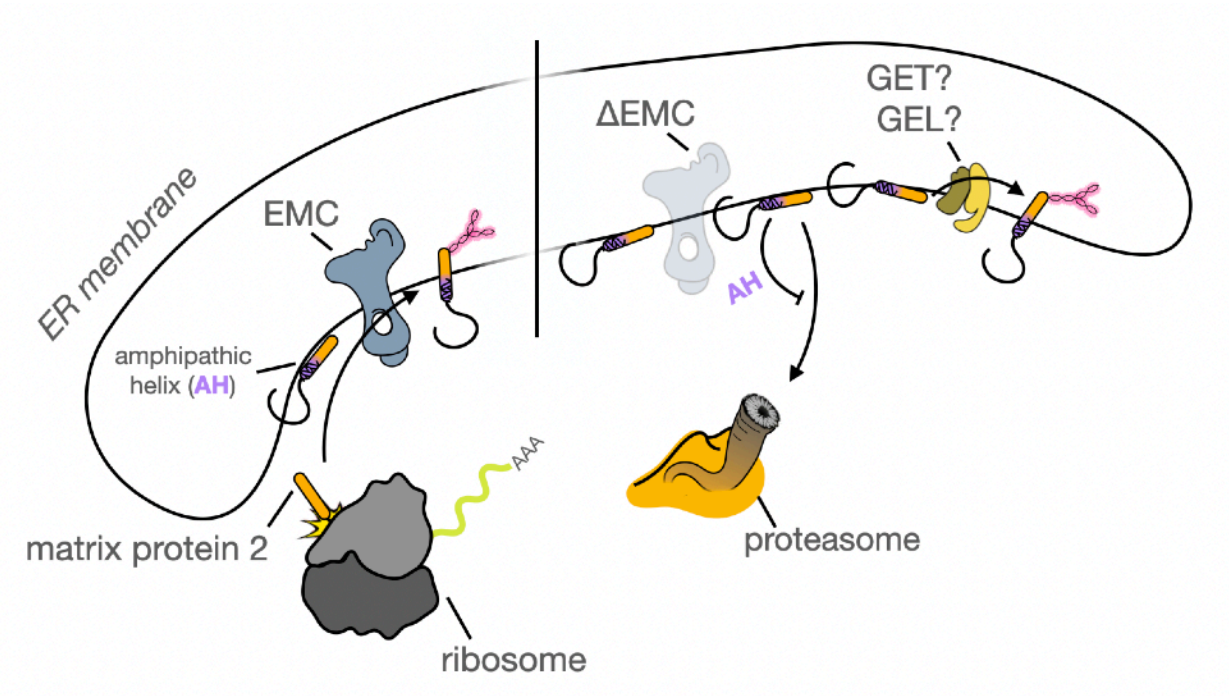
2024

Molecular Pathways and Mechanisms in the Biogenesis of Influenza A Virus Matrix Protein 2

Christian Diwo



Dissertation presented to obtain the

Ph.D degree in Integrative Biology and Biomedicine

Oeiras, December, 2024

ITQB NOVA, Avenida da República, 2780-157 Oeiras, Portugal | Tel +351 214 469 100 | itqb.unl.pt

SUMMARY

Membrane protein biogenesis is an essential cellular process, ensuring the proper synthesis, targeting, and insertion of proteins into biological membranes. This thesis investigates the molecular mechanisms governing membrane protein biogenesis, with a particular focus on how the Influenza A Virus (IAV) hijacks host cellular pathways to support viral protein production. The study primarily explores the role of the Endoplasmic Reticulum Membrane Protein Complex (EMC) in the biogenesis of the IAV M2 protein, an integral membrane protein that functions as a proton channel critical for viral replication, budding, and host cell remodeling.

Using a targeted CRISPR/Cas9 knockout screen, we identify EMC as a key factor in M2 biogenesis, supporting its correct membrane insertion, oligomerization and trafficking. Our findings reveal that while EMC facilitates efficient M2 biogenesis, alternative pathways, most likely other members of the Oxa1 family of insertases, can partially compensate for its absence. However, IAV replication is delayed in EMC deficient cells, leading to reduced cell-to-cell transmission. Structural and biochemical analyses further demonstrate that M2's transmembrane domain and C-terminal regions contain sequence features that dictate its EMC dependence, highlighting the importance of nascent chain properties in translocon selection.

Additionally, we find that M2, when not properly inserted into membranes, remains partly stable in the cytosol due to its amphipathic helix, which promotes membrane association and protects it from degradation. This suggests a viral adaptation that buffers against defects in membrane targeting and insertion, ensuring robust M2 availability during infection.

Moreover, our work provides broader insights into the mechanisms of membrane protein biogenesis and quality control, with relevance to other viruses and diseases associated with ER stress and protein misfolding. By elucidating the intricate interplay between viral infection and host cellular machinery, this thesis contributes to a deeper understanding of membrane protein homeostasis and potential vulnerabilities that can be exploited for antiviral strategies.

SUMÁRIO

A biogénese das proteínas de membrana é um processo celular essencial, garantindo a síntese, o direcionamento e a inserção adequados das proteínas nas membranas biológicas. Esta tese investiga os mecanismos moleculares que regulam a biogénese das proteínas de membrana, com um foco particular em como o Vírus da Gripe A (IAV) sequestra as vias celulares do hospedeiro para favorecer a produção de proteínas virais. O estudo explora principalmente o papel do Complexo de Membrana do Retículo Endoplasmático (EMC) na biogénese da proteína M2 do IAV, uma proteína de membrana integral que funciona como um canal de protões, sendo essencial para a replicação viral, o brotamento e a remodelação da célula hospedeira.

Através de um rastreio direcionado de knockout com CRISPR/Cas9, identificamos o EMC como um fator chave na biogénese da M2, apoiando a sua correta inserção na membrana, oligomerização e transporte. Os nossos resultados revelam que, embora o EMC facilite a biogénese eficiente da M2, vias alternativas – provavelmente outros membros da família Oxa1 – podem compensar parcialmente a sua ausência. No entanto, a replicação do IAV é retardada em células deficientes em EMC, resultando numa transmissão reduzida de célula para célula. Análises estruturais e bioquímicas demonstram ainda que o domínio transmembranar e as regiões C-terminais da M2 contêm características de sequência que determinam a sua dependência do EMC, destacando a importância das propriedades da cadeia nascente na seleção da translocação.

Adicionalmente, descobrimos que a M2, quando não corretamente inserida nas membranas, permanece parcialmente estável no citosol devido à sua hélice anfipática, que promove a associação à membrana e a protege da degradação. Isto sugere uma adaptação viral que minimiza defeitos no direcionamento e inserção na membrana, garantindo a disponibilidade robusta da M2 durante a infeção.

Além disso, o nosso trabalho oferece uma visão mais ampla sobre os mecanismos de biogénese e controlo de qualidade das proteínas de membrana, com relevância para outros vírus e doenças associadas ao stress do retículo endoplasmático e ao dobramento incorreto de proteínas. Ao esclarecer a complexa interação entre a infeção viral e a maquinaria celular do hospedeiro, esta tese contribui para uma compreensão mais profunda da homeostasia das proteínas de membrana e de potenciais vulnerabilidades que podem ser exploradas para o desenvolvimento de estratégias antivirais.

DECLARATION

I, Christian Diwo hereby declare that this dissertation and the data herein contained, are the results of my own work conducted between September 2019 and August 2024 in the laboratory of Dr. Maria João Amorim at Instituto Gulbenkian de Ciência (IGC), in Oeiras, Portugal, and under co-supervision of Dr. Colin Adrain at the Queen's University Belfast, Belfast, Ireland.

Financial support for this work was granted by a Doctoral Fellowship awarded to Christian Diwo by the IGC Integrative Biology and Biomedicine (IBB) program for the first year and since September 2020 by the Fundação para a Ciência e Tecnologia (FCT): DOCTORATES 4 COVID-19 2020.10234. BD.

The Amorim Lab was financially supported by an ERC grant project LOFlu (101001521).

Dr. Colin Adrain was financially supported by the Worldwide Cancer Research (grant reference 14–1289), Fundação para a Ciência e Tecnologia (FCT, grant references PTDC/BEX-BCM/3015/2014 and LISBOA-01–0145-FEDER-031330), the European Crohn's and Colitis organization (ECCO) and by a Marie Curie Career Integration Grant (grant reference PCIG13-GA-2013-618769) and by the La Caixa Banking Foundation (grant reference LCF/PR/HR17/52150018).

Work developed during this period, resulted the in following publications:

Etibor, T. A., O'Riain, A., Alenquer, M., **Diwo, C.**, Vale-Costa, S., Amorim, M. J.. Challenges in Imaging Analyses of Biomolecular Condensates in Cells Infected with Influenza A Virus. International Journal of Molecular Sciences, 24(20), (2023): 15253.

Faro-Viana J., Bergman ML., Gonçalves LA., Duarte N., Coutinho TP., Borges PC., **Diwo C.**, et al.. Population homogeneity for the antibody response to COVID-19 BNT162b2/Comirnaty vaccine is only reached after the second dose across all adult age ranges. Nature Communications 13, no. 1 (2022): 140.

ACKNOWLEDGEMENT

“Learn how to learn from those who disagree with you or even offend you. See if you can find truth in what they believe.”

Kevin Kelly, from 68 bits of unsolicited advice

First, I want to thank my parents and my family. I consider myself extremely privileged to be given the support I experienced, and the choice to waste opportunities in order to find my way and myself along that way. I want to thank my brother Matthias for lending me his confidence when I had little, and sharing the will and effort for constant improvement. I want to thank my grandma Erika, for sharing her wisdom and her ease of being.

I arrived in Lisbon the year before the pandemic, to visit two outstanding people and friends from my masters, and ended up staying for the opportunity to do a PhD. Mo who lives with his lovely family south in a bamboo forest, and Phil who organized my evening programs for that summer. I was fortunate to be sent to IGC when I expressed my desire to work in evolution in emails to researchers around Lisbon. There Claudia Bank took me in and I collaborated on a project to assess how the interaction of environment and genomic background shape the fitness landscape. I am immensely grateful for the space I was given and the smart and kind heads I was surrounded with to learn evolutionary theory and explore ideas. Because Claudia wanted to continue her work on viral evolution with a collaboration at IGC I was introduced to Maria João who became my supervisor when I joined the IBB PhD program. MJ was full of enthusiasm and ideas which convinced me that we could work together on a project, even before she won the ERC. The original project should have spanned molecular virology, cell biology and evolution in a project with MJ, Claudia and Colin Adrain. However we quickly realized that starting a new project needs agile planning and new tools, which took up most of the time in the project. MJ and Colin stayed as supervisors and together we explored how M2 interacts with EMC. Colin ended up moving to Belfast and I lost valuable partners in Catarina Gaspar and Miguel Cavadas when his lab members left. Luckily, Erika Carvalho was absorbed into our lab and we continued supporting our EMC-adjacent projects. Already before starting the PhD project Filipe Ferreira took on the challenge to teach me the ways of mammalian cell culture, viral infections, plaque assays and experimental planning. Although I may have stretched his patience thin at times he was always supportive. The same is true for Marta Alenquer, Temitope Etibor, Nuno

Santos, Silvia Costa and Daniela Bras who got to deal with my very excited, but little experienced self and directed me onto a good path to become self-sufficient in the lab, and I am grateful for that. To Camila Mariano, Aidan O'Riain, Ana Henriques, Oleg Chertkov, Marta Gallis, Isabel Sesifredo, Rita Pinto and my IBB colleagues who I will leave behind, we know there are no words to make it easier. Stay sane, bon chance!

Lastly, I want to thank my friends and flatmates who through all these years anchored me to the real world, or with whom I could escape it all together. In no special order but with a full heart Davide, Roberta, Marta, Ana, Fergus, Lucrezia, Gaeton, Eicul, Egroj, Gaby, Maja, Johnny, Bernie, Dela, Ivan, Clara, Maria, Michele, Riccardo and those who I forgot to mention in a hurry. A special thanks to Laura, who tolerated my recent decline of hygiene, and without whom I would have starved.

TABLE OF CONTENT

TABLE OF FIGURES

CHAPTER 1: INTRODUCTION

Under fortunate yet elusive conditions, biochemical systems evolved mechanisms to confine and separate ions and molecules into compartments. These compartments are either enclosed by membranes or created by the chemical processes within two immiscible liquid phases. Each strategy comes with its specific advantages, which together achieve control over the basic biochemical processes that are fundamental to cellular life. Here, we focus on the biology of membranes, while recognizing that valuable insights will be gained from studying emergent functions that stem from the distinct physical environments at the interface of membranes and phase separated compartments.

Biological membranes are highly specialized and selective barriers. The specific lipid and protein composition of these membranes establishes a compartment identity, which is tightly controlled and dynamically regulated to adjust its function. The membrane identity of hundreds of individual compartments in a cell is monitored by sensitive homeostatic regulatory mechanisms that balance the continuous synthesis, conversion and degradation of its components. Obligate intracellular parasites, like viruses, hijack these homeostatic systems with the introduction of as few as 1 to 20 genes, which reprogram membrane identities and remodel the cell into viral factories. The following introduction compiles research findings with the aim to understand the capacity of the membrane protein synthesis machinery to accommodate the flux of viral membrane proteins during influenza A virus (IAV) infection, a relevant human pathogen with the potential to cause yearly epidemics and occasional devastating pandemics of zoonotic origin.

1.1 The dynamic membrane identity

The lipid and protein composition of biological membranes determines their function. The human genome comprises roughly 20,000 genes, with about one third coding for proteins associated with membranes [1]. Half of these proteins are expressed in every cell, which together with an estimated 1,000 lipid species, compose membrane identities of a single cell [2], although vastly more lipid species are known [3]. The variety of functions these membranes perform is essential for the proper operation of hundreds of specialised cell types in the body, which change due to environmental cues and according to the cell cycle, circadian rhythm and aging. Membrane identities are quickly modulated through transport reactions, or the in-situ modification of proteins and lipids with methyls, alkyls, phosphates, carbohydrate or

ribonucleotide chains [4,5]. The identification of a cell type through its membrane identity can be exploited by viruses for cell entry and modulated to orchestrate viral replication and assembly, where the injection of a few genes may convert the cell into a viral factory.

Cells maintain their function through compartmentalised chemistry, for which membranes are essential. These membranes act as barriers to small molecules and ions, creating distinct chemical environments. Lipid bilayers form spontaneously in aqueous media. Their main constituents are phospholipids, which are long, slender molecules with polarised hydrophilic and hydrophobic characteristics. The hydrophilic phosphate head group is bonded to long apolar hydrocarbon chains. In an aqueous environment, the apolar tails interact with each other to avoid contact with water and assemble into a planar structure with a hydrophobic core devoid of water, and two hydrophilic surfaces, facing the aqueous media [6].

Membrane identity is determined by both the lipid and protein composition and their spatial organisation, which may differ between leaflets of a single bilayer. Lipid bilayers may be densely populated with proteins [7] which interact with lipids to shape membranes, define their mechanical properties, sense deformation, mechanical, electric and electromagnetic forces, regulate signalling and trafficking across membranes, and titrate molecular affinities. Continuous membranes can phase separate into sections of distinct organisation that limit lateral diffusion of proteins and lipids, termed lipid rafts [8]. Maintaining the identity of both the outer cell membrane and the internal membranes of organelles (the endomembrane system) is crucial for the cell's ability to respond to internal and external changes [9].

1.2 The flux of membrane components

1.2.1 ER targeting signals of proteins entering the secretory route

All proteins are synthesised in the cytosol, but those destined for the inside of an endomembrane compartment, the outside of the cell, or for insertion into membranes (transmembrane proteins) need to be targeted and translocated across membranes. Mammalian cells contain distinct endomembrane systems, each with specialised functions. The endoplasmic reticulum (ER) is the largest endomembrane system, stretching from the nucleus to the periphery of the cell. The vast network of membrane sheets and tubules are composed of subdomains for specialised function [10]. Importantly, the ER is the entry point for proteins that will either be distributed to organelles or secreted outside the cell [11]. Proteins that enter the secretary route can be soluble or membrane bound and have as defining features one or several

hydrophobic stretches of amino acids that function as signal for targeting [12], or to anchor proteins within the lipid bilayer [13]. Signal sequences that are located at the very N-terminus of a protein emerge first from the ribosomal exit tunnel during translation and are cleaved once a protein has been translocated across the ER membrane [14,15]. These signals have a short hydrophobic core of only 7-9 amino acids. Transmembrane proteins contain one or more transmembrane anchors depending on how often they weave in and out of the membrane with a length of approximately 20 amino acids. Signal sequences and transmembrane domains are mainly composed of hydrophobic amino acids, albeit they may also contain hydrophilic or even charged residues which are important for their function [16].

1.2.2 Transport of proteins and lipids throughout the cell

From the ER, proteins can reach the nuclear membrane through lateral diffusion [17] or traffic to the Golgi apparatus through vesicular transport [11]. From the Golgi proteins reach the plasma membrane, which is the continuous barrier between the cell and its environment, or lysosomes, which are organelles specialised in degradation. In addition to the ER and Golgi, other important organelles include endosomes, which coordinate the transport of engulfed molecules to different cellular destinations, peroxisomes, which are involved in lipid metabolism, and mitochondria, which convert ionic gradients into chemical energy, and have many functions in cellular metabolism [18].

Lipid synthesis takes place in many different organelles with mitochondria, ER, and Golgi synthesising the main diversity of membrane lipids and peroxisomes synthesising specialised membrane lipids [19]. The distributed synthesis requires an elaborate transport system to maintain membrane identity. Lipids are trafficked throughout the cell through vesicular transport, organelle to organelle contact sites, or lipid droplets. The synthesis and trafficking of lipids is finely controlled, as small percentages of differences in lipid composition can have dramatic effects on function. For example, small changes in the plasma membrane concentration of the lipid PIP2 changes the activity of membrane embedded potassium channels [20]. Furthermore, the lipid composition between leaflets of the lipid bilayer is finely controlled and maintained by proteins that can flip lipids between leaflets.

The lipid environment shapes the localization of its associated proteins [21]. Membrane thickness, lipid charge and lipid packing change throughout the secretory pathway [6]. The membrane thickness varies between different organelles along the secretory route, due to differences in lipid composition and the thickness of the hydrophobic core. Progressing from thin to thick from the ER → Golgi → plasma

membrane, membrane thickness influences the resident transmembrane proteome by matching the thickness of the membrane to the lengths of individual transmembrane domains [22]. The vast majority of transmembrane domains in eukaryotic cells are alpha helices, acting alone or in bundles of helices. A long hydrophobic helix may be tilted in a thin membrane to limit the contact between hydrophobic residues and the aqueous media, whereas a short hydrophobic helix may be restricted from diffusing into a membrane with a thicker hydrophobic core [23]. As the ER membrane is the beginning of the secretory route, its thin and soft membrane accommodates all transmembrane proteins.

The localization of proteins in the cell can also be determined by retrieval signals. Constant transport of proteins between the ER and the Golgi requires recycling of components that mediate transport, and return proteins that have been transported away from their intended resident organelle [24]. The proteins involved in vesicle formation and target membrane fusion need to be retrieved to the ER in order to maintain trafficking. Chemical inhibition of Golgi to ER transport, also blocks the export of proteins from the ER demonstrating the necessity of recycling of components [25].

Proteins and lipids are synthesised and trafficked throughout the cell and can be damaged through unfolding or oxidation and need to be degraded. The degradation of membrane proteins is an important aspect to maintain membrane function. Cells have the ability to degrade large portions of membranes or entire organelles through autophagy [26]. This process involves the sequestration of components in double membrane-bound vesicles, which subsequently fuse with lysosomes to degrade their content. Individual proteins, however, are degraded by the ubiquitin-proteasome system [27]. Such proteins that are recognised misfolded or non-functional are selectively tagged with a small protein named ubiquitin. Ubiquitinylated membrane proteins are extracted from the membrane and delivered to a protein degrading machinery called the proteasome, which unfolds the protein structure and hydrolyzes the protein into small peptides. The continuous synthesis, trafficking and degradation of membrane components has to be tightly regulated through feedback mechanisms in order to maintain organelle function or rapidly alter membrane identity in reaction to internal or environmental signals [9].

It is hard to imagine how the constant varying flux of thousands of components can maintain a finely-tuned membrane identity in hundreds of individual membranes in a cell and be maintained over the lifespan of an organism. Some of the complex feedback mechanisms have been elucidated which play a major role in organizing the flux of membrane components.

1.3 The unfolded protein response (UPR) as homeostatic sensor of the secretory pathway

Protein synthesis and degradation are tightly regulated depending on the functional needs of the cell. A mismatch between a demand in protein synthesis and degradation creates stress which is sensed in order to elicit countermeasures. The ER is the central entry point for proteins that are distributed to many other organelles and stress manifests in the accumulation of unfolded proteins in the ER lumen or membrane. Secretory demands are monitored through a system of 3 ER resident sensor proteins, that respond to ER stress with the regulation of proteins aiding in protein folding (chaperones), the expansion of the ER membrane [28] and upregulation of lipid biosynthesis [29–31], increased ER export [32], the degradation of misfolded ER proteins, and the downregulation of translation [33,34]. UPR is a major homeostatic system that is required for physiological adaptation to secretory demands and persistent dysregulation of UPR is associated with disease [35]. Importantly, if the cellular stresses cannot be resolved and the UPR is prolonged or severe, programmed cell death (apoptosis) is induced. Other organelles, like mitochondria, are monitored through analogous and interconnected protein quality control systems [36].

1.3.1 Protein folding at the ER

A major cause of stress in a cell is caused by misfolding of proteins. Protein folding of secondary and tertiary structure can occur spontaneously for simple proteins [37], but in most cases is assisted by chaperons that bind to the folding protein to guide the process [38]. In the ER a specialised system to monitor protein folding has evolved that monitors folding of soluble proteins or soluble domains of proteins through asparagine linked (N-linked) glycans and disulphide bond formation. Proteins that present a 3 amino acid motif NxT/S to the ER lumen are glycosylated by the OST complex [39]. The initially added glycan is trimmed and handed over to the calnexin-calreticulin complex that assists folding. The protein is released from the complex and the glycan is trimmed again to signal export from the ER. For correct folding a protein may require the crosslinking of a peptide chain through disulphide bonds between two cysteine residues. The enzyme protein disulphide isomerase (PDI) forms and rearranges these bonds in the oxidative environment of the ER lumen [40]. If a protein has not folded correctly, surface exposed hydrophobic stretches of amino acids, or rapid conformational changes in the structure will be detected by UDP-glucose glucosyltransferase (UGGT) and the glycan will be modified to reenter

calnexin-calreticulin assisted folding cycle [41,42]. A correctly folded glycoprotein may exit the ER. Besides the role of glycosylation in protein folding in the ER, glycans are further modified along the secretory route to regulate protein trafficking. Finally, glycosylation of membrane proteins on the plasma membrane can be interpreted by the immune system, or act as sensors of the environment [43,44]. New findings suggest that glycan modifications can be adapted in response to the environment, and thus contribute to the dynamic membrane identity [45].

A transmembrane protein is a continuous chain of amino acids with parts of the protein reaching the ER lumen, parts that are exposed to the membrane core, and parts that are located in the cytosol. The correct folding of all these domains has to be monitored and assisted. The soluble domains that reach the ER lumen may be folded as described above. The factors that are involved in aiding the correct folding of transmembrane domains after their insertion into the membrane have just recently been discovered. Membrane integral chaperones may associate with individual transmembrane helices to assemble bundles of transmembrane domains. The correct folding of some transmembrane domains of proteins that cross the membrane multiple times is mediated by the just recently discovered PAT complex [46]. Furthermore, the endoplasmic reticulum membrane protein complex (EMC) has been assigned a function in the assembly of complex ion channels [47].

1.3.2 Protein degradation at the ER

The accumulation of misfolded proteins in the ER needs to be controlled to prevent the induction of stress. The folding of a protein may fail terminally, which will commonly lead to their degradation through ER associated degradation (ERAD). ERAD is a process where an aberrant protein is ubiquitinated, extracted from the ER and sent for degradation by the proteasome in the cytosol [48]. The location at which the misfolding of a protein is recognised determines the downstream factors that are involved in the extraction process. ERAD-L substrates are recognised in the lumen, ERAD-M substrates are recognised in the membrane and ERAD-C substrates are recognised in the cytosol. Importantly, subunits of protein complexes which are not incorporated in their complex (orphaned) are sensed by ERAD-M and degraded. Marginally hydrophobic membrane proteins may translocate into the ER lumen when their binding partners are not present [49]. In mammals the recognition and ubiquitination machinery is highly diverse, and may consist of more than a dozen ER membrane bound E3 ubiquitin ligases that form retrotranslocation complexes with HRD1 or Derlins. Once ubiquitinated, a substrate is pulled out of the membrane by the highly conserved AAA-ATPase p97/VCP and handed over to the 26S proteasome

[48]. A glycosylated substrate may be deglycosylated by an enzyme called NGLY prior to extraction, but this does not seem to be essential for the retrotranslocation process [50]. Individual misfolded membrane proteins could also be directed to lysosomes that associate directly with ER membranes [51].

Besides the extraction of individual proteins, entire segments of the ER can be isolated for lysosomal degradation, a process termed ER-phagy [52]. ER receptors such as TEX264 [53], SEC62 [54] and others [55] recruit cytosolic components that sequester ER membranes and degrade its content through the fusion with lysosomes. A hallmark of these receptors are cytosolic, long unstructured domains and a LC3-interacting region (LIR) to interact with autophagic membranes. It remains unclear how the function of these constitutively expressed receptors is activated during UPR. Proteins that become faulty after leaving the ER may be as well degraded by the lysosomal system. The accumulation of aberrant proteins in the ER signals for unmet secretory requirements and triggers a cascade of transcriptional programs to limit the influx of proteins and increase the folding capacity of the ER via the UPR.

1.3.3 Sensing and response to ER stress

Evidence for the existence of the UPR first stems from experiments that overexpressed a misfolded mutant form of IAV hemagglutinin (HA) which triggered a transcriptional response of the major ER chaperone HSPA5 (BiP/GRP78) [56]. UPR is a transcriptional response to the accumulation of unfolded proteins, or the sensing of an aberrant lipid composition at the ER. The response can be varied depending on the exact type of stress, which are sensed by either of the three ER resident membrane proteins PERK, ATF6 or IRE1 [31,57,58].

Upon activation, PERK has an immediate effect on protein translation by phosphorylating eukaryotic translation initiation factor 2 subunit-α (eIF2α), which transiently attenuates translation and only allows the translation of specific transcripts. One of these transcripts is ATF4, a transcription factor regulating the UPR response partly through the induction of BiP [59]. ATF6 is another ER localised stress sponsor that is held in the ER by strongly binding to BiP. A so far unidentified mechanism dissociates BiP from ATF6 upon ER stress, upon which ATF6 trafficks to the Golgi where cleavage of a soluble domain fragment occurs, that subsequently translocates to the nucleus to act as transcription factor [60]. Cleaved ATF6 is the major transcription factor of BiP during UPR [61] in complex with important cofactors [62,63].

The ER resident transmembrane protein IRE1 senses ER stress in two distinct ways. IRE1 can bind unfolded proteins directly with its luminal domain, which induces dimerization and downstream signalling [57]. Furthermore, IRE1 can sense the lipid composition of the ER membrane. IRE1's single transmembrane helix has an adjacent amphipathic helix that interacts with the membrane. Amphipathic helices have a specific positioning of charged and apolar amino acids to form a hydrophilic and a hydrophobic face when coiled up in a helix structure. These chemical properties are ideal to interact with the amphipathic phospholipid molecules of the bilayer. IRE1's amphipathic helix may induce membrane curvature, dependent on the sterol content of a membrane, which may lead to its dimerization [21,64]. ERAD pathways have been implicated in the regulation of sterol synthesis and may be in this way linked to lipid composition induced UPR [48]. Activated IRE1 cleaves and promotes the editing of a cytosolic mRNA encoding XBP1. This triggers a shift in the reading frame which encodes for the transcription factor XBP1s. XBP1 mRNA expression is controlled by ATF6, which allows for a time-dependent response to elicit measures counteracting ER stress [65]. XBP1 mRNA is efficiently cleaved by the ER membrane bound IRE1, because it is localized to the ER membrane through its nascent chain, which can utilize the co-translational targeting machinery, although the translation product remains in the cytosol [66]. IRE1 acts independently of the transcriptional response, directly degrading ER localised mRNAs [67]. Furthermore, recent in situ cryo-electron tomography reconstructions reveal that activated IRE1 forms a double helix that constricts the ER membrane focally throughout the ER network [68]. As such, IRE1 may act as a diffusion barrier to isolate ER domains with unfolded proteins from the rest of the network.

UPR is a highly complex, regulated program with many levels of response to physiological conditions. In animals, IRE1 is regulated throughout the day in the liver by the circadian rhythm to meet the metabolic demand of organisms [69]. Chronic UPR response may be the cause of disease, especially when affecting terminally differentiated cell types, or organs of low regenerative capacity [70]. The inability of UPR to resolve chronic ER stress may have consequences for tissue function. Although not formally demonstrated, protein secretion may be enhanced during the UPR, given that ER export components [71] and exocytosis machinery components [71,72] are upregulated during UPR. This mechanism may distribute the stress across a tissue, which may either aid the resolution of the stress or contribute to disease [73].

UPR constitutes an important homeostatic sensor that can act fast by downregulating translation initiation and degrading mRNA at the ER membrane, when misfolded proteins accumulate in the ER. In order to enter the ER a protein that

is newly synthesized in the cytosol needs to be targeted to the ER and inserted into the membrane. This process may take place for as much as 2500 diverse membrane proteins translated at the same time [1], which includes the synthesis of membrane bound biogenesis factors themselves [255]. In the following section, I will present our current mechanistic understanding of how membrane proteins are synthesised, recognised for targeting to the ER membrane, handed over to translocation machinery and translocated across the ER membrane.

1.4 Membrane protein biogenesis

1.4.1 Targeting pathways for proteins entering the secretory route converge at the ER

Soluble and membrane bound proteins destined for the secretory route contain water insoluble stretches of hydrophobic amino acids that act as signals for targeting and membrane anchors, which make these proteins prone to aggregate in the aqueous cytosol. ER targeted proteins that fail to enter the ER are degraded to avoid the accumulation of toxic aggregates [74]. These errors may increase during aging and cause disease [75]. Thus, a high fidelity in targeting membrane proteins and secretory proteins to the endoplasmic reticulum is required for cellular health [75,76].

Targeting of a membrane protein is either initiated while its translation process is ongoing (co-translational targeting), or after translation has already finished and the protein is fully formed (post-translational targeting). A ribosome will initiate translation with a newly transcribed mRNA that translocates from the nucleus to the cytosol. Most mRNAs initiate translation from several ribosomes at the same time (polysomes). These large complexes may be targeted to the ER membrane, where new rounds of translation initiate close to the membrane until the mRNA is degraded. The protein encoded features that determine its targeting pathway and the rules for translocation at the ER are diverse, and need to encode which parts of a protein stays in the cytosol, and which parts are translocated across the ER membrane. The spatially distributed and stochastic nature of the targeting process leads to the same translated protein sequence encountering the ER membrane in different ways. This leads to a diversity of possible hand-over reactions of a translated protein to the ER translocation machinery happening simultaneously with many other hand-over reactions at the ER membrane (Figure 1.1).

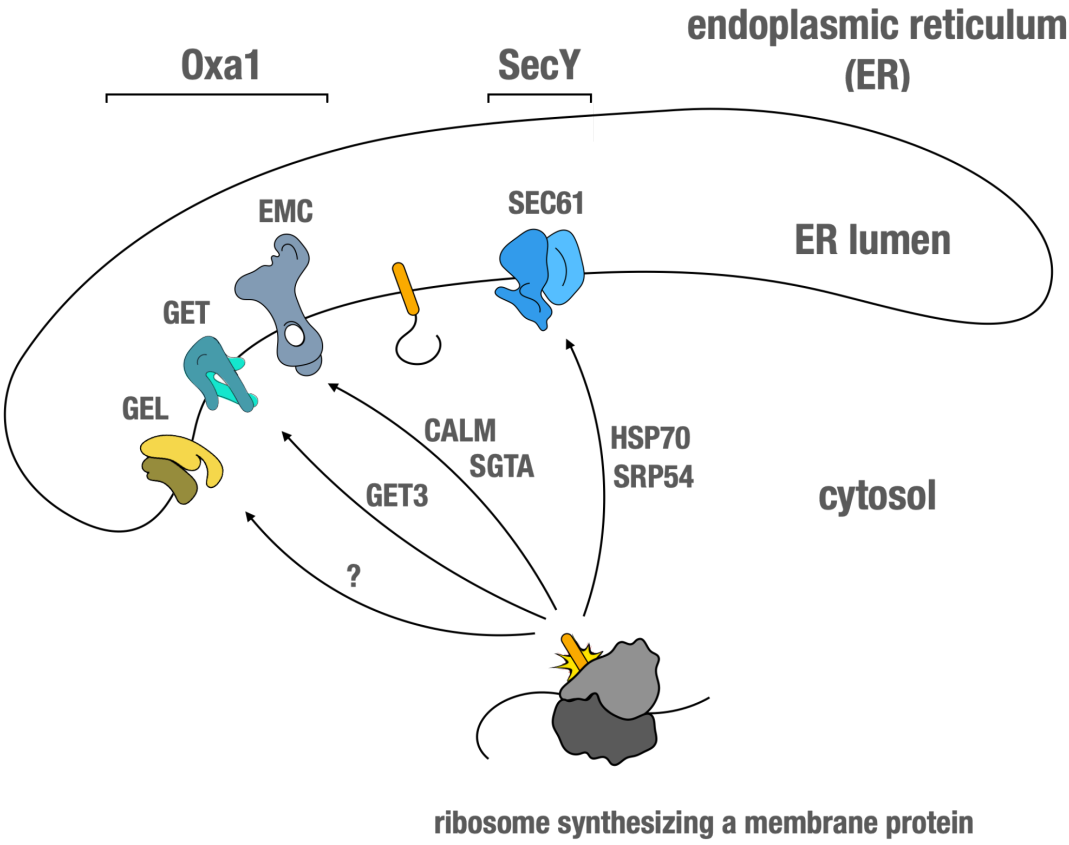


Figure 1.1 ER membrane protein biogenesis pathways

Membrane protein targeting is triaged depending on the hydrophobicity of a transmembrane domain sequence and the duration of translation after the emergence of the transmembrane domain from the ribosomal exit tunnel. A variety of ER targeting factors have been described that facilitate the shielding of hydrophobic stretches in the cytosol and the handover to a suitable translocon at the ER membrane.

1.4.2 Topogenic features of membrane proteins entering the secretory route

The interactions of a polypeptide with the ER membrane determines which transmembrane domain flanking region is translocated across the membrane and establishes the topology of the membrane protein. Two conformations are possible, either the N-terminus remains in the cytosol (Ncyto) or the N-terminus translocates across the membrane and the C-terminus remains in the cytosol (Nexo). The topology of the first transmembrane domain determines the topology of following transmembrane helices in multipass membrane proteins that weave in and out of the membrane. Topogenic features are flanking domain charge, flanking domain length, and the transmembrane domain length and hydrophobicity [77]. Charged residues are a strong determinant of protein topology. Positive residues tend to be retained on the cytosolic side, a phenomenon which is conserved from bacteria to eukaryotes but is to date not fully understood. In bacteria, the proton motive force is a major contributor

to the phenomenon, where charge is built up on the membrane due to a maintained proton gradient between the inside and outside of the cell. As the ER membrane does lack such an extensive force, several aspects are thought to contribute to the effect. Positively charged residues interact with negative charges within the ribosomal exit tunnel, which may retain the very N-terminus in the exit tunnel and extrude a nascent chain in a looped conformation [78]. Positively charged residues may interact with the charged phospholipid bilayer [254] or charge filters within the translocation machinery [140] may contribute to the retention of positive charges in the cytosol. As explained above, the length of the transmembrane domain flanking sequence is critical in the selection of a translocon and may contribute to membrane protein topology. The hydrophobicity of a transmembrane domain may influence the targeting route, or the energetic cost associated with partitioning into the membrane. The combination of these protein features determines how a membrane protein encounters the translocation machinery and how narrow or broad its requirements for specific factors are.

A specific classification of membrane proteins based on their sequence features and topology in the membrane has been established (Figure 1.2). TypeI proteins are proteins that are targeted via a cleavable N-terminal signal sequence and end up in an Nexo topology in the membrane. These proteins usually have large translocated domains. TypeII proteins may also have a large translocated domain and utilize as a targeting signal their N-terminally positioned transmembrane domain. This transmembrane domain is not cleaved and ends up in an Ncyto orientation in the membrane, where the sequence following the transmembrane domain is translocated across the membrane. TypeIII proteins also only have a transmembrane domain as targeting signal, which is positioned less than 50 amino acids from the N-terminus and end up Nexo in the membrane, translocating only a short hydrophilic segment. TypeIV proteins have as only targeting signal C-terminally located transmembrane domain. The protein is inserted in an Ncyto topology in the membrane, translocating only a short hydrophilic C-terminal domain of less than 50 amino acids. This scheme describes the topology of transmembrane proteins with a single transmembrane domain. The first transmembrane domain of multipass membrane proteins can be of typeI, typeII and typeIII and the remaining transmembrane domains weave in and out of the membrane in alternating topologies.

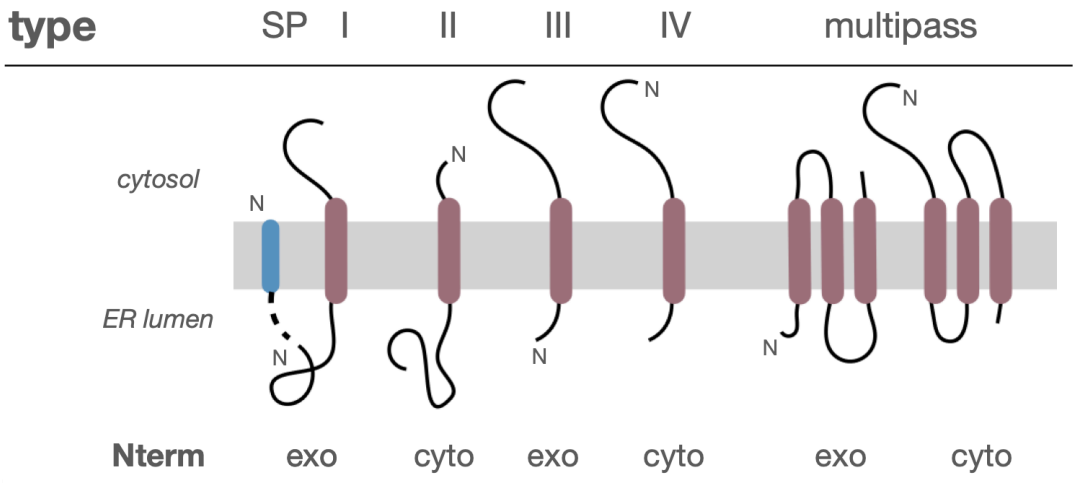


Figure 1.2 Membrane protein topologies

Membrane proteins are categorised depending on which transmembrane domain flanking region translocates across the ER membrane, and for the presence of a cleavable N-terminal signal peptide (SP). Multipass membrane proteins thread through the membrane in alternating topology of the transmembrane domain.

1.4.3 Co-translational targeting of proteins to the ER

Most proteins entering the secretory route are co-translationally targeted to the ER membrane, by a cascade of factors that assemble onto translating ribosomes, some of which have only recently been comprehensively described [76]. The nascent polypeptide-associated complex (NAC, NACA-NACB complex) is a central hub for the recruitment of factors that act co-translationally at the ribosomal exit tunnel [79]. NAC is present to high concentrations in cells, and associates stoichiometrically with ribosomes [80]. NAC scans the chemical character of the nascent chain within the ribosomal exit tunnel and recruits the far less abundant chaperone signal recognition particle (SRP) upon the detection of a hydrophobic stretch of amino acids [81]. SRP assembles out of 6 protein subunits and a long non-coding RNA [82]. The SRP54 subunit forms a flexible methionine-rich groove which positions itself directly at the ribosomal exit tunnel. The association of SRP54 with hydrophobic stretches of amino acids shields them from the cytosol and initiates membrane targeting of SRP in complex with a translating ribosome. SRP targeting is energy dependent. In its GTP bound state, SRP binds the bipartite ER receptor (SRPRA and SRPRB). Once bound, GTP hydrolysis compacts the receptor and subsequent binding to the small membrane protein TMEM208 releases SRP from the nascent chain close to the membrane to hand it over to a now proximal translocon [83,84]. NAC and SRP are

essential for the fidelity of protein targeting to organelles and the depletion of either leads to the mislocalization of proteins [76].

