Identification and Analysis of the Heparan Sulfate-Binding Domain and Cellular Factors Involved in the Entry of Human Endogenous Retrovirus K HERV-K (HML-2)

Inaugural-Dissertation

to obtain the academic degree

Doctor rerum naturalium (Dr. rer. nat.)

submitted to the Department of Biology, Chemistry, Pharmacy

of Freie Universität Berlin

by

Alaa Ramadan

2022

This work was carried out under the supervision of Prof. Dr. Norbert Bannert at Robert-Koch Institute, Division for HIV and other retroviruses (formerly)- Berlin, from Jan 2017 to Jul 2021.

This work was funded by the German Academic Exchange Service (DAAD) (Research grant: Doctoral Programmes in Germany, 2016/17)

DAAD
Deutscher Akademischer Austausch Dienst
German Academic Exchange Service

ROBERT KOCH INSTITUT

First Reviewer: Prof. Dr. Norbert Bannert,

Robert Koch-Institute

Unit 18- Sexually transmitted bacterial Pathogens (STI) and HIV

Nordufer 20, 13353 Berlin

Second Reviewer: Prof. Dr. Haike Antelmann,

Freie Universität Berlin

Institute of Biology- Microbiology

Königin-Luise-Straße 12 – 16, 14195 Berlin

Date of defense: 05.Oct.2022

Acknowledgements

I am most thankful to the Almighty Allah for His mercy and grace for granting me this opportunity and lending me the strength, patience, and guidance to undertake this challenging journey.

Secondly, my gratitude, thanks, and sincere appreciation go to my supervisor Prof. Dr. Norbert Bannert, for admitting me to work in his lab as a member of his team, providing constant assistance and supervision, and for the insightful comments and suggestions on this dissertation. Special thanks also go to Prof. Dr. Reinhard Kunze for his advice during my study, as well as to Prof. Dr. Haike Antelmann for privileging me to be my dissertation's second reviewer.

In addition, my heartfelt and earnest gratitude is extended to Dr. Oliver Hohn for all the support and counsel he gave me during my study. I also want to express my profound thanks to Dr. Oya Cingöz for the valuable conversations, helpful comments, and assistance proofreading the dissertation. My genuine thanks go to Dr. Vladimir Morozov for his advice during my study.

Many special thanks to Ms. Anja Richter, Ms. Martina Keller, Ms. Monika Jaensch, Ms. Ewelina Caspers, Mr. Philipp Rische, and Mr. Maik Kronberg for invaluable technical assistance. Big thanks also go to Ms. Sabina Reichert and Ms. Kornelia Gericke for their brilliant lab service. Thanks should also go to my office mates, colleagues, and all team members in FG18, including the former members who impacted and inspired me.

I would like to thank my friend Prof. Hilary Silver who assisted with proofreading the manuscript.

Thanks to the DAAD, who has made my study possible, funded my entire research, and supported my family and me in every possible way.

Finally, I owe a great deal of gratitude to my lovely husband, Mohammad Ali Sheiban, for his patience and understanding, as well as my kids, Omar and Zahr Al-Yasmin, who cheered me up during the difficult times. I am also grateful to my parents, who kept praying for me. My dissertation was not possible without their support and love.

Dedication

To my mom and dad, for unlimited love, endless support, praying for me, and pushing me to seek knowledge.

To my precious husband and soulmate Mohammad Ali, for your great love, patience, sacrifices, wiping my tears, ongoing support, and sharing my passion.

To my wonderful kids, Omar and Zahr Al-Yasmin, for your sweet love, support, taking responsibility, and blooming my life with joy and happiness.

To my late great friend and mentor, Dr. Linda S. Fawzi

Declaration of Independence

Herewith I certify that I have prepared and written my thesis independently and that I have not used any sources and aids other than those indicated by me. I also certify that my thesis has not been previously submitted to any other institution.

Berlin 29.06.2022

_____

Alaa Ramadan

Table of Contents

List of Figure

List of Tables

List of Abbreviations

Abbr.	abbreviation(s), abbreviated
ALS	amyotrophic lateral sclerosis
1KGP	1000 Genomes Project
aa	amino acid
Arg or R residue	arginine
BAC	bacterial artificial chromosome
BCA assay	Bicinchoninic acid assay
C12ORF59	chromosome 12 open reading frame 59
CA	capsid
CCR5	C-C chemokine receptor type 5
CD4	cluster of differentiation 4
cDNA	complementary DNA
coEnv	codon-optimized envelope protein
CpG islands	region of DNA that contains many CG sites in linear sequence
CT	cytoplasmic tail
C-terminus	Carboxy-terminus
CXCR4	C-X-C chemokine receptor type 4
ddNTPs	2', 3'-dideoxynucleotide triphosphate mix
DMEM	Eagle's minimal essential medium
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
dNTP	Deoxy-nucleoside triphosphate
DS	dermatan sulfate
dsDNA	double-stranded DNA
E.coli	Escherichia coli
ECL	enhanced chemiluminescence
EDTA	Ethylenediaminetetraacetic acid
EGFR	epidermal growth factor receptor
ELISA	enzyme-linked immunosorbent assay
ENTV	enzootic nasal tumor virus
Env	envelope glycoprotein (Env)

env	the envelope-gene
ER	endoplasmic reticulum
ERV	endogenous retrovirus
FACS	fluorescence-activated cell sorting
FBS	fetal bovine serum
FT	fibritin
gag	group-specific antigen encoding gene
Gag	group-specific antigen- proteins
GAGs	glycosaminoglycans
Gal	d-galactose
GalNAc	N-acetyl-d-galactosamine
GFP	green fluorescent protein
GlcA	d-glucuronic
GlcNAc	N-acetyl-d-glucosamine
GLUT1	Glucose transporter 1
GM-CSF	Granulocyte–macrophage colony-stimulating factor
GOLPH3	Golgi membrane phosphoprotein 3
gp120	HIV envelope glycoprotein
GPI	glycosylphosphatidylinositol anchor
H1N1	influenza A virus subtype Hemagglutinin Type 1 and Neuraminidase Type 1
H3N2	influenza A virus subtype Hemagglutinin Type 3 and Neuraminidase Type 2
H5N1	influenza A virus subtype Hemagglutinin Type 5 and Neuraminidase Type 1
HBD	HP/HS binding domains
HBV	hepatitis B virus
HCV	hepatitis C virus
HERV	human endogenous retrovirus
HIV	human immunodeficiency virus
HML	human MMTV-like
HP/HS	heparin/heparan sulfate
HRP	horseradish peroxidase

HSPG	heparan sulfate proteoglycans
HSV-1	human herpes simplex virus 1
HTLV-1	Human T-cell lymphotropic virus type 1
Hyal2	hyaluronidase 2
ICTV	The International Committee on Taxonomy of Viruses
IdoA	l-iduronic acid
IL	Interleukin
IL-1R1	IL-1 receptor 1
IL1RAP	interleukin 1 receptor accessory protein
IL1RL2	IL-1 receptor-like 2
IN	integrase
JSRV	Jaagsiekte sheep retrovirus
KS	keratan sulfate
LacdiNAc	N,N'-diacetyllactosamine
LB-Medium	Lysogeny Broth
LETMD1	LETM1 domain-containing protein 1
LINE	long-terminal interspersed elements
LPS	Lipopolysaccharide
LTR	long terminal repeats
LY6E	lymphocyte antigen-6
Lys or (K)	lysine
MA	matrix
MAC	Membrane Attack Complex
MACPF	Membrane Attack Complex/Perforin
M-CSF	Macrophage colony-stimulating factor
mitogen-activated protein kinase	MAP kinases
MMTV	mouse mammary tumor virus
MOI	multiplicity of infection
MPEG1	macrophage-expressed gene 1
MPMV	Mason-Pfizer monkey virus
mRNA	messenger ribonucleic acid
MyD88	myeloid differentiation primary response 88

Myr	million years
NC	nucleocapsid
NF-κB	nuclear factor "kappa-light-chain-enhancer" of activated B-cells
NRP-1	neuropilin-1
ORF	open reading frame
oricoEnvΔ659-699/ oricoEnvΔC1	The reconstituted, codon-optimized, and CT truncated envelope protein
OX-GP73	overexpression GP73
PAGE	polyacrylamide gel electrophoresis
PBS	primer binding site
PBS	*Phosphate-buffered saline*
PCR	polymerase chain reaction
PEI	polyethyleneimine
PERT	product-enhanced reverse transcriptase
PFA	paraformaldehyde
PIC	pre-integration complex
PMA	phorbol myristate acetate
pol	RT and IN encoding gene
PPT	polypurine tract
PR	protease
pro	protease coding gene
PSCA	prostate stem cell antigen
pSIVgml	prosimian immunodeficiency virus
RBS	receptor-binding site
RELIK	rabbit endogenous lentivirus type K
Rev/Rex/Rec	regulator of expression of virion proteins
RPMI 1640	Roswell Park Memorial Institute 1640 medium
RT	reverse transcriptase
RTC	reverse-transcription complex
SDS	sodium dodecyl sulfate
SINE	short-terminal interspersed elements
SLC1A5	Solute Carrier Family 1 Member 5

SM	SU+TM with eliminated cleavage site
SOC medium	super optimal broth
SP	signal peptide
SRVs	simian type D retroviruses
ssRNA	single-stranded RNA
SU	surface unit
TAE buffer	Tris-acetate-EDTA buffer
Tat	Trans-Activator of Transcription
TfR1	transferrin receptor 1
TGN	Trans-Golgi network
TLRs	Toll-like receptors
TM	transmembrane unit
TMEM9	transmembrane protein 9
TRIF	TIR-domain–containing adapter-inducing IFN-b
tRNA	Transfer ribonucleic acid
U3	unique 3' end
U5	unique 5' end
uPAR	urokinase-type plasminogen activator receptor
V- ATPase	vacuolar-ATPase
VSV-G	vesicular stomatitis virus envelope glycoprotein
XMRV	Xenotropic murine leukemia virus–related virus
XPR1	Xenotropic and Polytropic retrovirus Receptor 1
XRVs	exogenous retrovirus
XSU	XMRV-SU fusion protein

Summary

The human endogenous retroviruses (HERVs) are remnant elements of ancient retroviruses that infected the human ancestors' germline and integrated into their genome, and thereby passed down to their descendants in a mendelian fashion. The HERV-K (HML-2) family includes the most intact and complete sequences of human-specific elements with conserved ORF for most of their proteins and the ability to produce viral particles. The viral envelope protein (Env) promoted the initial endogenization activities and was shown to have a broad tropism. It mediates the viral entry through interaction with the cell surface heparan sulfates (HS).

In this study, we demonstrated first that the specified receptor binding site (RBS) of HML-2 Env is not involved in binding HS. Further, to identify the involved domain in binding HS, we aligned the protein sequence of HML-2 Env to that of the mouse mammary tumor virus (MMTV). This way, we identified an HS-binding domain (HBD) at the N-terminus of the HML-2 Env between residues 216 and 236, corresponding to MMTV HBD. Generated mutations in all positively charged residues of this domain impaired binding to heparin-coated beads and blocked viral entry. Moreover, mutations in the residues R_{216}, K_{219}, and K_{223}, were influential in HS binding and the viral entry.

In the second and third parts of the study, we aimed to identify the cellular factors or requirements for HML-2 Env in its early entry pathway utilizing two approaches. In the first approach, we successfully generated a trimer fusion protein for HML-2 consisting of the SU subunit fused with fibritin, a trimerization domain derived from the bacteriophage T4. The trimer fusion protein bound HS similarly to the native Env. However, it did not block HML-2 Env viral entry. Using this trimer fusion protein, we specifically co-purified the Golgi membrane protein 73 (GP73) as a potential attachment factor for HML-2 Env. The second approach included using HML-2 Env pseudotyped viral particles in conducting a functional screening of a cDNA library for transmembrane proteins in a non-permissive cell line. Screening results identified six transmembrane proteins (TMEM9, C12ORF59, IL1RAP, PSCA, LETMD1, and MPEG1) involved in the entry pathway of HML-2 that confer the susceptibility to it. We also found that only IL1RAP and MPEG1 proteins confer susceptibility in a different non-permissive cell line. Our results affirmed the role of HS in the HML-2 Env attachments and identified the specific HBD involved. They also affirmed the contribution of another molecule(s) that serves as the receptor(s). In those lines seven transmembrane proteins involved in the entry pathway of HML-2 Env were identified.

Zusammenfassung

Die Humanen Endogenen Retroviren (HERVs) sind Restbestandteile alter Retroviren, die die Keimbahn der menschlichen Vorfahren infizierten und in ihr Genom integriert wurden. Sie werden seither nach den Mendelschen Regeln an ihre Nachkommen weitergegeben. Die HERV-K (HML-2)-Familie umfasst die intaktesten und vollständigsten Sequenzen mit konservierten Leserahmen für die meisten ihrer Proteine und die Fähigkeit, virale Partikel zu produzieren. Das virale Hüllprotein (Env) förderte die anfängliche Endogenisierung und zeigt einen breiten Tropismus. Es vermittelt die Infektion durch erste Wechselwirkung mit Heparansulfaten (HS) auf der Zelloberfläche. In dieser Studie zeigten wir, dass die Rezeptorbindungsstelle (RBS) von HML-2 Env nicht an der Bindung von HS beteiligt ist. Um die daran beteiligte Domäne zu identifizieren, haben wir die Proteinsequenz von HML-2 Env mit der des Maus-Mammatumorvirus (MMTV) verglichen. Am N-Terminus des HML-2 Env wurde so eine potentielle homologe HS-bindende Domäne (HBD) zwischen den Positionen 216 und 236 identifiziert. Mutationen in allen positiv geladenen Resten dieser Domäne beeinträchtigten die Bindung an Heparin-beschichtete Beads und blockierten die Infektion. Darüber hinaus waren Mutationen an den Positionen R216, K219 und K223 für die HS-Bindung und den viralen Zelleintritt maßgeblich. Im zweiten und dritten Teil der Studie wollten wir die zellulären Faktoren für den HML-2 Env vermittelten Zelleintritt mit zwei Ansätzen identifizieren. Im ersten Ansatz ist es gelungen, ein Trimer-Fusionsprotein für HML-2 zu erzeugen, das aus der SU-Untereinheit besteht, die mit der Trimerisierungs-domäne fusioniert ist. Sie ist vom Fibritin des Bakteriophagen T4 abgeleitet. Das Trimer-Fusionsprotein bindet HS ähnlich wie das native Env. Es blockierte jedoch nicht den Eintritt von HML-2 in die Zellen. Mit diesem Fusionsprotein wurde das Golgi-Membranprotein 37 (GP73) als potenzieller Bindungsfaktor für HML-2 Env gefunden. Der zweite Ansatz umfasste die Verwendung von HML-2 Env-Pseudoviren bei der Durchführung eines funktionellen Screenings einer cDNA-Bibliothek mit einer nicht-permissiven Zelllinie. Damit wurden 6 Transmembranproteine (TMEM9, C12ORF59, IL1RAP, PSCA, LETMD1 und MPEG1) identifiziert, die am Infektionsprozess von HML-2 beteiligt sind. Nur IL1RAP und MPEG1 konnten allerding den Eintritt in eine andere nicht-permissiven Zelllinie vermitteln. Die Arbeiten bestätigen die Rolle von HS bei der Zelladhärenz von HML-2 Env und identifizierten die spezifische HBD. Sie bestätigten auch die Beteiligung eines oder mehrerer anderer Moleküle, die als Rezeptor(en) dienen. In diesem Zusammenhang wurden 7 Transmembranproteine identifiziert, die am Eintrittsweg von HML-2 beteiligt sind.

1. Introduction

1.1. Retroviruses

Retroviruses constitute a large family of diverse enveloped RNA viruses. They share standard structural, compositional, and replicative features (Coffin, 1992a, 1992b, 1996; Coffin *et al.*, 1997). The virions are 80-100 nm in diameter, with an outer membrane displaying viral spikes of glycoproteins. The location and shape of the viral protein core are distinctive characteristics of each genus of this family. The viral genome is single-stranded RNA (ssRNA), which is linear and non-segmented, with 7-12 kb in size. Although ssRNA has a positive polarity, the virion RNA is reverse transcribed into linear double stranded DNA (dsDNA) throughout the replication cycle and then integrated into the vertebrate host genome. This replication strategy is a distinctive characteristic of the family, which is where the family's name comes from (Chaitanya, 2019; Coffin *et al.*, 1997; Hayward, 2017).

