly folded [134]. In addition, Amontillado is required during pupae development [151].

Similarly, little information exists about the mosquito PCs. Only two studies addressed the annotation of these proteases [75], [78] but no functional evidence has been pursued. The maturation of flaviviruses has traditionally been attributed to furin, but which furin and the differences that occur in mosquitoes, remains an obscure topic. In this research, we aimed to identify the mosquito PCs involved in Flavivirus maturation. For that, we used CRISPr to silence furin1 and furin2 in *Aedes aegypti*, Aag2 cells. Our results indicate that furin1 is an active protease displaying typical PC activity and is responsible for the maturation of Flavivirus in Aag2 cells.

5.3   Methods

5.3.1   Cells and viruses

Aag2 cells were kindly provided by Prof. Raul Andino (UCSF). Aag2 cells were cultured in Schneider's *Drosophila* Medium supplemented with 10% FBS, non-essential amino acids (NEAA) and PenStrep at 27 ℃ without CO_2. S2 cells Aag2 cells were cultured in the same conditions, but without NEAA. Vero cells were cultured in Dulbecco's Modified Eagle Media (DMEM) and supplemented with 10% FBS. BHK cells were cultured in Minimum Essential Media (MEM) and supplemented with 10% FBS. C6/36 cells were cultured in MEM with 10% FBS at 30 ℃ with 5% CO_2. Low passage DENV2 strain 16681, and ZIKV strain H/PF/2013, were grown in C6/36 cells in the presence of MEM and 2% FBS.

5.3.2   siRNA silencing of furin1 in C6/36 cells

The DsiRNAs was produced by IDT, resuspended in DEPC water to 100 µM and diluted to 20 µM and stored at -80 ℃. Sub-confluent C6/36 cells were transfected with 10 nM of DsiRNA using Lipofectamine 2000 following the manufacturer instructions. Cellular RNA was extracted at different timepoints and checked for expression of furin1 by RT-qPCR using the delta delta Ct method. For the experiments checking viral infectivity, the DsiRNA treated cells were infected with the corresponding virus at an MOI=10 at 24 hours post transfection. The DsiRNAs were designed to target conserved regions in the catalytic region.

5.3.3   CRISPr mediated Knockout in Aag2 cells

gRNAs were designed using CHOPCHOP and CRISPOR, targeting the genes before the catalytic domain. The gRNAs were cloned into the pAC-sgRNA-Cas9 plasmid under the U6 promoter. An extra nucleotide, G, was added to the gRNAs that did not have it in the 5' end to aid with transcription initiation. The plasmids containing the gRNAs were transfected into Aag2 cells using Lipofectamine 3000 (Cat. L3000015) following manufacturer instructions. 48 hours post transfection, cells were plated in a 24-well plate at a density of 5×10^4 cells/well. The next day, puromycin was added to a final concentration of 625 ng/mL. On the 7th day after selection with puromycin, the cells were seeded at a density of 0.5 cells/well in a 96 well plate. Two weeks after,

the wells that contained single colonies were expanded for subsequent screening purposes. First genomic DNA was extracted (PureLink™ Genomic DNA Mini Kit, Cat. K182001). For detection of heterozygous, T7E1 was used (NEB Cat. M0302) and sequenced by the Sanger method in GeneWiz Inc.

5.3.4 Site directed mutagenesis and production of recombinant mosquito *Aa*FUR1

The soluble portion of *Aa*FUR1 was codon optimized and cloned into the pMT-Puro vector (Addgene #17923) (Bio Basic Inc). For the generation of the AaFUR_E311del mutant, a site directed mutagenesis was performed as follows. Briefly, original plasmid was used as template for a PCR using Q5 Polymerase (NEB M0491) using the primers F-ATGTGGTACCTGAACCGCG and R-GCCCCACTTCGGGTCGTT. Then the PCR product was DpnI digested (NEB R0176), in vitro phosphorylated with T4 Polynucleotide Kinase PNK (NEB M0201) and ligated using T4 DNA ligase (NEB M0202). Then, dh5alpha cells were transformed with 2 µL of ligated product. Single colonies were screened for the deletion and sequenced (GeneWiz Inc.)

Recombinant protein was produced from stably transfected S2 cells as has been previously reported [136]. Briefly, plasmids were transfected into S2 cells using Lipofectamine 3000 following manufacturer instructions. 48 hours after transfection, cells were selected with puromycin at a concentration of 10 µg/mL. Media was changed every 3-4 days. After three weeks, the cells were expanded, and media was changed for serum-free EX-CELL® 420 (Sigma Cat. 14420C) containing 600 µM of $CuSO_4$. 48 hours later, media was collected, and recombinant protein was purified in a two-step batch purification: first with overnight binding to the Ni-NTA beads (Cat. R90115), eluted with 100 mM Imidazole. Then, the buffer was exchanged for 20 mM MES, 150 mM NaCl and 2 mM $CaCl_2$ and the sample was bound for 2 hours to the ANTI-FLAG® M2 Affinity Gel (Cat. A2220) and eluted with 150 µg of FLAG® Peptide (Cat. F3290). Samples were quantified by BSA standard curve on an SDS-PAGE gel.

5.3.5 Infection of Aag2 cells and trans-complementation experiments

Aag2 cells were seeded overnight at a density of 5 x10^5 cells/well in a 24-well plate. The next day, cells were infected with ZIKV or DENV at a MOI=10 at room temperature for 2 hours.

Virus inoculum was removed, and cells were washed three times with PBS. 48 hours post infection, supernatant was collected for titration, western blots and real time qRT-PCR. Cells were washed once again three times with PBS and total RNA was extracted from cells for quantification of intracellular viral genomes.

For the trans-complementation experiments, the extracellular domain of furin1 was cloned into a plasmid for overexpression in Aag2 cells. Briefly, total RNA from Aag2 cells was extracted using the RNeasy Minikit (Cat. Qiagen 74104). First strand of cDNA was synthetized using the iScript Select cDNA Synthesis kit (BioRad 1708896). The full-length extracellular region of furin1 was PCR amplified with Q5 polymerase and cloned using Gibson Assembly (NEB) into the PSL1180polyUBdsRED (Addgene #49327) under the pUB promoter and removing the dsRED sequence. For Hfurin, the cloning procedure started from a pcDNA3-cDNA clone rather than total RNA. Three tags were included at the C-terminal region: FLAG, HiBiT and 6x-Histidine (Figure 5.7).

Mutant Aag2 cells were reverse transfected with pUB-Hfurin or -furin1 using lipofectamine 3000. Briefly, 500 ng of DNA were mixed with lipofectamine and P3000 reagents and incubated at room temperature for 15 minutes. Then, the DNA transfection mix was added to empty 24-well plates and 5×10^5 cells were added to each well. 24 hours post transfection the media were removed, and cells were infected as described above.

Table 5.1 gRNAs used to knockout *Aa*FUR1 and *Aa*FUR2 in Aag2 cells

Gene	gRNA	PAM	Strand	Exon	Motif	Clones with indels
*Aa*FUR1	G-ATGTTGCCCCCCGTCCACGA	CGG	Anti-sense	2	X-region	0
	G-ACCCCAAGTGGGGCGAGATG	TGG	Sense	3	Prosegment/Catalytic	1
*Aa*FUR2	GACGACTCCAGCGCACTAAC	GGG	Anti-sense	2	X-region	1
	G-CGGGTTAACTATCGGGACAC	CGG	Sense	4	Prosegment/Catalytic	0

5.3.6 Protein detection, Western Blots and SDS-PAGE

For rapid detection of expression of Hfurin and furin1 in the trans-complementation experiment, HiBiT signal was detected using the Nano-Glo® HiBiT Extracellular Detection

System (Promega N2420). All SDS-PAGE gels were run as 10% acrylamide gels under denaturing conditions. The Western blots for prM and E were done under non-reducing conditions. prM antibodies DENV2 (Invitrogen PA5-34967) and ZIKV (Genetex: GTX133584). E antibody 4G2.

5.3.7 Real time RT-qPCR and virus titration

A standard curve for DENV2 and ZIKV was developed as follow. First, DENV2 (PD2ICMO, DENV-2 16681) and ZIKV (pFLZIKV, ZIKV FSS13025) cDNA clones were linearized with XbaI and ClaI, respectively. Then, a T7 in vitro transcription reaction was performed following manufacturer instructions. The resulting RNA was treated with DNaseI (Cat. NEB M0303), purified using RNeasy Minikit, serially diluted 1 to 10 and an RT-qPCR was performed using SuperScript™ III Platinum™ One-Step RT-qPCR Kit (Invitrogen Cat. 11732020). Primers used for the RT-qPCR: DENV2: F-TTGCGGTGTCAATGGCTAACA, R-CCAATGCGTTCAATCGGCT; ZIKV: F-CCGCTGCCCAACACAAG, R-CCACTAACGTTCTTTTGCAGACAT.
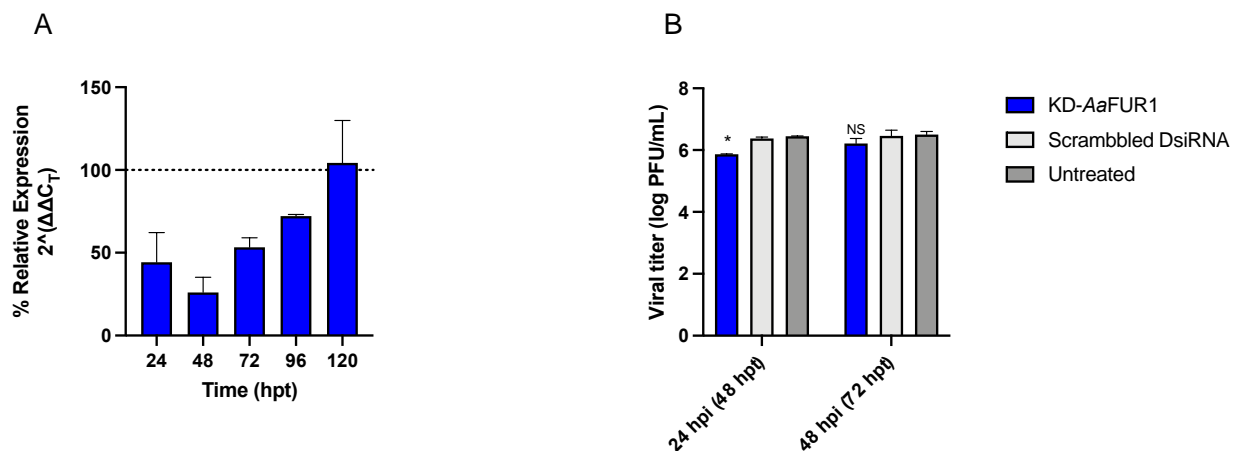
Figure 5.1 Effect of DsiRNA knockout on DENV2 infection of C6/36 cells
A. Relative expression of furin1 overtime after transfection of the DsiRNA. The cells were treated with an scrambled DsiRNA as a control. Rps17 was used as a reference. B. Viral titer of DENV2 at 24 and 48 hours post infection, which correspond to 48 and 72 hours post transfection with the DsiRNA. Statistical significance was determined using a two-tailed t-test, where $P<0.05$ was significant. N=2 independent experiments.

Virus titration was performed by standard plaque assays in a 12-well plate format. Briefly, monolayers of Vero or BHK cells were infected with 200 µL of six 1:10 dilutions of ZIKV or DENV, respectively. After infection at room temperature for 2 hours, an overlay of 1% agarose

plus media was added to the cells. For counting plaques, cells were stained with 4% neutral red at 3- and 5-days post infection for ZIKV and DENV2, respectively. Titer was calculated as plaques/(dilution factor x 0.2) = PFU/mL.
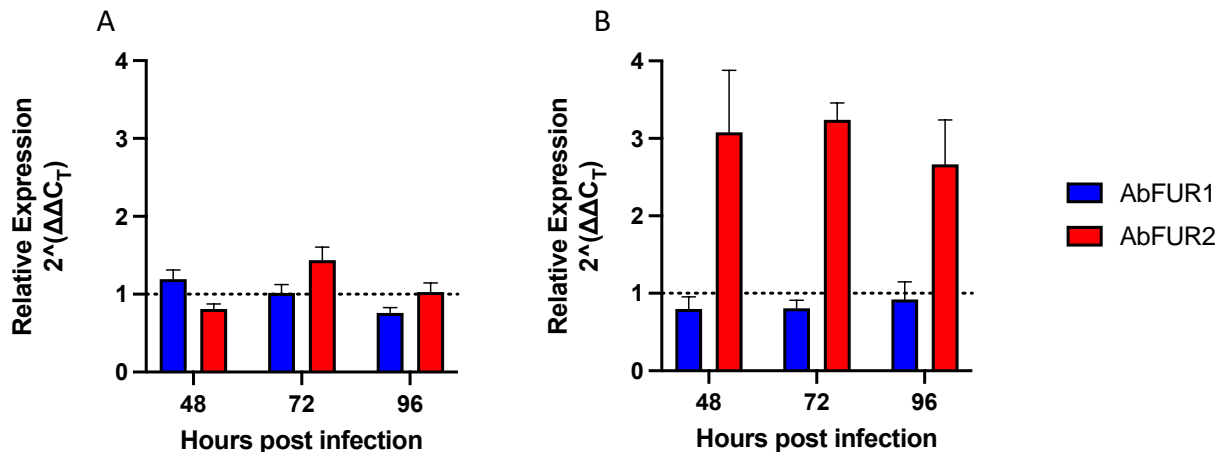
Figure 5.2 Relative expression of furin1 and furin2 after infection with DENV2 or ZIKV
C6/36 cells were infected with DENV2 (A) or ZIKV (B). 48 hours later the cellular mRNA was extracted for quantification of furin1 and furin2 expression. Data is shown relative to uninfected cells and using the housekeeping Rps17 as a reference. N=2 independent experiments.

5.4 Results

5.4.1 RNAi silencing of *Ab*FUR1 and expression levels upon infection in C6/36 cells

C6/36 cells were initially selected as a mosquito model to study Flavivirus maturation because it is a robust cell line that produces high titer of different viruses. However, this cell line was originally derived from *Aedes albopictus*[152], which is not the preferred vector for viruses like DENV2 and ZIKV. In addition, this cell line has been shown to be deficient in Dicer activity[153]. Despite these facts, knockdown with siRNA against AbFUR1 -*Aedes albopictus* furin1- was attempted. The dicer substrate, DsiRNA, against *Aa*FUR1 showed the maximum reduction at 48 hours post transfection in C6/36 cells (Figure 5.1A), however the RNA levels went almost back to normal at 90 hours later. Infectivity of DENV2 was tested in cells with a reduction in expression of *Ab*FUR1. At 24 hours post infection, which correspond to 48 hours post transfection with the DsiRNA, there was a slight reduction in titer of only half a log, which corresponds to around 1.3 million PFU of reduction (Figure 5.1B). The difference at 48 hours post infection was not significant. The DsiRNA inhibition in C6/36 cells was also attempted for

*Ab*FUR2 and SPCA, however for both genes there was no reduction in the transcript levels as was seen for *Ab*FUR1.

The next question was to see if the levels of expression of these genes would change in the presence of viral infection. For that, two viruses were tested: DENV2 and ZIKV. Unlike ZIKV, DENV2 is known for having an inefficient furin cleavage. Upon infection with DENV2, no difference in expression was detected for either *Ab*FUR1 nor *Ab*FUR2 (Figure 5.2A). However with ZIKV, the expression of *Ab*FUR2 increased by three fold (Figure 5.2B).
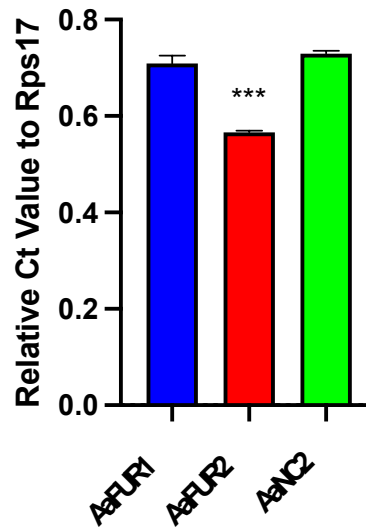
Figure 5.3 Relative expression of furin1, furin2 and NC2 in Aag2 cells
RNA was extracted from sub-confluent cells. RNA levels were quantified using Rps17 as reference. Statistical significance was determined using a two-tailed t-test. N=3 independent replicates.

