

**Understanding respiratory syncytial virus entry requirements by mapping the interaction
between the viral fusion protein and the insulin-like growth factor-1 receptor**

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree
of Master's of Science (M.Sc.)

Rachel S. Hayes

Department of Biochemistry

McGill University, Montreal

August 2022

© Rachel S. Hayes 2022

ABSTRACT

Human respiratory syncytial virus (RSV) is a global health threat with approximately 33 million people infected and approximately 120,000 deaths each year. Respiratory syncytial virus (RSV) has two main surface glycoproteins, the attachment glycoprotein (G) and the fusion (F) protein, which together mediate viral attachment and entry. Recent studies have revealed a new model for RSV entry. In this model, attachment is mediated by non-specific interactions between the RSV-G protein and glycosaminoglycans, and subsequently the RSV-F protein can interact specifically with the insulin-like growth factor 1 receptor (IGF1R). This interaction triggers IGF1R activation, culminating in the activation of PKC ζ signaling and the recruitment of an additional RSV-F receptor, nucleolin, from the nucleus to the cell surface. At the cell surface, RSV-F, IGF1R, and nucleolin form a trimeric complex that triggers RSV fusion and completes the RSV entry process. However, the specific residues of IGF1R that are important for interactions with RSV-F are still unclear. As such, we performed molecular docking analyses and identified 35 residues in IGF1R that may be important for interactions with RSV-F. We used alanine scanning mutagenesis to generate 35 mutants of IGF1R and we assessed their expression, maturation, as well as the effect of each mutation on RSV entry. We identified several mutations that appear to inhibit IGF1R maturation, but surprisingly these mutations did not have significant effects on RSV entry. This suggests that IGF1R maturation may not be required for RSV entry. Additionally, we identified one residue, S788A, that significantly reduced RSV infection, revealing a potentially critical residue for RSV-F-mediated entry. Future studies will be needed to further explore cell surface localization of IGF1R alanine mutants and verify the importance of Ser788 in RSV entry. Finally, to develop IGF1R-based biological decoy receptors for RSV entry, we generated a mutant IGF1R library using sequence saturation mutagenesis, targeting the region of IGF1R identified by our

molecular docking analysis. In future, we plan to establish a directed evolution platform and flow cytometry-based screening assay to develop biological inhibitors of RSV entry. We anticipate that this research will help to map the RSV-F:IGF1R interaction interface, and will aid in the development of targeting strategies that inhibit RSV entry. We hope that the establishment of the directed evolution platform will lead to the development of novel biological inhibitors of RSV entry, and that this approach can be applied to the development of antivirals for a variety of virus-host interactions.

RÉSUMÉ

Le virus respiratoire syncytial humain (VRS) est une menace pour la santé mondiale, avec environ 33 millions de personnes infectées et environ 120 000 décès chaque année. Le virus respiratoire syncytial (VRS) possède deux glycoprotéines de surface principales, la glycoprotéine d'attachement (G) et la protéine de fusion (F), qui ensemble assurent la médiation de l'attachement et de l'entrée du virus. Des études récentes ont révélé un nouveau modèle d'entrée du VRS. Dans ce modèle, l'attachement est facilité par des interactions non spécifiques entre la protéine G du VRS et les glycosaminoglycanes, puis la protéine F du VRS peut interagir spécifiquement avec le récepteur du facteur de croissance analogue à l'insuline de type 1 (IGF1R). Cette interaction déclenche l'activation de l'IGF1R, qui aboutit à l'activation de la signalisation de la PKC ζ et au recrutement d'un autre récepteur du VRS-F, la nucléoline, du noyau vers la surface cellulaire. À la surface des cellules, le VRS-F, l'IGF1R et la nucléoline forment un complexe trimérique qui déclenche la fusion du VRS et complète le processus d'entrée du VRS. Cependant, les résidus spécifiques de l'IGF1R qui sont importants pour les interactions avec le VRS-F ne sont toujours pas clairs. Ainsi, nous avons réalisé des analyses de docking moléculaire et avons identifié 35 résidus dans l'IGF1R qui pourraient être importants pour les interactions avec le VRS-F. Nous avons utilisé la mutagenèse à balayage d'alanine afin de générer 35 mutants d'IGF1R et nous avons évalué leur expression, leur maturation, ainsi que l'effet de chaque mutation sur l'entrée du VRS. Nous avons identifié plusieurs mutations qui semblent inhiber la maturation de l'IGF1R, mais étonnamment, ces mutations n'ont pas eu d'effets significatifs sur l'entrée du RSV. Ceci suggère que la maturation de l'IGF1R n'est peut-être pas nécessaire pour l'entrée du VRS. De plus, nous avons identifié un résidu, le S788A, qui réduit significativement l'infection par le VRS, révélant ainsi un résidu potentiellement critique pour l'entrée médiée par le VRS-F. Des études futures

seront nécessaires pour explorer davantage la localisation à la surface cellulaire des mutants alanine d'IGF1R et pour vérifier l'importance du résidu Ser788 dans l'entrée du VRS. Enfin, pour développer des leurres biologiques basés sur l'IGF1R pour l'entrée du VRS, nous avons généré une librairie de mutants de l'IGF1R en utilisant la mutagenèse par saturation de séquence, ciblant la région de l'IGF1R identifiée par notre analyse de docking moléculaire. À l'avenir, nous prévoyons d'établir une plateforme d'évolution dirigée et un test de criblage basé sur la cytométrie en flux pour développer des inhibiteurs biologiques de l'entrée du VRS. Nous pensons que cette recherche aidera à schématiser l'interface d'interaction VRS-F:IGF1R, et qu'elle contribuera au développement de stratégies de ciblage qui inhibent l'entrée du VRS. Nous espérons que l'établissement de la plateforme d'évolution dirigée mènera au développement de nouveaux inhibiteurs biologiques de l'entrée du VRS et que cette approche sera appliquée au développement d'antiviraux pour une variété d'interactions virus-hôte.

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my supervisor Dr. Selena Sagan. Thank you for showing me the amazing world of virology, first through your lectures, and then through taking me on in your lab. Your passion and love for science kept me motivated and curious during my degree. I will always be grateful for the guidance and mentorship you gave me, both inside and outside of the lab. Finally, thank you for fostering an amazing lab environment, your weekly homemade ice cream will be greatly missed. I am excited to see how the lab continues to grow along your new journey!

A big thank you to my past and present lab mates. Thank you, Jasmin, Trisha, and Sophie, for always being happy to answer my questions or sit down and talk something out. Thank you, Carolina, for always being able to resolve my crises. Marylin, for being my bay mate and going through this journey by my side. To Quinn, Qi, Breanna, and Manolya, you guys always made the hard days bearable.

I am grateful to my Research advisory committee members, Dr. Jerry Pelletier and Dr. Jacques Archambault, for their guidance and insight. I would also like to acknowledge the generous financial support provided by CIHR Frederick Banting and Charles Best Award and FRQ- S Master's Training Award, which made my research possible.

Thank you to my friends, Gaby, Jackie, Amanda, and Iris, for always being there when I needed to take a break and laugh.

Thank you to my family. Mom, for always supporting me, feeding me when I was too focused, and loving me unconditionally. To my big sister, Erin, for always looking out for me and pulling me out of my den of a room when I needed to see the light of day. To the both of you for listening to my presentations even when you had no idea what I was talking about. To my

dad, Rupert, who will never get to read my thesis, or see what I have done. None of it would have happened if you had not instilled such a passion for science and knowledge in me from a young age.

Finally, thank you to my partner, Emery. You have supported me in this process since day one and I know I wouldn't be here without you. Thank you for always being there to listen. Thank you for being my in-house editor. Thank you for being my person. You and our dog, Fern, will always make coming home the best part of my day.

PREFACE

This thesis was written according to the McGill University’s “Guidelines for Thesis Preparation”. The candidate has chosen to present in a “Manuscript-based thesis” format by following these guidelines:

“As an alternative to the traditional format, a thesis may be presented as a collection of scholarly papers of which the student is the first author or co-first author. A manuscript-based Master’s thesis must include the text of one or more manuscripts.”

“Manuscripts for publication in journals are frequently very concise documents. A thesis, however, is expected to consist of more detailed, scholarly work. A manuscript-based thesis will be evaluated by the examiners as a unified, logically coherent document in the same way a traditional thesis is evaluated.”

The information of the manuscript in preparation is listed below. The contributions of the authors are detailed in the “Contributions of authors” section.

Manuscript presented in this thesis:

Rachel S. Hayes, Ahmed K. Oraby, Carolina Camargo, David J. Marchant, and Selena M. Sagan. Understanding respiratory syncytial virus entry requirements by mapping interactions between the RSV-F and IGF1R. *Manuscript in Preparation*. 2022.

CONTRIBUTIONS OF AUTHORS

Chapter 2:

Rachel S. Hayes, Ahmed K. Oraby, Carolina Camargo, David J. Marchant, and Selena M. Sagan. Understanding respiratory syncytial virus entry requirements by mapping interactions between the RSV-F and IGF1R. *Manuscript in Preparation*. 2022.

R.S.H and S.M.S conceptualized and designed the study. D.J.M provided essential reagents and council throughout the study. A.K.O. performed the molecular docking experiment and analysis. R.S.H performed the alanine scanning expression experiments and analysis. C.C. performed the RSV infection experiments, C.C. and R.S.H performed the analysis. R.S.H. performed the SeSaM and library preparation experiments and analysis. R.S.H drafted the manuscript; R.S.H. and S.M.S revised and edited the final version

TABLE OF CONTENTS

ABSTRACT	II
RÉSUMÉ	IV
ACKNOWLEDGEMENTS	VI
PREFACE	VIII
TABLE OF CONTENTS	X
LIST OF FIGURES	XIV
LIST OF TABLES	XV
LIST OF ABBREVIATIONS	XVI
CHAPTER 1: INTRODUCTION	1
1.1 DISCOVERY AND EPIDEMIOLOGY OF RESPIRATORY SYNCYTIAL VIRUS	1
<i>1.1.1 Discovery</i>	1
<i>1.1.2 Epidemiology</i>	1
<i>1.1.3 Pathology and treatment</i>	2
1.2 MOLECULAR BIOLOGY OF RSV	2
<i>1.2.1 RSV structure and genome organization</i>	2
<i>1.2.2 RSV life cycle</i>	5
1.3 VIRAL ATTACHMENT AND ENTRY	7
<i>1.3.1 RSV-G and attachment</i>	7
<i>1.3.2 RSV-F and viral fusion</i>	8
1.4 IDENTIFICATION OF RSV ENTRY RECEPTORS.....	11

1.4.1 Nucleolin.....	11
1.4.2 IGF1R	14
1.4.3 IGF1R signaling.....	17
1.4.4 Current model for RSV attachment and entry.....	19
1.5 TOWARDS THE DEVELOPMENT OF BIOLOGICAL INHIBITORS OF RSV ENTRY	21
1.5.1 Viral Entry as a target of antiviral therapy	21
1.5.2 Directed evolution for the development of biological inhibitors of viral entry	21
1.6 RATIONALE AND HYPOTHESIS	23
1.7 REFERENCES.....	25
CHAPTER 2: UNDERSTANDING RSV ENTRY REQUIREMENTS BY MAPPING	
INTERACTIONS BETWEEN RSV-F AND IGF1R.....	35
2.1 PREFACE TO CHAPTER 2	36
2.2 ABSTRACT.....	37
2.3 INTRODUCTION.....	38
2.4 MATERIALS & METHODS.....	41
2.4.1 Cell lines and viruses	41
2.4.2 Preparation of Viral Stocks and Focus-Forming Unit (FFU) Assays.....	41
2.4.3 Molecular docking analyses	42
2.4.4 Plasmids and molecular cloning.....	43
2.4.5 IGF1R mutant expression and RSV infections.....	46
2.4.6 Flow Cytometry.....	46
2.4.7 Western blotting.....	47
2.4.8 Sequence saturation mutagenesis	47

2.4.9 <i>SeSaM Library preparation</i>	48
2.4.10 <i>Data and statistical analysis</i>	49
2.5 RESULTS	49
2.5.1 <i>Molecular docking of RSV-F with IGF1R</i>	49
2.5.2 <i>Expression analysis of immature and mature forms of IGF1R alanine mutants</i>	50
2.5.3 <i>With the exception of Ser788, all alanine mutants supported RSV infection to a similar extent as WT IGF1R.</i>	55
2.5.4 <i>SeSaM reaction and mutant library preparation</i>	57
2.6 DISCUSSION	61
2.7 ACKNOWLEDGMENTS.....	67
2.8 FUNDING INFORMATION	67
2.9 AUTHOR CONTRIBUTIONS	67
2.10 REFERENCES.....	68
CHAPTER 3: DISCUSSION	73
3.1 SUMMARY	73
3.2 FURTHER UNDERSTANDING OF THE ALANINE MUTANTS OF IGF1R.....	73
3.2.1 <i>Cell surface expression of alanine mutants of IGF1R</i>	73
3.2.2 <i>IGF1R activation</i>	74
3.2.3 <i>IGF1R:RSV-F binding affinity</i>	75
3.2.4 <i>Post-translational modifications of IGF1R</i>	76
3.3 DIRECTED EVOLUTION ASSAY	77
3.3.1 <i>Directed Evolution using SeSAM</i>	77
3.3.2 <i>Assay considerations and challenges</i>	79

3.3.3 <i>Application of the directed evolution platform to other virus-host interactions</i>	80
3.4 CONCLUDING REMARKS	80
3.5 REFERENCES	82
APPENDIX	85

LIST OF FIGURES

Figure 1.1. RSV particle structure and genome organization.....	3
Figure 1.2. <i>Pneumoviridae</i> viral life cycle.....	6
Figure 1.3. Domain organization and structure of RSV-F.....	9
Figure 1.4. Structures and antigenic sites of pre- and postfusion RSV-F.....	10
Figure 1.5. Domain organization of nucleolin.....	13
Figure 1.6. Domain organization and structure of IGF1R.....	15
Figure 1.7. IGF1R rearrangement upon IGF1 binding.....	18
Figure 1.8. RSV entry model.....	20
Figure 1.9. Directed evolution for development of biological inhibitors of viral entry.....	22
Figure 2.1. IGF1R domain organization and molecular docking of IGF1R:RSV-F.....	40
Figure 2.2. Expression of alanine mutants of IGF1R.....	52
Figure 2.3. Quantification of immature and mature IGF1R expression of alanine mutants.	53
Figure 2.4. RSV infection of IGF1R KO HAE cells expressing alanine mutants.....	56
Figure 2.5. Distribution of mutations across IGF1R target sequence.....	59
Figure 2.6. Flow cytometry-based screening assay to select for IGF1R receptors with a desired activity.....	65

LIST OF TABLES

Table 2.1 Primers used in SDM.	44
Table 2.2 IGF1R residues that interact with RSV-F.	51
Table 2.3 Key expected performance parameters of SeSaM compared to obtained results.	58
Table 2.4 Key expected performance parameters for a mutant library preparation compared to obtained results.	60

LIST OF ABBREVIATIONS

α -DG	Alpha-dystroglycan
α CT	Alpha chain c-terminal
A647	Alexa fluor 647
APC	Adenomatous polyposis coli
CCA	Chimpanzee coryza agent
cDNA	Complementary DNA
CFU	Colony forming units
CR	Cysteine-rich domain
CREB	cAMP response element-binding protein
dPTP α S	6-(2-deoxy- β -D-ribofuranosyl)-3,4-dihydro-8H- pyrimido-[4,5-C][1,2]oxazin-7-one
dRTP	1- β -D-ribofuranosyl-1,2,4-triazole-3-carboxamide
EBNA	Epstein-Barr virus nuclear antigen-1
epPCR	Error prone polymerase chain reaction
FACS	Fluorescent activated cell sorting
FBS	Fetal Bovine Serum
FFU	Focus-forming unit
FnIII	Type III fibronectin domain
FP	Fusion peptide
GAG	Glycosaminoglycan
GAR	Glycine/arginine-rich
GFP	Green fluorescent protein

GPI	Glycosylphosphatidylinositol
HAE	Human air epithelial
Hek	Human embryonic kidney
HIV	Human immunodeficiency virus
HR	Heptad repeat
HS	Heparan sulfate
HSPG	Heparin sulfate proteoglycans
ICAM-1	Intercellular adhesion molecule 1
ID	Insert domain
IGF	Insulin-like growth factor
IGF1R	Insulin-like growth factor 1 receptor
IR	Insulin receptor
KO	Knockout
LAMP1	Lysosomal-associated membrane protein 1
LCMV	Lymphocytic choriomeningitis virus
MAPK	Mitogen-activated protein kinase
NCL	Nucleolin
PBS	Phosphate-buffered saline
PI3K	Phosphoinositide-3 kinase
PKC ζ	Protein kinase C zeta
RBD	RNA binding domains
RdRp	RNA-dependent RNA polymerase
RNP	Ribonucleoprotein

RSV	Respiratory syncytial virus
SDS-PAGE	Sodium dodecyl sulfate-polyacrylamide
SeSaM	Sequence saturation mutagenesis
StEP	Staggered extension process
TLR4	Toll-like receptor-4
TK	Tyrosine kinase domain
VOPBA	Virus overlay protein binding assay
WT	Wild-type

CHAPTER 1: INTRODUCTION

1.1 DISCOVERY AND EPIDEMIOLOGY OF RESPIRATORY SYNCYTIAL VIRUS

1.1.1 Discovery

Respiratory syncytial virus (RSV) is a negative-sense RNA virus of the *Pneumoviridae* family. RSV was first discovered in 1956 in a group of chimpanzees suffering from cold symptoms, and was thus originally named *chimpanzee coryza agent* (CCA) (1). However, CCA was later isolated in two children suffering from respiratory illnesses, and researchers also found antibodies that could neutralize CCA in most school-aged children (1). Once they realized its ability to cause human illness, CCA was renamed “respiratory syncytial virus”.

1.1.2 Epidemiology

RSV is the leading cause of severe lower respiratory tract disease globally, with approximately 33 million cases annually (2). Of these, 3 million cases require hospitalization and approximately 118,200 fatalities occur per year (2, 3). By the age of 2, virtually all children have had at least one RSV infection (4). RSV also accounts for high rates of hospitalization and morbidity in the elderly population (5). RSV infections follow a seasonal transmission pattern. In temperate climates, most epidemics occur during the winter, while in tropical climates this time frame is typically elongated, with infections beginning midsummer (6, 7). There are two RSV subtypes, subtypes A and B, which are differentiated by their antigenic reactivity to monoclonal antibodies (7). Typically, RSV subtype A is more prevalent, but models have been proposed for alternating seasonal transmission cycles of subtypes A and B (7). Transmission of RSV occurs as a result of contact with droplets, either through direct contact with an infected individual or through contact

with contaminated fomites (8). Recently, it was shown that infectious virus can be found in aerosols produced by infants with RSV-positive bronchiolitis (9).

1.1.3 Pathology and treatment

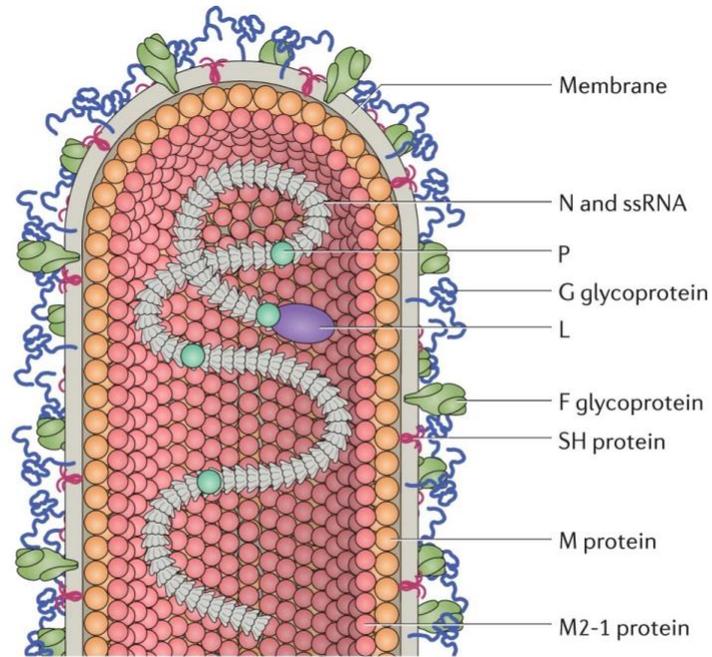
For the majority of children and infants, RSV symptoms are no more serious than a cold, but in more severe cases, life-threatening bronchiolitis and pneumonia can develop (4). Certain high-risk groups, including premature babies, immunocompromised infants, and those born with lung or heart disease, have a greater chance of developing severe infections (4, 10). Additionally, RSV bronchiolitis in infancy has been associated with the development of asthma and allergy in childhood (11). Currently, the only available treatment for RSV is a prophylactic monoclonal antibody, Palivizumab, which is reserved for high-risk patients, such as premature and/or immunocompromised infants (12). However, Palivizumab is costly, has low neutralizing activity, and is not effective if administered during an active RSV infection (13). As such, there is a pressing need for novel RSV vaccines and antiviral treatment strategies (12).

1.2 MOLECULAR BIOLOGY OF RSV

1.2.1 RSV structure and genome organization

RSV is an enveloped, non-segmented, negative-sense RNA virus of the *Pneumoviridae* family, genus orthopneumovirus (14). RSV particles have been described as spherical, asymmetric, and filamentous, but most RSV particles are filamentous (15, 16) (**Figure 1.1A**). The cellular plasma membrane-derived lipid bilayer envelope of RSV is decorated with the viral surface glycoproteins (1, 16). The RSV genome is 15.2 kb in length and encodes 10 genes, in the order 3' NS1-NS2-N-P-M-SH-G-F-M2-L 5' (1, 16).

A



B



Figure 1.1. RSV particle structure and genome organization. (A) Schematic of an RSV infectious viral particle. The attachment glycoprotein (G) and fusion (F) protein, as well as the small hydrophobic (SH) proteins are embedded in the viral envelope. The matrix (M) protein lies underneath the viral envelope. The M2-1 protein interacts with both the M protein and the nucleoprotein (N) encasing the viral negative-sense genomic RNA. The polymerase subunit (L) and the phosphoprotein (P) polymerase co-factor are also associated with N. (B) Organization of the RSV genome, with the 10 viral genes encoding 11 mature proteins. *Figure from (17).*

Each gene encodes a corresponding mRNA, with each encoding a single major protein, except for M2, which has two overlapping open reading frames (**Figure 1.1B**) (1, 16). The viral nucleoprotein (N) is involved in replication organelle biogenesis and serves as the nucleocapsid protein, which complexes with the viral RNA in a helical ribonucleoprotein (RNP) complex (16). The RSV polymerase complex is formed by the viral L, P, and M2-1 proteins, with the L protein being the RNA-dependent RNA polymerase (RdRp) responsible for viral mRNA synthesis and genome replication. The phosphoprotein (P) is a polymerase co-factor that also acts as an adaptor by binding to the N, M2-1, and the L protein to mediate the interaction between the nucleocapsid and polymerase complex (16). M2-1 is an essential transcription processivity factor, primarily involved in the complex as a transcription terminator (16, 18). The nonstructural (NS) proteins NS1 and NS2 inhibit interferon induction, signaling, and apoptosis (16). The role of M2-2 is unknown, but deletion of this protein results in delayed and reduced viral RNA replication (16).