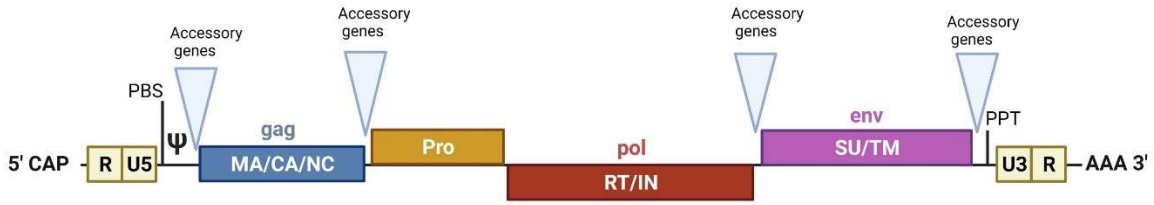

Figure 1-1 Retroviral genome structure.
The viral ssRNA shows the four essential genes; gag encodes MA and CA; pro encodes the PR; pol encodes RT and IN; The env encodes SU and TM. Complex retroviruses encode in their genome one or more additional accessory genes. The arrows indicate the possible locations of the accessory genes. Other essential viral replication sequences, including U3, R, and U5, are distinct regions from the non-coding ends of the extracellular RNA genome; (PBS) the primer-binding site; (ψ) the packaging sequences; (PPT) polypurine tract; (AAA) poly(A) tail (Gifford & Tristem, 2003). Created with BioRender.com based on (Gifford & Tristem, 2003)

The genome of retroviruses comprises two copies of ssRNA, each with at least four genes (Figure 1-1)**.** These genes are in the order 5' to 3', group-specific antigen (*gag*) that encodes the matrix (MA), capsid (CA), and nucleocapsid (NC) proteins, *pro* encodes the protease (PR), *pol* encodes the reverse transcriptase (RT) in addition to integrase (IN) and finally the *env* that encodes the envelope glycoprotein (Env) (Bannert & Kurth, 2006). However, some retroviruses also carry additional genes besides these basic genes. Based on the genome organization, complex and simple retroviruses can be distinguished (Coffin, 1992b; Murphy *et al.*, 2012). The simple retroviruses carry the previously mentioned genes, whereas, in the complex

retroviruses, additional regulatory proteins are produced through multiple splicing (Coffin, 1996) Figure 1-1.

The International Committee on Taxonomy of Viruses (ICTV) puts the family *Retroviridae* under order *Ortervirales*, class *Revtraviricetes*, phylum *Artverviricota*, and Kingdom *Pararnavirae* in the Realm *Riboviria* (ICTV, 2021). The family comprises two subfamilies *Orthoretrovirinae*, including six genera, *lentivirus,* and *Alpha-*, *Beta-*, *Delta-*, *Epsilon-*, and *Gamma-retroviruses*. And the second subfamily is *Spumaretrovirinae* consisting of five genera *Bovi-, Equi-, Feli-, Prosimii-* and *Simii-spumaviruses* (ICTV, 2021; Lefkowitz *et al.*, 2018; Walker *et al.*, 2020).

When retroviruses infect somatic cells, they insert their viral DNA genome into the chromosomal DNA of those cells. As a result, all offspring from infected cells will have identical insertions in their genomic DNA. However, when germline cells are infected, the integrated viral DNA (provirus) is passed down to the offspring, who will carry the provirus in every cell of the host's body. By that, the descendant will inherit it according to Mendel's laws. This kind of vertical transmission of the endogenized proviruses is the unique feature of the so-called endogenous retroviruses (ERVs). In contrast to the exogenous retroviruses (XRVs) that are transmitted horizontally among hosts (Bannert & Kurth, 2006). Some retroviruses, like Mouse mammary tumor virus (MMTV) and Jaagsiekte sheep retrovirus (JSRV), are found in both forms and can be transmitted in both ways (Löwer *et al.*, 1996).

While the exogenous retroviruses are classified as mentioned above, there is no classification system for endogenous retroviruses covered by ICTV (Hayward *et al.*, 2015; Jern *et al.*, 2005).

1.2. Endogenous retroviruses

ERVs are remnants of the ancient XRVs that infected the host germline and integrated into its genome at some point in the history of an organism before being passed to the descendants in a normal Mendelian fashion (Bannert & Kurth, 2004; Flint *et al.*, 2015). Their first discovery was back in 1960 in the genomes of birds and mammals, Years before Howard Temin and David Baltimore discovered the reverse transcriptase and even before retroviruses got their name. The first discovered ERVs were the avian leukosis virus, murine leukemia virus, and MMTV (Aaronson *et al.*, 1971; Bentvelzen *et al.*, 1970; Weiss, 1969). Researchers later discovered ERVs in the genomes of humans and other mammals (Martin *et al.*, 1981). Now, it is well established that ERVs exist in the genomes of various vertebrates, including fish, reptiles, and amphibians (Herniou *et al.*, 1998; Xu *et al.*, 2018). Their integration in multiple

loci on different chromosomes results from the horizontal and vertical transmission accompanied by proliferation capability. Most loci are nonfunctional due to accumulated mutations. Some, however, still have functional domains. Consequently, host cells still express their transcripts (Bannert & Kurth, 2004, 2006).

Endogenized retroelements account for over half of the human and mammalian genomes (Deininger & Batzer, 2002; Van de Lagemaat *et al.*, 2003). The integrated elements in the human genome make up the Human ERVs (HERVs). This name, however, does not reflect their host specificity but rather the genome host in which the ERV sequence is found. So, HERV refers to the endogenous retroviruses discovered in the human genome. They are, however, integrated into the genomes of Old-World monkeys and apes, as well as other mammalian species (Hayward *et al.*, 2015; Jern & Coffin, 2008; Mariani-Costantini *et al.*, 1989). The classification of HERVs into families depends on the complementary proviral primer binding site (PBS) sequence, where the cellular tRNA is bound to initiate the reverse transcription (Bannert & Kurth, 2006; Cohen & Larsson, 1988; Larsson *et al.*, 1989; Peters & Glover, 1980). However, this classification is inaccurate since proviruses belonging to one phylogenic linage may have different PBS sequences (Gifford & Tristem, 2003). On the other hand, the phylogenetic analyses have revealed relative relations between ERVs and present XRVs. Based on this evolutionary relatedness, ERVs are classified into three classes. Class I comprises the related ERVs to *Gamma-* and *Epsilon-retrovirus* genera, while those related to the *Alpha-, Beta-, and Delta-retrovirus* genera are in class II, and elements similar to *Spumaviruses* are in class III (Bannert & Kurth, 2006; Gifford *et al.*, 2005; Johnson, 2015). It has been long thought that there were no ERVs closely related to lentiviruses until the first rabbit endogenous lentivirus type K (RELIK) was discovered in the European rabbit's genome (Katzourakis *et al.*, 2007), followed by the discovery of the prosimian immunodeficiency virus in the genome of the gray mouse lemur (pSIVgml) (Gifford *et al.*, 2008). Class II, however, comprises the lenti-related ERVs (Gifford *et al.*, 2005; Johnson, 2015).

1.3. Human endogenous retroviruses (HERVs)

Retroelements account for 42 % of the human genome (Deininger & Batzer, 2002; Van de Lagemaat *et al.*, 2003). Although, they are not all complete and entirely intact. The genetic structures and sequences of these elements have been severely damaged over time by mutations, deletions, insertions, or even homologous recombination between the two flanking long terminal repeats (LTR), leading to total excision (Benachenhou *et al.*, 2009; Löwer *et al.*,

1996; Stoye, 2012). There are two types of retrotransposon elements, non-LTR and LTR elements. The long-terminal and short interspersed elements (LINE) and (SINE), respectively, are non-LTR elements (Medstrand *et al.*, 2002). LTR-elements, on the other hand, account for 8% of the human genome and include HERVs and the retrotransposons that are identical to HERVs but lack the *env* gene (Bannert & Kurth, 2004; Löwer *et al.*, 1996).

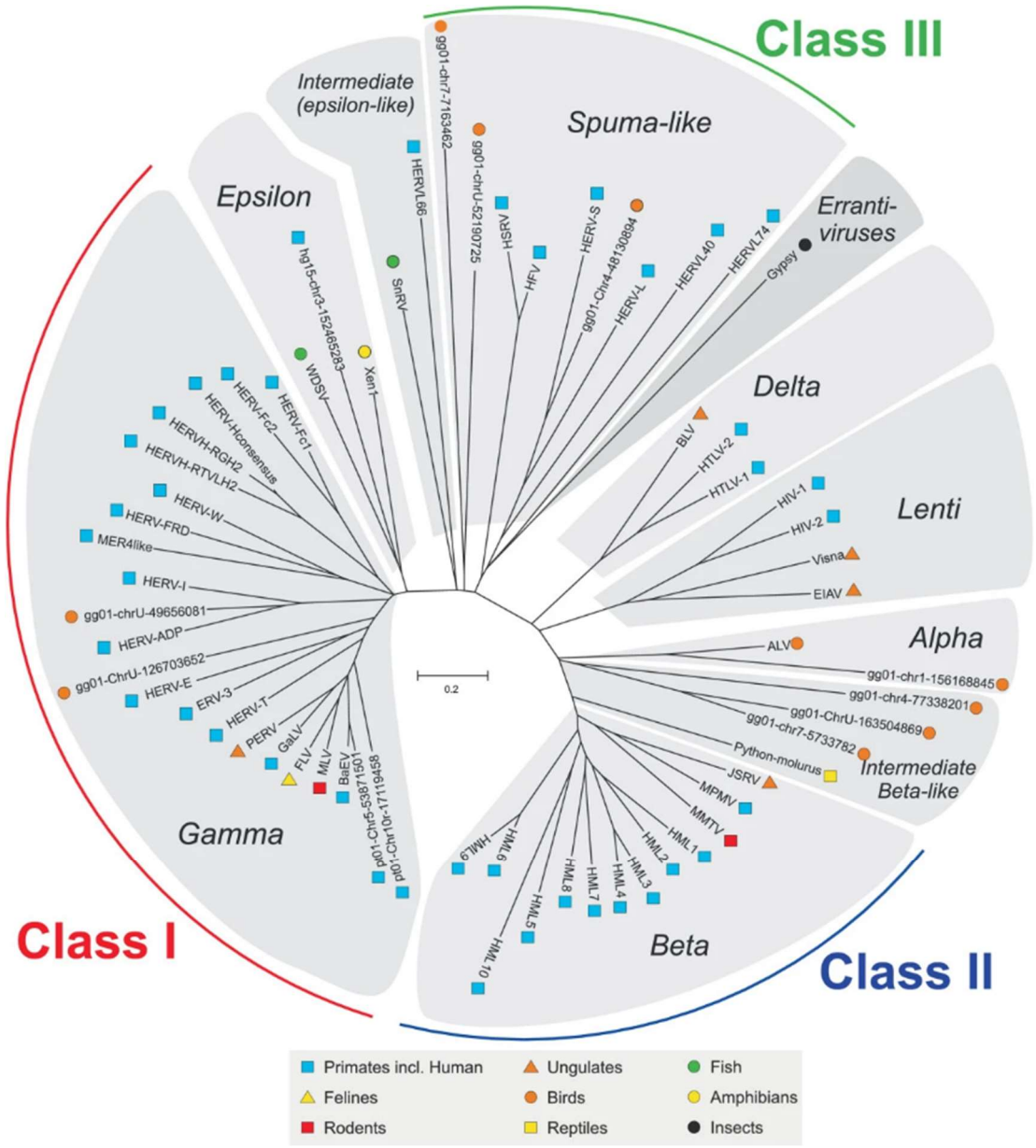

Figure 1-2 Unrooted Dendrogram of different classes of HERVs and their relations to the XRVs genera, based on phylogenetic neighbor-joining of the pol region.
Types of ERVs are indicated in the periphery. Taken from (Jern et al., 2005).

HERVs are heterogeneous and distributed among the three classes of ERVs (Figure 1-2). Based on the PBS sequence, they are grouped into 50 families (Hayward *et al.*, 2015; Jern *et*

al., 2005). More than 500 elements in the human genome belong to the HERV-K group (Mager & Medstrand, 2003). The PBS sequence in these elements is complementary to lysine-tRNA. Hence the (K) letter is given in the name (Peters & Glover, 1980). The first discovered element of this group was reported in the early 1980s, with sequence similarity to MMTV (Callahan *et al.*, 1982). Since all elements of this group phylogenetically are beta-related retroviruses, they belong to class II (Hohn *et al.*, 2013; Jern *et al.*, 2005). In addition to the HERV-K name, proviruses resemble MMTV, known as human MMTV-like (HML) (Franklin *et al.*, 1988). Depending on the estimated integration time of these elements (Bannert & Kurth, 2006) and the sequence of the conserved RT region (Medstrand & Blomberg, 1993), they are grouped into ten families carrying the name HML and numbered (1-10) (Bannert & Kurth, 2006; Garcia-Montojo *et al.*, 2018; Mager & Medstrand, 2003; Vargiu *et al.*, 2016).

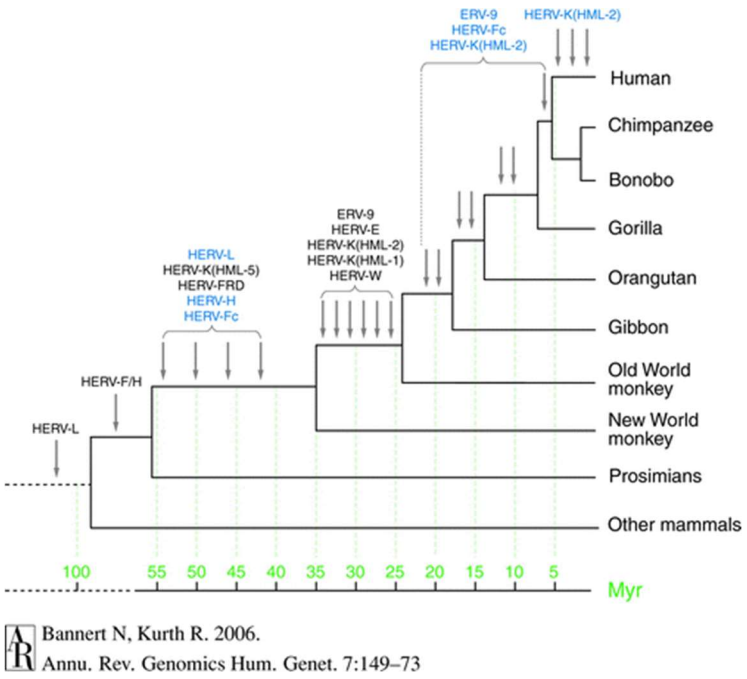

Figure 1-3 Phylogeny of primates indicates the estimated integration time for most HERV families. Taken from (Bannert & Kurth, 2006).

HERV-K (HML-2) family includes the youngest integrations among HERV subgroups, and it is the most characterized and studied family (Belshaw *et al.*, 2005; Subramanian *et al.*, 2011; Turner *et al.*, 2001). The elements of the family are well conserved, with more than 40 full-length proviruses (Macfarlane & Simmonds, 2004; Mager & Medstrand, 2003). Some of them are still active, as evidenced by their ability to generate intact viral particles (Boller *et al.*, 1993; Boller *et al.*, 2008). The HML-2 family is one of few subgroups of HERVs that has active transcript elements (Bronson *et al.*, 1978; Löwer *et al.*, 1984; Turner *et al.*, 2001). Functional

viral proteins from about 90 proviruses were reported (Subramanian *et al.*, 2011), particularly in tumors and cell lines derived from melanomas and teratocarcinoma (Löwer *et al.*, 1984; Schmitt *et al.*, 2013). The estimated time for the first integration of family elements is about 35 Myr ago and was into the germline of Old-World primates (Mager & Medstrand, 2003; Tönjes *et al.*, 1996). However, the recent integrations occurred less than 6 Myr ago when the lineages of humans and chimpanzees diverged (Figure 1-3), Resulting in human-specific HERVs with polymorphic insertions (Belshaw *et al.*, 2005), which makes HML-2 the only known family with human-specific integrations (Subramanian *et al.*, 2011).

HERV-K113 on chromosome 19p13.11 is one of these full-length proviruses with human-specific polymorphic insertions and the ability to produce intact non-infectious viral particles (Boller *et al.*, 2008; Turner *et al.*, 2001). Its distribution varies among human populations, with a frequency of 11.8% in Poland (Zwolińska *et al.*, 2013) and 27 % of all samples in the 1KGP database. However, its frequency was 52% of samples of African origin in the same project (Wildschutte *et al.*, 2016). It was also detected in 21.8% of breast cancer patients in Sub-Saharan Africa (Jha *et al.*, 2009). Depending on the LTR sequence, it was inserted in the genome 0.8 Myr ago, long before modern *Homo sapiens* appeared (Jha *et al.*, 2009).