5.4.2   Generation and characterization of *Aa*FUR1 and *Aa*FUR2 mutant cells.

Given that the genome of *Aedes aegypti* is better annotated and is the preferred vector of viruses like DENV2 and ZIKV, the Aag2 cells were selected for further genomic mutagenesis studies. First, the expression levels of the proprotein convertases were checked with relative qRT-PCR. AaFUR2 showed the lowest expression levels (Figure 5.3), whereas both *Aa*FUR1 and *Aa*NC2 remained at higher expression. To produce mosquito Aag2 cells deficient in furin activity, we used CRISPr technology to induce mutagenesis with NHEJ. We designed two gRNAs per gene (Table 5.1) and cloned each of them under the U6 promoter of the pAc-sgRNA-Cas9 plasmid. We transfected Aag2 cells with the corresponding plasmids, selected with puromycin for seven days and isolated single clones with limiting dilution (Figures 5.4 and 5.5A). Clones were expanded

and screened with T7E1 endonuclease and sequencing. The *Aa*FUR1-gRNA2-clone-6, from now on called *Aa*FUR1_mut, is homozygous and has a deletion of amino acid E311 (Figure 5.5B and 5.6B), which is adjacent to the predicted N-terminal region of the catalytic domain. Since the mutation in *Aa*FUR1 was a deletion of a complete codon, the protein did not truncate and can theoretically still be expressed. Conversely, *Aa*FUR2-gRNA1-clone-4, from now on called *Aa*FUR2_mut, is heterozygous (Figures 5.5C and 5.6C), producing a truncated protein of <102 amino acids.
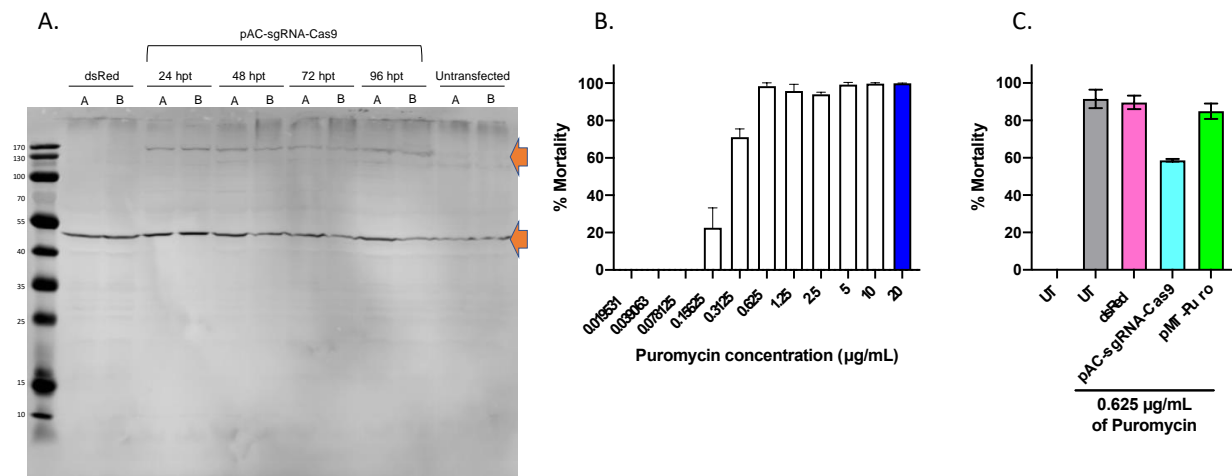
Figure 5.4 Puromycin selection and Cas9 expression in Aag2 cells
(A) Western Blot against FLAG to detect expression of Cas9 in Aag2 cells transfected with the pAc-sgRNA-Cas9 vector. The top orange arrow represents the expected size for Cas9. Bottom orange arrow represents the GAPDH cargo control. (B) Puromycin selection curve of Aag2 cells. The concentration selected was the lowest that would kill 100% of untransfected cells, in this case 0.625 µg/mL. (C) Percent mortality of cells transfected with different plasmids and treated with puromycin. UT is untransfected. As shown here, the cells transfected with the pAc-sgRNA-Cas9, which carries a puromycin selection marker has 40% reduction in mortality. N=3 independent replicates.

To determine whether the deletion of E311 in AaFUR1_mut would make an inactive protease, we expressed recombinant soluble protein in S2 cells. In addition to the wild type *Aa*FUR1, we generated two recombinant *Aa*FUR1_mut (delE311). All recombinant proteins were cloned under the regulation of the inducible promoter metallothionein and tagged with FLAG, HiBiT and 6xHis (Figure 5.7). 24 hours post transfection, cells were induced with $CuSO_4$. Supernatant was collected 48 hours post induction for analysis. WT and mutant AaFUR1 showed HiBiT signal in the supernatant, indicating that the delE311 mutation does not interfere with expression (Figure 5.5D), but the furin activity is significantly compromised (Figure 5.5E). This indicates that the cell line *Aa*FUR1_mut is deficient in furin1 activity.
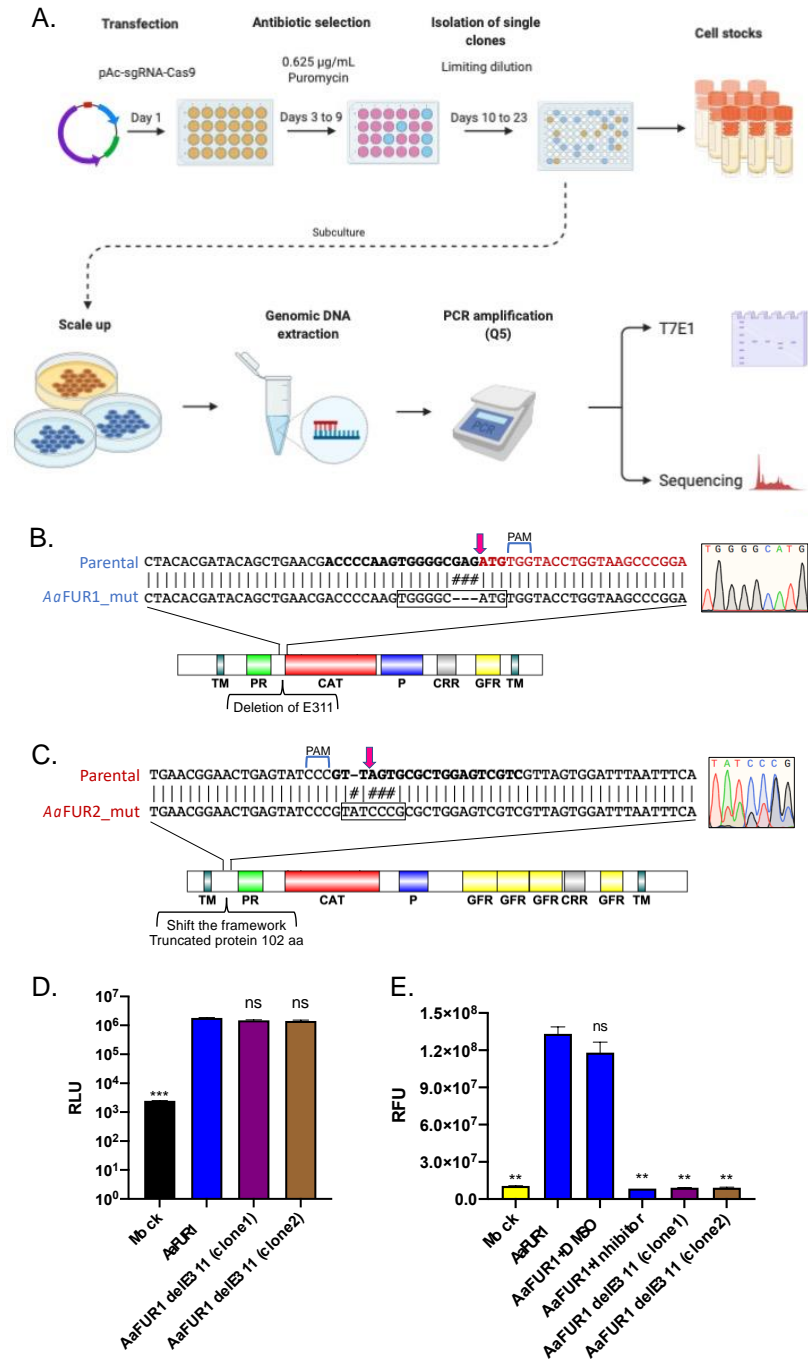
Figure 5.5 Generation and characterization of AaFUR1 and AaFUR2 mutant cells
(A) Schematic of the procedure to generate CRISPr-KO mutant Aag2 cells. (B & C) Sequence alignment of the region spanning the introduced mutations in AaFUR1 (B) and AaFUR2 (C). The gRNA sequence is bolded and the Cas9 cleavage site is indicated by a red arrow, followed by the corresponding Protospacer Adjacent Motif (PAM). (b) The location of the codon GAG is indicated in the protein models. (D) HiBiT signal from supernatant of S2 cells transfected with recombinant soluble AaFUR1. (E) furin activity assay of supernatant of S2 cells transfected with either wild type AaFUR1 or two biological clones with deletion of E311. The furin inhibitor I (Decanoyl-RVKR-CMK) was used as a control to demonstrate that the activity in the supernatant was coming from furin1 alone and no other contaminants. RLU: Relative light units. RFU: Relative fluorescent units. Statistical significance was determined using a two-tailed t-test where $P<0.05$ is significant. N=2.

109

5.4.3 Infectivity of DENV2 and ZIKV is reduced in AaFUR1_mut cells.

Next, we aimed at determining the lack of furin activity in the maturation of flaviviruses. For that, we compared two flaviviruses for efficient and partial maturation efficiency: ZIKV and DENV2, respectively. First, we performed a growth curve in the AaFUR1_mut and AaFUR2_mut cells. Viruses produced AaFUR1_mut cells showed reduced titer levels (Figures 5.8 A and F), suggesting that knockout of AaFUR2 does not affect the titer. Then, we performed a larger infection of parental and AaFUR1_mut cells and a sucrose cushion 72 hours post inoculation (Figures 5.8B and 3G). A western blot with the sucrose cushion indicated that DENV displays prM from both parental and AaFUR_mut cells, but ZIKV prM content is doubled in AaFUR1_mut cells (Figures 5.8C-F and H-I). Finally, we performed a trans-complementation experiment to restore furin activity. For that, we transfected different recombinant proteases under the constitutive promoter pUB. 24 hours post transfection, we infected cells at an MOI=10. For DENV2, introduction of recombinant AaFUR1 increases the titer (Figure 5.8E), whereas the titer for ZIKV is increased in the presence of human furin, AaFUR1 and AaFUR2 (Figure 5.8J).

PCR amplicon	Expected size (bp)
*Aa*FUR1-gRNA1	993
*Aa*FUR1-gRNA2	829
*Aa*FUR2-gRNA1	1005
*Aa*FUR2-gRNA2	951
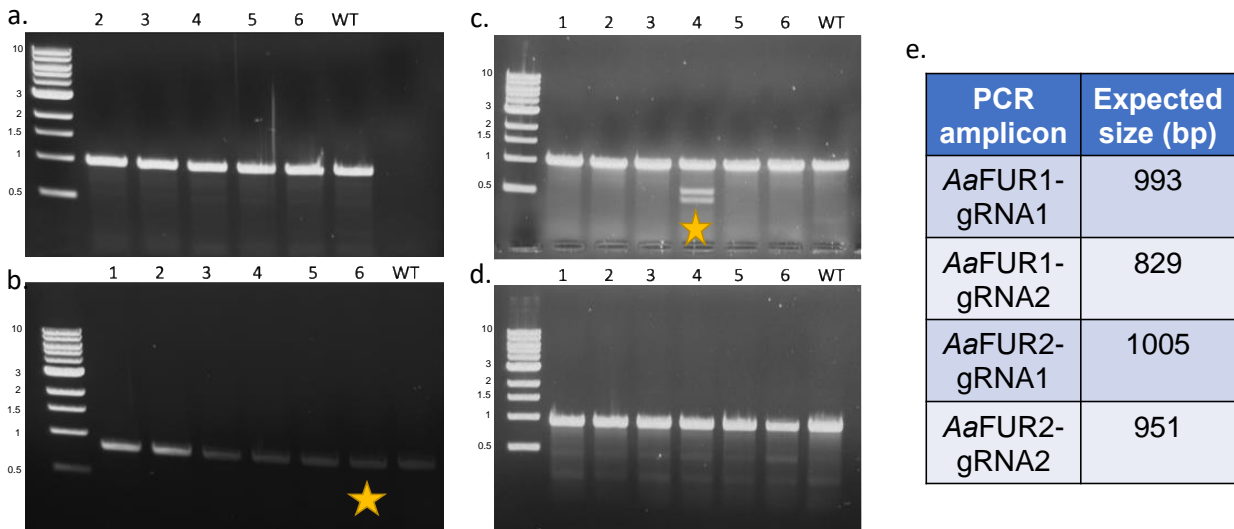
Figure 5.6 T7E1 screening of single cell isolates
(A) AaFUR1-gRNA1 (B) AaFUR1-gRNA2 (C) AaFUR2-gRNA1 (D) AaFUR2-gRNA2. The yellow stars highlight the mutants that were shown to have mutations, detected by Sanger sequencing. (E) expected sizes for each PCR amplicon.

5.4.4   Infectivity of SINV is reduced in AaFUR1_mut cells in a MOI dependent manner

Next, we aimed to check whether the lack of furin activiy in Aag2 cells would affect the infectivity of other viruses, the alphaviruses. For that, SINV was chosen given its high infectivity and versatility of its genome. In addition, as mentioned in Chapter 4, SINV has an optimized furin cleavage site. First, RNA from SINV was produced using the toto64 cDNA clone [64]. The RNA was transfected into Aag2 (both parental and mutant) and supernatant was collected every 24 hours for 5 days. As expected, the *Aa*FUR1_mut Aag2 cells showed a reduced viral titer, of about two logarithmic values of difference (Figure 5.9A). The viral RNA genomes were quantified using an standard curve. The viral genomes detected in the supernatant decreased over time in both types of cells, but with no significant differences between parental and *Aa*FUR1_mut Aag2 cells (Figure 5.9B). In addition, a trans-complementation experiment was performed to determine whether transient transfection with recombinant DNA would rescue the viral titer in the mutant cells. The experiment was done in parental cells as a control to check whether the transfection was having an influence in titer. There was a reduction in titer in the transfected parental cells, however, all the comparisons between the *Aa*FUR1_mut and parental Aag2 cells were made based on the dsRed transfection control. Transfection with human furin, *Aa*FUR1 and *Aa*FUR2 was able to recover the titer in mutant cells to levels equivalent to the parental cells transfected with dsRed, serving as a proof that the mutant *Aa*FUR1 is the only factor affecting the viral infectivity.
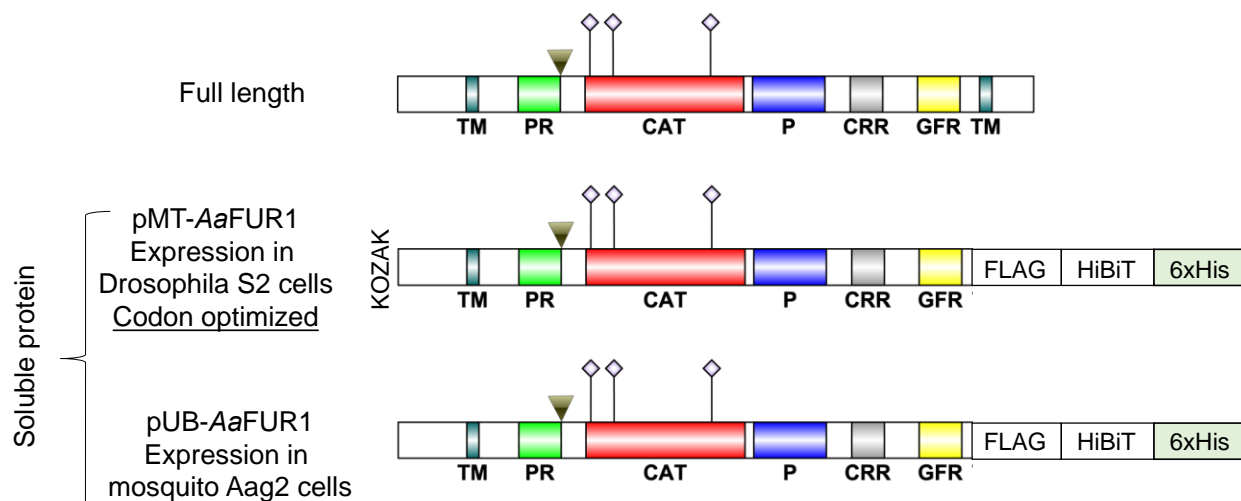
Figure 5.7 Recombinant *Aa*FUR1 constructs
Constructs created for recombinant expression and trans-complementation assays. TM: transmembrane region, PR: prosegment, CAT: catalytic domain, P: P-domain, CRR: cysteine-rich region, GFR, tyrosine-like growth factor receptor.