Co-translational targeting requires translation to be ongoing for a certain amount of time after the first targeting signal emerges from the ribosomal exit tunnel. This imposes a minimal C-terminal length on co-translationally targeted proteins, although there is no hard cut-off due to variation in elongation speed and the spatial context of translation. Proteome wide measurements in yeast determine that proteins with C-terminal length shorter than 100 amino acids are unlikely to be co-translationally targeted [85]. However, gene-specific aspects may attenuate the speed of translation to prolong targeting time such as rare codons that can be positioned downstream of a transmembrane domain to make SRP binding more likely [86]. Also stretches of positively charged amino acids or proline have been found to stall translation [87,88]. Furthermore, SRP binding itself slows down translation through interaction of its Alu domain with the elongation factor binding site of the ribosome [89].

Given the variety of factors that influence SRP targeting it is tempting to speculate that some proteins may be targeted co- and post-translationally, respectively, if the length of the C-terminus does not fully determine the translocation route, or when weak hydrophobic signals inefficiently engage SRP. The targeting time of a specific nascent chain may differ drastically, depending on the concentration and spatial distribution of targeting factors. The muscle-specific variant of NACA (NACAM) has an extremely long unstructured arm (2000 amino acids longer than NACA) to recruit SRP for co-translational targeting, which is likely to prolong targeting time. To date there is no direct measurement of the variation in targeting time of different membrane protein types. A study that attempts to estimate SRP mediated targeting time of a single typeIII reporter construct, finds that targeting is completed 3-4 seconds after signal emergence [90]. However, it is not clear if the obtained measurements are from translation that initiated in the cytosol, and includes the time for SRP mediated targeting and translocation, or if translation may have initiated at polysomes already anchored at the ER membrane and just includes translocation. Alternative approaches are needed that could specifically measure the different rates of targeting and translation in a variety of membrane protein topologies in different cell types.

1.4.4 Post-translational targeting of proteins to the ER

Post-translational targeting of proteins to the ER is needed if translation finishes before a hydrophobic signal has emerged from the ribosomal exit tunnel, or if

translation has proceeded for more than 140 amino acids before a signal emerges, and SRP recruitment becomes inefficient [91]. A network of cytosolic chaperones has been described to mediate post-translational ER membrane targeting [92–95].

Proteins with only a single transmembrane domain close to their C-terminus (typeIV/tail-anchored) are targeted to the ER membrane via a network of factors. Similar to SRP mediated targeting, the initial recognition of a hydrophobic segment is performed by factors associated with the ribosomal exit tunnel, like BAG6 [96] and HSP70s [97]. Once captured the protein destined for membrane insertion at the ER is handed over to GET3, SGTA or calmodulin depending on the hydrophobicity of the transmembrane domain [93]. HSP70-mediated folding as well as the handover reaction of a substrate from GET3 to its membrane receptor at the ER is ATP dependent [93]. In the absence of GET3 a HSC70/HSP40 dependent pathway has been described to target tail-anchored membrane proteins to a yet unidentified ER membrane receptor [98]. SGTA or calmodulin have been described to associate with relatively hydrophilic transmembrane domains, but we do not understand how the association of these factors with a membrane protein leads to membrane targeting, as no cognate ER receptors are known [99]. SRP independent targeting pathways are more active in yeast where the SRP independent targeting pathway (SND pathway) has been characterised in molecular detail [100]. The pathway in humans lacks molecular detail but is described to be important for proteins with centrally located transmembrane domains and for GPI anchored proteins [95]. Post-translational targeting reactions also makes it possible for cytosolically translated proteins to reach mitochondrial membranes [101].

Besides, these mechanisms for targeting that act on the protein itself mRNA features can localise translation to the ER membrane. The localised translation of mRNAs at the ER has even been observed for membrane proteins with C-terminal signal anchor, which should only be post-translationally targeted to the ER [102]. Recently a phase separated compartment that is associated with the ER has been described to possibly mediate localised translation [103]. However, it is clear that many targeting pathways converge at the ER membrane which need to be simultaneously handled at the ER and bear the potential for regulation, hierarchical processing, and flexible assignment of translocation machinery. For this, the ER membrane harbours highly conserved translocases to receive the diversity of targeting pathways.

1.4.5 The diversity of membrane translocating complexes (translocases) at the ER

Membrane protein translocation is a prerequisite for cellular life and the mechanisms that facilitate membrane insertion in modern eukaryotes are as old as the last common universal ancestor [104]. Two main families of translocon types are conserved as far back as the last universal common ancestor and can handle the diversity of membrane protein topologies described above. The Oxa1 superfamily with eukaryotic members residing in the bacterial inner membrane (YidC), mitochondrial inner membrane (OXA1), chloroplasts (ALB3) and the ER membrane (GET1/WRB, EMC3 and TMCO1) and the SecY superfamily with its eukaryotic homolog SEC61A1 and bacterial counterpart [105]. SecY pseudosymmetric halves may have evolved from a gene duplication of the Oxa1 family members [104]. The diversity of translocation reactions can be divided between these archetypes. The Oxa1 superfamily can only translocate short transmembrane domain flanking regions (hydrophilic stretches of amino acids flanking the hydrophobic transmembrane domain), or short loops between two transmembrane domains across the ER membrane. The SecY can translocate long terminal hydrophilic domains or loops [106].

The difference in translocation capacity stems from their structural differences. Oxa1-type translocases create a hydrophilic vestibule deep in the core of the membrane, which locally collapses the lipid bilayer and thus lowers the energetic barrier for the translocation of short hydrophilic sequences. Oxa1 family members have a conserved transmembrane domain motif, a bundle of 3 transmembrane helices that associate with accessory proteins to form the hydrophilic vestibule [107]. GET1 forms this vestibule with its obligatory partner GET2/CAMLG (GET complex), EMC3 with EMC6 (the entire EMC complex has 9 subunits in humans) and TMCO1 with OPTI (GEL complex). The conserved function is impressively demonstrated by EMC3 and EMC6 being able to rescue OXA1L deficiency in the mitochondrial inner membrane [108]. *In vitro* studies purifying the intact 9 subunit human EMC for reconstitution in liposomes, deliver the best evidence to date for EMC mediating membrane translocation of typeIV and typeIII proteins [109,110]. On the other hand, SecY-type translocons (in humans SEC61A1 with its obligatory cofactors SEC61B and SEC61G, from here on the trimeric translocon is just called Sec61) have two pseudo symmetric domains that form a protein conducting pore, for the translocation of soluble proteins and long transmembrane domain flanking regions into the ER lumen. A lateral gate to the membrane plane can open for the insertion of transmembrane domains that thread through the pore [106].

There may be several reasons why two translocation architectures have been conserved through evolution. SecY needs to be gated open by a hydrophobic signal, whereas Oxa1 family members are thought to be constitutively available for translocation. SecY translocation also requires the assembly of ER luminal chaperons like BiP to aid in translocation to generate an energy dependent pulling force, whereas the translocation through Oxa1 is energy independent. Furthermore, the gating of Sec61 in the ER membrane leads to a leak of calcium ions which have to be pumped back into the ER storage, costing energy [111]. Another reason for conserving two biogenesis pathways may be that SecY gating is slow compared to Oxa1 translocation [112]. Besides the time needed for the gating of the translocon, translocation of a polypeptide chain may proceed at 40 amino acids per second through the bacterial SecYEG (Sec61 homolog) [113], whereas the translocation of short N-terminus of only 40 amino acids through the bacterial YidC, may be accomplished at a timescale up to a magnitude shorter requiring only 50ms [114]. The mammalian EMC has been found to translocate two transmembrane domains connected by a short hairpin [106]. The short spacing of transmembrane domains may require rapid translocation through Oxa1 before transmembrane domains build up on the cytosolic face, which may aggregate and become incompetent for translocation. In humans EMC, GEL and Sec61 have been shown to cooperate for the translocation of polytopic membrane proteins where select transmembrane domains, or bundles of transmembrane domains, are translocated through EMC and long loops through Sec61 [106]. The Sec61 translocon has been shown to assemble with other factors to serve the diversity of membrane proteins [115,116]. Secretory proteins with marginally hydrophobic signal sequences or short overall length of less than 100 amino acids are not detected by SRP and require additional factors to efficiently engage Sec61. The SEC62-SEC63 complex in humans has been identified to be important for the translocation of membrane proteins of these signal peptide containing proteins [117–120]. Recent cryoEM maps suggest a stepwise gating function, of SEC63 and SEC62, engaging with Sec61 in order to receive a posttranslationally targeted signal sequence [121]. Furthermore, SEC62 has been shown to interact with a typeII protein to influence its final topology after an initial insertion in an inverted topology [122]. Although SEC62 is suggested to act post-translationally due to SRP not being required for SEC62 client protein biogenesis, it presents a ribosomal binding site and may be involved in cotranslational membrane insertion. The full client range and the determinants for the selection of a specific translocon is subject of ongoing investigation, as is potential redundancy between the pathways or the assembly of

specialised complexes from Oxa1 or SecY subunits with each other or other ER proteins [123].

1.4.6 Less well understood factors in membrane targeting and translocation

Genetic studies in bacterial systems have pioneered the discovery of many highly conserved membrane protein biogenesis pathways. Besides the targeting and translocation machinery discussed above, bacteria have evolved several other mechanisms which have either been lost in eukaryotes or have not yet been discovered. The bacterial complex SecDF has long been described to be able to conduct proteins across bacterial membranes [124]. Recent structural analysis identifies its association with the bacterial SecYEG to act as a motor that aids in protein translocation [125]. So far, the eukaryotic counterpart of SecDF has not been formally identified, but modern computational tools to compare structural homology suggest PTCH1 and the complex of NPC1-NPC2 to be structurally highly similar eukaryotic membrane proteins despite having only 11-15% sequence homology. These proteins are however located at the plasma membrane and lysosomes respectively and a role in protein translocation has not been described.

Another fascinating bacterial mechanism of membrane protein translocation involves the glycolipozyme MPIase [126,127]. Bacterial membrane inner leaflets present glycolipids that have been shown to target membrane proteins to YidC (bacterial Oxa1-type translocon) and cooperate in their translocation across the membrane [126,127]. There are no cytoplasmic facing glycolipids described today, however the enzyme that catalyses the glycan addition to diacylglycerol CDS2 is present in eukaryotes, resides in the ER and its active site is facing the cytosol.

Another protein which has long been identified to bind ribosomes at the ER is the large 180kDa typeIII protein RRBP1 [128,129]. Although RRBP1 is implicated in many disease phenotypes, the lack of mechanistic studies leaves its function elusive. It remains unclear if this binding function has a role in membrane protein biogenesis or if non-translating ribosomes can be stored close to the ER membrane in this way [124].

Post-translational targeting of peripheral membrane proteins specifically to the ER membrane has been described to be mediated by the selective affinity of an amphipathic helix to a certain lipid composition [130]. A similar targeting mechanism may be possible for ER translocated proteins or membrane proteins, although this has not been described.

There may be additional relevant signal and receptor mechanisms beyond the canonical targeting functions that can regulate membrane protein biogenesis and contribute to fidelity and robustness in the system.

1.4.7 Handover of ER targeted proteins to the ER translocation machinery

The specific modalities on how a protein destined for the secretory pathway interacts with the ER membrane creates a diversity of handover reactions from the targeting machinery to the translocation machinery. Specificity between the targeting reaction and the translocation machinery may be established by the targeting factor and a cognate receptor, or by the nascent chain interacting with the translocation machinery.

SRP mediated targeting ends with the translating ribosome being released from the SRP with the exit tunnel located around 5nm from the ER membrane [131], or even closer if SRP engages with TMEM208 [84]. At this point the interaction of the nascent chain with proximal translocons, or the surrounding lipids may tether the ribosome to the membrane, before translocation [114,131,132]. The sampling of the proximal membrane may be slow compared to the translocation step suggesting a potential for regulation [131]. TypeI and typeII N-terminal hydrophobic signals have been exclusively associated with SEC61 mediated translocation and crosslink to subunits of this translocon [133]. TypeIII proteins seem to have a more flexible choice for translocation [110,131]. A nascent chain for which translation has initiated further away from the ER will have a different interaction with the ER membrane than a nascent chain for which translation has initiated after the mRNA has been targeted to the ER membrane. The length of the typeIII nascent-chain protruding from the ribosomal exit tunnel, that interacts with the ER membrane, has been shown to affect its ability to engage EMC or SEC61 [131]. Synthetic typeIII constructs that are stalled in translation at least 50 residues after the N-terminal signal anchor are dependent on EMC for translocation, and crosslink to EMC3. Constructs stalled at shorter length, where the hydrophobic signal anchor barely emerges from the ribosome are sensitive to Sec61 inhibition and do not crosslink to EMC [131]. However, even when EMC and Sec61 are inhibited, a residual translocation reaction takes place through a yet unidentified mechanism. The typeIII protein BCMA has been shown to be entirely independent of EMC and Sec61 for translocation, however also here the exact translocation mechanism is unknown [134]. BCMA forms a beta sheet structure at its N-terminus, which may influence the selection of a translocon.

Specific protein features, like length and charge of the translocated domain and the hydrophobicity of the transmembrane domain may require special handling, where handover to unsuitable translocases need to be prohibited. Hydrophilic transmembrane domains may be inefficient in partitioning through the Sec61 translocon lateral gate into the lipid bilayer, which leads to the complete translocation

of a protein into the ER lumen [135]. Thus, mechanisms need to be in place to prevent unwanted handover reactions. The N-terminal secondary and tertiary structure of a membrane protein probing the ER membrane influences translocon selection. Sec61 has been described as being inefficient in translocating intrinsically disordered proteins (proteins that lack secondary structure), albeit having a functional signal peptide fused to their N-terminus [136–138]. The mechanism of how unstructured domains prevent translocation remain elusive. It is possible that the positively charged exposed backbone may hinder translocation, or that an interaction with membrane surfaces prevents efficient translocon engagement. Tertiary structure folding can prevent translocation, simply by steric hindrance. The ribosomal exit tunnel may be able to contain folded nascent chain structures that are too large to be translocated by Sec61 [131,139]. N-terminal charges are important signals, affecting how the ER membrane is probed for a suitable translocon. The positively charged hydrophilic vestibule of the EMC can act as a charge filter to reject certain unsuitable client proteins [140]. The same charge filter may be active in other Oxa1-family members [141]. A recent study shows that the charge in the N-terminal domain of multi-pass membrane proteins can make translocation via EMC inefficient, but that the Sec61 or GEL complex are able to partly compensate for the inefficient translocation reaction [123]. Transmembrane domain features influence translocon selection as demonstrated by typeIII reporter constructs harbouring different transmembrane domains [110]. How these transmembrane domains influence translocon selection remains reclusive.

For post-translationally targeted proteins, the specificity for a translocation machinery may be established by the cytosolic targeting factor to its cognate machinery at the ER. However, only the GET complex is targeted by its cognate targeting factor GET3. GET3 interacts mainly with hydrophobic transmembrane domains of typeIV proteins that are loaded onto GET3 by other cytosolic chaperones [92,142]. The transfer of a membrane protein to GET3 may be prohibited by low hydrophobicity of the transmembrane domain [109]. However most post-translationally targeted membrane proteins may have no strong deterring properties. The typeIV protein SEC61B has been described to be targeted to both GET and EMC for translocation [123], demonstrating a redundant role of both pathways. Handover of membrane proteins to EMC via SGTA or calmodulin is thought to occur not through direct interaction, but through proximal release of the protein and subsequent capture by EMC [99]. Similarly, it is not clear how targeting via HSP70s can be specific, as they engage proteins targeted to the ER and mitochondria [93]. A subset of post-translationally targeted mitochondrial proteins first associates with the ER

membrane and is subsequently transferred to mitochondria, via ER-mitochondrial contact sites [143,144]. The peripheral association with the ER membrane may lead to unwanted insertion into the ER, which is corrected through ERAD-independent retrotranslocation via ATP13A1 [144,145].

The GEL complex has only been described to aid in the translocation of polytopic membrane proteins [106,146]. No handover reactions from the targeting machinery are known.

Handover reactions may be time sensitive, either due to limited targeting time before the termination of translation, or the increasing probability of spontaneous hydrolysis of triphosphates in energy dependent processes. Specific protein features like the gating efficiency of signal peptides may lead to differences in sampling rate of the ER membrane for translocation [147]. In order to maintain quality in membrane protein translocation, redundancies and buffering mechanisms are necessary, especially under high load scenarios, where many proteins destined for the secretory route are translated and targeted simultaneously. Certain biogenesis pathways have been shown to be entirely redundant depending on the cellular context. The loss of the EMC pathway has a wide distribution of effects on growth when knocked out in a large number of cancer cell lines, ranging from essential to no effect (depmap database [148]). Interestingly, the depletion of EMC and GET together becomes lethal in cancer cell lines, suggesting a potential client overlap between the two pathways [149].

A multiomics approach finds that components of the membrane protein biogenesis machinery are regulated during UPR. Ribosomal footprinting data and shotgun proteomics suggest that SEC61 components, SEC63, and EMC3 are upregulated 6h post induction of ER stress [71]. Similarly SRP targeting components are increased, as well as SEC61 and EMC cofactors TRAP1 and CCDC47. The increase in membrane protein targeting and translocation factors during UPR may reflect the increased need for chaperone translocation into the ER lumen and ER membrane which are SEC61 and EMC dependent.

Systematic studies that assess the rate limiting step in the synthesis of transmembrane protein types with various physico-chemical characters are lacking. Further studies are needed to understand the capacity of the membrane protein biogenesis network to dynamically regulate handover reactions to distribute loads between translocation pathways and cope with secretory demands. The identification of bottlenecks in the system may give new insights into conditions where the loss of proteostasis is the cause of disease, or where the cellular biosynthesis machinery is hijacked by viral pathogens.

1.5 Influenza A virus (IAV) demands on host protein synthesis machinery

Viruses are obligate intracellular parasites and need to transmit in order to persist in a population. Respiratory viral infections are usually cleared rapidly by the hosts, which gives these types of viruses a limited window for transmission [150,151]. Virus particles that bear a lipid membrane (enveloped viruses) are not stable on surfaces outside the host. Although large deposits of virus on surfaces (fomites) may achieve contact transmission, the transmission via droplets and aerosols is thought to be dominant [152,153]. Enveloped respiratory viruses such as viruses of the family *Orthomyxoviridae*, *Paramyxoviridae*, or *Coronaviridae* generate enormous amounts of infectious virus particles in a short time in order to maximise the chances of transmitting from host to host [150,154–158]. This in turn requires extraordinary mechanisms of transforming cells into viral factories to rapidly produce their components. Of special interest for human health is the respiratory enveloped virus, influenza A virus (IAV) of the *Orthomyxoviridae* family. IAV has caused several pandemics in the past and continues to pose this threat due to recurring seasonal infections in humans and many animal hosts, frequent spillover and genetic reassortment mechanisms. These characteristics may lead to abrupt changes in genetic diversity and the emergence of virus variants that circumvent preexisting immunity or increase transmission rate. In the following sections, we investigate IAV's mechanisms of replication, that allow the virus to maximise virion production with a specific focus on viral protein synthesis.

1.5.1 IAV infection cycle

Infectious virus particles (virions) that enter a naive host organism need to follow similar stages in order to infect and transmit. Virions need to enter a cell, release the genome, replicate the genome and transcribe mRNA, synthesise protein components, assemble the components and the replicated genome and egress the cell (reviewed for IAV in Carter et al., 2024). Numbering corresponds to numbers in Figure 1.3. IAV invades the cell through receptor mediated endocytosis (1) [160,161], where it escapes late endosomes by fusing its envelope with the endosomal membrane (2) [162,163]. Upon fusion of the membranes the viral genome is released from the virion into the cytosol (3) [164–166] and transported to the nucleus (4) [167], which concludes the entry phase (reviewed in [168,169]). IAV's single stranded RNA genome is segmented into 8 segments [170], which encode the viral proteins in a negative sense as a template to transcribe mRNA [171]. In the nucleus, the genome is initially transcribed into mRNA by the viral polymerase which is complexed with each

segment [172]. mRNA transcription initiates with 5' fragments cleaved from capped host mRNAs [173] and adds a polyadenylated sequence through stuttering (5) [174,175]. In the cytosol the mRNA is captured by the host translation machinery (6). Some viral proteins enter the nucleus to assist viral replication to synthesize full length positive sense RNA (complementary RNA) [176] as a template to amplify the progeny negative sense genomic RNA (vRNA) (7). vRNAs assemble into ribonucleoproteins (vRNPs) with one copy of the trimeric viral RNA dependent RNA polymerase that sits at base paired 3' and 5' termini [177] and the remainder of the RNA is coated by the nucleoprotein (NP) [178]. Newly formed vRNPs exit the nucleus via a regulated process involving a daisy chain reaction that recruits sequentially the viral proteins M1 and NS2 and the host cell factor CRM1 (reviewed in [168]). Virus assembly takes place at budzones at the plasma membrane [179], where the structural viral proteins cluster and combine with a full set of 8 individual genomic vRNPs (8) [180,181]. The assembly process drives the budding of the membrane to pinch off fully assembled virions ready to infect the neighbouring cells [182].

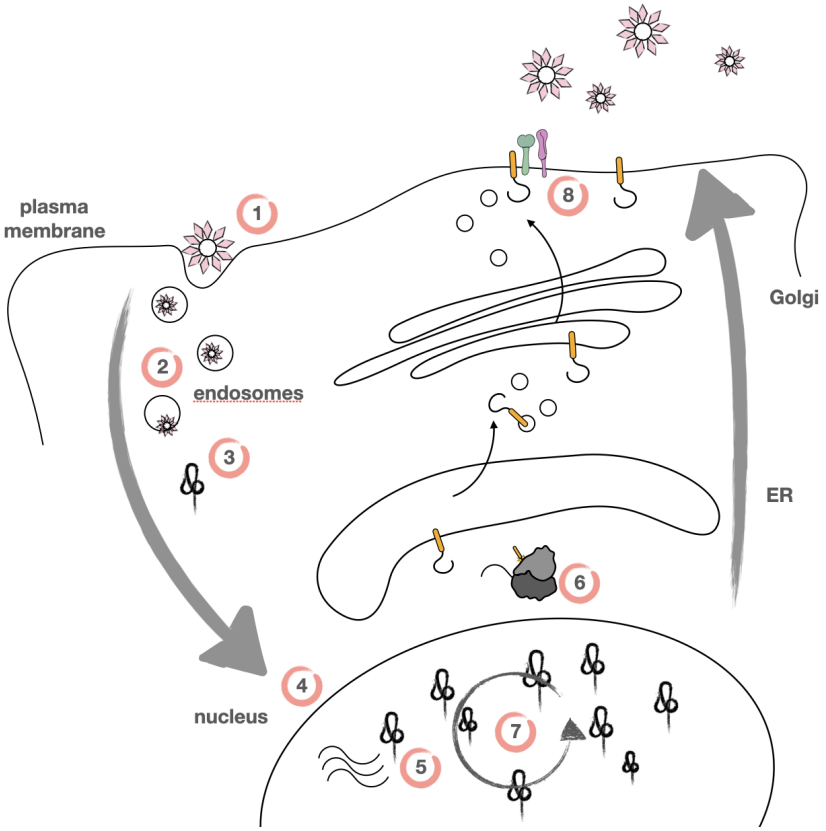


Figure 1.3 IAV infection cycle

IAV binds to sialic acid moieties presented at the cell surface and will be endocytosed (1). The acidification of endosomes triggers membrane fusion between the viral envelope and the endosome membrane (2) and the dissociation of the M1 capsid to release the segmented genome into the cytosol (3). The viral genomic segments are transported to the nucleus (4),

where mRNA translation of viral proteins takes place (5). Viral mRNAs are translated by cytosolic ribosomes (6) and imported into the nucleus to aid in viral genome replication (7). Viral structural proteins and a full set of 8 vRNPs assemble into budding virions at the plasma membrane (8).

1.5.2 IAV remodelling of the host cell

IAV genome replication in the nucleus of infected cells is an exception for negative sense RNA viruses, but gives the mRNA access to the splicing machinery. The 8 genomic segments encode 10 essential proteins and in addition, depending on the strain, up to 11 accessory proteins [159]. Three of these proteins are translated from spliced mRNAs and other mRNAs produce alternative reading frames by ribosomal frameshift or leaky scanning to initiate translation at a downstream site. The expression of only a few viral genes extensively remodels the transcriptional program of the infected cell [183]. The concerted mechanisms to reduce host mRNA, lead to viral mRNA contributing 60% to all transcripts by 8h post infection [184]. However, the dynamics and magnitude depends on the number of infectious particles entering a cell (multiplicity of infection, MOI). The above-mentioned process by which viral mRNAs acquire their 5' caps, termed cap-snatching, contributes to the reduction in host mRNA [185]. Cap-snatching globally reduces transcription elongation [185,186], but active host mRNA transcription seems to be required for viral mRNA synthesis [187]. Interestingly, viral mRNA expression peaks at around 3h after infection and completely subsides by 5h [188,189]. Novel sequencing methods suggest that only polymerases that are delivered with the genome during infection can synthesise viral mRNA [189]. Another virally encoded factor acting on host RNA is the endonuclease PA-X, a frameshift product of IAV segment 3. PA-X selectively degrades host mRNA in the nucleus, preserving viral transcripts and host transcripts that maintain cellular functions important for viral replication [190]. The multifunctional viral protein NS1 is the main countermeasure to interferon mediated host defences [191]. The effect is mainly mediated by interfering with 3' processing of host transcripts [192], however this is not conserved in all viral strains [193].

IAV infection generally leads to a decrease in the total rate of protein mass synthesised. Various studies report drastically different effect sizes from 80% decrease in total protein synthesis [194] to approximately only 20% decrease [184,195]. This may be due to differences in the magnitude of host shutoff of different viral strains, or due to differences in the ability to circumvent host defences that shut down translation [196,197]. However, it seems that IAV subverts cellular resources for viral protein production, rather than increasing the overall translational burden. On the

other hand, studies measuring the metabolic activity report increased glycolysis and overall metabolic flux, necessary to sustain high viral replication [198–200]. In the absence of increased overall protein synthesis activity, an increase in RNA synthesis and lipid anabolism may result in an overall increase in respiration. Nevertheless, viruses depend on the efficient translation of structural and non-structural proteins, and especially membrane proteins with their complex biogenesis pathways may be bottlenecks in viral replication.

1.5.3 IAV membrane protein expression and the role of UPR

The activation of specific arms of the UPR may be beneficial for viral replication, whereas others may be detrimental and actively countered by infection. UPR can be induced with the use of small molecules which has an overall detrimental effect of IAV replication, suggesting that IAV infection controls UPR processes [201,202].

Influenza A virus expresses 3 essential single pass membrane proteins, which are all integral to the viral envelope and need to localize to viral budding sites at the plasma membrane: hemagglutinin (HA), neuraminidase (NA), and matrix protein 2 (M2). HA and NA are large glycoproteins of typeI and typeII topology respectively, which require SRP mediated targeting and Sec61 mediated translocation at the ER [203–207]. M2 is a small typeIII membrane protein with less well understood biogenesis requirements. HA, NA and M2 are expressed to similar amounts in cells [194], but HA and NA incorporate with more copies into virions (960 copies of HA to 220 copies of NA to 60 copies of M2 for spherical virions [208–211]).

Despite the overall decrease in protein synthesis, the expression of HA and NA with large translocated domains and complex folding, may require an increase in ER chaperones. However, the reported effects of infection on BiP expression are inconsistent. Infection has been found to either drastically upregulate BiP 12h post infection [212], or have no effect at all [213,214]. BiP expression could be a double edged sword, that increases the folding capacity of the ER, but also may increase cytokine production and signal innate immune responses, which may result in increased tissue damage [215,216]. UPR activation is linked to pathogenicity during viral infection, which differs between strains. Recently emergent strains have evolved a higher number of glycosylation sites on HA to assist folding, which is associated with less UPR and pathogenicity [215], whereas HA with less glycans may be subject to rapid degradation [212]. Besides the demand on ER folding capacity, HA and NA's newly ascribed function of remodelling the ER through a zippering mechanism (in preparation), may induce ER stress by effectively reducing ER size.

As viruses are obligated to use the host translation machinery for viral protein expression, PERK mediated attenuation of translation should be avoided. Indeed, IAV infection is thought to limit eIF2a phosphorylation [195,197], and PERK. EIF2a phosphorylation may also be initiated by antiviral defences through the sensing of double stranded RNA [197], and may not necessarily indicate that viral gene expression burdens the ER.

XBP1 splicing may be induced during IAV infection. The inhibition of XBP1 splicing results in potent inhibition of virus particle production [214]. However, the study used inhibitors with insufficiently described mechanisms of action and unclear specificity [217]. Another study finds that the inhibition of XBP1 splicing has a more moderate effect on viral replication [213], using an inhibitor with a better defined mechanism of action [218]. It remains unclear if IRE1 mediated degradation of ER localised mRNAs is affected by these chemical interventions. The initial degradation of host mRNAs at the ER could be important for a rapid switch in the translation of viral membrane proteins.

IAV induced host shutoff contributes to alleviation of ER stress. NS1 mediated inhibition of host mRNA adenylation can reduce UPR measured by XPB1 mRNA splicing [219]. Importantly, not all IAV strains retain this function. The lab adapted strain PR8 encodes NS1, which has lost the ability to modulate host mRNA maturation, and infection with PR8 virus induces UPR [213,219]. Despite IAV infection reducing overall protein synthesis, the diversity in the expression of different proteins may be majorly reduced and lead to limitation of specific biogenesis factors and ER stress. Publicly available ribosomal footprinting data may answer how the demand on membrane protein biogenesis shifts during IAV infection [184,220].

IAV induced changes in host cell transcription seem to be necessary to free up resources for efficient viral membrane protein synthesis. To date it remains unclear how the massive shift in protein expression can maintain membrane identities necessary for membrane protein synthesis and viral assembly. Proteomic and lipidomic approaches with subcellular resolution may be necessary to identify how compartment functions are reprogrammed during infection. M2 is of special interest for its function as a proton channel, as organellar pH is tightly regulated along the secretory route and changes may subvert membrane trafficking to benefit infection. M2 has been shown to elicit damage responses that may subvert membrane trafficking [221] and induce inflammasome activation [222].

1.5.4 IAV M2 protein

So far no specific translational burden has been attributed to M2, also because its specific requirements for membrane protein biogenesis remain elusive. With two large glycoproteins that share the similar biogenesis requirements, M2 may have evolved alternative requirements in order to avoid competition.

M2 is a small 97 amino acid long protein of typeIII topology that needs to localize to the plasma membrane during infection to be incorporated into budding virions [211,223]. M2 is incorporated into virions with only very few copies, but may be expressed in large amounts depending on the viral strain [194,224]. M2 serves multiple functions during the IAV infection cycle. M2 forms a homotetrameric pH gated proton channel [225] with essential function for genome release during viral entry [164]. Furthermore, M2 has a role in viral assembly [226], aid in ESCRT-independent viral particle release from the plasma membrane [182], and stabilise host factors through direct interaction [227].

M2 is synthesized from a spliced gene, and variation in the splice site is adapted to the host, regulating the onset of M2 expression [194,224]. The M2 mRNA is spliced from the full length M1 gene of IAV segment 7. Segment 7 has been shown to produce up to 4 mRNAs depending on the viral strain (mRNA1 = M1, mRNA2 = M2, mRNA3 = not identified, mRNA4 = M42) [228]. M2 expression is essential [209], but an M2 variant with an alternative N-terminal domain M42, may be expressed by some viral strains that can rescue failure to express M2 in mice [229]. The role of mRNA3 has not been conclusively determined to date.

Structurally, M2 is divided into four domains. M2's N-terminus (residues 1-21) is highly conserved and harbours 2 cysteines (C17 and C19), which can stabilise homodimers and homotetramers [230,231]. M2's transmembrane domain (residues 22-46), which is sufficient for M2 tetramer formation and its pH gated ion channel function through H37 and W41 [232]. M2's amphipathic helix (residues 47-67) interacts with the cytosolic face of the lipid bilayer, which can induce membrane curvature depending on the cholesterol content, and contribute to viral budding [64,182]. Residue C50, residing within the amphipathic helix can be anchored to the membrane through palmitoylation [233]. M2's C-terminal domain (residues 68-97), which interacts with the virion matrix protein M1 during viral budding [226,234], and harbours a C-terminal LC3 interacting motif, which is important to subvert autophagy related processes [235].

M2 is the smallest known pH gated proton conductive pore protein and has been extensively studied, as M2 inhibition has been used as potent antiviral therapy [236]. M2 transmembrane domain and amphipathic helix (residues 18-60) are sufficient

for this function [237], however many parts of the protein in conjunction with the membrane environment contribute to the acid gating sensitivity and proton flux [236]. Hydrophilic residues within M2's single transmembrane domain facilitate the pH gating and ion channel functions. M2's transmembrane domain sequence $_{24}$DPL**T**IAA**N**II**G**IL**H**L**T**LWILDR$_{45}$, harbors several residues which may be incompatible with the apolar lipid bilayer core in a monomeric form. Threonine (T27,T39), Asparagine (N31) have polar side chains and Histidine (H37) may be polar or even partially protonated and charged at neutral pH. Glycine (G34) is not a polar group, but may induce kinks in the helix due to its limited ability to form hydrogen bonds in an alpha helix and thus expose the charged backbone [238]. Kinks increase the helix surface area and increase helix packing, which may be important for oligomerization. G34 has been proposed to allow conformational changes in the tetramer, important for proton channel function [239], and may be important to associate with the palmitoylating enzyme ZDHHC20 [233]. These transmembrane domain characteristics may not be well tolerated within the hydrophobic membrane core, and may require specialised biogenesis machinery for the membrane insertion of M2.

M2's ER membrane targeting requirements have been investigated in early *in vitro* studies. SRP targeting is necessary to insert M2 onto ER derived microsomes [240]. Beyond these findings, M2 has been found not to be affected by Sec61 inhibition [206].

Other viroporins have been investigated in more detail. The human immunodeficiency virus (HIV) encoded typeIII ion channel protein VPU, which has recently been investigated for its biogenesis requirements. VPU is similar to M2 in length, but only encodes one polar residue within its relatively hydrophobic transmembrane domain in the sequence used in this study ($_{6}$IAIAALVVAIIIAIVVW**S**IVII$_{27}$) [134]. VPU has been found to be dependent on SRPRA and EMC, but not Sec61 [134].

Some typeIII proteins have been reported to require EMC for their ER membrane insertion, however the physicochemical characteristics that warrant EMC processing remain elusive [110,134]. EMC has been shown to aid the insertion or assembly of membrane proteins with hydrophilic character such as ion channels [134,241–244], receptors [110,245,246], and transporters [247].

Furthermore, typeIII proteins have been reported to be able to access either EMC or Sec61 *in vitro* [131]. Thus, typeIII membrane proteins may be flexible in their biogenesis requirements.

1.6 Hypothesis and knowledge gap

Cells provide multiple machinery to insert the diversity of membrane proteins. We do not fully understand how membrane proteins select their preferred translocation route and the capacity of the membrane protein biogenesis network to regulate membrane protein biogenesis in response to changes in transcriptional programs.

Membrane proteins carry linear sequence motifs, encoded in stretches of hydrophobic amino acids, that signal for the recruitment of specific biogenesis machinery potentially already within the ribosomal exit tunnel [106]. Many biogenesis factors have been described, that assemble onto translating ribosomes to endow them with affinity for the ER membrane, chemically or proteolytically modify the newly emergent membrane protein sequence [79,248], resulting in yet unclear effects on decision making in membrane translocation. Once a membrane protein encounters the ER membrane, a handover reaction from the targeting machinery to the translocation machinery takes place. Depending on the signals presented, the translocation machinery is selected that suits the needs of the protein [106], and is potentially influenced by the occupancy of simultaneous handover reactions. Membrane proteins with hydrophilic residues within their transmembrane domains present a specific challenge for the translocation machinery, as the membrane they are inserting in is strongly hydrophobic and does not solvate hydrophilic amino acids [249,250]. However, membrane proteins often require hydrophilic residues within their transmembrane domains to fulfill their biological functions, for example, to form aqueous or ion conducting pores through the membrane [251]. Ion channels are an important class of membrane proteins involved in every aspect of cellular functions. Furthermore, many viruses encode ion channels, with essential functions during the infection cycle, and the capacity to contribute to pathogenicity [252].

Influenza A virus (IAV), a fast replicating respiratory pathogen, with great relevance for human health, encodes the smallest known proton selective pore protein, matrix protein 2 (M2) [225,252]. M2 is uncommon in its sequence characteristics compared to membrane proteins of the human proteome in several respects. M2 is rather short with only 97 amino acids and has a single transmembrane domain close to its N-terminus, which is composed of amino acids that occur rarely in transmembrane domains of single-pass transmembrane proteins. Little is known about the specific biogenesis requirements of M2 [225,240,252]. M2 is highly expressed during infection, yet it is incorporated with only a few copies in infectious virus particles [211]. Besides M2's essential role during viral entry as trigger for the viral genome release [164], M2 has been found to mediate the localization of other viral

components to the budding site at the plasma membrane [226], deacidify endomembrane compartments [221], which may lead to inflammasome activation [222], aid in ESCRT-independent viral particle release from the plasma membrane [182], and stabilise host factors through direct interaction [227]. M2 mRNA is spliced from the full length M gene that produces M1, and may begin synthesis as early as 2.5h post infection [253]. The ratio of M1 and M2 is tightly regulated and splicing sites are adapted to the host species and may differ between subtypes [194,224]. M2's specific expression dynamic may thus contribute to how the virus replicates within a cell and spreads from cell to cell within a host, which may ultimately influence transmissibility of a virus.

Understanding M2's specific biogenesis requirements may uncover novel regulatory mechanisms that can inform us about host-species barriers, or M2 pathogenicity, uncover new drug targets, and shine light on the capacity of the eukaryotic membrane protein biogenesis machinery to handle unconventional clients.

1.7 Objectives of this thesis

This study is conducted with the aim to identify factors that control the biogenesis of IAV M2, and to quantify the effect of the loss of these factors on IAV replication. Furthermore, we aim to identify M2's specific sequence characteristics that warrant the action of specific biogenesis factors. Lastly, we aim to measure the dynamics of M2 membrane targeting and insertion in order to assess the impact of biogenesis factors on intermediate states in the process.

1.8 References

1. Uhlén, M. *et al.* Proteomics. Tissue-based map of the human proteome. *Science* **347**, 1260419 (2015).

2. Sarmento, M. J. *et al.* The expanding organelle lipidomes: current knowledge and challenges. *Cell Mol Life Sci* **80**, 237 (2023).

3. Conroy, M. J. *et al.* LIPID MAPS: update to databases and tools for the lipidomics community. *Nucleic Acids Res* **52**, D1677–D1682 (2024).

4. Li, Y., Zhang, R. & Hei, H. Advances in post-translational modifications of proteins and cancer immunotherapy. *Front Immunol* **14**, 1229397 (2023).

5. Li, Z. *et al.* Cell-Surface RNA Associates with Heparan Sulfate and RNA-Binding Proteins to Modulate Receptor-Ligand Interactions. *bioRxiv* 2024.07.23.604836 (2024) doi:10.1101/2024.07.23.604836.

6. Harayama, T. & Riezman, H. Understanding the diversity of membrane lipid

composition. *Nature Reviews Molecular Cell Biology* **19**, 281–296 (2018).

7. Dupuy, A. D. & Engelman, D. M. Protein area occupancy at the center of the red blood cell membrane. *Proceedings of the National Academy of Sciences* **105**, 2848–2852 (2008).

8. Sakamoto, K. *et al.* Heterogeneous biological membranes regulate protein partitioning via fluctuating diffusivity. *PNAS Nexus* **2**, gad258 (2023).

9. Nunnari, J. & Walter, P. Regulation of organelle biogenesis. *Cell* **84**, 389–394 (1996).

10. Obara, C. J., Moore, A. S. & Lippincott-Schwartz, J. Structural Diversity within the Endoplasmic Reticulum-From the Microscale to the Nanoscale. *Cold Spring Harb Perspect Biol* **15**, (2023).

11. Palade, G. Intracellular aspects of the process of protein synthesis. *Science* **189**, 867 (1975).

12. Blobel, G. & Dobberstein, B. Transfer of proteins across membranes. II. Reconstitution of functional rough microsomes from heterologous components. *J Cell Biol* **67**, 852–862 (1975).

13. Spiess, M. & Lodish, H. F. An internal signal sequence: the asialoglycoprotein receptor membrane anchor. *Cell* **44**, 177–185 (1986).

14. Milstein, C., Brownlee, G. G., Harrison, T. M. & Mathews, M. B. A possible precursor of immunoglobulin light chains. *Nat New Biol* **239**, 117–120 (1972).