1.4. The retroviral infection cycle

Adsorption of retroviral particles on the cell surface, like other enveloped viruses, is accomplished by interactions with the viral Env (Figure 1-4). The earliest contacts between the virus and the cell surface are electrostatic interactions with glycoproteins or glycolipids. They allow the virus to recruit specific interactions with the receptor/s activating the fusion process, ultimately resulting in the virus entry (Marsh & Helenius, 2006; Mothes & Uchil, 2010; Young, 2001). The Env mediates both binding and fusion processes. Two subunits make up the Env, the surface unit (SU), which faces the outside, and the transmembrane (TM), the anchored protein to the viral membrane. They combine to form heterotrimers (Hunter, 1997; Overbaugh *et al.*, 2001). SU protein interacts with cell surface receptors through one or more domains (Mothes & Uchil, 2010; Overbaugh *et al.*, 2001). The viral receptor can be one molecule or multiple molecules. An example of a virus using multiple receptors is the human T-cell lymphotropic virus type 1 (HTLV-1). Three cell surface proteins function as its receptors the glucose transporter 1 (GLUT1), neuropilin-1 (NRP-1), and heparan sulfate proteoglycans (HSPG) (Ghez *et al.*, 2006; Jones *et al.*, 2005; Manel, Kim*, et al.*, 2003). The human immunodeficiency virus (HIV), on the other hand, engages two receptors, the cluster

of differentiation 4 (CD4) as a primer receptor (McDougal *et al.*, 1986), and either the C-C chemokine receptor type 5 (CCR5) or C-X-C chemokine receptor type 4 (CXCR4) as coreceptors (Bleul *et al.*, 1997; Feng *et al.*, 1996). Furthermore, Kameoka et al., 1993 and Arthos et al., 2008 have established that CD26 and integrin α4β7 can operate as cofactors for HIV (Arthos *et al.*, 2008; Kameoka *et al.*, 1993). In betaretroviruses, MMTV employs transferrin receptor 1 (TfR1) for cell entrance and HSPG as an attachment factor (Ross *et al.*, 2002; Zhang *et al.*, 2003). JSRV engages hyaluronidase 2 (Hyal2) as its entry receptor (Rai *et al.*, 2000). HML-2, on the other hand, recruits HSPG as an attachment factor (Robinson-McCarthy *et al.*, 2018). Its entry receptor, however, has yet to be discovered.

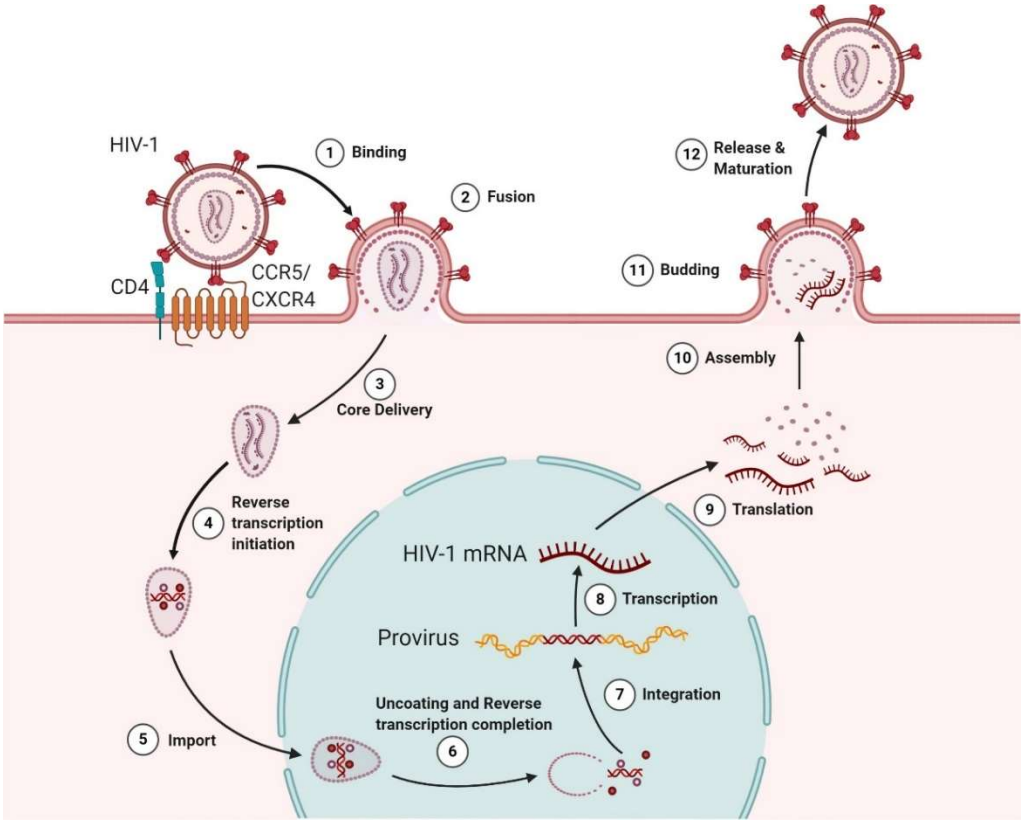

Figure 1-4 HIV as a model for retroviruses infection cycle.

The infection begins by binding the CD4 receptor and the coreceptors (CXCR4/CCR5)(Step 1), allowing the viral particle to enter and fuse with the target cell (step 2). Following core delivery and initiation of the uncoating process (step 3). The initiation of reverse transcription (step 4) before importing the PICs into the nucleus (step 5), where the reverse transcription is completed with the uncoating (step 6), followed by the integration of the viral DNA into the host chromosome (step 7). Transcription of the provirus (step 8) and transporting the viral RNAs to the cytoplasm for protein synthesis (step 9). Packaging the viral proteins and genomic RNA into virions (step 10. Virus budding from the cell (step11). The released virus then sprouted (step 12) (Ramdas et al., 2020). The figure is taken from (Ramdas et al., 2020)

The ectodomain of TM contains the fusion peptide. It brings the two membranes (the viral and the cellular) together, consequently allowing the fusion process (Mothes & Uchil, 2010; Overbaugh *et al.*, 2001). Fusion is not an energetically favorable process (Pecheur *et al.*, 1999). The virus can overcome unfavorable barriers by binding the specific receptor. Because of that, only certain cell surface proteins can function as viral receptors (Overbaugh *et al.*, 2001).

After engaging the specific receptor, for most retroviruses, neutral pH is sufficient for the fusion to occur at the plasma membrane (Overbaugh *et al.*, 2001). However, the fusion in gammaretroviruses and some betaretroviruses require a brief exposure to low pH to trigger the thermodynamically stable conformation. MMTV, JSRV, and HML-2 utilize this fusion process, which usually correlates with viral internalization (Bertrand *et al.*, 2008; Flint *et al.*, 2015; Mothes & Uchil, 2010; Redmond *et al.*, 1984; Robinson & Whelan, 2016). The acidic pH is maintained inside the endosome through the endocytosis uptake pathway, and viruses employ it in addition to other features to set off fusion (Mothes & Uchil, 2010).

The fusion process results in releasing of the viral core into the cytoplasm. The uncoating process directly follows it. The packed RT is used to reverse-transcribe the genomic RNA into DNA. This process occurs during the progression of uncoating in the subviral particles called reverse-transcription complexes (RTCs) (Desfarges & Ciuffi, 2012; Zhang *et al.*, 2000). The resulting DNA is linear, double-stranded, and ready for integration. It forms with IN and other subviral proteins, the pre-integration complex (PIC) (Bukrinsky *et al.*, 1993; Coffin *et al.*, 1997). The PIC then traffics to the nucleus, the replication site for retroviruses, by hijacking the intracellular trafficking machinery (Flint *et al.*, 2015; Ploubidou & Way, 2001). In most retroviruses, PIC is unable to pass through the nuclear membrane. It can reach the nucleus only during mitosis (Desfarges & Ciuffi, 2012; Lewis & Emerman, 1994; Roe *et al.*, 1993). However, some retroviruses like MMTV and HIV-1 can infect non-dividing and dividing cells (Ho *et al.*, 1986; Konstantoulas & Indik, 2014; Sherman & Greene, 2002). After reaching the nucleus, integration of the viral DNA ends, catalyzed by the IN, takes place into specific sites of the host genome and can be in many locations. The selection of the integration site is considered a distinctive characteristic of each retrovirus (Desfarges & Ciuffi, 2012; Engelman, 2010; Flint *et al.*, 2015). There are different preferences in the retroviral integration sites (Desfarges & Ciuffi, 2012; Schröder *et al.*, 2002). Promoters and transcription start sites of active genes are preferred integration sites of *Gamma-* and *Spuma-retroviruses* as well as HERVs (Brady *et al.*, 2009; Mitchell *et al.*, 2004). *Alpha-* and *Delta-retroviruses* integrate into transcription units and CpG islands (Derse *et al.*, 2007; Mitchell *et al.*, 2004). For *Lentiviruses*,

however, the integration preference is into active genes and the transcription unit with specific chromosomal features (Mitchell *et al.*, 2004; Schröder *et al.*, 2002). MMTV is the only virus with no preferences. It tends to integrate randomly into the host genome (Faschinger *et al.*, 2008).

The integration can cause disruption of host genes in the insertion site or even alter the activation of some genes influenced by the 3' LTR promoter. On the other hand, the chromosomal environment of the integration site and cellular status influence the proviral transcriptional activity. As a result, the provirus will either be active or latent (inactive) (Burnett *et al.*, 2009; Dar *et al.*, 2012; Singh *et al.*, 2010). The proviral DNA relies on the cellular gene-expression machinery after integration. It is transcribed as a cellular gene and controlled by cis-acting elements at the LTR's U3 region (Flint *et al.*, 2015; Rabson & Graves, 2011).

The viral full-length RNA product can function as both genomic and messenger RNA. It can be packaged into virions or translated into proteins after being processed by the cellular RNA processing machinery. In complex retroviruses, the expression regulatory proteins such as Rev (in HIV), Rex (in HTLV), and Rec (in HML-2) regulate the relative amount of unspliced or partially spliced RNA by binding them and facilitating their transport out of the nucleus to the cytoplasm (Ahmed *et al.*, 1990; Löwer *et al.*, 1995; Rabson & Graves, 2011). The alternative splicing generates multiple transcripts translated later into the viral proteins. Gag polyproteins, however, are translated on the same viral unspliced mRNA utilizing the ribosomal frameshift (Flint *et al.*, 2015; Rabson & Graves, 2011), or stop-codon readthrough (Csibra *et al.*, 2014).

The synthesis of viral proteins takes place in the cytoplasm. At the same time, the Env precursor is synthesized in the endoplasmic reticulum (ER). It is then transported to the Golgi apparatus and processed by cellular Furin protease to form the mature Env. The latter is transported to the plasma membrane, the assembly location for lentiviruses and HML-2 (Coffin *et al.*, 1997; Flint *et al.*, 2015; Lee & Bieniasz, 2007; Ono, 2010). It forms with the other components (two copies of ssRNA, the Gag and Gag-Pol precursors) the immature virions that turn to mature status by the PR activity that cleaves the Gag precursors at specific sites, causing the condensed core appearance in mature viruses (Flint *et al.*, 2015; Göttlinger & Weissenhorn, 2010; Lee & Bieniasz, 2007).

1.5. HML-2 family and K113

1.5.1 Genomic structure and morphology

The structure of K113, a member of the HML-2 family, is similar to that of betaretroviruses (Jern *et al.*, 2005). The enveloped viral particles are spherical with a diameter of about 110 nm (Figure 1-5; A) (Morozov & Morozov, 2021; Tönjes *et al.*, 1997). The viral membrane is a cell-derived lipid bilayer (Fenner *et al.*, 1987) that displays the viral envelope glycoproteins as heterotrimers of SU and TM subunits that are non-covalently associated (Henzy & Coffin, 2013). The MA, the first layer of the Gag, is right beneath the lipid membrane and bound to it by myristic acids (Coffin *et al.*, 1997). Underneath MA, the icosahedral capsid is located (Acton *et al.*, 2019), which represents the outer layer of the core. The NC and two copies of (+) ssRNA, as well as a few molecules of RT and IN, are all found within the core (Coffin *et al.*, 1997) (Figure 1-5; B).

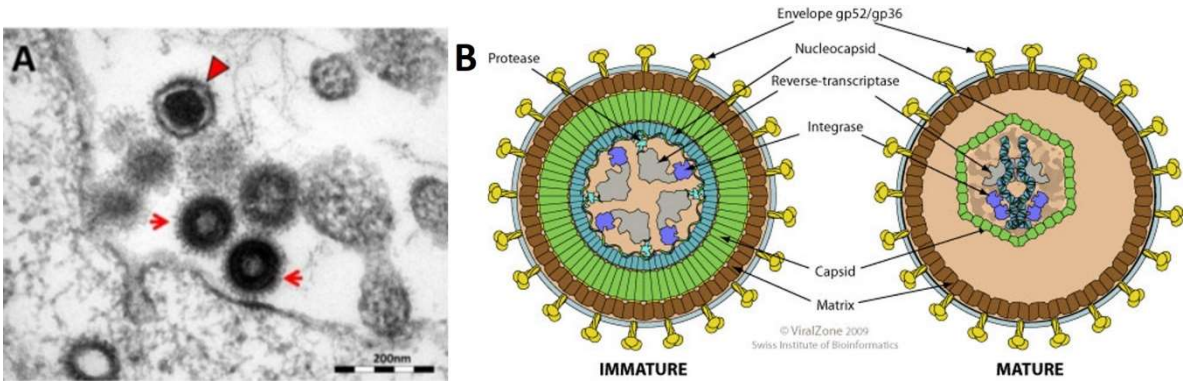

Figure 1-5 Structure of HML-2 viral particles

***A**. The mature and immature K113 viral particles budding from HEK293T cells, adapted from (Hohn et al., 2014). Arrows indicate the immature particles, and the arrowhead indicates the mature particle with the condensed core. **B**. schematic structure of Betaretrovirus genus's mature and immature viral particles, adapted from (ViralZone (RRID:SCR_006563)).*

The four ORF *gag*, *pro*, *pol*, and *env*, are encoded by the 9.5 kb provirus (Figure 1-6). K113 has an additional encoded accessory protein Rec, making it a type II HML-2. Based on the accessory protein, two types of HML-2 are distinguished. Type II encodes Rec, and type I encodes Np9 instead of Rec due to a 292-bp deletion at the *pol-env* genes (Figure 1-6) (Armbruester *et al.*, 2002; Bannert & Kurth, 2004; Löwer *et al.*, 1995; Subramanian *et al.*, 2011).