As stated above, the experiment with SINV were initially performed from transfection of viral RNA produced by *in vitro* transcription. However, the transfection efficiency was less than 5%, suggesting that the initial amount of virus produced was too little. To address this issue, the cells were infected this time with virus at different MOIs (Figure 5.10). Similar to what was observed with viral RNA transfection, the viral titer did not increase over time and stayed at lower levels than the parental cells. Still, the titer was much higher but the difference between parental and mutant cells increased overtime, especially in the cells infected at lower MOIs.
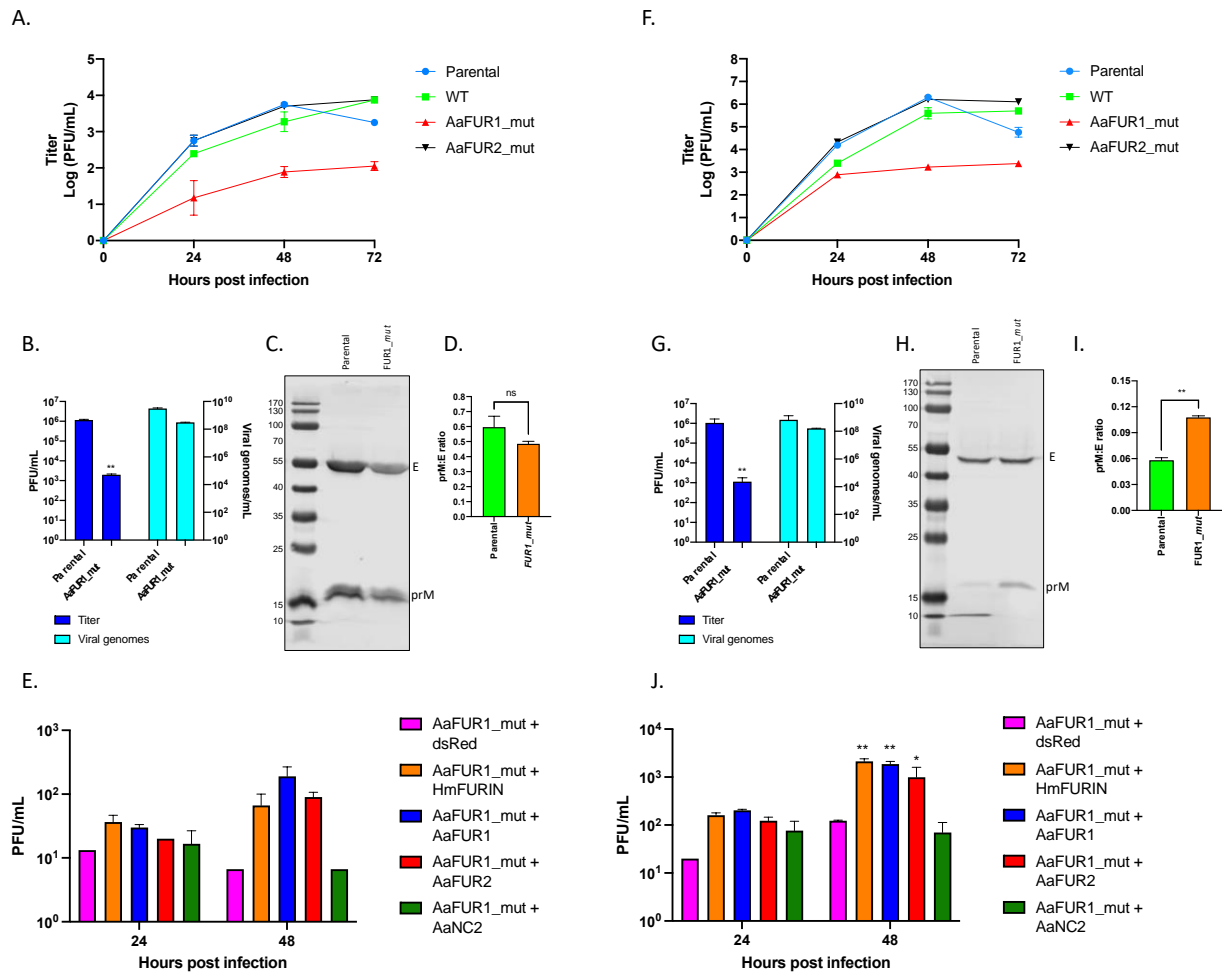
Figure 5.8 Infectivity of DENV-2 and ZIKV is reduced in AaFUR1_mut cells
(A-E) DENV2 (F-J) ZIKV: One step growth curve of DENV2 (A) and ZIKV (F) in different Aag2 cells at a MOI=10. WT: Wild type, is a cell line transfected with the gRNA-AaFUR1 but with no induced mutagenesis. (B & G) Viral titer and genomes 72 hours post infection (MOI=10). (c & h) Western blot for prM and E from partially purified virus of infected parental and AaFUR1_mut cells. (D & I) prM to E ratio calculated from the Western Blot. (E and J) Trans-complementation of different proteases in AaFUR1_mut cells. UT: untransfected cells. dsRED worked as a transfection efficiency control (=40%). Statistical significance was determined using a two-tailed t-test where P<0.05 is significant. N=2, two individual replicates.

5.5   Discussion

Mosquitoes are the primary route of transmission of multiple medically relevant arboviruses including several flaviviruses and alphaviruses. The most common vector is *Aedes aegypti,* a highly anthropophilic mosquito, responsible for transmitting dengue, Zika, yellow fever and chikungunya viruses. These viruses have caused significant disease burden in human history and continue to be a problem today. For this reason, studying mosquito-virus interactions are a valuable tool to understand special elements that can be used for development of strategies to control the disease before it is transmitted to the humans.
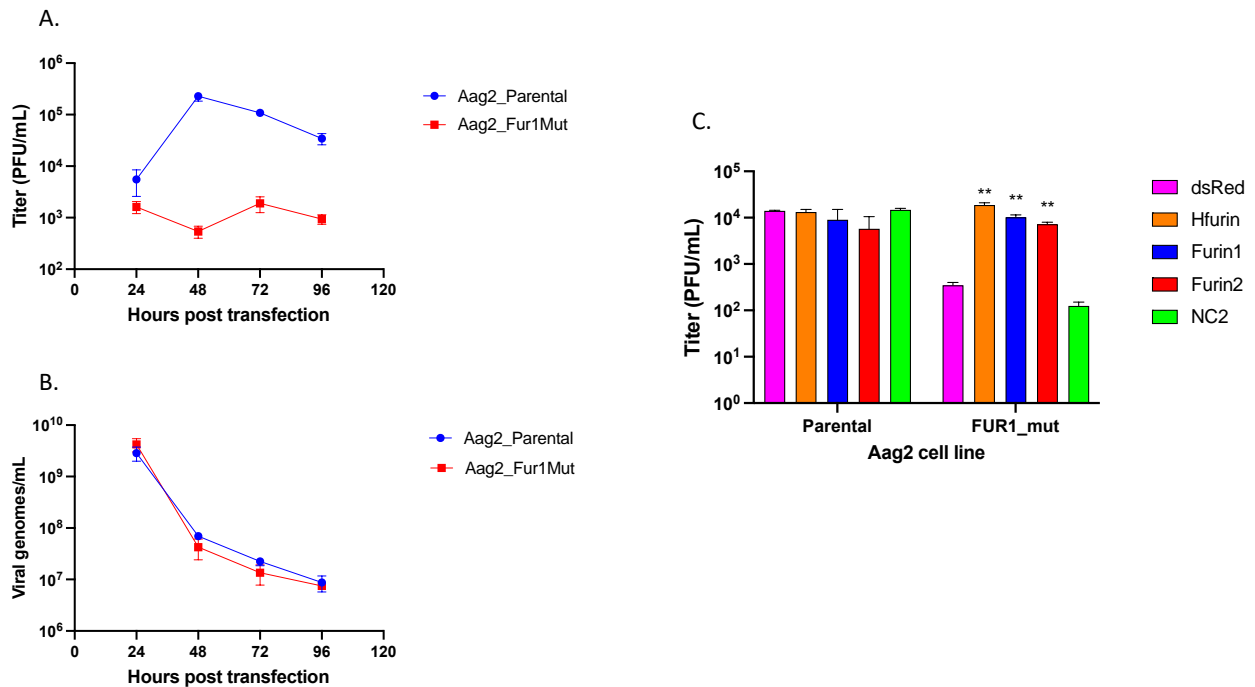
Figure 5.9 Infectivity of SINV virus is reduced in AaFUR1_mut cells
(A) Titer of SINV after viral RNA transfection. (B) Viral genomes. (C) Trans-complementation of different proteases in AaFUR1_mut cells. UT: untransfected cells. dsRED worked as a transfection efficiency control (=40%). N=2. The RNA transfection efficiency (A & B) was less than 5%. The transfection efficiency of plasmid DNA for trans-complementation was 60%. Statistical significance was determined using a two-tailed t-test where P<0.05 is significant. N=2. Two individual replicates.

The maturation of different viruses occurs after translation and can happen before or after assembly of the viral particles. The flaviviruses are assembled as immature particles, and they transit through the trans-Golgi network, where they get cleaved at the prM junction by the host protease furin. This protease is implicated in the processing not only of flaviviruses, but also of many other viral families. However, the features of the mosquito furin or furin-like proteases have

113

not been addressed. Identifying the protease in the mosquito that is responsible for maturation of flaviviruses such as DENV or ZIKV, would provide a target for antiviral strategies that prevent the disease transmission.
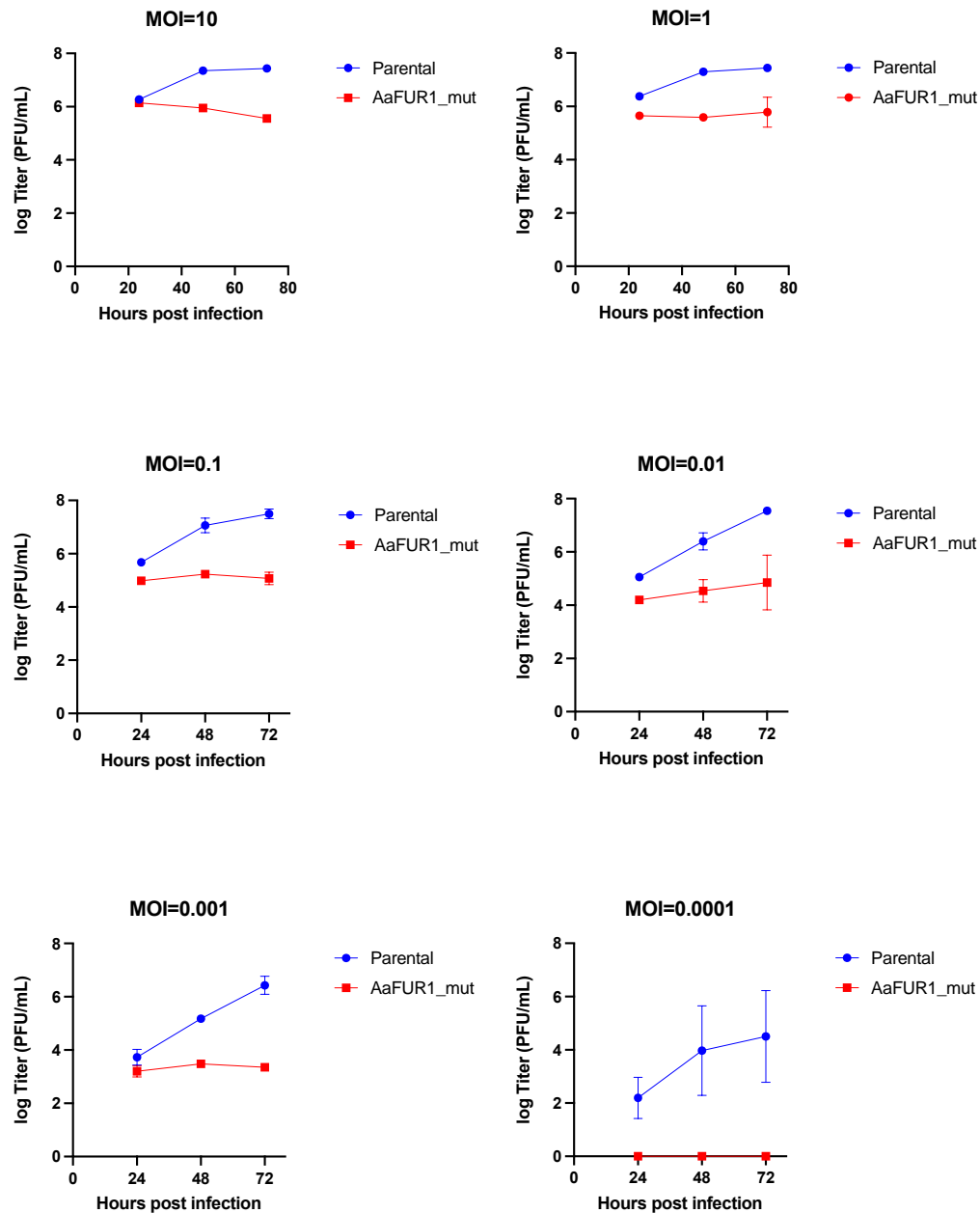
Figure 5.10 Growth curve of SINV at different MOIs in Parental and AaFUR1_mut cells
Aag2 cells were infected with SINV at different MOIs. Supernatant was collected a different timepoints and checked for viral titer. N=2 individual biological replicates.

In this chapter, two gene silencing approaches were used to do a loss of function assay. First the C6/36 cells, originally derived from *Aedes albopictus,* the alternate vector of ZIKV[154], were treated with DsiRNA, a reagent produced by Integrated DNA Technologies (IDT). The siRNA normally functions by targeting mRNA in the cells and directing it towards degradation with the RISC complex. In the case of DsiRNA, these are longer RNA molecules that anneal to the target RNA but before going to the RISC complex, it becomes substrate of DICR for prior processing. In the case of C6/36, it has been shown that these cells lack DICR activity and the siRNA pathway [153].

Therefore, gene silencing using RNA in these cells is extremely challenging and it explains why *Aa*FUR2 was not able to be knocked down. However, *Aa*FUR1 was successfully silenced during a very short period of time, before the expression levels ramped up back to normal. Despite, a small reduction in titer of DENV2 was observable at 24 hours post infection, suggesting that *Aa*FUR1 might have a role in the infectivity.

Interestingly, the expression of both *Aa*FUR1 and *Aa*FUR2 remained unchanged overtime during infection with DENV2. However, infection with ZIKV caused an increase of *Aa*FUR2 expression. DENV2 is known for having a reduced furin cleavage efficiency because it has a negatively charged amino acid at position P3, whereas ZIKV has an uncharged residue in the same site [47]. The results displayed here would suggest that infection with ZIKV requires more furin to process the optimized site and therefore the host would try to compensate by producing more protein as a way to process accumulated precursor protein in the TGN. This fact needs to be carefully addressed experimentally. One strategy would be to identify the promoter of *Aa*FUR2 and determine whether it is activated in the presence of ZIKV or other flaviviruses with optimized cleavage site.