15. Liaci, A. M. *et al.* Structure of the human signal peptidase complex reveals the determinants for signal peptide cleavage. *Mol Cell* **81**, 3934–3948.e11 (2021).

16. Von Heijne, G. The membrane protein universe: what's out there and why bother? *Journal of Internal Medicine* **261**, 543–557 (2007).

17. Wang, Y.-N., Yamaguchi, H., Hsu, J.-M. & Hung, M.-C. Nuclear trafficking of the epidermal growth factor receptor family membrane proteins. *Oncogene* **29**, 3997–4006 (2010).

18. Alberts, B. *Molecular Biology of the Cell*. (Garland Science, 2017).

19. Cockcroft, S. Mammalian lipids: structure, synthesis and function. *Essays Biochem* **65**, 813–845 (2021).

20. Levental, I. & Lyman, E. Regulation of membrane protein structure and function by their lipid nano-environment. *Nat Rev Mol Cell Biol* **24**, 107–122 (2023).

21. Radanović, T., Reinhard, J., Ballweg, S., Pesek, K. & Ernst, R. An Emerging Group of Membrane Property Sensors Controls the Physical State of Organellar Membranes to Maintain Their Identity. *BioEssays* **40**, 1700250 (2018).

22. Bretscher, M. S. & Munro, S. Cholesterol and the Golgi apparatus. *Science* **261**, 1280–1281 (1993).

23. Park, S. H. & Opella, S. J. Tilt angle of a trans-membrane helix is determined by hydrophobic mismatch. *J Mol Biol* **350**, 310–318 (2005).

24. Liu, L., Doray, B. & Kornfeld, S. Recycling of Golgi glycosyltransferases requires direct binding to coatomer. *Proc Natl Acad Sci U S A* **115**, 8984–8989 (2018).

25. Klausner, R. D., Donaldson, J. G. & Lippincott-Schwartz, J. Brefeldin A: insights into the control of membrane traffic and organelle structure. *J Cell Biol* **116**, 1071–1080 (1992).

26. Anding, A. L. & Baehrecke, E. H. Cleaning House: Selective Autophagy of Organelles. *Dev Cell* **41**, 10–22 (2017).

27. Wilkinson, K. D. The discovery of ubiquitin-dependent proteolysis. *Proc Natl Acad Sci U S A* **102**, 15280–15282 (2005).

28. Zhu, H. *et al.* Ufbp1 promotes plasma cell development and ER expansion by modulating distinct branches of UPR. *Nat Commun* **10**, 1084 (2019).

29. Integrated Functions of Membrane Property Sensors and a Hidden Side of the Unfolded Protein Response. *Molecular Cell* **71**, 458–467 (2018).

30. Moncan, M. *et al.* Regulation of lipid metabolism by the unfolded protein response. *J Cell Mol Med* **25**, 1359–1370 (2021).

31. Hetz, C., Zhang, K. & Kaufman, R. J. Mechanisms, regulation and functions of the unfolded protein response. *Nat Rev Mol Cell Biol* **21**, 421–438 (2020).

32. Shaheen, A. Effect of the unfolded protein response on ER protein export: a potential new mechanism to relieve ER stress. *Cell Stress Chaperones* **23**, 797–806 (2018).

33. Wek, R. C., Jiang, H.-Y. & Anthony, T. G. Coping with stress: eIF2 kinases and translational control. *Biochem. Soc. Trans.* **34**, 7–11 (2006).

34. Ng, D. T., Spear, E. D. & Walter, P. The unfolded protein response regulates multiple aspects of secretory and membrane protein biogenesis and endoplasmic reticulum quality control. *J Cell Biol* **150**, 77–88 (2000).

35. Di Conza, G., Ho, P.-C., Cubillos-Ruiz, J. R. & Huang, S. C.-C. Control of immune cell function by the unfolded protein response. *Nat Rev Immunol* **23**, 546–562 (2023).

36. Shpilka, T. & Haynes, C. M. The mitochondrial UPR: mechanisms, physiological functions and implications in ageing. *Nat Rev Mol Cell Biol* **19**, 109–120 (2018).

37. Anfinsen, C. B. Principles that govern the folding of protein chains. *Science* **181**, 223–230 (1973).

38. Horwich, A. L. & Fenton, W. A. Chaperonin-assisted protein folding: a chronologue. *Q Rev Biophys* **53**, e4 (2020).

39. Ramírez, A. S. & Locher, K. P. Structural and mechanistic studies of the

N-glycosylation machinery: from lipid-linked oligosaccharide biosynthesis to glycan transfer. *Glycobiology* **33**, 861–872 (2023).

40. Adams, B. M., Canniff, N. P., Guay, K. P. & Hebert, D. N. The Role of Endoplasmic Reticulum Chaperones in Protein Folding and Quality Control. *Prog Mol Subcell Biol* **59**, 27–50 (2021).

41. Wiseman, R. L., Mesgarzadeh, J. S. & Hendershot, L. M. Reshaping endoplasmic reticulum quality control through the unfolded protein response. *Mol Cell* **82**, 1477–1491 (2022).

42. Modenutti, C. P. *et al.* Clamping, bending, and twisting inter-domain motions in the misfold-recognizing portion of UDP-glucose: Glycoprotein glucosyltransferase. *Structure* **29**, 357–370.e9 (2021).

43. Belliveau, N. M. *et al.* Galvanin is an electric-field sensor for directed cell migration. *bioRxiv* (2024) doi:10.1101/2024.09.23.614580.

44. Pinho, S. S., Alves, I., Gaifem, J. & Rabinovich, G. A. Immune regulatory networks coordinated by glycans and glycan-binding proteins in autoimmunity and infection. *Cell Mol Immunol* **20**, 1101–1113 (2023).

45. Ma, M. *et al.* Regulated N-glycosylation controls chaperone function and receptor trafficking. *Science* (2024) doi:10.1126/science.adp7201.

46. Chitwood, P. J. & Hegde, R. S. An intramembrane chaperone complex facilitates membrane protein biogenesis. *Nature* **584**, 630–634 (2020).

47. Chen, Z. *et al.* Structure of an EMC chaperone: CaV complex reveals an ion channel assembly intermediate and a holdase function for the EMC. *Biophys. J.* **123**, 14a (2024).

48. Christianson, J. C. & Carvalho, P. Order through destruction: how ER-associated protein degradation contributes to organelle homeostasis. *EMBO J* **41**, e109845 (2022).

49. Feige, M. J. & Hendershot, L. M. Quality control of integral membrane proteins by assembly-dependent membrane integration. *Mol Cell* **51**, 297–309 (2013).

50. Grotzke, J. E., Lu, Q. & Cresswell, P. Deglycosylation-dependent fluorescent proteins provide unique tools for the study of ER-associated degradation. *Proc Natl Acad Sci U S A* **110**, 3393–3398 (2013).

51. Kennedy, A., Ren, H. Y., Madden, V. J. & Cyr, D. M. Lysosome docking to WIPI1 rings and ER-connected phagophores occurs during DNAJB12- and GABARAP-dependent selective autophagy of misfolded P23H-rhodopsin. *Mol Biol Cell* **33**, ar84 (2022).

52. Yang, M. *et al.* ER-Phagy: A New Regulator of ER Homeostasis. *Front Cell Dev Biol* **9**, 684526 (2021).

53. Chino, H., Hatta, T., Natsume, T. & Mizushima, N. Intrinsically Disordered Protein

TEX264 Mediates ER-phagy. *Mol Cell* **74**, 909–921.e6 (2019).

54. Fumagalli, F. *et al.* Translocon component Sec62 acts in endoplasmic reticulum turnover during stress recovery. *Nat Cell Biol* **18**, 1173–1184 (2016).

55. Ishimura, R. *et al.* The UFM1 system regulates ER-phagy through the ufmylation of CYB5R3. *Nat Commun* **13**, 7857 (2022).

56. Kozutsumi, Y., Segal, M., Normington, K., Gething, M. J. & Sambrook, J. The presence of malfolded proteins in the endoplasmic reticulum signals the induction of glucose-regulated proteins. *Nature* **332**, 462–464 (1988).

57. Gardner, B. M. & Walter, P. Unfolded proteins are Ire1-activating ligands that directly induce the unfolded protein response. *Science* **333**, 1891–1894 (2011).

58. Activation of the Unfolded Protein Response by Lipid Bilayer Stress. *Molecular Cell* **67**, 673–684.e8 (2017).

59. Luo, S., Baumeister, P., Yang, S., Abcouwer, S. F. & Lee, A. S. Induction of Grp78/BiP by translational block: activation of the Grp78 promoter by ATF4 through and upstream ATF/CRE site independent of the endoplasmic reticulum stress elements. *J Biol Chem* **278**, 37375–37385 (2003).

60. Shen, J., Snapp, E. L., Lippincott-Schwartz, J. & Prywes, R. Stable binding of ATF6 to BiP in the endoplasmic reticulum stress response. *Mol Cell Biol* **25**, 921–932 (2005).

61. Adachi, Y. *et al.* ATF6 is a transcription factor specializing in the regulation of quality control proteins in the endoplasmic reticulum. *Cell Struct Funct* **33**, 75–89 (2008).

62. Racek, T. *et al.* Transcriptional repression of the prosurvival endoplasmic reticulum chaperone GRP78/BIP by E2F1. *J Biol Chem* **283**, 34305–34314 (2008).

63. Baumeister, P. *et al.* Endoplasmic reticulum stress induction of the Grp78/BiP promoter: activating mechanisms mediated by YY1 and its interactive chromatin modifiers. *Mol Cell Biol* **25**, 4529–4540 (2005).

64. Martyna, A. *et al.* Membrane remodeling by the M2 amphipathic helix drives influenza virus membrane scission. *Sci Rep* **7**, 44695 (2017).

65. Yoshida, H. *et al.* A time-dependent phase shift in the mammalian unfolded protein response. *Dev Cell* **4**, 265–271 (2003).

66. Plumb, R., Zhang, Z.-R., Appathurai, S. & Mariappan, M. A functional link between the co-translational protein translocation pathway and the UPR. *Elife* **4**, (2015).

67. Hollien, J. & Weissman, J. S. Decay of endoplasmic reticulum-localized mRNAs during the unfolded protein response. *Science* **313**, 104–107 (2006).

68. Tran, N.-H. *et al.* The stress-sensing domain of activated IRE1α forms helical filaments in narrow ER membrane tubes. *Science* **374**, 52–57 (2021).

69. Cretenet, G., Le Clech, M. & Gachon, F. Circadian clock-coordinated 12 Hr period rhythmic activation of the IRE1alpha pathway controls lipid metabolism in mouse liver. *Cell Metab* **11**, 47–57 (2010).

70. Ajoolabady, A., Lindholm, D., Ren, J. & Pratico, D. ER stress and UPR in Alzheimer's disease: mechanisms, pathogenesis, treatments. *Cell Death Dis* **13**, 706 (2022).

71. Reich, S. *et al.* A multi-omics analysis reveals the unfolded protein response regulon and stress-induced resistance to folate-based antimetabolites. *Nat Commun* **11**, 2936 (2020).

72. Zhou, T. *et al.* Occludin protects secretory cells from ER stress by facilitating SNARE-dependent apical protein exocytosis. *Proc Natl Acad Sci U S A* **117**, 4758–4769 (2020).

73. Vonk, W. I. M. *et al.* Differentiation Drives Widespread Rewiring of the Neural Stem Cell Chaperone Network. *Mol Cell* **78**, 329–345.e9 (2020).

74. Wang, L. *et al.* UFMylation of RPL26 links translocation-associated quality control to endoplasmic reticulum protein homeostasis. *Cell Res* **30**, 5–20 (2020).

75. Gonzalez-Garcia, M., Fusco, G. & De Simone, A. Membrane Interactions and Toxicity by Misfolded Protein Oligomers. *Front Cell Dev Biol* **9**, 642623 (2021).

76. Hsieh, H.-H. & Shan, S.-O. Fidelity of Cotranslational Protein Targeting to the Endoplasmic Reticulum. *Int J Mol Sci* **23**, (2021).

77. Higy, M., Junne, T. & Spiess, M. Topogenesis of membrane proteins at the endoplasmic reticulum. *Biochemistry* **43**, 12716–12722 (2004).

78. Mercier, E., Wintermeyer, W. & Rodnina, M. V. Co-translational insertion and topogenesis of bacterial membrane proteins monitored in real time. *EMBO J* **39**, e104054 (2020).

79. Lentzsch, A. M. *et al.* NAC guides a ribosomal multienzyme complex for nascent protein processing. *Nature* **633**, 718–724 (2024).

80. Raue, U., Oellerer, S. & Rospert, S. Association of protein biogenesis factors at the yeast ribosomal tunnel exit is affected by the translational status and nascent polypeptide sequence. *J Biol Chem* **282**, 7809–7816 (2007).

81. Jomaa, A. *et al.* Mechanism of signal sequence handover from NAC to SRP on ribosomes during ER-protein targeting. *Science* **375**, 839–844 (2022).

82. Walter, P. & Blobel, G. Signal recognition particle contains a 7S RNA essential for protein translocation across the endoplasmic reticulum. *Nature* **299**, 691–698 (1982).

83. Lee, J. H. *et al.* Receptor compaction and GTPase rearrangement drive SRP-mediated cotranslational protein translocation into the ER. *Sci Adv* **7**, (2021).

84. Wang, H. & Hegde, R. S. Identification of a factor that accelerates substrate release from the signal recognition particle. *Science* **386**, 996–1003 (2024).

85. Chartron, J. W., Hunt, K. C. L. & Frydman, J. Cotranslational signal-independent SRP preloading during membrane targeting. *Nature* **536**, 224–228 (2016).

86. Pechmann, S., Chartron, J. W. & Frydman, J. Local slowdown of translation by nonoptimal codons promotes nascent-chain recognition by SRP in vivo. *Nat Struct Mol Biol* **21**, 1100–1105 (2014).

87. Lu, J. & Deutsch, C. Electrostatics in the ribosomal tunnel modulate chain elongation rates. *J Mol Biol* **384**, 73–86 (2008).

88. Woolstenhulme, C. J., Guydosh, N. R., Green, R. & Buskirk, A. R. High-precision analysis of translational pausing by ribosome profiling in bacteria lacking EFP. *Cell Rep* **11**, 13–21 (2015).

89. Lakkaraju, A. K. K., Mary, C., Scherrer, A., Johnson, A. E. & Strub, K. SRP keeps polypeptides translocation-competent by slowing translation to match limiting ER-targeting sites. *Cell* **133**, 440–451 (2008).

90. Goder, V., Crottet, P. & Spiess, M. In vivo kinetics of protein targeting to the endoplasmic reticulum determined by site-specific phosphorylation. *EMBO J* **19**, 6704–6712 (2000).

91. Siegel, V. & Walter, P. The affinity of signal recognition particle for presecretory proteins is dependent on nascent chain length. *EMBO J* **7**, 1769–1775 (1988).

92. Guna, A. & Hegde, R. S. Transmembrane Domain Recognition during Membrane Protein Biogenesis and Quality Control. *Curr Biol* **28**, R498–R511 (2018).

93. Shan, S.-O. Role of Hsp70 in Post-Translational Protein Targeting: Tail-Anchored Membrane Proteins and Beyond. *International Journal of Molecular Sciences* **24**, 1170 (2023).

94. Ast, T., Cohen, G. and Schuldiner, M. A Network of Cytosolic Factors Targets SRP-Independent Proteins to the Endoplasmic Reticulum. *Cell* **152**, 1134–1145 (2013).

95. Haßdenteufel, S. *et al.* hSnd2 protein represents an alternative targeting factor to the endoplasmic reticulum in human cells. *FEBS Lett* **591**, 3211–3224 (2017).

96. Mariappan, M. *et al.* A ribosome-associating factor chaperones tail-anchored membrane proteins. *Nature* **466**, 1120–1124 (2010).

97. Ziegelhoffer, T. *et al.* NAC and Zuotin/Hsp70 chaperone systems coexist at the ribosome tunnel exit in vivo. *Nucleic Acids Res* **52**, 3346–3357 (2024).

98. Rabu, C., Wipf, P., Brodsky, J. L. & High, S. A precursor-specific role for Hsp40/Hsc70 during tail-anchored protein integration at the endoplasmic reticulum. *J Biol Chem* **283**, 27504–27513 (2008).

99. Hegde, R. S. The Function, Structure, and Origins of the ER Membrane Protein Complex. *Annu Rev Biochem* **91**, 651–678 (2022).

100. Aviram, N. *et al.* The SND proteins constitute an alternative targeting route to the endoplasmic reticulum. *Nature* **540**, 134–138 (2016).

101. Song, J. & Becker, T. Fidelity of organellar protein targeting. *Curr Opin Cell Biol* **75**, 102071 (2022).

102. Cui, X. A. *et al.* mRNA encoding Sec61β, a tail-anchored protein, is localized on the endoplasmic reticulum. *J Cell Sci* **128**, 3398–3410 (2015).

103. Ma, W. & Mayr, C. A Membraneless Organelle Associated with the Endoplasmic Reticulum Enables 3'UTR-Mediated Protein-Protein Interactions. *Cell* **175**, 1492–1506.e19 (2018).

104. Lewis, A. J. O. & Hegde, R. S. A unified evolutionary origin for the ubiquitous protein transporters SecY and YidC. *BMC Biol* **19**, 266 (2021).

105. Chen, Y. & Dalbey, R. E. Oxa1 Superfamily: New Members Found in the ER. *Trends Biochem Sci* **43**, 151–153 (2018).

106. Hegde, R. S. & Keenan, R. J. A unifying model for membrane protein biogenesis. *Nat Struct Mol Biol* **31**, 1009–1017 (2024).

107. Anghel, S. A., McGilvray, P. T., Hegde, R. S. & Keenan, R. J. Identification of Oxa1 Homologs Operating in the Eukaryotic Endoplasmic Reticulum. *Cell Rep* **21**, 3708–3716 (2017).

108. Güngör, B., Flohr, T., Garg, S. G. & Herrmann, J. M. The ER membrane complex (EMC) can functionally replace the Oxa1 insertase in mitochondria. *PLoS Biol* **20**, e3001380 (2022).

109. Guna, A., Volkmar, N., Christianson, J. C. & Hegde, R. S. The ER membrane protein complex is a transmembrane domain insertase. *Science* **359**, 470–473 (2018).

110. Chitwood, P. J., Juszkiewicz, S., Guna, A., Shao, S. & Hegde, R. S. EMC Is Required to Initiate Accurate Membrane Protein Topogenesis. *Cell* **175**, 1507–1519.e16 (2018).

111. Cavalié, A. & Zimmermann, R. Editorial: The evolving picture of Ca2+ leak from endoplasmic reticulum in health and diseases. *Front. Physiol.* **14**, 1182455 (2023).

112. Kim, S. J., Mitra, D., Salerno, J. R. & Hegde, R. S. Signal sequences control gating of the protein translocation channel in a substrate-specific manner. *Dev Cell* **2**, 207–217 (2002).

113. Fessl, T. *et al.* Dynamic action of the Sec machinery during initiation, protein translocation and termination. *Elife* **7**, (2018).

114. Laskowski, P. R. *et al.* Monitoring the binding and insertion of a single transmembrane protein by an insertase. *Nat Commun* **12**, 7082 (2021).

115. Gemmer, M., Chaillet, M. L. & Förster, F. Exploring the molecular composition of the multipass translocon in its native membrane environment. *Life Sci Alliance* **7**, (2024).

116. Gemmer, M. & Förster, F. A clearer picture of the ER translocon complex. *J Cell Sci* **133**, (2020).

117. Ng, D. T., Brown, J. D. & Walter, P. Signal sequences specify the targeting route to the endoplasmic reticulum membrane. *J Cell Biol* **134**, 269–278 (1996).

118. Weng, T.-H. *et al.* Architecture of the active post-translational Sec translocon. *EMBO J* **40**, e105643 (2021).

119. Jung, S.-J. *et al.* Expression level of Sec62 modulates membrane insertion of marginally hydrophobic segments. *Biochim Biophys Acta Biomembr* **1864**, 184051 (2022).

120. Lakkaraju, A. K. K. *et al.* Efficient secretion of small proteins in mammalian cells relies on Sec62-dependent posttranslational translocation. *Mol Biol Cell* **23**, 2712–2722 (2012).

121. Itskanov, S., Kuo, K. M., Gumbart, J. C. & Park, E. Stepwise gating of the Sec61 protein-conducting channel by Sec63 and Sec62. *Nat Struct Mol Biol* **28**, 162–172 (2021).

122. Jung, S.-J., Kim, J. E. H., Junne, T., Spiess, M. & Kim, H. Cotranslational Targeting and Posttranslational Translocation can Cooperate in Spc3 Topogenesis. *J Mol Biol* **433**, 167109 (2021).

123. Page, K. R. *et al.* Role of a holo-insertase complex in the biogenesis of biophysically diverse ER membrane proteins. *Mol Cell* **84**, 3302–3319.e11 (2024).

124. Rapoport, T. A., Jungnickel, B. & Kutay, U. Protein transport across the eukaryotic endoplasmic reticulum and bacterial inner membranes. *Annu Rev Biochem* **65**, 271–303 (1996).

125. Tsukazaki, T. Structure-based working model of SecDF, a proton-driven bacterial protein translocation factor. *FEMS Microbiol Lett* **365**, (2018).

126. Nishiyama, K.-I. *et al.* MPIase is a glycolipozyme essential for membrane protein integration. *Nat Commun* **3**, 1260 (2012).

127. Nomura, K. *et al.* Role of a bacterial glycolipid in Sec-independent membrane protein insertion. *Sci Rep* **12**, 12231 (2022).

128. Savitz, A. J. & Meyer, D. I. Identification of a ribosome receptor in the rough endoplasmic reticulum. *Nature* **346**, 540–544 (1990).

129. Wanker, E. E., Sun, Y., Savitz, A. J. & Meyer, D. I. Functional characterization of the 180-kD ribosome receptor in vivo. *J Cell Biol* **130**, 29–39 (1995).

130. Arad, S. *et al.* A new targeting motif for endoplasmic reticulum surface proteins. *bioRxiv* (2024) doi:10.1101/2024.04.22.590521.

131. Wu, H. & Hegde, R. S. Mechanism of signal-anchor triage during early steps of membrane protein insertion. *Mol Cell* **83**, 961–973.e7 (2023).

132. Nicchitta, C. V., Murphy, E. C., 3rd, Haynes, R. & Shelness, G. S. Stage- and ribosome-specific alterations in nascent chain-Sec61p interactions accompany translocation across the ER membrane. *J Cell Biol* **129**, 957–970 (1995).

133. High, S. *et al.* Sec61p is adjacent to nascent type I and type II signal-anchor proteins during their membrane insertion. *J Cell Biol* **121**, 743–750 (1993).

134. O'Keefe, S. *et al.* An alternative pathway for membrane protein biogenesis at the endoplasmic reticulum. *Commun Biol* **4**, 828 (2021).

135. Xie, K. *et al.* Features of transmembrane segments that promote the lateral release from the translocase into the lipid phase. *Biochemistry* **46**, 15153–15161 (2007).

136. Gonsberg, A. *et al.* The Sec61/SecY complex is inherently deficient in translocating intrinsically disordered proteins. *J Biol Chem* **292**, 21383–21396 (2017).

137. Jung, S. & Tatzelt, J. Impaired transport of intrinsically disordered proteins through the Sec61 and SecY translocon; implications for prion diseases. *Prion* **12**, 88–92 (2018).

138. Miesbauer, M. *et al.* alpha-Helical domains promote translocation of intrinsically disordered polypeptides into the endoplasmic reticulum. *J Biol Chem* **284**, 24384–24393 (2009).

139. Conti, B. J., Elferich, J., Yang, Z., Shinde, U. & Skach, W. R. Cotranslational folding inhibits translocation from within the ribosome-Sec61 translocon complex. *Nat Struct Mol Biol* **21**, 228–235 (2014).

140. Pleiner, T. *et al.* A selectivity filter in the ER membrane protein complex limits protein misinsertion at the ER. *J Cell Biol* **222**, (2023).

141. McDowell, M. A. *et al.* Structural Basis of Tail-Anchored Membrane Protein Biogenesis by the GET Insertase Complex. *Mol Cell* **80**, 72–86.e7 (2020).

142. Ngosuwan, J., Wang, N. M., Fung, K. L. & Chirico, W. J. Roles of cytosolic Hsp70 and Hsp40 molecular chaperones in post-translational translocation of presecretory proteins into the endoplasmic reticulum. *J Biol Chem* **278**, 7034–7042 (2003).

143. Hansen, K. G. *et al.* An ER surface retrieval pathway safeguards the import of mitochondrial membrane proteins in yeast. *Science* **361**, 1118–1122 (2018).

144. Koch, C. *et al.* The ER-SURF pathway uses ER-mitochondria contact sites for protein targeting to mitochondria. *EMBO Rep* **25**, 2071–2096 (2024).

145. McKenna, M. J., Adams, B. M., Chu, V., Paulo, J. A. & Shao, S. ATP13A1 prevents ERAD of folding-competent mislocalized and misoriented proteins. *Mol Cell* **82**, 4277–4289.e10 (2022).

146. Wu, H., Smalinskaitė, L. & Hegde, R. S. EMC rectifies the topology of multipass membrane proteins. *Nat Struct Mol Biol* **31**, 32–41 (2024).

147. Hegde, R. S. & Bernstein, H. D. The surprising complexity of signal sequences. *Trends Biochem Sci* **31**, 563–571 (2006).

148. Dempster, J. M. *et al.* Chronos: a cell population dynamics model of CRISPR experiments that improves inference of gene fitness effects. *Genome Biol* **22**, 343 (2021).

149. Guna, A. *et al.* A dual sgRNA library design to probe genetic modifiers using genome-wide CRISPRi screens. *BMC Genomics* **24**, 651 (2023).

150. Pulit-Penaloza, J. A. *et al.* Kinetics and magnitude of viral RNA shedding as indicators for Influenza A virus transmissibility in ferrets. *Commun Biol* **6**, 90 (2023).

151. Kieran, T. J., Sun, X., Maines, T. R. & Belser, J. A. Optimal thresholds and key parameters for predicting influenza A virus transmission events in ferrets. *Npj Viruses* **2**, 64 (2024).

152. Wang, C. C. *et al.* Airborne transmission of respiratory viruses. *Science* **373**, (2021).

153. Weber, T. P. & Stilianakis, N. I. Inactivation of influenza A viruses in the environment and modes of transmission: a critical review. *J Infect* **57**, 361–373 (2008).

154. Burke, C. W. *et al.* Illumination of parainfluenza virus infection and transmission in living animals reveals a tissue-specific dichotomy. *PLoS Pathog* **7**, e1002134 (2011).

155. Escaffre, O. *et al.* Henipavirus pathogenesis in human respiratory epithelial cells. *J Virol* **87**, 3284–3294 (2013).

156. Iyaniwura, S. A. *et al.* The kinetics of SARS-CoV-2 infection based on a human challenge study. *Proc Natl Acad Sci U S A* **121**, e2406303121 (2024).

157. Puhach, O., Meyer, B. & Eckerle, I. SARS-CoV-2 viral load and shedding kinetics. *Nat Rev Microbiol* **21**, 147–161 (2023).

158. Canini, L., Woolhouse, M. E. J., Maines, T. R. & Carrat, F. Heterogeneous

shedding of influenza by human subjects and its implications for epidemiology and control. *Sci Rep* **6**, 38749 (2016).

159. Carter, T. & Iqbal, M. The Influenza A Virus Replication Cycle: A Comprehensive Review. *Viruses* **16**, (2024).

160. Shinya, K. *et al.* Avian flu: influenza virus receptors in the human airway. *Nature* **440**, 435–436 (2006).

161. Matlin, K. S., Reggio, H., Helenius, A. & Simons, K. Infectious entry pathway of influenza virus in a canine kidney cell line. *J Cell Biol* **91**, 601–613 (1981).

162. Lamb, R. A. & Jardetzky, T. S. Structural basis of viral invasion: lessons from paramyxovirus F. *Curr Opin Struct Biol* **17**, 427–436 (2007).

163. White, J., Matlin, K. & Helenius, A. Cell fusion by Semliki Forest, influenza, and vesicular stomatitis viruses. *J Cell Biol* **89**, 674–679 (1981).

164. Takeda, M., Pekosz, A., Shuck, K., Pinto, L. H. & Lamb, R. A. Influenza a virus M2 ion channel activity is essential for efficient replication in tissue culture. *J Virol* **76**, 1391–1399 (2002).

165. Bui, M., Whittaker, G. & Helenius, A. Effect of M1 protein and low pH on nuclear transport of influenza virus ribonucleoproteins. *J Virol* **70**, 8391–8401 (1996).

166. Banerjee, I. *et al.* Influenza A virus uses the aggresome processing machinery for host cell entry. *Science* **346**, 473–477 (2014).

167. Hudson, J. B., Flawith, J. & Dimmock, N. J. Early events in influenza virus infection. III. The formation of a nucleoplasmic ribonucleoprotein complex from the input virion. *Virology* **86**, 167–176 (1978).

168. Eisfeld, A. J., Neumann, G. & Kawaoka, Y. At the centre: influenza A virus ribonucleoproteins. *Nat Rev Microbiol* **13**, 28–41 (2015).

169. Sempere Borau, M. & Stertz, S. Entry of influenza A virus into host cells - recent progress and remaining challenges. *Curr Opin Virol* **48**, 23–29 (2021).

170. Palese, P. & Schulman, J. L. Differences in RNA patterns of influenza A viruses. *J Virol* **17**, 876–884 (1976).

171. Kingsbury, D. W. & Webster, R. G. Cell-free translation of influenza virus messenger RNA. *Virology* **56**, 654–657 (1973).

172. Bishop, D. H. L., Roy, P., Bean, W. J., Jr & Simpson, R. W. Transcription of the influenza ribonucleic acid genome by a virion polymerase III. Completeness of the transcription process. *J. Virol.* **10**, 689–697 (1972).

173. Plotch, S. J., Bouloy, M., Ulmanen, I. & Krug, R. M. A unique cap(m7GpppXm)-dependent influenza virion endonuclease cleaves capped RNAs to generate the primers that initiate viral RNA transcription. *Cell* **23**, 847–858 (1981).

174. Robertson, J. S., Schubert, M. & Lazzarini, R. A. Polyadenylation sites for influenza virus mRNA. *J Virol* **38**, 157–163 (1981).

175. Zheng, H., Lee, H. A., Palese, P. & García-Sastre, A. Influenza A virus RNA polymerase has the ability to stutter at the polyadenylation site of a viral RNA template during RNA replication. *J Virol* **73**, 5240–5243 (1999).

176. Barrett, T., Wolstenholme, A. J. & Mahy, B. W. Transcription and replication of influenza virus RNA. *Virology* **98**, 211–225 (1979).

177. Hsu, M. T., Parvin, J. D., Gupta, S., Krystal, M. & Palese, P. Genomic RNAs of influenza viruses are held in a circular conformation in virions and in infected cells by a terminal panhandle. *Proc Natl Acad Sci U S A* **84**, 8140–8144 (1987).

178. Compans, R. W., Content, J. & Duesberg, P. H. Structure of the ribonucleoprotein of influenza virus. *J Virol* **10**, 795–800 (1972).

179. Murphy, J. S. & Bang, F. B. Observations with the electron microscope on cells of the chick chorio-allantoic membrane infected with influenza virus. *J Exp Med* **95**, 259–268 (1952).

180. Takeda, M., Leser, G. P., Russell, C. J. & Lamb, R. A. Influenza virus hemagglutinin concentrates in lipid raft microdomains for efficient viral fusion. *Proc Natl Acad Sci U S A* **100**, 14610–14617 (2003).

181. Schmitt, A. P. & Lamb, R. A. Influenza virus assembly and budding at the viral budozone. *Adv Virus Res* **64**, 383–416 (2005).

182. Rossman, J. S., Jing, X., Leser, G. P. & Lamb, R. A. Influenza virus M2 protein mediates ESCRT-independent membrane scission. *Cell* **142**, 902–913 (2010).

183. Levene, R. E. & Gaglia, M. M. Host Shutoff in Influenza A Virus: Many Means to an End. *Viruses* **10**, (2018).

184. Bercovich-Kinori, A. *et al.* A systematic view on influenza induced host shutoff. *Elife* **5**, (2016).

185. Chan, A. Y., Frank, V. T., Matt, S., Othmar, E. G. & Ervin, F. Influenza virus inhibits RNA polymerase II elongation. *Virology* **351**, 210–217 (2006).

186. Bauer, D. L. V. *et al.* Influenza Virus Mounts a Two-Pronged Attack on Host RNA Polymerase II Transcription. *Cell Rep* **23**, 2119–2129.e3 (2018).

187. Lamb, R. A. & Choppin, P. W. Synthesis of influenza virus polypeptides in cells resistant to alpha-amanitin: evidence for the involvement of cellular RNA polymerase II in virus replication. *J Virol* **23**, 816–819 (1977).

188. Shapiro, G. I., Gurney, T., Jr & Krug, R. M. Influenza virus gene expression: control mechanisms at early and late times of infection and nuclear-cytoplasmic transport of virus-specific RNAs. *J Virol* **61**, 764–773 (1987).

189. Phan, T. *et al.* Segment-specific kinetics of mRNA, cRNA and vRNA

accumulation during influenza infection. *J Virol* **95**, (2021).

190. Khaperskyy, D. A., Schmaling, S., Larkins-Ford, J., McCormick, C. & Gaglia, M. M. Selective Degradation of Host RNA Polymerase II Transcripts by Influenza A Virus PA-X Host Shutoff Protein. *PLoS Pathog* **12**, e1005427 (2016).

191. García-Sastre, A. *et al.* Influenza A virus lacking the NS1 gene replicates in interferon-deficient systems. *Virology* **252**, 324–330 (1998).

192. Nemeroff, M. E., Barabino, S. M., Li, Y., Keller, W. & Krug, R. M. Influenza virus NS1 protein interacts with the cellular 30 kDa subunit of CPSF and inhibits 3'end formation of cellular pre-mRNAs. *Mol Cell* **1**, 991–1000 (1998).

193. Krug, R. M. Functions of the influenza A virus NS1 protein in antiviral defense. *Curr Opin Virol* **12**, 1–6 (2015).

194. Bogdanow, B. *et al.* The dynamic proteome of influenza A virus infection identifies M segment splicing as a host range determinant. *Nat Commun* **10**, 5518 (2019).

195. Khaperskyy, D. A. *et al.* Influenza a virus host shutoff disables antiviral stress-induced translation arrest. *PLoS Pathog* **10**, e1004217 (2014).

196. Becker, A. C. *et al.* Influenza A Virus Induces Autophagosomal Targeting of Ribosomal Proteins. *Mol Cell Proteomics* **17**, 1909–1921 (2018).

197. Lu, Y., Wambach, M., Katze, M. G. & Krug, R. M. Binding of the influenza virus NS1 protein to double-stranded RNA inhibits the activation of the protein kinase that phosphorylates the eIF-2 translation initiation factor. *Virology* **214**, 222–228 (1995).

198. Ren, L. *et al.* Influenza A Virus (H1N1) Infection Induces Glycolysis to Facilitate Viral Replication. *Virol Sin* **36**, 1532–1542 (2021).

199. Smallwood, H. S. *et al.* Targeting Metabolic Reprogramming by Influenza Infection for Therapeutic Intervention. *Cell Rep* **19**, 1640–1653 (2017).

200. Zhang, Y. *et al.* Influenza A virus-induced glycolysis facilitates virus replication by activating ROS/HIF-1α pathway. *Free Radic Biol Med* **225**, 910–924 (2024).

201. Slaine, P. D. *et al.* Thiopurines activate an antiviral unfolded protein response that blocks influenza A virus glycoprotein accumulation. *J Virol* **95**, (2021).

202. Al-Beltagi, S. *et al.* Thapsigargin Is a Broad-Spectrum Inhibitor of Major Human Respiratory Viruses: Coronavirus, Respiratory Syncytial Virus and Influenza A Virus. *Viruses* **13**, (2021).

203. Dou, D., da Silva, D. V., Nordholm, J., Wang, H. & Daniels, R. Type II transmembrane domain hydrophobicity dictates the cotranslational dependence for inversion. *Mol Biol Cell* **25**, 3363–3374 (2014).

204. Wang, N., Glidden, E. J., Murphy, S. R., Pearse, B. R. & Hebert, D. N. The

cotranslational maturation program for the type II membrane glycoprotein influenza neuraminidase. *J Biol Chem* **283**, 33826–33837 (2008).

205. Hebert, D. N., Foellmer, B. & Helenius, A. Calnexin and calreticulin promote folding, delay oligomerization and suppress degradation of influenza hemagglutinin in microsomes. *EMBO J.* **15**, 2961–2968 (1996).

206. Heaton, N. S. *et al.* Targeting Viral Proteostasis Limits Influenza Virus, HIV, and Dengue Virus Infection. *Immunity* **44**, 46–58 (2016).

207. Chen, W., Helenius, J., Braakman, I. & Helenius, A. Cotranslational folding and calnexin binding during glycoprotein synthesis. *Proc Natl Acad Sci U S A* **92**, 6229–6233 (1995).

208. McMahon, A. *et al.* High-throughput super-resolution analysis of influenza virus pleomorphism reveals insights into viral spatial organization. *PLoS Pathog* **19**, e1011484 (2023).

209. Chen, B. J., Leser, G. P., Jackson, D. & Lamb, R. A. The influenza virus M2 protein cytoplasmic tail interacts with the M1 protein and influences virus assembly at the site of virus budding. *J Virol* **82**, 10059–10070 (2008).

210. Inglis, S. C., Carroll, A. R., Lamb, R. A. & Mahy, B. W. Polypeptides specified by the influenza virus genome I. Evidence for eight distinct gene products specified by fowl plague virus. *Virology* **74**, 489–503 (1976).

211. Zebedee, S. L. & Lamb, R. A. Influenza A virus M2 protein: monoclonal antibody restriction of virus growth and detection of M2 in virions. *J Virol* **62**, 2762–2772 (1988).

212. Frabutt, D. A., Wang, B., Riaz, S., Schwartz, R. C. & Zheng, Y.-H. Innate Sensing of Influenza A Virus Hemagglutinin Glycoproteins by the Host Endoplasmic Reticulum (ER) Stress Pathway Triggers a Potent Antiviral Response via ER-Associated Protein Degradation. *J Virol* **92**, (2018).

213. Marques, M. *et al.* Influenza A virus propagation requires the activation of the unfolded protein response and the accumulation of insoluble protein aggregates. *iScience* **27**, 109100 (2024).

214. Hassan, I. H. *et al.* Influenza A viral replication is blocked by inhibition of the inositol-requiring enzyme 1 (IRE1) stress pathway. *J Biol Chem* **287**, 4679–4689 (2012).

215. Hrincius, E. R. *et al.* Acute Lung Injury Results from Innate Sensing of Viruses by an ER Stress Pathway. *Cell Rep* **11**, 1591–1603 (2015).

216. Smith, J. A. A new paradigm: innate immune sensing of viruses via the unfolded protein response. *Front Microbiol* **5**, 222 (2014).

217. Volkmann, K. *et al.* Potent and selective inhibitors of the inositol-requiring enzyme 1 endoribonuclease. *J Biol Chem* **286**, 12743–12755 (2011).

218. Sanches, M. *et al.* Structure and mechanism of action of the hydroxy-aryl-aldehyde class of IRE1 endoribonuclease inhibitors. *Nat Commun* **5**, 4202 (2014).

219. Mazel-Sanchez, B. *et al.* Influenza A viruses balance ER stress with host protein synthesis shutoff. *Proc Natl Acad Sci U S A* **118**, (2021).

220. Machkovech, H. M., Bloom, J. D. & Subramaniam, A. R. Comprehensive profiling of translation initiation in influenza virus infected cells. *PLoS Pathog* **15**, e1007518 (2019).

221. Ulferts, R. *et al.* Subtractive CRISPR screen identifies the ATG16L1/vacuolar ATPase axis as required for non-canonical LC3 lipidation. *Cell Rep* **37**, 109899 (2021).

222. Ichinohe, T., Pang, I. K. & Iwasaki, A. Influenza virus activates inflammasomes via its intracellular M2 ion channel. *Nat Immunol* **11**, 404–410 (2010).

223. Lamb, R. A., Zebedee, S. L. & Richardson, C. D. Influenza virus M2 protein is an integral membrane protein expressed on the infected-cell surface. *Cell* **40**, 627–633 (1985).

224. Han, Y.-J. *et al.* Targeting influenza A virus by splicing inhibitor herboxidiene reveals the importance of subtype-specific signatures around splice sites. *Journal of Biomedical Science* **30**, 1–22 (2023).

225. Pinto, L. H., Holsinger, L. J. & Lamb, R. A. Influenza virus M2 protein has ion channel activity. *Cell* **69**, 517–528 (1992).