Together with the N protein, the M, SH, G, and F proteins are the structural proteins that form the filamentous RSV particle (16). The matrix (M) protein associates with the cytoplasmic domain of the F protein and is responsible for the filamentous appearance of the virus (17). The M protein also interacts with M2-1, which mediates association with the polymerase complex (17). The viral envelope contains the fusion (F), attachment glycoprotein (G), and small hydrophobic (SH) proteins. The G and F proteins mediate viral attachment and entry, while the SH protein is a pentameric ion channel that is hypothesized to be involved in delaying apoptosis in infected cells (17, 19).

1.2.2 RSV life cycle

RSV preferentially infects the most apical ciliated cells of the airway epithelium (17, 20). Non-specific viral attachment is mediated by the RSV-G protein via its heparin-binding domains, which interact with cell surface glycosaminoglycans (GAGs), mainly heparin sulfate proteoglycans (HSPG) (7, 17, 21). Several studies also suggest that RSV-G interacts with the CX3CR1 receptor on ciliated cells of the airway epithelium (19, 22). Additionally, the F protein has been shown to interact with intercellular adhesion molecule-1 (ICAM-1), toll-like receptor-4 (TLR4), nucleolin, and more recently, the insulin-like growth factor 1 receptor (IGF1R) (23-27). RSV-F interactions with IGF1R and nucleolin trigger F protein-mediated fusion with host cell membranes, and current models suggest that RSV can fuse with either the plasma membrane or with endocytic vesicles (17, 26, 27). Differing efficiencies of these events might depend on the environmental conditions and the target cell (17).

After entry into the host cell, the viral RNP and the polymerase complex are released into the cell cytoplasm (**Figure 1.2**). Viral replication takes place in the cytoplasm at replication complexes which form on internal membranes (7). These complexes form by the coalescence of the RSV L, P, M2-1 proteins, and the N-encapsidated genomic RNA (7). The L protein (RdRp) transcribes the viral mRNAs and is responsible for the production of the full-length, positive-sense antigenomic intermediate (17). The current model for the switch between transcription and replication depends on the availability of the viral N protein and on which nucleotide(s) is/are pre-loaded in the viral polymerase upon non-templated RNA synthesis. If pre-loaded with G, the polymerase will transcribe mRNA, but if pre-loaded with AC, the polymerase will synthesize the full-length antigenome and subsequently full-length genomic RNAs (7, 28).

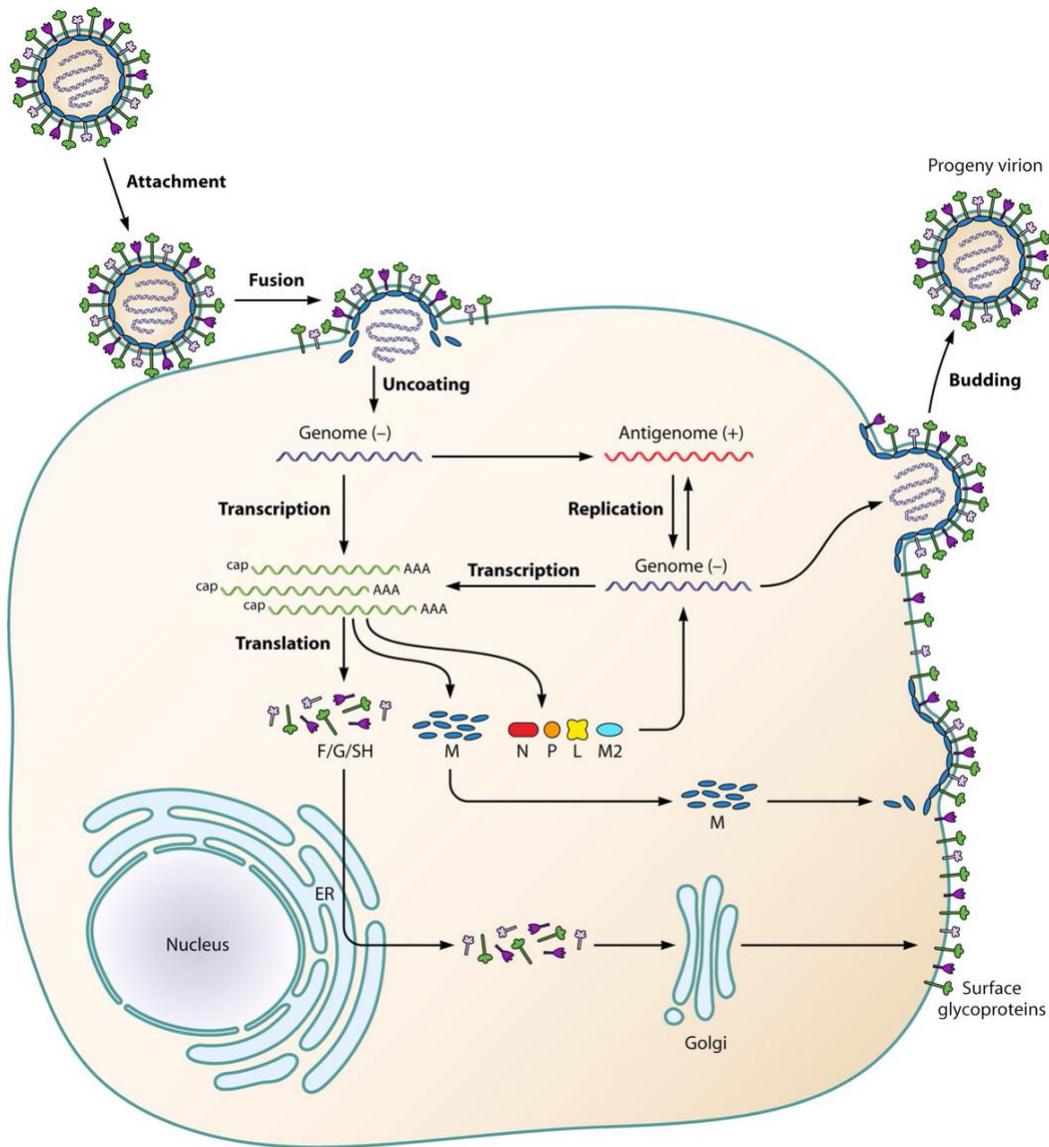


Figure 1.2. *Pneumoviridae* viral life cycle. The RSV life cycle initiates with attachment to the host cell membrane. In the case of RSV, fusion may occur at the plasma membrane (depicted) or within endocytic vesicles. Upon fusion, the virion is uncoated, and the negative-sense genomic RNA coated by the viral nucleoprotein (N) is released into the cytoplasm. The genome is first used to transcribe viral mRNA and subsequently to produce antigenomic and progeny negative-sense genomic RNAs. The progeny genomic RNAs are also incorporated into progeny virions or may act as a template for secondary transcription. Once translated, the M proteins and ribonucleoprotein complexes are transported to the plasma membrane. Viral glycoproteins F, G, and SH are synthesized on the ER and transit through the Golgi apparatus to the plasma membrane. Finally, new progeny virions are assembled and released from the plasma membrane by budding. *Figure from (29).*

During RSV infection, the primary sites of viral RNA synthesis, known as inclusion bodies, are formed and these play a role in modifying the antiviral response (7, 30). The viral RNPs are transported to the plasma membrane for virion assembly, where actin-dependent filament budding is driven by the RSV-F and M proteins associating with lipid rafts (16, 17). However, there is evidence that these filaments might form before reaching the plasma membrane and then merge with the plasma membrane through an unknown mechanism (17, 31). The viral RNP and associated polymerase complex interacts with the M protein, and infectious viral particles bud from the plasma membrane gaining a viral envelope in the process containing the SH, G and F proteins (**Figure 1.2**).

1.3 VIRAL ATTACHMENT AND ENTRY

1.3.1 RSV-G and attachment

RSV-G is a heavily glycosylated type II integral membrane glycoprotein that facilitates initial attachment to host cells during infection (19, 32). It resembles mucins produced in airways, as it has upwards of 30 O-linked glycans and 4-5 N-linked glycans (19). RSV-G also has a region similar to the CX3C motif of the chemokine CX3CL1, also known as fractalkine (33). Several host factors have been identified as RSV-G binding partners. While RSV-G is not absolutely required for infection, *in vitro* assays suggest that deletion of RSV-G leads to a 10-fold reduction in infectivity (34). On immortalized cells, it first appeared that heparan sulfate (HS) proteoglycans were the receptor for RSV-G (19). However, these are not present on the apical surface of human air epithelial (HAE) cells, and therefore HS is most likely not the true physiological receptor (19). Some studies have since indicated the fractalkine receptor, CX3CR1,

as the receptor for RSV-G. Possible evidence of this has been shown in HAE cultures and animal models (22, 33, 35). However, the role of CX3CR1 in the RSV life cycle remains controversial.

1.3.2 RSV-F and viral fusion

The structure of RSV-F is highly conserved in both RSV subtypes (19). RSV-F is a type I viral fusion protein, similar to the influenza hemagglutinin protein or the human immunodeficiency virus (HIV) envelope protein (36). The RSV-F glycoprotein is first produced as a 574 amino acid inactive precursor, called F₀ (37). A trimer of three F₀ monomers then assembles and is cleaved by a furin-like host protease (19). Each monomer is cleaved twice after amino acid 109 and 136, producing three fragments (38). The smaller N-terminal fragment (F₂) is attached to the larger C-terminal fragment (F₁) by two disulfide bonds (**Figure 1.3**) (36). The intervening segment dissociates after cleavage and is not present as a part of the mature F protein (39). F₂ is predicted to have two N-glycosylation sites, whereas F₁ is predicted to have one N-glycosylation site (19, 36). The mature form of the RSV-F protein is a functional trimer (19).

RSV-F-protein mediated fusion occurs due to the difference in energies between two F protein trimer states, specifically the prefusion and postfusion states (**Figure 1.3 and 1.4**) (40). The prefusion state is a metastable, untriggered version adopted before virion interactions with the host cell, whereas the postfusion state is a stable rearrangement that typically forms after fusion of the viral and host membranes (19, 40). It has also been shown that both pre- and postfusion states of RSV-F exist on the surface of RSV virions, further supporting the hypothesis that the prefusion state is metastable (41). Interestingly, both the prefusion and postfusion RSV-F states have a variety of epitopes that are targets of neutralizing antibodies (42).

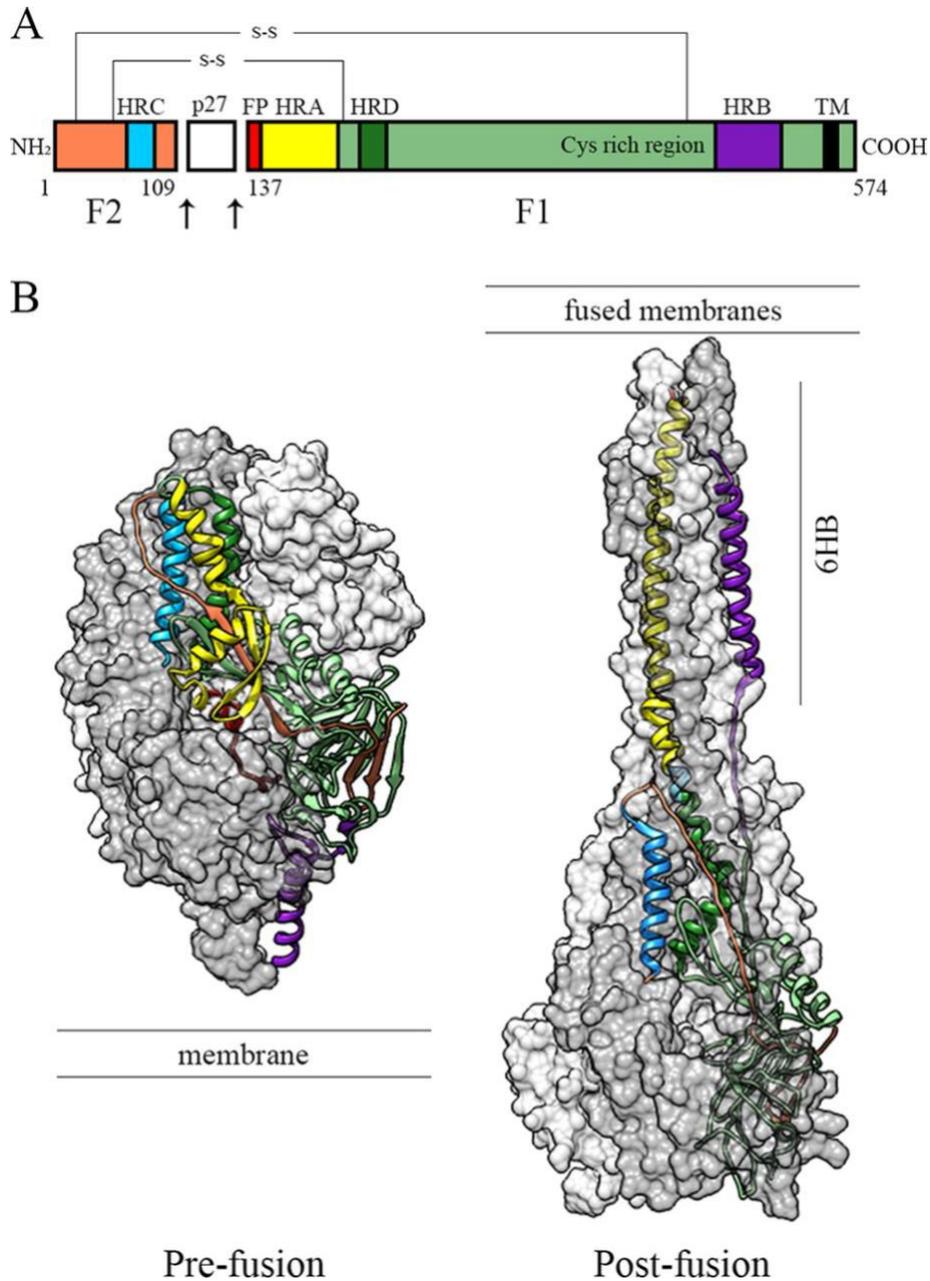


Figure 1.3. Domain organization and structure of RSV-F. (A) Domain organization of the RSV-F monomer. The heptad repeat C (HRC), 27 amino-acid fragment (p27), fusion peptide (FP), heptad repeat A (HRA), heptad repeat D (HRD), heptad repeat B (HRB) are indicated. The two cleavage sites are indicated with black arrows. Cleavage results in the release of p27 and the F1 and F2 regions, which become disulfide-linked (S-S). (B) Trimer of the pre-fusion (PBD ID:4JHW) and the post-fusion (PBD ID:3RK1) structures. Two protomers are shown in grey with surface representation, one protomer is shown as a ribbon diagram. Domains are coloured as in (A). *Figure from (43).*

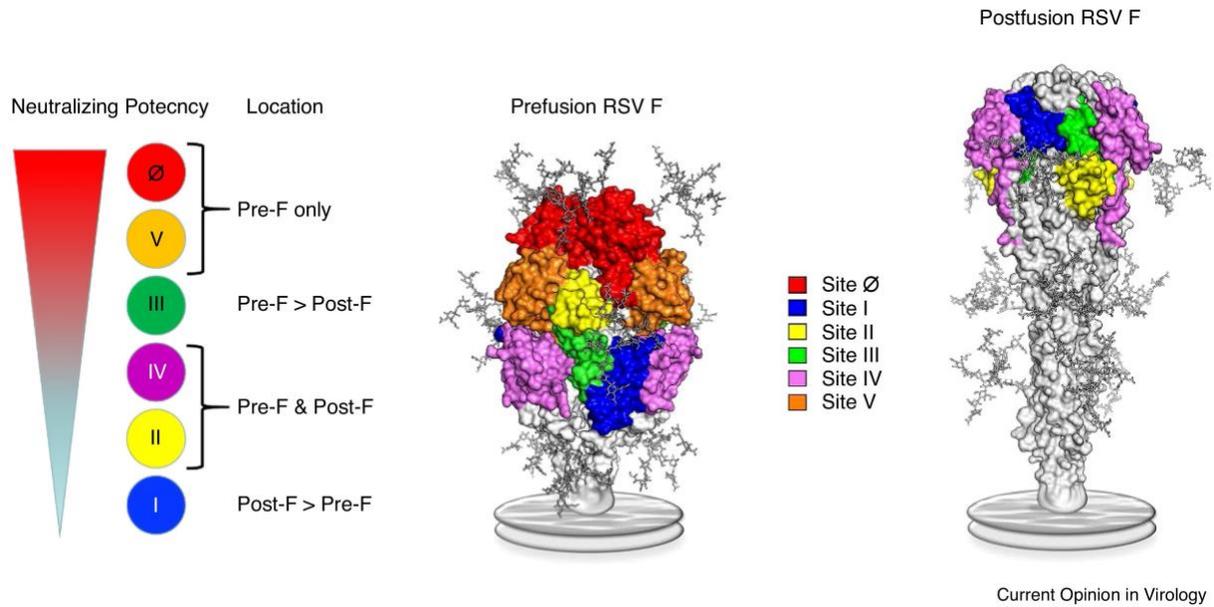


Figure 1.4. Structures and antigenic sites of pre- and postfusion RSV-F. Prefusion and postfusion RSV-F (pre-F and post-F) are shown as molecular surfaces, with the N-linked glycans (sticks), and the viral membrane is represented as a gray disc. There are two pre-F-specific antigenic sites (Ø and V), and two sites which are present on both conformations (II and IV). Antibodies against site III generally bind with greater affinity to the pre-F conformation, whereas antibodies against site I bind with greater affinity to the post-F conformation. Neutralizing potency of the antigenic sites (Ø, I, II, III, IV and V) are indicated (left). *Figure from (42).*

There are currently six main antigenic sites mapped (termed Ø, I, II, III, IV, V), and their availability can differ between the two RSV-F states (**Figure 1.4**). The prefusion state is an ideal target for the development of small molecule drugs, as preventing conversion to the postfusion state could block viral fusion and hence inhibit RSV infection (19). The prophylactic monoclonal antibody, Palivizumab, used in high risk patients, targets antigenic site II on the prefusion and postfusion states of RSV-F (44). The rearrangement from the prefusion to the postfusion form of RSV-F can occur at high temperature and low osmolarity, but the mechanism of rearrangement is currently unknown (45). While it is unclear what physiological event triggers this rearrangement, it is hypothesized that it is triggered upon interaction with one or more host cell receptors (45). However, although the trigger for fusion is still uncertain, the mechanism of membrane fusion has been elucidated (19). Briefly, in the prefusion state, the hydrophobic fusion peptide (FP) is present at the N-terminus of F₁, followed by heptad repeat A (HRA) and heptad repeat B (HRB) (**Figure 1.3**) (19). When triggered, the pre-HRA domains extend and form a trimer which extends the FP into the host cell membrane. The HAR trimer then hairpins and forms a stable 6-helix bundle, which allows the virion and cell membranes to join, resulting in virion fusion (19). Even with a complete understanding of how fusion occurs, the receptor(s) of RSV-F and the trigger of viral fusion is still a point of on-going research.

1.4 IDENTIFICATION OF RSV ENTRY RECEPTORS

1.4.1 Nucleolin

Nucleolin has been established as a receptor for RSV-F, and has been shown to be necessary for optimal RSV entry (26). The identification of nucleolin as a receptor for RSV entry was identified using a virus overlay protein binding assay (VOPBA). This assay involves separating

proteins by electrophoresis on a membrane, incubating the proteins with whole viruses, and then determining if the virus interacted with a specific protein via a virus-specific antibody (26). Samples from susceptible human, dog, and hamster cell lines, combined with mass spectrometry analyses revealed nucleolin as the consistent hit between samples (26). Subsequently, it was shown that RSV-F can co-precipitate nucleolin during infection, and these proteins were also shown to colocalize at the cell surface *in vitro*. Additionally, nucleolin knockdown or pre-incubation of virus with soluble nucleolin was shown to inhibit RSV infection *in vitro*, and expression of nucleolin in a nonpermissive *Spodoptera Frugiperda* Sf9 insect cell line was sufficient to make these cells permissive for RSV infection (26). Finally, in a mouse model, silencing of nucleolin in the lung was shown to decrease RSV infection *in vivo* (26). Thus, taken together, these results implicate nucleolin as an important cellular receptor for RSV-F during viral entry.

Nucleolin is a phosphoprotein that can be found primarily in the nucleolus but also in the nucleus, cytoplasm, and cell membrane. It is implicated in rRNA synthesis and ribosome biogenesis, as well as in a variety of activities related to regulation of chromatin structure and nucleogenesis (46, 47). These functions occur through its three different domains (**Figure 1.5**). The N-terminal domain contains acidic stretches and some of its roles include rRNA processing, interacting with histone H1, and regulating chromatin condensation (48). The central domain contains four RNA binding domains (RBDs) and is responsible for pre-mRNA splicing, mRNA stability, and poly-A tail synthesis and maturation (48). Finally, the C-terminal domain contains a glycine/arginine-rich (GAR) domain that has DNA helicase activity, in addition to being responsible for nuclear import of ribosomal proteins (48, 49). More recently, RBDs 1 and 2 have been implicated in interactions with the RSV-F protein (50).

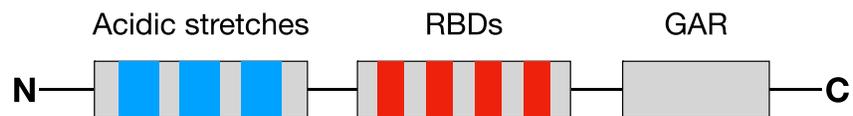


Figure 1.5. Domain organization of nucleolin. Domain organization of nucleolin with three acidic stretches, followed by four RNA binding domains (RBDs), and a glycine/arginine-rich (GAR) domain. RBDs 1 and 2 are implicated in interactions with RSV-F.