The second gene silencing strategy utilized in this study was CRISPr-Cas9. This technology allows the manipulation the genome by introducing indels (insertions/deletions) that can disrupt the framework of the coding sequence, rendering a protein inactive by encountering a premature stop codon or producing a sequence useless to the cell. In this occasion, the CRISPr-Cas9 genome editing was performed in Aag2 cells, originally derived from *Aedes aegypti*, which has a better annotated genome[155]. The gRNAs were designed to target a region early in the

coding sequence of the protein, so that the catalytic domain would be disturbed, and the cells would be deficient in furin activity. It is important to point out that the mutation obtained in *Aa*FUR1 corresponded to the deletion of three nucleotides, which together would make up for a triplet encoding the Glutamic acid residue at position 311 (E311). This amino acid residue is right before the predicted catalytic domain and to this date, had not been addressed. In LoVo cells, derived from carcinoma epithelial tissue, the W547R mutation in the P-domain rendered furin inactive in one of the alleles[156], [157]. In the case of Aag2 cells, the E311del mutation is homozygous, present in both chromatids, and as shown in recombinant protein assays, this amino acid residue is necessary for activity.

It was also shown that this mutation reduced the infectivity of three different viruses: DENV2, ZIKV and SINV. In the case of DENV2, the maturation is normally inefficient, as is shown that the virus produced in parental cells still had prM content. However further reduction in the cleavage by lack of furin activity reduced the viral titer to even lower levels. In the case of ZIKV, the reduction of prM content is very clear and the reduction in titer is also more noticeable. Similarly, despite having a different maturation pathway, SINV titer was also reduced in *Aa*FUR1_mut Aag2 cells. Moreover, in the case of SINV, the reduction in infectivity was more evident at lower MOIs, suggesting that the virus produced in furin deficient cells is not able to spread to other furin deficient cells. But, when infecting the BHK cells for the plaque assays, the maturation can happen at entry because BHKs have intact furin activity. These results are consistent with Chapter 4.

It is also important to point out that the knockout of *Aa*FUR2 did not have any effect in DENV2 and ZIKV infection. This is because in parental cells, the expression levels of *Aa*FUR2 in Aag2 cells is lower than *Aa*FUR1. Therefore, if *Aa*FUR2 is silenced, the virus can still rely on *Aa*FUR1 without any change. Once again, this study was done in Aag2 cells, which does not necessarily translates to the actual organism. In the mosquitoes, it would be necessary to determine whether *Aa*FUR2 plays a role in the infection of this viruses, because as shown in the trans-complementation experiments, the addition of *Aa*FUR2 to *Aa*FUR1_mut Aag2 cells recovered the titer to similar levels as the parental control.

In conclusion, the silencing and functional studies shown in this study suggest that the furin1 of mosquitoes plays a key role in the maturation and infectivity of DENV2, ZIKV and SINV. In addition, the first insect cell line deficient in furin activity was generated in this study and demonstrates that a single amino acid deletion before the catalytic domain plays a role in viral infection. Combined, these results point out a potential target to develop strategies to control diseases in the insect vector before they can be transmitted to the next human host.

CHAPTER 6.   METHOD DEVELOPMENT FOR SCREENING OF INHIBITORY DRUGS AGAINST SARS-COV-2

6.1   Chapter Summary

The coronavirideae family are a group of medically relevant viruses that have caused several global concerns; the current COVID-19 (coronavirus disease 2019) pandemic is an example of the harm that these viruses can cause. The highly pathogenic coronaviruses such as SARS-CoV, MERS-CoV and SARS-CoV-2 require proteolytic cleavage at the S2' site of the spike protein. This cleavage is necessary for exposure of the fusion peptide and virus entry into the cytoplasm of the cell. While this event is not fully understood, there is consensus that either host TMPRSS2 (transmembrane serine protease 2) cleaves at the plasma membrane or Cathepsin cleaves in endosomes. In addition to the S2' cleavage site, there is an insertion of a furin cleavage site at the junction of S1 and S2 domains of the spike protein of SARS-CoV-2, the functional implication of which is far from clear. In this chapter, we show that the spike protein is cleaved at the furin site (S1/S2) before exiting the cell. The importance of this cleavage for viral biology and its potential for antiviral therapeutics is being explored. In Part 2 of the chapter, an assay was optimized for testing antiviral drugs against the virus, and it was used to screen inhibitory compounds against the viral protease M^{PRO} (main protease). Inhibiting M^{PRO} will prevent the cleavage of the viral polyprotein and hence reduce viral replication. The assay relies on cell death or cytopathic effect (CPE) that SARS-CoV-2 induces in Vero E6 cells; if a compound inhibits the virus, the cells will stay alive upon exposure to the virus and therefore cell survival is used as the assay read-out. Finally, to identify physiological cell culture models for studying SARS-CoV-2, lung progenitor cells derived from induced pluripotent stem cells (iPSCs) and kindly provided by Eyestem (India) were used for testing their permissiveness to SARS-CoV-2. We show that these cells support SARS-CoV-2 infection though the titer obtained was lower than what is observed in Vero E6. However, despite having low viral infection and inherent challenges in growing these cells, the human progenitor cells are more biologically relevant and could be used in confirmation of the promising lead compounds.

6.2   Introduction

The coronaviruses belong to the order of nidovirales and contain a subfamily called coronavirinae. This subfamily is divided into four genera, namely alpha-coronavirus, beta-coronavirus, gamma-coronavirus and delta-coronavirus. The alpha- and beta-coronaviruses include human pathogens, whereas gamma and delta coronaviruses infect other animals. The three highly pathogenic human coronaviruses: SARS-CoV, MERS-CoV and SARS-CoV-2 are beta-coronaviruses. With over 29 kb, the genome of the coronaviruses is among the largest of the RNA viruses. As was discussed in chapter 1, first two thirds of this genome is dedicated to viral replication and host evasion, and the latter one third is dedicated to structural proteins that make the viral particle.
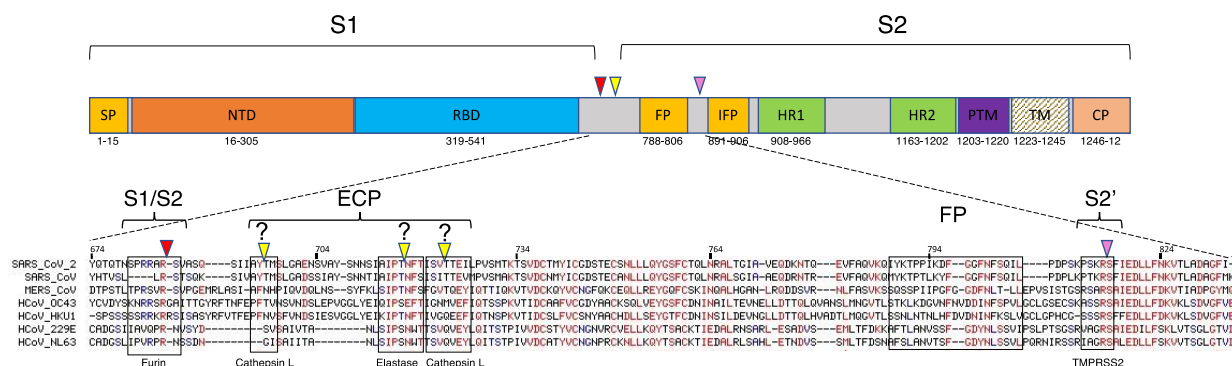
Figure 6.1 Representation of SARS-CoV-2 Spike and the cleavage sites
Domains and cleavage sites were determined from different approaches. SignalP 5.0 (http://www.cbs.dtu.dk/services/SignalP/) was used to predict the Signal Peptide (SP). The homology to SARS-CoV was used to map the N-Terminal Domain (NTD), Receptor Binding Domain (RBD), Fusion Peptide (FP), Internal Fusion Peptide (IFP), Heptad Repeat 1 and 2 (HR1 and HRP2), Pre-transmembrane domain (PTM) and Cytoplasmic domain (CP) for SARS-CoV-2. The transmembrane region (TM) was predicted using SOSUI server (http://harrier.nagahama-i-bio.ac.jp/sosui/sosui_submit.html). S1/S2 and S2' sites were mapped based on homology to other human coronaviruses and are highlighted with red and pink arrowheads, respectively. The ECP site(s) are highlighted with yellow arrowhead. ECP site(s) were predicted by homology to other coronaviruses (site AY-TM) or by prediction (sites AIPT-NFT and SV-TTEI) using iProt-Sub (http://iprot-sub.erc.monash.edu.au). Sequences of the maturation region was aligned with other human beta- (SARS-CoV, MERS-CoV, OC43 and HKU1) and alpha-coronaviruses (229E and NL63).

The spike protein is present as a homotrimer on the surface of the virus and is structurally divided into two functional domains: S1 and S2. The S1 domain contains the receptor binding domain (RBD) which provides the attachment specificity depending on the receptor of the virus. In contrast, the S2 contains the fusion peptide, which is necessary for viral fusion with the host membrane. The S2' site, which is located upstream of the fusion peptide, must be cleaved to

activate fusion with the host membrane and is therefore critical for viral entry. The exposure of S2' site during viral entry is not understood. Host proteases like TMPRSS2 and cathepsin are believed to drive this cleavage, based on the sub-cellular context (plasma membrane vs endosomes, respectively). As such, their inhibition results in reduction of viral infectivity. However, in the case of SARS-CoV-2, there is an insertion of a multi-basic region which has been suggested to be processed by host furin protease. The consequence of S1/S2 cleavage by furin on viral biology and pathogenesis is not fully understood and whether it plays a role in exposing the S2' site or fusion peptide is not clear. Gaining insight into this process will lead to the development of inhibitory drugs that prevent the processing and activation of the spike protein.

Analogous to the proteolytic processing of the spike protein, the polyproteins pp1a and pp1ab are processed by two viral cysteine proteases, PL^{PRO} (within nsp3) and M^{PRO} (also known as 3CL or nsp5). M^{PRO} is known as the main protease because it is implicated in cleavage of most of the polyprotein into functional proteins. Therefore, if the M^{PRO} is inhibited, the replication and transcription machinery will not be formed, thus halting the viral life cycle. For this reason and the substrate specificity of M^{PRO}, it is an ideal candidate for the development of lead compounds for antiviral therapeutics.

Figure 6.2 Representation of the clones of the Spike protein
Transmembrane Spike represents the protein as it is in the virus. The soluble Spike lacks the transmembrane region and has a TEV protease site, a Foldon Trimerization Domain, Flag-tag, HiBiT peptide and 8xHis tag. The S1/S2 clone is has the SPRR site deleted. S2' clone contains a mutation (K814A).

This chapter details work done together with Dr. Devika Sirohi (Kuhn Laboratory, Purdue University) in collaboration with Dr. Andrew Mesecar & Dr. Arun Ghosh (M^{PRO} compounds), Dr. Ramaswamy Subramanian, Dr. Michael Poderycki and Eyestem Research Private Limited (iPSC derived lung progenitor cells). The furin cleavage site of SARS-CoV-2 was probed using a heterologous expression system, corroborating data that others have now published. In addition,

we optimized a method for testing inhibitory drugs against SARS-CoV-2. The method utilizes a luciferase-based signal for assessment of cytopathic effect caused by the virus. This method can be adapted for different compounds that target various aspects of the viral life cycle. We have tested compounds against the Spike protein, M^{PRO} and various host proteins, provided by multiple academic and industry collaborators. Only representative data of inhibitors against M^{PRO} are shown here and given their proprietary nature, the identities are not disclosed. Finally, we also tested the permissiveness to SARS-CoV-2 of progenitor stem cells and conclude that these cells support SARS-CoV-2 infection, albeit the titer is low compared to the titer obtained from VeroE6.
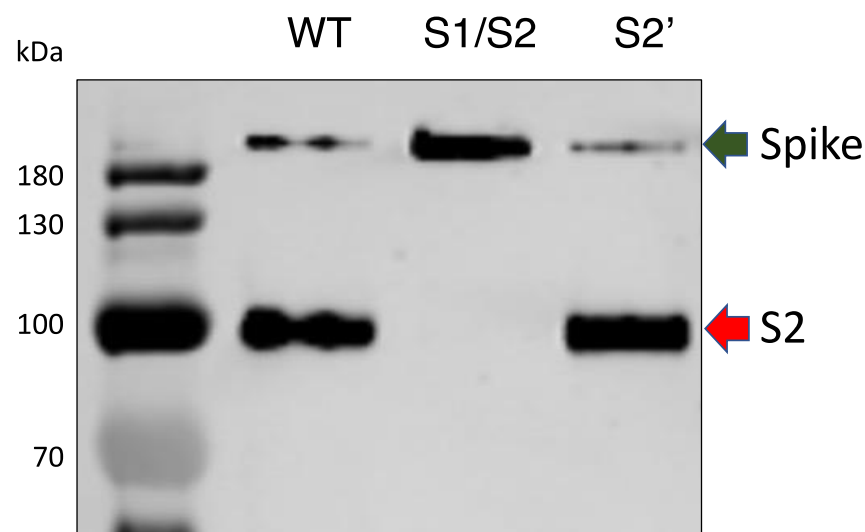
Figure 6.3 The Spike protein of SARS-CoV-2 is cleaved at the S1/S2 site before entry
The Spike protein mutants were transfected into HEK293T cells in a T75 flask. 24 hours post transfection, the supernatant was collected and filtered. Protein concentration was determined with the FLAG® Immunoprecipitation Kit, following manufacturer instructions. The samples were run on a SDS-PAGE under reducing conditions and transferred to a nitrocellulose membrane of 0.45 µm (BioRad). The membrane was then blotted with the primary antibody mouse-anti-Flag overnight. Secondary antibody goat-anti-mouse 800 was added for one hour. The membrane was then imaged in an Odyssey machine (LI-COR) and annotated manually.

6.3 Materials and Methods

6.3.1 Cells and virus

Vero E6 and HEK293T cells were grown in EMEM or DMEM media, respectively, supplemented with 10% FBS at 37 ℃ in an atmosphere of 5% CO_2. Huh7.5 cells were cultured in the same conditions as HEK293T. The virus utilized in this study was produced from RNA provided by Dr. Michael Diamond (Washington University School of Medicine in St. Louis) The

viral RNA was transfected into Huh7.5 cells using the TransiT-mRNA reagent (Mirus 2550) following manufacturer instructions. 72 hours post transfection, the supernatant was removed and tittered. All virus handling was performed in a Biosafety cabinet in BSL3 with proper personal protective equipment: PAPR, N95 and disposable suits/gown/sleeves.

6.3.2 Plasmids

The expression construct from spike protein was obtained from BEI resources (NR-52310, produced by Florian Krammer, Ph.D., Fatima Amanat and Shirin Strohmeier from the Department of Microbiology, Icahn School of Medicine at Mount Sinai, New York, New York, USA). For ease of gene manipulation, the region encoding the spike ectodomain was sub-cloned into the pcDNA3.0 plasmid along with foldon trimerization motif and FLAG, HiBiT and 8xHis tags using Gibson assembly. For the generation of the spike mutants, site directed mutagenesis was performed as described in chapter 5.

6.3.3 Plaque assays

Virus titration was performed by standard plaque assays in a 6-well plate format. Briefly, monolayers of VeroE6cells were infected with 250 µL of six 1:10 dilutions of SARS-CoV-2. After infection at room temperature for 1 hours, an overlay of 1% agarose in media was added to the cells. For counting plaques, cells were stained with 4% neutral red at 3 days post infection. Titer was calculated as plaques/(dilution factor x 0.25) and reported as PFU/mL

6.4   Results

6.4.1   furin cleavage site on the SARS-CoV-2 Spike protein

There are three sites that have been suggested to be involved in the proteolytic cleavage of the Spike protein of SARS-CoV-2: S1/S2, ECP and S2' (Figure 6.1). The S1/S2 site is the least conserved; present in MERS-CoV and the current SARS-CoV-2 but absent in SARS-CoV. Previous work done on the S1/S2 site of MERS-CoV suggest that the cleavage might happen during secretion, but whether or not furin is the main player, remains a topic of controversy [54]–[56].