226. McCown, M. F. & Pekosz, A. The influenza A virus M2 cytoplasmic tail is required for infectious virus production and efficient genome packaging. *J Virol* **79**, 3595–3605 (2005).

227. Huang, Y., Urban, C., Hubel, P., Stukalov, A. & Pichlmair, A. Protein turnover regulation is critical for influenza A virus infection. *Cell Syst* **15**, 911–929.e8 (2024).

228. Pinto, R. M., Lycett, S., Gaunt, E. & Digard, P. Accessory Gene Products of Influenza A Virus. *Cold Spring Harb Perspect Med* **11**, (2021).

229. Wise, H. M. *et al.* Identification of a novel splice variant form of the influenza A virus M2 ion channel with an antigenically distinct ectodomain. *PLoS Pathog* **8**, e1002998 (2012).

230. Castrucci, M. R. *et al.* The cysteine residues of the M2 protein are not required for influenza A virus replication. *Virology* **238**, 128–134 (1997).

231. Holsinger, L. J. & Lamb, R. A. Influenza virus M2 integral membrane protein is a homotetramer stabilized by formation of disulfide bonds. *Virology* **183**, 32–43 (1991).

232. Lamb, R. A. The Structure, Function, and Pathobiology of the Influenza A and B

Virus Ion Channels. *Cold Spring Harb Perspect Med* **10**, (2020).

233. Meng, X., Templeton, C., Clementi, C. & Veit, M. The role of an amphiphilic helix and transmembrane region in the efficient acylation of the M2 protein from influenza virus. *Sci Rep* **13**, 18928 (2023).

234. Petrich, A., Dunsing, V., Bobone, S. & Chiantia, S. Influenza A M2 recruits M1 to the plasma membrane: A fluorescence fluctuation microscopy study. *Biophys J* **120**, 5478–5490 (2021).

235. Beale, R. *et al.* A LC3-interacting motif in the influenza A virus M2 protein is required to subvert autophagy and maintain virion stability. *Cell Host Microbe* **15**, 239–247 (2014).

236. Aledavood, E., Selmi, B., Estarellas, C., Masetti, M. & Luque, F. J. From Acid Activation Mechanisms of Proton Conduction to Design of Inhibitors of the M2 Proton Channel of Influenza A Virus. *Front Mol Biosci* **8**, 796229 (2021).

237. Ma, C. *et al.* Identification of the functional core of the influenza A virus A/M2 proton-selective ion channel. *Proc Natl Acad Sci U S A* **106**, 12283–12288 (2009).

238. Dong, H., Sharma, M., Zhou, H.-X. & Cross, T. A. Glycines: role in α-helical membrane protein structures and a potential indicator of native conformation. *Biochemistry* **51**, 4779–4789 (2012).

239. Georgieva, E. R. *et al.* Conformational Response of Influenza A M2 Transmembrane Domain to Amantadine Drug Binding at Low pH (pH 5.5). *Front Physiol* **7**, 317 (2016).

240. Hull, J. D., Gilmore, R. & Lamb, R. A. Integration of a small integral membrane protein, M2, of influenza virus into the endoplasmic reticulum: analysis of the internal signal-anchor domain of a protein with an ectoplasmic NH2 terminus. *J Cell Biol* **106**, 1489–1498 (1988).

241. Richard, M., Boulin, T., Robert, V. J. P., Richmond, J. E. & Bessereau, J.-L. Biosynthesis of ionotropic acetylcholine receptors requires the evolutionarily conserved ER membrane complex. *Proc Natl Acad Sci U S A* **110**, E1055–63 (2013).

242. Talbot, B. E., Vandorpe, D. H., Stotter, B. R., Alper, S. L. & Schlondorff, J. S. Transmembrane insertases and -glycosylation critically determine synthesis, trafficking, and activity of the nonselective cation channel TRPC6. *J Biol Chem* **294**, 12655–12669 (2019).

243. Coelho, J. P. L. *et al.* A network of chaperones prevents and detects failures in membrane protein lipid bilayer integration. *Nat Commun* **10**, 672 (2019).

244. Chen, Z. *et al.* EMC chaperone-Ca structure reveals an ion channel assembly intermediate. *Nature* **619**, 410–419 (2023).

245. Satoh, T., Ohba, A., Liu, Z., Inagaki, T. & Satoh, A. K. dPob/EMC is essential for

biosynthesis of rhodopsin and other multi-pass membrane proteins in Drosophila photoreceptors. *Elife* **4**, (2015).

246. Miller-Vedam, L. E. *et al.* Structural and mechanistic basis of the EMC-dependent biogenesis of distinct transmembrane clients. *Elife* **9**, (2020).

247. Shurtleff, M. J. *et al.* The ER membrane protein complex interacts cotranslationally to enable biogenesis of multipass membrane proteins. *Elife* **7**, (2018).

248. Addya, S. *et al.* Targeting of NH2-terminal-processed microsomal protein to mitochondria: a novel pathway for the biogenesis of hepatic mitochondrial P450MT2. *J Cell Biol* **139**, 589–599 (1997).

249. Duong, F. & Wickner, W. Sec-dependent membrane protein biogenesis: SecYEG, preprotein hydrophobicity and translocation kinetics control the stop-transfer function. *The EMBO Journal* **17**, 696 (1998).

250. Sääf, A., Wallin, E. & von Heijne, G. Stop-transfer function of pseudo-random amino acid segments during translocation across prokaryotic and eukaryotic membranes. *Eur J Biochem* **251**, 821–829 (1998).

251. Green, W. N. & Andersen, O. S. Surface charges and ion channel function. *Annu Rev Physiol* **53**, 341–359 (1991).

252. Nieva, J. L., Madan, V. & Carrasco, L. Viroporins: structure and biological functions. *Nature Reviews Microbiology* **10**, 563–574 (2012).

253. Shih, S. R., Nemeroff, M. E. & Krug, R. M. The choice of alternative 5’ splice sites in influenza virus M1 mRNA is regulated by the viral polymerase complex. *Proc Natl Acad Sci U S A* **92**, 6324–6328 (1995).

254. Eisenberg, S. *et al.* Mapping the electrostatic profiles of cellular membranes. *Mol Biol Cell* **32**, 301–310 (2021).

255. Gaspar, C.J., Vieira, L.C., Santos, C.C., Christianson, J.C., Jakubec, D., Strisovsky, K., Adrain, C. and Domingos, P.M. EMC is required for biogenesis of Xport‐A, an essential chaperone of Rhodopsin‐1 and the TRP channel. EMBO reports, 23(1), p.e53210. (2022)

CHAPTER 2: IAV M2 biogenesis partially depends on EMC, and other insertases of the Oxa1 family

2.1 Chapter Summary

Viruses depend on the host cellular machinery to synthesise their protein components, which can limit the replication speed of rapidly growing viruses such as influenza A virus (IAV). Many viruses express essential membrane proteins that are particularly challenging to synthesise, due to the complex manoeuvre of inserting these proteins into the ER membrane. Thus, membrane protein biogenesis may become rate limiting to supply these components for viral assembly. IAV encodes three essential membrane proteins: hemagglutinin (HA), neuraminidase (NA), and matrix protein 2 (M2). While HA and NA biogenesis have been well characterised, whereas M2's biogenesis remains incompletely understood. M2 is the smallest known proton-selective pH gated ion channel and its single transmembrane domain is sufficient for this function when assembled into a homo-tetramer. M2's remarkable characteristics make it an interesting model for a class of poorly characterised membrane proteins. M2's function as an ion channel is facilitated by hydrophilic residues within its single transmembrane domain, which are not well tolerated within the hydrophobic core of the lipid bilayer, and likely require special machinery during biogenesis. The objective of this chapter is to identify host factors that are important for M2 biogenesis. Using the well-established glycosylation assay to measure ER membrane translocation, we identify host machinery that plays a role in targeting M2 to the endoplasmic reticulum (ER) membrane, and is responsible for inserting M2 into the ER. We identify all ER resident members of the Oxa1 superfamily as potential M2 biogenesis factors, implying a cooperative function. The ER membrane protein complex (EMC) has been described to catalyse ER translocation of membrane proteins with terminal transmembrane domains as well as internal transmembrane domains of hydrophilic character. As a downstream measure of the efficiency of M2 biogenesis, we quantify how EMC ablation affects the rate at which M2 accumulates at the plasma membrane in infection. Surprisingly, M2 can traffic efficiently to the plasma membrane without EMC, albeit at a lower rate. The slower accumulation rate of M2 at the plasma membrane correlates with a delayed onset of virion production, but does not affect virion production later in infection. The delayed virion production does however affect the rate at which IAV transmits from cell to cell.

We hypothesise that EMC partly mediates M2's ER membrane insertion which optimises M2's rate of accumulation at the plasma membrane, but that M2 can possibly access multiple ER membrane translocons of the Oxa1-family in the absence of the EMC. We quantify for the first time the efficiency of redundant pathways in membrane protein biogenesis and their relevance for an essential viral component. Uncovering the molecular details of M2 biogenesis allows us to study mechanisms of translocon selection, or the potential for adaptive responses and load sharing between ER membrane translocation routes. This may be important for viral replication of many human and animal pathogens, as well as for conditions where a disruption of proteostasis is the cause of disease.

2.2 Introduction

Membrane protein biogenesis requires first, the recognition of a membrane protein in the cytosol, which initiates targeting to the ER membrane and subsequent translocation through a preferred translocon [1]. Membrane targeting is initiated by the emergence of a hydrophobic stretch of amino acids from the ribosomal exit tunnel. Upon signal emergence, most ribosomes associate with the signal recognition particle (SRP) to create a platform for membrane attachment and handover of the nascent membrane protein to the ER membrane [2]. SPR is assembled from a 300 nucleotide RNA and 6 proteins and positions SRP54's hydrophobic groove right in front the ribosomal exit tunnel to engage emerging hydrophobic signals from nascent polypeptides [3]. SRP attachment leads to subsequent capture of the translating ribosome by the bipartite SRP receptor (soluble SRPRA, membrane bound SRPRB) at the ER membrane, which creates the required proximity of a transmembrane domain with the ER membrane to be captured by the ER resident translocons [4,5]. It is possible for membrane protein targeting to proceed independent of the SRP system through the association of a fully translated membrane protein with cytosolic chaperons [1]. Post-translational membrane protein targeting is necessary, if the first hydrophobic signal is located far away from the N-terminus, or if the translation reaction terminates before targeting can be achieved [6]. In this way, organellar membranes other than the ER can be supplied with their respective membrane proteins [7]. At the ER membrane, a translating ribosome, or a fully translated protein encounters several ER resident membrane protein complexes, that are available to catalyse the translocation of hydrophilic stretches of amino acids across the membrane and insert hydrophobic transmembrane domains into the lipid bilayer [1].

M2 is a 97 amino acid long membrane protein, with a single transmembrane helix at its amino terminus (N-terminus), that anchors M2 in the membrane with its N-terminus facing the ER lumen (Nexo, typeIII) [8]. M2's short C-terminus might allow co-translational or post-translational translocation at the ER [9]. The biogenesis of M2-like proteins with terminal signal anchors have been focused upon in recent years. *In vitro* translocation assays of typeIII nascent chains stalled at different length show that the timing of targeting and translation can determine the insertion route either via the EMC for long nascent chains exposed at the ribosome, or Sec61 for short peptides, demonstrating for the first time how variation in translation and targeting dynamics can be buffered by different components of the ER insertion machinery [10].

The EMC is a 9-subunit complex of which EMC6 and EMC3 form a highly conserved motif of the Oxa1 superfamily (Oxa1 insertase for mitochondrial encoded membrane proteins) of protein translocating complexes, which are thought to locally collapse the lipid bilayer, which reduces the energetic barrier enough for the translocation of short hydrophilic terminal peptides or loops [11]. Other members of the same family present in the ER membrane are the GET complex (WRB and CAMLG receptor complex), which is important for post-translational translocation of tail anchored proteins, and the GEL complex (TMCO1 and OPTI complex), which is important for the insertion of multipass transmembrane proteins [1]. The Sec61 complex creates an aqueous protein conducting pore across the ER membrane that is capable of conducting long hydrophilic amino acid stretches or soluble proteins for secretion and insert transmembrane domains through its lateral gate [1]. Membrane proteins can be targeted post-translationally to Sec61, which requires association of Sec61 with the Sec62 and Sec63 complex [12]. In yeast a further SRP independent pathway (SND-pathway) for proteins with centrally located signal anchors has been characterised, which similar to the GET pathway consists of a cytosolic chaperone and a ER membrane bound receptor [13]. The pathway is active in mammalian cells, however, to date only a homolog of the ER bound receptor TMEM208 has been identified [14]. TMEM208 recently has been found to accelerate substrate release from SRP during co-translational ER membrane targeting [5].

2.3 A selective cas9 KO screen identifies M2 biogenesis factors

Many proteins that enter the secretory route will be linked to large glycan molecules that are successively modified during trafficking [15]. N-linked glycosylation has been long used to assess ER translocation and trafficking as the covalent addition of a

glycan chain can be easily monitored in a western blot [16]. The presentation of a 3 amino acid motif (NXT/S) to the oligosaccharyltransferase complex (OST) at the luminal face of ER membrane will lead to the rapid covalent linkage of a glycan chain. OST can act either cotranslationally, through a tight association with Sec61 during protein extrusion, or post-translationally. Once a protein has been glycosylated in the ER these glycan chains will be modified along the secretory route when proteins pass through the golgi apparatus (Figure 2.1A) [15].

M2 is not natively glycosylated, but a single amino acid substitution (L3N) at M2's N-term creates the glycosylation motif in our reporter construct M2.Y.F (Y for glycosylation site and F for a C-terminal Flag-tag, Figure 2.1B). The reporter construct can be transfected into A549 lung epithelial carcinoma cells to analyse its trafficking behaviour by western blot. M2 traffics along the secretory route to reach the plasma membrane. 3 distinct populations can be detected by western blot corresponding to different trafficking intermediates (Figure 2.1A). We can determine these glycosylation intermediates by glycosidase digestion. All N-linked glycans are susceptible to PNGase digestion, but modifications that occur later in the Golgi network are resistant to EndoH digestion [17]. This means that the EndoH sensitive band with an 6kDa shift (+Y) relative to the unglycosylated form (-Y), corresponds to ER localised or cis-Golgi localised M2, and that the hyperglycosylated smear around 35kDa (++Y) which is only sensitive to PNGase treatment corresponds to golgi or plasma membrane localised M2. Besides being glycosylated, M2.Y.F seems to be additionally post-translationally modified judging by the faint, PNGase resistant band at around 30kDa (Figure 2.1A).

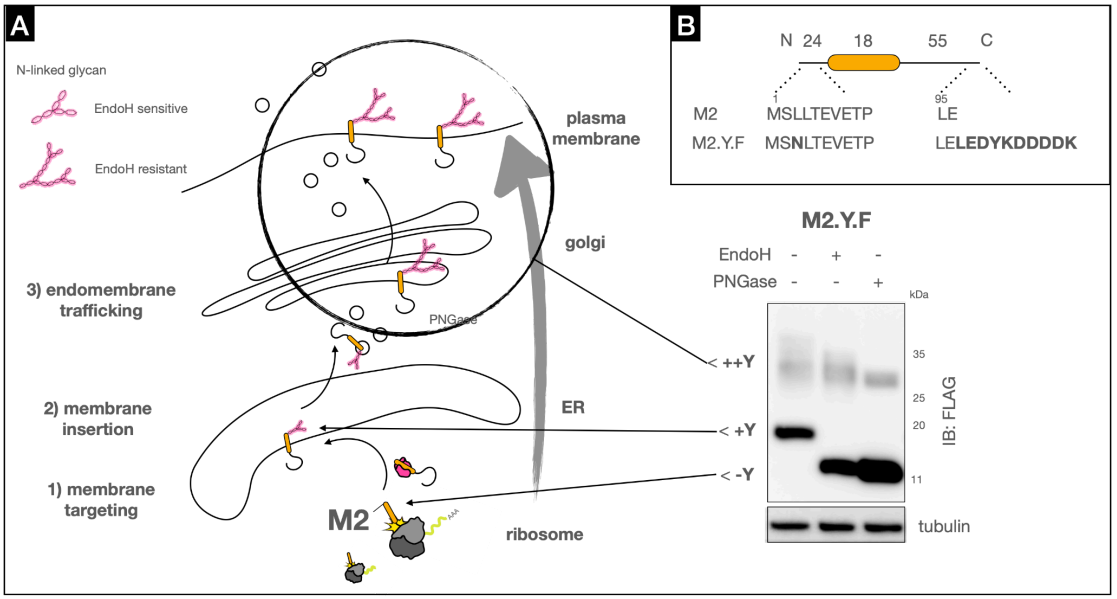


Figure 2.1 Glycosylation competent reporter construct M2.Y.F

A549 cells were transfected with the L3N mutant of PR8 M2. The cell lysates were either mock treated or deglycosylated with PNGase or EndoH before analysing by western blot. The resulting bands reveal trafficking intermediates of M2 (N=2).

In order to identify M2 biogenesis factors, we perform a selective cas9 mediated knock out screen in 293T cells targeting known membrane protein biogenesis factors (Figure 2.2C). Each gene was targeted by 2 *in vitro* transcribed guides and M2 glycosylation was tested 72h after guide transfection. This approach is designed to minimise the time cells could adapt to the knockout conditions and shunt clients from one pathway into another, as well as to minimise cytotoxic effects from the loss of essential genes. The CHRONOS database is a fantastic resource that quantified the effect of most single gene knockouts in over 1000 cancer cell lines as part of the DEPMAP project [18] (Figure 2.2B). Our approach is inspired by the method developed in the Liu lab, where cas9-sgRNA complexes are synthesised ex vivo and transfected together to achieve a brief spike of nuclease activity [19]. A negative charge of the molecules is important to escape endosomes efficiently. Based on this observation we opted to transfect the guide RNA by itself and let the cell synthesise cas9 from a plasmid. To do so, 293T cells are first transfected with cas9 to express high amounts of the protein from multiple plasmid copies per cell. In order to deliver a high amount of guide RNA for a short amount of time, the guides are *in vitro* transcribed and transfected like siRNAs. We transfect M2.Y.F 24h after guide transfection and collect lysates 48h after transfection. 72h after guide RNA transfection should be enough time to deplete most protein from targeted cells, however we might not be able to transfect all cells which might result in a mixed population of cells having lost expression or expressing to WT levels. We control for the knock down on a population level by western blot (Figure 2.2A).

For the assessment of M2 dependence on a targeted factor we assessed the glycosylation efficiency as the ratio of a glycosylated band (+Y) over the sum of +Y and an unglycosylated band (-Y) (-Y/SUM(-Y,+Y)) (Figure 2.2A). Ratios smaller than for the non-targeting condition are evidence of a biogenesis defect due the accumulation of a non-glycosylated band. This could be due to M2 not inserting into the ER membrane and accumulating cytosolically, M2 changing topology and inserting Cterm first in a non-functional configuration, or M2 inserting into another organelle.

The screen identifies at least one component of the SRP targeting system SRPRA as a clear hit. This is in agreement with an earlier study that found that M2 is

not inserting into alkaline-extracted rough rabbit microsomes [20], but requires a soluble factor that is likely SRP [21]. SRPRA has been previously identified to act in the biogenesis of other typeIII proteins of similar length as M2 [22]. Surprisingly, M2 glycosylation seems to be strongly affected by several Oxa1 family members, namely EMC (EMC3 and obligatory cofactor EMC6), GET (CAMLG) and GEL (TMCO1) components, but not for the GET pathway-specific targeting factor GET3. Although the targeting of TMCO1 shows a clear phenotype, we cannot confirm the depletion due to limitations in the immuno-detection. Previous reports refer to EMC as the preferred insertion pathway for typeIII proteins, but that insertion via Sec61 is the alternative [10]. Sec61 targeting does not show a defect in M2 glycosylation, however due to limitations in the immuno-detection of Sec61 in our screen and the high impact on cell survival we cannot conclude if Sec61 is important for M2 biogenesis. Similarly, the immuno-detection of SRP54 and TMEM208 proved to be challenging and limit conclusions about the involvement in M2 biogenesis. Besides this, Sec63 targeting did not deplete the protein and so we cannot assess the effect on M2 biogenesis.

The result of the screen hints at M2 biogenesis being sensitive to the loss of multiple Oxa1 family members. The possible access of a membrane protein to several Oxa1 family members has been demonstrated for post-translationally targeted tail-anchored proteins [23], however a cooperative function in the biogenesis of single pass membrane proteins has not been described before. As the EMC has been described as the preferred translocation route for typeIII proteins and a central hub in membrane protein biogenesis of many protein topologies that can be targeted co-translationally and post-translationally [1,24], we focus on the role of EMC in M2 biogenesis and consider the other Oxa1 family members as possible alternative pathways.

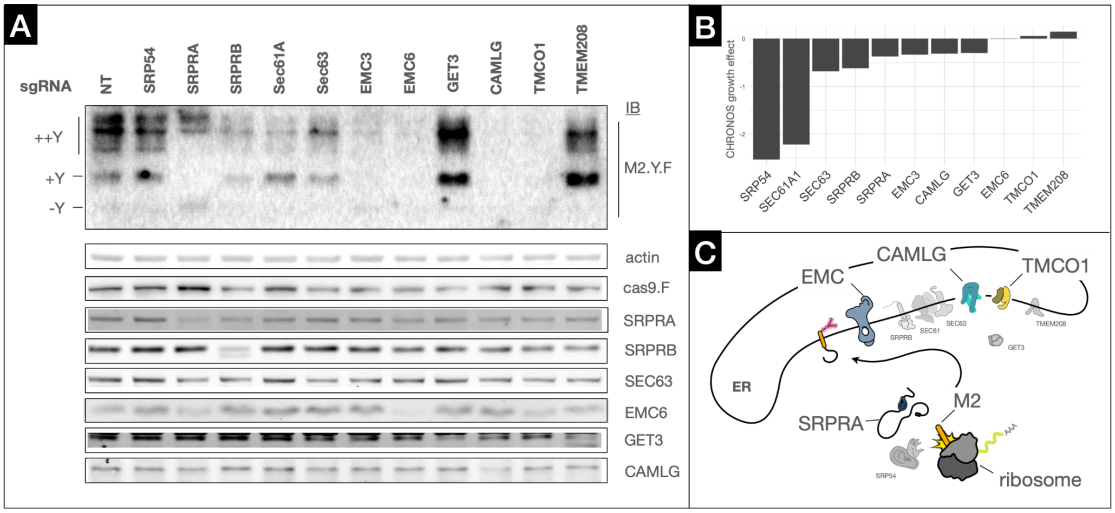


Figure 2.2 Targeted KO screen for M2 biogenesis factors in 293T cells

A) 293T cells were transfected with cas9 and subsequently with *in vitro* transcribed guides against a specific part of the membrane protein targeting or translocation machinery. KO cells were transfected with M2.Y.F and cell lysates analysed by western blot (N=2) B) The measured growth defect of a single gene KO on A549 cells from the CHRONOS database. C) Scheme of factors involved in M2 biogenesis.

2.4 EMC aids in M2 biogenesis and trafficking in transfection and infection

To investigate the importance of EMC and the alternative Oxa1 translocons for M2 biogenesis, we generated A549 cell lines lacking subunit EMC6 which abrogates EMC function but does not affect Sec61 and GET function (Figure 2.3C) and is potentially less toxic than EMC3 KO (Figure 2.2B). Due to EMC6 being one of the core components assembling the complex, EMC5 is also lost in EMC6 KO cells which is a commonly observed mechanism of subunit regulation by stoichiometry [25,26] (Figure 2.3B). The polyclonal A549 EMC6 KO cells (Δ6) do glycosylate HLA-A (N-terminal FLAG-tag, Sec61 client), and SEC61B (C-terminal HA-tag, GET client), but do not glycosylate a known EMC client SQS (C-terminal HA-tag, Figure 2.3C). Reconstituting EMC6 expression in the Δ6 cells, through an inducible promoter rescues SQS glycosylation (+6). WT and Δ6 cells do have a similar doubling time, although we did not quantify this.

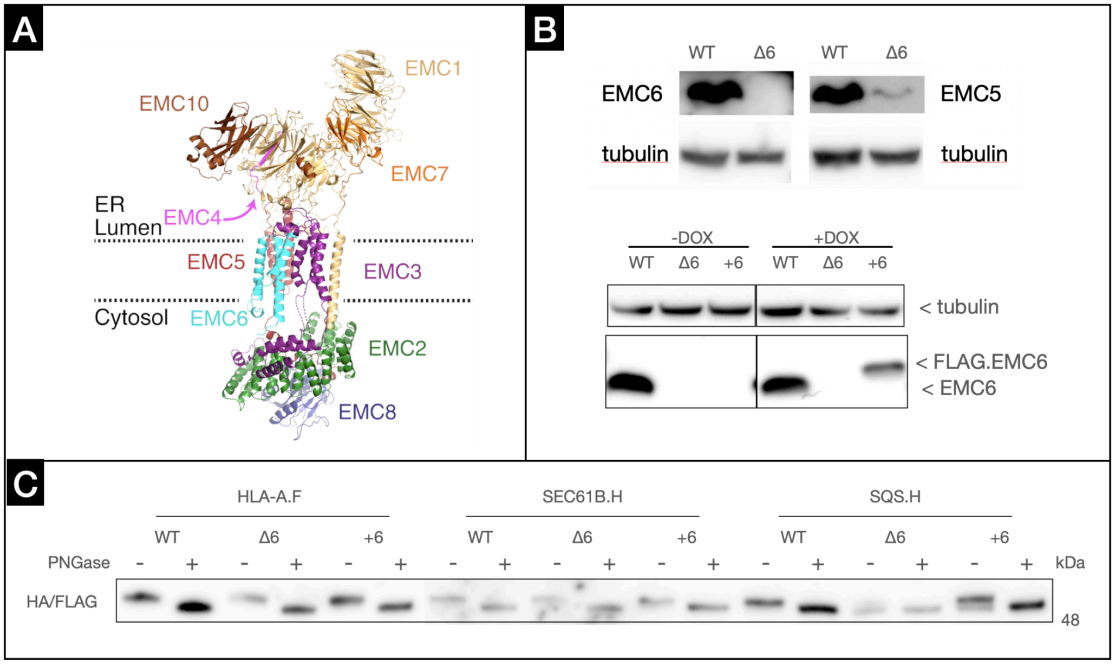


Figure 2.3 A549 EMC6 cell lines

A) The EMC is a 10 subunit complex with EMC6 being one of its core membrane proteins, forming part of the hydrophilic vestibule (reproduced from [27]). B) A549 EMC6 KO cell lines (Δ6) were produced by CRISPR cas9 guided double strand break through the transfection of PX330 with 3 different guides and selection with puromycin. The loss of EMC6 leads to a substantial reduction in EMC5. EMC6 rescue cell lines (+6) with doxycycline inducible EMC6 were produced by lentiviral transduction of FLAG.EMC6 into Δ6. WT and Δ6 were transduced with empty vector. Cell lysates were analysed by western blot (N=2) C) A549 WT, Δ6 or +6 cells were transfected with glycosylation competent proteins of different ER translocation machinery. Cell lysates were analysed by western blot (N=2). SEC61B.H and SQS.H were obtained from the Christianson lab.

A549 Δ6 cells recapitulate the glycosylation and expression phenotype observed in the KO screen in 293T cells (Figure 2.4A). We confirm the effect on M2 expression in IAV infection. A549 Δ6 cells infected for 8 and 12 hours with the lab-adapted IAV strain A/Puerto Rico/8/34 (PR8) show a sustained loss of M2 relative to virally expressed nuclear protein (NP) normalised to loading control (Figure 2.4B), which results in lower levels of M2 at the plasma membrane (Figure 2.4D). The +6 cell line expresses very low amounts of EMC6 (Figure 2.4C) which rescues overall M2 expression. However, the rescue population is not homogenous and cells deliver M2 to the plasma membrane to WT levels or Δ6 levels, suggesting that EMC6 might not be sufficiently expressed in all cells to rescue M2 biogenesis (Figure 2.4D). However, these results clearly show that M2 can traffic to the plasma membrane in the right topology without EMC, as detected by the N-terminal specific antibody at the cell surface. The data collectively indicates that EMC plays a role in M2 biogenesis, and that M2 cannot accumulate to the same levels without EMC, but that plasma membrane trafficking is possible without EMC.

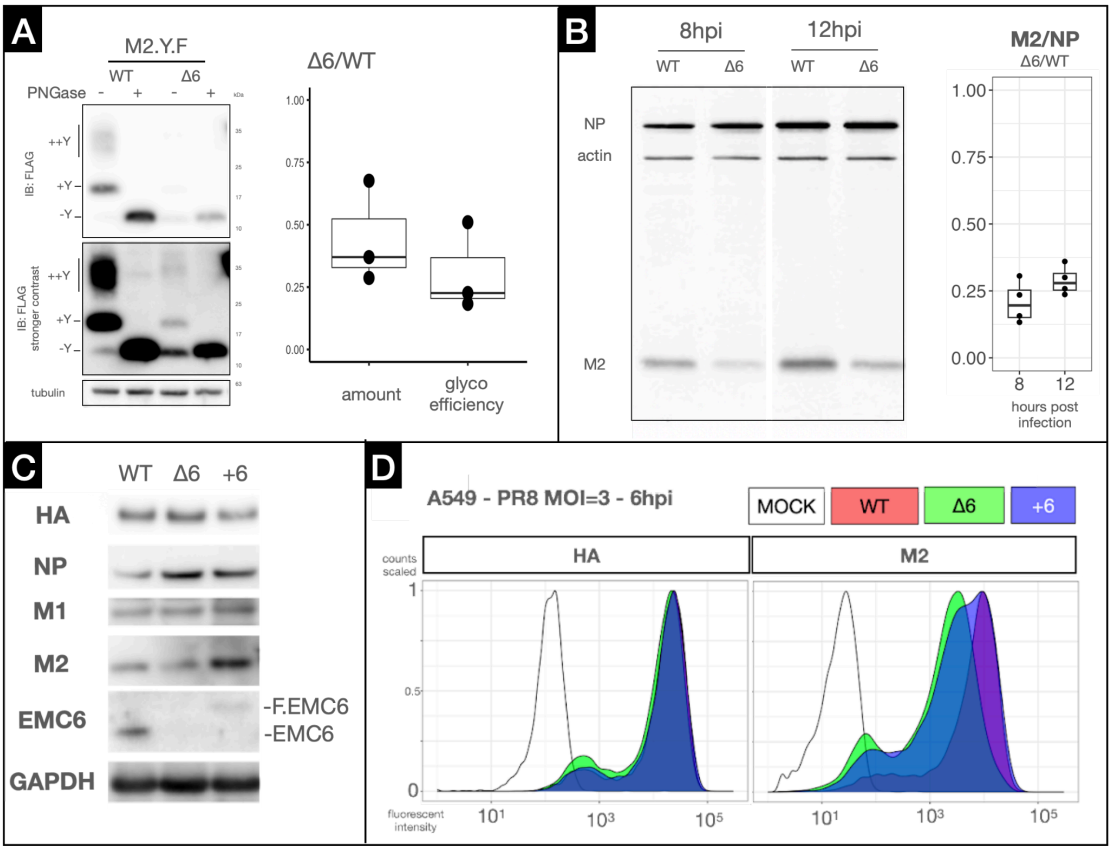


Figure 2.4 Effect of EMC loss of function on M2 expression in transfection and infection

A) A549 WT or Δ6 cells were transfected with M2.Y.F and incubated for 24h, cell lysates were analysed by western blot. The glycosylation efficiency (+Y/SUM(+Y,-Y) and the total protein amount was quantified from western blots (N=3). B) A549 WT or Δ6 cells were infected with PR8 at MOI of 3 and cell lysates were harvested 8 or 12 hours post infection. The protein amount was quantified from western blots (N=4). C) A549 WT, Δ6 or +6 cells were induced with 1ug/ml doxycycline and infected with PR8 at MOI 3 and cell lysates were harvested 6 hours post infection. The lysates were probed for different viral proteins and EMC6 by western blot (N=2). D) A549 WT, Δ6 or +6 cells were infected with PR8 at MOI of 3 and cells were detached 6h post infection, fixed, stained for HA and M2 and measured by flow cytometry (N=3).

2.5 The lack of EMC affects M2 trafficking dynamics, but the redundant pathways are sufficient for infection

The level of viral proteins at the plasma membrane is a function of the amount of protein trafficking to the plasma membrane and the amount lost from the plasma membrane through retrograde transport, budding virions, or degradation [28,29].

Complicated non-linear effects of viral protein trafficking over time of infection are to be expected, due to specific viral protein expression dynamics, modulation of host protein expression [30], host cell immune defences that modulate protein synthesis or degradation [31], remodelling of secretory organelles [31], or interaction between viral proteins. These complex relationships make it difficult to formulate a mechanistic model of M2 trafficking, however we hypothesise that EMC acts during M2 biogenesis, and thus affects M2 trafficking reactions downstream. We aimed to capture the dynamics of how M2 accumulates at the plasma membrane over time in WT and EMC6 ablated conditions by infecting cells with PR8 virus and measuring M2 accumulation at the plasma membrane over a 15h infection time course in a flow cytometry assay. HA accumulation served as a control between WT and ∆6 cells. PR8 expresses higher amounts of M2 than other IAVs [32], which might make the EMC more or less relevant for M2 trafficking depending on the amount needed at the plasma membrane and the efficiency of EMC and the redundant pathways. To test the effect of M2 synthesis levels on the relevance of EMC we needed to modulate M2 expression in the virus. However, modulating M2 expression independently is not easy since it is expressed from a spliced version of the full length M1 mRNA. We used a previously characterised chimeric virus expressing M2 from segment 7 (M gene) of A/Udorn/307/1972 (H3N2) (hereafter named Udorn) in the PR8 background previously named MUd. This virus expresses a different M1 and M2 carrying 7 and 10 amino acid changes in its sequence respectively (Figure 2.5C). Importantly, this virus expresses much less M2 compared to PR8, whereas the expression of other viral proteins is not majorly affected (Figure 2.5D).

Measuring M2 and HA at the plasma membrane results in a distribution of cells that contains two populations, which represents either uninfected cells, or cells that are infected but do not express the protein yet, or infected cells expressing viral proteins. According to the poisson distribution 5% of the cells should be uninfected at an MOI of 3. We separated the viral protein expressing cells by fitting two skew-normal distributions to accurately assess the mean fluorescent intensity of the protein expressing population at every time point (Figure 2.5A and 2.5B). Comparing the plasma membrane accumulation of HA over time, shows no difference between WT and ∆6 in PR8 or MUd, albeit with HA reaching two times higher levels in MUd infected cells compared to PR8 (Figure 2.5E). M2 accumulation at the plasma membrane shows clear differences between PR8 and MUd infected cells, with MUd infected cells accumulating M2 much slower than PR8 infected cells, which is congruent with a lower expression of M2 in MUd. EMC affects the accumulation of M2 at the plasma membrane differently in PR8 and MUd. In PR8 infected cells, M2

accumulation is most affected at early time points in absence of EMC (22% Δ6/wt at 5hpi), but catches up to 81% of WT levels at 15 hours post infection. In MUd infection, the loss of EMC affects the plasma membrane levels of M2 to reach about 45% of the levels in WT cells constantly over the time of infection measured (Figure 2.5E). The differences between WT and Δ6 at 5 and 6 hours post infection are very low, and measurement noise outweighs the effect size. Thus, these time points are not informative.

M2 is found only at very low levels in virions compared to HA (15 copies of M2 [33] to 300 copies of HA [34] for spherical virions). We test if the effect of EMC on M2's plasma membrane accumulation has an effect on virion production. We only observe a growth defect at 5 hours post infection in PR8 infected cells, where 2 out of 3 biological replicates have titers below the level of detection in Δ6 conditions (<10 PFU/ml), however at 15 hours post infection, virion production is indistinguishable (Figure 2.5F). Even in MUd infected cells where M2 levels at the plasma membrane are more severely affected we see only a strong effect at 6 hours post infection, where all measured supernatants of the Δ6 condition are below the level of detection. However at later time points we also observe that MUd infection produces infectious virus particles in EMC ablated conditions so that at 15 hours post infection there is no significant difference to the WT (Figure 2.5F).

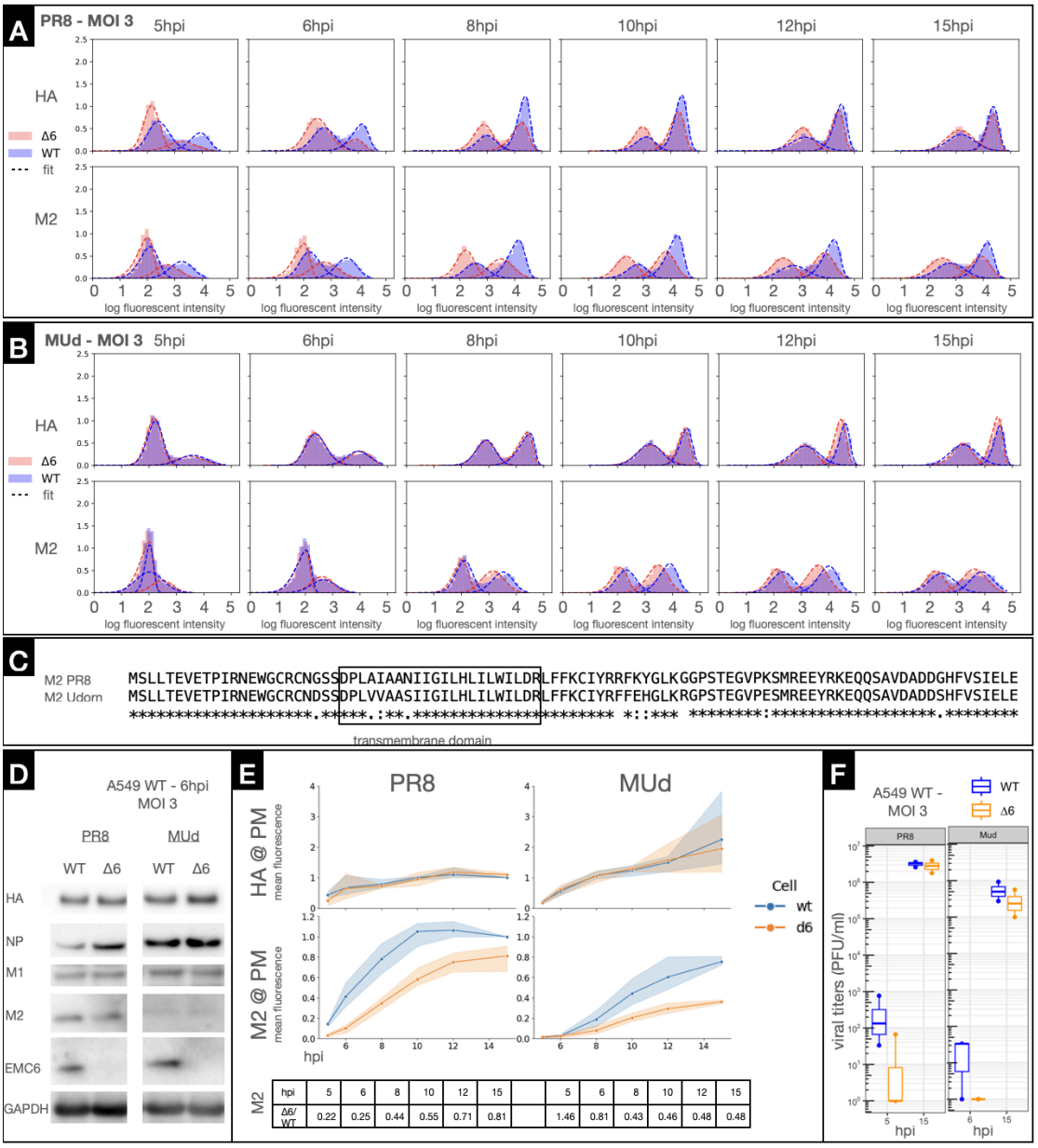


Figure 2.5 Dynamics of M2 accumulation at the plasma membrane in PR8 and MUd infected cells

A) A549 WT or Δ6 cells were infected with PR8 at MOI 3 for cell harvest at the indicated time points and subsequent surface expression of viral HA and M2 at the cell surface by flow cytometry. The mean of an HA or M2 positive population of single cells was accurately estimated by fitting a skew normal distribution (N=3). B) Same as for A) but cells were infected with MUd virus, a chimeric virus of PR8 carrying Udorn segment 7 (N=3). C) Viral protein expression of cells infected at MOI of 3 with either PR8 or MUd at 6 hours post infection analysed by western blot (N=2). D) Time series of the estimated means from A) and B) for HA and M2 normalised to the value of PR8 in WT conditions at 15hpi for each protein in each replicate so that values between viruses remain comparable as they were measured at

the same time. E) Viral titers analysed from supernatants of the 5h and 15h time points for PR8 and 5h and 15h time points for MUd of the same experiment (N=3).