Two identical LTRs flank the provirus at the 3' and 5' ends. Each has three regions, the unique 3' end (U3), the repeated region (R), and the unique 5' end (U5). Elements in U3 at the 5' LTR function as a promoter and facilitate the binding of RNA polymerase II (Pol II) to the

transcription starting site at the R region. The PBS is directly located after the U5, in the untranslated region, where the tRNA- Lys binds and primes the reverse transcription (Bannert *et al.*, 2010). The polypurine tract (PPT), which primes the synthesis of the plus DNA strand,

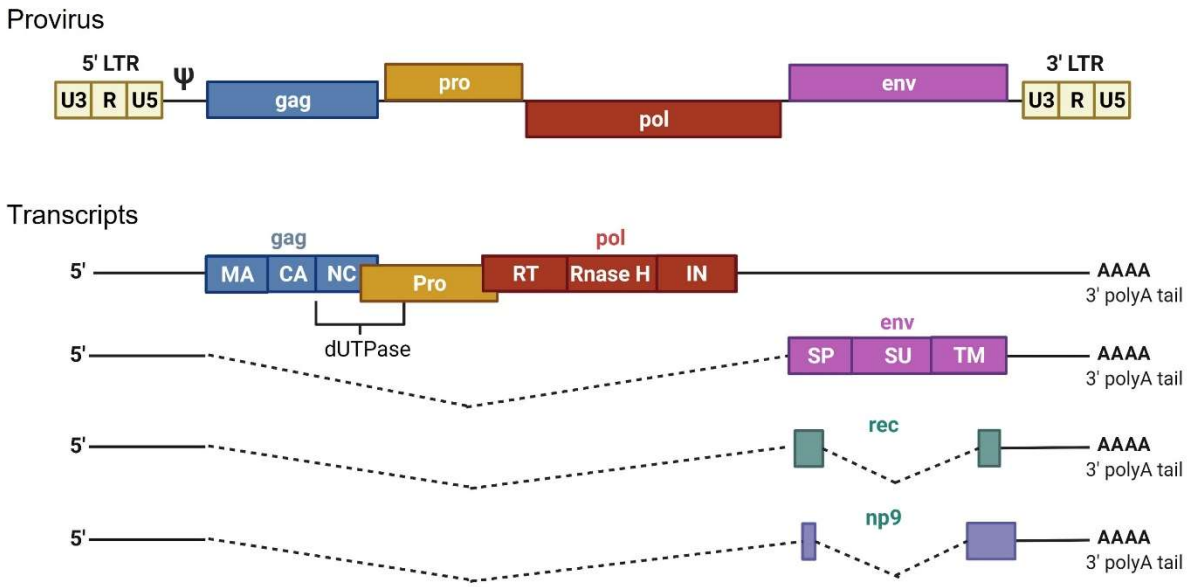

Figure 1-6 Organization and transcripts of HML-2 provirus.
The four ORF sequences are colored in the proviral form of HML-2. The R segment in LTRs separates the U3 and U5 regions. The first transcript has three ORFs that code for Gag, Pro, and Pol proteins. Only the gag has a start codon in this transcript; pro and pol translation is mediated by two ribosomal frameshifts (–1). Both the gag and pro reading frames encoded UTPase (Barabás et al., 2003; Hizi & Herzig, 2015; Köppe et al., 1994). The Env is encoded by the second transcript, which contains three domains signal peptide (SP), surface (SU), and transmembrane (TM). The third transcript encodes Rec generated by type 2 HML-2 proviruses and results from alternate splicing of the env ORF. Type 1 proviruses produce transcript 4 (np9) that results from a 292 bp deletion in the pol–env junction (Bannert & Kurth, 2004; Garcia-Montojo et al., 2018; Hohn et al., 2013). Created with BioRender.com based on (Garcia-Montojo et al., 2018; Hohn et al., 2013).

is located downstream before U3 at the opposite LTR (Engelman, 2010). The *gag* encodes 74 kDa precursor protein that is processed into the p15-MA, the spacer peptide (SP1), p15, the p27-CA, the p10-NC, and other two small proteins (George *et al.*, 2011). Furthermore, Gag-Pro and Gag-Pro-Pol polyproteins result from two slippery sites and single base ribosomal frame-shifting. They are processed into PR, RT, and IN (Coffin *et al.*, 1997; Hohn *et al.*, 2013). Additionally, dUTPase, the feature for betaretroviruses and non-primate lentiviruses, is encoded near the N-terminus of the *pro* (Figure 1-6). The enzyme degrades dUTPs and prevents their incorporation into the newly synthesized proviral DNA (Hizi & Herzig, 2015; Jern *et al.*, 2005; Mayer & Meese, 2003). The *env* gene encodes a classical envelope of retroviruses. The

protein is produced first as a precursor (gp95), which is cleaved later into the SU~ 44 kDa. And the TM ~26 kDa proteins (Hanke *et al.*, 2009).

1.5.2 The envelope protein

K113 Env is translated from a single spliced mRNA (Figure 1-6), as previously stated. It shows the typical structures of retroviral Envs (Figure 1-7). The 11 kDa signal peptide (SP) encoded by the long leader sequence guides the Env to the ER, the initial step in the secretory pathway. Here, the entire polypeptide chain is translocated into the lumen. After that, the SP is cleaved from the rest of the chain (Hanke *et al.*, 2009; Hunter & Swanstrom, 1990; Ruggieri *et al.*, 2009). In the lumen, a co-translationally N-linked glycosylation process occurs utilizing

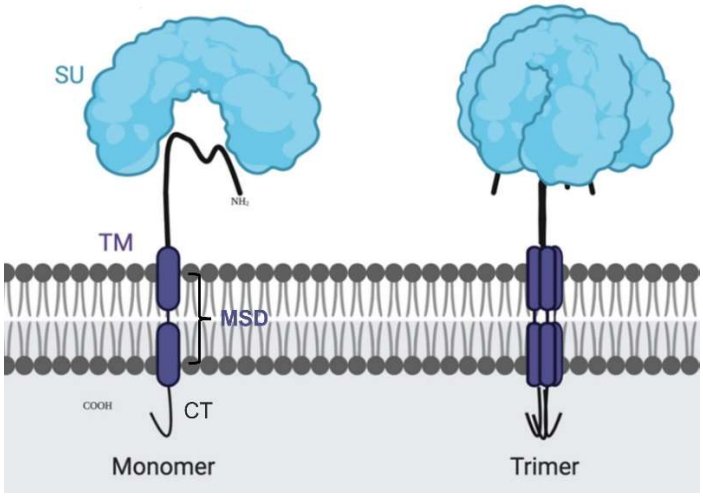

Figure 1-7 Schematic structure of the retroviral Env.

The viral Env is cleaved into the surface glycoprotein (SU) and the transmembrane protein (TM) during transit to the cell membrane by a cellular furin protease. Then, they are assembled via non-covalent bonds. TM has extracellular and intracellular domains. It is attached to the membrane via a membrane-spanning domain (MSD). The cytoplasmic tail refers to the C-terminal portion of the protein within the cell. The folded monomer (left) is oligomerized in the secretory pathway to form the mature heterotrimer (right) (Gallaher et al., 1989). The figure is modified from (Vergara Bermejo et al., 2020).

cellular oligosaccharyltransferase enzymes that attach mannose-rich oligosaccharides to the asparagine residues at the motifs N-X-S/T (X is any amino acid (aa) except proline) (Flint *et al.*, 2015; Hunter & Swanstrom, 1990). The glycosylation of the K113 Env is considered to be moderate compared to other retroviruses. It has nine N-linked glycosylation sites at the ratio SU/TM= 5/4, which is similar to other retroviruses (Hanke *et al.*, 2009). This modification step is crucial for proper protein folding, and the latter allows the protein to exit and travel further along the secretory pathway (Coffin *et al.*, 1997; Delwart & Panganiban, 1990).

Assembly of Envs into an oligomeric structure occurs in the ER. This process is essential for intracellular transport (Coffin *et al.*, 1997; Tsai *et al.*, 2002). The implicated factors in the oligomerization process are not fully understood (Hunter & Swanstrom, 1990). However, specific highly conserved areas of the Env may be involved in the oligomerization process, such as the leucine zipper-like motifs that form a coiled-coil region in the ectodomain of the TM proteins in many retroviruses (Coffin *et al.*, 1997; Delwart *et al.*, 1990; Lu *et al.*, 1995).

The adequately folded oligomeric Env precursor is then translocated into the Golgi apparatus. It enters through the cis-Golgi network and undergoes several specific reactions, resulting in the maturation of the oligosaccharide chains. The further step includes Env cleavage into SU and TM in the late Golgi compartment. This process is achieved by the resident furin protease and is accompanied by conformational changes that sequester the fusion peptide inside the oligomer. Such a step is crucial for the infectivity and the stability of the Env (Flint *et al.*, 2015; Hanke *et al.*, 2009; Henzy & Coffin, 2013; Hunter & Swanstrom, 1990).
Betaretroviruses members exhibit a noncovalent association between the two cleaved subunits, similar to the lentiviruses (Henzy & Coffin, 2013). This kind of non-covalent interaction between SU-TM is labile, causing shading off some SU proteins (Robey *et al.*, 1987). As a result, a nonfunctional Env may form (Yang *et al.*, 2000).

Following the previous two essential modifications, the mature Env complex travels in a transport vesicle that buds from the Trans-Golgi network (TGN) to the assembly site at the plasma membrane (Coffin *et al.*, 1997; Flint *et al.*, 2015; Lee & Bieniasz, 2007). The trafficking and the localization at the plasma membrane are both regulated by the cytoplasmic tail (CT), a part of the TM C-terminus that stays within the cytoplasm near the inner faces of the plasma membrane. It also plays a role in viral incorporation in a cell type-dependent requirement (Murakami & Freed, 2000; Muranyi *et al.*, 2013; Qi *et al.*, 2013). Though, it is not implicated in receptor binding or viral entry processes (Tedbury & Freed, 2015). CTs of retroviruses vary in their length. While lentiviruses possess long CTs (150 aa in HIV-1), other retroviruses' CTs comprise less than 50 aa (Mammano *et al.*, 1997; Schnierle *et al.*, 1997). In K113, however, the CT is relatively short, consisting of 44 aa (Dewannieux *et al.*, 2005). Finally, the assembled immature viral particles bud from the cells bearing the mature Env (Figure 1-7) (Vogt, 1996). Although the K113 *env* ORF is intact and translated into complete Env, the latter has post-insertional mutations that compromise its functionality and prevent its incorporation into viral particles (Beimforde *et al.*, 2008; Turner *et al.*, 2001).

1.5.3 Reconstruction of the functional Env of K113

According to a characterization study of a K113 allele in the BAC library, several nucleotides in the K113 sequence were altered during evolution. Env, reverse transcriptase, and Gag proteins were all affected. As a result, the Env was poorly expressed, partially processed, and failed to incorporate into the viral particles produced by this element. Consequently, the viral particles were replication-incompetent (Beimforde *et al.*, 2008; Turner *et al.*, 2001).

Hanke et al. 2009, showed that 40% of GC content and the rare codons contribute to the poor expression of the Env. Other factors such as the unstable RNA and the internal splicing that generates Rec protein at the expense of Env protein are also involved. She with her coworkers were able to boost the expression to more than 50-fold through a codon optimization process. The codon-optimized Env (coEnv) had nearly 70% of its codons synonymously exchanged,

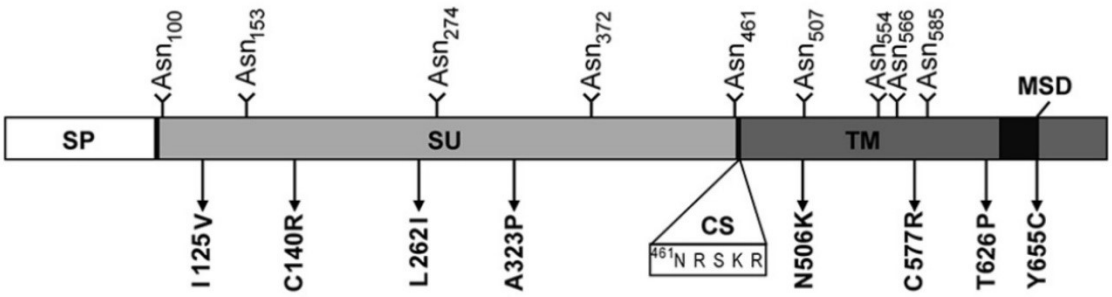

Figure 1-8 Schematic representation of the reconstituted K113 oricoEnv.
By comparing the sequences of 10 highly conserved HML-2 elements, eight post-insertional mutations (the arrows) within the Env amino acid sequence were identified and reversed to generate oricoEnv. Four amino acid substitutions are in the SU and four in the TM. Positions of the potential N- linked glycosylation sites are also indicated (above). The reversed mutation in N506K reduces the potential glycosylation sites to nine. CS: cleavage site (Hanke et al., 2009).The figure is adapted from (Hanke et al., 2009).

raising its GC content to 63%. Furthermore, eight post-insertional nonsynonymous mutations were identified and reversed (Figure 1-8) by aligning the Env sequence with other well-conserved envelope sequences belonging to elements from the HML-2 family. As a result, the processing efficiency of the Env precursor was improved. The reconstituted Env (oricoEnv) represents the Env of K113 at the timepoint of viral integration into the human genome (Hanke *et al.*, 2009).

In several retroviruses, truncation of the CT improves the fusogenic properties of the Env and its incorporation into viral particles (Celma *et al.*, 2007; Côté *et al.*, 2008). It is also true for

HML-2 Env since successful incorporation of the Env was achieved after truncation in the region 659-699 of the Env sequence, leaving only four aa of CT. The reconstituted, codon-optimized, and successfully incorporated Env was called (oricoEnvΔ659-699/ oricoEnvΔC1) (Hanke *et al.*, 2009).

1.5.4 HML-2 Env tropism and entry

The successful incorporation of the reconstituted and CT truncated Env (oricoEnvΔC1) into viral particles establishes the infection. Through pseudotyping, viral entry could be achieved. Moreover, since Env mediated entry implies the expression of the appropriate specific receptor, the tropism of HML-2 Env could be estimated (Hanke *et al.*, 2009; Kramer *et al.*, 2016).

Viral tropism is the capability of a virus to infect a specific group of host cells (Chappell and Dermody, 2015). It is determined by the expression of the viral entry receptor and the multiple post-entry factors that regulate the viral replication cycle in the infected cell (Chappell & Dermody, 2015). The expression of a functional entry receptor is a crucial determinant (Chappell & Dermody, 2015; Kramer *et al.*, 2016)

Kramer et al. 2016 demonstrated the capability of oricoEnvΔC1 to mediate viral entry in a broad range of cell types (16 types) belonging to ten species, including human, feline, canine, mouse, rat, porcine, hamster, guinea pig, simian, and rabbit (Kramer *et al.*, 2016). Other groups could achieve viral entry mediated by different reconstituted HML-2 Envs, such as HERV-K Phoenix (Dewannieux *et al.*, 2006) and HERV-K con (Lee & Bieniasz, 2007) in additional species, including monkeys, bats, chickens, and snakes, broadening the tropism of HML-2 Env to mammals, birds, and reptiles. Such a broad tropism suggests ubiquitously expressed entry receptors conserved among amniotes. It also may suggest that HML-2 Env can utilize multiple entry pathways (Dewannieux *et al.*, 2006; Garcia-Montojo *et al.*, 2018; Lee & Bieniasz, 2007; Robinson & Whelan, 2016). Besides the receptor, HML-2 entry requires endosomal acidification and a membrane scission mediated by dynamin. These factors can also influence HML-2 entry (Robinson & Whelan, 2016).

1.5.5 The receptor and receptor binding site

The interaction with the cellular receptor is the critical step in the viral infection cycle. It defines viral pathogenesis and regulates the host range and tissue tropism (Grove & Marsh, 2011; Heise, 2014). HML-2 is now known to possess a broad viral tropism. Though, its entry

receptor itself has not been identified yet (Dewannieux *et al.*, 2006; Kramer *et al.*, 2016; Robinson & Whelan, 2016).

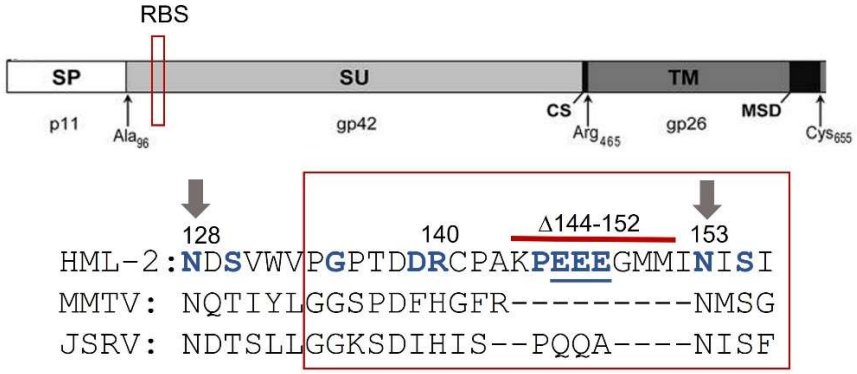

Figure 1-9 RBS of HML-2.
Up: Schematic representation of oricoEnvΔC1 indicates the position of RBS (the red square)
Down: part of amino acid sequence alignment of K113, MMTV, and JSRV at the region of RBS.
RBS mutational analysis included eight single-site mutations (blue) P145A, N128A, N153A, S130A, S155A, G135A, D138A, and R140A, Glutamic acids stretch mutation EEE146-148 AAA (underlined), and Δ144-152 deletion (residues indicated by the red line above). Arrows indicate the glycosylation sites. The figure is modified from the source (Bannert, 2017).

Within most animal virus families, there is minimal to no difference in receptor utilization across closely related viruses (Coffin *et al.*, 1997). Xenotropic and Polytropic murine leukemia viruses (X- and P-MLVs), for example, both use the same receptor XPR1 (Xenotropic and Polytropic retrovirus Receptor 1) (Battini *et al.*, 1999; Overbaugh *et al.*, 2001; Tailor, Nouri, Lee*, et al.*, 1999; Yang *et al.*, 1999). The betaretroviruses Mason-Pfizer monkey virus (MPMV) and simian type D retroviruses (SRVs) (types 1, 2, 4, and 5) share with the gamma-related HERV-W the same receptor SLC1A5 (Solute Carrier Family 1 Member 5) (Blond *et al.*, 2000; Tailor, Nouri, Zhao*, et al.*, 1999). The Envs for these viruses are relatively similar and form a single phylogenetic cluster (Overbaugh *et al.*, 2001). However, this cannot be applied to HML-2 and its closely related MMTV and JSRV; since neither the cell expressing TfR1 nor Hyal2 allows viral entry mediated by HML-2 Env (Kramer *et al.*, 2016). Furthermore, it has been clearly shown that HML-2 Env does not utilize the receptors of the other functional envelopes of HERV-W and HERV-FRD (Geppert, 2019).