10 µL 10 µL 10 µL 10 µL 10 µL 10 µL 10 µL 10 µL

10 µL of DMSO (×8, one per tube)

Stock = 20 mM | 10 mM | 5 mM | 2.5 mM | 1.25 mM | 0.625 mM | 0.3125 mM | 0.15625 mM | 0.078 mM

8µL (per tube)

800µL EMEM + 2% FBS (per tube)

2X 200 µM | 100 µM | 50 µM | 25 µM | 12.5 µM | 6.25 µM | 3.125 µM | 1.5625 µM | 0.78 µM

100 µL (per tube)

100 µL of media OR 2X Virus (MOI=0.02)

1X 100 µM | 50 µM | 25 µM | 12.5 µM | 6.25 µM | 3.125 µM | 1.5625 µM | 0.78 µM | 0.39 µM
MOI=0.01 (each)

Add compound + virus mix to the VeroE6 cells
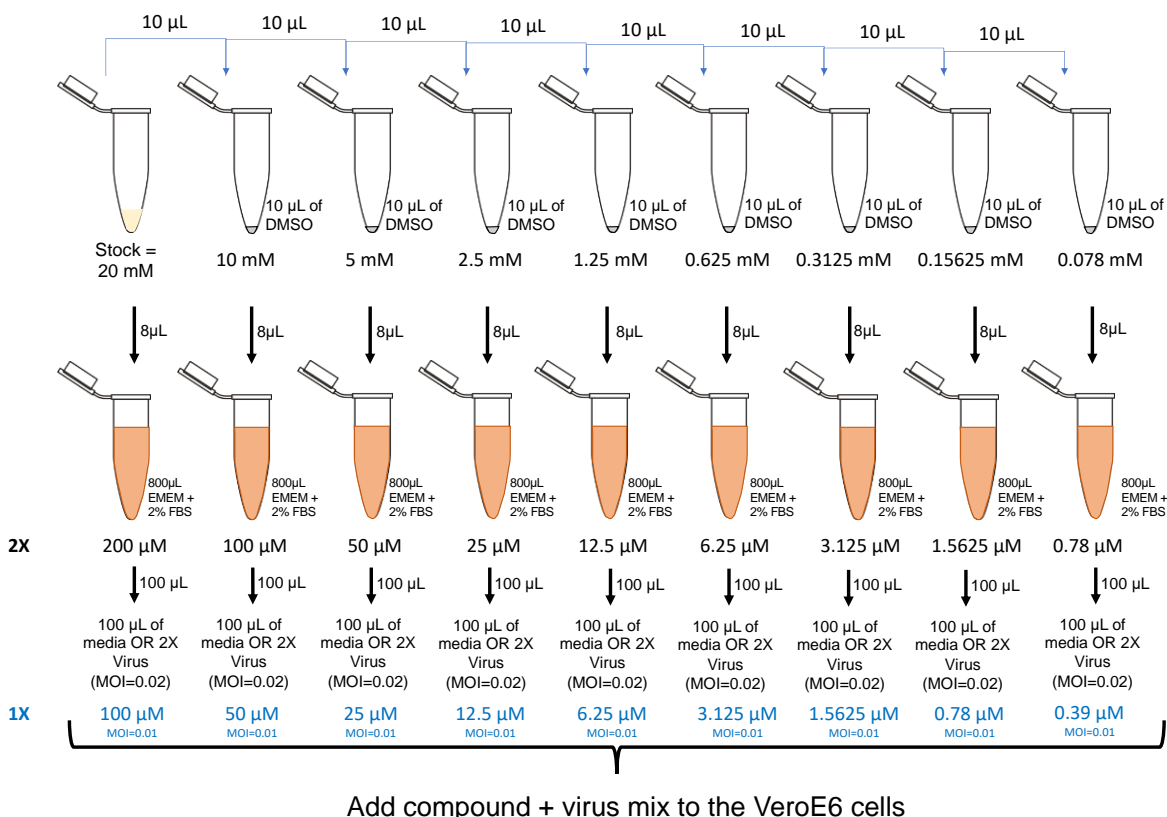
Figure 6.4 Dilution strategy for inhibitory compounds. Compounds are initially diluted in DMSO in a 1:2 dilution ratio. Then, the corresponding dilution is added to medium for reaching concentration of µM

Here, we generated a soluble version of the Spike protein which contains a foldon trimerization domain and three tags: FLAG, HiBiT and 8xHis (Figure 6.2). Versions with mutations in the spike proteins at the S1/S2 and the S2' were also created. Deletion of the furin site at the S1/S2 site prevents the protein from being processed before exiting the cell (Figure 6.3). This cleavage has also been addressed by other studies that used pseudotyped viral particles[158]–[160].

Studies on the ECP and S2' sites are limited. It is known that at least cleavage at the S2' site is essential for viral entry, possibly by mediating fusion of host and viral membranes [53]. Exposure of the fusion peptide (FP) and Internal Fusion Peptide (IFP) (Figure 6.1) is crucial for membrane fusion [53]. Coronavirus entry has been proposed to occur in a sequential process: an initial cleavage happens at the ECP site (either by elastase or Cathepsin L), and results in the exposure of the disordered FP, which primes for the plasma membrane. Then, a second cleavage happens at the S2' site, resulting in the exposure of the IFP and insertion into the plasma membrane

[53], [161]. This model has not been experimentally studied, but the role of Cathepsin L in entry is crucial in certain cell lines. However, the site where the Spike protein is cleaved by Cathepsin L is not known. The motif AYTM has been suggested to be the Cathepsin L site for SARS-CoV, at least by *in vitro* peptide cleavage assays [162]. The same site is slightly different in MERS-CoV (AFNH). However, mutations in this site do not affect entry of pseudotyped viral particles of MERS-CoV, but inhibition of Cathepsin L deprive entry, indicating that Cathepsin L is deprive at other positions in the Spike [163], [164].
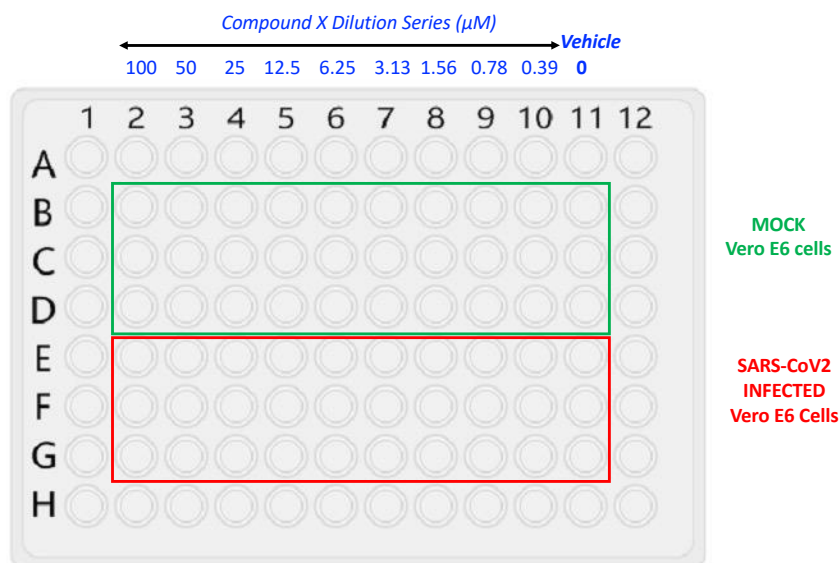
Figure 6.5 Example of plate set up for cytotoxicity-based assay with VeroE6 cells infected with SARS-CoV-2

An infectious cDNA clone of SARS-CoV-2 was generated by the Pei-Yong Shi lab [165]. This clone can be used for reverse genetics in which deletions, insertions or substitutions can be added to the genome of the virus. We are currently working on generating a HiBiT tagged clone, which will allow for rapid detection of using the luminescence-based Nano-Glo® HiBiT assay.

6.4.2   Assay development for testing inhibitory drugs against SARS-CoV-2

Vero E6 cells are seeded in a 96 well plate at a density of 25000 cells per well and grown overnight at 37 C in an atmosphere of 5% CO_2. The compounds, dissolved in DMSO, are serially diluted 1:2. Then, 8 uL of the corresponding compound concentration is further diluted into 800 uL of working media (EMEM supplemented with 2% FBS). This results in a 1 to 100 dilution. At

this point, the concentration of the compound in the media corresponds to 2X the concentration of the desired final concentration. The top three wells are diluted into 1X using working media, whereas the bottom three were diluted using virus, which was also added at a initial concentration of MOI=0.02 for a final of 0.01 (Figures 6.4 and 6.5).
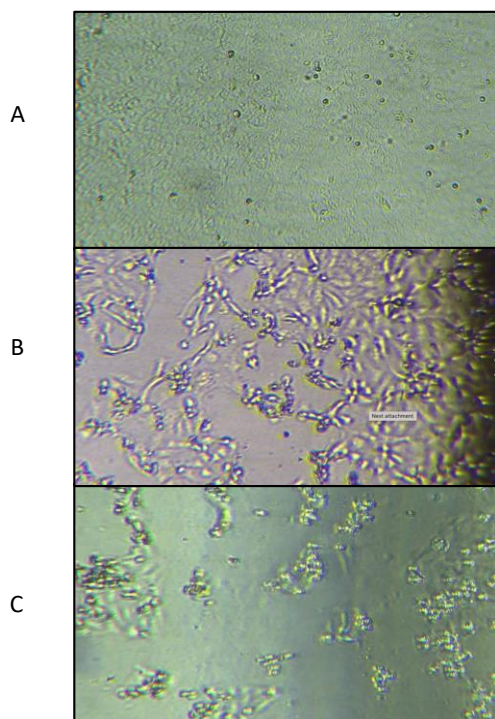
Figure 6.6 Example of cells treated with inhibitory compounds and challenged with SARS-CoV-2
A. Highest concentration on inhibitory compound. B. Suboptimal concentration of inhibitory compound. C. DMSO control

The diluted compound plus media or virus is added to the VeroE6 cells. The plates are incubated at 37 °C for 72 hours. The cells are checked under the microscope for qualitative analysis of compound-based protection (Figure 6.6). Then, a cell viability assay is performed using the CellTiter-Glo 2.0 kit (Promega), following manufacturer instructions. Briefly, cell plates are equilibrated at room temperature for 30 min. The CellTiter-Glo 2.0 is also equilibrated at room temperature and added to wells at a 1:1 v/v ratio (i.e. 100 µL to 100 µL of medium containing cells). The contents are mixed for 2 min on an orbital plate shaker. Then, the plate is incubated for 10 minutes at room temperature before being read on a SpectraMax L luminometer with an integration time of 1 second per well. The signal obtained reflects the ATP present in metabolically active cells. Therefore, if a compound is preventing viral replication, this will

translate into less cell cytotoxicity and a higher luminescence signal. The cell viability of compound-treated and virus exposed cells is compared to vehicle control (i.e. cells treated with 0.5% DMSO and virus). The cells treated with compound but not virus (top three wells in figure 6.5) informs about the compound cytotoxicity and is also used for comparison. Table 6.1 is an example of different compounds screened using this method. In this case, the compounds target the M^{PRO} of SARS-CoV-2, which is part of the replication of the virus life cycle. In instances where the inhibitory drug targets the spike protein to prevent attachment and entry, the assay is slightly modified: briefly, the virus is incubated with the compound for 1 hour at room temperature before adding it to the cells. The results shown in Table 6.1 reflect the potency of different compounds at inhibiting viral infection, where compound A has the highest potency with an EC50 in the low micromolar concentration.

Table 6.1 EC$_{50}$ of M^{PRO} inhibitors against SARS-CoV-2

Compound	EC $_{50}$ (µM)
A	0.831
B	2.767
C	3.645
D	5.912
E	8.472
F	9.839
G	8.91
H	8.56
I	7.482

6.4.3 Testing permissiveness of progenitor cells to SARS-CoV-2

Given that VeroE6 cells are a kidney cell line derived from non-human primates, the results of antiviral testing in these cells are not necessarily translated into the humans. For this reason, there is an urgent need of identifying better cell culture models that closely recapitulate the SARS-CoV-2 infection in humans. Using lung progenitor cells is a strategy that we pursued in this set of experiments. In collaboration with the Eyestem Research Private Limited (Bangalore, India), Dr.

Michael Poderycki provided the cells in a 24 well plate format (Figure 6.7). The cells were infected with SARS-CoV-2 at different MOIs= 0.01, 0.1 and 1. The highest viral titer is detected between 48 and 96 hours post infection but does not reach values above 10.000 PFU/mL (Figure 6.8). This suggests that the progenitor cells can be used for experiments with SARS-CoV-2 mimicking the infection in human cells, however not for production of large amounts of virus.
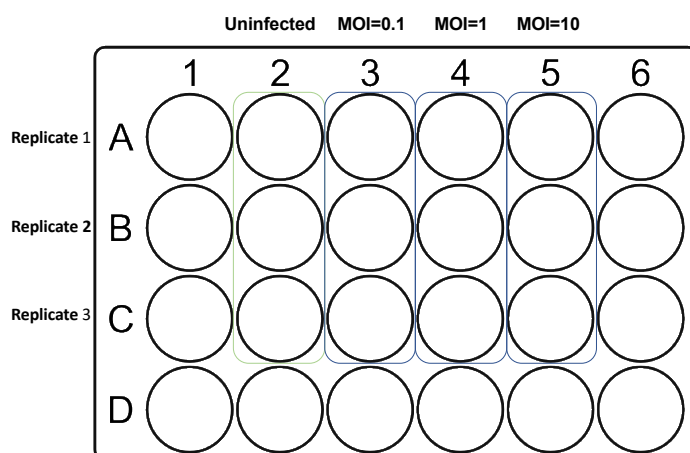
Figure 6.7 Plate set up of iPSC-derived lung progenitor cells

6.5 Discussion

To date, the furin cleavage site of SARS-CoV-2 remains a highly investigated and enigmatic topic. As shown in here, the deletion of the furin site present at S1/S2 prevents the protein from being processed before exiting the cells. The importance of this cleavage has already been addressed by others in animal models. It has been shown that the deletion of furin site in the spike attenuates the infection with SARS-CoV-2 but induces protection against re-challenge with wild type virus [166], as well as reduction of infection in the upper respiratory track of ferrets [167]. Further studies are necessary to determine whether furin is necessary for this cleavage in humans and the implications that it can have in terms of viral transmissibility and pathogenicity outcome.

In addition, the only available antiviral to combat SARS-CoV-2 is Remdesivir, which is a drug originally produced against hepatitis C and that acts by inhibiting the RNA-dependent RNA polymerase. Therefore, other compounds need to be developed to treat acute infection with SARS-CoV-2. The M^{PRO} has been suggested to be a good target for development of antivirals. The assay

development shown in this chapter shows a robust strategy to screen antiviral compounds against SARS-CoV-2 using a luminescence dependent cytotoxicity-based signal.
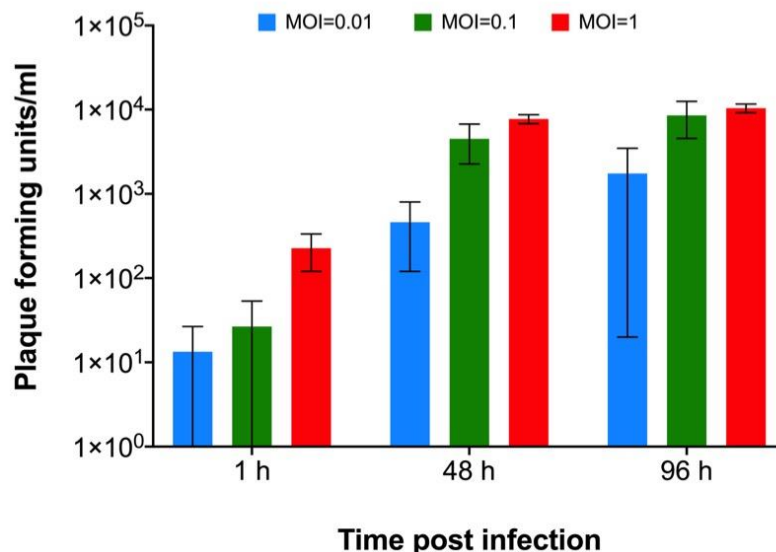
Figure 6.7 SARS-CoV-2 growth in iPSC-derived lung progenitor cells provided by Eyestem. iPSC-derived lung progenitor cells were infected at different MOIs. Supernatant was collected at the specified time points and tittered in Vero E6 cells. N=1, three technical replicates.

Finally, the selection of an ideal cell line to perform assays with SARS-CoV-2 is another field of investigation. The Vero E6 cells are very useful at testing compounds against viral proteins. However, when addressing host factors that the virus relies on, Vero E6 are not useful and other models need to be tested to elucidate more biologically relevant strategies. The progenitor cells tested in this chapter reflect that they can be infected with virus but at lower level than the Vero E6. However, top hit compounds should be tested in progenitor cells in a plaque reduction assay.