Instead of just comparing individual time points we compared the entire accumulation dynamics between the different conditions. In the absence of a mechanistic understanding of M2 trafficking in infection, we modeled the dynamics with a logistic growth model in order to obtain interpretable parameters. We estimated the maximal level reached (parameter L), the maximal slope of the curve (parameter k), with which we calculate the maximal rate of accumulation (k*L/4), and the point at which the maximal accumulation rate occurs (parameter x0). First, we normalise all values to the estimated parameter L of the PR8 infected WT cells for each repeat (Figure 2.6A). Most interestingly, the maximal rate of accumulation (L*k/4) of M2 at the plasma membrane over the course of infection is similarly reduced to 56% ∆6/WT for PR8 and to 46% ∆6/WT for MUd (Figure 2.6B), arguing for a similar effect of EMC ablation on M2 trafficking independent of the expression strength or strain dependent amino acid substitutions.

These results show that EMC affects M2 accumulation dynamics at the plasma membrane, and in the case of MUd leads to a much reduced surface expression. However, significant amounts of residual M2 can traffic to the plasma membrane in a functional state without EMC. For PR8 replication in A549 cells, M2 is not rate limiting for virion production (Figure 2.5F). We reasoned that in a cell line that produces higher amounts of virions M2 might become rate-limiting in the absence of EMC. Thus, we turned to another model influenza cell line. MDCK cells produce high amounts of virions in a short time and have been used to propagate IAV in the lab for decades.

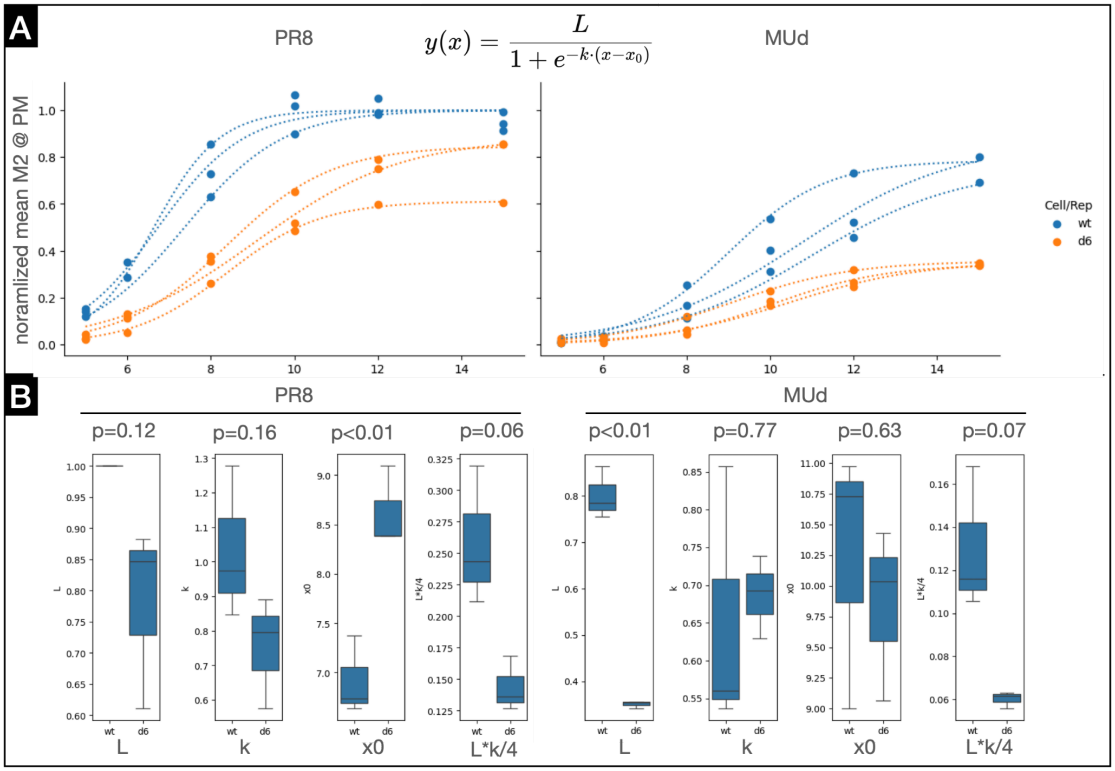


Figure 2.6 Comparison of growth dynamics with parameters obtained from logistic growth fit

A) A logistic function was fitted to each repeat of time series (data from 2.5A and 2.5B), normalised to the estimated parameter L of PR8 in WT conditions and fit again to the normalised data. B) Quantification of the estimated parameters from A) with p-values assessed using t-test statistics.

2.6 EMC function is important for fast replication and cell to cell transmission

MDCK cells are canine kidney cells which have been used to propagate influenza virus in the lab. We knock out EMC6 from these cell lines using the same RNA guides as for A549 cells as the gene is highly conserved. We produce EMC6 rescue cell lines with the same strategy as before. MDCK Δ6 infected with PR8 virus, selectively loses most M2 expression, without affecting other viral proteins, and we observe a small growth defect at 6hpi although the difference is not significant (Figure 2.7A). MDCK cells are strongly adherent and can be used to observe cell to cell transmission, which requires the protease trypsin in the media. To test whether the small defect in early replication has an effect on viral cell to cell transmission we infected monolayers of MDCK WT or Δ6 cells with a low quantity of virus of 60 PFU per well and quantified the spread of single infection foci 30h post infection. We determined the spread of the infection by cytopathic effect (Figure 2.7B) and by anti

HA staining (Figure 2.7C). PR8, Udorn and a previously circulating strain of seasonal influenza A/USSR/90/1977 (USSR, H1N1) were all significantly affected in their cell to cell transmission, as determined by the plaque size (Figure 2.7B). A fast growing, filamentous virus Udorn demonstrates a larger defect in cell to cell transmission than PR8 in this assay.

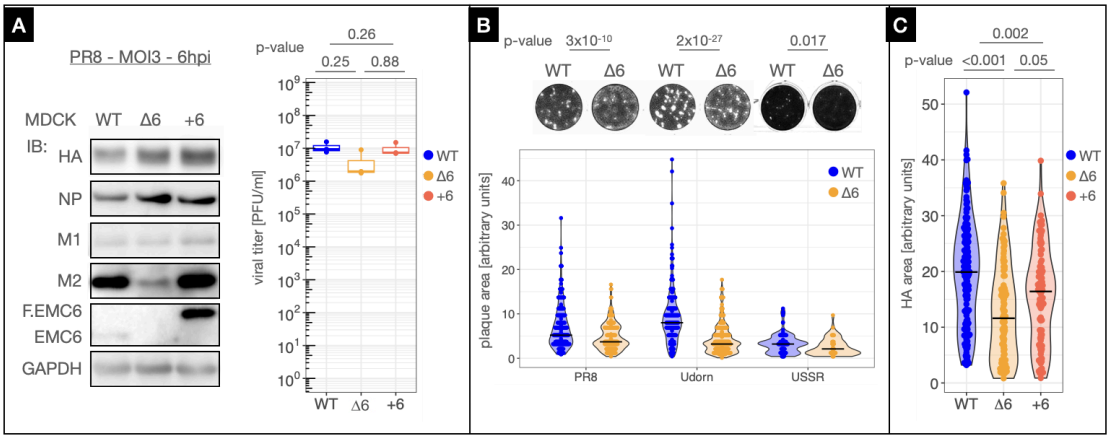


Figure 2.7 Effect of EMC loss on cell to cell transmission in MDCK cells

A) MDCK WT, Δ6 or +6 cells infected with PR8 at MOI of 3 for 6 hours. Cell lysates are analysed by western blot (N=2). Supernatants were titrated by plaque assay (N=3). P-values by one-way ANOVA B) A confluent monolayer of MDCK WT or Δ6 cells was infected with 60 PFU of PR8 virus incubated with viscous AVICEL overlay for 30h. Cells were fixed, stained, imaged and plaque area was quantified in imageJ (N=3). P-values by t-test. C) A549 WT, Δ6 and +6 cells were infected as in B), fixed and stained with HA antibody and imaged. Area was quantified in imageJ and p-values assessed by one-way ANOVA (N=3).

2.7 Discussion

Successful IAV replication requires M2 to be localised to the plasma membrane to incorporate into budding virions [8,33] and aid in virion release by inducing membrane curvature [35]. M2's biogenesis requirements have not been resolved before and are of interest due to M2's essential role during infection and its pathogenic effect. Besides the interest for virology, M2 may serve as a model protein, due to functionally relevant hydrophilic residues within its transmembrane domain [36].

Our initial screen for M2 biogenesis factors targets components of many of the known membrane protein targeting routes or membrane inserting translocons. M2's short C-terminus may allow for co- or post-translational targeting. We find at least one of the components of the cotranslational membrane protein targeting machinery SRPRA to reduce M2's glycosylation efficiency. This is in agreement with earlier studies, demonstrating the necessity of M2's SRP targeting *in vitro* [21]. SRP has been shown to target tail-anchored membrane proteins for post-translational

insertion *in vitro* [37], however the association could so far not been confirmed in living cells, where tail-anchored proteins either target via calmodulin, SGTA or GET3 [23]. Surprisingly, our screen shows similar defects in M2 biogenesis for all known members of ER resident Oxa1 family insertases. While EMC can translocate proteins that have been targeted post- or co-translationally, the GET pathway has been described to only be targeted post-translationally [38]. The GEL complex has been characterised as functioning in multipass membrane protein biogenesis but not for the insertion of single pass membrane proteins. GEL inserts later TMDs of multipass membrane proteins, after EMC or Sec61 have inserted the first TMD [1]. As for EMC targeting of tail-anchored membrane proteins, a targeting by translocase-proximal substrate release could be plausible [39]. Earlier studies have identified that also HIV typeIII protein VPU is SRPRA and EMC5 dependent in an *in vitro* translation assay [22]. VPU is similar to M2 in length, but its transmembrane domain harbors less hydrophilic residues. It is unclear if VPU as well can access redundant pathways and function in infection without EMC. A mechanistic understanding of redundancy between the Oxa1 family of insertases is lacking at this point. However, the existence of redundant access to membrane protein targeting pathways [40] or ER translocation pathways [24] has been observed. GET and EMC may have an overlapping client range [41], while inefficient translocation via EMC can be compensated via the SEC61 or GEL complex [24]. This redundancy between translocation pathways may cause competition for insertion via the remaining Oxa1 family translocons, this way affecting M2 insertion. If resource limits are the reason for loss of M2 in absence of EMC is testable, by overexpressing the potentially limiting components (GET2 and CAMLG) in absence of EMC. Furthermore, in our screen we observe that the loss of TMCO1 slightly affects EMC6 expression. TMCO1 has been predicted to interact with EMC6 in a recently published computational human interactome [42]. The predicted structure of TMCO1 and EMC6 creates a similar positively charged vestibule like the complex of EMC3 and EMC6, making a function of EMC6 outside the EMC complex plausible. However, biochemical evidence for this hypothesis is lacking.

Although it is very well known that different IAVs express varying amounts of M2 [30,32,43], it is unclear how this affects viral replication. Previous research tries to link high M2 expression to maladaptation [44]. M segment splice sites are adapted to their host, so that avian M segment produces a suboptimal ratio of M1/M2 in mammalian cells [30]. Uncoupling M2 expression from other M gene functions by genetic means is difficult, as M1 expression level regulates nuclear export of newly synthesized genomic segments, and synonymous mutations outside of splice sites can have growth defects [45]. M2 needs to reach the plasma membrane to be incorporated into

virions, however previous research shows that a small plasma membrane localized fraction may be sufficient for efficient replication [46]. In line with these observations, our results show that a delay in M2 accumulation at the plasma membrane, through the ablation of M2 biogenesis factors (EMC6 KO), only affects the onset of infectious virus particle production, without affecting later time points of infection. Another study that limits M2 plasma membrane localization through increased M2 degradation finds similar effect sizes to our study [28]. The E3 ligase MARCH8 has been shown to remove M2 from the plasma membrane [28]. Overexpression of MARCH8 results in reduced M2 plasma membrane levels in WSN (IAV H1N1) infected cells, which results in an increasing defect in viral titers with the progression of infection to reach a maximum of 5 fold difference at 12h [28]. Surprisingly, knockdown of MARCH8 increases viral titers, arguing that M2 is rate limiting in their system. However, M2 levels at the plasma membrane have not been measured under MARCH8 knockdown conditions, thus we cannot conclude that increased M2 plasma membrane levels are correlated with an increased viral titer. Another study depleted splicing factors that limit the amount of M2 mRNA without affecting other M gene splicing products, achieving a similar constant decrease of M2 expression as in our study. In contrast to our results, the decrease of M2 levels through the depletion of a splicing factor resulted in a 20 fold difference 24h post infection with WSN virus [47]. The vastly different effects on virion production in this study may be due to knock-on effects upon depletion of a host splicing factor.

Consistent with an effect on the onset of virion release, we show that the lack of EMC function can limit cell to cell transmission in a cell line that favors fast replication. A similar inhibition of viral cell to cell transmission has been observed when targeting M2 at the infected cell surface with antibodies [33,48]. The loss of EMC function has a larger effect on Udorn cell to cell transmission than PR8 in our assay, which might be due to Udorn forming larger filamentous virions compared to PR8's spherical or bacillus type morphology. Filamentous IAVs might spread from cell to cell through protrusions [49,50]. However, it remains to be tested if the lack of EMC affects virion morphology.

Functions other than M2's structural role might require a specific expression dynamic to allow optimal replication and in host infection dynamics. Besides M2's essential role during viral entry as trigger for the viral genome release [51], M2 has been found to mediate the localization of other viral components to the budding site at the plasma membrane [46], deacidify endomembrane compartments [52], which may lead to inflammasome activation [53], aid in ESCRT-independent viral particle release from the plasma membrane [35], and stabilise host factors through direct interaction [54].

Thus interfering with M2's expression dynamics may be a viable path for therapeutic treatment against severe influenza.

2.8 Methods

Cell lines

A549 (ATCC CCL-185), HEK293T and Madin-Darby Canine Kidney (MDCK-I, acquired from ATCC) were obtained from Prof. Paul Digard, Roslin Institute, UK, and cultured in Dulbecco's Modified Eagle Medium (DMEM) supplemented with 10% fetal bovine serum (FBS), 2 mM L-glutamine, and 1% (v/v) penicillin-streptomycin in a humidified incubator at 37°C and 5% CO2.

Antibodies

Flag-tag (Sigma-Aldrich F1804-200UG), HA-tag, M2 14C2 (Abcam ab5416), IAV-HA hybridoma (gift from Jonathan Yewdell), NP (abcam ab128193), M1 (Abcam ab20910), EMC6 (Abcam ab84902), EMC5 (Abcam ab122202), GAPDH, tubulin YL1-2 hybridoma (homemade), streptavidin-HRP (Thermo Fisher S911), actin (Sigma-Aldrich A5441), SRPRA (Sigma-Aldrich SAB1409707), SRPRB (Thermo Fisher PA5-87107), SEC63 (Proteintech 13978-1-AP), GET3 (Abnova H00000439 M03), CAMLG (Cell Signalling Technology 16713913S), LC3B

Reagents

Formaldehyde (Acros 10231622), JetPrime (Polyplus 101000046), Fugene HD (Promega PROME2313), ECL Select (GE Healthcare RPN2235), BSA (Sigma-Aldrich A9418-100G), PNGase F (New England Biolabs 174P0704S), Endo H (New England Biolabs P0702), nitrocellulose membrane (GE Healthcare 10600003), FBS dialyzed (GE Healthcare HYCLSH30079.02), culture plates (Corning 734-1597), poly-D-lysine (Sigma-Aldrich P1024-100MG)

Cell line construction

Polyclonal knockouts of EMC6 were obtained by transfecting A549 cells with 3 plasmids pX330 encoding the respective on-target sgRNA, or non-targeting sgRNA together with a plasmid for puromycin selection using Fugene HD transfection reagent. 24h post transfection, cells were selected with 3ug/ml puromycin for 72h or until all control cells not receiving selection plasmid died. For A549 WT and EMC6 KO cells the cells were subjected to another round of transfection and selection to obtain a polyclonal population. To obtain oligoclonal MDCK WT and EMC6 KO cells, the cells were sorted into 96-well plates, and clones were selected following verification of protein depletion by western blotting. 10 WT and EMC6 KO clones

were combined to obtain an oligoclonal population. Rescue cell lines were obtained by lentiviral transduction with a pLEX either containing a FLAG-tag open reading frame, or N-terminally tagged FLAG.EMC6 and selected in hygromycin for 72h or until all control cells not receiving lentivirus died.

Viruses

A/Puerto Rico/8/34 (PR8) virus were obtained via reverse genetics plasmid system (de Wit et al., 2007). 8 plasmids were transfected into 293T cells, and supernatants were used to expand in MDCK cells. The final viral stock solutions were obtained by using the supernatants from infected MDCK cells to infect embryonated chicken eggs. A/Udorn/307/1972, MUd (PR8 + Udorn segment 7) and A/USSR/90/1977 were a gift from Prof. Paul Digard, Roslin Institute, UK.

Plasmids

Plasmids for SEC61B.H and SQS.H were a gift from John Christianson's lab [23]. HLA-A.F was amplified from human cDNA and cloned into pcDNA3 with a C-terminal Flag-tag and an AVI-tag following the endogenous signal peptide. M2 was amplified from cDNA generated from PR8 infected A549 cells and cloned into pcDNA3 with a C-terminal Flag tag. The L3N mutation generating the glycosylation site was introduced by quick change mutagenesis.

CRISPR guides

name	**target**	**sequence 5' to 3'**	**backbone**
sgEMC6.1px	EMC6	TTGTTCCGCGAGAAAGCGAG	PX330
sgEMC6.2px	EMC6	GCCGCCTCGCTGATGAACGG	PX330
sgEMC6.3px	EMC6	GAACGTCCAGAACAGGACGT	PX330
sgNT.1	non-targeting	ACGGAGGCTAAGCGTCGCAA	*in vitro* transcription
sgNT.2	non-targeting	CGCTTCCGCGGCCCGTTCAA	*in vitro* transcription
sgSRP54.1	SRP54	CAAAATGTGATTATGTTTGT	*in vitro* transcription
sgSRP54.2	SRP54	CACATTTTGTTTTCCTTTAG	*in vitro* transcription
sgSRPRA.1	SRPRA	ATTTCCAAAATGACTTCCTG	*in vitro* transcription
sgSRPRA.2	SRPRA	AATTGATAGATGACGTGCAT	*in vitro* transcription
sgSRPRB.1	SRPRB	GACTCGCGCCGGGTGGCAGA	*in vitro* transcription
sgSRPRB.2	SRPRB	AGCCCTACCTAGACACCTTG	*in vitro* transcription
sgSEC61A1.1	SEC61A1	TCCGGCAGGATGACACAGAA	*in vitro* transcription
sgSEC61A1.2	SEC61A1	CGGAAATTCAGAAGCCAGAG	*in vitro* transcription
sgSEC63.1	SEC63	TCCTTATGAAGTATTAAATT	*in vitro* transcription
sgSEC63.2	SEC63	AAATAGTTCTGCTTGCAGGA	*in vitro* transcription

sgEMC3.1	EMC3	GACCACCCAGAGGCGGATGT	*in vitro* transcription
sgEMC3.2	EMC3	GCAGCAGGATGGACACGTAG	*in vitro* transcription
sgEMC6.1	EMC6	GGTCCGGCAATAATCCAGGA	*in vitro* transcription
sgEMC6.2	EMC6	AGACGTGCACCATGCCGTAG	*in vitro* transcription
sgGET3.1	GET3	TGGCGGCAGGGGTGGCCGGG	*in vitro* transcription
sgGET3.2	GET3	TCGATGATGTTGCTAAGTGT	*in vitro* transcription
sgCAMLG.1	CAMLG	GCAGCGCATCAACCGGATCAT	*in vitro* transcription
sgCAMLG.2	CAMLG	GTCGCTACCGACGGCGGGGAG	*in vitro* transcription
sgTMCO1.1	TMCO1	AGACAAGTACAAGAGACTGA	*in vitro* transcription
sgTMCO1.2	TMCO1	ACAAAAGCCAATAGCAAACA	*in vitro* transcription

Transfection, western blots and deglycosylation

A549 WT or Δ6 cells were seeded in 24 well plates (6x10^4 cells/well) and transfected 16h later with plasmid (each 150ng/well) with Jetprime reagent and incubated for 20h. Cells were lysed in 70ul TX100 lysis buffer (50mM Tris-HCl pH 7.4, 150mM NaCl, 1% (v/v) Triton X-100, 5mM EDTA and protease inhibitors) on ice, scraped and transferred to 1.5ml test tubes and spun at 1000xg and 4°C. The supernatant was transferred to new tubes, mixed with 6x LDS sample buffer (550 mM Tris-HCl pH 7.4, 200 mM LDS, 0.13 mM EDTA, 250 mM DTT, 15% v/v glycerol, 0.025% SERVA Blue G250 and 0.025% Phenol Red) and to a 95°C hot plate for 3min. Samples were stored at -20°C until analysis by western blot.

If needed the 10ul of denatured samples were deglycosylated with NEB PNGaseF (0.3ul per reaction) or EndoH (0.3ul per reaction) as per manufacturer's description. 10% Acrylamide gels containing SDS and sucrose were loaded with samples and run in MOPS running buffer (50 mM MOPS, 50 mM TrisBase, 3.5 mM SDS, and 0.8mM EDTA) at 150V for 50min. Gel was transferred in Biorad transfer system onto nitrocellulose membranes, blocked in 5% milk in PBS + 0.1%TX100 and incubated in primary antibody overnight at 4°C or 1:10000 streptavidin-HRP for 1h at room temperature. 1:10000 secondary antibody incubation was 1h at room temperature. Chemiluminescent detection with Amersham AI600, or fluorescent antibody detection with Biorad Odyssee. Western blots were quantified by densitometry using FIJI ImageJ.

Viral infection with samples for flow cytometry, plaque assay and western blot

Cells plated the previous day were washed in PBS and infected with viral solutions in serum free DMEM at MOI of 3. At the intended time points, cell supernatants were collected and frozen at -80°C until use. For western blot samples cells were washed

in PBS, immediately lysed in 2x LDS sample buffer and stored at -20°C. For flow cytometry samples cells were detached in trypsin at 37°C for 7min and transferred to a 96 well plate with conical bottom to spin down at 300xg 4°C for 3min each spin. Cells were washed once in PBS + 1% FBS and once in PBS and fixed in PBS + 4% formaldehyde for 10min at room temperature. Cells were spun down at 500xg 4°C for 5min and washed twice in PBS and stored at 4°C for a maximum of 2 days. Cells were stained for 1h in primary antibody solution or PBS for unstained controls at room temperature. Cells were spun down and washed twice in PBS and stained with 1:1000 isotype specific secondary antibody solution Alexa 488 or Alexa 647 for 30min at room temperature. Cells were spun down and washed twice in PBS and measured at Fortessa X-20.

Flow cytometry analysis

Data acquired on Fortessa X-20 was gated for cells and single cells using FlowJo. Single cells gated data for the two measured channels M2 and HA was exported and further processed in python with custom scripts. Optimal parameters were fitted by maximum likelihood estimation for a mixture model of two skew-normal distributions using global optimization via differential evolution. Only the distributions with the larger mean at each time point were analysed further. The data was processed in two different ways. First, the data for each channel was normalized to the 15h time point for each protein in the WT cells and the mean at each time point was compared between WT and Δ6 cells. Second, each repeat for each protein and WT or Δ6 condition was fit to a logistic growth model using a differential evolution algorithm and normalized to the parameter L of the WT condition for each protein and repeat. The data was fit again to obtain the final optimized parameters of the normalized data.

Plaque assay

MDCK cells were seeded in 12 well plates to reach confluency the next day. Cells were washed in PBS and infected with 300ul of serially diluted supernatants in serum free DMEM. After 1h cells were washed in acid wash solution, washed in PBS and overlayed with serum free DMEM plus 0.14% BSA and AVICEL solution. Cells were incubated at 37°C + 5% CO2 for 30h.

Selective cas9 KO screen

First two sgRNAs against cellular targets were synthesized. Gene name of human targets: NT (non-targeting), SRP54, SRPRA, SRPRB, SEC61, SEC63, EMC3, EMC6, GET3, CAMLG, TMCO1, TMEM208. Genomic target sequences were designed with chopchop server. 2 primers were PCR amplified (25ul PCR reaction containing 0.5uM each) to build the double stranded DNA template for the *in vito* transcription using NEB Phusion polymerase. The forward primer contained a 5’ T7

promoter sequence (taatacgactcactatag, +G if target does not start with G), the 20 nucleotide specific genomic target sequence, and 20 nucleotides Spyro-cas9 sgRNA scaffold 3' of the genomic target sequence. The reverse primer was the reverse final nucleotides of Spyro-cas9 sgRNA scaffold overlapping 20 nucleotides with 3' end of the forward primer. *In vitro* transcription was performed with Promega T7 Ribomax Express as per manufacturers instructions in a 5ul reaction. RNA was purified on columns, eluted in RNAse free water and checked for purity and to obtain concentration on a 2% agarose gel containing 1% bleach.

To perform the screen 293T cells were seeded in two wells of a 6 well plate (3x10^5 cells/well) and transfected 8h later with cas9-BLAST plasmid (1ug/well) using Jetprime reagent. 24h after transfection cells were detached with trypsin, combined and seeded in a 24 well plate (10^5 cells/well) coated with poly-D-lysine. The next morning cells were transfected with *in vitro* transcribed purified sgRNA (2 sgRNAs per target, each 1ug sgRNA/well in 200ul total) with Dharmafect reagent and incubated for 8h. Cells were transfected in fresh medium with M2.Y.F plasmid (300 ng/well) using Jetprime reagent and incubated for 48h. Cells were lysed in TX100 lysis buffer, separated from nuclei, measured for protein concentration with Bradford assay and normalized to the lowest concentration.

2.9 References

1. Hegde, R. S. & Keenan, R. J. A unifying model for membrane protein biogenesis. *Nat Struct Mol Biol* **31**, 1009–1017 (2024).

2. Jomaa, A. *et al.* Mechanism of signal sequence handover from NAC to SRP on ribosomes during ER-protein targeting. *Science* **375**, 839–844 (2022).

3. Walter, P. & Blobel, G. Signal recognition particle contains a 7S RNA essential for protein translocation across the endoplasmic reticulum. *Nature* **299**, 691–698 (1982).

4. Lee, J. H. *et al.* Receptor compaction and GTPase rearrangement drive SRP-mediated cotranslational protein translocation into the ER. *Sci Adv* **7**, (2021).

5. Wang, H. & Hegde, R. S. Identification of a factor that accelerates substrate release from the signal recognition particle. *Science* **386**, 996–1003 (2024).

6. Siegel, V. & Walter, P. The affinity of signal recognition particle for presecretory proteins is dependent on nascent chain length. *EMBO J* **7**, 1769–1775 (1988).

7. Song, J. & Becker, T. Fidelity of organellar protein targeting. *Curr Opin Cell Biol* **75**, 102071 (2022).

8. Lamb, R. A., Zebedee, S. L. & Richardson, C. D. Influenza virus M2 protein is an integral membrane protein expressed on the infected-cell surface. *Cell* **40**, 627–633 (1985).

9. Chartron, J. W., Hunt, K. C. L. & Frydman, J. Cotranslational signal-independent SRP preloading during membrane targeting. *Nature* **536**, 224–228 (2016).

10. Wu, H. & Hegde, R. S. Mechanism of signal-anchor triage during early steps of membrane protein insertion. *Mol Cell* **83**, 961–973.e7 (2023).

11. Chen, Y. & Dalbey, R. E. Oxa1 Superfamily: New Members Found in the ER. *Trends Biochem Sci* **43**, 151–153 (2018).

12. Wu, X., Cabanos, C. & Rapoport, T. A. Structure of the post-translational protein translocation machinery of the ER membrane. *Nature* **566**, 136–139 (2019).

13. Aviram, N. *et al.* The SND proteins constitute an alternative targeting route to the endoplasmic reticulum. *Nature* **540**, 134–138 (2016).

14. Yang, J. *et al.* Human SND2 mediates ER targeting of GPI-anchored proteins with low hydrophobic GPI attachment signals. *FEBS Lett* **595**, 1542–1558 (2021).

15. Ramírez, A. S. & Locher, K. P. Structural and mechanistic studies of the N-glycosylation machinery: from lipid-linked oligosaccharide biosynthesis to glycan transfer. *Glycobiology* **33**, 861–872 (2023).

16. Spiess, M. & Lodish, H. F. An internal signal sequence: the asialoglycoprotein receptor membrane anchor. *Cell* **44**, 177–185 (1986).

17. Freeze, H. H. & Kranz, C. Endoglycosidase and glycoamidase release of N-linked glycans. *Curr Protoc Mol Biol* **Chapter 17**, Unit 17.13A (2010).

18. Dempster, J. M. *et al.* Chronos: a cell population dynamics model of CRISPR experiments that improves inference of gene fitness effects. *Genome Biol* **22**, 343 (2021).

19. Zuris, J. A. *et al.* Cationic lipid-mediated delivery of proteins enables efficient protein-based genome editing in vitro and in vivo. *Nat Biotechnol* **33**, 73–80 (2015).

20. Walter, P. & Blobel, G. Translocation of proteins across the endoplasmic reticulum. II. Signal recognition protein (SRP) mediates the selective binding to microsomal membranes of in-vitro-assembled polysomes synthesizing secretory protein. *J Cell Biol* **91**, 551–556 (1981).

21. Hull, J. D., Gilmore, R. & Lamb, R. A. Integration of a small integral membrane protein, M2, of influenza virus into the endoplasmic reticulum: analysis of the internal signal-anchor domain of a protein with an ectoplasmic NH2 terminus. *J Cell Biol* **106**, 1489–1498 (1988).

22. O'Keefe, S. *et al.* An alternative pathway for membrane protein biogenesis at the

endoplasmic reticulum. *Commun Biol* **4**, 828 (2021).

23. Guna, A., Volkmar, N., Christianson, J. C. & Hegde, R. S. The ER membrane protein complex is a transmembrane domain insertase. *Science* **359**, 470–473 (2018).

24. Page, K. R. *et al.* Role of a holo-insertase complex in the biogenesis of biophysically diverse ER membrane proteins. *Mol Cell* **84**, 3302–3319.e11 (2024).

25. Mueller, S. *et al.* Protein degradation corrects for imbalanced subunit stoichiometry in OST complex assembly. *Mol. Biol. Cell* **26**, 2596–2608 (2015).

26. Volkmar, N. *et al.* The ER membrane protein complex promotes biogenesis of sterol-related enzymes maintaining cholesterol homeostasis. *J. Cell Sci.* **132**, (2019).

27. Pleiner, T. *et al.* Structural basis for membrane insertion by the human ER membrane protein complex. *Science* **369**, 433–436 (2020).

28. Liu, X. *et al.* MARCH8 inhibits influenza A virus infection by targeting viral M2 protein for ubiquitination-dependent degradation in lysosomes. *Nat Commun* **12**, 4427 (2021).

29. Bhowmick, S., Chakravarty, C., Sellathamby, S. & Lal, S. K. The influenza A virus matrix protein 2 undergoes retrograde transport from the endoplasmic reticulum into the cytoplasm and bypasses cytoplasmic proteasomal degradation. *Arch Virol* **162**, 919–929 (2017).

30. Bogdanow, B. *et al.* The dynamic proteome of influenza A virus infection identifies M segment splicing as a host range determinant. *Nat Commun* **10**, 5518 (2019).

31. Martin-Sancho, L. *et al.* Restriction factor compendium for influenza A virus reveals a mechanism for evasion of autophagy. *Nat Microbiol* **6**, 1319–1333 (2021).

32. Han, Y.-J. *et al.* Targeting influenza A virus by splicing inhibitor herboxidiene reveals the importance of subtype-specific signatures around splice sites. *J Biomed Sci* **30**, 10 (2023).

33. Zebedee, S. L. & Lamb, R. A. Influenza A virus M2 protein: monoclonal antibody restriction of virus growth and detection of M2 in virions. *J Virol* **62**, 2762–2772 (1988).

34. McMahon, A. *et al.* High-throughput super-resolution analysis of influenza virus pleomorphism reveals insights into viral spatial organization. *PLoS Pathog* **19**, e1011484 (2023).

35. Rossman, J. S., Jing, X., Leser, G. P. & Lamb, R. A. Influenza virus M2 protein mediates ESCRT-independent membrane scission. *Cell* **142**, 902–913 (2010).

36. Lamb, R. A. The Structure, Function, and Pathobiology of the Influenza A and B Virus Ion Channels. *Cold Spring Harb Perspect Med* **10**, (2020).

37. Abell, B. M., Pool, M. R., Schlenker, O., Sinning, I. & High, S. Signal recognition particle mediates post-translational targeting in eukaryotes. *EMBO J* **23**, 2755–2764 (2004).

38. Shan, S.-O. Role of Hsp70 in Post-Translational Protein Targeting: Tail-Anchored Membrane Proteins and Beyond. *International Journal of Molecular Sciences* **24**, 1170 (2023).

39. Hegde, R. S. The Function, Structure, and Origins of the ER Membrane Protein Complex. *Annu Rev Biochem* **91**, 651–678 (2022).

40. Hann, B. C. & Walter, P. The signal recognition particle in S. cerevisiae. *Cell* **67**, 131–144 (1991).

41. Guna, A. *et al.* A dual sgRNA library design to probe genetic modifiers using genome-wide CRISPRi screens. *BMC Genomics* **24**, 651 (2023).

42. Zhang, J. *et al.* Computing the Human Interactome. *bioRxiv* (2024) doi:10.1101/2024.10.01.615885.

43. Zhu, Y. *et al.* Human TRA2A determines influenza A virus host adaptation by regulating viral mRNA splicing. *Sci Adv* **6**, eaaz5764 (2020).

44. Calderon, B. M. *et al.* Dysregulation of M segment gene expression contributes to influenza A virus host restriction. *PLoS Pathog* **15**, e1007892 (2019).

45. Hutchinson, E. C., Curran, M. D., Read, E. K., Gog, J. R. & Digard, P. Mutational analysis of cis-acting RNA signals in segment 7 of influenza A virus. *J Virol* **82**, 11869–11879 (2008).

46. McCown, M. F. & Pekosz, A. The influenza A virus M2 cytoplasmic tail is required for infectious virus production and efficient genome packaging. *J Virol* **79**, 3595–3605 (2005).

47. Tsai, P.-L. *et al.* Cellular RNA binding proteins NS1-BP and hnRNP K regulate influenza A virus RNA splicing. *PLoS Pathog* **9**, e1003460 (2013).

48. Kolpe, A. *et al.* Super-resolution microscopy reveals significant impact of M2e-specific monoclonal antibodies on influenza A virus filament formation at the host cell surface. *Sci Rep* **9**, 4450 (2019).

49. Roberts, K. L., Manicassamy, B. & Lamb, R. A. Influenza A virus uses intercellular connections to spread to neighboring cells. *J Virol* **89**, 1537–1549 (2015).

50. Roberts, P. C. & Compans, R. W. Host cell dependence of viral morphology. *Proc Natl Acad Sci U S A* **95**, 5746–5751 (1998).

51. Takeda, M., Pekosz, A., Shuck, K., Pinto, L. H. & Lamb, R. A. Influenza a virus

M2 ion channel activity is essential for efficient replication in tissue culture. *J Virol* **76**, 1391–1399 (2002).

52. Ulferts, R. *et al.* Subtractive CRISPR screen identifies the ATG16L1/vacuolar ATPase axis as required for non-canonical LC3 lipidation. *Cell Rep* **37**, 109899 (2021).

53. Ichinohe, T., Pang, I. K. & Iwasaki, A. Influenza virus activates inflammasomes via its intracellular M2 ion channel. *Nat Immunol* **11**, 404–410 (2010).

54. Huang, Y., Urban, C., Hubel, P., Stukalov, A. & Pichlmair, A. Protein turnover regulation is critical for influenza A virus infection. *Cell Syst* **15**, 911–929.e8 (2024).

RESULTS CHAPTER 3: Nascent chain features of EMC dependence

3.1 Chapter Summary

The cis-acting membrane protein features that influence the targeting dynamics of a protein, and the subsequent selection of a translocon at the ER membrane, are incompletely understood. Although much is known about generic signals of membrane protein targeting and topogenesis, we lack a detailed understanding of sequence features that narrow or broaden the repertoire of possible handover reactions of a membrane protein from the cytosol to the ER membrane. In the previous chapter we report that the loss of EMC largely affects M2 accumulation in the cell and M2's rate of accumulation at the plasma membrane. In this chapter we explore M2 sequence features that make M2 EMC dependent. M2 is potentially cotranslationally targeted to the ER membrane, and may require special factors to mediate membrane insertion due to the presence of hydrophilic residues within its transmembrane domain. We identify SRPRB as an EMC independent typeIII membrane protein, which we use to assess M2's features of EMC dependence by constructing protein domain chimeras between M2 and SRPRB. We find that M2's TMD and C-terminus carry signals of EMC dependence, but that C-terminal and N-terminal modifications can overwrite these signals to become EMC independent. We confirm that the typeIII protein topology has an unmet flexibility in the selection of a translocation pathway and that the targeting dynamics matter for a selection of a translocon. Ultimately, our research will support the identification of biogenesis factors from sequence, which could be a major asset in the development of pan-viral therapeutics.

3.2 Introduction

For most membrane proteins, the specific biogenesis pathway starts with the emergence of a hydrophobic stretch of amino acids from the ribosomal exit tunnel. In this chapter, we focus on cis-acting nascent chain encoded signals of M2 protein that influence its dependence on EMC processing at the ER.

Membrane proteins with a single transmembrane domain, as well as the first transmembrane domain of multipass membrane proteins have been historically

categorised into four types (Type I - IV, Figure 3.1A). The type is based on their final orientation in the membrane (also called topology), the presence of an N-terminal cleavable hydrophobic signal peptide, and the location of the transmembrane domain relative to the length of the sequence. Based on these observations many mechanistic studies have been conducted in order to describe signals that would influence topology, such as transmembrane domain flanking charges, folding status and length of transmembrane flanking domains, transmembrane domain length, or transmembrane domain hydrophobicity [1,2]. These signals either recruit different machinery for membrane targeting, or directly interact with the ER membrane and its protein components to be directed to the right machinery and bias the insertion topology [1,2].

Membrane protein biogenesis facilitates the insertion of a membrane protein in the correct topology, but also serves the current metabolic requirements of the cell. Signal peptides and transmembrane domains have a reduced amino acid repertoire due to the enrichment of hydrophobic residues, but the diversity of existing sequences has been shown to regulate the efficiency of insertion [3], or triage post-translationally targeted proteins for the translocation machinery, either targeting the GET pathway or EMC [4]. Results described in Chapter 2 show that M2 can traffic to the plasma membrane in absence of EMC, albeit at a lower rate. M2 might be promiscuous towards other members of the Oxa1 superfamily of insertases, despite its rather hydrophilic transmembrane domain which should require special handling at the ER. Thus M2 sequence might harbour interesting insights into signals for redundant processing at the ER.

3.3 Comparison of viral and human single transmembrane proteins

Transmembrane domain hydrophobicity and C-terminal length have been previously described as potential signals that make a typeIII protein an EMC client [5,6]. We compared M2 to the human proteome and viral proteomes of genera which contain viruses that infect humans.

To do so, we generated a dataset of predicted single pass membrane proteins of viruses infecting humans (1226 proteins of 65 genera) and of predicted single pass membrane proteins of the human proteome (2083 proteins). We manually annotate the transmembrane type according to the predicted topology and position of the transmembrane domain relative to the protein length. Within this dataset, we find that 22 genera express 93 unique typeIII protein, namely Alphainfluenzavirus,

Alphapapillomavirus, Alphavirus, Betacoronavirus, Betainfluenzavirus, Centapoxvirus, Cytomegalovirus, Jeilongvirus, Lentivirus, Marseillevirus, Mastadenovirus, Molluscipoxvirus, Orthopneumovirus, Orthopoxvirus, Orthoreovirus, Parapoxvirus, Paslahepevirus, Pegivirus, Rhadinovirus, Rotavirus, Tibrovirus, Yatapoxvirus.