The cell surface receptor interacts with the viral Env through one or more specific domains found on the SU subunit, which serves as the key to "unlock" the cell (Marsh & Helenius, 2006; Overbaugh *et al.*, 2001; Pecheur *et al.*, 1999). Hanke and coworkers 2009 discovered that single-site mutation arginine-to-cysteine R140C in the SU blocked viral entry completely.

Interestingly, this mutation is one of the eight reversed substitutions in the consensus K113 Env sequence that renders the nonfunctional Env (Hanke *et al.*, 2009). R_{140} was later identified as a receptor-binding site (RBS) component, a crucial domain for receptor binding (Bannert, 2017).

Mapping the RBS of the HML-2 SU was done through sequence alignment of the reconstituted oricoEnvΔC1 with the closely related MMTV and JSRV. The identified RBS of HML-2 corresponded to those of both viruses and extended between N_{128} and N_{153} residues at the N-terminus of SU (Figure 1-9) (Bannert, 2017; Wamara, 2020). Residues belonging to the domain were tested for involvement in the viral entry by mutational analysis, including eight single-site mutations P145A, N128A, N153A, S130A, S155A, G135A, D138A, and R140A. In addition to a triple mutation in three Glutamic acids stretch (EEE146-148AAA) and a deletion Δ144-152. Most envelopes bearing these mutations were successfully incorporated into viral particles, although they failed to mediate efficient viral entry (Bannert, 2017; Priesnitz, 2019; Wamara, 2020).

1.5.6 Glycosaminoglycans and HML-2 Env

The glycosaminoglycans (GAGs) are complex linear polysaccharides ubiquitously expressed in all cells. They can be found in the plasma membrane, the intracellular compartments, and the extracellular matrix (Aquino & Park, 2016). They are composed of repeated disaccharide units. Each unit is made up of an amino sugar, usually acetylated (N-acetyl-d-glucosamine (GlcNAc) or N-acetyl-d-galactosamine (GalNAc)), bound to a uronic acid (d-glucuronic (GlcA) or l-iduronic acid (IdoA)) or d-galactose (Gal) (Esko, 1999; Stick & Williams, 2009). The disaccharide unit and the glycosidic bond are unique for each type (Jinno & Park, 2015). There are five types of GAGs, including chondroitin sulfate $(GlcA\beta1\text{-}3GalNAc\beta1\text{-}4)_n$, dermatan sulfate (DS) $(GlcA/IdoA\beta1\text{-}3GalNAc\beta1\text{-}4)_n$, keratan sulfate (KS) $(Gal\beta1\text{-}4GlcNAc\beta1\text{-}3)_n$, hyaluronic acid (HA) $(GlcA\beta1\text{-}3GlcNAc\beta1\text{-}4)_n$, and heparin/heparan sulfate (HP/HS) $(GlcA/IdoA\beta1\text{-}4GlcNAc\alpha1\text{-}4)_n$ (Aquino & Park, 2016; Esko, 1999). Because of the carboxyl group of the uronic acid and sulfation of the sugar units, GAGs are highly negatively charged (Frevert & Wight, 2006).

The biosynthetic processes of GAG vary depending on the cell type and involve numerous biosynthetic enzymes. However, the synthetic pathways do not always go to completion. As a result, diverse structures of GAGs, including the length and the extent of modification, are found in the cell (Aquino & Park, 2016; Jinno & Park, 2015).

HP and HS are structurally heterogeneous (Salmivirta *et al.*, 1996; Sugahara & Kitagawa, 2002). There are 23 different disaccharide units identified in the structure of HP/HS. They result from different epimerization and modification processes (Esko & Selleck, 2002). HS chains are long and have an average molecular weight of ~30 kDa. In contrast, HP chains are short and have an average molecular weight of ~15 kDa. The major epimer of the uronic acid in HS is d-glucuronic acid. However, L-iduronic acid predominates in HP (Lindahl *et al.*, 1998; Shriver *et al.*, 2012). Also, HS is less sulfated than HP; it has about one sulfate group per unit, whereas HP has about 2.7 sulfate groups per unit, making it the highest negatively charged macromolecule (Shriver *et al.*, 2012; Toida *et al.*, 1997). Furthermore, HP is primarily intracellular, while HS is common at the cell surface and within the extracellular matrix (Bishop *et al.*, 2007; Esko & Selleck, 2002).

All GAGs, except HA, exist as proteoglycans, in which GAGs are covalently attached to a specific Ser residue (except KS) in a Ser-Gly repeated sequence of a core protein (Aquino & Park, 2016; Jinno & Park, 2015; Kjellén & Lindahl, 1991). Syndecans and Glypicans are two classes of HS-proteoglycans (HSPG) that are found at the cell surface (Cagno *et al.*, 2019; Nugent *et al.*, 2013). In the four types of Syndecans, HS and sometimes chondroitin sulfate chains are attached to a single transmembrane core protein (Alexopoulou *et al.*, 2007). On the other hand, Glypicans (1 through 6) are GPI-anchored HSPG, in which the protein core is anchored to the plasma membrane via a glycosylphosphatidylinositol anchor (Saunders *et al.*, 1997).

GAGs play vital roles in cell signaling, which includes regulation of cell growth and proliferation, anticoagulation, wound repair, and promotion of the cellular adhesion (Casale & Crane, 2019; Raman *et al.*, 2005). Additionally, GAGs (particularly HS) may impact pathogenicity through interaction with many pathogens, like viruses, bacteria, parasites, and fungi (Bartlett & Park, 2010; Rostand & Esko, 1997; Spillmann, 2001).

Viruses initially contact the cell surface through low-affinity interactions, primarily electrostatic (Van der Waals forces), with GAGs. However, these low affinity and high avidity interactions do not imply a lack of specificity. The high-affinity and more specific interactions with secondary or tertiary receptors occur after the low-affinity binding events (Maginnis, 2018; Stencel-Baerenwald *et al.*, 2014; Yamauchi & Helenius, 2013).

Viruses increase their density at the cell surface through the low-affinity interactions with the ever-presence negatively charged GAGs. Subsequently, their interaction with the specific

entry receptors are facilitated (Aquino & Park, 2016; Cagno *et al.*, 2019; Rusnati *et al.*, 2009). Lots of virus families are known to bind HS, including the flaviviruses hepatitis B and C (HBV and HCV) (Leistner *et al.*, 2008; Shi *et al.*, 2013), the human papillomavirus (HPV) (Johnson *et al.*, 2009), the human herpes simplex virus 1 (HSV-1) (Shukla *et al.*, 1999), the retroviruses HIV (Saphire *et al.*, 2001), HTLV (Jones *et al.*, 2005), HML-2 (Geppert, 2019; Robinson-McCarthy *et al.*, 2018), and MMTV (Zhang *et al.*, 2003). In addition to the most recent coronavirus severe acute respiratory syndrome virus (SARS-CoV-2) (Clausen *et al.*, 2020). Binding the ligands to the HSPG results in tethering on the cell surface, sequestration and protection, oligomerization, and allosteric activation. It can also induce internalization since HSPGs can act as endocytic receptors (Cagno *et al.*, 2019; Gómez Toledo *et al.*, 2021).

HIV binds the Syndecans of all types (1, 2, 3, and 4) (Saphire *et al.*, 2001), although, Syndecans -2 and -3 are the major attachment receptors for HIV. The implication of HSPG has been found to prolong and increase HIV infectivity in dendritic cells. It also promotes HIV transmission to CD4+ T cells (Bobardt *et al.*, 2003; Cameron *et al.*, 1992; De Witte *et al.*, 2007). Furthermore, HIV Tat (Trans-Activator of Transcription), a regulatory protein released from the infected cells, binds HSPG on different cells, including endothelial cells, lymphocytes, monocytes, macrophages, and neurons (Helland *et al.*, 1991). This binding depends on HS polysaccharide chain size and sulfation degree (Rusnati *et al.*, 1997; Rusnati *et al.*, 1999), and it is needed to internalize Tat into cells. Once it enters the cytoplasm through endocytosis, it can modulate the transcriptional activities in the cell (Debaisieux *et al.*, 2012; Huigen *et al.*, 2004; Rusnati & Presta, 2002; Tyagi *et al.*, 2001). Additionally, the interaction of Tat dimer with HS results in forming a tetrameric structure of HS-Tat-Tat-HS that directly promotes the extravasation caused by binding lymphoid cells to the endothelium (Rusnati & Urbinati, 2009).

HML-2 Env also specifically implicates HS as its direct attachment factor. The binding has been found to promote viral infectivity (Geppert, 2019; Robinson-McCarthy *et al.*, 2018). For HTLV, HSPG functions more as an entry coreceptor than an attachment factor. It has been found to play a critical role in HTLV -1 binding and entry into adherent non-T cells and $CD4^+$ T cells (Jones *et al.*, 2005; Okuma *et al.*, 2003; Pinon *et al.*, 2003). A similar role for HSPG has been found to play in the entry of HSV-1 (Shukla *et al.*, 1999).

The interaction with HS or HP is mostly ionic (Capila & Linhardt, 2002; Tyrrell *et al.*, 1995). It depends on the patterns and orientations of the sulfate and carboxyl groups in the polysaccharide chain. Consequently, specific domains on the viral Env or the protein ligands

with correct fit patterns of positively charged residues are required to ensure appropriate affinity and specificity in the interaction (Hileman *et al.*, 1998; Munoz & Linhardt, 2004). These HP/HS binding domains (HBD) are rich in basic amino acids such as arginine and lysine (Munoz & Linhardt, 2004). HBDs in some HP/HS binding proteins are characterized by two common consensus sequences, XBBXBX and XBBBXXBX, where B is a basic residue and X is any non-basic residue. They are arranged in *β*-strand and *α*-helix conformations, respectively. The B residues are oriented on one side of the protein generating a positively charged surface. (Cardin & Weintraub, 1989; Munoz & Linhardt, 2004). However, not all HP/HS binding proteins show the consensus sequences. Instead, they have patterns of frequent clusters of basic residues (one to three B) separated by spacers of non-basic residues (one or two X) (Fromm *et al.*, 1997). In addition, a spatial structural motif similar to the previously described patterns could be used to bind HP/HS (Margalit *et al.*, 1993). Among basic residues, arginine and lysine are the most common residues found in HBDs. Both have positive charges at neutral pH. However, arginine forms stronger and stabler bonds with sulfate groups than lysine (Fromm *et al.*, 1995). The arginine to lysine ratio is believed to be tailored for optimum binding with HP/HS (Fromm *et al.*, 1997). Additionally, some non-basic residues in the domain might play a role in HP/HS interaction (Caldwell *et al.*, 1996). Viral Envs may involve more than one HBD in interacting with HS chains. HIV, for example, has four HBDs on its gp120 subunit (Crublet *et al.*, 2008). HTLV and MMTV Envs, on the other hand, have each one HBD to bind HS (Krilleke *et al.*, 2007; Zhang *et al.*, 2003). The HBD of HML-2, however, has not been identified yet.

Because of structural relatedness, most HP-binding proteins can bind HS in biological systems (Munoz & Linhardt, 2004; Ori *et al.*, 2011). On the other hand, HS binding proteins can also bind HP. Since HP is commercially available in large quantities and in many forms (as salts and a chromatography resin), it can be a surrogate for HS in testing HS-binding proteins. A bead-based assay using HP-coated beads is a standard method in this regard (Xu & Esko, 2014).

Chondroitin sulfate and DS also share structural similarities with HS due to epimerization and changes in sulfation patterns. They are found at the cell surface and extracellular matrix. Consequently, the capability of HS-binding proteins to bind those two GAGs in addition to HS is expected but with less affinity (Djerbal *et al.*, 2017; Olson & Esko, 2013; Xu & Esko, 2014).

1.6. The role of HERV proteins in human health and diseases

Nearly all human tissues express HERVs even at minor levels, and it is hard to find a human tissue that lacks any HERV transcripts (Bannert & Kurth, 2004; Muradrasoli *et al.*, 2006; Seifarth *et al.*, 2005). However, not all HERVs loci are transcriptionally active. The post-insertional mutations and the different cellular silencing mechanisms, including epigenetic modifications, all contribute to repressing their transcription. As a result, only about 30% of loci are active (Fuchs *et al.*, 2011; Oja *et al.*, 2007; Pérot *et al.*, 2012). The transcripts of HML-2 function as double-edged swords, they can be beneficial for the host, but they can also be pathogenic. For example, they are found at high levels in the embryonic stem cells and contribute to inducing and maintaining the pluripotency of the undifferentiated stem cells (Fuchs *et al.*, 2013). Additionally, LTRs function as alternative promoters with bidirectional activity. They can enhance the expression of genes downstream and upstream (Dunn *et al.*, 2006; Feuchter & Mager, 1990). In the hippocampus, the LTR of an HML-2 locus on chromosome 22 acts as a neuron-specific enhancer to regulate the expression of the proline dehydrogenase gene. Based on that, HERVs might have a role in the evolution of the human central nervous system (Suntsova *et al.*, 2013). Moreover, the fusogenic Envs of HERV-W (Syncytin-1) and HERV-FRD (Syncytin-2) are highly expressed in syncytiotrophoblast and cytotrophoblast cells. Together with their receptors SLC1A5 and MFSD2 (major facilitator superfamily domain containing 2), they form the placenta (Blaise *et al.*, 2003; Blond *et al.*, 2000; De Parseval & Heidmann, 2005; Esnault *et al.*, 2008; Malassine *et al.*, 2005). However, HML-2 Env is expressed in the villous and the extravillous of the cytotrophoblast cells but has no fusogenic activity (Kämmerer *et al.*, 2011). Most recently, HML-2 *env*, *gag*, and *pol* were found to be activated and expressed in COVID-19 patients. The elevated levels of these transcripts are believed to activate interferon production in moderate to severe COVID-19 patients (Guo *et al.*, 2022).

On the other hand, increasing interest in HERVs studies as potential consequences on human health has revealed the implication of HERV-K proteins in cancer development and autoimmunity diseases, especially rheumatoid arthritis and systemic lupus erythematosus. In addition to amyotrophic lateral sclerosis (ALS) (Alfahad & Nath, 2013; Herve *et al.*, 2002; Krzysztalowska-Wawrzyniak *et al.*, 2011; Kurth & Bannert, 2010). ALS is a neurodegenerative disorder that affects the motor neurons and causes fatal paralysis within a few years. The majority of cases are sporadic (Kiernan *et al.*, 2011; Masrori & Van Damme, 2020; Van Rheenen *et al.*, 2016). 20% are caused by increased levels of HML-2 transcripts in the brain of

ALS patients (Tam *et al.*, 2019), particularly RT and Env proteins (Douville *et al.*, 2011; Garcia-Montojo *et al.*, 2019; Li *et al.*, 2015). Moreover, the severity of the disease is linked to the high levels of antibodies against HML-2 (Arru *et al.*, 2018). A recent study showed that treating patients with a combination of HIV-antiretroviral drugs reduced HML-2 transcripts levels and slowed the disease progression (Garcia-Montojo *et al.*, 2021).