CHAPTER 7.   CONCLUSIONS AND FURTHER DIRECTIONS

Mosquitoes are the primary transmission route of multiple arboviruses of clinical relevance such as DENV2, ZIKV and SINV. These viruses rely on host proteases during several steps in their life cycle, especially during the maturation process. Furin, a member of the family of proprotein convertases, is the main host factor that is involved in the maturation of multiple viral families as it performs the proteolytic cleavage of the structural proteins prM in flaviviruses or E3-E2 in alphaviruses. Despite the importance of mosquitoes in the transmission of these pathogens, little is known about the furin homologs that carry out the maturation. Understanding these proteases in the context of mosquito infection and transmission would provide targets for development of novel strategies aiming to control spread of arboviruses.

The objectives of this research were to identify the proprotein convertase homologs in mosquitoes, characterize their enzymatic activity and determine their biological function in the context of flavivirus and alphavirus infection. Using bioinformatic analyses, three proprotein convertases were identified in the genome of *Aedes aegypti:* furin1, furin2 and NC2. They all display the canonical domains of these type of proteases, including the catalytic domain, the prosegment and the P-domain. In addition, they all showed expression among different tissues of mosquitoes as well as in the cell line Aag2.

Attempts to express the proprotein convertases of *Aedes albopictus* were challenging given incomplete gene models as reported on NCBI, and the inacurrate sequence at the 5' end of the gene. In addition, attempts to replace the N-terminal region with BiP or human furin signal peptide did not result in secreted protein. This suggests that the region at the N-terminal is specific to these proteases. However, expression of the proprotein convertases of *Aedes aegypti* did produce recombinant protein that was purified and characterized. A two-step purification process with histidine affinity purification, followed by FLAG-immunoprecipitation, resulted in highly pure samples as detected with silver stain.

The enzymatic activity assays demonstrated that mosquito furin1 and furin2 are active proteases. Furin1 displayed half the activity that was observed with human furin. Furin2 showed

only one third of the activity detected for furin1. Both proteases showed the highest enzymatic activity at neutral pH, but only furin1 increased activity at different temperatures. Similarly, furin1 showed to be dependent on calcium, whereas furin2 did not, which can be explained by the lack of a calcium binding site as was observed in the iTasser predicted structure. Furthermore, the enzymatic constants, Vmax and Km, for furin1 and furin2 were lower than human furin, which suggests that these have higher affinity for the substrate. In contrast, NC2 did not show activity under the experimental conditions performed.

The competitive inhibitor, furin inhibitor-I, showed a lower IC_{50} for furin1 than for human furin. This inhibitor was originally designed against furin and other human proprotein convertases and is widely used in the field of virology when the furin activity is being addressed in the context of viral maturation. This compound, a peptide that produces irreversible inhibition, was able to reduce the titer of DENV2 and ZIKV produced in Aag2 cells. For alphavirus, SINV, the reduction in viral titer was only observed when the initial infection was performed at lower MOIs, suggesting that the inhibition of furin prevents the spread of this virus.

The gene silencing studies demonstrated that knockout of furin1 reduced the maturation of DENV, ZIKV and SINV. In addition, it was also shown that deletion of a single amino acid residue, E311, produced inactive furin1. This amino acid is outside of the active site and binding pockets of furin1. Similar to what was observed with the furin inhibitor-I, the reduction in titer of SINV was dependent on the initial MOI. The explanation for this is that at lower MOIs, the virus that is being tittered corresponds to multiple cycles of replication where the spread is the major contributor of newly produced virus, rather than initial infection in high MOIs. However, spread in furin-deficient cells, either compound-inhibited or CRISPr-knockout, SINV is unable to mature at entry. But, when the virus is being tittered in the BHK cells, it can use furin during entry. For this reason, at high MOIs (=10) and early time points, there is no difference in the titer of virus produced in wild type versus furin-1 mutant Aag2 cells.

Interestingly, the prM content of DENV2 did not change in furin1 mutant Aag2 cells. However, the titer is compromised by several logs of difference. In the case of ZIKV, there was a significant increase in the prM content of the virus produced in mutant furin1 cells. The cleavage

site of DENV2 is known for having a suboptimal cleavage but still retains infectivity[1]. However, this suboptimal maturation might explain the high prM content in wild type cells, but once it reaches suboptimal maturation levels, as happens in the furin1 mutant cells, the virus is non-infectious.

The next steps in this project would be to resolve the crystal structure of the mosquito proprotein convertases. For this, the stably selected S2 cells should be expanded to several liters of suspension culture and the purification would require a different protein tag. A Strep-tag would probably be better for large amounts of protein, rather than FLAG-immunoprecipitation. In addition, it would be helpful to resolve the structure of the furin1 with the deletion of E311. This would provide insights into the importance of the E311 residue in the context of the protein folding, structure and activity. Another route to undertake with this project, would be the generation of CRISPr knockout mosquitoes. A CRISPr approach with homology directed repair, HDR, would be the best strategy to generate the mutant mosquitoes. Similarly, the treatment of mosquitoes with different furin inhibitors would be a best hit strategy for development of disease control strategies.

Combined, the results obtained in this dissertation reveal that mosquitoes have a repertoire of proprotein convertases relevant for arboviral maturation. These proteins can be exploited in the future as targets for development of strategies aiming to control disease transmission. Two main approaches can be used. The first one would be to generate genetically modified mosquitoes that do not have functional furin1 and therefore would hypothetically be deficient at transmitting viruses. This approach would require the engineering of mosquitoes with gene-drive systems that would ensure that this genotype and phenotype is maintained in the wild. The second approach would involve the development and production of antivirals that could be sprayed in the field. This second strategy would require multiple layers of research, including potential effects on other species, the stability of the compound and the length in which such spraying can endure.

REFERENCES

[1]   J. S. Bond, "Proteases: History, discovery, and roles in health and disease," *J. Biol. Chem.*, vol. 294, no. 5, pp. 1643–1651, 2019.

[2]   A. J. Barrett, N. D. Rawlings, and E. A. O'brien, "The MEROPS database as a protease information system," *J. Struct. Biol.*, vol. 134, no. 2–3, pp. 95–102, 2001.

[3]   N. D. Rawlings, A. J. Barrett, and R. Finn, "Twenty years of the MEROPS database of proteolytic enzymes, their substrates and inhibitors," *Nucleic Acids Res.*, vol. 44, no. D1, pp. D343–D350, 2016.

[4]   E. Di Cera, "Serine proteases," *IUBMB Life*, vol. 61, no. 5, pp. 510–515, 2009.

[5]   J. Kraut, "Serine proteases: structure and mechanism of catalysis.," *Annu. Rev. Biochem.*, vol. 46, pp. 331–358, 1977.

[6]   E. S. Radisky, J. M. Lee, C. J. K. Lu, and D. E. Koshland, "Insights into the serine protease mechanism from atomic resolution structures of trypsin reaction intermediates," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 103, no. 18, pp. 6835–6840, 2006.

[7]   L. Hedstrom, "Serine protease mechanism and specificity," *Chem. Rev.*, vol. 102, no. 12, pp. 4501–4523, 2002.

[8]   N. D. Rawlings, A. J. Barrett, and A. Bateman, "MEROPS: The peptidase database," *Nucleic Acids Res.*, vol. 38, no. SUPPL.1, pp. 227–233, 2009.

[9]   N. G. Seidah and A. Prat, "The biology and therapeutic targeting of the proprotein convertases," *Nat. Rev. Drug Discov.*, vol. 11, no. 5, pp. 367–383, 2012.

[10]   M. J. Page and E. Di Cera, "Serine peptidases: Classification, structure and function," *Cell. Mol. Life Sci.*, vol. 65, no. 7–8, pp. 1220–1236, 2008.

[11]   E. Di Cera, "Serine Proteases," *Int. Union Biochem. Mol. Biol. Life*, vol. 61, no. 5, pp. 510–515, 2009.

[12]   U. Shinde and G. Thomas, "Insights from bacterial subtilases into the mechanisms of intramolecular chaperone-mediated activation of furin," *Methods Mol. Biol.*, vol. 768, pp. 59–106, 2011.

[13]   M. Zhong *et al.*, "The prosegments of furin and PC7 as potent inhibitors of proprotein convertases. In vitro and ex vivo assessment of their efficacy and selectivity," *J. Biol. Chem.*, vol. 274, no. 48, pp. 33913–33920, 1999.

[14] N. G. Seidah, M. S. Sadr, M. Chrétien, and M. Mbikay, "The multifaceted proprotein convertases: Their unique, redundant, complementary, and opposite functions," *J. Biol. Chem.*, vol. 288, no. 30, pp. 21473–21481, 2013.

[15] S. Henrich *et al.*, "The crystal structure of the proprotein processing proteinase furin explains its stringent specificity (Nature Structural Biology (2003) 10 (520-526))," *Nat. Struct. Biol.*, vol. 10, no. 8, p. 669, 2003.

[16] O. Lenz, J. Ter Meulen, H. D. Klenk, N. G. Seidah, and W. Garten, "The Lassa virus glycoprotein precursor GP-C is proteolytically processed by subtilase SKI-1/S1P," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 98, no. 22, pp. 12701–12705, 2001.

[17] A. Hyrina, F. Meng, S. J. McArthur, S. Eivemark, I. R. Nabi, and F. Jean, "Human Subtilisin Kexin Isozyme-1 (SKI-1)/Site-1 Protease (S1P) regulates cytoplasmic lipid droplet abundance: A potential target for indirect-acting anti-dengue virus agents," *PLoS One*, vol. 12, no. 3, pp. e0174483–e0174483, Mar. 2017.

[18] J. P. Gorski, N. T. Huffman, C. Cui, E. P. Henderson, R. J. Midura, and N. G. Seidah, "Potential role of proprotein convertase SKI-1 in the mineralization of primary bone.," *Cells. Tissues. Organs*, vol. 189, no. 1–4, pp. 25–32, 2009.

[19] W. P. Sahng, Y. A. Moon, and J. D. Horton, "Post-transcriptional regulation of low density lipoprotein receptor protein by proprotein convertase subtilisin/kexin type 9a in mouse liver," *J. Biol. Chem.*, vol. 279, no. 48, pp. 50630–50638, 2004.

[20] N. A. Taylor, W. J. M. Van De Ven, and J. W. M. Creemers, "Curbing activation: proprotein convertases in homeostasis and pathology," *FASEB J.*, vol. 17, no. 10, pp. 1215–1227, 2003.

[21] M. Uehara, Y. Yaoi, M. Suzuki, K. Takata, and S. Tanaka, "Differential localization of prohormone convertases PC1 and PC2 in two distinct types of secretory granules in rat pituitary gonadotrophs.," *Cell Tissue Res.*, vol. 304, no. 1, pp. 43–49, Apr. 2001.

[22] V. Thimon, M. Belghazi, J. L. Dacheux, and J. L. Gatti, "Analysis of furin ectodomain shedding in epididymal fluid of mammals: Demonstration that shedding of furin occurs in vivo," *Reproduction*, vol. 132, no. 6, pp. 899–908, 2006.

[23] E. Braun and D. Sauter, "furin-mediated protein processing in infectious diseases and cancer," *Clin. Transl. Immunol.*, vol. 8, no. 8, pp. 1–19, 2019.

[24] C. Gyamera-Acheampong and M. Mbikay, "Proprotein convertase subtilisin/kexin type 4 in mammalian fertility: a review," *Hum. Reprod. Update*, vol. 15, no. 2, pp. 237–247, Mar. 2009.

[25] W. Garten, "Characterization of Proprotein Convertases and Their Involvement in Virus Propagation," *Act. Viruses by Host Proteases*, pp. 205–248, Feb. 2018.

[26] A. S. Peterson, L. G. Fong, and S. G. Young, "PCSK9 function and physiology," *J. Lipid Res.*, vol. 49, no. 6, pp. 1152–1156, Jun. 2008.

[27] G. Thomas, "furin at the cutting edge: From protein traffic to embryogenesis and disease," *Nat. Rev. Mol. Cell Biol.*, vol. 3, no. 10, pp. 753–766, 2002.

[28] S. Tian, Q. Huang, Y. Fang, and J. Wu, "furinDB: A database of 20-residue furin cleavage site motifs, substrates and their associated drugs," *Int. J. Mol. Sci.*, vol. 12, no. 2, pp. 1060–1065, Feb. 2011.

[29] S. Takahashi, T. Nakagawa, T. Banno, T. Watanabe, K. Murakami, and K. Nakayama, "Localization of furin to the trans-Golgi network and recycling from the cell surface involves Ser and Tyr residues within the cytoplasmic domain," *J. Biol. Chem.*, vol. 270, no. 47, pp. 28397–28401, 1995.

[30] J. Shapiro, N. Sciaky, J. Lee, H. Bosshart, R. H. Angeletti, and J. S. Bonifacino, "Localization of endogenous furin in cultured cell lines.," *J. Histochem. Cytochem. Off. J. Histochem. Soc.*, vol. 45, no. 1, pp. 3–12, Jan. 1997.

[31] M. Teuchert, S. Berghöfer, H. D. Klenk, and W. Garten, "Recycling of furin from the plasma membrane. Functional importance of the cytoplasmic tail sorting signals and interaction with the AP-2 adaptor medium chain subunit," *J. Biol. Chem.*, vol. 274, no. 51, pp. 36781–36789, 1999.

[32] A. J. Roebroek *et al.*, "Failure of ventral closure and axial rotation in embryos lacking the proprotein convertase furin.," *Development*, vol. 125, no. 24, pp. 4863–4876, Dec. 1998.

[33] Y. Chen *et al.*, "Mutations within a furin consensus sequence block proteolytic release of ectodysplasin-A and cause X-linked hypohidrotic ectodermal dysplasia," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 98, no. 13, pp. 7218–7223, Jun. 2001.

[34] G. Zhao *et al.*, "Influence of a coronary artery disease-associated genetic variant on furin expression and effect of furin on macrophage behavior," *Arterioscler. Thromb. Vasc. Biol.*, vol. 38, no. 8, pp. 1837–1844, 2018.

[35] T. A. Y. Ayoubi, J. W. M. Creemers, A. J. M. Roebroek, and W. J. M. Van De Ven, "Expression of the dibasic proprotein processing enzyme furin is directed by multiple promoters," *J. Biol. Chem.*, vol. 269, no. 12, pp. 9298–9303, 1994.

[36] F. Blanchette, R. Day, W. Dong, M. H. Laprise, and C. M. Dubois, "TGFbeta1 regulates gene expression of its own converting enzyme furin.," *J. Clin. Invest.*, vol. 99, no. 8, pp. 1974–1983, Apr. 1997.

[37] R.-N. Chen *et al.*, "Thyroid Hormone Promotes Cell Invasion through Activation of furin Expression in Human Hepatoma Cell Lines," *Endocrinology*, vol. 149, no. 8, pp. 3817–3831, Aug. 2008.

[38] M. Pesu, L. Muul, Y. Kanno, and J. J. O'Shea, "Proprotein convertase furin is preferentially expressed in T helper 1 cells and regulates interferon gamma," *Blood*, vol. 108, no. 3, pp. 983–985, Aug. 2006.

[39] Z. Ortutay, A. Grönholm, M. Laitinen, M. Keresztes-Andrei, I. Hermelo, and M. Pesu, "Identification of Novel Genetic Regulatory Region for Proprotein Convertase FURIN and Interferon Gamma in T Cells," *Front. Immunol.*, vol. 12, p. 630389, Feb. 2021.

[40] E. D. Anderson, S. S. Molloy, F. Jean, H. Fei, S. Shimamura, and G. Thomas, "The ordered and compartment-specific autoproteolytic removal of the furin intramolecular chaperone is required for enzyme activation," *J. Biol. Chem.*, vol. 277, no. 15, pp. 12879–12890, 2002.

[41] M. Vey, W. Schäfer, S. Berghöfer, H. D. Klenk, and W. Garten, "Maturation of the trans-Golgi network protease furin: compartmentalization of propeptide removal, substrate cleavage, and COOH-terminal truncation.," *J. Cell Biol.*, vol. 127, no. 6 Pt 2, pp. 1829–1842, Dec. 1994.

[42] P. Jaaks and M. Bernasconi, "The proprotein convertase furin in tumour progression," *Int. J. Cancer*, vol. 141, no. 4, pp. 654–663, Aug. 2017.

[43] S. O. Dahms *et al.*, "The structure of a furin-antibody complex explains non-competitive inhibition by steric exclusion of substrate conformers," *Sci. Rep.*, vol. 6, no. September, pp. 1–7, 2016.