We find no difference in the amino acid composition of viral and human transmembrane domains (Figure 3.1B) and overall hydrophobicity, calculated by an experimentally derived hydrophobicity scale in a 19 amino acid window along the protein sequence ΔGapp in kcal/mol [7] (Figure 3.1C). However, viral proteins have shorter C-terminal lengths compared to human proteins. M2 is rather hydrophilic and has a very short C-terminal length compared to viral and human proteins (Figure 3.1B and C, red dashed line). M2's transmembrane domain 11 amino acid core $_{29}$AA**N**IIGIL**H**LI$_{39}$ has an asparagine (N) and histidine (H), which are very rare in the transmembrane domain core of single pass membrane proteins in general (Figure 3.1B). We conclude that M2 has a hydrophilic transmembrane domain and a short C-terminus relative to most human proteins and viral proteins of genera that infect humans.

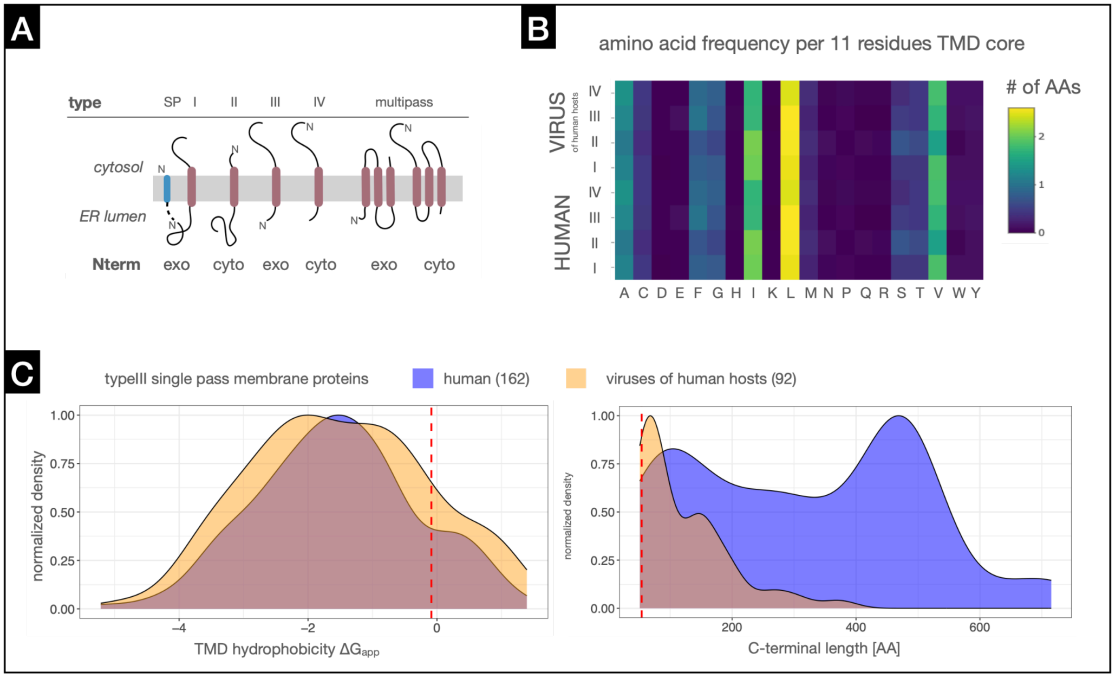


Figure 3.1 Comparing M2 sequence characteristics to human and viral proteomes. Database of viral single pass membrane proteins of different hosts and their frequency by type and host category. 25613 viral proteins of 115 genera with viruses of human hosts were acquired from the of the NCBI refSeq database and clustered into 10715 similar proteins with MMSeq2. Sequences were clustered if they at least had 60% sequence identity, and covered more than 80% of an aligned sequence in the cluster (--min-seq-id 0.6 -c 0.8 --cov-mode 1).

The TMBed language model was used to predict transmembrane domains, their topology, and signal peptides, and predictions were refined with ΔGapp calculations (< 1.5 kcal/mol, higher energy means less favorable to insert into membrane), resulting in a dataset of 1226 single pass membrane proteins. Protein types were manually annotated based on signal peptide presence, predicted topology, position of the transmembrane domain and length of translocated domain. A) Schematic of membrane protein topologies. B) Comparison of transmembrane domain amino acid composition (11 amino acids around the ΔGapp minimum), by type of single pass membrane proteins of the human proteome and viral proteomes of human hosts. C) Comparison of human and viral typeIII membrane proteome in TMD ΔGapp and Cterminal length (red dotted line are values for M2).

3.4 M2's transmembrane domain and C-terminus contain signals of EMC dependence

M2's sequence has notable characteristics which could act as signals during its biogenesis. To identify which domains in M2 render it EMC dependent, we searched the human proteome for a typeIII proteins that differ in one or both characteristics to find a protein which is EMC independent. Swapping the domains between M2 and an EMC independent typeIII protein would lead us to isolate domains that are important for EMC processing. Searching the database, we found TEMP and SRPRB as suitable candidate proteins. TEMP has a similar C-terminal length to M2 (69 amino acids), similarly unstructured terminal domains, a natural glycosylation site at its N-terminus and importantly, a hydrophobic transmembrane domain (ΔGapp -1.48 kcal/mol). TEMPs function in the cell is not well defined. SRPRB has been selected for its hydrophobic transmembrane domain (ΔGapp -1.8 kcal/mol) and its long C-terminal domain (226 amino acids). SRPRB is an important component of the eukaryotic co-translational targeting machinery, but has not formally been investigated for its biogenesis requirements, and is thus highly interesting. We engineered a glycosylation site at SRPRB's N-terminus and fused a FLAG tag to SRPRB's and TEMP's C-terminus for easier detection in a western blot (constructs TEMP.Y.F and SRPRB.Y.F).

We transfect the constructs into A549 WT and Δ6 cells to observe their glycosylation and the overall amount of PNGase treated lysates (Figure 3.2B). As before M2 levels and glycosylation is drastically reduced. We find that TEMP.Y.F's glycosylation is also affected. TEMP.Y.F seems to undergo further post translational modifications as judged by the 2 bands that present in PNGase treated conditions. Interestingly, we can observe a strong unglycosylated band persisting in Δ6 conditions. Lastly, we find that SRPRB glycosylation and overall amounts are not

affected by the loss of EMC (Figure 3.2B). We conclude that SRPRB is EMC independent and thus select SRPRB to construct domain chimeras with M2.

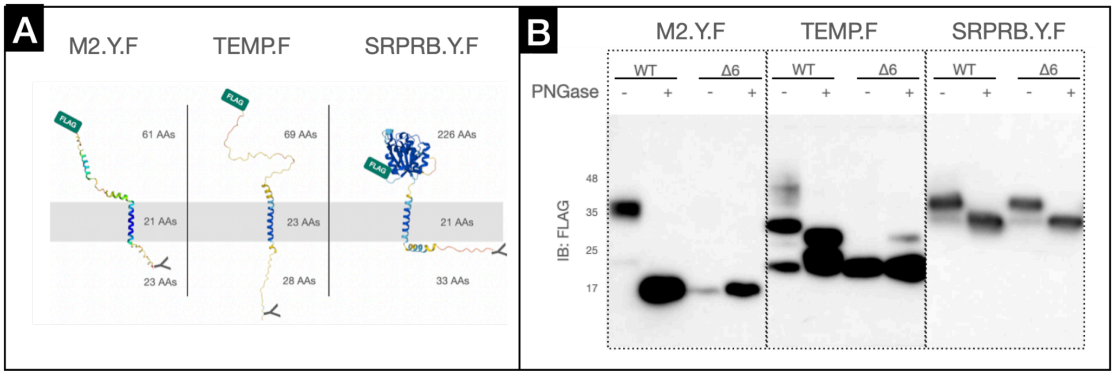


Figure 3.2 Selection of typeIII proteins to test for EMC dependence

A) Alphafold2 [8] predicted structures of M2, TEMP and SRPRB, their domain lengths and their topology in the membrane. B) A549 WT and Δ6 cells were transfected with M2.Y.F, TEMP.F or SRPRB.Y.F and cell lysates were mock treated or treated with PNGase and analysed by western blot (N=2).

In order to test if any of M2's domains carry signals for EMC dependence we assess the glycosylation efficiency (+Y/SUM(+Y,-Y)) and the total protein levels in a transfection assay of the chimeric proteins. We selected the stretch of amino acids containing the transmembrane helix plus 3 amino acids for the flanking charges (topology determining elements) as the transmembrane domain of M2 and SRPRB and created chimaeras swapping the N-terminus, transmembrane domain and C-terminus between them. We transfect the resulting panel of chimeras into A549 WT and Δ6 cells to monitor the glycosylation efficiency and total levels at steady state. First, we tested if a single domain of M2 can transfer EMC dependence onto SRPRB. SRPRB and MS1 (M2 N-terminus onto SRPRB) remain EMC independent, as judged by the glycosylation efficiency and overall stability. MS2 (M2 TMD) and MS3 (M2 C-terminus) become EMC dependent as measured by the reduction in glycosylation efficiency. However, despite the defect in glycosylation efficiency, only MS2 is reduced in total protein levels but MS3 is stable. These results show that M2 TMD and M2 C-terminus carry signals of EMC dependence, and the reduction in glycosylation efficiency hints at a biogenesis defect. The other way around, with chimeras MS4, MS5 and MS6 we can test if we can make M2 EMC independent by exchanging one of M2's domains with SRPRB signals. Confirming that indeed M2's TMD and C-terminus both carry signals of EMC dependence, none of these chimeras becomes independent of EMC judged by the overall amount of protein being lost in similar amounts to the WT M2 sequence (Figure 3.3). However, we can observe that

MS4 restores the apparent glycosylation efficiency. At this point we cannot judge if a non-glycosylated band is unstable and degraded in this construct before membrane insertion takes place, or if the protein product is efficiently inserted, but unstable in the membrane (Figure 3.4).

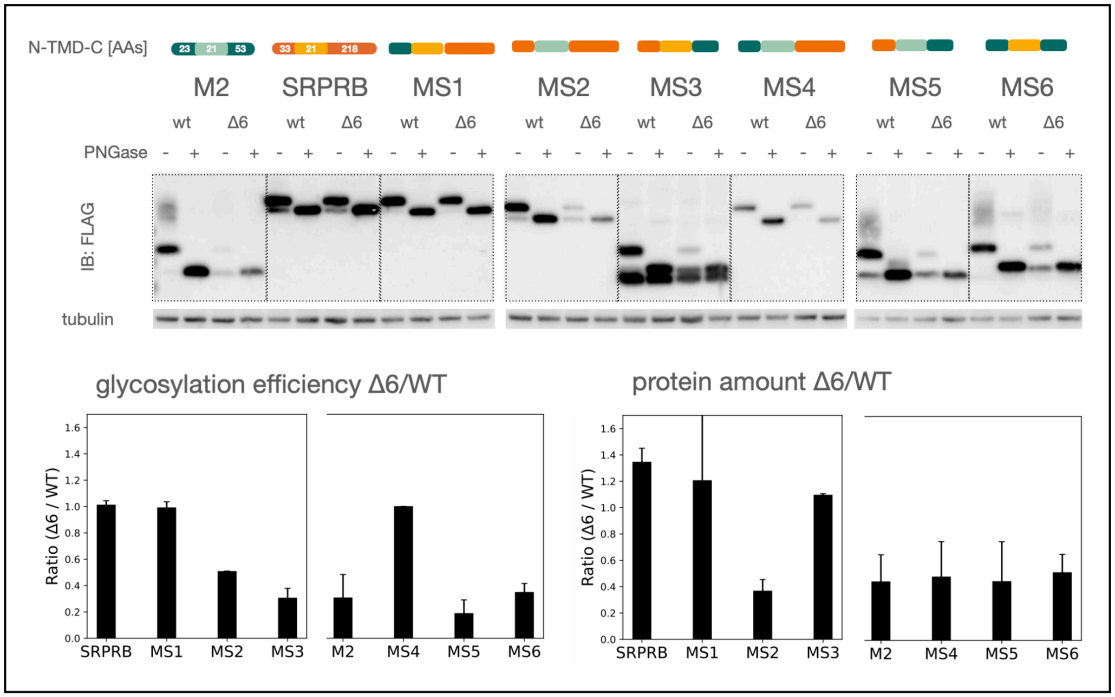


Figure 3.3 Panel of M2 SRPRB domain chimaeras to test for signals of EMC dependence

A549 WT or Δ6 cells were transfected with each of the constructs and cell lysastes were mock treated or treated with PNGase and analysed by western blot. The glycosylation efficiency (+Y/SUM(+Y,-Y)) and total amounts were analysed in imageJ by densitometry (N=3).

Previous research found C-terminal length to affect EMC dependence of SYT1, a typeIII protein involved in synaptic vesicle fusion [6]. We wanted to test if C-terminal length M2 might influence its EMC dependence. For this, we constructed M2 mutants where we concatenate the same C-terminal 37 amino acids (residue 60 to 97, conserving the same nucleotide sequence) to achieve a longer C-terminus that conserves potential motifs (Figure 3.4A). Astonishingly, we can observe that C-terminal extension of M2 progressively gains glycosylation efficiency (Figure 3.4B). The gain in EMC independence due to C-terminal extension might be due to the elongation of the mRNA, however the EMC dependence of MS2 and MS4 would argue against mRNA length playing a role.

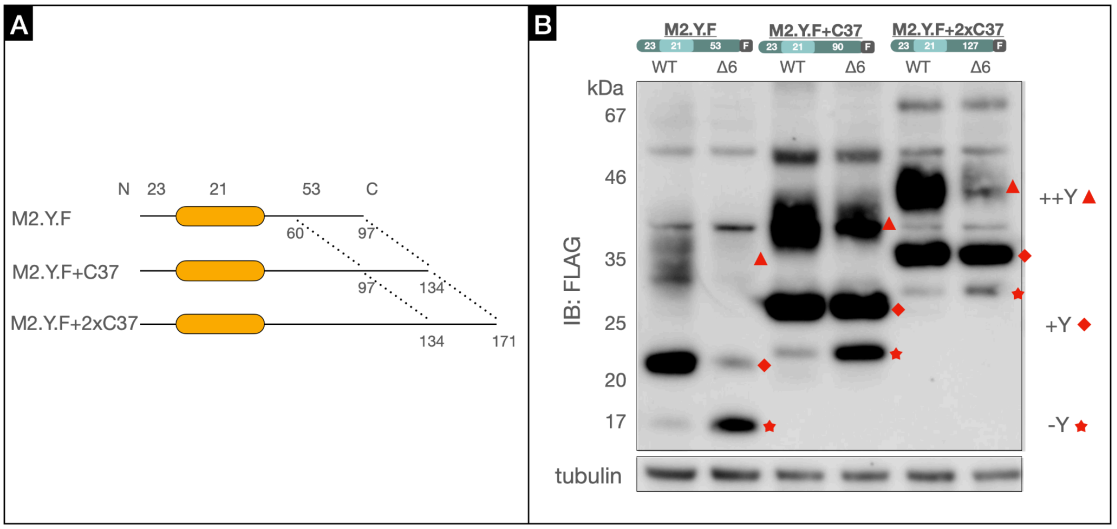


Figure 3.4 M2 C-terminal concatenation mutants restore glycosylation efficiency in absence of EMC

A549 WT or Δ6 cells were transfected with each of the constructs and analysed by western blot (N=3 with 2 repeats from Figure 3.5B)

Sec61 has been described to be able to engage typeIII proteins for translocation. It may be possible that C-terminal extension mutants have more time during targeting and can access the Sec61 translocon co-translationally. We can test this hypothesis by inhibiting Sec61 with the strong lateral gate inhibitor Ipomoeassin F [9]. Low concentrations of the inhibitor in the medium completely abrogate typeI protein expression. Transfection of IAV HA shows a dramatic decrease in levels in presence of Ipomoeassin F (Figure 3.5A). Inhibiting Sec61 in WT cells did not have a strong effect on M2 or the C-terminal extension mutants. Inhibiting Sec61 in absence of EMC had a small effect on the glycosylation efficiency of M2.Y.F+C37 and M2.Y.F+C74, which may be a direct or indirect effect (Figure 3.5B). We additionally tested the effects of the double inhibition of Sec61 and EMC on SRPRB with no apparent effect on glycosylation efficiency. It remains uncertain which ER translocon mediates membrane insertion of these typeIII proteins. Other members of the Oxa1 family at the ER may play a role in translocating these M2 constructs or SRPRB. Interestingly, the hyperglycosylated species (++Y) of M2 and the C-terminal extension mutants seem to be affected by EMC6 KO at steady state, hinting at a defect subsequent to ER insertion along the trafficking route. The effect of Sec61 inhibition on ++Y size is difficult to rationalise. Sec61 inhibition results in the depletion of many cellular proteins that might be important for glycan modifications, or M2 trafficking.

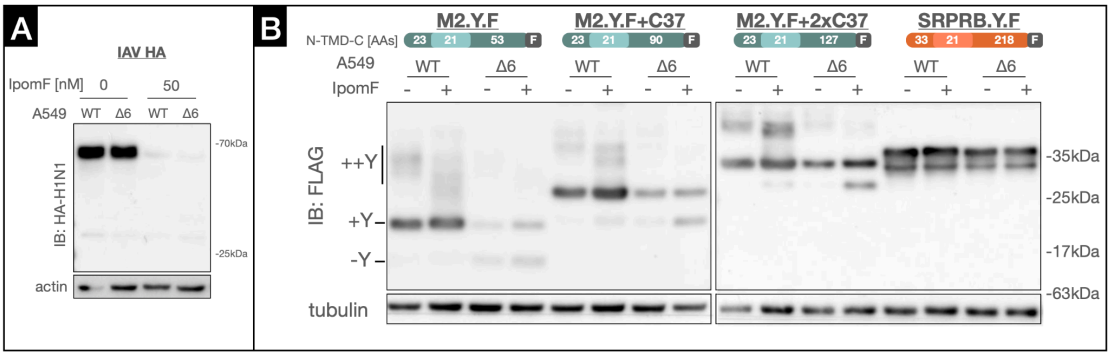


Figure 3.5 M2 C-terminal extension mutants become partially Sec61 dependent in the absence of EMC

A) A549 WT or Δ6 cells were transfected with HA from IAV PR8 and incubated with 50nM Ipomoeassin F or equivalent volume fraction of DMSO and lysates were analysed by western blot (N=1). B) A549 WT or Δ6 cells were transfected with M2 C-terminal extension mutants or SRPRB and incubated with 50nM Ipomoeassin F or equivalent volume fraction of DMSO and lysates were analysed by western blot (N=3).

3.5 M2 N-terminal mutations can mitigate the loss of EMC

Although the SRPRB and M2 chimeras did not show any effect of N-terminal replacement on EMC dependence we decided to test if the N-terminal fusion of M2 with a signal peptide, effectively converting M2 into a typeI protein, could mitigate the biogenesis defect of EMC loss. For this purpose we tested two different signal peptides, one from calreticulin (SP.M2.Y.F), an ER resident luminal protein, which can also be found in the cytosol, and the signal peptide of IAV's own typeI protein HA (HA-SP.M2.Y.F). Interestingly, we find that both signal peptide fusion constructs can restore glycosylation efficiency (Figure 3.6A and 3.6B), however only the signal peptide of calreticulin rescues the protein amount (Figure 3.6A). This result hints at EMC being important for insertion, however, that a signal sequence may carry additional information that may stabilize the protein or lead to its degradation. Surprisingly, we find that SP.M2.Y.F produces a stable unglycosylated band which has been cleaved from its signal peptide running at the same height as unglycosylated M2 in an acrylamide gel (Figure 3.6A). This means that SP.M2.Y.F has entered the ER, is processed to remove the signal peptide, but is not glycosylated. At this point it is unclear if the unglycosylated product of SP.M2.Y.F at steady state is inserted into a membrane or has been retrotranslocated and is stable in the cytosol. Very unexpectedly, Sec61 inhibition has no effect on the glycosylation efficiency of SP.M2.Y.F (Figure 3.6C), describing a signal peptide dependent, but Sec61 independent pathway that has not been reported before.

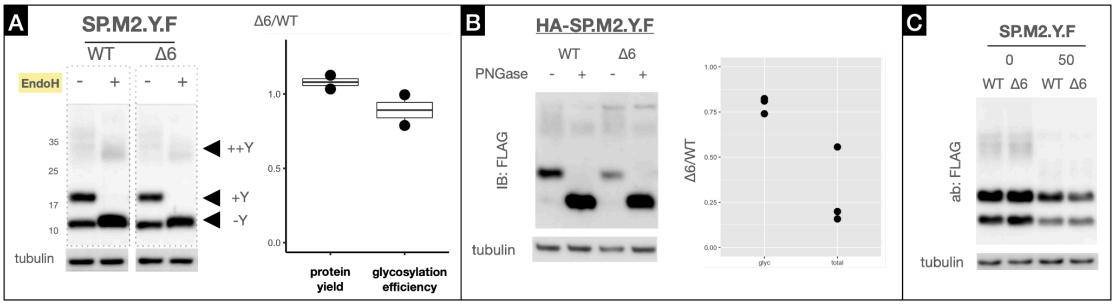


Figure 3.6 Signal Peptide at M2's N-terminus can restore glycosylation efficiency in absence of EMC

A) A549 WT or Δ6 cells were transfected with SP.M2.Y.F and lysates were mock treated or treated with EndoH and analysed by western blot. Glycosylation efficiency and total amounts were quantified by densitometry in imageJ (N=2). B) A549 WT or Δ6 cells were transfected with HA-SP.M2.Y.F and lysates were mock treated or treated with PNGase and analysed by western blot. Glycosylation efficiency and total amounts were quantified by densitometry in imageJ (N=3). C) A549 WT or Δ6 cells were transfected with SP.M2.Y.F and incubated with 50nM Ipomoeassin F or equivalent volume fraction of DMSO and lysates were analysed by western blot (N=3).

Finally, we stumble upon N-terminal M2 mutants that modify EMC dependence without the sequence of a classical signal peptide. The substitution of the N-terminus with a sequence of similar length (AVI-tag + HA-tag), or the partial substitution of the first 15 amino acids of the N-terminus (AVI-tag) result in an EMC independent construct (Figure 3.7A). Although the AVItag and HA sequences are highly negatively charged the AVI.MMM sequence has conserved transmembrane domain flanking charges with the M2 WT construct (Figure 3.7B). How the very N-terminal sequence might determine translocon dependence is completely unknown at this point. Although our results lack the power to determine the exact sequence features of EMC dependence, we can show that even the very N-terminal sequence of a typeIII protein can overwrite the EMC dependent signals present in the transmembrane domain and C-terminus.

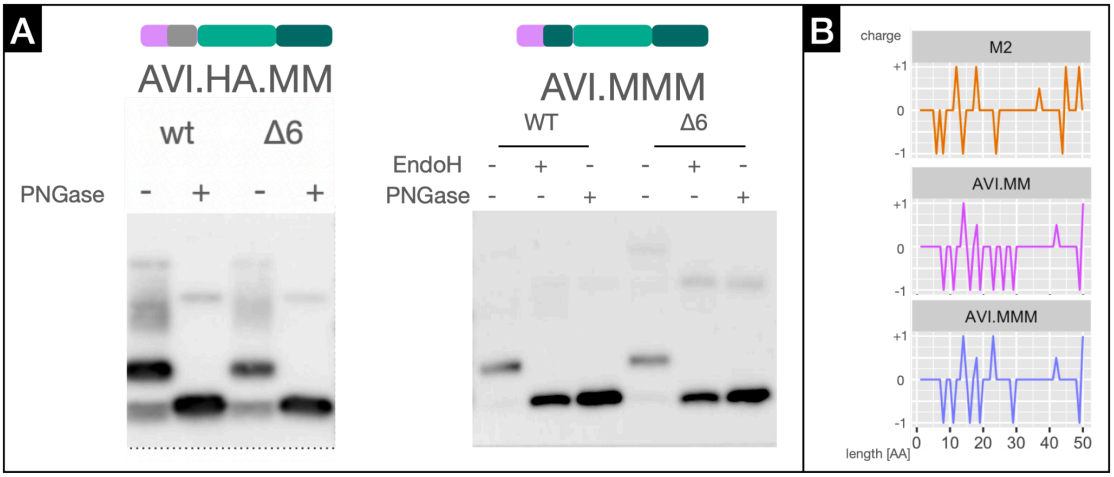


Figure 3.7 N-terminal mutants affect EMC dependence

A) A549 WT or Δ6 cells were transfected with the indicated construct (AVI-tag pink, HA-tag grey, M2 sequence green). Cell lysates were mock treated or treated with PNGase, or EndoH and analysed by western blot (N=1). B) Charges along the sequences were visualized giving E and D a negative charge of 1 and K and R a positive charge of 1, and H a positive charge of 0.5.

3.6 Discussion

Nascent chain features that determine biogenesis pathways are complex and can be found in many aspects of their chemical and physical characteristics. Here we show that viral and human typeIII proteins differ most significantly in the length of the C-terminal domain. This is interesting as our results hint at C-terminal length of typeIII proteins playing a role for EMC dependence, which might make EMC function important for many viral replication cycles. Beyond these candidates, the EMC has been implicated in the replication cycles of flaviviruses, either in its canonical translocation function for the flavivirus polyprotein [10], or in a process where EMC components seem to be repurposed during dengue virus entry [11].

Transmembrane domain hydrophobicity has been found initially to triage post-translationally inserted typeIV proteins for either the GET pathway or EMC pathway [12], however, this may depend on access to the targeting chaperone rather than an a direct selection by a translocon. The choice of a cytosolic chaperone for membrane targeting is dependent on the hydrophobicity of the transmembrane domain. Very hydrophobic transmembrane domains of typeIV proteins are captured by GET3 to be inserted by the GET pathway and hydrophilic transmembrane domains are captured by SGTA for EMC translocation [13].

M2 harbours unconventional amino acids within its transmembrane domain. Asparagines are rare in the transmembrane core of single pass membrane proteins.

Interestingly, this specific residue is commonly a Serine in other IAVs. The substitution can be traced back to have emerged as a resistance mutation to the first antiviral drug amantadine, targeting M2's ion channel function [14].

However, we find that transmembrane domain hydrophobicity is not the sole determining factor of M2 biogenesis pathway. TEMP has a conventionally hydrophobic transmembrane domain, but both M2 and TEMP are dependent on EMC processing. By constructing domain chimeras with SRPRB we find that M2's C-terminus carries features of EMC dependence. We show that C-terminal length to be important by concatenating M2's C-terminus onto itself. The C-terminal extension, rescues glycosylation efficiency, but is still partially defective in accumulating hyperglycosylated M2. The domain chimeras MS2 and MS4 both are constructs that carry M2's TMD, but SRPRB's long C-terminal domain. These constructs do show a defect in expression levels in absence of EMC, but no increase in an unglycosylated band. This may be due to very efficient degradation of uninserted product, or a defect subsequent to membrane insertion. These results show that C-terminal signals can act independently of TMD signals for EMC dependence, maybe during different stages of M2 biogenesis.

Possibly the way M2 is targeted to the ER membrane seems to have an effect on EMC translocation dependence. We see that by providing a strong ER targeting signal with an N-terminal signal peptide we can rescue glycosylation efficiency in the absence of EMC, without making the protein Sec61 dependent. The signal peptide mutants and the C-terminal extension mutants could function in a similar manner, effectively extending the time of targeting after SRP recognition in the cytosol. It is very surprising that these mutants do not become inhibited by Ipomoeassin F, but some function of the signal peptide seems to be enough to regain glycosylation efficiency. Furthermore, the large unglycosylated band of signal peptide cleaved SP.M2 could be evidence for retrotranslocation. We do not know at this point if the unglycosylated band is inserted into a membrane or stable in the cytosol. These results hint at EMC having a role in efficient protein targeting, or being able to process membrane proteins where cotranslational targeting time is critically short, which could be the case for M2 with a relatively short C-terminus of 56 amino acids. C-terminal extension mutants of M2 have been investigated before. The fusion of a large ER translocated domain (~400 amino acids) partially inverted M2 topology, an effect which we could not reproduce here [15]. Additional features of C-terminal domains might play a role. SRPRB has a globular C-terminal domain, while the concatenated M2 C-terminal domains are predicted to be alpha helical or disordered.

Lastly we show that N-terminal sequence matters and that signals present in the very N-terminus, more than 10 amino acids away from the beginning of the transmembrane domain can affect the EMC dependence of M2. N-terminal modifying enzymes that act at the ribosomal exit tunnel have recently been implicated in having a role in membrane protein biogenesis in bacteria [16]. Similar N-terminal modifications take place in eukaryotes [17–19]. Further studies need to be conducted to assess which role protein modification at the ribosomal exit tunnel like N-terminal acetylation plays in ER membrane targeting and translocation.

3.7 Methods

Cell lines

A549 (ATCC CCL-185), were cultured in Dulbecco's Modified Eagle Medium (Gibco Thermo Fisher 21969035) supplemented with 10% fetal bovine serum (Thermo Fisher A5256801), 2 mM L-glutamine (Thermo Fisher 25030024), and 1% (v/v) penicillin-streptomycin (Biowest L0022-100) in a humidified incubator at 37°C and 5% CO2.

Plasmids

TEMP and SRPRB were amplified from human and mouse cDNA respectively and cloned into pcDNA3 with a C-terminal Flag-tag using KpnI and XhoI restriction sites. The glycosylation site at SRPRB's N-terminus was introduced so that the protein sequence starts with MNGT. Domain chimeras between SRPRB and M2 were achieved by overlap-PCR and cloned into pcDNA3. The M2 C-terminal extension mutants were cloned with a strategy where one of the restriction sites was destroyed after cloning using type IIS restriction enzymes to successfully introduce first one extension then two extensions. SP.M2.Y.F and HA-SP.M2.Y.F were cloned by adding the calreticulin or IAV haemagglutinin signal peptide to M2.Y.F by overlap-PCR. Similarly, AVI.MM and AVI.MMM were constructed by adding AVI-tag and HA-tag, replacing the entire M2 N-terminus, or just adding AVI-Tag replacing the terminal 15 amino acids by overlap-PCR. IAV-HA was amplified from the reverse genetics system and cloned into pLEX.

Computational analysis of human and viral proteomes

The NCBI virus refSeq protein database (downloaded on 18.02.2025) contained 686474 protein entries. All known virus genera were obtained from ICTV (ICTV_Master_Species_List_2023_MSL39.v4) and filtered for those genera infecting humans (115 genera), by filtering for genera of virus species known to infect humans from virus-host-db [20] (downloaded on 11.11.2024). Phages were removed from the

dataset. RefSeq entries were filtered for 115 genera to obtain 25613 protein sequences. These sequences were clustered into 10715 similar proteins with MMSeq2 [21]. Sequences were clustered if they at least had 60% sequence identity, and covered more than 80% of an aligned sequence in the cluster (--min-seq-id 0.6 -c 0.8 --cov-mode 1). The TMBed language model [22] was used to predict transmembrane domains, their topology, and signal peptides, and transmembrane domain predictions were refined with ΔGapp calculations [7] in a 19 residue window. Transmembrane domains were considered when the location coincided with a minimum in the ΔGapp calculation and was lower than 1.5 kcal/mol, resulting in a dataset of 1226 single pass membrane proteins. Protein types were manually annotated based on signal peptide presence, predicted topology, position of the transmembrane domain and length of translocated domain (typeII N-terminal length < 150 and C-terminal length > 60; typeIII N-terminal length < 60 and C-terminal length > 60; typeIV C-terminal length < 60; other are proteins not fitting into any of these categories).

Transfection, western blots and deglycosylation

A549 WT or Δ6 cells were seeded in 24 well plates (6x10^4 cells/well) and transfected 16h later with plasmid (each 150ng/well) with Jetprime reagent and incubated for 20h. Cells were lysed in 70ul TX100 lysis buffer (50mM Tris-HCl pH 7.4, 150mM NaCl, 1% (v/v) Triton X-100, 5mM EDTA and protease inhibitors) on ice, scraped and transferred to 1.5ml test tubes and spun at 1000xg and 4°C. The supernatant was transferred to new tubes, mixed with 6x LDS sample buffer (550 mM Tris-HCl pH 7.4, 200 mM LDS, 0.13 mM EDTA, 250 mM DTT, 15% v/v glycerol, 0.025% SERVA Blue G250 and 0.025% Phenol Red) and to a 95°C hot plate for 3min. Samples were stored at -20°C until analysis by western blot.

If needed the 10ul of denatured samples were deglycosylated with NEB PNGaseF (0.3ul per reaction) or EndoH (0.3ul per reaction) as per manufacturer's description. 10% Acrylamide gels containing SDS and sucrose were loaded with samples and run in MOPS running buffer (50 mM MOPS, 50 mM TrisBase, 3.5 mM SDS, and 0.8mM EDTA) at 150V for 50min. Gel was transferred in Biorad transfer system onto nitrocellulose membranes, blocked in 5% milk in PBS + 0.1%TX100 and incubated in primary antibody overnight at 4°C or 1:10000 streptavidin-HRP for 1h at room temperature. 1:10000 secondary antibody incubation was 1h at room temperature. Chemiluminescent detection with Amersham AI600, or fluorescent antibody detection with Biorad Odyssee. Western blots were quantified by densitometry using FIJI ImageJ.

3.8 References

1. Higy, M., Junne, T. & Spiess, M. Topogenesis of membrane proteins at the endoplasmic reticulum. *Biochemistry* **43**, 12716–12722 (2004).

2. Hegde, R. S. & Keenan, R. J. A unifying model for membrane protein biogenesis. *Nat. Struct. Mol. Biol.* **31**, 1009–1017 (2024).

3. Hegde, R. S. & Bernstein, H. D. The surprising complexity of signal sequences. *Trends Biochem. Sci.* **31**, 563–571 (2006).

4. Guna, A., Hazu, M., Pinton Tomaleri, G. & Voorhees, R. M. A TAIe of two pathways: Tail-anchored protein insertion at the endoplasmic reticulum. *Cold Spring Harb. Perspect. Biol.* **15**, a041252 (2023).

5. Chitwood, P. J., Juszkiewicz, S., Guna, A., Shao, S. & Hegde, R. S. EMC Is Required to Initiate Accurate Membrane Protein Topogenesis. *Cell* **175**, 1507–1519.e16 (2018).

6. O'Keefe, S. *et al.* An alternative pathway for membrane protein biogenesis at the endoplasmic reticulum. *Commun. Biol.* **4**, 828 (2021).

7. Hessa, T. *et al.* Molecular code for transmembrane-helix recognition by the Sec61 translocon. *Nature* **450**, 1026–1030 (2007).

8. Jumper, J. *et al.* Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).

9. Itskanov, S. *et al.* A common mechanism of Sec61 translocon inhibition by small molecules. *Nat. Chem. Biol.* **19**, 1063–1071 (2023).

10. Ngo, A. M. *et al.* The ER membrane protein complex is required to ensure correct topology and stable expression of flavivirus polyproteins. *Elife* **8**, (2019).

11. Bagchi, P., Speckhart, K., Kennedy, A., Tai, A. W. & Tsai, B. A specific EMC subunit supports Dengue virus infection by promoting virus membrane fusion essential for cytosolic genome delivery. *PLoS Pathog* **18**, e1010717 (2022).

12. Guna, A., Volkmar, N., Christianson, J. C. & Hegde, R. S. The ER membrane protein complex is a transmembrane domain insertase. *Science* **359**, 470–473 (2018).

13. Shan, S.-O. Role of Hsp70 in Post-Translational Protein Targeting: Tail-Anchored Membrane Proteins and Beyond. *International Journal of Molecular Sciences* **24**, 1170 (2023).

14. Tzitzoglaki, C. *et al.* Amantadine variant – aryl conjugates that inhibit multiple M2 mutant – amantadine resistant influenza a viruses. *Eur. J. Med. Chem. Rep.* **6**, 100083 (2022).

15. Dou, D., da Silva, D. V., Nordholm, J., Wang, H. & Daniels, R. Type II transmembrane domain hydrophobicity dictates the cotranslational dependence

for inversion. *Mol Biol Cell* **25**, 3363–3374 (2014).

16. Yang, C.-I. *et al.* System-wide analyses reveal essential roles of N-terminal protein modification in bacterial membrane integrity. *iScience* **25**, 104756 (2022).

17. Lentzsch, A. M. *et al.* NAC guides a ribosomal multienzyme complex for nascent protein processing. *Nature* **633**, 718–724 (2024).

18. Gamerdinger, M. & Deuerling, E. Cotranslational sorting and processing of newly synthesized proteins in eukaryotes. *Trends Biochem Sci* **49**, 105–118 (2024).

19. Aksnes, H. *et al.* An organellar nα-acetyltransferase, naa60, acetylates cytosolic N termini of transmembrane proteins and maintains Golgi integrity. *Cell Rep* **10**, 1362–1374 (2015).

20. Mihara, T. *et al.* Linking Virus Genomes with Host Taxonomy. *Viruses* **8**, 66 (2016).

21. Steinegger, M. & Söding, J. MMseqs2 enables sensitive protein sequence searching for the analysis of massive data sets. *Nat Biotechnol* **35**, 1026–1028 (2017).

22. Bernhofer, M. & Rost, B. TMbed: transmembrane proteins predicted through language model embeddings. *BMC Bioinformatics* **23**, 326 (2022).

RESULTS CHAPTER 4: Membrane association protects non-inserted M2 from degradation and buffers compromised M2 biogenesis in infection

4.1 Chapter Summary

Membrane proteins are translated in the cytosol and must be efficiently targeted and translocated across a membrane to avoid toxic protein aggregates. Proteins with an Nexo N-terminal signal anchor (typeIII) have been shown to be flexible in their choice of ER translocation machinery, which may buffer against inefficient insertion and avoid protein stress. The mechanistic basis of this flexibility is not well understood. We previously found that M2's transmembrane and C-terminal domains make M2 abundance and glycosylation efficiency EMC-dependent. EMC function is however not an absolute requirement for M2 biogenesis, and M2 can enter the membrane in the correct topology, localise to the plasma membrane, and function in infection without EMC, albeit at lower levels than in EMC-sufficient cells. Furthermore, pharmacologic inhibition of Sec61 in an EMC knockout background does not additionally affect M2 biogenesis. EMC may be one of several redundant translocation pathways for M2, and may be involved in stabilising and shielding M2's transmembrane domain after insertion.

To determine EMC's function in M2 biogenesis we sought to measure M2's ER translocation kinetics with high temporal resolution. We used cytosol-specific biotinylation of an AVI-tagged M2 protein to measure M2's ER translocation dynamics and determine the efficiency of M2 insertion in absence of EMC. This N-terminally AVI-tagged M2 mutant retains EMC dependence comparable to the WT protein. We find that EMC KO reduces the rate of M2 glycosylation and dimerization, however, unglycosylated M2 can accumulate to a significant amount. We find that an amphipathic helix in M2's C-terminus can stabilise non-inserted M2 at the cytosolic membrane face when ER membrane insertion is compromised, and possibly allows retrotranslocated M2 to be reinserted into the membrane and trafficked to the plasma membrane. Importantly, without membrane association, M2 cannot accumulate cytosolically, and is drastically attenuated in the membrane in the absence of EMC. We validate the effect seen in the reporter construct in IAV replication, which is more affected by EMC loss when M2's amphipathic helix is mutated. We propose that

membrane association is an intermediate step in membrane protein biogenesis, where membranes act as chaperones to buffer slow insertion kinetics by shielding non-inserted membrane proteins from degradation. This premise of membrane association-mediated stabilisation could be the molecular basis for flexibility of typeIII proteins in their choice of ER translocation routes.

4.2 Introduction

Membrane protein targeting reactions are tasked with directing a newly synthesised membrane protein to the correct membrane for insertion and shielding the hydrophobic sequences from the aqueous cytosol prior to insertion [1]. Soluble or membrane bound proteins destined for the secretory pathway are targeted to the ER either during translation or after the protein is fully synthesised. The interaction of a hydrophobic signal sequence, or transmembrane domain with the ER membrane, can sample the available translocation machinery [2]. A suitable translocon is engaged to translocate hydrophilic stretches of amino acids N or C-terminal of the hydrophobic sequence across the membrane and establish the topology of the inserted domain. Membrane protein biogenesis may fail and result in the transient cytosolic accumulation of uninserted product, which is rapidly degraded or risks forming aggregates [3,4]. Topological error during the insertion process could potentially be corrected before an inserted membrane protein is degraded [56,7]. By contrast, terminally misfolded proteins need to be extracted from the membrane and are degraded by the proteasome. The membrane extraction process or retrotranslocation of aberrant membrane proteins, is mediated by a process called ERAD (ER-Associated Degradation), with its main motor component the AAA-ATPase p97 [8]. Retrotranslocation usually requires ubiquitination and results in cytosolic degradation, however ubiquitin independent retrotranslocation has been shown to be important for calreticulin function in the cytosol or nucleus [9]. Recently a ER resident complex that mediates ubiquitination independent retrotranslocation, ATP13A1, has been shown to be able to extract mislocalized tail-anchored mitochondrial proteins from the ER. Extracted proteins are not targeted for proteasomal degradation and remain competent for mitochondrial translocation [10].