Many relative exogenous retroviruses are causative agents of different types of cancers. Such as HTLV-1 in humans, which causes adult T-cell leukemia/lymphoma in 2-5% of the infected cases, and MMTV, which causes breast cancer in mice (Graff *et al.*, 1949; Zhang *et al.*, 2017). The expression of HML-2 proteins, especially the Env and the accessory proteins Rec and Np9, were associated with different kinds of tumors, including teratocarcinoma, melanoma, ovarian, prostate, breast cancers, germ cell tumors, and leukemia blood lymphocytes (Armbruester *et al.*, 2002; Büscher *et al.*, 2005; Kurth & Bannert, 2010; Löwer *et al.*, 1984; Wang-Johanning *et al.*, 2001; Wang-Johanning *et al.*, 2007). LTR also plays a role in inducing cancer by activating the oncogenes. The chromosomal rearrangement caused by homologous recombination can also cause cancer (Garcia-Montojo *et al.*, 2018). Furthermore, HML-2 Env acts directly or indirectly in pathogenicity. It can modulate cellular proliferation and immunosuppression through a conserved region of the TM subunit that has been shown to change cytokine expression, especially interleukin 10 (IL10) (Morozov *et al.*, 2013). Moreover, increasing levels of HML-2 mRNA were found to associate with HIV infection, which induces the transcripts in peripheral blood mononuclear cells (PBMCs) (Bhardwaj *et al.*, 2014; Contreras-Galindo, Almodóvar-Camacho*, et al.*, 2007). However, some studies confirmed the presence of detectable levels of mRNA in the plasma of HIV patients (Contreras-Galindo, Almodóvar-Camacho*, et al.*, 2007; Contreras-Galindo *et al.*, 2006; Contreras-Galindo *et al.*, 2012; Contreras-Galindo, López*, et al.*, 2007; Esqueda *et al.*, 2013), others could not (Bhardwaj *et al.*, 2014; Karamitros *et al.*, 2016). Nevertheless, high titers of HML-2 specific antibodies have been detected in the blood of HIV patients, especially the elite controllers, who do not get any antiviral treatment but still have controlled levels of HIV (Michaud *et al.*, 2014); in addition, HERV specific T-cell response to HIV infection has been reported and found to control HIV viral load (Garrison *et al.*, 2007; Jones *et al.*, 2012; SenGupta *et al.*, 2011; Tandon *et al.*, 2011).

1.7. The importance of the early entry pathway of HML-2

None of the current known HML-2 elements produce infectious particles due to post-insertional mutations or the absence of the Env (Boller *et al.*, 1993; Boller *et al.*, 2008; Hohn *et al.*, 2013; Muster *et al.*, 2003; Tönjes *et al.*, 1997). The polymorphisms of some intact HML-2 proviruses, on the other hand, suggest that bearing those mutations varies among ethnic groups. As a result, an infectious provirus could exist even at low incidence in the human population (Belshaw *et al.*, 2004; Hohn *et al.*, 2013; Mayer & Meese, 2003; Turner *et al.*, 2001). Recombination events can also result in the production of chimeric infectious particles. For example, the gammaretrovirus XMRV is the consequence of a recombination event between two defective mouse endogenous retroviruses (Paprotka *et al.*, 2011; Urisman *et al.*, 2006). Dewannieux et al., 2006, demonstrated recombination in vitro by recombining three partially activated HML-2 elements to reassemble an infectious Phoenix HML-2 provirus (Dewannieux *et al.*, 2006). HML-2, mature and immature, viral particles were found in the plasma of HIV-infected patients, according to one study. Recombination between some HML-2 loci that become active following HIV infection is most likely the source of these viral particles (Contreras-Galindo *et al.*, 2012).

HML-2 transcripts in response to HIV infection (Contreras-Galindo, López*, et al.*, 2007; Gray *et al.*, 2019; Vincendeau *et al.*, 2015), on the other hand, imply that HIV replication may be modulated (Garcia-Montojo *et al.*, 2018). The potential interactions between HIV-1 and HML-2 could significantly impact HIV features, including infectivity, tropism, and cytopathic harm it causes to host cells (Brinzevich *et al.*, 2014). For example, the ability of functional integrase from HML-2 loci to act on HIV LTR has been described (Kitamura *et al.*, 1996). As a result, it can compensate for the defective HIV integrase under certain circumstances (Garcia-Montojo *et al.*, 2018). Similarly, the HML-2 Env K18 expressed by primary lymphocytes can rescue Env-deficient HIV in a co-transfection in vitro study, permitting assembly and release of infectious viral particles (Brinzevich *et al.*, 2014). Given the broad tropism of HML-2 Env (Kramer *et al.*, 2016; Robinson & Whelan, 2016), HIV virions pseudotyped with HML-2 Env would have a different cellular tropism. However, there is no indication that HML-2 Env can rescue HIV virions that are deficient in vivo (Garcia-Montojo *et al.*, 2018). Still, it raises the importance of identifying the HML-2 Env receptor and any cellular factors involved in its early entry pathway. Identifying these host cell requirements would give a deep insight into the conservation of virus-receptor interactions and a better understanding of the viral evolutionary history (Maginnis, 2018). It also could be of benefit for HML-2 Env implicated-cancer studies.

2. Aim of the study

The HML-2 Env mediates viral entry into cells of many species, including mammals, birds, and reptiles, demonstrating a broad tropism (Dewannieux *et al.*, 2006; Kramer *et al.*, 2016; Robinson & Whelan, 2016). Previous studies showed that HSPG is the first requirement in the HML-2 entry pathway (Robinson-McCarthy *et al.*, 2018). The cell surface HSPGs are vital in binding many viruses to the cell surface. They can serve as an entry receptor as well as an attachment factor (Cagno *et al.*, 2019). For HML-2 Env, HSPG was found to bind the Env as an attachment factor and facilitate the viral entry without precluding the involvement of other receptor or cellular proteins (Robinson-McCarthy *et al.*, 2018). Binding HSPG requires particular domains on the viral Env (Hileman *et al.*, 1998), which are unknown for HML-2 Env.

The first part of this study aimed to identify and characterize the HS-binding domain on the HML-2 Env by aligning the Env sequence to the relative MMTV Env, which binds HS using a specified domain. Then, performing site-directed mutagenesis on all positively charged residues in the identified domain, aiming to study their influence on the Env binding and the entry.

The second and third parts of the study aimed to identify and characterize the other cellular factors or requirements in the early entry pathway of HML-2 using two different approaches.

The first approach included designing and generating a soluble trimer fusion protein of the HML-2 that binds to permissive cells that express the elusive receptor molecules at the cell surface. Then, using the produced trimer to pulldown the receptor molecule based on the affinity between them in order to isolate the latter and identify it by mass-spectrometry analysis.

The second approach included a functional screening of integral membrane proteins cDNA library in non-permissive cell lines to identify the receptor and/or other cellular factors involved in the early entry events of HML-2.

3. Materials and methods

3.1. Materials

3.1.1. Devices and equipment

1- devices that are used in this study

Table 3-1 Devices list

Device	Supplier
Avanti J-20 XP centrifuge	Beckman Coulter, USA
BD FACSCalibur	BD, USA
Centrifuge 5415 D	Eppendorf, Germany
Centrifuge 5804R	Eppendorf, Germany
Centrifuge 5810R	Eppendorf, Germany
CFX96 Touch Real-Time System	Bio-Rad Laboratories, USA
CO2 Incubator	Binder, Germany
Coulter particle count and size analyzer Z2	Beckman Coulter, USA
Curix 60 Film processor	AGFA, Belgium
DNA Engine Thermal cycler (PCR)	Bio-Rad Laboratories, USA
Eclipse TS100-F Inverted Fluorescence Microscope	Nikon, Japan
ELISA Reader Sunrise	Tecan, Switzerland
FlexCycler PCR	Analytik Jena, Germany
Fume hood	Köttermann, Germany
Gel-documentation system	Phase, Germany
HB-500 MinidizerTM Hybridization Oven	UVP, USA
Horizontal electrophoresis system	Bio-Rad Laboratories, USA
Ice machine	Ziegra, Germany
Incubator	Heraeus, Germany
Incubator shaker C24	New Brunswick Scientific, USA
Iincubator shaker ecotron	Infors-HT, Germany
Infinite F200 Microplate reader	Tecan, Switzerland
Inverted Phase Contrast Light Microscope ID03	Zeiss, Germany
KAISER Darkroom Safelight with Red-Filter	Fotoimpex, Germany
Lab dancer vortex	IKA, China
Laminar flow	Karl Bleymehl-reinraumtechnik, Germany
Luminator 2RE	Rex-Leuchtplatten, Germany
LUMIstar Omega	BMG LABTECH, Germany
Magnetic stirrer Combimag-REO	IKA, China
Microplate Spectrophotometer Multiskan GO	Thermo Fisher Scientific, Germany

Microwave	Bosch, Germany
Mikro 200R microliter centrifuge -cooled.	Hettich, Germany
Microplate Luminometer Centro LB 960	Berthold Technologies, Germany
Mini Trans-Blot Module	Bio-Rad Laboratories, USA
Minifuge	Neolab, Germany
Mini-PROTEAN Tetra Vertical Electrophoresis system	Bio-Rad Laboratories, USA
Mini-PROTEAN Tetra Handcast Systems	Bio-Rad Laboratories, USA
Multifuge 1S-R	Thermo Fisher Scientific, Germany
Nanodrop 1000	Thermo Fisher Scientific, Germany
Nunc- Aura PCR cabinet	Nunc, Denmark
Odyssey Infrared Imaging System	LI-COR, Germany
Optima L-100K ultracentrifuge	Beckman Coulter, USA
Optima XL-90 ultracentrifuge	Beckman Coulter, USA
Overhead mixer REAX 2	Heidolph Instruments, Germany
PH meter MP 220	Mettler Toledo, Germany
Power supply PAC 300/ 3000	Bio-Rad Laboratories, USA
Precision balance Kern EG	Sartorius, Germany
Precision balance PR803	Mettler Toledo, Germany
Pump-P1	Amersham biosciences, UK
Roller mixer RM5	Cat-Ing, Germany
Rotator	IKA, China
Safety Cabinets HERAsafe KS12	Kendro, Germany
Shaker 3006	GFL, Germany
Thermomixer 5436	Eppendorf, Germany
Trans-Blot SD Semi-Dry Electrophoretic Transfer	Bio-Rad Laboratories, USA
Tube strip picofuge	Stratagene, USA
UV Transilluminator	Vilber Lourmat, France
VACUSAFE aspiration system	Integra-Biosciences, Germany
Vortex Mixer	VWR international
Water bath	PolyScience, USA

2- Tools and equipment

Table 3-2 Equipment list

Equipment	Manufacture
5 mL Round Bottom Polystyrene FACS Tubes	Carl Roth, Germany
Amicon Ultra-15 centrifugal filter Ultracel 30K	Merck Millipore, Ireland
Black/Clear Flat Bottom Polystyrene Microplate 96-well	Corning, USA
Blot Absorbent Filter Papier	Bio-Rad Laboratories, USA
Cell scraper 24cm	TPP, Switzerland

CL-XPosure Film	Thermo Fisher Scientific, Germany
Costar 3896 plates	Corning, USA
Cryo.S tubes	Merck, Germany
Deep well plates 96	Carl Roth, Germany
Extra Thick Blot Paper	Bio-Rad Laboratories, USA
Falcon Cell strainer 40um, 100um	Corning, USA
Kodak cassette	Kodak, USA
Magnetic Particle Separator	Thermo Fisher Scientific, Germany
Minisart filter unit 0.45um, 0,22um	Sartorius, Germany
Neubauer-cell counter (0,0025 mm²/0,1 mm)	Carl Roth, Germany
Nitrocellulose Membranes, 0.45um	Bio-Rad Laboratories, Germany
Nunc MaxiSorp flate 96plates	Thermo Fisher Scientific, Germany
Nunclon Delta surface white cell culture test plates 96 F	Thermo Fisher Scientific, Germany
Pipettes of all sizes	Eppendorf, Germany
Rotilab-sealing film micro-test plates	Carl Roth, Germany
Scienceware replicator, 96-well	Sigma-Aldrich, Germany
Single-use glass pipettes (2 bis 50 ml)	TPP, Switzerland
Single-use injects (5, 20 ml)	B, Braun, Germany
Tissue culture dishes 100, 150	TPP, Switzerland
Tissue culture Flasks 25, 75, 150, 300 cm2	TPP, Switzerland
Tissue culture test plates F 6, 12, 24, 48, 96-wells	TPP, Switzerland
Ultra-clear centrifuge Tubes	Beckman Coulter, USA
V96 microwell plate, clear	Thermo Fisher Scientific, Germany
Vac-Man 96 Vacuum Manifold	Promega, USA

3- Software

Table 3-3 The used software

Software	Company
CFX Maestro	Bio-Rad, USA
FlowJo V10	BD, USA
Geneious Prime 2020.2.3	Biomatters, New Zealand
Graphpad Prism V 9.1	Graphpad Software, USA
Magellan	Tecan, Switzerland
MARS Data Analysis Software	BMG LABTECH GmbH, Germany

3.1.2. Chemicals, reagents, enzymes, and kits

1- Chemicals and reagents

Table 3-4 Chemicals and reagents list

Name	Manufacture
2-propanol	Carl Roth, Germany
6X DNA *Loading Dye*	Thermo Fisher Scientific, Germany
Acetic acid	Carl Roth, Germany
Agar-Agar, bacteriological	Carl Roth, Germany
Agarose HR-PLUS	Carl Roth, Germany
Ambion Nuclease-Free water	Thermo Fisher Scientific, Germany
Ammonium Persulfate (APS) 10 %	Carl Roth, Germany
Ampicillin Sodium salt	Carl Roth, Germany
Bromophenol blue indicator	Carl Roth, Germany
complete Mini Protease Inhibitor Cocktail- Roche	Sigma-Aldrich, Germany
COULTER CLENZ cleaning agent	Beckman Coulter, USA
DEPC treated Water	Ambion Life Technologies, USA
Dimethyl sulfoxide (DMSO)	Carl Roth, Germany
dNTPs	Thermo Fisher Scientific, Germany
Double-distilled water(ddH_20)	Robert Koch Institute
eBioscience Fixable Viability Dye eFluor 660	Thermo Fisher Scientific, Germany
Ethanol	Carl Roth, Germany
Ethidium bromide	Carl Roth, Germany
Ethylenediaminetetraacetic acid (EDTA)	AppliChem, Germany
FBS IgG Stripped	PAA Laboratories, USA
GeneRuler 100bp Ladder Plus	Thermo Fisher Scientific, Germany
GeneRuler 1kb Ladder Plus	Thermo Fisher Scientific, Germany
Gibco Fetal Bovine Serum FBS	Thermo Fisher Scientific, Germany
Gibco Pen Strep (Penicillin/Streptomycin) 100	Thermo Fisher Scientific, Germany
Glycerol	Carl Roth, Germany
Glycine	Carl Roth, Germany
Heparan sulfate	AMS Biotechnology, UK
Heparin sodium	Sigma-Aldrich, Germany
Hydrogen peroxide, 30 %	Sigma-Aldrich, Germany
Kanamycin	Carl Roth, Germany
L-Glutamine 200 mM	Sigma-Aldrich, Germany
Methanol	Carl Roth, Germany
MS2 RNA Template	Sigma-Aldrich, Germany

Nonidet- P40	AppliChem, Germany
O-phenylenediamine dihydrochloride (OPD) tablet, 5mg	Sigma-Aldrich, Germany
PageRuler Prestained Protein Ladder	Thermo Fisher Scientific, Germany
Paraformaldehyde (PFA)	Carl Roth, Germany
Polybrene Infection / Transfection Reagent	Sigma-Aldrich, Germany
Polyethyleneimine (PEI)	Polysciences, USA
Precision Plus Protein Dual Color Standards	Bio-Rad Laboratories, USA
Quick Coomassie staining	Generon, UK
RNase OUT (40 U/µl)	Thermo Fisher Scientific, Germany
Rotiphorese 50X TAE buffer	Carl Roth, Germany
Rotiphorese Gel 30 (Acrylamide 30%/Bis)	Carl Roth, Germany
SERVA Native Marker	SERVA Electrophoresis, Germany
Skimmed milk powder	Carl Roth, Germany
Sodium azide 99,5%	SERVA Electrophoresis, Germany
Sodium chloride	Carl Roth, Germany
Sodium dodecyl sulfate SDS pellets	Carl Roth, Germany
ß-Mercaptoethanol	Sigma-Aldrich, Germany
Sucrose	Sigma-Aldrich, Germany
Sulfuric acid, 30%	Carl Roth, Germany
TEMED	Roth, Karlsruhe
Terrific broth modified	Carl Roth, Germany
Tris	Carl Roth, Germany
Triton-X 100	Carl Roth, Germany
Trypan blue	Sigma-Aldrich, Germany
Trypsin-EDTA solution 1X	Sigma-Aldrich, Germany
Tween-20	Carl Roth, Germany

2- Kits

Table 3-5 List of kits

Kit	Manufacture
Applied biosystem BigDye Terminator v3.1	Thermo Fisher Scientific, Germany
HS-Mg RT Activity Kit (Cavidi-Assay)	Cavidi AB
Luciferase assay system	Promega, USA
Mini-PROTEAN TGX Precast Protein Gels 4–15%	Bio-Rad Laboratories, USA
Mini-PROTEAN TGX Precast Protein Gels 4–20%	Bio-Rad Laboratories, USA
NucleoBond PC 500, Maxi kit	Macherey Nagel
Pierce BCA Protein assay Kit	Thermo Fisher Scientific, Germany
Platinum SuperFi DNA Polymerase	Thermo Fisher Scientific, Germany

Protein Deglycosylation Mix II	New England Biolabs, Germany
ProteoSilver Silver Stain plus	Sigma-Aldrich, Germany
Qiagen Plasmid Maxi Kit	Qiagen, Germany
Qiaprep Spin Miniprep Kit	Qiagen, Germany
QIAquick Gel Extraction Kit	Qiagen, Germany
QIAquick PCR Purification Kit	Qiagen, Germany
Quant-iT PicoGreen dsDNA Assay Kits	Thermo Fisher Scientific, Germany
QuikChange Lightning Site-Directed Mutagenesis Kit	Agilent Technologies, USA
SensiFAST Probe No-ROX Master Mix One Step Kit	Bioline, USA
SuperSignal West Pico PLUS Chemiluminescent Substrate	Thermo Fisher Scientific, Germany
V5-tagged protein purification kit Ver.2	MBL, Japan
Wizard(R) SV 96 Plasmid DNA Purification System	Promega, USA

3-Enzymes

Table 3-6 List of Enzymes

Name	Manufacture
Accutase - Enzyme Cell Detachment Medium	Thermo Fisher Scientific, Germany
DpnI	New England Biolabs, UK
Dream Taq DNA Polymerase	Thermo Fisher Scientific, Germany
FastDigest EcoRI	Thermo Fisher Scientific, Germany
FastDigest HindIII	Thermo Fisher Scientific, Germany
Heparinase I and III Blend	Sigma-Aldrich, Germany
NdeI High Fidelity	New England Biolabs, UK
NotI	Thermo Fisher Scientific, Germany
Pfu DNA Polymerase	Thermo Fisher Scientific, Germany
T4 DNA Ligase	Thermo Fisher Scientific, Germany

3.1.3. Buffers and beads

1- Buffers

Table 3-7 List of used buffers.