Interestingly, beyond RSV, nucleolin has also been implicated in several other viral infections. Specifically, nucleolin has been implicated in replication of hepatitis delta virus and poliovirus and presents at the cell surface to facilitate coxsackie B virus and HIV entry (51-54). Interestingly, nucleolin does not have a transmembrane domain or glycosylphosphatidylinositol (GPI) anchor, but has been shown to be present at the cell surface (25). This phenomenon is possible because nucleolin is part of a protein complex that contains membrane-bound proteins (25). Curiously, despite the clear importance of nucleolin in RSV-F-mediated cell entry, nucleolin generally has low cell surface expression in relevant cell types (27). Thus, it was hypothesized that perhaps there is an upstream cellular entry receptor or event in the RSV entry process that is important for “calling” nucleolin to the cell surface.

1.4.2 IGF1R

In addition to nucleolin, a recent study has identified IGF1R as an important RSV-F entry factor (27). In this study, researchers determined that several inhibitors of cellular signaling pathways could reduce RSV infection (27). Specifically, an IGF1R inhibitor reduced infection, and subsequently, it was shown that anti-IGF1R antibodies or knockout of IGF1R blocks RSV infection (27). Binding studies also revealed that although RSV-G was able to interact with IGF1R non-specifically, RSV-F forms a specific complex with IGF1R, with a dissociation constant of 13.6 nM (27). These results established IGF1R as a receptor for RSV-F.

IGF1R is implicated in cell growth, differentiation, and proliferation (55). It is a receptor tyrosine kinase, which is expressed as a ~180 kDa precursor, and is proteolytically processed into an α and β subunit to produce the mature protein (**Figure 1.6**) (55, 56).

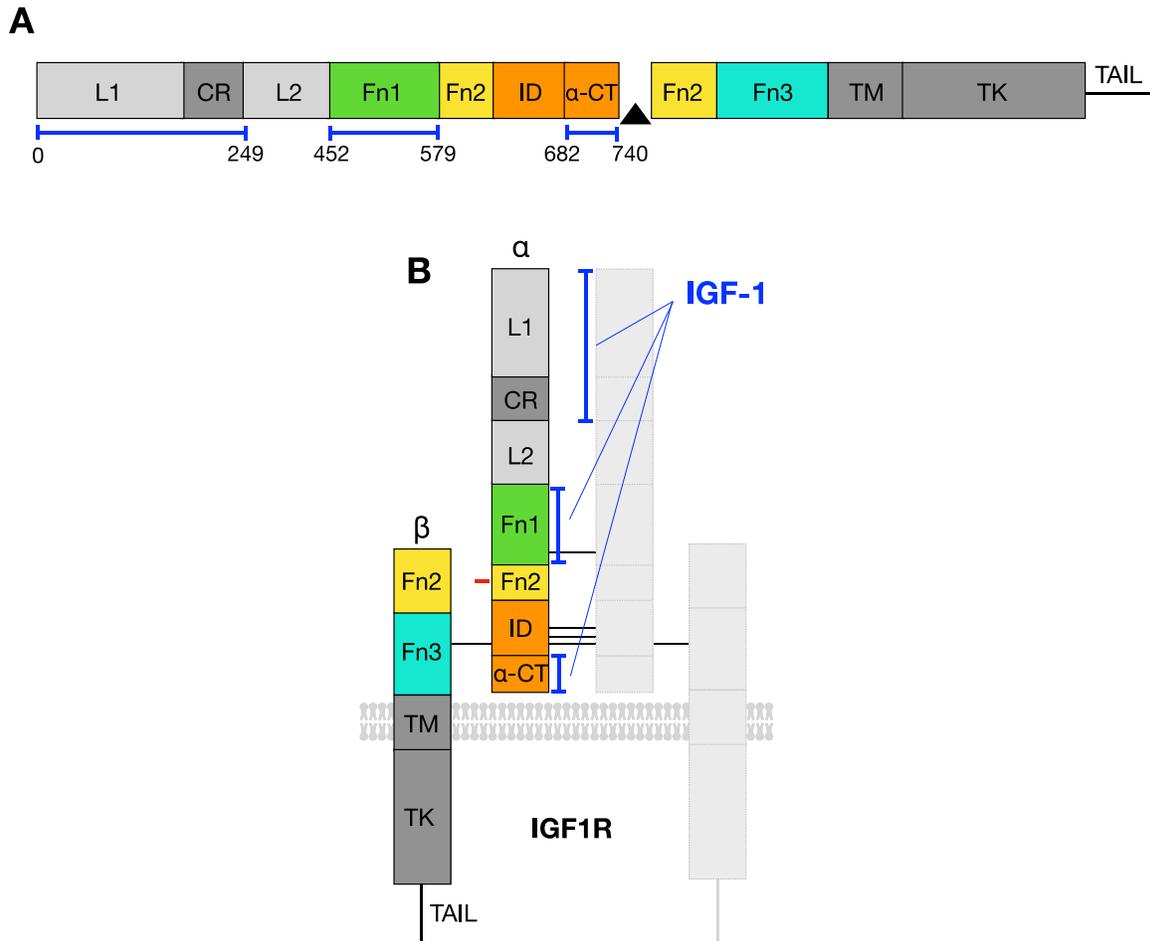


Figure 1.6. Domain organization and structure of IGF1R. (A) Domain organization of the IGF1R protein. The L1 and L2 (leucine-rich globular domains); CR (cysteine-rich domain); Fn1, Fn2, Fn3 (fibronectin III repeats 1-3); ID (insert domain, within FnIII-2); α -CT (c-terminal region of the α -chain); TM (transmembrane); TK (Tyrosine kinase domain); and TAIL (c-terminal tail) domains are indicated. Domain boundaries as well as the IGF1 ligand binding sites (blue) are indicated (55). The arrowhead denotes the cleavage site between the α and β subunits. (B) Mature IGF1R is a disulfide-linked homodimer of two $\alpha\beta$ subunits. Disulphide linkages are denoted by horizontal black lines between the two $\alpha\beta$ subunits, and domains and interaction sites are coloured as in (A).

Multiple N-glycosylation sites in IGF1R are vital to ensure proper folding and oligomerization, along with its trafficking to the cell surface (56). The mature IGF1R protein is a disulfide-linked homodimer composed of two α and β subunits (**Figure 1.6**) (57, 58). The disulfide bonds between the monomers occur at Cys514, Cys669, Cys670, and Cys672 (59). One exists between the α and β subunits at Cys633 of the α unit and Cys849 of the β unit (59). Possible disruption of these residues, and therefore disruption of the disulfide bonds, could affect the regulation of protein folding, stability, and IGF1R activity (59, 60).

IGF1R is a transmembrane protein, and similar to other transmembrane proteins, has a modular architecture composed of different, and sometimes repeated, structural units (61). One IGF1R $\alpha\beta$ monomer contains the following domains (starting from the most extracellular portion): leucine-rich domain (L1), a cysteine-rich domain (CR), another leucine-rich domain (L2), three type III fibronectin domains (FnIII-1, 2, 3), a transmembrane domain, a tyrosine kinase domain (TK), and a C-terminal tail (**Figure 1.6**) (62, 63). The L1 and L2 domains provide structural framework to support protein-protein interactions (64). The cysteine-rich domain is mainly responsible for the differentiating affinities between the IGF1R ligands, insulin-like growth factor 1 (IGF1), insulin-like growth factor 2 (IGF2), and insulin (65). The FnIII domains are also implicated in protein-protein interactions, but they also act as spacers in order to correctly position important regions of the ectodomain (56). The C-terminus of the α chain (α CT) lies within insert domain (ID) of FnIII-2, and assembles an α helix on the central β sheet of L1 when the protein folds at the membrane (**Figure 1.6**) (62). The ID is a 110 amino acid nonglobular insertion, which contains the cleavage site and three cysteines that participate in inter-subunit disulfide bonds (63). The α CT motif is important in ligand binding, which will be explored in the following section.

1.4.3 IGF1R signaling

As discussed above, IGF1R has roles in cell growth, differentiation, and proliferation and its broad expression means that it plays important roles in several tissues (55). For example, IGF1R stimulation promotes neuronal survival and myelination, mammary development and lactation, as well as bone and muscle formation and differentiation (66, 67). Additionally, IGF1R activation is implicated in multiple signaling cascades during stress and nutrient loss to shift energy use from growth and reproduction to preservation and conservation (67). Dysregulation of IGF1R signaling can lead to inhibition of apoptosis and changes in metabolism; and IGF1R overexpression is associated with invasion and metastasis of cancer cells (66). As such, dysregulation of IGF1R can lead to cancer and growth retardation (55).

The two main ligands of IGF1R are IGF1 and IGF2 (57). IGF1R is also able to interact with insulin, but with a 50- to 100-fold lower affinity, as IGF1R and the insulin receptor (IR) have 57% sequence identity (55, 57). The cryo-EM structure of the full-length IGF1R–IGF1 complex in its active state suggests that the IGF1 binding site involves the L1 and CR domains of one IGF1R monomer and the α -CT and FnIII-1 domains of the other (**Figures 1.6 and 1.7**) (55). In the working model for IGF1-induced activation, IGF1 binds L1 of one monomer and the α -CT region of another IGF1R monomer (55). Thus, for IGF1 binding, the α -CT region of one IGF1R monomer needs to be available. Upon IGF1 binding, the unoccupied α -CT becomes unavailable due to the conformational rearrangement of the receptor (55). This conformational change brings the IGF1R transmembrane domains together, triggering auto-phosphorylation of the cytoplasmic domains of the two β -subunits (63).

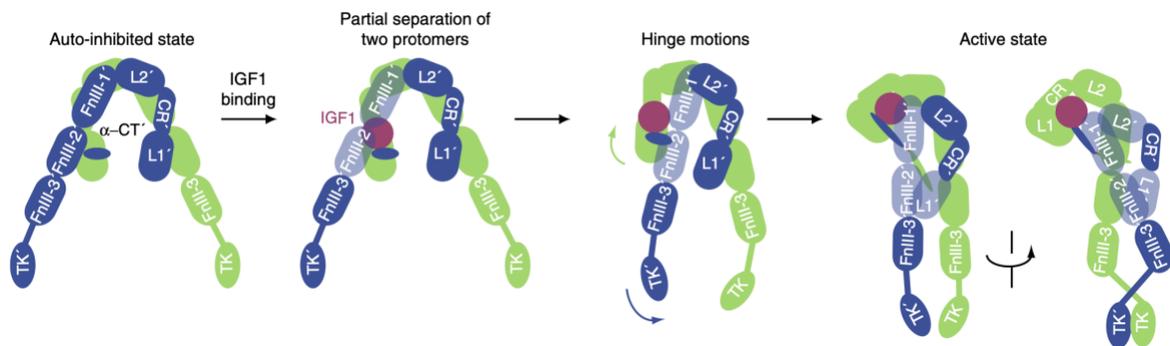


Figure 1.7. IGF1R rearrangement upon IGF1 binding. Working model for IGF1-induced IGF1R activation. IGF1R monomers are colored in green and blue, and IGF1 is colored in pink. The domains L1 and L2 (leucine-rich repeat domains), CR (cysteine-rich domain), FnIII-1, FnIII-2, and FnIII-3 (fibronectin type III domains), α-CT (the C-terminus of the α-subunit); TK (tyrosine kinase domain) are indicated. *Figure from (55).*

In turn, this leads to downstream activation of two main signaling pathways: the mitogen-activated protein kinase (MAPK) cascade and the phosphoinositide-3 kinase (PI3K) cascade, which in turn triggers the expression of a variety of cell growth factors and other aforementioned cellular functions (57). IGF1R signaling and activation has been well characterized in terms of normal physiological functions, but the downstream pathway involved in RSV and IGF1R-induced nucleolin trafficking to the cell surface had yet to be investigated (27). To determine if kinases were important for the process of entry, a screen of small molecules that inhibit the activation of specific kinases was performed (27). Inhibition of protein kinase C zeta (PKC ζ) was shown to result in reduced RSV infection (27). It was subsequently shown that IGF1R activation resulted in activation of PKC ζ , and that this activation calls nucleolin to the cell surface (27). This demonstrated the importance of PKC ζ signaling in RSV entry. Taken together, these results implicate IGF1R, PKC ζ , and nucleolin in the RSV entry pathway.

1.4.4 Current model for RSV attachment and entry

The identification of IGF1R, PKC ζ signaling, and nucleolin in RSV entry has led to the current working model for RSV entry (**Figure 1.8**). Specifically, RSV-G is thought to initially attach to cells through non-specific contacts with GAGs, and possibly also CX3CR1 (19, 22, 33, 35). Next, the RSV-F protein is thought to bind specifically to IGF1R. This interaction triggers IGF1R activation, culminating in the activation of PKC ζ signaling and the subsequent recruitment of nucleolin from the nucleus to the cell surface. At the cell surface, RSV-F, IGF1R, and nucleolin interact to form a trimeric complex that facilitates RSV fusion and entry (**Figure 1.8**).

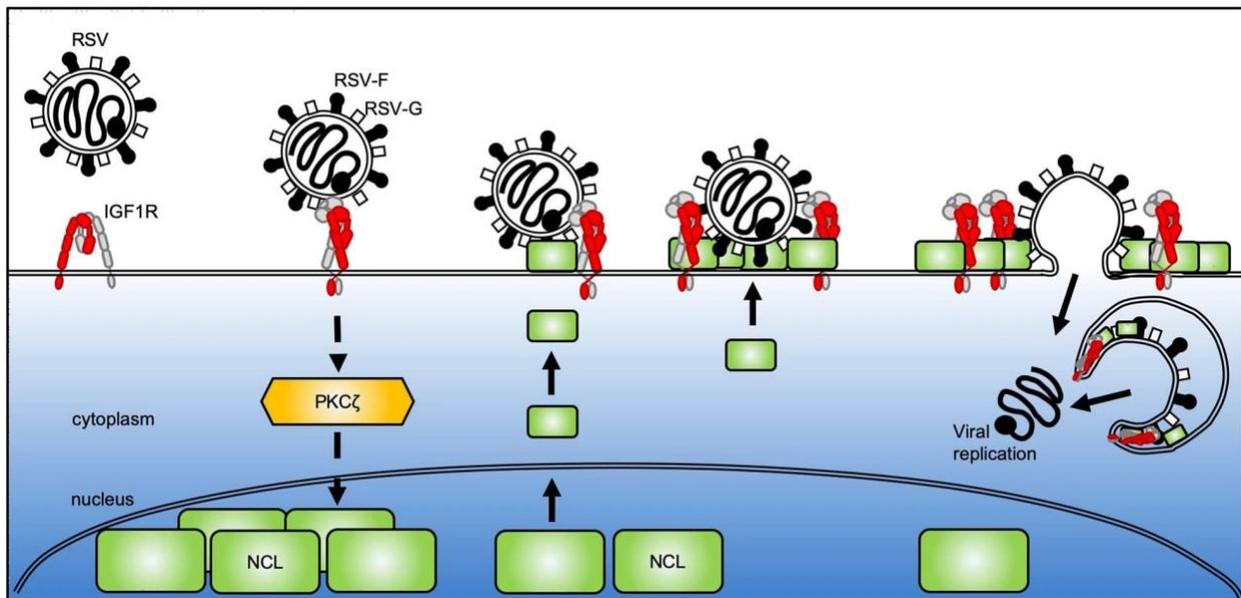


Figure 1.8. RSV entry model. RSV-F on the surface of the RSV virion binds to the insulin-like growth factor 1 receptor (IGF1R) on the cell surface. Binding of RSV-F activates IGF1R, which activates the PKC ζ signalling pathway. Activation of PKC ζ recruits nucleolin from the nucleus to the cell membrane. At the cell membrane, RSV-F, IGF1R, and nucleolin interact to facilitate fusion of the viral envelope and the host membrane. *Figure from (27).*

1.5 TOWARDS THE DEVELOPMENT OF BIOLOGICAL INHIBITORS OF RSV ENTRY

1.5.1 Viral Entry as a target of antiviral therapy

Entry, and more specifically fusion, has always been a lucrative target for antiviral development. Due to its pivotal role in fusion, RSV-F is an attractive candidate for the development of antivirals. However, targeting the interaction between RSV-F and specific host factors could also have therapeutic potential. Of particular interest is the interaction between IGF1R and RSV-F. As we described herein, prefusion RSV-F binds the IGF1R, which subsequently activates PKC ζ signalling, triggering nucleolin translocation to the cell surface culminating in RSV fusion and completing the viral entry process (27). Considering the importance of IGF1R:RSV-F interactions for initiating a viral infection, interfering with this interaction is likely to be a lucrative avenue for the development of novel RSV antivirals.

1.5.2 Directed evolution for the development of biological inhibitors of viral entry

As alluded to already, virus-host interactions are an extremely important targets in the development of antiviral treatments. Herein, we hypothesized that we could exploit host-virus interactions to develop biological inhibitors of viral entry. To this end, we have designed a directed evolution platform to develop and screen biological inhibitors of viral entry. Briefly, we plan to use directed evolution to evolve host receptor molecules to have greater affinity for the viral receptor, while simultaneously screening for those with reduced affinity for the natural ligand (**Figure 1.9**). In this way, we may select for biological inhibitors of viral entry that do not interfere with canonical host receptor-ligand interactions.

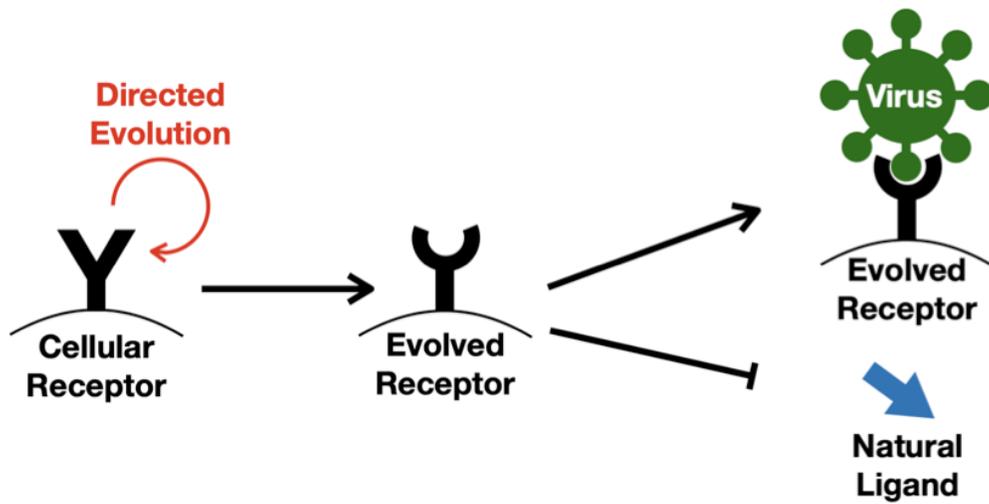


Figure 1.9. Directed evolution for development of biological inhibitors of viral entry. Cellular receptors are subjected to directed evolution. Evolved receptors can subsequently be screened for enhanced viral entry and reduced binding to the natural ligand.

This is similar to the use of receptor decoy molecules, which have been widely applied in HIV-1 studies using soluble CD4 molecules, but our strategy provides an additional advantage by simultaneously selecting against host-ligand interactions to limit off-target effects (**Figure 1.9**) (68). The directed evolution strategy employed herein is inspired by natural selection, which can implement a variety of genetic diversification techniques to introduce mutations into a specific gene (69, 70). This strategy enables the creation of a diverse library of mutant receptors, which in turn permits screening and identification of mutations that produce a desired phenotype (69, 70). As a proof of principle, we plan to exploit the interaction between RSV-F and the IGF1R protein to develop biological inhibitors of RSV entry.

1.6 RATIONALE AND HYPOTHESIS

With the recent identification of IGF1R as a receptor for RSV-F there is now a new target for potential antiviral development. However, it is still unclear which domains of IGF1R specifically interact with RSV-F. Determining the specific interaction domains of IGF1R will be essential to informing the design of potential decoy-based inhibitors of RSV entry. Thus, we hypothesize that through molecular docking and alanine scanning mutagenesis, we might gain insights into the residues in IGF1R important for RSV-F-mediated viral entry. Furthermore, we hypothesize that this will allow us to identify the regions of IGF1R suitable for directed evolution as a first step in the development process for biological inhibitors of RSV entry.

In *Chapter 2*, we performed molecular docking analysis and identified 35 residues of IGF1R implicated in interactions with RSV-F. We subsequently performed alanine-scanning mutagenesis on these residues and investigated their expression and processing, as well as their ability to facilitate RSV infection. We found that some mutations disrupted expression or

maturation of IGF1R, and that one mutation severely impaired RSV entry. Additionally, we used sequence-saturation mutagenesis to generate a random mutant IGF1R library targeted to the region of IGF1R implicated in RSV-F binding.

Finally, in *Chapter 3* we further discuss IGF1R and RSV-F interactions and propose future experiments to support or clarify our findings. We also discuss our progress towards a directed evolution platform and its potential further application.

1.7 REFERENCES

1. Walsh EE, Hall CB. 2015. 160 - Respiratory Syncytial Virus (RSV), p 1948-1960.e3. *In* Bennett JE, Dolin R, Blaser MJ (ed), Mandell, Douglas, and Bennett's Principles and Practice of Infectious Diseases (Eighth Edition) doi:<https://doi.org/10.1016/B978-1-4557-4801-3.00160-0>. W.B. Saunders, Philadelphia.
2. Shi T, McAllister DA, O'Brien KL, Simoes EAF, Madhi SA, Gessner BD, Polack FP, Balsells E, Acacio S, Aguayo C, Alassani I, Ali A, Antonio M, Awasthi S, Awori JO, Azziz-Baumgartner E, Baggett HC, Baillie VL, Balmaseda A, Barahona A, Basnet S, Bassat Q, Basualdo W, Bigogo G, Bont L, Breiman RF, Brooks WA, Broor S, Bruce N, Bruden D, Buchy P, Campbell S, Carosone-Link P, Chadha M, Chipeta J, Chou M, Clara W, Cohen C, de Cuellar E, Dang D-A, Dash-yandag B, Deloria-Knoll M, Dherani M, Eap T, Ebruke BE, Echavarria M, de Freitas Lázaro Emediato CC, Fasce RA, Feikin DR, Feng L, et al. 2017. Global, regional, and national disease burden estimates of acute lower respiratory infections due to respiratory syncytial virus in young children in 2015: a systematic review and modelling study. *The Lancet* 390:946-958.
3. Nair H, Simões EA, Rudan I, Gessner BD, Azziz-Baumgartner E, Zhang JSF, Feikin DR, Mackenzie GA, Moïsi JC, Roca A, Baggett HC, Zaman SM, Singleton RJ, Lucero MG, Chandran A, Gentile A, Cohen C, Krishnan A, Bhutta ZA, Arguedas A, Clara AW, Andrade AL, Ope M, Ruvinsky RO, Hortal M, McCracken JP, Madhi SA, Bruce N, Qazi SA, Morris SS, El Arifeen S, Weber MW, Scott JAG, Brooks WA, Breiman RF, Campbell H, Severe Acute Lower Respiratory Infections Working G. 2013. Global and regional burden of hospital admissions for severe acute lower respiratory infections in young children in 2010: a systematic analysis. *Lancet (London, England)* 381:1380-1390.