Not much is known about promiscuity in translocon selection and possible buffering mechanisms that may act against inefficient insertion. N-terminal signal anchors of Nexo topology (typeIII) seem to encode some flexibility and have been found to engage Sec61 or EMC for translocation, depending on the lengths of the

sequence that is exposed from the ribosomal exit tunnel during co-translational translocation [11].

M2 is an essential IAV encoded ion channel of typeIII topology. We previously found that M2's transmembrane and C-terminal domains make M2 abundance and glycosylation efficiency EMC dependent. EMC function is however not an absolute requirement for M2 biogenesis, and M2 can enter the membrane in the correct topology, localise to the plasma membrane, and function in infection without EMC, albeit with reduced efficiency. This residual M2-targeting capacity does not require Sec61 since inhibition of Sec61 in an EMC KO background does not additionally affect M2 glycosylation efficiency at steady state. In a targeted genetic screen we identify several Oxa1-superfamily members to affect M2 glycosylation, indicating that M2 may be able to access multiple translocation routes through a yet unidentified mechanism.

Insertases of the Oxa1 family have in common the ability to locally collapse the lipid bilayer, which lowers the energetic barrier for short hydrophilic stretches of amino acids to translocate across [12]. Three ER-resident membrane protein complexes, GET, EMC, and GEL have been identified as part of the family, each with their specific targeting reaction. GET is post-translationally targeted to insert type IV membrane proteins with hydrophobic signal anchors, whereas GEL specifically aids in the translocation of multipass membrane proteins. The EMC has been characterised as an insertase for post-translationally targeted tail anchored membrane proteins, with hydrophilic transmembrane domains, as well as an insertase for co-translationally targeted proteins with Nexo signal anchor, and Cterminal transmembrane domains of multipass membrane proteins [1]. A client overlap for typeIV proteins has been described between GET and EMC [2] and only KO of both complexes becomes lethal for cells [13]. Furthermore, EMC has been described to be able to shield hydrophilic residues within transmembrane domains during folding of polytopic ion channels [14]. M2's transmembrane domain harbors hydrophilic residues which are facing the inside of a channel when M2 is assembled as a homotetramer. Hence, M2 may be dependent on EMC function for membrane insertion or for the shielding of these hydrophilic residues within its transmembrane domain for oligomerization.

The ability to assess the rate of conversion between transient intermediate states during biogenesis requires time-resolved measurements. Compartment-specific biotinylation has been previously used to study transient states in membrane protein biogenesis [7,15]. Here, we have adapted this approach to study the dynamics of M2 biogenesis in WT cells and in absence of EMC.

4.3 Time resolved glycosylation dynamics of M2, assessed through compartment specific biotinylation

Our approach to measure M2's ER membrane translocation dynamics is inspired by an experimental design that made use of an endogenous cytosolic kinase which can modify its cognate peptide motif with radioactive phosphates to observe the glycosylation kinetics of membrane proteins *in vivo* [16]. Instead of using an endogenous kinase we turned to the AVI tag birA system, which allows for the specific biotinylation of a 15 amino acid peptide with respect to the localization of the biotin ligase birA (Figure 4.1A). We aim to label a subpopulation of M2 for which the N-terminus is located in the cytosol, which should only be the case during the initial ER targeting reaction, before translocation occurs. By expressing birA in the cytosol, we can label the AVI-tag that has been N-terminally fused to M2 (AVI.M2), when AVI.M2 is emerging from the ribosomal exit tunnel, but not AVI.M2 that has its N-terminus already translocated across the ER membrane (Figure 4.1B). Membrane translocation is assessed by glycosylation of M2's N-terminus, where the unglycosylated species (-Y) is expected to be located in the cytosol, the glycosylated species (+Y) is expected to be located in the ER, and the hyperglycosylated species (++Y) is either located to the Golgi network, or the plasma membrane (Figure 4.1B). Pulsed labelling requires cell culture medium without biotin once birA and AVI.M2 are transfected. Continuous culture in the presence of biotin results in all AVI.M2 glycosylation species being labelled in WT and EMC6 KO cells, similar to the total amounts assessed by FLAG staining (Figure 4.1C). Omitting birA from the reaction will reveal endogenously biotinylated proteins, and omitting AVI.M2 produces a band that corresponds to birA bound to biotin, both of which act as specificity controls (Figure 4.1C). As has been previously shown for the kinase system, we can confirm that cytosolic birA cannot biotinylate AVI-tag that follows the signal peptide of the cotranslationally targeted HLA-A, because SRP54 and Sec61 translocon engagement occlude the access to AVI-tag during cotranslational targeting (Figure 4.1D). AVI.HLA-A can however be biotinylated by birA that is targeted to the ER lumen (Figure 4.1D). This control proves that cytosolically targeted birA does not enter the secretory route (ER lumen) and stays exclusively in the cytosol.

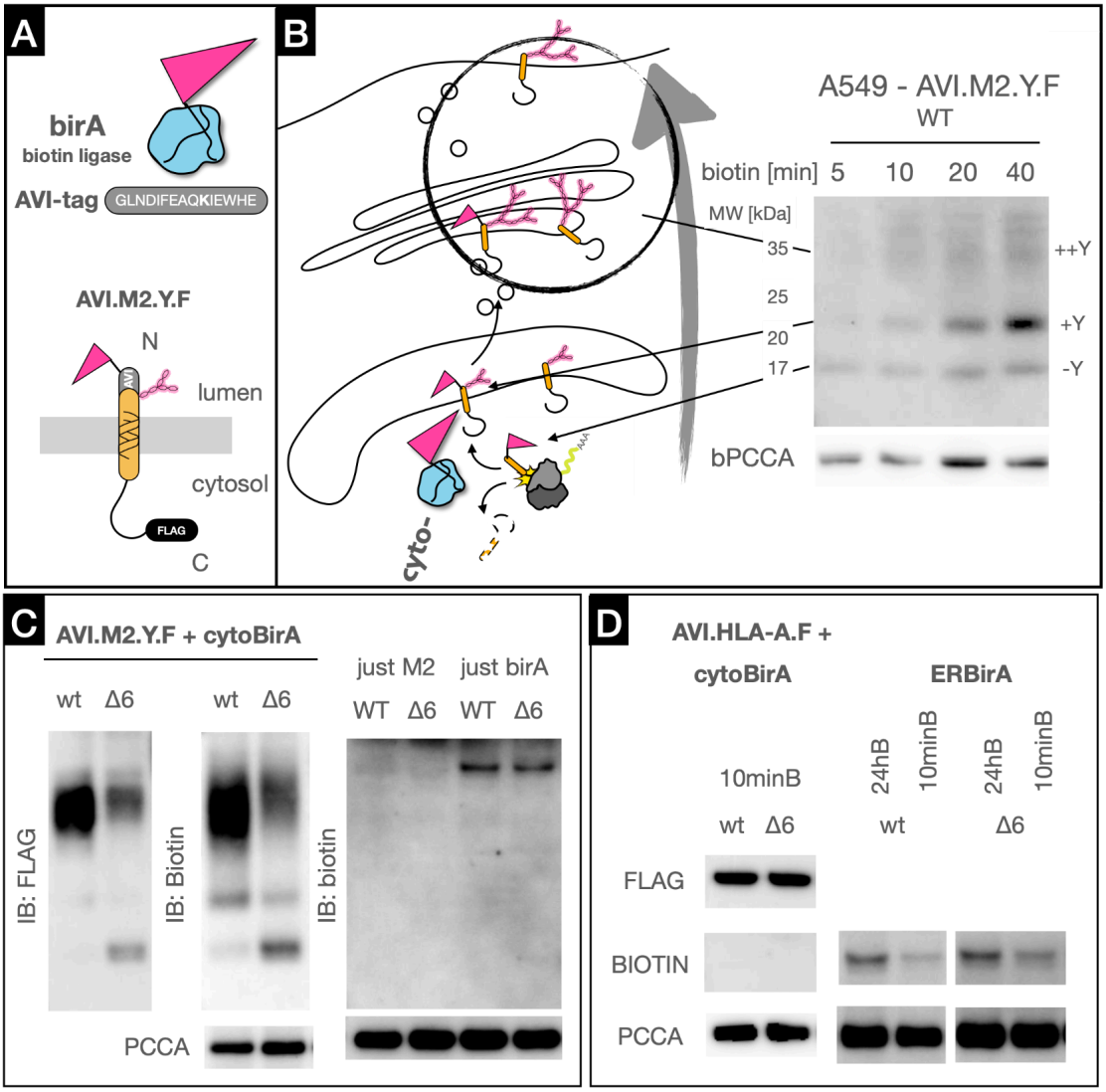


Figure 4.1 Pulsed biotinylation for time resolved glycosylation assay

A) A 15 amino acid AVI-tag is fused to M2.Y.F's N-terminus (AVI.M2) which creates a highly specific biotinylation site for birA. B) A549 WT cells are transfected with cytoBirA plus AVI.M2 in biotin free media for 24h, labelled with 100uM biotin for the indicated time interval and harvested on ice to quench the labelling reaction. The lysates are analysed by western blotting, biotinylated PCCA (mitochondrial carboxylase) serves as loading control. C) A549 WT or Δ6 cells are transfected with cytoBirA plus AVI.M2, or just with AVI.M2 or just with birA for 24h in biotin containing media, lysed and analysed by western blot. D) A549 WT cells are transfected with cytoBirA plus AVI.HLA-A or ERBirA plus AVI.HLA-A for 24, labelled with 100uM biotin for the indicated time interval and harvested on ice to quench the labelling reaction. The lysates were analysed by western blotting.

4.4 The effect of EMC on M2's glycosylation dynamics

We hypothesise that the diminished accumulation rate of M2 at the plasma membrane is due to a defect in M2's efficiency of inserting into the ER membrane in the absence of EMC. Several processes contribute to the steady-state levels of different glycosylated species of M2.Y.F during its biogenesis and trafficking along the secretory pathway (Figure 4.2A). M2 is unglycosylated in the cytosol (-Y), which could be subject to degradation or translocation across the ER membrane and glycosylation (+Y). M2 is further modified upon trafficking beyond the ER (++Y). In the ER, M2 could potentially be subject to deglycosylation and extraction from the ER membrane into the cytosol (retrotranslocation) and subsequent degradation, or potentially cleaved by intramembrane proteolysis. As translocation, and retrotranslocation occur at faster time scales than we can resolve with this assay, it is not possible to distinguish these rates without specific inhibition of either.

A time course of pulsed biotinylation in WT and ∆6 cells revealed a pronounced difference in the glycosylation dynamics in absence of EMC (Figure 4.2B +DTT -PNGase). We quantified a striking reduction in glycosylation efficiency (34% ∆6/WT) and maximal +Y accumulation rate (44% ∆6/WT) in ∆6 cells (Figure 4.2C), strongly correlating with the reduction in M2's plasma membrane accumulation rate. Disulphide-linked dimerization of M2 occurs at positions C17 and C19 through the action of ER luminal isomerases [17]. The lack of EMC resulted in reduced dimerization (Figure 4.2B -DTT +PNGase). Surprisingly, our data showed that -Y accumulated at a rate that kept the total biotin labelled fraction close to equal between WT and ∆6 cells (Figure 4.2B +DTT +PNGase). Pulsed biotinylation of AVI.M2 in the cytosol can resolve the ER membrane translocation dynamics of M2 assessed by glycosylation and dimerization. We hypothesize that the loss of EMC either affects M2 insertion rate, its insertion topology into the ER membrane, or M2's mislocalization to a different membrane, eg. peroxisomes or mitochondria.

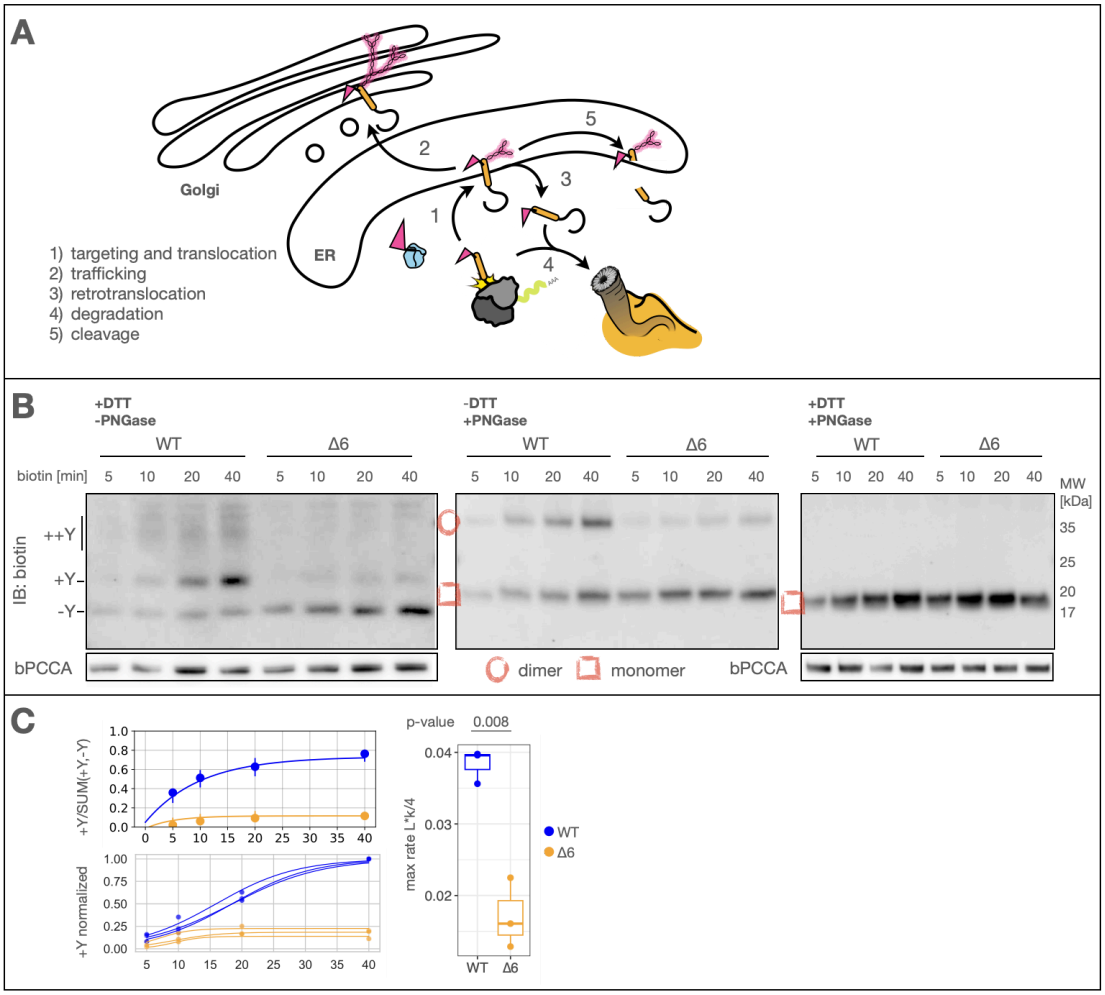


Figure 4.2 M2 glycosylation dynamics through pulsed biotinylation

A) Scheme of possible reactions during M2 biogenesis acting on unglycosylated (-Y) and glycosylated M2 (+Y). B) A549 WT and Δ6 cells were transfected with cytoBirA plus AVI.M2.Y.F for 24, labelled with biotin for the indicated time intervals before harvesting on ice to quench the labelling reaction. The lysates were treated ±DTT and ±PNGase and analysed by western blot (N=3). C) Glycosylation efficiency (+Y/SUM(+Y,-Y)) and +Y levels were quantified by densitometry. An exponential model was fitted to the mean glycosylation efficiency of WT or Δ6. A logistic growth model was fitted to the mean +Y accumulation of WT or Δ6 of each repeat and data was normalized to the maximal estimated value and the maximal slope was determined as max accumulation rate.

4.5 Unglycosylated M2 is stable in the cytosol and partly inserted into membranes

In order to elucidate where in the cell unglycosylated M2 accumulates, we selectively permeabilized the plasma membrane with digitonin to extract the cytosol from membranes and found that in Δ6 cells unglycosylated M2 can be found in the cytosol (Figure 4.3A). Unglycosylated M2 is also found in the washed fraction. In this

experiment it remains unclear if -Y is membrane inserted, or if this fraction contains non-permeabilized cells, as it also contains GAPDH, which is expected for mild permeabilization conditions. In order to test if unglycosylated M2 is inserted into membranes, we mechanically lysed cells to fractionate the membranes and extract the membranes with sodium carbonate (Na_2CO_3) to wash away peripheral membrane proteins. Following this approach, we obtained an unglycosylated band (-Y) in the membrane fraction, which is resistant to the salt wash, as compared to the ER luminal peripheral membrane protein PDI (Figure 4.3C), indicating that that a fraction of unglycosylated M2 (-Y) is inserted into a membrane.

M2 might enter the membrane topologically inverted so that the N-terminal glycosylation motif is not presented to the ER lumen. We generated a reporter to test this possibility by adding a glycosylation acceptor motif to both the N- and C-termini of M2 (NYCY). Under the assumption that addition of the second glycosylation acceptor motif does not impact on topogenesis, the lack of glycosylated species of the NYCY mutant in a 10 minute biotin pulse in ∆6 cells (Figure 4.3B lane 6) shows that topological inversion is not prevalent and that complete translocation of M2 into the ER lumen does also not occur (Figure 4.3B). We can selectively detect which glycosylated species are present on the plasma membrane by biotinylating M2's N-terminus with purified birA added extracellularly into the culture media. We control for selective plasma membrane labelling by confirming that cytosolic AVI.GFP and ER localized AVI.SRPRB cannot be biotinylated in the same experiment. Indeed, we see that a significant fraction of unglycosylated M2 can reach the plasma membrane in ∆6 cells (Figure 4.3D). In order to test if infection changes M2 insertion efficiency, we transfect cells with birA and AVI.M2 and infect 24h later with PR8 at MOI 3. The amount of M2 at the plasma membrane is slightly increased with infection (Figure 4.3D lane 4,5,9,10), which is likely due to the effect of NS1. Cotransfection of NS1 and M2 resulted in a higher level of M2 at the plasma membrane measured by flow cytometry (results not shown).

Unglycosylated M2 in the membrane may be a product of inefficient glycosylation, or retrotranslocation from the membrane and retargeting. Before a glycoprotein is extracted from the membrane it is usually deglycosylated. This changes the glycan-linked asparagine to aspartate which abolishes glycosylation competence [7]. M2 that is extracted and retargeted for ER insertion could therefore not be glycosylated, but trafficked unglycosylated to the plasma membrane. At the moment we cannot distinguish between the two scenarios.

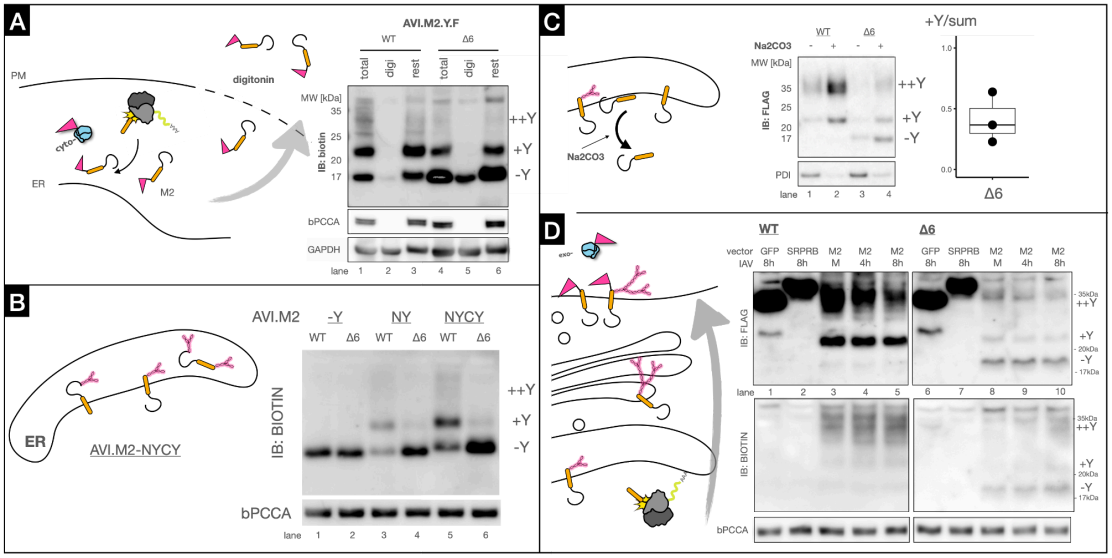


Figure 4.3 Unglycosylated M2 can be found inserted into membranes and partially reaches the plasma membrane

A) A549 WT and Δ6 cells were transfected with cytoBirA plus AVI.M2 for 24h, labelled with biotin for 40min before harvesting on ice with either 1% triton (total) or 0.02% digitonin. The digitonin solubilized cells were fractionated by centrifugation in supernatant (digi) and pellet (rest). Lysates were analysed by western blot relative to the mitochondrial membrane protein PCCA (biotinylated bPCCA) and the cytosolic GAPDH (N=2). B) A549 WT and Δ6 cells were transfected with cytoBirA plus AVI.M2-Y (no glycosylation site), or AVI.M2-NY (N-terminal glycosylation site (regular AVI.M2)) or AVI.M2-NYCY (N- and C-terminal glycosylation site) for 24h, labelled with biotin for 5min before harvesting on ice to quench the labelling reaction. Lysates were analysed by western blot (N=3). C) A549 WT and Δ6 cells were transfected with M2.Y.F for 24h, detached by trypsin treatment, mechanically lysed, fractionated by ultracentrifugation (membrane pellet lane 1 and 3) washed in Na_2CO_3 and fractionated again by ultracentrifugation (membrane pellet lane 2 and 4). Lysates were analysed by western blot, with PDI as peripheral membrane protein control (N=3). D) A549 WT and Δ6 cells were transfected with AVI.GFP, AVI.SRPRB or AVI.M2 for 20h without cytoBirA. Cells were either mock infected (M), or infected with PR8 at MOI of 3 for 4 or 8h prior to labelling. Cell surface exposed AVI-tag was labelled with purified birA added to the media for 5min before harvesting on ice to quench the labelling reaction. Lysates were analysed by western blot (N=2).

4.6 M2's amphipathic helix mediates stability of uninserted cytosolic M2

The accumulation of uninserted M2 in the cytosol is somewhat unexpected, as quality control mechanisms should rapidly degrade this fraction [3,4]. M2 has an amphipathic helix directly following its transmembrane domain, which has been shown to be capable of targeting soluble mCherry to membranes [18]. We asked if M2's amphipathic helix is responsible for unglycosylated M2 stability in the cytosol. M2's

amphipathic helix has been extensively studied in the context of viral egress [19,20]. We opted to construct a previously characterised amphipathic helix mutant, which maintains the secondary structure but abolishes the hydrophobic moment by mutating 5 hydrophobic residues to alanine (AVI.M2-AHmut.Y.F, Figure 4.4A) [20].

In a 5 minutes biotinylation reaction we observed that similar to AVI.M2, AVI.M2-AHmut is glycosylated (+Y) in WT cells, and that +Y accumulated less efficiently in Δ6 cells. In both conditions we can observe proportionally less accumulation of an unglycosylated band for M2 +Y versus M2 AHmut (Figure 4.4A). The overall levels assessed by anti-FLAG antibody reveal that none of the species of AVI.M2-AHmut may efficiently accumulate in Δ6 conditions (Figure 4.4A).

We assessed the same parameters as before, glycosylation and dimerization dynamics, this time of AVI.M2-AHmut.Y.F in a time course of pulsed biotinylation (Figure 5C +DTT -PNGase). AVI.M2-AHmut.Y.F is affected similar to AVI.M2.Y.F in glycosylation efficiency (38% Δ6/WT), but even more affected in the maximal +Y accumulation rate (19% Δ6/WT) (Figure 4.4C). In contrast, AVI.M2-AHmut.Y.F did not produce a stable unglycosylated band in Δ6 conditions (Figure 4.4B +DTT -PNGase). Interestingly, the capacity for AVI.M2-AHmut to form dimers after membrane insertion was compromised in both cell lines (Figure 4.4B -DTT +PNGase). Strikingly, we observed that in contrast to AVI.M2.Y.F, the total amount of AVI.M2-AHmut.Y.F is drastically decreased in Δ6 compared to WT cells (Figure 4.4B +DTT +PNGase).

We assessed if the residual unglycosylated band is in the cytosol or in the membrane fraction, by extracting the cytosol with digitonin (Figure 4.4D). This experiment was performed in parallel with AVI.M2.Y.F and processed on the same blot (Figure 4.3A). In both WT and Δ6 cells -Y can be detected in the digitonin extracted fraction, and only a very faint residual band is detected in the digitonin-washed fraction (Figure 4.4D). The data shows that M2's amphipathic helix mediates the stability of an unglycosylated fraction in the cytosol.

These results suggest that M2's amphipathic helix is important for stabilising the unglycosylated band in the cytosol and that the membrane insertion in Δ6 cells is further compromised without the amphipathic helix.

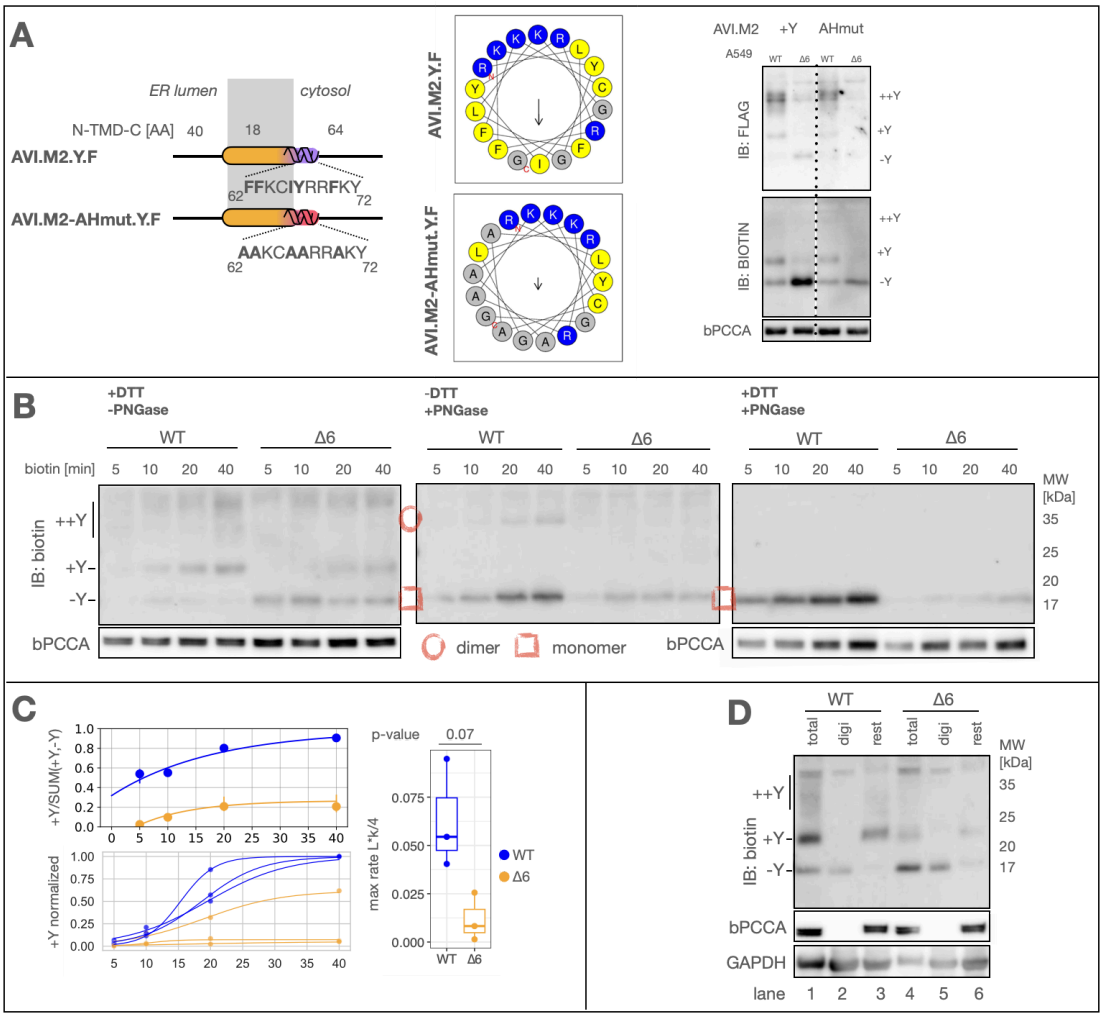


Figure 4.4. M2's amphipathic helix mediates cytosolic stability of unglycosylated M2

A) Sequence modifications of AVI.M2-AHmut and helical wheel representation. A549 WT and Δ6 cells were transfected with cytoBirA plus AVI.M2 and AVI.M2-AHmut for 24h, labelled with biotin for 5min before harvesting on ice to quench the labelling reaction. Lysates were analysed by western blot first revealing the total amount of M2 variants by FLAG, and revealing the biotinylated bands after stripping the blot by streptavidin staining (N=2). B) A549 WT and Δ6 cells were transfected with cytoBirA plus AVI.M2-AHmut for 24h, labelled with biotin for the indicated time intervals before harvesting on ice to quench the labelling reaction (N=3). The lysates were analysed by western blot and the glycosylation efficiency was quantified by densitometry. C) Glycosylation efficiency (+Y/SUM(+Y,-Y)) and +Y levels were quantified by densitometry. An exponential model was fitted to the mean glycosylation efficiency of WT or Δ6. A logistic growth model was fitted to the mean +Y accumulation of WT or Δ6 of each repeat and data was normalized to the maximal estimated value and the maximal slope was determined as max accumulation rate. D) A549 WT and Δ6 cells were transfected with cytoBirA plus AVI.M2-AHmut for 24h, labelled with biotin for 40min before harvesting on ice with either 1% triton (total) or 0.02% digitonin (N=2). The digitonin solubilized cells were fractionated by centrifugation in supernatant (digi) and pellet (rest). Lysates were analysed by western blot.

4.7 Membrane association increases M2's translocation efficiency

In order to assess if membrane association is required for the stabilization in the cytosol we made use of another mutation in M2 that affects its post-translational modification. M2 palmitoylation occurs at a cysteine at position 50, located within its amphipathic helix [21]. Palmitoylation takes place at the cytosolic face of membranes and is catalysed by enzymes that localise to the cytosolic side of the ER membrane, the Golgi, and the plasma membrane [18]. In agreement with membrane association, we detected palmitoylation as a shift in +Y and -Y fractions of AVI.M2.Y.F in both WT and EMC ablated conditions (Figure 4.5A lane 1 and 7), compared to the DTT treated controls (Figure 4.5A lane 2 and 8), or when the cysteine at position 50 is mutated to a serine (Figure 4.5A lane 3,4 and 9,10). Strikingly, this band is almost entirely lost when M2's amphipathic helix is disrupted (Figure 4.5A lane 5,6 and 11,12). This data suggests that membrane association of fully translated M2 precedes translocation. To test this scenario, we assessed if AVI.M2.Y.F can insert post-translationally, by inhibiting translation with a high dose of cycloheximide for 30 minutes, before pulse labelling cytosolic AVI.M2.Y.F with biotin. Indeed, we can observe glycosylation under these conditions in WT and Δ6 cells, confirming that at least a part of cytosolic -Y can insert post-translationally (Figure 4.5B).

Increased stability of non-inserted M2 in the cytosol may increase M2's chance of translocation and thus buffer inefficient insertion. We measured the impact of membrane association on cytosolic stability and the impact on translocation rate. To do so, we pulse-labelled M2 in the cytosol and either mock-treated cells with DMSO or inhibited proteasomal degradation or autophagy with MG132 or bafilomycin A1, respectively, for 6h. LC3 lipidation (LC3 II) confirmed the action of MG132 and bafilomycin A1 over DMSO-treated cells (LC3 II, Figure 4.5C). Furthermore, we assessed transfection efficiency to be equal between the conditions with birA expression from the co-transfected plasmid, arguing against differences in transfection efficiency to be the cause of the observed effects (birA, Figure 4.5C).

In agreement with the previous experiments, we found a drastic decrease in the +Y levels at 20 minutes in the absence of EMC (31% Δ6/WT, Figure 4.5C). Strikingly, we found that a lack of the amphipathic helix similarly decreased the +Y levels at 20 min (22% WT^{AHmut}/WT) and that the combination of EMC ablation and amphipathic helix mutation was additive, decreasing the +Y levels at 20 min even further (7% $\Delta6^{AHmut}$/WT, Figure 4.5C). These data establish a pronounced effect of membrane association on the translocation efficiency of M2, which is presumably especially important when insertion sites are scarce.

If the cytosolic stabilization of an unglycosylated band aids in translocation efficiency, there should be an increase in +Y levels through the inhibition of degradation. Such increase in +Y levels would also argue against the amphipathic helix mutant suffering a topogenesis defect. Indeed, proteasomal inhibition increased the -Y levels over DMSO treatment in all conditions (2x WT^{M2}, 2x $\Delta6^{M2}$, 3.5x WT^{AHmut}, 9x $\Delta6^{AHmut}$) and resulted in higher +Y levels relative to the DMSO treatment in all conditions except in WT cells expressing AVI.M2.Y.F (1x WT^{M2}, 1.5x $\Delta6^{M2}$, 2x WT^{AHmut}, 2.5x $\Delta6^{AHmut}$, Figure 4.5C). Interestingly, lysosomal inhibition only increased AVI.M2.Y.F's +Y levels in WT cells (3x WT^{M2}, Figure 4.5C), which is in agreement with previous research that links strong M2 expression with its lysosomal degradation [22]. We conclude that M2 is able to adhere to the cytosolic face of membranes via its amphipathic helix to escape proteasomal degradation, increasing translocation efficiency.

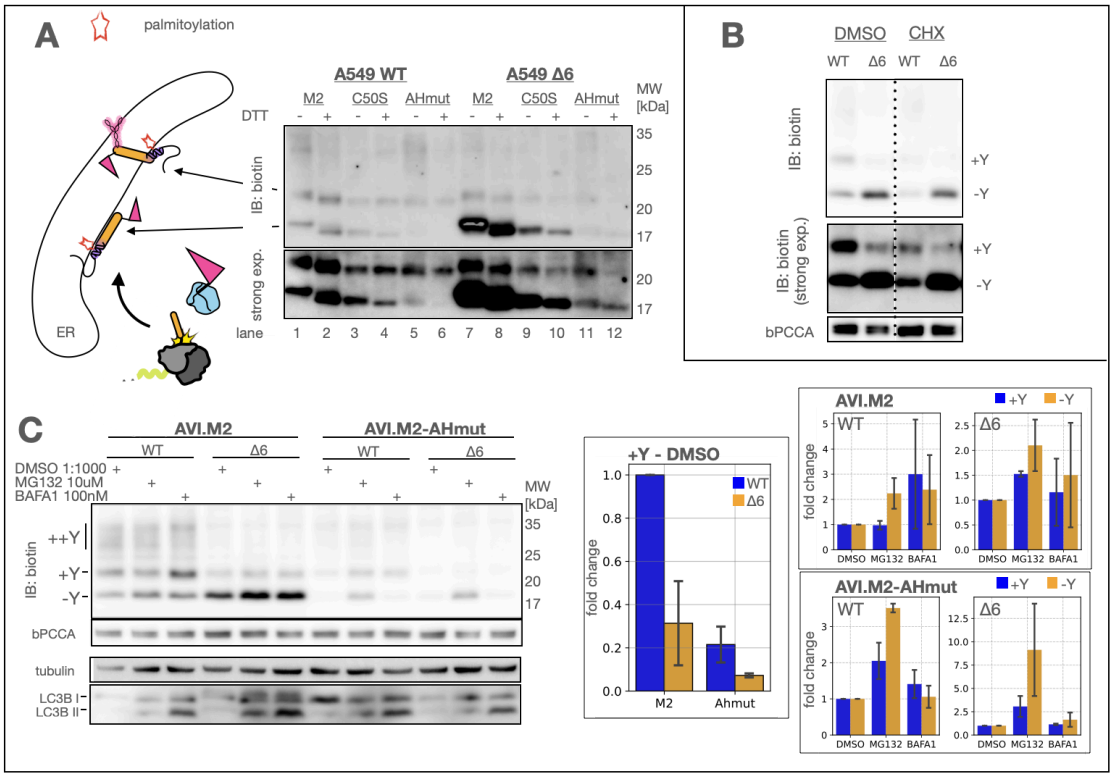


Figure 4.5 M2's amphipathic helix mediates cytosolic stability through membrane association

A) A549 WT or Δ6 cells were transfected with cytoBirA plus AVI.M2, or AVI.M2-C50S or AVI.M2-AHmut for 24h and labelled with biotin for 10min before harvesting on ice to quench the labelling reaction. Lysates were diluted in loading buffer with or without DTT and analysed by western blot (N=3). B) A549 WT or Δ6 cells were transfected with cytoBirA plus AVI.M2 for 20h and treated with either 50ug/ml CHX, or the equivalent volume fraction of DMSO for 30 min, and labelled with biotin plus drugs for 20min before harvesting on ice to quench the labelling reaction. Lysates were analysed by western blot (N=2). C) A549 WT or Δ6 cells were

transfected with cytoBirA plus AVI.M2 or AVI.M2-AHmut for 20h and treated with either MG132, Bafilomycin A1 or the equivalent volume fraction of DMSO for 6h and labelled with biotin plus drugs for 20min before harvesting on ice to quench the labelling reaction. Lysates were analysed by western blot and quantified by densitometry (N=3).

In our transfection model AHmut cannot efficiently accumulate in cells. In infection, M2 is highly expressed and only needed to small amounts in infectious virus particles and at the plasma membrane. We tested if M2's amphipathic helix is important to buffer compromised ER insertion in infection. We created a mutant virus in the PR8 background carrying the same mutations in M2 that abolishes the hydrophobic feature in the amphipathic helix (AHmut). We infected WT and Δ6 cells and measured HA and M2 accumulation at the plasma membrane and viral titers 8 hours post infection. As before, we assessed the magnitude of plasma membrane expression in non-permeabilized cells by modelling the expression profile with two skew normal distributions (Figure 4.6A). We see that M2 is expressed efficiently on the cell surface in both viruses in WT cells, but is significantly reduced to 46% and 28% of the levels in WT cells in PR8 and AHmut respectively (Figure 4.6B). As before, we find that the loss of EMC function does not have an effect on viral titers when comparing PR8 infected Δ6 to +6 cells (+6 = Δ6+EMC6 rescue cells), but strikingly we find that loss of EMC shows a significant effect (25% Δ6/+6) on viral replication when M2's amphipathic helix is ablated in AHmut replication (Figure 4.6C). The results confirm that M2's amphipathic helix plays an important role in M2 biogenesis, when ER membrane insertion is compromised or retrotranslocation affects M2's stability in the membrane.

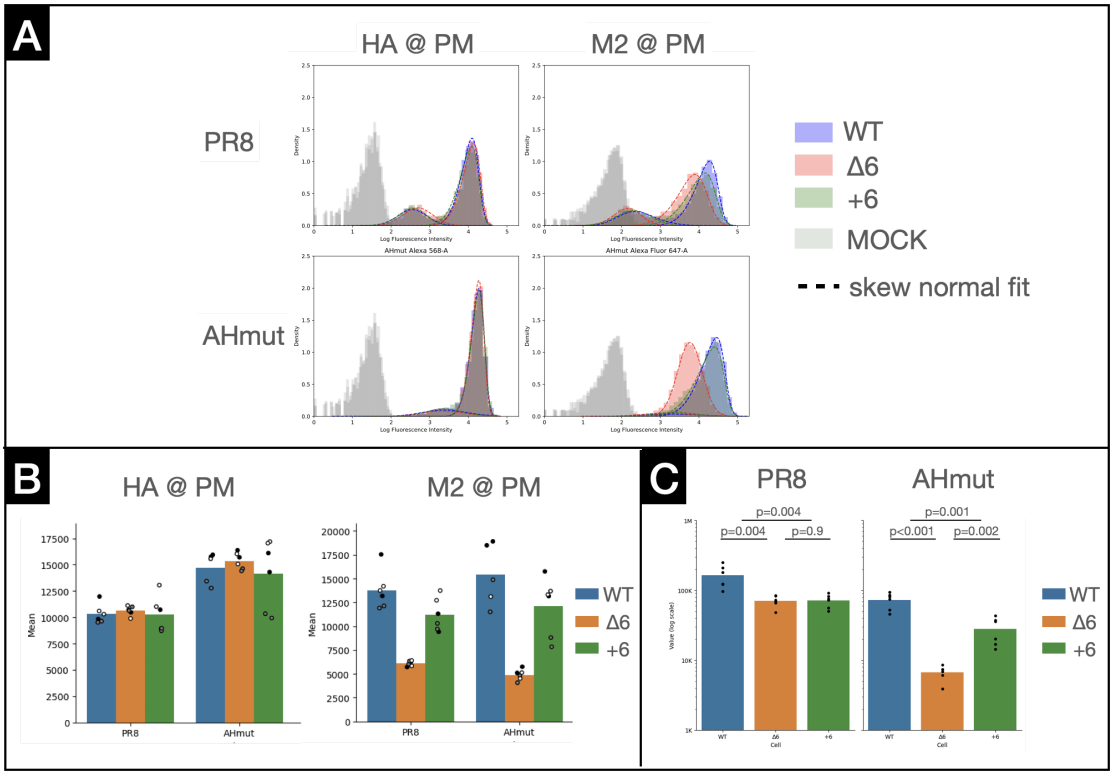


Figure 4.6 Role of M2 amphipathic helix in access to alternative membrane insertion pathways for M2

A) Expression of HA and M2 at the surface of non-permeabilized infected cells. A549 WT, Δ6 or +6 cells were infected with PR8 or AHmut at MOI 3 for cell harvest at 8 hours post infection, and subsequent detection of viral HA and M2 by flow cytometry. The mean of an HA or M2 positive population of single cells was accurately estimated by fitting a skew normal distribution (N=3). B) Quantification of mean viral protein expression at the cell surface of A). C) Viral titers measured in supernatant in the same experiment described in A).