Buffer	components
10x Tris-Glycine buffer	30.3g Tris, 114.2g Glycine, Add to 1L with ddH20
1x PBS w/o Ca2+, w/o Mg2+	Robert Koch Institute
2x Tris-Glycine Native Sample Buffer	200 mM Tris HCl, 20% Glycerol, 0.005% Bromophenol Blue pH 8.6
Carbonate-Bicarbonate Buffer	Sigma-Aldrich, Germany
Cell culture lysis buffer 5x	Promega, USA
ELISA substrate	12.5 ml of Phosphate-Citrate Buffer, 5 mg of OPD, 12 µl H2O2 (30%) solution.
FACS fixation buffer	3 % paraformaldehyde in 1x PBS.
FACS staining buffer	2% (v/v) FBS, 1mM NaN3 in 1x PBS.

Heparinase reaction buffer	20mM Tris- HCl, 50mM NaCl, 4mM CaCl2, 0,01% (w/v) BSA, pH 7.5
Homogenizing buffer	50 mM Tris, 150 mM NaCl, 1 mM EDTA, 1% NP-40, 1% glycerol, protease inhibitor cocktail, pH7.4
Native Running Buffer	1x Tris-Glycine buffer
NP-40 lysis buffer	150 mM NaCl, 1.0% NP-40,50 mM Tris, protease inhibitor pH 8.0
PERT 2x Lysis buffer	0,25% Triton X-100, 50mM KCl, 100mM TrisHCl, 40% Glycerol, pH 7,4
Phosphate-citrate buffer	Sigma-Aldrich, Germany
PM buffer	2% Skimmed milk in 1x PBS
PMT buffer	2% Skimmed milk in 1x PBS, 0,1% Tween-20
Protein collecting buffer	1 M Tris-HCl, pH 9.0
Protein Elution buffer	0.1 M glycine-HCl, pH 2.7
PT buffer	0,05 % Tween-20 in 1x PBS
Resolving Buffer, 1.5M Tris-HCl pH 8.8	Bio-Rad Laboratories, USA
Restore western blot stripping buffer	Thermo Fisher Scientific, Germany
Roti-Free Stripping Buffer 2.0	Carl Roth, Germany
SDS Running buffer	1x Tris-Glycine buffer, 10% SDS
SDS Sample Buffer, Laemmli 2× Concentrate	Sigma-Aldrich, Germany
Stacking Buffer, 0.5 M Tris-HCl pH 6.8	Bio-Rad Laboratories, USA
TN buffer (for pull-down purified proteins)	10mMTris-HCl, 150 mM NaCl, 0.2% TritonX-100 and a protease inhibitor cocktail
TNE buffer	50mM Tris, 2mM EDTA, 150mM NaCl, 1% tritonX-100
Transfer buffer	1x Tris-Glycine buffer, 20% Methanol
WB Blocking buffer	5% Skimmed milk in 1x PBS
WB washing buffer	0,001% Tween-20 in 1x PBS

2- Beads

Table 3-8 List of used beads in purification and pull-down assays

Name	manufacture
Anti-V5 Agarose Affinity Gel	Sigma-Aldrich, Germany
Dynabeads® Protein G	Thermo Fisher Scientific, Germany
Heparin-agarose type I	Sigma-Aldrich, Germany
HiTrap Protein G HP (5 X 1mL)	GE Healthcare, USA
Protein A agarose beads	Sigma-Aldrich, Germany
Protein G on Sepharose 4B fast flow	Sigma-Aldrich, Germany
Streptavidin−agarose	Sigma-Aldrich, Germany

3.1.4. Antibodies

Table 3-9 List of used antibodies with their working dilution

Antibody	Host/class	Producer	Dilution
α-β-Actin	Mouse, monoclonal	Sigma-Aldrich, Germany	1:1000
α-GAPDH	Mouse, polyclonal	Sigma-Aldrich, Germany	1:1000
α-(HERV K) Env (HERM-1811-5)	Mouse, monoclonal	Austral Biologicals, USA	1:5000
α-V5 Tag - HRP	Mouse, monoclonal	Thermo Fisher Scientific, USA	1:5000
α-Human IgG (γ-chain specific)-Peroxidase	Goat, Polyclonal	Sigma-Aldrich, Germany	1:1000 ELISA 1:5000 WB
α-X-receptor antibody	Rabbit, polyclonal	Abcam, UK	1:1000
α-Rabbit IgG - HRP	Goat, polyclonal	Southern Biotech, USA	1:5000
α-Mouse IgG -HRP	Goat, Polyclonal	Southern Biotech, USA	1:10000
α- GOLPH2 IgG1	Mouse, monoclonal	Thermo Fisher Scientific, USA	1:50 FACS 1:1000 WB
α-Human IgG (H+L)- Alexa Fluor 488	Goat, Polyclonal	Thermo Fisher Scientific, USA	1µg/ml
α-Mouse IgG (H+L)- Alexa Fluor 488	Goat, Polyclonal	Thermo Fisher Scientific, USA	1µg/ml
Mouse IgG1 kappa Isotype Control - FITC	Mouse, monoclonal	Thermo Fisher Scientific, USA	1µg/ml
α- VSV-G mouse	Mouse, monoclonal	Abcam, UK	1:1000
α-Mouse IgG -IR Dye 680LT	Donkey, Polyclonal	LI-COR Biosciences - Germany	1:25000
α-GP37	Mouse, monoclonal	Santa Cruz Biotechnology, USA	1:1000
α-V5 Tag	Mouse, monoclonal	Bio-Rad, USA	1:5000
α-human IgG -IR Dye 680LT	Goat, Polyclonal	LI-COR Biosciences - Germany	1:25000
α-Gag HIV (AG3.0)	Mouse, Polyclonal	Stephen Norley	1:800
HIV plasma pool	human	Stephen Norley	1:10000

3.1.5. Cells and medium

1- Human cell lines

Table 3-10 List of used Cell lines.

Cell Line Name	Tissue	Medium
HEK293T	kidney; Embryo	DMEM c
HCT116	Large intestine; Colon	RPMI-1640 c
NCI-H23	Non-Small Cell; Lung	RPMI-1640 c
NCI-H226	Non-Small Cell; Lung	RPMI-1640 c
NCI-H522	Non-Small Cell; Lung	RPMI-1640 c
SK-MEL-2	Melanoma; Skin	RPMI-1640 c
SK-MEL-28	Melanoma; melanocyte Skin	RPMI-1640 c
SK-MEL-5	Melanoma; melanocyte Skin	RPMI-1640 c
HEp-2	carcinoma; Larynx	DMEM c
MALME-3M	Melanoma; Fibroblast Skin	RPMI-1640 c

DU-145	Prostate	RPMI-1640 c
U251	CNS	RPMI-1640 c

2- Bacterial strains

Table 3-11 List of used E. coli strains.

Strain	Supplier
One-Shot TOP10 Chemically Competent E. coli	Thermo Fisher Scientific, Germany
XL10-Gold Ultracompetent cells	Agilent Technologies, USA

3- Mediums for cell lines and bacteria

Table 3-12 List of cell lines and bacterial mediums.

Medium	components
RPMI 1640	Robert-Koch-Institute
DMEM	Robert-Koch-Institute
DMEM c	DMEM + 10% FBS+ 2mM L-Gln
RPMI-1640 c	RPMI 1640 + 10% FBS+ 2mM L-Gln
LB-Medium	Robert-Koch-Institute
LB-Agar	20g/l of LB-Medium
Terrific Medium	47.6 g/l ddH2O, 4 ml glycerol
S.O.C. Medium	Thermo Fisher Scientific, Germany

3.1.6. Plasmids and primers

1- Plasmids

Table 3-13 List of Plasmids.

Name	Source/Publication
pTH-IP10LP-synHIVgp120-Fibritin (K3)	Nadine Beimforde, n.p.
pcDNA-oricoEnvΔC1 /-V5(HERV-K (HML-2) Env Plasmid)	(Hanke *et al.*, 2009)
pcDNA-oricoEnvCS$^-$-V5	(Hanke *et al.*, 2009)
SM-FT-V5 in pTH	Alaa Ramadan, n.p.
pTH-SM-FT-hFC	Alaa Ramadan, n.p.
pTH-SU-FT-hFC	Alaa Ramadan, n.p.
pTH-SU-FT-V5	Alaa Ramadan, n.p.
pCSI-XSU-Fc (XSU-hFc)	Dusty Miller (Battini *et al.*, 1999)
psPAX2	Trono Lab/ addgene
pWPXL-Luci	(Wamara, 2020)
pCMV-VSV-G (Env plasmid)	(Bannert *et al.*, 2000)
pcDNA- oricoEnvR140C Δ659-699 (ΔC1) (R140C Env. plasmid)	(Hanke *et al.*, 2009)
pTH-XMRV coEnv (XMRV Env plasmid)	Oliver Hohn (Hong *et al.*, 2009)
pcDNA-JSRV coEnv (JSRV Env plasmid)	(Kramer *et al.*, 2016)
pMAX-GFP	Oliver Hohn/ Lonza, Switzerland
pCMV6-XL5- GOLPH2 (NM_016548) Human Untagged Clone GOLM1	Origene, USA
pLX304- library	Transomic co./ (Yang *et al.*, 2011)

pWPXL-ovalbumin	Oliver Hohn, n.p.
pcDNA-oricoEnvR216A-ΔC1 (R216A Env plasmid)	Alaa Ramadan, n.p.
pcDNA- oricoEnvK219A/R221A-ΔC1 (K219A/R221A Env Plasmid)	Alaa Ramadan, n.p.
pcDNA-oricoEnvR221A/K223A-ΔC1 (R221A/K223A Env plasmid)	Alaa Ramadan, n.p.
pcDNA-oricoEnvK223A/K225A-ΔC1 (K223A/K225A Env plasmid)	Alaa Ramadan, n.p.
pcDNA-oricoEnvK223A-ΔC1 (K223A Env plasmid)	Alaa Ramadan, n.p.
pcDNA-oricoEnvK225A-ΔC1 (K225A Env plasmid)	Alaa Ramadan, n.p.
pcDNA-oricoEnvK225A/K229A-ΔC1 (K225A/K229A Env plasmid)	Alaa Ramadan, n.p.
pcDNA-oricoEnvK229A-ΔC1 (K229A Env plasmid)	Alaa Ramadan, n.p.
pcDNA-oricoEnvK229A/K233A-ΔC1 (K229A/K233A Env plasmid)	Alaa Ramadan, n.p.
pcDNA-oricoEnvK233A/K236A-ΔC1 (K233A/K236A Env plasmid)	Alaa Ramadan, n.p.
pcDNA-oricoEnv HBD all Sites-ΔC1 (HBD-1 Env plasmid)	Alaa Ramadan/ Diana Andus, n.p.
pcDNA-oricoEnvN128A-ΔC1 (N128A Env plasmid)	(Wamara, 2020)
pcDNA-oricoEnvN153A-ΔC1 (N153A Env plasmid)	(Wamara, 2020)
pcDNA-oricoEnvEEE146-148AAA-ΔC1 (EEE146-148AAA Env plasmid)	(Wamara, 2020)
pcDNA-oricoEnvΔ144-152-ΔC1 (Δ144-152 Env plasmid)	(Wamara, 2020)
pcDNA-oricoEnvS138A-ΔC1-V5 (S138A Env plasmid)	(Priesnitz, 2019)
pcDNA-oricoEnvS155A-ΔC1-V5 (S155A Env plasmid)	(Priesnitz, 2019)
pcDNA-oricoEnvG135A-ΔC1-V5 (G135A Env plasmid)	(Priesnitz, 2019)
pcDNA-oricoEnvD138A-ΔC1-V5 (D138A Env plasmid)	(Priesnitz, 2019)
pcDNA-oricoEnvP145A-ΔC1-V5 (P145A Env plasmid)	(Priesnitz, 2019)
pcDNA-oricoEnvR140A-ΔC1-V5 (R140A Env plasmid)	(Priesnitz, 2019)

2- Primers

All ordered primers were received as a lyophilized powder, resuspended in ddH2O to a final concentration of 100µM as a master stock. The working stock was prepared by 1:10 dilution to a final concentration of 10µM.

- Cloning primers

Table 3-14 List of used primers for PCR based cloning with their annealing temperature (AT) and the restriction site at the 5' end

Primer name	Sequences 5'to 3'	REase site	AT °C.
SU domain cloning (Thermo Fisher Scientific)			
ClonHERV-SU-F	GTACGTAAGCTTGCCACCATGAACCCCAGCGAGATGC	HindIII	63
ClonHERV-SU-R	ACGTACGAATTCGTTCAGCACGCCCTTCAG	EcoRI	
SU/TM (SM) cloning (Thermo Fisher Scientific)			
ClonHERV-SU-F	GTACGTAAGCTTGCCACCATGAACCCCAGCGAGATGC	HindIII	65
ClonHERV-SU/TM-R	ACGTACGAATTCTTTCACCCATGTCACGGGGTT	EcoRI	

V5 tag cloning (Thermo Fisher Scientific)			
clonV5-tag-F	AATCTAGCTAGCGGTAAGCCTATCCCTAACCCTCTC	NheI	62
clonV5-tag-R	AAAAGGAAAAGCGGCCGCCTACGTAGAATCGAGACCGAG	NotI	
Human IgG- Fc tag cloning (Thermo Fisher Scientific)			
clonIgG-hFC-F	AATCTAGCTAGCCCGTGCCCAGCACCTGAA	NheI	59
clonIgG-hFC-R	AAAAGGAAAAGCGGCCGCTCATTTACCCGGAGACAG	NotI	
GOLM1(GP73) cloning (Integrated DNA Technology)			
GOLM1-ORF-F	GTACGTAAGCTTGCCACCATGATGGGCTTGGGAAAC	HindIII	55
GOLM1-ORF-R	AATCTAGCTAGCGAGTGTATGATTCCGCTTTTC	NheI	

- Mutagenesis primers

Table 3-15 List of primers used with QuikChange Lightning Site-Directed Mutagenesis Kit.