4. Piedimonte G, Perez MK. 2014. Respiratory syncytial virus infection and bronchiolitis. *Pediatrics in review* 35:519-530.
5. Tin Tin Htar M, Yerramalla MS, Moisi JC, Swerdlow DL. 2020. The burden of respiratory syncytial virus in adults: a systematic review and meta-analysis. *Epidemiol Infect* 148:e48.
6. Staadegaard L, Caini S, Wangchuk S, Thapa B, de Almeida WAF, de Carvalho FC, Fasce RA, Bustos P, Kyncl J, Novakova L, Caicedo AB, de Mora Coloma DJ, Meijer A, Hooiveld M, Huang QS, Wood T, Guiomar R, Rodrigues AP, Lee VJM, Ang LW, Cohen C, Moyes J, Larrauri A, Delgado-Sanz C, Demont C, Bangert M, Dückers M, van Summeren J, Paget J. 2021. Defining the seasonality of respiratory syncytial virus around the world: National and subnational surveillance data from 12 countries. *Influenza Other Respir Viruses* 15:732-741.
7. Griffiths C, Drews SJ, Marchant DJ. 2017. Respiratory Syncytial Virus: Infection, Detection, and New Options for Prevention and Treatment. *Clinical microbiology reviews* 30:277-319.
8. Leung NHL. 2021. Transmissibility and transmission of respiratory viruses. *Nature Reviews Microbiology* 19:528-545.
9. Kulkarni H, Smith CM, Lee Ddo H, Hirst RA, Easton AJ, O'Callaghan C. 2016. Evidence of Respiratory Syncytial Virus Spread by Aerosol. Time to Revisit Infection Control Strategies? *Am J Respir Crit Care Med* 194:308-16.
10. Ginsberg GM, Somekh E, Schlesinger Y. 2018. Should we use Palivizumab immunoprophylaxis for infants against respiratory syncytial virus? - a cost-utility analysis. *Israel journal of health policy research* 7:63-63.

11. Sigurs N, Bjarnason R, Sigurbergsson F, Kjellman B. 2000. Respiratory Syncytial Virus Bronchiolitis in Infancy Is an Important Risk Factor for Asthma and Allergy at Age 7. *American Journal of Respiratory and Critical Care Medicine* 161:1501-1507.
12. Resch B. 2017. Product review on the monoclonal antibody palivizumab for prevention of respiratory syncytial virus infection. *Human vaccines & immunotherapeutics* 13:2138-2149.
13. Geskey JM, Thomas NJ, Brummel GL. 2007. Palivizumab: a review of its use in the protection of high risk infants against respiratory syncytial virus (RSV). *Biologics : targets & therapy* 1:33-43.
14. Rima B, Collins P, Easton A, Fouchier R, Kurath G, Lamb RA, Lee B, Maisner A, Rota P, Wang L, Consortium IR. 2017. ICTV Virus Taxonomy Profile: Pneumoviridae. *Journal of General Virology* 98:2912-2913.
15. Ke Z, Dillard RS, Chirkova T, Leon F, Stobart CC, Hampton CM, Strauss JD, Rajan D, Rostad CA, Taylor JV, Yi H, Shah R, Jin M, Hartert TV, Peebles RS, Jr., Graham BS, Moore ML, Anderson LJ, Wright ER. 2018. The Morphology and Assembly of Respiratory Syncytial Virus Revealed by Cryo-Electron Tomography. *Viruses* 10.
16. Collins PL, Fearn R, Graham BS. 2013. Respiratory syncytial virus: virology, reverse genetics, and pathogenesis of disease. *Curr Top Microbiol Immunol* 372:3-38.
17. Battles MB, McLellan JS. 2019. Respiratory syncytial virus entry and how to block it. *Nature reviews Microbiology* 17:233-245.
18. Gao Y, Cao D, Pawnikar S, John KP, Ahn HM, Hill S, Ha JM, Parikh P, Ogilvie C, Swain A, Yang A, Bell A, Salazar A, Miao Y, Liang B. 2020. Structure of the Human

- Respiratory Syncytial Virus M2-1 Protein in Complex with a Short Positive-Sense Gene-End RNA. *Structure* 28:979-990.e4.
19. McLellan JS, Ray WC, Peeples ME. 2013. Structure and function of respiratory syncytial virus surface glycoproteins. *Current topics in microbiology and immunology* 372:83-104.
 20. Zhang L, Peeples ME, Boucher RC, Collins PL, Pickles RJ. 2002. Respiratory Syncytial Virus Infection of Human Airway Epithelial Cells Is Polarized, Specific to Ciliated Cells, and without Obvious Cytopathology. *Journal of Virology* 76:5654-5666.
 21. Feldman SA, Audet S, Beeler JA. 2000. The fusion glycoprotein of human respiratory syncytial virus facilitates virus attachment and infectivity via an interaction with cellular heparan sulfate. *Journal of virology* 74:6442-6447.
 22. Chirkova T, Lin S, Oomens AGP, Gaston KA, Boyoglu-Barnum S, Meng J, Stobart CC, Cotton CU, Hartert TV, Moore ML, Ziady AG, Anderson LJ. 2015. CX3CR1 is an important surface molecule for respiratory syncytial virus infection in human airway epithelial cells. *The Journal of general virology* 96:2543-2556.
 23. Behera AK, Matsuse H, Kumar M, Kong X, Lockey RF, Mohapatra SS. 2001. Blocking Intercellular Adhesion Molecule-1 on Human Epithelial Cells Decreases Respiratory Syncytial Virus Infection. *Biochemical and Biophysical Research Communications* 280:188-195.
 24. Marr N, Turvey SE. 2012. Role of human TLR4 in respiratory syncytial virus-induced NF- κ B activation, viral entry and replication. *Innate Immunity* 18:856-865.
 25. Mastrangelo P, Hegele RG. 2012. The RSV fusion receptor: not what everyone expected it to be. *Microbes and Infection* 14:1205-1210.

26. Tayyari F, Marchant D, Moraes TJ, Duan W, Mastrangelo P, Hegele RG. 2011. Identification of nucleolin as a cellular receptor for human respiratory syncytial virus. *Nature Medicine* 17:1132-1135.
27. Griffiths CD, Bilawchuk LM, McDonough JE, Jamieson KC, Elawar F, Cen Y, Duan W, Lin C, Song H, Casanova J-L, Ogg S, Jensen LD, Thienpont B, Kumar A, Hobman TC, Proud D, Moraes TJ, Marchant DJ. 2020. IGF1R is an entry receptor for respiratory syncytial virus. *Nature* 583:615-619.
28. Noton SL, Fearn R. 2015. Initiation and regulation of paramyxovirus transcription and replication. *Virology* 479-480:545-54.
29. Schildgen V, Hoogen Bvd, Fouchier R, Tripp RA, Alvarez R, Manoha C, Williams J, Schildgen O. 2011. Human Metapneumovirus: Lessons Learned over the First Decade. *Clinical Microbiology Reviews* 24:734-754.
30. Rincheval V, Lelek M, Gault E, Bouillier C, Sitterlin D, Blouquit-Laye S, Galloux M, Zimmer C, Eleouet J-F, Rameix-Welti M-A. 2017. Functional organization of cytoplasmic inclusion bodies in cells infected by respiratory syncytial virus. *Nature Communications* 8:563.
31. Vanover D, Smith DV, Blanchard EL, Alonas E, Kirschman JL, Lifland AW, Zurla C, Santangelo PJ. 2017. RSV glycoprotein and genomic RNA dynamics reveal filament assembly prior to the plasma membrane. *Nature Communications* 8:667.
32. Wertz GW, Collins PL, Huang Y, Gruber C, Levine S, Ball LA. 1985. Nucleotide sequence of the G protein gene of human respiratory syncytial virus reveals an unusual type of viral membrane protein. *Proc Natl Acad Sci U S A* 82:4075-9.

33. Tripp RA, Jones LP, Haynes LM, Zheng H, Murphy PM, Anderson LJ. 2001. CX3C chemokine mimicry by respiratory syncytial virus G glycoprotein. *Nature Immunology* 2:732-738.
34. Techaarpornkul S, Barretto N, Peeples ME. 2001. Functional analysis of recombinant respiratory syncytial virus deletion mutants lacking the small hydrophobic and/or attachment glycoprotein gene. *J Virol* 75:6825-34.
35. Johnson SM, McNally BA, Ioannidis I, Flano E, Teng MN, Oomens AG, Walsh EE, Peeples ME. 2015. Respiratory Syncytial Virus Uses CX3CR1 as a Receptor on Primary Human Airway Epithelial Cultures. *PLoS pathogens* 11:e1005318.
36. McLellan JS, Yang Y, Graham BS, Kwong PD. 2011. Structure of Respiratory Syncytial Virus Fusion Glycoprotein in the Postfusion Conformation Reveals Preservation of Neutralizing Epitopes. *Journal of Virology* 85:7788-7796.
37. Collins PL, Huang YT, Wertz GW. 1984. Nucleotide sequence of the gene encoding the fusion (F) glycoprotein of human respiratory syncytial virus. *Proceedings of the National Academy of Sciences of the United States of America* 81:7683-7687.
38. Zimmer G, Budz L, Herrler G. 2001. Proteolytic activation of respiratory syncytial virus fusion protein. Cleavage at two furin consensus sequences. *J Biol Chem* 276:31642-50.
39. Begoña Ruiz-Argüello M, González-Reyes L, Calder LJ, Palomo C, Martín D, Saíz MJ, García-Barreno B, Skehel JJ, Melero JA. 2002. Effect of proteolytic processing at two distinct sites on shape and aggregation of an anchorless fusion protein of human respiratory syncytial virus and fate of the intervening segment. *Virology* 298:317-26.
40. McLellan JS, Chen M, Joyce MG, Sastry M, Stewart-Jones GBE, Yang Y, Zhang B, Chen L, Srivatsan S, Zheng A, Zhou T, Graepel KW, Kumar A, Moin S, Boyington JC,

- Chuang G-Y, Soto C, Baxa U, Bakker AQ, Spits H, Beaumont T, Zheng Z, Xia N, Ko S-Y, Todd J-P, Rao S, Graham BS, Kwong PD. 2013. Structure-based design of a fusion glycoprotein vaccine for respiratory syncytial virus. *Science (New York, NY)* 342:592-598.
41. Liljeroos L, Krzyzaniak MA, Helenius A, Butcher SJ. 2013. Architecture of respiratory syncytial virus revealed by electron cryotomography. *Proc Natl Acad Sci U S A* 110:11133-8.
 42. Graham BS. 2017. Vaccine development for respiratory syncytial virus. *Curr Opin Virol* 23:107-112.
 43. Bermingham IM, Chappell KJ, Watterson D, Young PR. 2018. The Heptad Repeat C Domain of the Respiratory Syncytial Virus Fusion Protein Plays a Key Role in Membrane Fusion. *Journal of Virology* 92:e01323-17.
 44. Patel N, Massare MJ, Tian J-H, Guebre-Xabier M, Lu H, Zhou H, Maynard E, Scott D, Ellingsworth L, Glenn G, Smith G. 2019. Respiratory syncytial virus prefusogenic fusion (F) protein nanoparticle vaccine: Structure, antigenic profile, immunogenicity, and protection. *Vaccine* 37:6112-6124.
 45. McLellan JS. 2015. Neutralizing epitopes on the respiratory syncytial virus fusion glycoprotein. *Current opinion in virology* 11:70-75.
 46. Abdelmohsen K, Gorospe M. 2012. RNA-binding protein nucleolin in disease. *RNA Biol* 9:799-808.
 47. Tajrishi MM, Tuteja R, Tuteja N. 2011. Nucleolin: The most abundant multifunctional phosphoprotein of nucleolus. *Commun Integr Biol* 4:267-75.

48. Jia W, Yao Z, Zhao J, Guan Q, Gao L. 2017. New perspectives of physiological and pathological functions of nucleolin (NCL). *Life Sciences* 186:1-10.
49. Tuteja N, Huang NW, Skopac D, Tuteja R, Hrvatic S, Zhang J, Pongor S, Joseph G, Faucher C, Amalric F, et al. 1995. Human DNA helicase IV is nucleolin, an RNA helicase modulated by phosphorylation. *Gene* 160:143-8.
50. Mastrangelo P, Chin AA, Tan S, Jeon AH, Ackerley CA, Siu KK, Lee JE, Hegele RG. 2021. Identification of RSV Fusion Protein Interaction Domains on the Virus Receptor, Nucleolin. *Viruses* 13.
51. Izumi RE, Valdez B, Banerjee R, Srivastava M, Dasgupta A. 2001. Nucleolin stimulates viral internal ribosome entry site-mediated translation. *Virus Research* 76:17-29.
52. Lee CH, Chang SC, Chen CJ, Chang MF. 1998. The nucleolin binding activity of hepatitis delta antigen is associated with nucleolus targeting. *J Biol Chem* 273:7650-6.
53. de Verdugo UR, Selinka HC, Huber M, Kramer B, Kellermann J, Hofschneider PH, Kandolf R. 1995. Characterization of a 100-kilodalton binding protein for the six serotypes of coxsackie B viruses. *J Virol* 69:6751-7.
54. Nisole S, Said EA, Mische C, Prevost M-C, Krust B, Bouvet P, Bianco A, Briand J-P, Hovanessian AG. 2002. The anti-HIV pentameric pseudopeptide HB-19 binds the C-terminal end of nucleolin and prevents anchorage of virus particles in the plasma membrane of target cells. *Journal of Biological Chemistry* 277:20877-20886.
55. Li J, Choi E, Yu H, Bai X-c. 2019. Structural basis of the activation of type 1 insulin-like growth factor receptor. *Nature Communications* 10:4567.

56. Adams TE, Epa VC, Garrett TPJ, Ward* CW. 2000. Structure and function of the type 1 insulin-like growth factor receptor. *Cellular and Molecular Life Sciences CMLS* 57:1050-1093.
57. Hakuno F, Takahashi S-I. 2018. 40 years of IGF1: IGF1 receptor signaling pathways. *Journal of molecular endocrinology* 61:T69-T86.
58. Khatib A-M, Siegfried G, Prat A, Luis J, Chrétien M, Metrakos P, Seidah NG. 2001. Inhibition of Proprotein Convertases Is Associated with Loss of Growth and Tumorigenicity of HT-29 Human Colon Carcinoma Cells: IMPORTANCE OF INSULIN-LIKE GROWTH FACTOR-1 (IGF-1) RECEPTOR PROCESSING IN IGF-1-MEDIATED FUNCTIONS *. *Journal of Biological Chemistry* 276:30686-30693.
59. Ward CW, Garrett TP, Lou M, Mckern NM, Adams TE, Elleman TC, Hoyne PA, Frenkel MJ, Cosgrove LJ, Lovrecz GO. 2013. The structure of the type 1 insulin-like growth factor receptor. *Madame Curie Bioscience Database* [Internet].
60. Ishikawa H, Kim S, Kwak K, Wakasugi K, Fayer MD. 2007. Disulfide bond influence on protein structural dynamics probed with 2D-IR vibrational echo spectroscopy. *Proceedings of the National Academy of Sciences* 104:19309-19314.
61. Ward CW, Garrett TP. 2001. The relationship between the L1 and L2 domains of the insulin and epidermal growth factor receptors and leucine-rich repeat modules. *BMC Bioinformatics* 2:4.
62. Xu Y, Kong GKW, Menting JG, Margetts MB, Delaine CA, Jenkin LM, Kiselyov VV, De Meyts P, Forbes BE, Lawrence MC. 2018. How ligand binds to the type 1 insulin-like growth factor receptor. *Nature Communications* 9:821.

63. Kavran JM, McCabe JM, Byrne PO, Connacher MK, Wang Z, Ramek A, Sarabipour S, Shan Y, Shaw DE, Hristova K, Cole PA, Leahy DJ. 2014. How IGF-1 activates its receptor. *Elife* 3.
64. Kobe B, Kajava AV. 2001. The leucine-rich repeat as a protein recognition motif. *Curr Opin Struct Biol* 11:725-32.
65. Hoyne PA, Elleman TC, Adams TE, Richards KM, Ward CW. 2000. Properties of an insulin receptor with an IGF-1 receptor loop exchange in the cysteine-rich region. *FEBS Lett* 469:57-60.
66. Yuan J, Yin Z, Tao K, Wang G, Gao J. 2018. Function of insulin-like growth factor 1 receptor in cancer resistance to chemotherapy. *Oncol Lett* 15:41-47.
67. Hartog H, Wesseling J, Boezen HM, van der Graaf WTA. 2007. The insulin-like growth factor 1 receptor in cancer: Old focus, new future. *European Journal of Cancer* 43:1895-1904.
68. Haim H, Si Z, Madani N, Wang L, Courter JR, Princiotta A, Kassa A, DeGrace M, McGee-Estrada K, Mefford M, Gabuzda D, Smith AB, 3rd, Sodroski J. 2009. Soluble CD4 and CD4-mimetic compounds inhibit HIV-1 infection by induction of a short-lived activated state. *PLoS Pathog* 5:e1000360.
69. Nannemann DP, Birmingham WR, Scism RA, Bachmann BO. 2011. Assessing directed evolution methods for the generation of biosynthetic enzymes with potential in drug biosynthesis. *Future medicinal chemistry* 3:809-819.
70. Packer MS, Liu DR. 2015. Methods for the directed evolution of proteins. *Nature Reviews Genetics* 16:379.

CHAPTER 2: UNDERSTANDING RSV ENTRY REQUIREMENTS BY MAPPING INTERACTIONS BETWEEN RSV-F AND IGF1R

Rachel S. Hayes¹, Ahmed K. Oraby², Carolina Camargo³, David J. Marchant², and Selena M. Sagan^{1,3}.

¹Department of Biochemistry, McGill University, Montreal, Canada

²Department of Medical Microbiology & Immunology, University of Alberta, Edmonton, Alberta, Canada

³Department of Microbiology & Immunology, McGill University, Montreal, Canada

Keywords

Respiratory Syncytial virus (RSV), fusion (F) protein, insulin-like growth factor 1 receptor (IGF1R), receptor, entry, directed evolution, biological inhibitor.

2.1 PREFACE TO CHAPTER 2

As discussed in *Chapter 1*, the RSV fusion protein (RSV-F) has been shown to bind and activate IGF1R during viral entry (1). However, it is unknown which domains of IGF1R are responsible for this interaction. Herein, we performed molecular docking analysis to identify the interaction interface between IGF1R and RSV-F. Subsequently, we used alanine-scanning mutagenesis to investigate the importance of the identified IGF1R residues in RSV entry. Additionally, we wanted to exploit the interaction between IGF1R and RSV-F to develop biological inhibitors of viral entry. To accomplish this, we used sequence saturation mutagenesis to generate a mutant IGF1R library, which will be used in future in a screening approach to develop IGF1R-based decoy inhibitors of RSV entry.

This chapter was adapted from the following manuscript: Understanding respiratory syncytial virus entry requirements by mapping interactions between the RSV-F and IGF1R.

Rachel S. Hayes, Ahmed K. Oraby, Carolina Camargo, David J. Marchant, and Selena M. Sagan. *Manuscript in preparation.*

2.2 ABSTRACT

Respiratory syncytial virus (RSV) has two main surface glycoproteins, the attachment glycoprotein (G) and the fusion (F) protein, which together mediate viral entry. Attachment is mediated by the RSV-G protein, while the RSV-F protein makes specific contact with the cellular insulin-like growth factor 1 receptor (IGF1R). This interaction leads to IGF1R phosphorylation and activation, triggering a signaling cascade which calls the co-receptor, nucleolin, from the nucleus to the cell surface where it can trigger viral fusion. We performed molecular docking analyses which provided a potential set of 35 residues in IGF1R that may be important for interactions with RSV-F. We used alanine scanning mutagenesis to generate these 35 mutants of IGF1R and assessed their expression, maturation, as well as the effect of each mutation on RSV entry. We identified several mutations that appear to inhibit IGF1R maturation, but surprisingly these mutations had no significant effect on RSV entry. This suggests that IGF1R maturation may not be required for RSV entry. Additionally, we identified one residue, S788A, that significantly reduced RSV infection, revealing a potentially critical residue for RSV-F-mediated entry. Finally, to develop IGF1R-based biological decoy receptors for RSV entry, we generated a mutant IGF1R library using sequence saturation mutagenesis targeting the region of IGF1R identified in our molecular docking analysis. We anticipate that this research will help to map the RSV-F:IGF1R interaction interface, and will aid in the development of targeting strategies that inhibit RSV entry.

2.3 INTRODUCTION

Human respiratory syncytial virus (RSV) is the leading cause of severe lower respiratory tract disease globally, with approximately 33 million cases annually (2). Certain high-risk groups have a greater chance of developing severe infections, including premature babies, immunocompromised infants, and those born with lung or heart disease (3, 4). Currently, there is no vaccine available and there are no treatments for active RSV infections. The only available treatment is a prophylactic monoclonal antibody, Palivizumab, which is reserved for high-risk patients such as premature and/or immunocompromised infants (5). However, Palivizumab is costly, has low neutralizing activity, and is not effective if administered during an active RSV infection (6).

RSV is an enveloped, non-segmented, negative-sense RNA virus of the *Pneumoviridae* family (7). It preferentially infects the ciliated cells of the airway epithelium (8). Infection is mediated by the two main RSV surface glycoproteins: the attachment glycoprotein (G) and the fusion (F) glycoprotein (9). Initial, non-specific virus binding is mediated by the RSV-G protein via its heparin-binding domains, which interact with cell surface glycosaminoglycans (GAGs), mainly heparin sulfate proteoglycans (10). RSV-F is responsible for viral envelope fusion with host cell membrane(s) (9, 10). Several studies also suggest that the RSV-G protein interacts with the CX3CR1 receptor on ciliated cells of the airway epithelium (9, 10). Additionally, the RSV-F protein has been shown to interact with intercellular adhesion molecule-1 (ICAM-1), toll-like receptor-4 (TLR4), nucleolin (NCL), and more recently, the insulin-like growth factor 1 receptor (IGF1R) (1, 11-14).

In the current model of RSV-F-mediated fusion, the prefusion metastable RSV-F protein must first interact with IGF1R (1). This triggers IGF1R activation and induction of the protein

kinase C zeta (PKC ζ) cellular signaling cascade (1). This signaling cascade triggers nucleolin translocation from the nucleus to the cell surface, where it is then able to also bind to RSV-F and IGF1R (1). This interaction triggers RSV-F protein-mediated fusion with host cell membranes, which can occur at the plasma membrane or within an endocytic vesicle (1, 14, 15). The location of fusion might be dependent on the environmental conditions as well as the target cell (15).