[44] A. Stieneke-Grober *et al.*, "Influenza virus hemagglutinin with multibasic cleavage site is activated by furin, a subtilisin-like endoprotease," *EMBO J.*, vol. 11, no. 7, pp. 2407–2414, 1992.

[45] H. H. Hoffmann *et al.*, "Diverse Viruses Require the Calcium Transporter SPCA1 for Maturation and Spread," *Cell Host Microbe*, vol. 22, no. 4, pp. 460-470.e5, 2017.

[46] T. Gary, "furin at the cutting edge: from protein traffic to embryogenesis and disease," *Mol. Cell*, vol. 3, no. 10, pp. 753–766, 2007.

[47] G. Izaguirre, "The Proteolytic Regulation of Virus Cell Entry by furin and Other Proprotein Convertases," *Viruses*, vol. 11, no. 9, 2019.

[48] R. J. Wool-Lewis and P. Bates, "Endoproteolytic Processing of the Ebola Virus Envelope Glycoprotein: Cleavage Is Not Required for Function," *J. Virol.*, vol. 73, no. 2, pp. 1419–1426, 1999.

[49] S. Matsuyama, N. Nagata, K. Shirato, M. Kawase, M. Takeda, and F. Taguchi, "Efficient Activation of the Severe Acute Respiratory Syndrome Coronavirus Spike Protein by the Transmembrane Protease TMPRSS2," *J. Virol.*, vol. 84, no. 24, pp. 12658–12664, 2010.

[50] K. Stadler, S. L. Allison, J. Schalich, and F. X. Heinz, "Proteolytic Activation of Tick-Borne Encephalitis Virus by furin," *J. Virol.*, vol. 71, no. 11, pp. 8475–8481, 1997.

[51] P. Decha *et al.*, "Source of high pathogenicity of an avian influenza virus H5N1: Why H5 is better cleaved by furin," *Biophys. J.*, vol. 95, no. 1, pp. 128–134, 2008.

[52] A. R. Fehr and S. Perlman, "Coronaviruses: An Overview of Their Replication and Pathogenesis," *Methods Mol Biol.*, pp. 1–23, 2016.

[53] G. Lu, Q. Wang, and G. F. Gao, "Bat-to-human : spike features determining ' host jump ' of MERS-CoV , and beyond," *Trends Microbiol.*, vol. 23, no. 8, pp. 468–478, 2015.

[54] J. K. Millet and G. R. Whittaker, "Host cell entry of Middle East respiratory syndrome coronavirus after two-step , furin-mediated activation of the spike protein," *Proc. Natl. Acad. Sci.*, pp. 1–6, 2014.

[55] J. Park *et al.*, "Proteolytic processing of Middle East respiratory syndrome coronavirus spikes expands virus tropism," *Proc. Natl. Acad. Sci.*, 2016.

[56] S. Matsuyama, K. Shirato, M. Kawase, Y. Terada, K. Kawachi, and S. Fukushi, "Middle East Respiratory Syndrome Coronavirus Spike Protein Is Not Activated Directly by Cellular furin during Viral Entry into Target Cells," *J. Virol.*, vol. 92, no. 19, pp. 1–12, 2018.

[57] S. Mukhopadhyay, R. J. Kuhn, and M. G. Rossmann, "A structural perspective of the Flavivirus life cycle," *Nat. Rev. Microbiol.*, vol. 3, no. 1, pp. 13–22, 2005.

[58] S. Apte-Sengupta, D. Sirohi, and R. J. Kuhn, "Coupling of replication and assembly in flaviviruses," *Curr. Opin. Virol.*, vol. 9, pp. 134–142, 2014.

[59] D. Sirohi and R. J. Kuhn, "Zika Virus Structure, Maturation, and Receptors," *J. Infect. Dis.*, vol. 216, no. Suppl 10, pp. S935–S944, 2017.

[60] C. M. R. Nicholls, M. Sevvana, and R. J. Kuhn, "Structure-guided paradigm shifts in Flavivirus assembly and maturation mechanisms," *Adv. Virus Res.*, vol. 108, no. January, pp. 33–83, 2020.

[61] P. Plevka *et al.*, "Maturation of flaviviruses starts from one or more icosahedrally independent nucleation centres," *EMBO Rep.*, vol. 12, no. 6, pp. 602–606, 2011.

[62] I. Yu *et al.*, "Structure of the immature dengue virus at low pH primes proteolytic Maturation," *Science (80-.).*, vol. 319, no. March, pp. 1834–1838, 2008.

[63] I.-M. Yu, H. A. Holdaway, P. R. Chipman, R. J. Kuhn, M. G. Rossmann, and J. Chen, "Association of the pr peptides with dengue virus at acidic pH blocks membrane fusion," *J. Virol.*, vol. 83, no. 23, pp. 12101–12107, 2009.

[64] J. Jose, J. Tang, A. B. Taylor, T. S. Baker, and R. J. Kuhn, "Fluorescent protein-tagged Sindbis virus E2 glycoprotein allows single particle analysis of virus budding from live cells," *Viruses*, vol. 7, no. 12, pp. 6182–6199, 2015.

[65] J. Jose, J. E. Snyder, and R. J. Kuhn, "A structural and functional perspective of alphavirus replication and assembly," *Future Microbiol.*, vol. 4, no. 7, pp. 837–856, Sep. 2009.

[66] J. Y.-S. Leung, M. M.-L. Ng, and J. J. H. Chu, "Replication of alphaviruses: a review on the entry process of alphaviruses into cells," *Adv. Virol.*, vol. 2011, p. 249640, 2011.

[67] S. J. Ryan, C. J. Carlson, E. A. Mordecai, and L. R. Johnson, "Global expansion and redistribution of Aedes-borne virus transmission risk with climate change," *PLoS Negl. Trop. Dis.*, vol. 13, no. 3, p. e0007213, Mar. 2019.

[68] C. Rückert and G. D. Ebel, "How Do Virus-Mosquito Interactions Lead to Viral Emergence?," *Trends Parasitol.*, pp. 1–12, 2018.

[69] S. Bhatt *et al.*, "The effect of malaria control on Plasmodium falciparum in Africa between 2000 and 2015," *Nature*, vol. 526, no. 7572, pp. 207–211, 2015.

[70] N. Harriott, "Mosquito Control and Pollinator Health," *Pestic. you*, vol. 36, no. 2, pp. 9–17, 2016.

[71] L. Alphey, "Genetic Control of Mosquitoes," *Annu. Rev. Entomol.*, vol. 59, no. 1, pp. 205–224, Jan. 2014.

[72] J. Fang, "Ecology: A world without mosquitoes," *Nature*, vol. 466, no. 7305, pp. 432–434, 2010.

[73] S. Aittomäki *et al.*, "Proprotein convertase furin1 expression in the Drosophila fat body is essential for a normal antimicrobial peptide response and bacterial host defense," *FASEB J.*, vol. 31, no. 11, pp. 4770–4782, 2017.

[74] A. Sohr, L. Du, R. Wang, L. Lin, and S. Roy, "Drosophila FGF cleavage is required for efficient intracellular sorting and intercellular dispersal," *J. Cell Biol.*, vol. 218, no. 5, pp. 1653–1669, 2019.

[75] G. L. Cano-Monreal, J. C. Williams, and H. W. Heidner, "An Arthropod Enzyme, Dfurin 1, and a Vertebrate furin Homolog Display Distinct Cleavage Site Sequence Preferences for a Shared Viral Proprotein Substrate," *J. Insect Sci.*, vol. 10, no. 29, pp. 1–16, 2010.

[76] J. Künnapuu, I. Björkgren, and O. Shimmi, "The Drosophila DPP signal is produced by cleavage of its proprotein at evolutionary diversified furin-recognition sites," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 106, no. 21, pp. 8501–8506, 2009.

[77] D. E. Siekhaus and R. S. Fuller, "A role for amontillado, the Drosophila homolog of the neuropeptide precursor processing protease PC2, in triggering hatching behavior.," *J. Neurosci.*, vol. 19, no. 16, pp. 6942–54, 1999.

[78] J. S. Chen and A. S. Raikhel, "Subunit cleavage of mosquito pro-vitellogenin by a subtilisin-like convertase," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 93, no. 12, pp. 6186–6190, 1996.

[79] N. G. Seidah and A. Prat, "The biology and therapeutic targeting of the proprotein convertases," *Nat. Rev. Drug Discov.*, vol. 11, no. 5, pp. 367–383, 2012.

[80] B. D. Bennett *et al.*, "A furin-like convertase mediates propeptide cleavage of BACE, the Alzheimer's β-secretase," *J. Biol. Chem.*, vol. 275, no. 48, pp. 37712–37717, 2000.

[81] S. Bhatt *et al.*, "The global distribution and burden of dengue," *Nature*, vol. 496, no. 7446, pp. 504–507, 2013.

[82] M. A. Tolle *et al.*, "Mosquito-borne diseases.," *Curr. Probl. Pediatr. Adolesc. Health Care*, vol. 39, no. 4, pp. 97–140, Apr. 2009.

[83] A. J. M. Roebroek, I. G. L. Pauli, Y. Zhang, and W. J. M. van de Ven, "cDNA sequence of a Drosophila melanogaster gene, Dfur1, encoding a protein structurally related to the subtilisin-like proprotein processing enzyme furin," *FEBS Lett.*, vol. 289, no. 2, pp. 133–137, 1991.

[84] A. J. Roebroek, T. A. Ayoubi, J. W. Creemers, I. G. Pauli, and W. J. Van de Ven, "The Dfur2 gene of Drosophila melanogaster: genetic organization, expression during embryogenesis, and pro-protein processing activity of its translational product Dfurin2.," *DNA Cell Biol.*, vol. 14, no. 3, pp. 223–34, 1995.

[85] A. J. M. Roebroek *et al.*, "Cloning and functional expression of Dfurin2, a subtilisin-like proprotein processing enzyme of Drosophila melanogaster with multiple repeats of a cysteine motif," *J. Biol. Chem.*, vol. 267, no. 24, pp. 17208–17215, 1992.

[86] J. Penney *et al.*, "LRRK2 regulates retrograde synaptic compensation at the Drosophila neuromuscular junction," *Nat. Commun.*, vol. 7, 2016.

[87] L. Zhang, Z. Ali Syed, I. Van Dijk H'rd, J. M. Lim, L. Wells, and K. G. Ten Hagen, "O-Glycosylation regulates polarized secretion by modulating Tango1 stability," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 111, no. 20, pp. 7296–7301, 2014.

[88] T. K. Johnson, M. A. Henstridge, A. Herr, K. A. Moore, J. C. Whisstock, and C. G. Warr, "Torso-like mediates extracellular accumulation of furin-cleaved Trunk to pattern the Drosophila embryo termini," *Nat. Commun.*, vol. 6, pp. 1–6, 2015.

[89] E. Ordan and T. Volk, "Amontillado is required for Drosophila Slit processing and for tendon-mediated muscle patterning," *Biol. Open*, vol. 5, no. 10, pp. 1530–1534, 2016.

[90] C. Wegener, H. Herbert, J. Kahnt, M. Bender, and J. M. Rhea, "Deficiency of prohormone convertase dPC2 (AMONTILLADO) results in impaired production of bioactive neuropeptide hormones in Drosophila," *J. Neurochem.*, vol. 118, no. 4, pp. 581–595, 2011.

[91] M. M. W. Smolenaars, M. A. M. Kasperaitis, P. E. Richardson, K. W. Rodenburg, and D. J. Van Der Horst, "Biosynthesis and secretion of insect lipoprotein: Involvement of furin in cleavage of the apoB homolog, apolipophorin-II/I," *J. Lipid Res.*, vol. 46, no. 3, pp. 412–421, 2005.

[92] N. Nour, A. Basak, M. Chrétien, and N. G. Seidah, "Structure-function analysis of the prosegment of the proprotein convertase PC5A," *J. Biol. Chem.*, vol. 278, no. 5, pp. 2886–2895, 2003.

[93] U. Shinde and M. Inouye, "Intramolecular chaperones and protein folding," *Trends Biochem. Sci.*, vol. 18, no. 11, pp. 442–446, 1993.

[94] A. Zhou, S. Martin, G. Lipkind, J. LaMendola, and D. F. Steiner, "Regulatory roles of the P domain of the subtilisin-like prohormone convertases," *J. Biol. Chem.*, vol. 273, no. 18, pp. 11107–11114, 1998.

[95] W. Liu *et al.*, "IBS: An illustrator for the presentation and visualization of biological sequences," *Bioinformatics*, vol. 31, no. 20, pp. 3359–3361, 2015.

[96] S. El-Gebali *et al.*, "The Pfam protein families database in 2019," *Nucleic Acids Res.*, vol. 47, no. D1, pp. D427–D432, 2019.

[97] C. J. A. Sigrist *et al.*, "New and continuing developments at PROSITE," *Nucleic Acids Res.*, vol. 41, no. D1, pp. 344–347, 2013.

[98] M. Goujon *et al.*, "A new bioinformatics analysis tools framework at EMBL-EBI," *Nucleic Acids Res.*, vol. 38, no. SUPPL. 2, pp. 695–699, 2010.

[99] J. J. Almagro Armenteros *et al.*, "SignalP 5.0 improves signal peptide predictions using deep neural networks," *Nat. Biotechnol.*, vol. 37, no. 4, pp. 420–423, 2019.

[100] T. Hirokawa, S. Boon-Chieng, and S. Mitaku, "SOSUI: classification and secondary structure prediction system for membrane proteins," *Bioinformatics*, vol. 14, no. 4, pp. 378–379, 1998.

[101] S. Möller, M. D. R. Croning, and R. Apweiler, "Evaluation of methods for the prediction of membrane spanning regions," *Bioinformatics*, vol. 17, no. 7, pp. 646–653, 2001.

[102] A. Dereeper *et al.*, "Phylogeny.fr: robust phylogenetic analysis for the non-specialist.," *Nucleic Acids Res.*, vol. 36, no. Web Server issue, pp. 465–469, 2008.

[103] I. Letunic and P. Bork, "Interactive Tree Of Life (iTOL): An online tool for phylogenetic tree display and annotation," *Bioinformatics*, vol. 23, no. 1, pp. 127–128, 2007.

[104] M. Schmid, E. Kauffman, A. Payne, E. Harris, and L. Kramer, "Preparation of Mosquito Salivary Gland Extract and Intradermal Inoculation of Mice," *Bio-Protocol*, vol. 7, no. 14, p. 8325, 2017.

[105] B. Yao, L. Zhang, S. Liang, and C. Zhang, "SVMTriP: A Method to Predict Antigenic Epitopes Using Support Vector Machine to Integrate Tri-Peptide Similarity and Propensity," *PLoS One*, vol. 7, no. 9, pp. 5–9, 2012.

[106] J. Yang, R. Yan, A. Roy, D. Xu, J. Poisson, and Y. Zhang, "The I-TASSER suite: Protein structure and function prediction," *Nat. Methods*, vol. 12, no. 1, pp. 7–8, 2014.

[107] J. B. Denault, L. Bissonnette, J. M. Longpré, G. Charest, P. Lavigne, and R. Leduc, "Ectodomain shedding of furin: Kinetics and role of the cysteine-rich region," *FEBS Lett.*, vol. 527, no. 1–3, pp. 309–314, 2002.

[108] W. Y. Tao, Y. C. Cheng, M. H. Song, D. A. Weisblat, and D. H. Kuo, "Diversification of metazoan Kexin-like proprotein convertases: Insights from the leech Helobdella," *bioRxiv*, pp. 1–46, 2019.

[109] M. N. Baeshen *et al.*, "Production of biopharmaceuticals in E. Coli: Current scenario and future perspectives," *J. Microbiol. Biotechnol.*, vol. 25, no. 7, pp. 953–962, 2015.

[110] N. Jenkins, "Modifications of therapeutic proteins: Challenges and prospects," *Cytotechnology*, vol. 53, no. 1–3, pp. 121–125, 2007.

[111] G. Walsh and R. Jefferis, "Post-translational modifications in the context of therapeutic proteins," *Nat. Biotechnol.*, vol. 24, no. 10, pp. 1241–1252, 2006.

[112] Y. Endo and T. Sawasaki, "Cell-free expression systems for eukaryotic protein production," *Curr. Opin. Biotechnol.*, vol. 17, no. 4, pp. 373–380, 2006.