The ability to buffer against limited translocation capacity may provide a significant advantage in viral infection. In our PR8 infection model, we show that M2 maintains sufficient plasma membrane expression to sustain viral replication in EMC-deficient cells. Analyzing publicly available M2 sequences, we find that the amphipathic helix is generally conserved across IAV strains, but the strength of its hydrophobic moment varies between viruses of different host species (Figure 4.7A). Interestingly, other viral ion channels may exploit similar strategies. For example, HIV's Vpu ion channel encodes two amphipathic helices in its C-terminus, suggesting that membrane association might be a conserved feature among viroporins [23].

We assessed how prevalent TMD-adjacent amphipathic helices are in single pass membrane proteins within the human proteome and in viral proteomes of human hosts (Figure 4.7B). We calculated the maximal hydrophobic moment along a

100 amino acid stretch centred around the transmembrane domain assuming a helical secondary structure. Interestingly, we find that many typeI and typeIII proteins have a significant hydrophobic moment (>0.35) close to the TMD at the C-terminal flanking region, whereas typeIV and typeII proteins have a tendency for a significant hydrophobic moment N-terminally to the TMD. It remains to be determined if these features can promote a similar buffering effect in these proteins that we find for M2. Signal sequences have a large diversity in their hydrophobic character, which has been linked to their likelihood to require SRP targeting [24] and postulated to hold regulatory potential [25]. The fine tuning of translocation efficiency has been demonstrated for the prion protein, which is rejected from entering the ER during ER stress [26]. However, we do not find a correlation between the presence of an amphipathic helix in typeI proteins and the strength strength of the signal sequence, nor a correlation with TMD hydrophobicity in all protein types (data not shown).

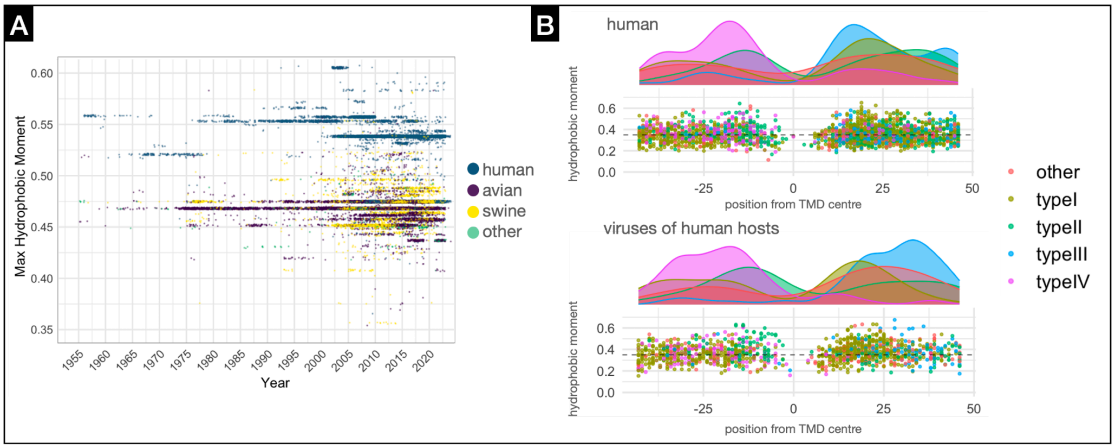


Figure 4.7. Hydrophobic moment analysis of viral and human proteins

A) 44621 M2 nucleotide sequences were aligned and translated. Protein sequences with consensus $_{41}$WILDRLFFKCIYRRFKYGLKRG$_{62}$ were analysed for the maximal hydrophobic moment in a 12 amino acid window, assuming 3.6 residues per turn of a helix and applying the Eisenberg hydrophobicity scale [27]. Data was plotted by year and organism the sequence was sampled from. A jitter around the y value was applied for better visualization. B) Sequences from the database described in Figure 3.1 were analysed in the same way as in A) along a 100 amino acid stretch centred around the TMD. The dotted horizontal line at 0.35 is the mean value for a large sample of random sequences.

4.8 Discussion

Fidelity in membrane protein targeting and insertion are critical for cellular homeostasis, ensuring that thousands of simultaneously synthesized proteins reach their correct organelles [28–30]. However, not all proteins translocate efficiently, and ambiguous sequence-encoded signals can result in toxic misfolded or non-inserted

translation products [3,31–33]. Viruses, which rely on the host membrane protein biogenesis machinery, must efficiently navigate these constraints to ensure rapid membrane protein synthesis of key viral proteins for virion assembly. We find that the small viral ion channel M2, employs previously unappreciated features to improve its membrane translocation efficiency.

By measuring M2's ER membrane translocation dynamics with pulsed biotinylation, we find that EMC ablation leads to M2 accumulating at the cytosolic membrane face through a mechanism that involves its transmembrane domain-adjacent amphipathic helix. Amphipathic helix-mediated membrane association protects cytosolic M2 from proteasomal degradation, and allows M2 to insert post-translationally through an alternative translocon. Very little M2 at the plasma membrane is sufficient for virion production and the defect in M2 expression due to EMC loss does not limit virion production at 8h post infection. However, a virus that expresses M2 amphipathic helix mutant is severely affected when replicating in EMC-deficient cells, strongly suggesting that M2's amphipathic helix buffers inefficient insertion.

Recently, the potential for the membrane protein biogenesis machinery to act redundantly has been investigated, and overlapping pathways have been identified for multipass membrane proteins with an Nexo first transmembrane domain. Depending on the folding state and the N-terminal charges, substrates were more or less dependent on EMC or BOS (3 subunit complex named "back of Sec61") which closely interact [2,33]. A redundancy in membrane protein translocation machinery for post-translationally inserted typeIV proteins has long been demonstrated in client overlap between Oxa family members EMC and GET [2]. However, the mechanism of redundant access remains insufficiently understood. The small phage coat protein Pf3, a model membrane protein for many decades, may interact with membranes before its insertion via YidC, the prokaryotic homolog of EMC [34]. Furthermore, crosslinking studies show that a membrane protein nascent chain might interact first with the membrane, before interacting with the EMC [11]. This membrane proximity might aid in the ER insertion process. Membrane proximity as a mechanism for the insertion, via a redundant translocation route, is an intriguing possibility as Oxa1-family members of insertases share a common translocation mechanism. M2, which is proximal to the membrane, could insert through similar hydrophilic cavities of Oxa1 family members that allow the translocation of short hydrophilic domains. The specificity for translocation through GET or EMC is thought to be established by the targeting pathway, which is irrelevant in this scenario. A critical experiment to test whether redundant pathways are responsible for the insertion of residual M2 in the

absence of EMC would be to overexpress these translocons in an EMC knockout background and observe a rescue of M2 glycosylation efficiency, or observe further reduction in glycosylation efficiency in a double KO. Furthermore, results in Chapter 3 clearly show that SRPRB biogenesis does not depend on EMC or Sec61, as double inhibition does not affect its glycosylation efficiency. SRPRB is likely to utilise an alternative translocon to efficiently translocate across the ER membrane in our overexpression system. Combining biotin pulse with crosslinking to proteins early in M2 or SRPRB biogenesis may enrich translocon factors that aid insertion.

M2's amphipathic helix may also aid in the initial membrane targeting step. We currently do not fully understand which features of amphipathic helices encode specificity for certain membranes in the cell. Recent studies however started comparative analysis to tease apart specific amphipathic helix features and their affinities for lipid composition, membrane curvature, and lipid packing [21,35,36]. We identify M2's membrane association through C50 palmitoylation, which is likely mediated by DHHC20 [21]. However, as palmitoylation enzymes are distributed throughout all endomembrane systems and the plasma membrane, we cannot derive any specific target membrane from these results.

We show that membrane association can buffer inefficient insertion. Furthermore, membrane association could lead to concentrating insertion competent substrates close to translocons to preferentially engage in translocation. This may be especially important in the early phase of viral replication, when host mRNAs are competing for translation with viral mRNAs. In infection, M2 lacking the amphipathic helix can accumulate at the plasma membrane in absence of EMC, whereas in transfection most M2 is degraded. Further experiments are needed to understand if the elimination of competing host mRNAs increases M2 biogenesis efficiency in this case. Thus, the translation dynamics of viral proteins may affect the dynamics of virion production, their spread through the infected host tissue and ultimately the likelihood of transmission between hosts.

Finally, our results suggest that M2 could be retrotranslocated in absence of EMC. M2's stability in the membrane may be dependent on homo-dimerization, or through binding to a chaperone. Both of these aspects may be affected by EMC, as EMC has been shown to act as chaperone in ion channel assembly [37], and the expression level of many proteins [38]. M2 may become sensitive to ERAD upon the loss of EMC, which involves retrotranslocation from the ER membrane and potential glycan cleavage in the cytosol [7]. The AAA-ATPase p97 is generating the pulling force for membrane protein extraction from the membrane [39]. We initially aimed to quantify the contribution of ERAD to -Y in Δ6 cells by pharmacologically inhibiting p97. Our

current results are not conclusive, because the lack of EMC results in an elevated basal level of unfolded proteins, and transient inhibition of ERAD leads to UPR induction. Furthermore, in order for retrotranslocated M2 to contribute to -Y would need to escape proteasomal degradation after extraction from the membrane, which our data cannot explain. It would be possible to prove the plasma membrane localization of retrotranslocated and reinserted M2 -Y with an antibody specific for the modification due to deglycosylation [7].

Hypothetical model of M2 biogenesis:

M2 may be preferentially co-translationally targeted to the ER membrane, but may also be associated with cytosolic chaperones like SGTA or calreticulin. However, M2's short C-terminus does not allow for a prolonged sampling time to engage a suitable translocon cotranslationally and is thus released close to the ER membrane upon dissociation of the ribosomal subunits and SRP release. M2 interacts post-translationally with the ER membrane through the action of the amphipathic helix and will be palmitoylated by a membrane embedded DHHC enzyme, which may strengthen membrane association. The highly expressed EMC preferentially engages with the fully translated, membrane associated M2 for translocation, and possibly holds M2 engaged to efficiently dimerize upon insertion. The lack of M2 dimerization or association with a chaperone may lead to retrotranslocation of M2 from the membrane. Membrane association upon retrotranslocation may partially result in inefficient proteasomal degradation and the retention of M2 at the membrane. Retrotranslocated, membrane associated M2 may thus have another chance at engaging a translocon for translocation. In the absence of membrane association, M2 has a shorter time window for engaging with EMC (or other Oxa1-family translocons) for translocation. M2 which fails to rapidly translocate without membrane association will be more likely captured by the proteasome for degradation.

4.9 Methods

Cell lines

A549 (ATCC CCL-185) and Madin-Darby Canine Kidney (ATCC MDCK-I) were cultured in Dulbecco's Modified Eagle Medium (DMEM, Gibco Thermo Fisher 21969035) supplemented with 10% fetal bovine serum (FBS, Thermo Fisher A5256801), 2 mM L-glutamine (GLUT, Thermo Fisher 25030024), and 1% (v/v) penicillin-streptomycin (P/S, Biowest L0022-100) in a humidified incubator at 37°C and 5% CO2.

Viruses

A/Puerto Rico/8/34 (PR8) virus and related mutant virus were obtained via reverse genetics plasmid system (de Wit et al., 2007).

Plasmids

M2.Y.F was amplified with a forward primer adding the N-terminal AVI-tag and cloned into pcDNA3 to obtain AVI.M2. The C50S mutation was introduced by quick change mutagenesis. The glycosylation site in AVI.M2 was mutated back to the original leucine to obtain AVI.M2.-Y by quick change mutagenesis. AVI.M2.NYCY was obtained by introducing amino acids NGT at AVI.M2's C-terminus before the Flag-tag by quick change mutagenesis. The mutations in M2's amphipathic helix to obtain AVI.M2.AHmut were introduced via overlap-PCR with primers containing the mutations and cloned into pcDNA3. The mutations in the IAV genomic segment 7 were introduced in the same way. birA was amplified from myc-birA (gift from Alfonso Bolado, Edinburgh) and tagged with an N-terminal V5-tag, and either tagged with a C-terminal nuclear export signal (cytoBirA), or tagged with an N-terminal signal peptide (ER-BirA) and cloned into pcDNA3. AVI.GFP and AVI.SRPRB were obtained by amplifying the sequences from plasmid with an N-terminal AVI-tag and cloned into pcDNA3.

Antibodies

Flag-tag (Sigma-Aldrich F1804-200UG), M2 14C2 (Abcam ab5416), IAV-HA hybridoma (gift from Jonathan Yewdell), GAPDH, tubulin YL1-2 hybridoma (homemade), streptavidin-HRP (Thermo Fisher S911), LC3B.

Reagents

JetPrime (Polyplus 101000046), cycloheximide (Alfa Aesar J66901.03), ECL Select (GE Healthcare RPN2235), BSA (Sigma-Aldrich A9418-100G), nitrocellulose membrane (GE Healthcare 10600003), FBS dialyzed (GE Healthcare HYCLSH30079.02), culture plates (Corning 734-1597), D-biotin (Supelco 000000000000047868), birA purified (gift from Margarida Silva, Oeiras ITQB).

Pulsed biotinylation

A549 WT or Δ6 cells were seeded in 24 well plates (6x10^4 cells/well) and transfected 16h later with cytoBirA plasmid and AVI.M2.Y.F or AVI.M2-AHmut.Y.F plasmid (each 150ng/well) with Jetprime reagent and incubated for 20h in pre-labelling media (DMEM +2mM GLUT +1%P/S +10%FBS dialysed). Medium was replaced with labelling media +100nM biotin for the time indicated in the experiment, quickly placed on ice and washed 2x with cold PBS to quench the biotinylation reaction. From here on all steps were performed on ice or 4°C. Cells were lysed in 70ul TX100 lysis buffer (50mM Tris-HCl pH 7.4, 150mM NaCl, 1% (v/v) Triton X-100,

5mM EDTA and protease inhibitors) on ice, scraped and transferred to 1.5ml test tubes and spun at 1000xg and 4°C. The supernatant was transferred to new tubes, mixed with 6x LDS sample buffer (550 mM Tris-HCl pH 7.4, 200 mM LDS, 0.13 mM EDTA, 250 mM DTT, 15% v/v glycerol, 0.025% SERVA Blue G250 and 0.025% Phenol Red) and transferred from ice to a 95°C hot plate for 3min. Samples were stored at -20°C until analysis by western blot.

Cytosol extraction with digitonin

Cells were treated following pulsed biotinylation for 40min after quenching the reaction on ice. From here on all steps were performed on ice or 4°C. For the total fraction, cells were lysed in 70ul TX100 lysis buffer on ice, scraped and transferred to 1.5ml test tubes. To obtain the digi and rest factions, 50ul PBS + 0.02% digitonin + 1mM PMSF was added to each well and incubated on ice for 20min. The supernatant was collected in a 1.5ml tube, and cells left in the well were lysed in 50ul TX100 lysis buffer and transferred to a 1.5ml tube. Total, digi and rest fractions were spun down 10min at max speed. The supernatant of all fractions were transferred to new tubes, mixed with 6x LDS buffer and transferred from ice to a 95°C hot plate for 3min. Samples were stored at -20°C until analysis by western blot.

Plasma membrane biotinylation

A549 WT or Δ6 cells were seeded in 24 well plates (6x10^4 cells/well) and transfected 16h later with AVI.M2.Y.F, or AVI.GFP, or AVI.SRPRB plasmid (each 150ng/well) with Jetprime reagent and incubated for 20h. Cells were infected with PR8 at MOI of 3 for 4 or 8h prior to labelling. Medium was replaced with labelling media +100nM biotin +15ug/ml purified birA for 5min at 37°C, and quickly placed on ice and washed 2x with cold PBS to quench the biotinylation reaction. From here on all steps were performed on ice or 4°C. Cells were lysed in 70ul TX100 lysis buffer on ice, scraped and transferred to 1.5ml test tubes and spun at 1000xg and 4°C. The supernatant was transferred to new tubes, mixed with 6x LDS sample buffer and transferred from ice to a 95°C hot plate for 3min. Samples were stored at -20°C until analysis by western blot.

Soduim carbonate extraction

A549 WT or Δ6 cells were seeded in 6 well plates (3x10^5 cells/well) and transfected 16h later with M2.Y.F plasmid (each 750ng/well) with Jetprime reagent and incubated for 20h. Cells were detached with trypsin 7min at 37°C, quenched with FBS and transferred to 1.5ml tubes on ice. From here on all steps were performed on ice or 4°C with cold solutions as membrane inserted single pass membrane proteins may otherwise dissociate from membranes. Cells were washed in PBS and spun down at 300xg 1min. Cell pellet was resuspended in 400ul Hypo buffer (10mM HEPES–KOH

pH 7.4, 1,5mM MgCl2, 10mM KCl + inhibitors) and mechanically lysed with 10 very strong strokes through a 30G needle, keeping the tube in contact with ice at all times. Completion of cell lysis can be observed under a regular light microscope. The solution was spun down at 1000xg 5min. The supernatant was carefully layered on top of a 1.5 ml ultracentrifuge tubes with 1ml of hypo buffer +5% sucrose and spun 20min at 100kxg in a static rotor. The tubes were marked for the direction of the pellet, which is not simply visible by eye. The supernatant was completely removed by pipetting and the pellet was resuspended in 100ul PBS with a 30G needle to break up the pellet. 20ul of this solution was kept as the unwashed membrane fraction and denatured in 6xLDS. The rest was resuspended in 400ul freshly prepared 100mM Na_2CO_3 plus inhibitors and incubated 30min on ice. The solution was layered on top of 1ml 100mM Na_2CO_3 + 5%sucrose and inhibitors in a 1.5ml ultracentrifuge tube and spun 30min 120kxg. The tube was marked and the supernatant was carefully removed by pipetting. The remaining pellet was directly resuspended and denatured in 50ul PBS + 8ul 6xLDS sample buffer. Samples were heated at 95°C for 3min and analysed by western blot.

Western blots and deglycosylation

If needed the 10ul of denatured samples were deglycosylated with NEB PNGaseF (0.3ul per reaction) or EndoH (0.3ul per reaction) as per manufacturer's description. 10% Acrylamide gels containing SDS and sucrose were loaded with samples and run in MOPS running buffer (50 mM MOPS, 50 mM TrisBase, 3.5 mM SDS, and 0.8mM EDTA) at 150V for 50min. Gel was transferred in Biorad transfer system onto nitrocellulose membranes, blocked in 5% milk in PBS + 0.1%TX100 and incubated in primary antibody overnight at 4°C or 1:10000 streptavidin-HRP for 1h at room temperature. 1:10000 secondary antibody incubation was 1h at room temperature. Chemiluminescent detection with Amersham AI600, or fluorescent antibody detection with Biorad Odyssee. Western blots were quantified by densitometry using FIJI ImageJ.

Viral infection with samples for flow cytometry, plaque assay and western blot

Cells plated the previous day were washed in PBS and infected with viral solutions in serum free DMEM at MOI of 3. At the intended time points, cell supernatants were collected and frozen at -80°C until use. For western blot samples cells were washed in PBS, immediately lysed in 2x LDS sample buffer and stored at -20°C. For flow cytometry samples cells were detached in trypsin at 37°C for 7min and transferred to a 96 well plate with conical bottom to spin down at 300xg 4°C for 3min each spin. Cells were washed once in PBS + 1% FBS and once in PBS and fixed in PBS + 4% formaldehyde for 10min at room temperature. Cells were spun down at 500xg 4°C for

5min and washed twice in PBS and stored at 4°C for a maximum of 2 days. Cells were stained for 1h in primary antibody solution or PBS for unstained controls at room temperature. Cells were spun down and washed twice in PBS and stained with 1:1000 isotype specific secondary antibody solution Alexa 488 or Alexa 647 for 30min at room temperature. Cells were spun down and washed twice in PBS and measured at Fortessa X-20.

Flow cytometry analysis

Data acquired on Fortessa X-20 was gated for cells and single cells using FlowJo. Single cells gated data for the two measured channels M2 and HA was exported and further processed in python with custom scripts. Optimal parameters were fitted by maximum likelihood estimation for a mixture model of two skew-normal distributions using global optimization via differential evolution. Only the distributions with the larger mean at each time point were analysed further. The data was processed in two different ways. First, the data for each channel was normalized to the 15h time point for each protein in the WT cells and the mean at each time point was compared between WT and Δ6 cells. Second, each repeat for each protein and WT or Δ6 condition was fit to a logistic growth model using a differential evolution algorithm and normalized to the parameter L of the WT condition for each protein and repeat. The data was fit again to obtain the final optimized parameters of the normalized data.

Plaque assay

MDCK cells were seeded in 12 well plates to reach confluency the next day. Cells were washed in PBS and infected with 300ul of serially diluted supernatants in serum free DMEM. After 1h cells were washed in acid wash solution, washed in PBS and overlayed with serum free DMEM plus 0.14% BSA and AVICEL solution. Cells were incubated at 37°C + 5% CO2 for 30h.

Computational analysis of amphipathic properties

44621 IAV segment 7 nucleotide sequences were obtained from NCBI (downloaded), aligned and translated. Residues 41-62 of the M2 open reading frame were analysed. The single pass protein database analysed in Chapter 3 was analysed for potential transmembrane domain proximal amphipathic helices along a 100 amino acid sequence centered around the transmembrane domain center assessed by the ΔGapp minimum. Hydrophobicity was assessed through the Eisenberg scale [27] assuming 3.6 residues per turn of a helix.

4.10 References

1. Hegde, R. S. & Keenan, R. J. A unifying model for membrane protein

biogenesis. *Nat Struct Mol Biol* **31**, 1009–1017 (2024).

2. Page, K. R. *et al.* Role of a holo-insertase complex in the biogenesis of biophysically diverse ER membrane proteins. *Mol Cell* **84**, 3302–3319.e11 (2024).

3. Rane, N. S., Yonkovich, J. L. & Hegde, R. S. Protection from cytosolic prion protein toxicity by modulation of protein translocation. *EMBO J* **23**, 4550–4559 (2004).

4. Drisaldi, B. *et al.* Mutant PrP is delayed in its exit from the endoplasmic reticulum, but neither wild-type nor mutant PrP undergoes retrotranslocation prior to proteasomal degradation. *J Biol Chem* **278**, 21732–21743 (2003).

5. Jung, S.-J., Kim, J. E. H., Junne, T., Spiess, M. & Kim, H. Cotranslational Targeting and Posttranslational Translocation can Cooperate in Spc3 Topogenesis. *J Mol Biol* **433**, 167109 (2021).

6. Kim, S. J., Rahbar, R. & Hegde, R. S. Combinatorial control of prion protein biogenesis by the signal sequence and transmembrane domain. *J Biol Chem* **276**, 26132–26140 (2001).

7. Ji, J. *et al.* An ATP13A1-assisted topogenesis pathway for folding multi-spanning membrane proteins. *Mol. Cell* **84**, 1917–1931.e15 (2024).

8. Lemberg, M. K. & Strisovsky, K. Maintenance of organellar protein homeostasis by ER-associated degradation and related mechanisms. *Mol Cell* **81**, 2507–2519 (2021).

9. Holaska, J. M. *et al.* Calreticulin Is a receptor for nuclear export. *J Cell Biol* **152**, 127–140 (2001).

10. McKenna, M. J., Adams, B. M., Chu, V., Paulo, J. A. & Shao, S. ATP13A1 prevents ERAD of folding-competent mislocalized and misoriented proteins. *Mol Cell* **82**, 4277–4289.e10 (2022).

11. Wu, H. & Hegde, R. S. Mechanism of signal-anchor triage during early steps of membrane protein insertion. *Mol Cell* **83**, 961–973.e7 (2023).

12. Wu, X. & Rapoport, T. A. Translocation of Proteins through a Distorted Lipid Bilayer. *Trends Cell Biol* **31**, 473–484 (2021).

13. Guna, A. *et al.* A dual sgRNA library design to probe genetic modifiers using genome-wide CRISPRi screens. *bioRxiv* (2023) doi:10.1101/2023.01.22.525086.

14. Chen, Z. *et al.* Structure of an EMC chaperone: CaV complex reveals an ion channel assembly intermediate and a holdase function for the EMC. *Biophys. J.* **123**, 14a (2024).

15. Emerman, A. B., Zhang, Z.-R., Chakrabarti, O. & Hegde, R. S. Compartment-restricted biotinylation reveals novel features of prion protein metabolism in vivo. *Mol Biol Cell* **21**, 4325–4337 (2010).

16. Goder, V., Crottet, P. & Spiess, M. In vivo kinetics of protein targeting to the endoplasmic reticulum determined by site-specific phosphorylation. *EMBO J.* **19**, 6704–6712 (2000).

17. Castrucci, M. R. *et al.* The cysteine residues of the M2 protein are not required for influenza A virus replication. *Virology* **238**, 128–134 (1997).

18. Thaa, B., Levental, I., Herrmann, A. & Veit, M. Intrinsic membrane association of the cytoplasmic tail of influenza virus M2 protein and lateral membrane sorting regulated by cholesterol binding and palmitoylation. *Biochem J* **437**, 389–397 (2011).

19. Rossman, J. S., Jing, X., Leser, G. P. & Lamb, R. A. Influenza virus M2 protein mediates ESCRT-independent membrane scission. *Cell* **142**, 902–913 (2010).

20. Martyna, A. *et al.* Membrane remodeling by the M2 amphipathic helix drives influenza virus membrane scission. *Sci Rep* **7**, 44695 (2017).

21. Meng, X., Templeton, C., Clementi, C. & Veit, M. The role of an amphiphilic helix and transmembrane region in the efficient acylation of the M2 protein from influenza virus. *Sci Rep* **13**, 18928 (2023).

22. Martin-Sancho, L. *et al.* Restriction factor compendium for influenza A virus reveals a mechanism for evasion of autophagy. *Nat Microbiol* **6**, 1319–1333 (2021).

23. Schubert, U. *et al.* The human immunodeficiency virus type 1 encoded Vpu protein is phosphorylated by casein kinase-2 (CK-2) at positions Ser52 and

Ser56 within a predicted alpha-helix-turn-alpha-helix-motif. *J Mol Biol* **236**, 16–25 (1994).

24. Ast, T., Cohen, G. & Schuldiner, M. A network of cytosolic factors targets SRP-independent proteins to the endoplasmic reticulum. *Cell* **152**, 1134–1145 (2013).

25. Hegde, R. S. & Kang, S.-W. The concept of translocational regulation. *J Cell Biol* **182**, 225–232 (2008).

26. Rane, N. S., Kang, S.-W., Chakrabarti, O., Feigenbaum, L. & Hegde, R. S. Reduced translocation of nascent prion protein during ER stress contributes to neurodegeneration. *Dev Cell* **15**, 359–370 (2008).

27. Eisenberg, D., Wilcox, W. & McLachlan, A. D. Hydrophobicity and amphiphilicity in protein structure. *J Cell Biochem* **31**, 11–17 (1986).

28. McKenna, M. J. *et al.* The endoplasmic reticulum P5A-ATPase is a transmembrane helix dislocase. *Science* **369**, (2020).

29. Song, J. & Becker, T. Fidelity of organellar protein targeting. *Curr Opin Cell Biol* **75**, 102071 (2022).

30. Hsieh, H.-H. & Shan, S.-O. Fidelity of Cotranslational Protein Targeting to the Endoplasmic Reticulum. *Int J Mol Sci* **23**, (2021).

31. Shaffer, K. L., Sharma, A., Snapp, E. L. & Hegde, R. S. Regulation of protein compartmentalization expands the diversity of protein function. *Dev Cell* **9**, 545–554 (2005).

32. Kang, S.-W. *et al.* Substrate-specific translocational attenuation during ER stress defines a pre-emptive quality control pathway. *Cell* **127**, 999–1013 (2006).

33. Guna, A., Volkmar, N., Christianson, J. C. & Hegde, R. S. The ER membrane protein complex is a transmembrane domain insertase. *Science* **359**, 470–473 (2018).

34. Laskowski, P. R. *et al.* Monitoring the binding and insertion of a single transmembrane protein by an insertase. *Nat Commun* **12**, 7082 (2021).

35. Nishimura, T., Lazzeri, G., Mizushima, N., Covino, R. & Tooze, S. A. Unique

amphipathic α helix drives membrane insertion and enzymatic activity of ATG3. *Sci Adv* **9**, eadh1281 (2023).

36. Arad, S. *et al.* A new targeting motif for endoplasmic reticulum surface proteins. *bioRxiv* (2024) doi:10.1101/2024.04.22.590521.

37. Chen, Z. *et al.* EMC chaperone-Ca structure reveals an ion channel assembly intermediate. *Nature* **619**, 410–419 (2023).

38. Tian, S. *et al.* Proteomic Analysis Identifies Membrane Proteins Dependent on the ER Membrane Protein Complex. *Cell Rep* **28**, 2517–2526.e5 (2019).

39. Christianson, J. C. & Carvalho, P. Order through destruction: how ER-associated protein degradation contributes to organelle homeostasis. *EMBO J* **41**, e109845 (2022).

CHAPTER 5: CONCLUSIONS AND PERSPECTIVES

Membrane protein targeting and translocation is a time sensitive stochastic process, for most proteins preceding at the speed of translation. Fidelity in targeting proteins to the correct organelle [1–3] and flexibility between translocation routes [4–7] are essential to maintain membrane identity and avoid excessive failure to insert into a membrane. Buffering mechanisms are necessary for robust insertion not only of a single substrate but for all diversity of simultaneously translated membrane proteins engaging with the ER membrane.

We show in Chapter 2 that this plasticity in membrane protein biogenesis exists for the influenza A virus encoded typeIII membrane protein M2, as M2 biogenesis is sensitive to the knockout of multiple members of the Oxa1 family of insertases (Figure 2.2), and that in absence of EMC, residual amounts of M2 can traffic to the plasma membrane and efficiently function in viral replication (Figure 2.5).

The diversity of ER translocons exists to serve the diversity of transmembrane proteins, but recent evidence demonstrates a significant overlap between translocation routes [6,7]. These biogenesis pathways are not exclusive to a specific type of membrane protein, but allow for significant overlap between pathways depending on the nascent chain encoded signals [6]. In Chapter 3, we show that M2's transmembrane domain and C-terminus harbor signals that require EMC processing (Figure 3.3), but that M2's EMC dependence can be overwritten by C-terminal extension (Figure 3.4). It may be possible that an increased time of translation during sampling for a translocon makes M2 more efficient in competing for translocation with simultaneously targeted membrane proteins. Possibly ribosome associated factors may increase translocation efficiency, over post-translationally targeted proteins.

It is tempting to speculate that IAV membrane proteins may evolve a balance in their competition for cellular resources in order to achieve optimal expression dynamics for viral particle production. Hierarchical processing of membrane proteins depending on their specific folding rate has been identified in prokaryotes [8]. This may similarly be conserved in eukaryotes, for example, through variations in signal strengths [9] and other cis-acting biogenesis signals [6]. *In vitro* translation of a panel of typeIII reporter constructs with different transmembrane domains are differentially affected by the absence of EMC as translocation sites for yet undefined reasons [10]. Early *in vitro* experiments show that co-translationally targeted membrane proteins can compete for the same translocation sites [11]. Competition for resources during membrane protein biogenesis has been demonstrated *in vivo* during antibody secretion, and can be computationally predicted [12]. The knockout of a highly

expressed surface marker increases antibody production in these cells, although it is unclear if this is due to energy demands or occupancy of biogenesis factors [12]. Our findings suggest that competition for ER insertion sites is a major factor influencing translocation efficiency. Several observations support this mechanism. 1) M2 signal peptide mutants enhance M2 translocation efficiency without involving Sec61 translocation. This suggests that enhanced targeting of M2 to the ER may allow it to outcompete post-translationally targeted proteins, improving insertion efficiency. 2) GET3 and TMEM208 knockouts (KO) increase M2 expression. GET3 targets tail-anchored proteins to the ER, while TMEM208 facilitates SRP-targeted membrane protein handover to translocons. In the absence of these factors, fewer competing proteins may reach the ER membrane, reducing insertion competition and indirectly enhancing M2 biogenesis. 3) The KO of several Oxa1-family members has comparable effect on M2 expression. Shunting of membrane proteins to residual translocation sites may increase competition regardless of the depleted translocon. 4) While translocation efficiency of an amphipathic helix mutant of M2 is strongly affected in transfected WT cells, a virus expressing the same mutant expresses M2 to WT levels at the plasma membrane. IAV infection is known to shut down host translation [13], which may reduce competition for ER translocation sites, increasing M2 insertion efficiency. We observe that M2-AHmut is not able to accumulate in absence of EMC in transfection, but can accumulate at the plasma membrane in infection (Figure 4.4 and 4.6). We hypothesize that the increase in efficiency may be due to viral-induced host shutoff, and elimination of competing translation and translocation events. The competition for certain translocation machinery under physiological conditions, needs to be buffered to make sure that costly translated proteins are not unnecessarily degraded, or perhaps form toxic cytosolic aggregates if outcompeted.

In Chapter 4, we demonstrate a possible mechanism for plasticity in membrane protein biogenesis, where cytosolic membrane association of a non-inserted, fully translated membrane protein, protects it from degradation until a translocon is engaged for translocation. Future studies need to assess the generality of this mechanism. We find that the typeIII protein TEMP has a stable unglycosylated band upon EMC ablation, although it does not have a significant amphipathic helix close to its transmembrane domain. Future studies will show how buffering mechanisms during membrane protein biogenesis protect from disease.

It is likely that membrane association of membrane proteins prior to translocation has evolutionarily preceded the complex co-translational insertion pathways that exist today. Oxa1 family of insertases present a highly conserved protein fold that can locally collapse the membrane. However, lipid bilayer collapse

may not be exclusive to this fold, and may have existed in a protocellular system. Modern molecular simulation approaches are well suited to screen various membrane protein folds for their effect on lipid bilayer stability and potentially uncover new membrane protein complexes with translocase activity.

At low multiplicity of IAV infection, viral protein translation is highly stochastic [14]. Nevertheless, relative amounts of viral protein components matter for the rate and amount of infectious virus particle production [13]. We show in Chapter 2 that a defect in the accumulation rate of M2 at the plasma membrane has an effect on the onset of viral replication. Viral shedding is highly heterogeneous between human hosts [15,16]. IAV-infected ferrets are the established model to investigate the parameters that contribute to viral transmission [17]. A metaanalysis of many transmission studies finds that high viral titers early after infection (1 day), are beneficial for transmission via droplets, or direct contact, but that high viral titers later only favours transmission via direct contact [17]. More studies need to be conducted in order to find the viral protein expression programs that contribute to high early titers. The high expression of M2 in lab-adapted viruses may be beneficial in a cell culture model, but detrimental for a complex infection in a host. At this point, we can only speculate how M2 expression dynamics may impact human to human transmission events.

My work establishes M2 membrane association as a crucial step in efficient M2 biogenesis. However, further questions within M2 biology and membrane protein biogenesis dynamics are important to answer in order to elucidate molecular events contributing to the effect and establish the importance of this mechanism for cellular function.

5.1 Open questions

1) Can the overexpression of GET or GEL compensate for loss of EMC in M2 biogenesis?
2) Does membrane attachment of other proteins with terminal signal anchor and small translocated domains allow flexibility in translocon selection?
3) Can hydrophobic amino acid substitutions in M2's transmembrane domain compensate for the disruption of the amphipathic helix?
4) Is M2 translocation rate influenced by competition with other translocated proteins, and if yes which ones?
5) Is M2 translocation efficiency affected by viral host shutoff?
6) Can membrane association increase translocation efficiency of membrane proteins known to be inefficient like prion protein?

7) How does lipid environment affect the membrane association of non-inserted membrane proteins?
8) Which physiological processes are facilitated by pre-translocation membrane association? Can the process be regulated or cause disease?
9) Are there other cytosolic molecular factors involved to stabilise M2 at the membrane face?
10) Is M2 flexible in its mode of targeting to the ER membrane and can be targeted co-translationally or post-translationally?
11) Does M2 C-terminal extension influence cytosolic targeting or other aspects of M2 biogenesis?
12) How does the amphipathic helix influence M2 dimer formation?

5.2 References

1. Backes, S. *et al.* The chaperone-binding activity of the mitochondrial surface receptor Tom70 protects the cytosol against mitoprotein-induced stress. *Cell Rep* **35**, 108936 (2021).

2. McKenna, M. J. *et al.* The endoplasmic reticulum P5A-ATPase is a transmembrane helix dislocase. *Science* **369**, (2020).

3. Hsieh, H.-H. & Shan, S.-O. Fidelity of Cotranslational Protein Targeting to the Endoplasmic Reticulum. *Int J Mol Sci* **23**, (2021).

4. Guna, A. *et al.* A dual sgRNA library design to probe genetic modifiers using genome-wide CRISPRi screens. *BMC Genomics* **24**, 651 (2023).

5. Cohen, N., Aviram, N. & Schuldiner, M. A systematic proximity ligation approach to studying protein-substrate specificity identifies the substrate spectrum of the Ssh1 translocon. *EMBO J* **42**, e113385 (2023).

6. Page, K. R. *et al.* Role of a holo-insertase complex in the biogenesis of biophysically diverse ER membrane proteins. *Mol Cell* **84**, 3302–3319.e11 (2024).

7. Wu, H. & Hegde, R. S. Mechanism of signal-anchor triage during early steps of membrane protein insertion. *Mol Cell* **83**, 961–973.e7 (2023).

8. Zhu, Z., Wang, S. & Shan, S.-O. Ribosome profiling reveals multiple roles of SecA in cotranslational protein export. *Nat Commun* **13**, 3393 (2022).

9. Chen, H., Kim, J. & Kendall, D. A. Competition between functional signal peptides demonstrates variation in affinity for the secretion pathway. *J Bacteriol* **178**, 6658–6664 (1996).

10. Chitwood, P. J., Juszkiewicz, S., Guna, A., Shao, S. & Hegde, R. S. EMC Is Required to Initiate Accurate Membrane Protein Topogenesis. *Cell* **175**, 1507–1519.e16 (2018).

11. Lingappa, V. R., Katz, F. N., Lodish, H. F. & Blobel, G. A signal sequence for the insertion of a transmembrane glycoprotein. Similarities to the signals of secretory proteins in primary structure and function. *J. Biol. Chem.* **253**, 8667–8670 (1978).

12. Gutierrez, J. M. *et al.* Genome-scale reconstructions of the mammalian secretory pathway predict metabolic costs and limitations of protein secretion. *Nat Commun* **11**, 68 (2020).

13. Bogdanow, B. *et al.* The dynamic proteome of influenza A virus infection identifies M segment splicing as a host range determinant. *Nat Commun* **10**, 5518 (2019).

14. Heldt, F. S., Kupke, S. Y., Dorl, S., Reichl, U. & Frensing, T. Single-cell analysis and stochastic modelling unveil large cell-to-cell variability in influenza A virus infection. *Nat Commun* **6**, 8938 (2015).

15. Shetty, N. *et al.* Influenza virus infection and aerosol shedding kinetics in a controlled human infection model. *J Virol* **98**, e0161224 (2024).

16. Hooker, K. L. & Ganusov, V. V. Impact of Oseltamivir Treatment on Influenza A and B Virus Dynamics in Human Volunteers. *Front Microbiol* **12**, 631211 (2021).

17. Kieran, T. J., Sun, X., Maines, T. R. & Belser, J. A. Optimal thresholds and key parameters for predicting influenza A virus transmission events in ferrets. *Npj Viruses* **2**, 64 (2024).