IGF1R is a receptor tyrosine kinase that is first expressed as a ~180 kDa precursor and is subsequently proteolytically processed (into two subunits) in the Golgi by furin during egress to the cell surface (16-18). The mature receptor is a disulfide-linked homodimer, composed of two α and β subunits (19, 20). The two main ligands of IGF1R are insulin-like growth factor 1 (IGF1) and insulin-like growth factor 2 (IGF2) (19). Ligand binding leads to a signaling cascade that triggers cell growth, differentiation, and proliferation (16). IGF1 interactions with IGF1R include the first leucine-rich domain (L1), the cysteine-rich domain (CR), the type III fibronectin domain 1 (FnIII-1), as well as the C-terminal of the α chain (α CT) of IGF1R (**Figure 2.1A-B**) (16). However, it is still unclear how RSV-F interacts with IGF1R.

Herein, we wanted to map RSV-F:IGF1R interactions and exploit these interactions to develop biological inhibitors of viral entry. To this end, we performed molecular docking analysis and then used alanine-scanning mutagenesis to investigate the residues of IGF1R required for RSV entry. We found that residues close to the IGF1R cleavage site (e.g. residues 744-746) impaired IGF1R maturation, but surprisingly had no significant effect on RSV entry. Additionally, we identified one residue, S788A, that severely impaired RSV entry. Finally, we used sequence saturation mutagenesis to generate a mutant IGF1R library, which will be used in future in a screening approach to develop IGF1R-based decoy inhibitors of RSV entry.

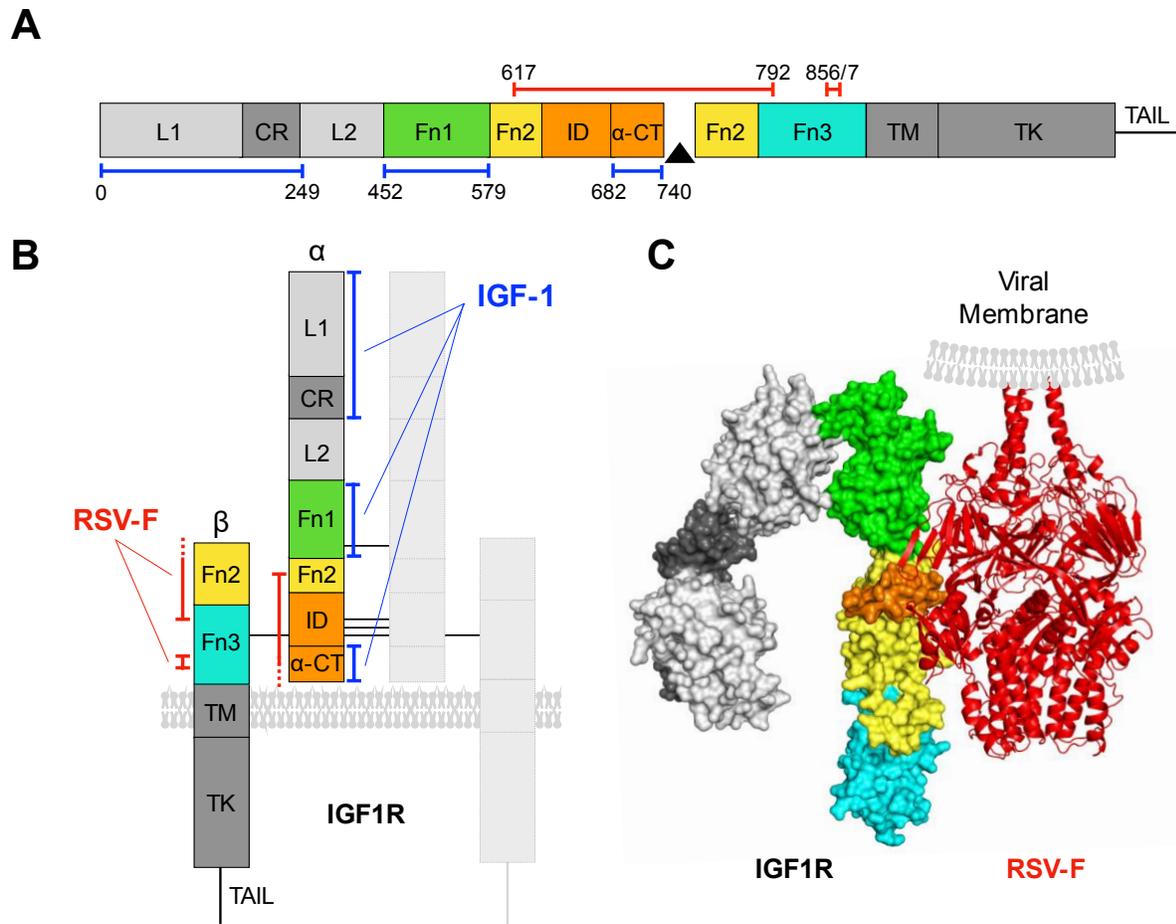


Figure 2.1. IGF1R domain organization and molecular docking of IGF1R:RSV-F. (A) Domain organization of the IGF1R protein. The L1 and L2 (leucine-rich globular domains); CR (cysteine-rich domain); Fn1, Fn2, Fn3 (fibronectin III repeats 1-3); ID (insert domain, within FnIII-2); α -CT (c-terminal region of the α -chain); TM (transmembrane); TK (Tyrosine kinase domain); and TAIL (c-terminal tail) domains are indicated. Domain boundaries as well as the IGF-1 ligand binding sites (blue) and the predicted RSV-F interaction sites (red) are indicated (16). The arrowhead denotes the cleavage site between the α and β subunits. (B) Mature IGF1R is a disulfide-linked homodimer of two $\alpha\beta$ subunits. Disulfide linkages are denoted by horizontal black lines between the two $\alpha\beta$ subunits, and domains and interaction sites for one of the IGF1R monomers is coloured as in (A). (C) Molecular docking of the IGF1R ectodomain (PDB ID: 5U8R) and RSV-F (PDB ID: 5UDC). IGF1R is shown as a surface representation and RSV-F is shown as a ribbon diagram. IGF1R domains are coloured as in (A).

2.4 MATERIALS & METHODS

2.4.1 Cell lines and viruses

Human embryonic kidney 293T (Hek 293T) cells were provided by Dr. Connie Krawczyk, Van Andel Institute, Michigan, USA. HeLa cells were kindly provided by from Dr. Martin Richer, Indiana University School of Medicine, IN, USA. Both 293T and HeLa cells were maintained in complete media [1X Dulbecco's Modified Eagle Medium (DMEM), 10% FBS, 2mM L-glutamine, 1X non-essential amino acids (Wisent)]. Human Airway Epithelial (HAE) cells and 1HAEo⁻ IGF1R knockout (KO) cells were obtained from Dr. David Marchant, University of Alberta, AB, Canada. All HAE cells were maintained in HAE media [1X Minimum Essential Media (MEM), 10% Fetal Bovine Serum (FBS), 2 mM L-glutamine, 1X non-essential amino acids], while the 1HAEo⁻ IGF1R KO cell media was additionally supplemented with 5 µg/mL blasticidin and 0.5 µg/mL puromycin. All cells were grown at 37°C, 5% CO₂. The laboratory-adapted, recombinant green fluorescent protein (GFP) RSV (rgRSV, type A2 strain, RW30²⁷) infectious complementary DNA (cDNA) system was obtained from Mark E. Peeples (Nationwide Children's Hospital, Columbus, OH, USA) (21).

2.4.2 Preparation of Viral Stocks and Focus-Forming Unit (FFU) Assays.

Infectious virus was prepared by transfection of sub-confluent HeLa cells with full-length rgRSV RW30 cDNA, with four support plasmids (RSV N, P, L, and M2-1), and a plasmid expressing T7 RNA polymerase (a gift from B. Lee; Addgene plasmid 65974) using TransIT-HeLa MONSTER (Mirus Bio, MIR 2900). After rescue, the virus was then propagated in HeLa cells in T75 flasks and collected in cell-free RSV-conditioned DMEM with 10% FBS (1). To prepare rgRSV stocks, we passaged RSV-conditioned media on HeLa cells that were seeded at $3-4 \times 10^6$

cells in 25 mL of DMEM/2% FBS in 150 mm tissue culture dishes 24 hours (h) before infection. Infections were performed at MOI 0.01 with an inoculum that contained virus diluted in DMEM (Wisent). For passage 3 (p:3), virus infections were performed for 3-4 h, for p:1 or p:2 inoculums were left overnight. After infection, the inoculum was removed and replaced with DMEM/2% FBS (Wisent). At 3 days post-infection cells were scraped, collected, and pelleted at 1200 rpm for 5 minutes (min) and the supernatant was recovered. Aliquots (500 μ L) of the supernatant were collected in cryovials, snap-frozen, and stored at -80°C until use.

rgRSV stocks were titrated by focus-forming unit (FFU) assay on HeLa cells as previously described (21). Briefly, 5-fold serial dilutions of virus stocks were incubated for 22 h at 37°C on a monolayer of HeLa cells (2×10^4 cells per well) in a 96-well plate. Viral dilutions were prepared in DMEM/5% FBS and infections were performed for 22 h at 37°C . To determine titers, wells containing between 10-100 GFP+ foci were analyzed by immunofluorescence using an inverted fluorescent microscope (Zeiss Axio Observer.A1), on the lowest magnification (10X). Viral titers are expressed as the number of FFU per mL (FFU/mL).

2.4.3 Molecular docking analyses

Crystal structures of RSV-F bound to the MEDI8897 antibody (PDB ID: 5UDC) and the IGF1R ectodomain (PDB ID: 5U8R) were used to prepare the proteins for docking. The bound MEDI8897 antibody was removed from the structure of RSV-F and the Protein Preparation Wizard implemented in Schrödinger Small Molecule Discovery Suite was used to add hydrogen atoms, perform energy minimization using the OPLS-3e forcefield, and create appropriate protonation states of the amino acid side chains (22, 23).

Molecular docking between the two proteins was investigated using two protein-protein docking servers to perform a blind docking. The first server was ClusPro 2.0 protein-protein docking webserver (<http://cluspro.bu.edu/>) using piper which employs FFT-based protein-protein docking (24). The second server was the HDOCK webserver (<http://hdock.phys.hust.edu.cn/>) (25). The best models were then selected based on the lowest energy clustered structures and the probability of the poses to match between the two webserver in which ClusPro outperformed HDOCK. Protein-protein interactions were visualized using PyMol (26).

2.4.4 Plasmids and molecular cloning

The IGF1R expression vector, pcDNA3.1-IGF1R-WT, was obtained from David Marchant (1). Alanine-scanning mutagenesis was performed via introduction of point mutations via site-directed mutagenesis using the QuikChange II XL Site-Directed Mutagenesis Kit (Agilent Technologies) according to the manufacturer's instructions. All primer sequences used to generate the alanine mutants are listed in **Table 2.1**. Each residue was mutated to alanine, except for residue 787 where the alanine was changed to a glycine residue. All plasmids were confirmed by Sanger Sequencing (Genome Québec).

To create the template for sequence saturation mutagenesis (SeSaM) we introduced a silent mutation to create a *BstBI* site within IGF1R in pcDNA3.1-IGF1R-WT using FOR_SDM_IGF1R_BstBI (5'-TGA GAA GAG GAG TTC GAA GCT GAA AGA ACG TCC AA -3') and REV_SDM_IGF1R_BstBI (5'-TTG GAC GTT CTT TCA GCT TCG AAC TCC TCT TCT CA- 3'). The IGF1R gene from pcDNA3.1-IGF1R-BstBI was then subcloned into pUC18 using *EcoRI* and *BamHI* to generate the mutant library storage vector, pUC18-IGF1R-BstBI.

Table 2.1 Primers used in SDM.

IGF1R Residue	Primer	Sequence (5'-3')
Arg617	FOR	GAG TTA CTA CAT TGT GGC CTG GCA GCG GCA GCC
	REV	GGC TGC CGC TGC CAG GCC ACA ATG TAG TAA CTC
Trp618	FOR	ACT ACA TTG TGC GCG CGC AGC GGC AGC CTC
	REV	GAG GCT GCC GCT GCG CGC GCA CAA TGT AGT
Gln619	FOR	CAT TGT GCG CTG GGC GCG GCA GCC TCA G
	REV	CTG AGG CTG CCG CGC CCA GCG CAC AAT G
Arg620	FOR	TGT GCG CTG GCA GGC GCA GCC TCA GGA C
	REV	GTC CTG AGG CTG CGC CTG CCA GCG CAC A
Tyr688	FOR	CCG AGA AGG AGG AGG CTG AAG CCC GCA AAG TCT TTG
	REV	CAA AGA CTT TGC GGG CTT CAG CCT CCT CCT TCT CGG
Phe692	FOR	GGC TGA ATA CCG CAA AGT CGC TGA GAA TTT CCT GCA CAA C
	REV	GTT GTG CAG GAA ATT CTC AGC GAC TTT GCG GTA TTC AGC C
Phe695	FOR	GCT GAA TAC CGC AAA GTC TTT GAG AAT GCC CTG CAC AAC TCC A
	REV	TGG AGT TGT GCA GGG CAT TCT CAA AGA CTT TGC GGT ATT CAG C
Leu696	FOR	GCA AAG TCT TTG AGA ATT TCG CGC ACA ACT CCA TCT TCG TGC
	REV	GCA CGA AGA TGG AGT TGT GCG CGA AAT TCT CAA AGA CTT TGC
Ser699	FOR	GAA TTT CCT GCA CAA CGC CAT CTT CGT GCC CAG
	REV	CTG GGC ACG AAG ATG GCG TTG TGC AGG AAA TTC
Ile700	FOR	ATT TCC TGC ACA ACT CCG CCT TCG TGC CCA GAC CTG
	REV	CAG GTC TGG GCA CGA AGG CGG AGT TGT GCA GGA AAT
Phe701	FOR	TTC AGG TCT GGG CAC GGC GAT GGA GTT GTG CAG G
	REV	CCT GCA CAA CTC CAT CGC CGT GCC CAG ACC TGA A
Val702	FOR	CAC AAC TCC ATC TTC GCG CCC AGA CCT GAA AGG
	REV	CCT TTC AGG TCT GGG CGC GAA GAT GGA GTT GTG
Pro703	FOR	CAA CTC CAT CTT CGT GGC CAG ACC TGA AAG GAA
	REV	TTC CTT TCA GGT CTG GCC ACG AAG ATG GAG TTG
Arg704	FOR	CTC CAT CTT CGT GCC CGC ACC TGA AAG GAA GCG G
	REV	CCG CTT CCT TTC AGG TGC GGG CAC GAA GAT GGA G
Glu744	FOR	GGA AGA GCT GGA GAC AGC GTA CCC TTT CTT TGA GA
	REV	TCT CAA AGA AAG GGT ACG CTG TCT CCA GCT CTT CC
Tyr745	FOR	GAA GAG CTG GAG ACA GAG GCC CCT TTC TTT GAG AGC AG
	REV	CTG CTC TCA AAG AAA GGG GCC TCT GTC TCC AGC TCT TC
Pro746	FOR	GCT GGA GAC AGA GTA CGC TTT CTT TGA GAG CAG
	REV	CTG CTC TCA AAG AAA GCG TAC TCT GTC TCC AGC
Phe748	FOR	GGA GAC AGA GTA CCC TTT CGC TGA GAG CAG AGT GGA TAA C
	REV	GTT ATC CAC TCT GCT CTC AGC GAA AGG GTA CTC TGT CTC C
Ser761	FOR	AGG AGA GAA CTG TCA TTG CTA ACC TTC GGC CTT TC

	REV	GAA AGG CCG AAG GTT AGC AAT GAC AGT TCT CTC CT
Asn762	FOR	GAT AAC AAG GAG AGA ACT GTC ATT TCT GCC CTT CGG CCT TTC A
	REV	TGA AAG GCC GAA GGG CAG AAA TGA CAG TTC TCT CCT TGT TAT C
Leu763	FOR	ACA AGG AGA GAA CTG TCA TTT CTA ACG CTC GGC CTT TCA CAT TC
	REV	CAA TGT GAA AGG CCG AGC GTT AGA AAT GAC AGT TCT CTC CTT G
Arg764	FOR	GGA GAG AAC TGT CAT TTC TAA CCT TGC GCC TTT CAC ATT GTA C
	REV	GTA CAA TGT GAA AGG CGC AAG GTT AGA AAT GAC AGT TCT CTC C
Tyr769	FOR	CCT TCG GCC TTT CAC ATT GGC CCG CAT CGA TAT CCA CAG
	REV	CTG TGG ATA TCG ATG CGG GCC AAT GTG AAA GGC CGA AGG
Arg770	FOR	CGG CCT TTC ACA TTG TAC GCC ATC GAT ATC CAC AGC T
	REV	AGC TGT GGA TAT CGA TGG CGT ACA ATG TGA AAG GCC G
Asp772	FOR	CAC ATT GTA CCG CAT CGC TAT CCA CAG CTG CAA CC
	REV	GGT TGC AGC TGT GGA TAG CGA TGC GGT ACA ATG TG
Phe747	FOR	CTG GAG ACA GAG TAC CCT GCC TTT GAG AGC AGA GTG GA
	REV	TCC ACT CTG CTC TCA AAG GCA GGG TAC TCT GTC TCC AG
Glu749	FOR	CAG AGT ACC CTT TCT TTG CGA GCA GAG TGG ATA ACA A
	REV	TTG TTA TCC ACT CTG CTC GCA AAG AAA GGG TAC TCT G
Ser750	FOR	AGA GTA CCC TTT CTT TGA GGC CAG AGT GGA TAA CAA GGA G
	REV	CTC CTT GTT ATC CAC TCT GGC CTC AAA GAA AGG GTA CTC T
His774	FOR	GTA CCG CAT CGA TAT CGC CAG CTG CAA CCA CGA G
	REV	CTC GTG GTT GCA GCT GGC GAT ATC GAT GCG GTA C
Ala787	FOR	CTG GGC TGC AGC GGC TCC AAC TTC GTC
	REV	GAC GAA GTT GGA GCC GCT GCA GCC CAG
Ser788	FOR	GGG CTG CAG CGC CGC CAA CTT CGT CTT T
	REV	AAA GAC GAA GTT GGC GGC GCT GCA GCC C
Phe790	FOR	CTG CAG CGC CTC CAA CGC CGT CTT TGC AAG GAC T
	REV	AGT CCT TGC AAA GAC GGC GTT GGA GGC GCT GCA G
Phe792	FOR	CGC CTC CAA CTT CGT CGC TGC AAG GAC TAT GCC C
	REV	GGG CAT AGT CCT TGC AGC GAC GAA GTT GGA GGC G
Arg856	FOR	GTC CAG ACA GGA ATA CGC GAA GTA TGG AGG GGC C
	REV	GGC CCC TCC ATA CTT CGC GTA TTC CTG TCT GGA C
Lys857	FOR	CCA GAC AGG AAT ACA GGG CGT ATG GAG GGG CCA AGC
	REV	GCT TGG CCC CTC CAT ACG CCC TGT ATT CCT GTC TGG

2.4.5 IGF1R mutant expression and RSV infections

One day prior to transfections, 6-well plates were seeded at 6×10^5 293T, HAE, or 1HAEo⁻ KO cells/well. Transfection complexes were prepared in Opti-MEM with 1 μ g pcDNA3.1-IGF1R-WT or mutant plasmids and 2 μ L Lipofectamine 2000 (ThermoFisher) according to the manufacturer's instructions. At 4 h post-transfection, transfection complexes were removed and replaced with complete media. For mutant expression, 293T cell lysates were harvested approximately 24 h post-transfection.

For infections, 24 h post-transfection HAE and 1HAEo⁻ KO cells were infected with rgRSV at an MOI 5 in 1 mL of virus inoculum containing MEM/2% FBS. One well was not infected for gating purposes. At 4 h post-infection, infection media was removed and replaced with 2 volumes of MEM complete media. At 20 h post-infection, the cells were washed with sterile 1X phosphate-buffered saline (PBS), trypsinized, collected and centrifuged at 3000 rpm at 4°C for 10 min. Cells were re-suspended in 500 μ L of 4% paraformaldehyde at room temperature for 30 min or overnight at 4°C to allow complete fixation.

2.4.6 Flow Cytometry

Cells were centrifuged at 3000 rpm at 4°C for 10 min and resuspended in 500 μ L of PBS. Approximately 100 μ L of cell suspension was combined with 100 μ L of 1X PBS in a 96-well plate and read immediately. Reporter mean fluorescence intensity was analyzed by flow cytometry to determine the percentage of GFP-expressing cells on a Guava easyCyte bench top flow cytometer (Millipore Guava EasyCyte 5HT flow cytometer). A set of 20,000 events were run for each sample. Data was analyzed with Guava Suite Software acquisition and analysis (Millipore).

2.4.7 Western blotting

To collect whole cell lysates, cells were harvested on ice in 1 mL of PBS, then resuspended in 50-200 μ L of radioimmunoprecipitation assay (RIPA) lysis buffer [150 mM sodium chloride, 1% NP-40, 0.5% sodium deoxycholate, 0.1% sodium dodecyl sulfate (SDS), 50 mM Tris, pH 8.0]. Protein concentration was quantified using the Pierce™ BCA Protein Assay Kit (ThermoFisher) according to the manufacturer's instructions. Approximately 10 μ g of protein from each sample was run out on a sodium dodecyl sulfate-polyacrylamide (SDS-PAGE) gel with a 5% stacking and 10% resolving gel (ProtoGel, National Diagnostics). The gel was transferred to an Immobilon-P PVDF transfer membrane (Millipore). The membranes were blocked in 5% milk in Tris-buffered saline with 0.1% Tween® 20 (TBS-T) for 1 h. Membranes were then incubated overnight with primary antibodies diluted in 5% BSA: monoclonal rabbit anti-IGF1R β (D23H3) antibody (Cell Signaling Technologies, 1:10,000) and rabbit anti-actin (A2066, Sigma, 1:10,000). Blots were incubated for 1 h with HRP-conjugated secondary antibody diluted in 5% milk: anti-rabbit (111-035-144, Jackson ImmunoResearch Laboratories, 1:25,000). Blots were visualized using enhanced chemiluminescence using the ECL™ Prime Western Blotting Detection Reagent (GE Healthcare).