[113] S. Geisse, H. Gram, B. Kleuser, and H. P. Kocher, "Eukaryotic expression systems: A comparison," *Protein Expr. Purif.*, vol. 8, no. 3, pp. 271–282, 1996.

[114] T. J. Griffin, G. Seth, H. Xie, S. Bandhakavi, and W. S. Hu, "Advancing mammalian cell culture engineering using genome-scale technologies," *Trends Biotechnol.*, vol. 25, no. 9, pp. 401–408, 2007.

[115] T. K. Kim and J. H. Eberwine, "Mammalian cell transfection: The present and the future," *Anal. Bioanal. Chem.*, vol. 397, no. 8, pp. 3173–3178, 2010.

[116] F. L. Graham, J. Smiley, W. C. Russell, and R. Nairn, "Characteristics of a human cell line transformed by DNA from human adenovirus type 5," *J. Gen. Virol.*, vol. 36, no. 1, pp. 59–72, 1977.

[117] Y. C. Lin *et al.*, "Genome dynamics of the human embryonic kidney 293 lineage in response to cell biology manipulations," *Nat. Commun.*, vol. 5, no. 11, 2014.

[118] R. B. DuBridge, P. Tang, H. C. Hsia, P. M. Leong, J. H. Miller, and M. P. Calos, "Analysis of mutation in human cells by using an Epstein-Barr virus shuttle system.," *Mol. Cell. Biol.*, vol. 7, no. 1, pp. 379–387, 1987.

[119] I. Berger, D. J. Fitzgerald, and T. J. Richmond, "Baculovirus expression system for heterologous multiprotein complexes," *Nat. Biotechnol.*, vol. 22, no. 12, pp. 1583–1587, 2004.

[120] D. J. Fitzgerald, P. Berger, C. Schaffitzel, K. Yamada, T. J. Richmond, and I. Berger, "Protein complex expression by using multigene baculoviral vectors," *Nat. Methods*, vol. 3, no. 12, pp. 1021–1032, 2006.

[121] C. Bieniossek, T. Imasaki, Y. Takagi, and I. Berger, "MultiBac: Expanding the research toolbox for multiprotein complexes," *Trends Biochem. Sci.*, vol. 37, no. 2, pp. 49–57, 2012.

[122] M. Rämet *et al.*, "Drosophila scavenger receptor CI is a pattern recognition receptor for bacteria," *Immunity*, vol. 15, no. 6, pp. 1027–1038, 2001.

[123] M. Rämet, P. Manfruelli, A. Pearson, B. Mathey-Prevot, and R. A. B. Ezekowitz, "Functional genomic analysis of phagocytosis and identification of a Drosophila receptor for E. Coli," *Nature*, vol. 416, no. 6881, pp. 644–648, 2002.

[124] J. Zitzmann, G. Sprick, T. Weidner, C. Schreiber, and P. Czermak, "Process Optimization for Recombinant Protein Expression in Insect Cells," in *New Insights into Cell Culture Technology*, 2017.

[125] T. Iwaki, M. Figuera, V. A. Ploplis, and F. J. Castellino, "Rapid selection of Drosophila S2 cells with the puromycin resistance gene," *Biotechniques*, vol. 35, no. 3, pp. 482–486, 2003.

[126] M. G. Santos, S. A. C. Jorge, K. Brillet, and C. A. Pereira, "Improving heterologous protein expression in transfected Drosophila S2 cells as assessed by EGFP expression," *Cytotechnology*, vol. 54, no. 1, pp. 15–24, 2007.

[127] M. K. Schwinn *et al.*, "CRISPR-Mediated Tagging of Endogenous Proteins with a Luminescent Peptide," *ACS Chem. Biol.*, vol. 13, no. 2, pp. 467–474, 2018.

[128] "NEBaseChanger." [Online]. Available: https://nebasechanger.neb.com. [Accessed: 07-Aug-2021].

[129] J. R. Miller *et al.*, "Analysis of the Aedes albopictus C6/36 genome provides insight into cell line utility for viral propagation," *Gigascience*, vol. 7, no. 3, pp. 1–13, 2018.

[130] J. Coleman, M. Inukai, and M. Inouye, "Dual functions of the signal peptide in protein transfer across the membrane," *Cell*, vol. 43, no. 1, pp. 351–360, 1985.

[131] H. Owji, N. Nezafat, M. Negahdaripour, A. Hajiebrahimi, and Y. Ghasemi, "A comprehensive review of signal peptides: Structure, roles, and applications," *Eur. J. Cell Biol.*, vol. 97, no. 6, pp. 422–441, 2018.

[132] B. Martoglio and B. Dobberstein, "Signal sequences: More than just greasy peptides," *Trends Cell Biol.*, vol. 8, no. 10, pp. 410–415, 1998.

[133] L. Muller, X. Zhu, and I. Lindberg, "Mechanism of the facilitation of PC2 maturation by 7B2: Involvement in ProPC2 transport and activation but not folding," *J. Cell Biol.*, vol. 139, no. 3, pp. 625–638, 1997.

[134] J. R. Hwang, D. E. Siekhaus, R. S. Fuller, P. H. Taghert, and I. Lindberg, "Interaction of Drosophila melanogaster prohormone convertase 2 and 7B2. Insect cell-specific processing and secretion," *J. Biol. Chem.*, vol. 275, no. 23, pp. 17886–17893, 2000.

[135] M. R. Dyson, "Selection of soluble protein expression constructs: The experimental determination of protein domain boundaries," *Biochem. Soc. Trans.*, vol. 38, no. 4, pp. 908–912, 2010.

[136] T. Iwaki and F. J. Castellino, "A single plasmid transfection that offers a significant advantage associated with puromycin selection in Drosophila Schneider S2 cells expressing heterologous proteins," *Cytotechnology*, vol. 57, no. 1, pp. 45–49, 2008.

[137] K. Lin and P. Gallay, "Curing a viral infection by targeting the host: The example of cyclophilin inhibitors," *Antiviral Res.*, vol. 99, no. 2013, pp. 86–77, 2020.

[138] R. Kumar, D. Mehta, N. Mishra, D. Nayak, and S. Sunil, "Role of host-mediated post-translational modifications (PTMS) in RNA virus pathogenesis," *Int. J. Mol. Sci.*, vol. 22, no. 1, pp. 1–26, 2021.

[139] J. Berg, J. Tymoczko, and L. Stryer, "Many Enzymes Are Activated by Specific Proteolytic Cleavage," in *Biochemistry*, New York: W H Freeman, 2002.

[140] E. Böttcher-Friebertshäuser, H. D. Klenk, and W. Garten, "Activation of influenza viruses by proteases from host cells and bacteria in the human airway epithelium," *Pathog. Dis.*, vol. 69, no. 2, pp. 87–100, 2013.

[141] S. S. Molloy, P. A. Bresnahan, S. H. Leppla, K. R. Klimpel, and G. Thomas, "Human furin is a calcium-dependent serine endoprotease that recognizes the sequence Arg-X-X-Arg and efficiently cleaves anthrax toxin protective antigen," *J. Biol. Chem.*, vol. 267, no. 23, pp. 16396–16402, 1992.

[142] S. O. Dahms, M. Arciniega, T. Steinmetzer, R. Huber, and M. E. Than, "Structure of the unliganded form of the proprotein convertase furin suggests activation by a substrate-induced mechanism," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 113, no. 40, pp. 11196–11201, 2016.

[143] C. Wu *et al.*, "furin: A Potential Therapeutic Target for COVID-19," *iScience*, vol. 23, no. 10, p. 101642, 2020.

[144] E. Bergeron *et al.*, "Implication of proprotein convertases in the processing and spread of severe acute respiratory syndrome coronavirus," *Biochem. Biophys. Res. Commun.*, vol. 326, no. 3, pp. 554–563, 2005.

[145] M. Imran *et al.*, "Decanoyl-Arg-Val-Lys-Arg-chloromethylketone: An antiviral compound that acts against flaviviruses through the inhibition of furin-mediated prM cleavage," *Viruses*, vol. 11, no. 11, 2019.

[146] S. Rajapakse, "Dengue shock," *J. Emergencies, Trauma Shock*, vol. 4, no. 1, pp. 120–127, 2011.

[147] M. K. Kindhauser, T. Allen, V. Frank, R. S. Santhana, and C. Dye, "Zika Virus: the origin and spread of a mosquito-borne virus.," *Bull. World Health Organ.*, vol. 94, no. 9, pp. 675-686C, Sep. 2016.

[148] N. L. Achee *et al.*, "Alternative strategies for mosquito-borne arbovirus control," *PLoS Negl. Trop. Dis.*, vol. 13, no. 1, p. e0006822, Jan. 2019.

[149] H. M. Van Der Schaar *et al.*, "Dissecting the cell entry pathway of dengue virus by single-particle tracking in living cells," *PLoS Pathog.*, vol. 4, no. 12, 2008.

[150] J. S. Hayflick, W. J. Wolfgang, M. A. Forte, and G. Thomas, "A unique Kex2-like endoprotease from Drosophila melanogaster is expressed in the central nervous system during early embryogenesis," *J. Neurosci.*, vol. 12, no. 3, pp. 705–717, 1992.

[151] L. Y. M. Rayburn, J. Rhea, S. R. Jocoy, and M. Bender, "The proprotein convertase amontillado (amon) is required during Drosophila pupal development," *Developmental Biology*, vol. 333, no. 1. pp. 48–56, 2009.

[152] T. Walker, C. L. Jeffries, K. L. Mansfield, and N. Johnson, "Mosquito cell lines: History, isolation, availability and application to assess the threat of arboviral transmission in the United Kingdom," *Parasites and Vectors*, vol. 7, no. 1, pp. 1–9, 2014.

[153] D. E. Brackney *et al.*, "C6 / 36 Aedes albopictus Cells Have a Dysfunctional Antiviral RNA Interference Response," vol. 4, no. 10, pp. 24–27, 2010.

[154] B. A. McKenzie, A. E. Wilson, and S. Zohdy, "Aedes albopictus is a competent vector of Zika virus: A meta-analysis," *PLoS One*, vol. 14, no. 5, pp. 1–16, 2019.

[155] B. J. Matthews *et al.*, "Improved reference genome of Aedes aegypti informs arbovirus vector control," *Nature*, vol. 563, no. 7732, pp. 501–507, 2018.

[156] S. Takahashi *et al.*, "A mutation of furin causes the lack of precursor-processing activity in human colon carcinoma LoVo cells," *Biochemical and Biophysical Research Communications*, vol. 195, no. 2. pp. 1019–1026, 1993.

[157] S. Takahashi *et al.*, "A second mutant allele of furin in the processing-incompetent cell line, LoVo. Evidence for involvement of the homo B domain in autocatalytic activation," *J. Biol. Chem.*, vol. 270, no. 44, pp. 26565–26569, 1995.

[158] J. Shang *et al.*, "Cell entry mechanisms of SARS-CoV-2," *Proc. Natl. Acad. Sci.*, vol. 117, no. 21, 2020.

[159] A. C. Walls *et al.*, "Structure , Function , and Antigenicity of the SARS- Structure , Function , and Antigenicity of the SARS-CoV-2 Spike Glycoprotein," *Cell*, pp. 1–12, 2020.

[160] M. Hoffmann, H. Kleine-weber, and S. Pöhlmann, "A Multibasic Cleavage Site in the Spike Protein of SARS-CoV-2 Is Essential for Infection of Human ll ll Short Article A Multibasic Cleavage Site in the Spike Protein of SARS-CoV-2 Is Essential for Infection of Human Lung Cells," *Mol. Cell*, pp. 779–784, 2020.

[161] B. Coutard, C. Valle, X. De Lamballerie, B. Canard, N. G. Seidah, and E. Decroly, "The spike glycoprotein of the new coronavirus 2019-nCoV contains a furin- like cleavage site absent in CoV of the same clade," *Antiviral Res.*, vol. 176, no. February, p. 104742, 2020.

[162] B. J. Bosch, W. Bartelink, and P. J. M. Rottier, "Cathepsin L Functionally Cleaves the Severe Acute Respiratory Syndrome Coronavirus Class I Fusion Protein Upstream of Rather than Adjacent to the Fusion Peptide," *J. Virol.*, vol. 82, no. 17, pp. 8887–8890, 2008.

[163] Y. Yang *et al.*, "Two Mutations Were Critical for Bat-to-Human Transmission of Middle East Respiratory Syndrome," *J. Virol.*, vol. 89, no. 17, pp. 9119–9123, 2015.

[164] H. Kleine-Weber, M. T. Elzayat, M. Hoffmann, and S. Pöhlmann, "Functional analysis of potential cleavage sites in the MERS- coronavirus spike protein," *Nat. Sci. Reports*, no. October, pp. 1–11, 2018.

[165] X. Xie *et al.*, "An Infectious cDNA Clone of SARS-CoV-2 ll," *Cell Host Microbe*, vol. 27, no. 5, pp. 841-848.e3, 2020.

[166] B. A. Johnson *et al.*, "Loss of furin cleavage site attenuates SARS-CoV-2 pathogenesis," *Nature*, vol. 591, no. 7849, pp. 293–299, 2021.

[167] T. P. Peacock *et al.*, "The furin cleavage site in the SARS-CoV-2 spike protein is required for transmission in ferrets," *Nat. Microbiol.*, vol. 6, no. 7, pp. 899–909, 2021.

VITA

Carlos Andres Brito Sierra was born in Villavicencio, Colombia in March of 1994. He attended high school in Colegio Departamental la Esperanza in Villavicencio. Then, he obtained a Bachelor of Science degree from Universidad Nacional de Colombia in Bogota. During his tenure as undergraduate student, Carlos participated in different research programs, including cytogenetics, malaria and avian haemoparasites. He worked with Professor Nubia Matta interrogating the vectorial capacity of biting midges and black flies at transmitting different types of avian malaria in the paramo ecosystems of Colombia. Carlos was fortunate to obtain a fellowship for an undergraduate internship program at Purdue university, where he joined the laboratory of Professor Catherine A. Hill. There, Carlos continued his interest in insect-borne diseases and started a research around the development of novel insecticides for mosquito control. Carlos maintained his interest in science and chose to accept a position as Graduate Student at Purdue University within the PULSe program. There, he continued his interest in mosquito-borne diseases and decided to pursue his PhD in the laboratory of Professor Richard J. Kuhn, where he got involved in the world of virology. During the pursuit of his PhD, Carlos interrogated the maturation process of flaviviruses in mosquitoes, and the results of his dissertation provided insights into the host factors involved in this process. Carlos's goal is to apply the knowledge he gained through his career to develop strategies that can control diseases using a multidisciplinary approach of disciplines.

PUBLICATIONS

- **Brito-Sierra, CA**., Sirohi, D., Hill, C.A., Kuhn, RJ. *In preparation.* Mosquito furin like proteases: implications during Flavivirus and alphavirus maturation.

- Nunya Chotiwan1, **Brito-Sierra CA**., Hill CA., Rushika Perera. *In preparation.* Expression of fatty acid synthase genes and their role in development and arboviral infection of Aedes aegypti

- Sirohi, D., **Brito-Sierra, CA**., Kuhn, RJ. *In preparation*. Role of host proprotein convertase in virion morphogenesis, infectivity and pathogenesis. *MMBR*.

- **Brito-Sierra, CA**., Kaur, J., Hill, C. A. Protocols for Testing the Toxicity of Novel Insecticidal Chemistries to Mosquitoes. J. Vis. Exp. (144), e57768, doi:10.3791/57768 (2019)

- Matthews BJ*, O Dudchenko*, S Kingan*, S Koren, I Antoshechkin, JE Crawford, WJ Glassford, M Herre, SN Redmond, NH Rose, GD Weedall, Y Wu, SS Batra, **CA. Brito-Sierra**, *et al*. Improved Aedes aegypti mosquito reference genome assembly enables biological discovery and vector control. Nature 563, 501–507 (2018). https://doi.org/10.1038/s41586-018-0692-z (2018)

- J. P. VanZee, J. A Schlueter, S. Schlueter, P. Dixon, **CA. Brito-Sierra** and C. A. Hill. Paralog analyses reveal gene duplication events and genes under positive selection in Ixodes scapularis and other ixodid ticks. BMC Genomics 17, 241 https://doi.org/10.1186/s12864-015-2350-2 (2016)