2.4.8 Sequence saturation mutagenesis

Sequence saturation mutagenesis (SeSaM) was carried out as previously described, with some modifications (27). Briefly, in the first step the target region was amplified in the presence of specific concentrations of nucleotides with a modified back bone that can be cleaved in the presence of iodine (phosphorothioate nucleotides), which we optimized in terms of concentration for each base, using A 35%, G 30%, T 35%, C 30%. The single-stranded products of this

reaction were then isolated by magnetic streptavidin beads, which resulted in single-stranded DNA fragments of varying lengths. In step two, a universal base [6-(2-deoxy- β -D-ribofuranosyl)-3,4-dihydro-8H-pyrimido-[4,5-C][1,2]oxazin-7-one) (dPTP α S) or [1- β -D-ribofuranosyl-1,2,4-triazole-3-carboxamide] (dRTP) was added onto the cleaved products. In step three, the universal base fragments were then elongated to the full gene length, with the modification that we ran two PCR reactions: the first reaction with the A forward, G forward, T reverse, and C reverse libraries, and the second reaction with the opposite bases. Approximately 80 ng of each of the four libraries was used in each reaction. In the final step, we performed a nested PCR with 20 ng of each of the two products generated in step 3 to specifically amplify the product and replace the universal base with standard nucleotides; such that, in theory every position in the target sequence has the potential to be mutated to each of the other nucleotides.

We then determined the mutational frequency of the reaction before proceeding to library preparation. The PCR product was digested with *BstBI* and *EcoRI* and ligated into pUC18-IGF1R-BstBI. Ligation reactions were transformed into Subcloning Efficiency™ DH5 α Competent Cells to create the sample library. To determine library diversity, plasmid DNA was isolated and sequenced from 30 colonies using Sanger Sequencing (Genome Québec). Mutational analyses of the sequences were performed using the Mutanalyst software (<http://www.mutanalyst.com>) to assess the mutation frequency of the complete SeSaM reaction (28).

2.4.9 SeSaM Library preparation

Once we obtained an appropriate mutation frequency, we carried out library preparation as previously described, with a few modifications (15, 29). Briefly, to prepare the final mutant

library, the SeSaM product and pUC18-IGF1R-BstBI was digested with *BstBI* and *EcoRI* and ligation reactions were performed using 500 ng of the insert and a 3-fold molar excess of the vector. Ligation reactions were transformed by electroporation using 10-beta electrocompetent *E. coli* (NEB), according to the manufacturer's instructions. The electroporation conditions were 2.0 kV, 200 Omega, and 25 μ F using 1-mm gap electroporation cuvettes. In our modified protocol, we pooled 4 electroporations and plated 1, 2.5, and 5 μ L of onto 10-cm LB-agar ampicillin plates and 1.5 mL onto two large LB-agar ampicillin square bioassay dishes (245 mm \times 245 mm, Corning). Plates were incubated for 16 h at 37°C (10-cm) or 20 h at 30°C (245 mm \times 245 mm). The colonies on control plates (10-cm) were counted, and the cell density was used to calculate library size. The final library was scraped and collected as described previously from the two large square bioassay dishes and stored in cryogenic vials with sterile glycerol (50% v/v) at -80°C (29).

2.4.10 Data and statistical analysis

All data are displayed as a mean of three or more independent experiments and error bars indicate the standard deviation of the mean. Statistical analysis was performed using GraphPad Prism v9. Statistical significance was determined by paired t-test with no correction for multiple comparisons.

2.5 RESULTS

2.5.1 Molecular docking of RSV-F with IGF1R

Previous studies indicated that RSV-F interactions with IGF1R trigger RSV entry and fusion (1). To identify the RSV-F interaction site on the IGF1R molecule, we performed molecular docking

analysis using RSV-F (PDB ID: 5UDC) and the IGF1R ectodomain (PDB ID: 5U8R). Molecular docking analyses revealed that RSV-F primarily interacts with IGF1R in type III fibronectin domains two and three (FnIII- 2, 3), the insert domain (ID), and the C-terminal of the α chain (α CT) (**Figure 2.1C**). These interactions involved hydrogen bonds, van der waals interactions, and two salt-bridges, with an average distance of 2.5 – 3.8 angstroms. The molecular docking suggested interaction between RSV-F and one monomer of the ectodomain of IGF1R. Based on this analysis, the IGF1 and RSV-F binding sites overlap in the α CT region of IGF1R (**Figure 2.1A-B**) and the RSV-F binding site spans the α and β subunits (**Figure 2.1A-C**). The molecular modeling identified 35 residues of IGF1R that are predicted to come within 2.5 – 3.8 angstroms of RSV-F (**Table 2.2**). These residues were therefore selected for site-directed mutagenesis and downstream analyses in RSV-F-mediated viral entry.

2.5.2 Expression analysis of immature and mature forms of IGF1R alanine mutants

Based on the molecular docking analyses (**Figure 2.1 and Table 2.2**), we performed alanine-scanning mutagenesis, creating 35 individual IGF1R mutants for analysis via site-directed mutagenesis (**Table 2.2**). Where the existing IGF1R residue already had an alanine, we mutated to a glycine. First, we tested WT and mutant IGF1R expression in Hek 293T cells by western blot (**Figure 2.2**). Transfection of wild-type (WT) IGF1R resulted in a banding pattern consistent with both immature (~105 kDa) and mature (~200 kDa) IGF1R. Except for S699A and R601A, each of the 35 alanine mutants were expressed to some extent after transfection into Hek 293T cells (**Figure 2.2**). However, upon quantification, the expression of the immature and mature forms varied across the mutants, indicating that some of the mutations influenced IGF1R gene expression or processing (**Figure 2.3A-B**).

Table 2.2 IGF1R residues that interact with RSV-F^a.

IGF1R Domains			
α Chain		β Chain	
FnIII-2	α CT	FnIII-2	FnIII-3
Arg617	Try688	Glu744	Arg856
Trp618	Phe692	Try745	Lys857
Gln619	Phe695	Pro746	
Arg620	Leu696	Phe747	
	Ser699	Phe748	
	Ile700	Glu749	
	Phe701	Ser750	
	Val702	Ser761	
	Pro703	Asn762	
	Arg704	Leu763	
		Arg764	
		Tyr769	
		Arg770	
		Asp772	
		His772	
		Ala787	
		Ser788	

^aBased on molecular docking data.

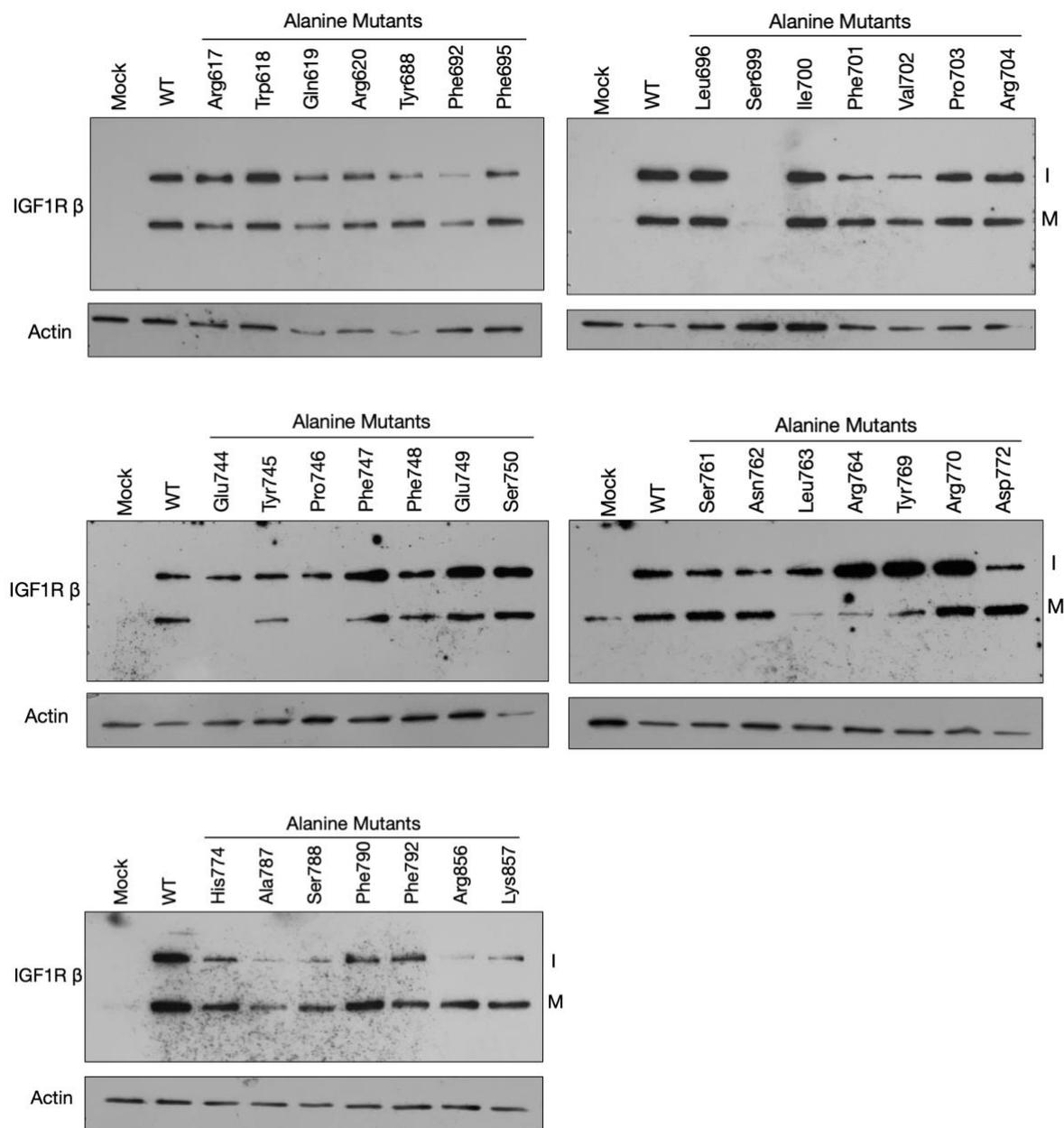


Figure 2.2. Expression of alanine mutants of IGF1R. Western blot analysis of mock, WT, and alanine mutant IGF1R-transfected Hek 293T cells at 24 h post-transfection. Immature (200 kDa) and mature (105kDa, β chain) are indicated. I, immature; M, mature.

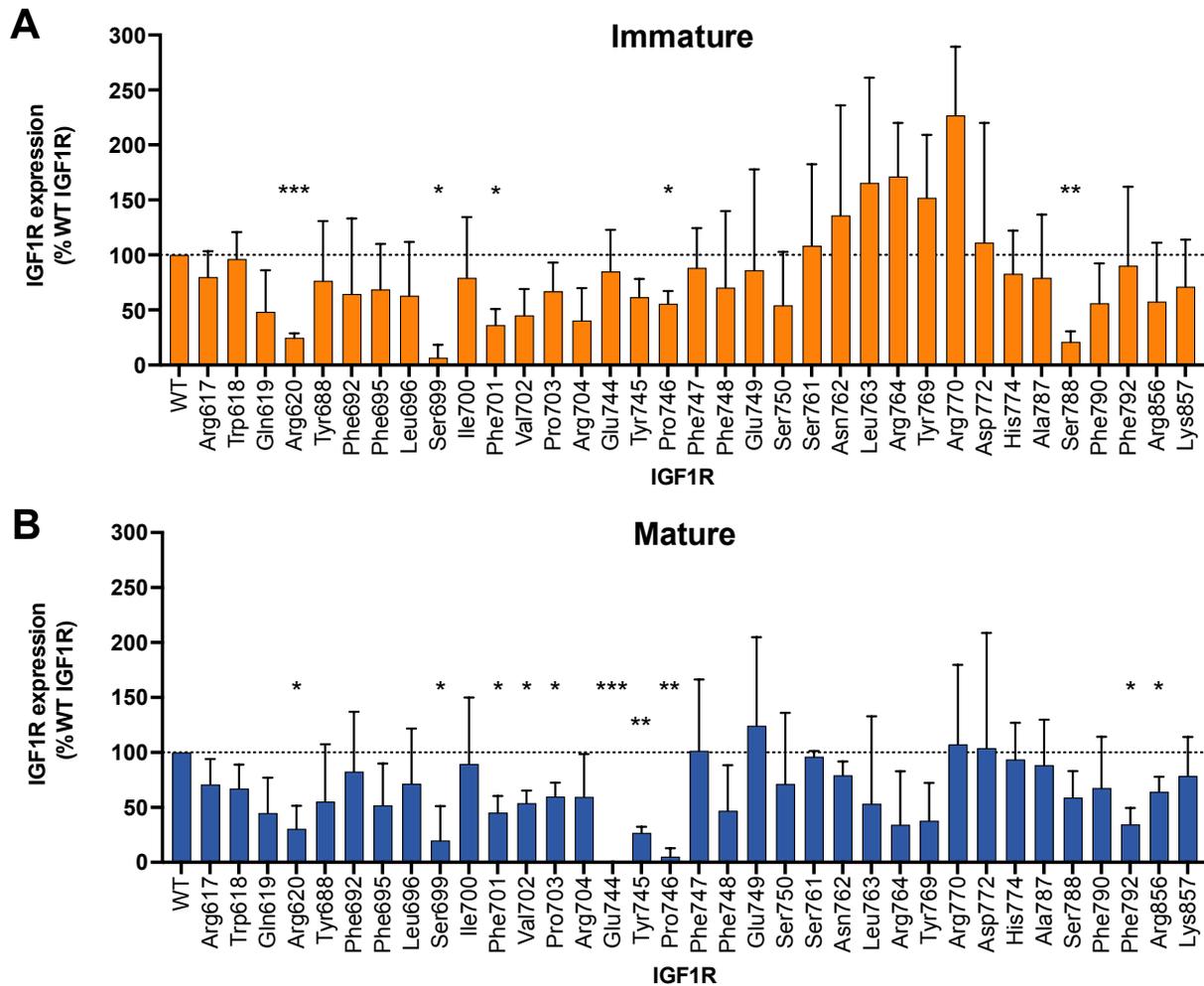


Figure 2.3. Quantification of immature and mature IGF1R expression of alanine mutants. (A) Densitometry analysis of immature IGF1R expression at 24 h post-transfection of the alanine mutants, normalized to actin and expressed as % WT IGF1R. (B) Densitometry analysis of mature IGF1R expression at 24 h post-transfection of the alanine mutants, normalized to actin and expressed as % WT IGF1R. All data are representative of three independent replicates, and error bars represent the standard deviation (SD) of the mean. P-values were calculated by paired t-test, * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

Briefly, the vast majority of the mutants had similar levels of expression of immature IGF1R to WT, with the exception of residues Arg620, Ser699, Phe701, Pro746, and Ser788, where we observed a significant decrease in the levels of immature IGF1R expression (**Figure 2.3A**). Additionally, mutations at residues Val702 and Arg 704 trended towards an overall decrease, while mutations at residues Ser762-Arg770 trended towards an overall increase in immature IGF1R expression, respectively (**Figure 2.3A**). Quantification of mature IGF1R, revealed a significant decrease in expression when mutations were made at the Arg620, Ser699, Phe701, Val702, Pro703, Glu744, Try745, Pro746, Phe792, and Arg856 residues (**Figure 2.3B**). Notably, although we were able to detect a small amount of mature IGF1R for the Arg620 and Ser699 mutant, these mutants appeared to be severely impaired in IGF1R expression. Additionally, we observed a trend toward a decrease in mature IGF1R expression when mutations were made at the Phe748, Leu763, Arg746, and Tyr769 residues (**Figure 2.3B**). Given that the IGF1R cleavage site is at position Glu740, it is not surprising that mutations at residues Glu744, Tyr745, and Pro746 result in a decrease in expression of mature IGF1R. Taken together, with the exception of Arg620 and Ser699, all of the 35 alanine mutant IGF1R constructs were expressed in Hek 293T cells, with the vast majority expressing at similar levels to WT. However, the Arg620 and Ser699 mutants were severely impaired in IGF1R expression and mutations near the IGF1R cleavage site (Glu740), as well as those at positions Phe701, Val702, Pro703, Phe792, and Arg856, had reduced mature IGF1R expression in Hek 293T cells.

2.5.3 With the exception of Ser788, all alanine mutants supported RSV infection to a similar extent as WT IGF1R.

After confirming expression of each of the alanine-scanning mutants, we next wanted to assess the ability of the alanine scanning mutants to support RSV infection (**Figure 2.4**). To do so, we developed an RSV infection assay using recombinant GFP-expression RSV (rgRSV) in WT and IGF1R knockout (KO) human airway epithelial (HAE) cells. Firstly, we compared GFP positive cells at 18 h post-infection of WT HAE cells, as well as mock or WT IGF1R-transfected IGF1R KO HAE cells (**Figure 2.4A**). While rgRSV readily infected WT HAE cells, KO HAE cells did not support rgRSV infection unless transfected with WT IGF1R (**Figure 2.4A**). Specifically, when compared with WT HAE cells, infection of KO HAE cells resulted in less than 2% eGFP positive cells, while transfection of WT IGF1R resulted in ~40% eGFP positive cells (**Figure 2.4A**). Notably, the transfection efficiency in the IGF1R KO HAE cells was on average >50% (data not shown).

Next, we assessed rgRSV infection after transfection of the 35 alanine-scanning mutants of IGF1R (**Figure 2.4B**). Interestingly, all of the alanine mutants were able to support rgRSV infection in IGF1R KO HAE cells, with the vast majority supporting rgRSV infection at a similar extent as WT IGF1R (**Figure 2.4B**). Notably, the Ser788 mutant demonstrated a significant decrease in infection (42.6% of WT), while the Leu696 or Pro703 mutants resulted in a small but significant increase in RSV infection (**Figure 2.4B**). Interestingly, since several mutants had reduced expression of mature IGF1R, at least when expressed in Hek 293T cells (**Figures 2.2 and 2.3**), these results may suggest that immature IGF1R is able to support rgRSV infection.

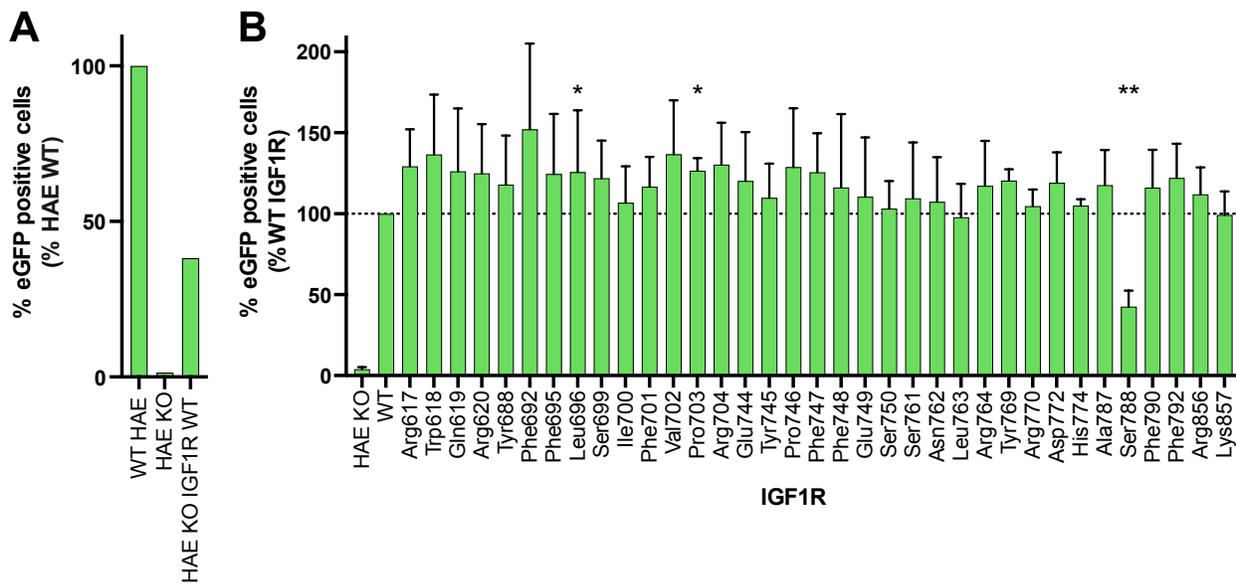


Figure 2.4. RSV infection of IGF1R KO HAE cells expressing alanine mutants. (A) WT HAE cells, as well as mock or WT IGF1R-transfected IGF1R KO HAE cells were infected with rgRSV at 24 h post-transfection, and RSV entry (eGFP positive cells) was assessed at 18 h post-infection by FACS analysis. Data is expressed as % HAE WT infection. (B) Similarly, alanine mutants were transfected into IGF1R KO HAE cells and at 24 h post-transfection cells were infected with rgRSV. Infected cells were harvested at 18 h post-infection to assess RSV entry (eGFP positive cells) by FACS analysis. Data is expressed as % eGFP positive cells normalized to the WT IGF1R-transfected. Untransfected IGF1R KO HAE cells indicate background levels of infection. All data are representative of three independent replicates, and error bars represent the SD of the mean. P-values were calculated by paired t-test, * $p < 0.05$; ** $p < 0.01$, *** $p < 0.001$.

2.5.4 SeSaM reaction and mutant library preparation

Since the molecular docking provided us with insights into the potential RSV-F binding site on IGF1R, we selected residues 619-857 for SeSaM (**Figure 2.1**). To generate a library of mutant IGF1R molecules, we performed SeSaM and then assessed mutation frequency and library diversity using the Mutanalyzer software after sequencing of ~30 clones per replicate (**Table 2.3**). Overall, we obtained an average mutation frequency of 2.43×10^{-3} substitutions/base (2.73×10^{-3} and 2.13×10^{-3} substitutions/base for replicates 1 and 2, respectively), which was in good agreement with the expected 2.8×10^{-3} substitutions/base (**Table 2.3**). As a percent of the total target sequences, we obtained deletions, frameshifts, and unmutated DNA at a ratio of 0%-7%-11.85% (with 0%-10.7%-7.1% and 0%-3.3%-16.6% for replicates 1 and 2, respectively) compared with the expected result of 0%–12.7%–5.4% (**Table 2.3**). The bias indicator, represented by the ratio of transitions to transversions, was on average 3.4 (with 2.8 and 4 for replicates 1 and 2, respectively), compared with the expected result of 1 (**Table 2.3**). Finally, we mapped out the distribution and frequency of mutations across the SeSaM target region of IGF1R, suggesting we had a relatively even distribution of mutations across the target region (**Figure 2.5**). Thus, the SeSaM performed as expected, providing mutation frequencies and distributions consistent with those previously reported, including no deletions and low frameshift and unmutated DNA frequencies. However, while we also observed a good distribution of mutations across the target region, we did observe a bias in transition vs. transversion mutations.

Based on these positive indicators, we generated a mutant IGF1R library. The efficiency of the library preparation was expected to be $>10^9$ colony forming units (CFU)/ μg and we obtained an efficiency of 22×10^{10} CFU/ μg (**Table 2.4**).

Table 2.3 Key expected performance parameters of SeSaM compared to obtained results.^a

Property	Expected	Average	Replicate^a	
			#1	#2
Deletions (%) – Frameshifts (%) – unmutated DNA (%)	0–12.7–5.4	0-7-11.85	0-10.7-7.1	0-3.3-16.6
Mutations frequency (substitution/base)	2.8×10^{-3}	2.43×10^{-3}	2.73×10^{-3}	2.13×10^{-3}
Fraction transition to transversion (bias indicator)	1.0	3.4	2.8	4

^aResults shown from two independent SeSaM reactions.

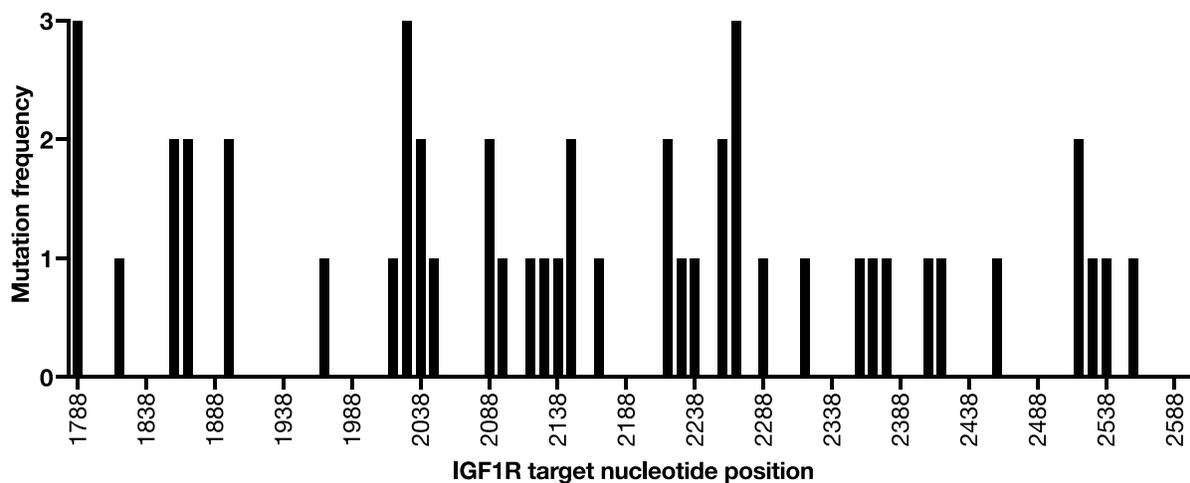


Figure 2.5. Distribution of mutations across IGF1R target sequence. The frequency of mutations at each nucleotide was mapped to the ~800 nucleotide target sequence (nucleotides 1788-2598, corresponding to amino acids 596-866) of IGF1R. Data is representative from the SeSaM replicate used to generate the IGF1R mutant library.

Table 2.4 Key expected performance parameters for a mutant library preparation compared to obtained results.

Property	Expected	Result
Efficiency	$>10^9$ range	22×10^{10} cfu/ μ g
Background	$<1\%$	3.8%
Library Size	$>2.5 \times 10^{10}$ cells/mL	5.4×10^{10} cells/mL

Additionally, based on previous studies, the background (vector only colonies) was expected to be <1%, and we observed a slightly higher background of approximately 3.8% (**Table 2.4**). Finally, the expected library size was $>2.5 \times 10^{10}$ cells/mL, and we obtained a library size of 5.4×10^{10} cells/mL (**Table 2.4**). Thus, our SeSaM-generated library of IGF1R mutants provided a library with efficiency and size consistent with previous studies; however, we observed a slightly higher background than expected.

2.6 DISCUSSION

To further understand the interaction between RSV-F and IGF1R in the context of RSV entry we performed molecular docking analyses. The docking revealed 35 residues of IGF1R that are predicted to come in close contact with RSV-F. We then employed alanine-scanning mutagenesis to examine the importance of these residues in RSV entry. Prior to testing the alanine mutants in viral entry, we confirmed expression of each of the mutants in Hek 293T cells. We chose Hek 293T cells for expression analyses due to their resilience, low background IGF1R expression, and high transfection efficiency; in contrast to the HAE KO cells, which are only able to be maintained for a few cell passages. We found that most of the alanine mutants were expressed in Hek 293T cells, with the exception of Ser699 and Arg620, which had severely reduced expression of both the immature and mature forms of IGF1R. Notably, with respect to the immature IGF1R, only Phe701, Pro746, and Ser788 had significantly reduced expression compared to WT IGF1R. It is possible that these residues impact the expression or stability of IGF1R.

Similarly, when we assessed mature IGF1R expression, we again observed that Arg620, Ser699, and Phe701 had significantly reduced expression compared with WT IGF1R. This was

not surprising given these residues resulted in a reduction in immature IGF1R expression. Interestingly, F701A is known to cause reduced binding to IGF1, and at least one study observed similar expression levels to WT IGF1R in 293 EBNA cells (constitutively expresses the Epstein-Barr virus nuclear antigen-1) (30). However, this study only expressed the ectodomain of IGF1R, and hence this may explain the discrepancy observed herein. Additionally, we observed decreases in mature IGF1R expression of Val702, Pro703, Glu744, Tyr745, Pro746, Phe792 and Arg856. Since the IGF1R cleavage site is at Glu740, mutations at residues 744-746 may thus alter IGF1R maturation. However, the reduced mature IGF1R expression of Pro703, Phe792 and Arg856 are unclear, but as these residues had similar immature IGF1R expression to WT. It is possible that they are impaired in processing, have defects in homodimer formation, or the mutations alter the stability of the mature IGF1R isoform. Interestingly, despite observing reduced expression of immature IGF1R, the S788A did not have significantly reduced levels of mature IGF1R to WT; suggesting that perhaps this mutant has accelerated IGF1R maturation or improves the stability of the mature IGF1R isoform.

After confirmation of expression, we assessed the ability of the alanine mutants to facilitate RSV infection in IGF1R KO HAE cells. We employed these cells because they are a physiologically relevant cell line, and as they are IGF1R KOs they had very low background RSV infection (**Figure 2.4A**) However, we did observe a low rate of infectivity (~2%), which is likely because this is a bulk KO cell line, and not a clonal IGF1R KO population (1). Surprisingly, despite differential expression of immature and mature IGF1R in several of our alanine mutants, with the exception of Ser788, all the mutants were able to support RSV infection to a similar extent as WT IGF1R (**Figures 2.3 and 2.4**). Notably, mutant Ser699 had low immature and mature expression when expressed in Hek 293T cells but was able to support

RSV infection at levels similar to WT IGF1R in HAE KO cells. This suggests that the defects in expression, stability, or processing that we observed in the Hek 293T cells may not carry over to the HAE KO cells. Alternatively, it is also possible that cleavage (maturation) of IGF1R is not required, and that the immature form of IGF1R is able to support RSV entry. However, to provide support for the latter, we will need to confirm immature and mature IGF1R expression levels in the HAE KO cells.

Interestingly, we observed an approximately 57.4% decrease in GFP positive cells upon expression of the Ser788 mutant compared with WT IGF1R in HAE KO cells, suggesting an impairment in RSV infection (**Figure 2.4B**). Notably, this mutant had a decrease in immature IGF1R expression in Hek 293T cells, but had close to WT levels of mature IGF1R expression (**Figure 2.3B**). The decreased infection of Ser788 could be a result of the fact that replacing a serine with an alanine exchanges a hydroxyl group for a methylenic hydrogen, thereby removing a possible region of hydrogen bonding provided by the hydroxyl group. Additionally, since serine is a phosphate group acceptor, it is possible that this residue is normally post-translationally modified by phosphorylation, which is important for viral entry. However, additional research will be needed to reveal whether this site is typically phosphorylated. It is also possible that the Ser788 mutant has a defect on cell surface expression of IGF1R. To resolve this, we plan on performing FACS analysis to assess Ser788 vs WT IGF1R cell surface expression in HAE KO cells. Finally, considering our molecular docking results indicated that all of the 35 residues of IGF1R explored herein make contact with RSV-F, it is surprising that we did not observe more mutants with significantly impaired viral entry. As such, it is possible that the molecular docking performed herein is not an accurate reflection of IGF1R:RSV-F

interactions, or simply that single amino acid mutations do not result in a drastic enough disruption of IGF1R interactions with RSV-F to impair viral entry.

Finally, we sought to exploit RSV-F:IGF1R interactions to develop IGF1R decoy-based biological inhibitors of RSV entry. To do so, we have devised a directed evolution strategy for the development of biological inhibitors of viral entry (**Figure 2.6A**). Briefly, this strategy involves directed evolution of the cellular receptor, and then a screening strategy that selects for evolved (mutant) receptors that improve viral entry but are impaired in binding to the natural ligand. We plan to use RSV-F:IGF1R interactions, as well as the IGF1 ligand in our directed evolution strategy as a proof-of-principle. To this end, we selected a region of IGF1R (nucleotides 1788-2598, corresponding to amino acids 596-866) for random mutagenesis via SeSaM. We found that the SeSaM provided a mutation distribution and frequency consistent with what was expected, and a low frequency of unmutated sequences (28). However, we did observe a higher ratio of transition-to-transversion mutations, which is not surprising given that transitions (purine to purine or pyrimidine to pyrimidine) occur more easily than transversions (purine to pyrimidine and *vice versa*). Subsequent library preparation and analysis indicated that our library size is also sufficient for screening purposes.

In future, we plan on screening our IGF1R mutant library to select for mutant receptors that have a higher affinity for RSV-F and a reduced affinity for IGF1 using a flow cytometry-based assay (**Figure 2.6**). Briefly, the IGF1R mutant will be expressed as a C-terminal fusion protein with mCherry (**Figure 2.6B**). In this way, any frameshifts or mutations which introduce a stop codon will be selected out from our screening assay as they will not produce mCherry. We will also take advantage of the rgRSV used herein, which produces eGFP, and will label the IGF1 ligand (commercially available) with A647.

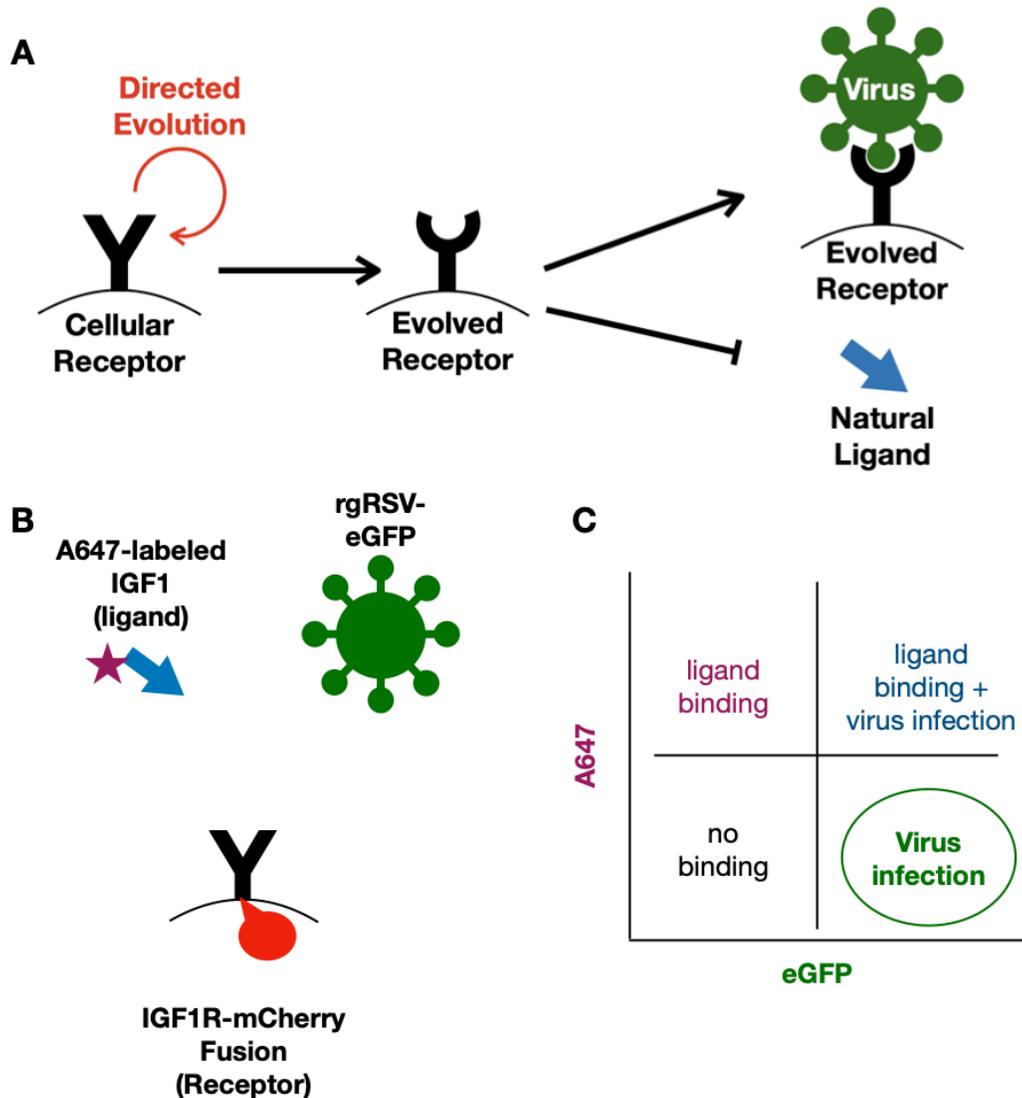


Figure 2.6. Flow cytometry-based screening assay to select for IGF1R receptors with a desired activity. (A) Directed evolution strategy for the development of biological inhibitors of viral entry. Briefly, cellular receptors are subjected to directed evolution. A library of evolved receptors are screened for enhanced viral entry and reduced binding to the natural ligand. (B) For the proof-of-principle directed evolution experiments exploiting interactions between IGF1R, RSV-F and IGF1, the IGF1R receptors are expressed as an mCherry fusion proteins. rgRSV expressing eGFP is used for infection and A647-conjugated IGF1 is used to stain cells. (C) Gating strategy used for biological inhibitor screening assay. A library of lentiviruses expressing mutant IGF1R-mCherry fusion proteins generated through SeSaM will be transduced into IGF1R KO HAE cells and subsequently infected with rgRSV. Cells will be stained with A647-IGF1 and subjected to cell sorting using flow cytometry. The sorting strategy will include gating on mCherry-expressing cells and then sorting cells that are eGFP positive (allow viral infection) but A647 negative (does not bind natural ligand).

To screen for mutant decoy receptors with increased affinity for RSV-F, but reduced affinity for IGF1, we plan to transduce cells with lentiviral particles carrying the IGF1R-mCherry mutants and infect with rgRSV (31). We will then stain with A647-IGF1 and analyze cells by flow cytometry. By gating on mCherry-expressing cells and selecting those which allow for viral infection (based on the eGFP signal), but do not allow natural ligand binding (based on the A647 signal), we will be able to screen for biological inhibitors of RSV entry that do not interfere with natural ligand-receptor interactions (**Figure 2.6C**). Subsequent rounds of directed evolution using identified mutations, will further refine and select for IGF1R decoy-based biologic inhibitors of RSV entry. Future research will thus focus on production of mutant IGF1R lentivirus production and optimization of the flow cytometry-based screening assay.

In summary, herein we used molecular docking and site-directed mutagenesis to provide insight into the IGF1R:RSV-F interface. Specifically, we found that mutation of Ser788 severely impairs viral entry and that mutations located near the IGF1R cleavage site (Glu740) lead to a decrease in mature IGF1R expression in Hek 293T cells, but did not impair RSV entry in HAE KO cells. Future work will be needed to confirm whether the latter mutants are similarly impaired in IGF1R maturation in the HAE KO cells. Additionally, we also generated a IGF1R mutant library using SeSaM that can be used in future to select for IGF1R decoy-based biological inhibitors of RSV entry. Once this directed evolution platform is established for IGF1R:RSV interactions, we hope to apply it to a variety of other important host-pathogen interactions.

2.7 ACKNOWLEDGMENTS

We would like to acknowledge Connie Krawczyk (Van Andel Institute) and Martin J. Richer (Indiana University School of Medicine) for kindly providing Hek 293T and HeLa cells, respectively. We would also like to acknowledge Mark E. Peeples (Nationwide Children's Hospital) and David Marchant (University of Alberta) for the rgRSV system and technical support/protocols to get it up and running in our group. We would also like to thank Leanne Bilawchuk, David J. Marchant, and Ahmed K. Oraby for many useful discussions over the course of this research.

2.8 FUNDING INFORMATION

This research was supported by seed funding from the McGill Interdisciplinary Initiative in Infection and Immunity (MI4). R.S.H. was supported by fellowships from the Canadian Institutes of Health Research (CIHR) Frederick Banting and Charles Best Canada Graduate Scholarship – Master's program and the Fonds de Recherche du Québec – Santé (FRQS). In addition, this research was undertaken, in part, thanks to the Canada Research Chairs program (S.M.S.).

2.9 AUTHOR CONTRIBUTIONS

R.H and S.M.S. designed the study. R.S.H., C.C. and A.K.O. performed the experiments and analyzed the data, and R.H., C.C. and S.M.S. wrote and edited the manuscript. S.M.S. and D.J.M. obtained funding for the research.

2.10 REFERENCES

1. Griffiths CD, Bilawchuk LM, McDonough JE, Jamieson KC, Elawar F, Cen Y, Duan W, Lin C, Song H, Casanova J-L, Ogg S, Jensen LD, Thienpont B, Kumar A, Hobman TC, Proud D, Moraes TJ, Marchant DJ. 2020. IGF1R is an entry receptor for respiratory syncytial virus. *Nature* 583:615-619.
2. Shi T, McAllister DA, O'Brien KL, Simoes EAF, Madhi SA, Gessner BD, Polack FP, Balsells E, Acacio S, Aguayo C, Alassani I, Ali A, Antonio M, Awasthi S, Awori JO, Azziz-Baumgartner E, Baggett HC, Baillie VL, Balmaseda A, Barahona A, Basnet S, Bassat Q, Basualdo W, Bigogo G, Bont L, Breiman RF, Brooks WA, Broor S, Bruce N, Bruden D, Buchy P, Campbell S, Carosone-Link P, Chadha M, Chipeta J, Chou M, Clara W, Cohen C, de Cuellar E, Dang D-A, Dash-yandag B, Deloria-Knoll M, Dherani M, Eap T, Ebruke BE, Echavarria M, de Freitas Lázaro Emediato CC, Fasce RA, Feikin DR, Feng L, et al. 2017. Global, regional, and national disease burden estimates of acute lower respiratory infections due to respiratory syncytial virus in young children in 2015: a systematic review and modelling study. *The Lancet* 390:946-958.
3. Ginsberg GM, Somekh E, Schlesinger Y. 2018. Should we use Palivizumab immunoprophylaxis for infants against respiratory syncytial virus? - a cost-utility analysis. *Israel journal of health policy research* 7:63-63.
4. Piedimonte G, Perez MK. 2014. Respiratory syncytial virus infection and bronchiolitis. *Pediatrics in review* 35:519-530.
5. Resch B. 2017. Product review on the monoclonal antibody palivizumab for prevention of respiratory syncytial virus infection. *Human vaccines & immunotherapeutics* 13:2138-2149.

6. Geskey JM, Thomas NJ, Brummel GL. 2007. Palivizumab: a review of its use in the protection of high risk infants against respiratory syncytial virus (RSV). *Biologics : targets & therapy* 1:33-43.
7. Rima B, Collins P, Easton A, Fouchier R, Kurath G, Lamb RA, Lee B, Maisner A, Rota P, Wang L, Consortium IR. 2017. ICTV Virus Taxonomy Profile: Pneumoviridae. *Journal of General Virology* 98:2912-2913.
8. Zhang L, Peeples ME, Boucher RC, Collins PL, Pickles RJ. 2002. Respiratory Syncytial Virus Infection of Human Airway Epithelial Cells Is Polarized, Specific to Ciliated Cells, and without Obvious Cytopathology. *Journal of Virology* 76:5654-5666.
9. McLellan JS, Ray WC, Peeples ME. 2013. Structure and function of respiratory syncytial virus surface glycoproteins. *Current topics in microbiology and immunology* 372:83-104.
10. Chirkova T, Lin S, Oomens AGP, Gaston KA, Boyoglu-Barnum S, Meng J, Stobart CC, Cotton CU, Hartert TV, Moore ML, Ziady AG, Anderson LJ. 2015. CX3CR1 is an important surface molecule for respiratory syncytial virus infection in human airway epithelial cells. *The Journal of general virology* 96:2543-2556.
11. Behera AK, Matsuse H, Kumar M, Kong X, Lockey RF, Mohapatra SS. 2001. Blocking Intercellular Adhesion Molecule-1 on Human Epithelial Cells Decreases Respiratory Syncytial Virus Infection. *Biochemical and Biophysical Research Communications* 280:188-195.
12. Marr N, Turvey SE. 2012. Role of human TLR4 in respiratory syncytial virus-induced NF- κ B activation, viral entry and replication. *Innate Immunity* 18:856-865.
13. Mastrangelo P, Hegele RG. 2012. The RSV fusion receptor: not what everyone expected it to be. *Microbes and Infection* 14:1205-1210.

14. Tayyari F, Marchant D, Moraes TJ, Duan W, Mastrangelo P, Hegele RG. 2011. Identification of nucleolin as a cellular receptor for human respiratory syncytial virus. *Nature Medicine* 17:1132-1135.
15. Battles MB, McLellan JS. 2019. Respiratory syncytial virus entry and how to block it. *Nature reviews Microbiology* 17:233-245.
16. Li J, Choi E, Yu H, Bai X-c. 2019. Structural basis of the activation of type 1 insulin-like growth factor receptor. *Nature Communications* 10:4567.
17. Adams TE, Epa VC, Garrett TPJ, Ward* CW. 2000. Structure and function of the type 1 insulin-like growth factor receptor. *Cellular and Molecular Life Sciences CMLS* 57:1050-1093.
18. Wallborn T, Wüller S, Klammt Jr, Kruis T, Kratzsch Jr, Schmidt G, Schlicke M, Müller E, Schmitz van de Leur H, Kiess W, Pfäffle R. 2010. A Heterozygous Mutation of the Insulin-Like Growth Factor-I Receptor Causes Retention of the Nascent Protein in the Endoplasmic Reticulum and Results in Intrauterine and Postnatal Growth Retardation. *The Journal of Clinical Endocrinology & Metabolism* 95:2316-2324.
19. Hakuno F, Takahashi S-I. 2018. 40 years of IGF1: IGF1 receptor signaling pathways. *Journal of molecular endocrinology* 61:T69-T86.
20. Khatib A-M, Siegfried G, Prat A, Luis J, Chrétien M, Metrakos P, Seidah NG. 2001. Inhibition of Proprotein Convertases Is Associated with Loss of Growth and Tumorigenicity of HT-29 Human Colon Carcinoma Cells: IMPORTANCE OF INSULIN-LIKE GROWTH FACTOR-1 (IGF-1) RECEPTOR PROCESSING IN IGF-1-MEDIATED FUNCTIONS *. *Journal of Biological Chemistry* 276:30686-30693.

21. Hallak LK, Spillmann D, Collins PL, Peeples ME. 2000. Glycosaminoglycan Sulfation Requirements for Respiratory Syncytial Virus Infection. *Journal of Virology* 74:10508-10513.
22. Shelley JC, Cholleti A, Frye LL, Greenwood JR, Timlin MR, Uchimaya M. 2007. Epik: a software program for pK prediction and protonation state generation for drug-like molecules. *Journal of Computer-Aided Molecular Design* 21:681-691.
23. Roos K, Wu C, Damm W, Reboul M, Stevenson JM, Lu C, Dahlgren MK, Mondal S, Chen W, Wang L, Abel R, Friesner RA, Harder ED. 2019. OPLS3e: Extending Force Field Coverage for Drug-Like Small Molecules. *J Chem Theory Comput* 15:1863-1874.
24. Desta IT, Porter KA, Xia B, Kozakov D, Vajda S. 2020. Performance and Its Limits in Rigid Body Protein-Protein Docking. *Structure* 28:1071-1081.e3.
25. Yan Y, Zhang D, Zhou P, Li B, Huang SY. 2017. HDOCK: a web server for protein-protein and protein-DNA/RNA docking based on a hybrid strategy. *Nucleic Acids Res* 45:W365-w373.
26. Schrödinger L, DeLano W. 2020. PyMOL, <http://www.pymol.org/pymol>.
27. Ruff AJ, Kardashliev T, Dennig A, Schwaneberg U. 2014. The Sequence Saturation Mutagenesis (SeSaM) Method, p 45-68. *In* Gillam EMJ, Copp JN, Ackerley D (ed), *Directed Evolution Library Creation: Methods and Protocols* doi:10.1007/978-1-4939-1053-3_4. Springer New York, New York, NY.
28. Ferla MP. 2016. Mutanalyst, an online tool for assessing the mutational spectrum of epPCR libraries with poor sampling. *BMC Bioinformatics* 17:152.
29. Hanson-Manful P, Patrick WM. 2013. Construction and Analysis of Randomized Protein-Encoding Libraries Using Error-Prone PCR, p 251-267. *In* Gerrard JA (ed),

Protein Nanotechnology: Protocols, Instrumentation, and Applications, Second Edition
doi:10.1007/978-1-62703-354-1_15. Humana Press, Totowa, NJ.

30. Mynarcik DC, Williams PF, Schaffer L, Yu GQ, Whittaker J. 1997. Identification of Common Ligand Binding Determinants of the Insulin and Insulin-like Growth Factor 1 Receptors: INSIGHTS INTO MECHANISMS OF LIGAND BINDING *. *Journal of Biological Chemistry* 272:18650-18655.
31. Vannucci L, Lai M, Chiuppesi F, Ceccherini-Nelli L, Pistello M. 2013. Viral vectors: a look back and ahead on gene transfer technology. *New Microbiol* 36:1-22.

CHAPTER 3: DISCUSSION

3.1 SUMMARY

As discussed in *Chapter 1*, a new model of RSV entry has been proposed based on recent findings (1, 2). In this model, the RSV-F protein is thought to bind specifically to IGF1R. This interaction triggers IGF1R activation, culminating in the activation of PKC ζ signaling and the subsequent recruitment of nucleolin from the nucleus to the cell surface. At the cell surface, RSV-F, IGF1R, and nucleolin interact to form a trimeric complex that facilitates RSV fusion and entry. However, the specific residues of IGF1R that are important for interacting with RSV-F are still unclear. In *Chapter 2* we performed molecular docking analysis and alanine-scanning mutagenesis to investigate the residues of IGF1R required for RSV entry. Additionally, we also wanted to exploit the interaction between IGF1R and RSV-F to develop biological inhibitors of viral entry. To accomplish this, we used SeSaM to generate a mutant IGF1R library, which will be used in a screening approach to develop IGF1R-based decoy inhibitors of RSV entry.

3.2 FURTHER UNDERSTANDING OF THE ALANINE MUTANTS OF IGF1R

3.2.1 Cell surface expression of alanine mutants of IGF1R

Herein, we analyzed the expression of the alanine mutants in Hek 293T cells, as well as their ability to facilitate RSV infection in IGF1R KO HAE cells. However, further investigation of the alanine mutants is likely still required. Firstly, in future it will be important to explore the expression of the alanine mutants in IGF1R KO HAE cells so that we can further confirm that any effect on RSV infection was a result of the specific residue change and not because of defects in IGF1R mutant expression in this cell type. Herein, we analyzed expression in Hek 293T cells due to their high transfection efficiency and robust nature, compared with the IGF1R

KO HAE cells, which are limited to only a few cell passages. Despite this, preliminary analyses suggest that they also have sufficient transfection efficiency (on average >50%, data not shown). Thus, in future it may be helpful to also explore expression of the alanine mutants in the IGF1R KO HAE cells. Additionally, while western blot analysis allowed us to assess mutant IGF1R expression levels and maturation, it may also be helpful to assess IGF1R cell surface expression by fluorescent activated cell sorting (FACS) analysis. This will allow us to confirm whether changes in RSV infection are a result of differences in cell surface expression of the IGF1R alanine mutants. Interestingly, preliminary flow cytometry analyses suggests that after transfection into IGF1R KO HAE cells, WT (68%) and the Ser788 mutant (73%) had similar levels of cell surface expression (data not shown). As such, further analysis will help confirm that changes in RSV infection observed in our alanine mutants are a result of changes in the RSV-F:IGF1R interface, rather than changes in expression, maturation, or cell surface expression of the mutants.

3.2.2 IGF1R activation

In addition to testing cell surface expression, we are also interested in understanding how the alanine mutations modulate IGF1R activation. As discussed in ***Chapter 1***, in the current model of RSV entry, IGF1R activation and signaling is thought to be important for “calling” nucleolin from the nucleus to the cell surface to facilitate viral fusion and complete the viral entry process. As such, it is also possible that one or more of the alanine mutants alters IGF1R signalling, thereby modulating RSV entry. Thus, in future, it will be useful to assess whether the IGF1R mutants have similar activation to WT IGF1R by directly testing phosphorylation of the tyrosine kinase domain of IGF1R. To do so, we could perform IGF1-mediated stimulation assays in cells

transfected with WT, a kinase dead (D1105N, negative control) and the alanine mutant IGF1Rs and assess receptor activation via western blot using a phospho-IGF1R antibody as previously described (3). Interestingly, several residues in the tyrosine kinase domain (including Tyr1131, Tyr1135, and Tyr1136) have already been identified as necessary for IGF1R kinase activation (4). Alternatively, since PKC ζ signaling is also activated downstream of IGF1R, we could also test activation of PKC ζ as a proxy for IGF1R stimulation (1). Additionally, mutations that alter IGF1-mediated activation of IGF1R may also be useful in informing the design of biological inhibitors of RSV entry (explored in more detail below in *section 3.3*). Moreover, as outlined in *Chapter 1*, IGF1R is also implicated in several disease states due to its roles in apoptosis and metabolism (5). As such, the alanine mutants generated herein may also be useful in furthering our understanding of IGF1R in other disease states, beyond RSV entry. Nonetheless, assessing whether the alanine mutants alter IGF1R activation and signaling will be informative for furthering our understanding of both ligand-receptor interactions as well as IGF1R requirements for RSV-F-mediated cell entry.

3.2.3 IGF1R:RSV-F binding affinity

In addition to assessing RSV entry, it may also be helpful in future to explore how mutations in IGF1R alter RSV-F binding affinity. To this end, microscale thermophoresis could be used to assess binding affinity, as described previously (1). This would allow us to more quantitatively assess differences in binding affinity due to our introduced alanine mutations. However, it is likely that to do so, we would need to use a modified form of RSV-F, known as DS-Cav1, that has two sets of mutations that stabilize the prefusion conformation of RSV-F (6). Specifically, DS denotes a double substitution (C155S and C290S), while the Cav1 substitutions (S190F and

V207L) fill hydrophobic cavities and prevent side-chain clashes, which together stabilize the prefusion form of RSV-F (6). Nonetheless, quantifying the binding affinities of alanine mutant IGF1Rs with RSV-F would be helpful in furthering our understanding of this important interaction and may provide further insights in the development therapeutic antibodies or RSV entry-targeted antivirals.

3.2.4 Post-translational modifications of IGF1R

In *Chapter 2*, we identified one mutation that when mutated to alanine was able to significantly reduce RSV infection, specifically Ser788. In future, it would be helpful to further investigate why mutation of this one residue was able to significantly impair RSV entry. Interestingly, it has been established that phosphorylation typically occurs at serine, threonine, and tyrosine residues (7). Therefore, it is possible that this residue is modified by phosphorylation (or even some other form of modification) which is important for RSV entry. In future, we plan to generate a phosphomimetic amino acid like aspartic acid and test the impact of this substitution on RSV entry. This may help implicate phosphorylation of this residue on RSV entry. Additionally, it is also possible that our introduced mutations induced or altered other post-translational modifications, and thus as more information becomes available about IGF1R post-translational modifications, we will explore whether differences in RSV entry are due to alterations in post-translational modification of IGF1R.

3.3 DIRECTED EVOLUTION ASSAY

3.3.1 *Directed Evolution using SeSAM*

Herein, we describe our progress towards a directed evolution assay for the development of biological inhibitors of viral entry (**Figure 2.6**). Specifically, as a proof-of-principle, we are exploiting host cell receptor interactions with the viral envelope protein, the interactions necessary for viral entry. As described in *Chapters 1 and 2*, we are using directed evolution to create a mutant library, and plan to subsequently select for mutants which are able to bind to the virus with high affinity, but are no longer able to bind to the natural ligand. In this way, we can develop biological decoy inhibitors of viral entry without disrupting natural ligand interactions. Herein, we have described our progress towards IGF1R-based decoy receptors for inhibition of RSV entry.

To perform directed evolution, we made use of SeSaM to introduce random mutations into a relevant region of IGF1R. We chose to use SeSaM as it employs the use of a universal base and provides each position in the target gene with equal likelihood of substitution with each of the other bases, thereby minimizing bias (8). Our results herein suggested that we had good mutation frequency and distribution, with minimal bias; and this allowed us to generate a large diverse library of mutant IGF1R sequences. However, alternative methods exist for the introduction of mutations, such as mutagenetic chemicals or physical damage, such as ultraviolet irradiation (9, 10). These techniques are useful for genome wide mutations; however, since we have a specific target region and gene in mind, SeSaM or error-prone PCR (epPCR) are more appropriate (10).

Interestingly, we also briefly explored epPCR which involves manipulating the PCR reaction conditions to increase the error rate of the polymerase, resulting in the amplification of a gene of interest, while introducing a variety of random point mutations (11). This is achieved by

using a low fidelity polymerase and increasing the error rate of the polymerase by increasing the polymerase concentration, extension time, MnCl₂ concentration, and altering the dNTP ratios in the PCR (11). However, the main downside of epPCR is that there is a mutational bias introduced by the polymerase, for example with *Taq* polymerase there are more AT→GC transitions and AT→TA transversions (12). While there have been attempts to cancel out bias using a combination of two polymerases with opposite mutational spectrums, when we tried this combination, we still observed a high mutational bias in addition to a low mutation frequency (data not shown) (12). As such, SeSaM proved to be the superior method with a low bias, and was a robust approach for the directed evolution assay developed herein (**Chapter 2**).

In future, should we feel that the SeSaM library is not diverse enough or contains too many wild-type (unmutated) sequences we could also consider additional methods. For example, previous studies have made use of *E. coli* mutator strains, such as XL-1-red, which is deficient in DNA repair pathways (13). Use of such mutator strains results in mutations introduced but not repaired when the plasmid is replicated. A possible downside is that deleterious mutations could arise in the bacteria themselves, decreasing their fitness and altering growth of the culture. However, it is most likely that natural selection will cause the fit mutants to outgrow any mutants with severe growth defects (14). Nonetheless, using this approach, we would need a way to ensure that we limit mutations to our region of interest. This could be accomplished by placing this region into a plasmid using a restriction enzyme fragment and then using those same sites to recover the mutants. Those mutants with mutations in the restriction enzyme site(s) would be lost during cloning, but this shouldn't affect final library preparation.

Finally, another possible alternative method to consider is DNA shuffling. In this method, a mutant library can be further mutated by randomly digesting the target sequence with DNase I

and then reassembling these fragments by overlapping PCR or staggered extension process (StEP) (15, 16). DNA shuffling could be an option if we need to increase the diversity of our library, as it could be performed on our established SeSaM library. However, a delicate balance needs to be achieved because as the mutation rate increases, so does the risk of introducing too many deleterious mutations (17). Nonetheless, several alternative methods are available for directed evolution, but the success thus far with SeSaM suggests that it may be a superior technique for the purposes of developing a diverse library of IGF1R-based decoy receptors to select for biological inhibitors of RSV entry.

3.3.2 Assay considerations and challenges

In our directed evolution platform, to select for receptors with higher affinity for RSV-F and a lower affinity for the natural ligand (IGF1), we may need to perform several rounds of SeSaM and selection (**Figure 2.6**). Briefly, after selecting for IGF1R-expressing (mCherry positive) cells, which allow for rgRSV infection (eGFP positive), but not IGF1 binding (A647 negative), we plan to sequence the IGF1R-transduced plasmids. Identified mutations will be introduced into IGF1R and subjected to an additional round(s) of SeSaM and screening. This will allow us to further refine IGF1R decoy receptors with the desired activity. This may also be further informed by molecular docking as well as our alanine scanning mutagenesis analyses, where we identified specific residues which may impact IGF1R:RSV-F interactions. Future analyses of our IGF1R mutants in stimulation assays (as described above in **section 3.2.2**) may also identify mutations that inhibit IGF1 interactions and/or signaling which may also be useful in the design of IGF1R-based decoy receptors.

3.3.3 Application of the directed evolution platform to other virus-host interactions

In future, we hope to apply the directed evolution platform for the development of biological inhibitors of other virus-host interactions, including but not limited to viral entry. In the case of RSV, viral entry and fusion is a lucrative target for the development of antivirals; however, this may also be true for other viruses that fuse at the plasma membrane, like HIV-1 and SARS-CoV-2 (18, 19), or those which fuse with the endosome, like influenza A and Ebola virus (20).

Moreover, the platform has the potential to be applied to any virus-host interaction and does not need to be limited to viral entry. However, the benefit of the approach is the ability to select for biological inhibitors that bind with a higher affinity to the viral protein, but with lesser affinity to the natural ligand. As such, the most lucrative targets will be those which have only one or a few well-characterized natural ligands, and where the ligand could be easily introduced into the screening assay, so as to minimize natural receptor-ligand interactions in the host. In summary, while directed evolution has traditionally been applied to improve enzyme fidelity or efficiency, we are working here towards a directed evolution platform for the development novel antivirals. We hope that in the future this will aid in the development of biological inhibitors of RSV entry but will also be applied to other virus-host interactions to develop antivirals for a variety of important viral pathogens.

3.4 CONCLUDING REMARKS

Our investigations herein centered on viral entry, specifically the interaction between IGF1R and RSV-F. Through molecular docking and alanine scanning mutagenesis, we were able to identify the potential interaction interface between IGF1R and RSV-F and identified at least one residue (Ser788) that, when mutated, severely impaired RSV entry. Additionally, we identified several

IGF1R residues that appear to be important for processing and maturation of IGF1R. Finally, we also made progress towards the development of an antiviral screening assay based on directed evolution. As a proof-of-principle, we generated a mutant IGF1R library using SeSaM that will be used in future in this assay to develop biological inhibitors of RSV entry. We anticipate that in future this platform can be applied to other important virus-host interactions, to develop biological inhibitors for a variety of important human and veterinary pathogens.

3.5 REFERENCES

1. Griffiths CD, Bilawchuk LM, McDonough JE, Jamieson KC, Elawar F, Cen Y, Duan W, Lin C, Song H, Casanova J-L, Ogg S, Jensen LD, Thienpont B, Kumar A, Hobman TC, Proud D, Moraes TJ, Marchant DJ. 2020. IGF1R is an entry receptor for respiratory syncytial virus. *Nature* 583:615-619.
2. Tayyari F, Marchant D, Moraes TJ, Duan W, Mastrangelo P, Hegele RG. 2011. Identification of nucleolin as a cellular receptor for human respiratory syncytial virus. *Nature Medicine* 17:1132-1135.
3. Li J, Choi E, Yu H, Bai X-c. 2019. Structural basis of the activation of type 1 insulin-like growth factor receptor. *Nature Communications* 10:4567.
4. Lopaczynski W, Terry C, Nissley P. 2000. Autophosphorylation of the Insulin-like Growth Factor I Receptor Cytoplasmic Domain. *Biochemical and Biophysical Research Communications* 279:955-960.
5. Yuan J, Yin Z, Tao K, Wang G, Gao J. 2018. Function of insulin-like growth factor 1 receptor in cancer resistance to chemotherapy. *Oncol Lett* 15:41-47.
6. McLellan JS, Chen M, Joyce MG, Sastry M, Stewart-Jones GBE, Yang Y, Zhang B, Chen L, Srivatsan S, Zheng A, Zhou T, Graepel KW, Kumar A, Moin S, Boyington JC, Chuang G-Y, Soto C, Baxa U, Bakker AQ, Spits H, Beaumont T, Zheng Z, Xia N, Ko S-Y, Todd J-P, Rao S, Graham BS, Kwong PD. 2013. Structure-based design of a fusion glycoprotein vaccine for respiratory syncytial virus. *Science (New York, NY)* 342:592-598.

7. Ardito F, Giuliani M, Perrone D, Troiano G, Lo Muzio L. 2017. The crucial role of protein phosphorylation in cell signaling and its use as targeted therapy (Review). *Int J Mol Med* 40:271-280.
8. Ruff AJ, Kardashliev T, Dennig A, Schwaneberg U. 2014. The Sequence Saturation Mutagenesis (SeSaM) Method, p 45-68. *In* Gillam EMJ, Copp JN, Ackerley D (ed), *Directed Evolution Library Creation: Methods and Protocols* doi:10.1007/978-1-4939-1053-3_4. Springer New York, New York, NY.
9. Nannemann DP, Birmingham WR, Scism RA, Bachmann BO. 2011. Assessing directed evolution methods for the generation of biosynthetic enzymes with potential in drug biosynthesis. *Future medicinal chemistry* 3:809-819.
10. Packer MS, Liu DR. 2015. Methods for the directed evolution of proteins. *Nature Reviews Genetics* 16:379.
11. Cadwell RC, Joyce GF. 1992. Randomization of genes by PCR mutagenesis. *Genome research* 2:28-33.
12. Vanhercke T, Ampe C, Tirry L, Denolf P. 2005. Reducing mutational bias in random protein libraries. *Analytical Biochemistry* 339:9-14.
13. Muteeb G, Sen R. 2010. Random Mutagenesis Using a Mutator Strain, p 411-419. *In* Braman J (ed), *In Vitro Mutagenesis Protocols: Third Edition* doi:10.1007/978-1-60761-652-8_29. Humana Press, Totowa, NJ.
14. Maharjan RP, Liu B, Li Y, Reeves PR, Wang L, Ferenci T. 2013. Mutation accumulation and fitness in mutator subpopulations of *Escherichia coli*. *Biology Letters* 9:20120961.

15. Stemmer WP. 1994. DNA shuffling by random fragmentation and reassembly: in vitro recombination for molecular evolution. *Proc Natl Acad Sci U S A* 91:10747-51.
16. Wong TS, Roccatano D, Schwaneberg U. 2007. Steering directed protein evolution: strategies to manage combinatorial complexity of mutant libraries. *Environmental Microbiology* 9:2645-2659.
17. Ferla MP. 2016. Mutanalyst, an online tool for assessing the mutational spectrum of epPCR libraries with poor sampling. *BMC Bioinformatics* 17:152.
18. Melikyan GB. 2014. HIV entry: a game of hide-and-fuse? *Curr Opin Virol* 4:1-7.
19. Tang T, Bidon M, Jaimes JA, Whittaker GR, Daniel S. 2020. Coronavirus membrane fusion mechanism offers a potential target for antiviral development. *Antiviral Res* 178:104792.
20. Kawaoka Y. 2005. How Ebola virus infects cells. *N Engl J Med* 352:2645-6.

APPENDIX

Permissions to use copyright material

Figure 1.1 From Battles, M. B. & McLellan, J. S. Respiratory syncytial virus entry and how to block it. *Nature Reviews Microbiology*, 2019, 17, 233-245. Permission was granted by Springer Nature and Copyright Clearance Center, provided the original work is properly cited and acknowledgement is included with the figure.

Figure 1.2 From Schildgen, V. et al. Human Metapneumovirus: Lessons Learned over the First Decade. *Clinical Microbiology Reviews*, 2011, 24, 734-754. Permission was granted by American Society for Microbiology - Journals; permission conveyed through Copyright Clearance Center, Inc.

Figure 1.3 From Bermingham IM, Chappell KJ, Watterson D, Young PR. The Heptad Repeat C Domain of the Respiratory Syncytial Virus Fusion Protein Plays a Key Role in Membrane Fusion. *Journal of Virology*, 2018, 92, 4. Permission was granted by American Society for Microbiology - Journals; permission conveyed through Copyright Clearance Center, Inc.

Figure 1.4 From Graham, B. S. Vaccine development for respiratory syncytial virus. *Current Opinion in Virology*, 2017, 23, 107-112. Permission was granted by Elsevier and Copyright Clearance Center.

Figure 1.7 From Li, J., Choi, E., Yu, H. & Bai, X.-c. Structural basis of the activation of type 1 insulin-like growth factor receptor. *Nature Communications*, 2019, 10, 4567. This is an open-access article under the terms of the Creative Commons License (Attribution 4.0 International), allowing reuse of the material, provided the original work is properly cited.

Entry model

Figure 1.8 From Griffiths, C. D. et al. IGF1R is an entry receptor for respiratory syncytial virus. *Nature*, 2020, 583, 615-619. Permission was granted by Springer Nature and Copyright Clearance Center, provided the original work is properly cited and acknowledgement is included with the figure.