Primer name	Sequences 5'to 3'	Manufacture
Mutagenesis in pTH vector		
EcoRI site upstream Fibritin motif		
ELF-F	ACCAAGTGGCTGTGGGAATTCAAGGGATACATTCCT	Thermo Fisher Scientific
ELF-R	AGGAATGTATCCCTTGAATTCCCACAGCCACTTGGT	
NheI downstream Fibritin motif		
NheI –Stop-F	GCTCAGCACCTTCCTGGCTAGCTAAGCGGCCGCTCGAGCATG	Thermo Fisher Scientific
NheI –Stop-R	CATGCTCGAGCGGCCGCTTAGCTAGCCAGGAAGGTGCTGAGC	
Heparin-binding site (HBD) mutagenesis		
R216A		
P1-F	GGACTTCAGCTACCAGGCGAGCCTGAAGTTCCGG	Integrated DNA Technology
P1-R	CCGGAACTTCAGGCTCGCCTGGTAGCTGAAGTCC	
R216A/K219A		
P2-F	C TAC CAG GCG AGC CTG GCG TTC CGG CCC AAG GGC	Integrated DNA Technology
P2-R	GCCCTTGGGCCGGAACGCCAGGCTCGCCTGGTAG	
K219A/R221A		
P3-F	GCTACCAGCGGAGCCTGGCGTTCGCGCCCAAGGGCAAGCCC	Integrated DNA Technology
P3-R	GGGCTTGCCCTTGGGCGCGAACGCCAGGCTCCGCTGGTAGC	
R221A/K223A		
P4-F	CGGAGCCTGAAGTTCGCGCCCGCGGGCAAGCCCTGCCCC	Integrated DNA Technology
P4-R	GGGGCAGGGCTTGCCCGCGGGCGCGAACTTCAGGCTCCG	
K223A/K225A		
P5-F	CTGAAGTTCCGGCCCGCGGGCGCGCCCTGCCCCAAAGAG	Integrated DNA Technology
P5-R	CTCTTTGGGGCAGGGCGCGCCCGCGGGCCGGAACTTCAG	
K225A		
P6-F	CGGCCCAAGGGCGCGCCCTGCCCCAAAG	Integrated DNA Technology
P6-R	CTTTGGGGCAGGGCGCGCCCTTGGGCCG	
K229A		

P7-F	GGGCGCGCCCTGCCCCGCAGAGATCCCCAAAGAG	Integrated DNA Technology
P7-R	CTCTTTGGGGATCTCTGCGGGGCAGGGCGCGCCC	
K229A		
P8-F	GGGCAAGCCCTGCCCCGCAGAGATCCCCAAAGAG	Integrated DNA Technology
P8-R	CTCTTTGGGGATCTCTGCGGGGCAGGGCTTGCCC	
K233A		
P9-F	CCCGCAGAGATCCCCGCAGAGTCCAAGAACACC	Integrated DNA Technology
P9-R	GGTGTTCTTGGACTCTGCGGGGATCTCTGCGGG	
K233A		
P10-F	GCCCCAAAGAGATCCCCGCAGAGTCCAAGAACACCG	Integrated DNA Technology
P10-R	CGGTGTTCTTGGACTCTGCGGGGATCTCTTTGGGGC	
K236A		
P11-F	GATCCCCGCAGAGTCCGCGAACACCGAGGTGCTGG	Integrated DNA Technology
P11-R	CCAGCACCTCGGTGTTCGCGGACTCTGCGGGGATC	
K223A		
P23-F	CTGAAGTTCCGGCCCGCGGGCAAGCCCTGCCCC	Integrated DNA Technology
P23-R	GGGGCAGGGCTTGCCCGCGGGCCGGAACTTCAG	
R216A for HBD all site		
P1m-F	GGACTTCAGCTACCAGGCGAGCCTGGCGTTCGCG	Integrated DNA Technology
P1m-R	CGCGAACGCCAGGCTCGCCTGGTAGCTGAAGTCC	
K219A/R221A for HBD all site		
P3m-F	CAGCGGAGCCTGGCGTTCGCGCCCGCGGGCGCG	Integrated DNA Technology
P3m-R	CGCGCCCGCGGGCGCGAACGCCAGGCTCCGCTG	

- PERT assay primers

Table 3-16 PERT assay Primers/probes (Horie et al., 2010)

Primer	Sequence 5'to 3'	Manufacture
Horie2010_PERT-F	TCCTGCTCAACTTCCTGTCGAG	Thermo Fisher Scientific
Horie2010_PERT-R	CACAGGTCAAACCTCCTAGGAATG	Thermo Fisher Scientific
MS2-PERT_Horie-Probe (FAM)	FAM-TCTTTAGCGAGACGCTACCATGGCTA-BHQ	Tib Molbiol

- Sequencing primers

Table 3-17 List of the sequencing primers and their used annealing temperature

Primer name	Sequence 5'to 3'	AT	Manufacture
HBD sequencing primers			
Seq SU-F.	GCGGCTGCCAACTACACCTACTG	55°C	IDT
Seq Env-SU-R	TGTGCTTGTGCTTGTCCAG	55°C	
Library sequencing primers			
CMV-F	CGCAAATGGGCGGTAGGCGTG	54°C	IDT
WPRE-R	CATAGCGTAAAAGGAGCAACA	48°C	
Fusion protein sequencing primers			
pTH-F	CTTTCCATGGGTCTTTTCTG	50°C	Thermo Fisher Scientific
Fibritin-F	GTGAGAAAGGACGGAGAGT	52°C	
BGH-R	CCTCGACTGTGCCTTCTA	52°C	

3.2. Methods

3.2.1. DNA analysis methods

3.2.1.1. PCR based cloning

We used the polymerase chain reaction (PCR) to amplify the gene of interest from the vector DNA before being cloned into an expression vector. The reaction mix included specific primers producing products flanked by the proper restriction sites and a high fidelity Taq DNA polymerase enzyme to create specific and error-free PCR products (the insert). We amplified SU and V5-tag genes from oricoEnv-ΔC1-V5 plasmid, SUTM (SM) from oricoEnvCS^{-}-V5 plasmid, Fc-tag from XSU-hFc, and GP73 from pCMV6-XL5 using primers (Table 3-14) and Platinum SuperFi DNA Polymerase or Pfu DNA Polymerase (Thermo Fisher Scientific).

Table 3-18 Master mix preparation for cloning PCR

Component	*Final Conc.*
5x SuperFi™ Buffer (7.5 mM MgCl)	1x
25 mM dNTPs mix	0.2 mM each
10µM Primer For	0.5 µM
10µM Primer-Rev	0.5 µM
Platinum SuperFi/ Pfu DNA Polymerase	1/1.5 U
H2O	to 50 µL
Template plasmid DNA	1 pg -10 ng
Final volume	50 µl

The PCR master mix was prepared according to Table 3-18. Furthermore, the thermal cycling parameters in Table 3-19 were applied.

Table 3-19 Thermal cycling profile for cloning PCR

Steps		***Temp. SuperFi /Pfu***	***Time SuperFi /Pfu***
Initial denaturation		98°C/95°C	30 sec/1 min
30 cycles	Denature	98°C/95°C	5sec/30 sec
	Anneal	according to primer	10 sec/30 sec
	Extend	72°C	20 sec/2:40 min
Final extension		72°C	5 min
		4°C	hold

3.2.1.2. Mutagenesis PCR

Using the QuikChange Lightning Site-Directed Mutagenesis kit (Agilent Technologies), we introduced EcoRI and NheI restriction sites upstream and downstream of the Fibritin motif in

pTH plasmid, as well as mutagenized the putative heparin-binding site on the K113 oricoEnvΔC1 plasmid. Mutagenesis primers were designed according to the kit's guidelines so that the desired mutation was in the middle of each primer and the melting temperature (T_m) was ≥78°C. Using the following formula, T_m was calculated:

T_m= 81.5 + 0.41(%GC) - 675/N - % mismatch

N is the primer length in bases, and values for %GC and % mismatch are whole numbers. Primers were ordered from two different sources (Thermo Fisher Scientific and IDT) (Table 3-15), and the PCR master mix was prepared according to the kit (Table 3-20).

Table 3-20 Master mix preparation for mutagenesis PCR

Component	***Final vol./con.***
10× reaction buffer	5µl (1x)
dsDNA template	(10–100 ng)
10µM Primer For	0.5 µl
10µM Primer-Rev	0.5 µl
of dNTP mix	1µl
H2O	to 50 µL
Quick-change Lightning Enzyme	1µl

The thermal cycling profile provided by the kit was followed (Table 3-21).

Table 3-21 Thermal cycling profile of mutagenesis PCR using QuikChange Lightning Site-Directed Mutagenesis kit

Step		***Temp.***	***Time***
Initial denaturation		95°C	2 min.
18 cycles	Denature	95°C	20 sec.
	Anneal	60°C	10 sec.
	Extend	68°C	3:45 min.
Final extension		68°C	5 min.

The methylated (parental) DNA template in PCR products was digested with 2 µl of DpnI (provided with the kit) at 37°C for 5 min. Then PCR products were transformed into chemically competent E. coli Top 10 or ultracompetent XL10-Gold cells and plated on selection plates. The successful mutagenesis test was confirmed by Sanger sequencing.

3.2.1.3. Sequencing PCR

We conducted Sanger-sequencing to check for the successful mutagenesis and the cloned inserts into vectors. Sequencing primers (Table 3-17) with BigDye Terminator v3.1 reaction

mix (Thermo Fisher Scientific) were used for the sequencing PCR, in which fluorescence-labeled 2', 3'-dideoxynucleotide triphosphate mix (ddNTPs) were included. The PCR master mix was prepared according to Table 3-22.

Table 3-22 Master mix preparation for Sanger-sequencing PCR

Component	***Final vol./ con.***
DNA template (plasmid DNA)	150 – 300 ng
5x Sequencing Buffer	2 µl
BigDye™ Terminator 3.1 ready reaction mix	0.5 µl
sequencing primer (10 pmol/µl)	0.5 µl
ddH2O	to 10 µl
final volume	10 µl

PCR thermal conditions were set as shown in Table 3-22. The sequence analysis was conducted in the central sequencing laboratory of the RKI, and the sequences were evaluated using Geneious Prime software.

Table 3-23 Thermal cycling profile of sequencing PCR.

Step		***Temp.***	***Time***
Initial denaturation		96°C	2 min.
25 cycles	Denature	96°C	10 sec.
	Anneal	according to primer*	5 sec.
	Extend	60°C	4 min.
End		4°C	hold

Annealing temperature for each sequencing primer is mentioned in (Table 3-17).

3.2.1.4. DNA restriction with enzymes

A restriction with bacterial type II endonucleases was conducted to verify vectors and clone genes (SU, SUTM, FC, V5, and GOLM1) in the pTH plasmid. Restriction enzymes that cut specifically and create a 5' overhang (Table 3-6) were employed with appropriate buffers, as directed by the manufacturer. For the Fast-digest restriction enzymes, the reaction mix was incubated at 37°C for ~ 30 min. However, for standard digestion, the incubation time was extended to 1h. We used electrophoresis to examine the digestion quality and products on an agarose gel.

3.2.1.5. Ligation

For ligation of the DNA fragment of cloned genes with 2-4 base sticky ends after endonuclease restriction, T4 DNA Ligase (Thermo Fisher Scientific) was used to join these fragments with

similar sticky ends digested vector. The reaction mix was set up according to the manufacturer's instructions, using 20 ng of the vector. The insert amount was calculated according to the following formula in a 3:1 insert: vector molar ratio:

ng of insert = 3× (kb of insert/kb of the vector) × ng of vector

The ligation mix, which included both the vector and insert, was incubated at room temperature for 10-20 min. Followed by an inactivation step at 65°C for 10 min. Then 5 µl of the mixture was transformed into One Shot Top 10 chemically competent E. coli (Thermo Fisher Scientific).

3.2.1.6. DNA-agarose gel electrophoresis

Agarose gel electrophoresis, the standard method to separate DNA fragments according to their sizes, was used to detect and separate the PCR products and plasmids or inserts digested with restriction enzymes. Agarose gel served as a matrix for DNA fragments to migrate through and was prepared in different concentrations (0.5 to 2%) according to the length of the analyzed DNA fragment in 1x TAE buffer with (0,5 µg/ml) Ethidium bromide (Carl Roth). DNA samples were mixed with 6x loading dye (Thermo Fisher Scientific) before loading into the casted gels. The electrophoretic separation took place in 1x TAE running buffer at 80 V for 40-60 min, depending on the fragment size. Different DNA markers (ladders) were used based on the length of analyzed DNA (GeneRuler 1kb Plus or GeneRuler 100 bp Plus) (Thermo Fisher Scientific). Analyzed DNA was visualized by UV and documented using the Professional Gel Documentation System (Phase).

3.2.1.7. Bacterial transformation

The One Shot TOP10 Chemically Competent E. coli cells (Thermo Fisher Scientific) were utilized to transform cloning products. Recombinant plasmids were introduced into chemically competent bacterial cells to replicate them. When single-strand mutagenesis was performed, mutagenized plasmids were transformed into XL10-Gold Ultracompetent cells (Agilent Technologies).

1–10 ng of the plasmid DNA was added to each cell's tube on the ice and incubated for 30 min. The cells were then exposed to 42°C for 30 sec, then directly incubated in ice for 2 min. The tubes were then filled with 250-500 µl of pre-warmed SOC medium (Thermo Fisher Scientific) and incubated at 37°C for 1h before being plated on antibiotic-selective LB agar (Table 3-12).

3.2.1.8. DNA isolation and purification

a) DNA Isolation

E. coli bacteria were cultured overnight in 3-5 ml of LB medium containing 100 mg/ml antibiotic (Ampicillin or Kanamycin). The plasmid DNA was isolated using the Qiaprep Spin Miniprep Kit (Qiagen). For a high yield plasmid DNA, we used Qiagen Plasmid Maxi Kit (Qiagen), and the bacteria were cultivated in a greater volume of LB medium (250-300 ml) with selective antibiotics.

We used the NucleoBond PC 500 Maxi kit (Macherey Nagel) to extract plasmid DNA from cultivated bacteria in the Terrific medium. After extraction, DNA was kept at -20°C. Restriction digestion or sequencing was used to verify the extracted DNA.

For library plasmids, overnight bacterial cultures were prepared in 96 deep well plates with 1.9 ml/ well of Terrific medium and Ampicillin. The culture plates were inoculated from the bacterial parents' plates using 96-well replicators (Sigma), then sealed with adhesive film and pierced in each well before being incubated at 37°C overnight. Wizard(R) SV 96 Plasmid DNA Purification System (Promega) was used to isolate the library plasmids according to the kit's protocol using Vac-Man 96 Vacuum Manifold (Promega). DNA plasmids were eluted with ddH2O 100ul/well, then plates were sealed and stored at -20°C.

b) DNA purification

Using QIAquick PCR Purification Kit (Qiagen), the amplified DNA fragments were purified from PCR reaction components according to the kit's protocol. The purified DNA was then subjected to restriction digestion and inserted into a vector. Similarly, the DNA fragments in agarose gels were extracted and purified from enzymes, salts, agarose, and ethidium bromide using QIAquick Gel Extraction Kit (Qiagen). The recovered DNA from gels was subjected to the ligation process.

3.2.1.9. DNA measurement

The Nanodrop-1000 (Thermo Fisher Scientific) was used to quantify DNA concentrations. The blank value was established with 1 µl ddH2O, and 1.5 µl of the sample was measured at automatically set wavelengths of 260 and 280 nm. The concentration of DNA was calculated automatically as well.

Measurement of Library DNA plasmids was done using Quant-iT PicoGreen dsDNA Assay Kit (Thermo Fisher Scientific) since the library plasmids were collected in 96-well plates. PicoGreen is a fluorescent probe that binds to the minor groves of dsDNA and forms a highly luminescent complex. Upon binding, its fluorescence increases >1000-fold, proportional to the quantity of dsDNA present (Dragan *et al.*, 2010; Singer *et al.*, 1997).

All samples were assayed at 200µl/well as the final volume in 96-well plates. Plasmid DNA samples were diluted 1:50 in 1x TE buffer by giving 2 µl of DNA sample to 98 µl of 1xTE buffer in each well of the reading microplate (Black/Clear Flat Bottom Microplate, Corning). Lambda DNA standard (100ng/µl) was diluted in 1xTE according to Table 3-24 to prepare 8 standards (A-H), then 100 µl/well of each standard was given to the reading microplate.

Table 3-24 Preparation of Diluted Lambda DNA standards

Vial	volume of diluent 1x TE (µl)	Volume and source of the DNA standard (µl)	Final DNA standard concentration ng/µl
A0	196	4 of stock	2
A	60	60 of vial A0	1
B	75	45 of vial A0	0,75
C	90	30 of vial A0	0,5
D	105	15 of vial A0	0,25
E	114	6 of vial A0	0,1
F	119	1,2 of vial A0	0,02
G	108	12 of vial F	0,002
H	100	0	0

100 µl of freshly diluted PicoGreen 1:200 in 1x TE buffer was given to each well, mixed, and incubated for 2-5 min at room temperature before being measured. Samples were excited at 480 nm, and the fluorescence intensity was measured at a wavelength of 520 nm using Infinite F200 Microplate reader (Tecan) and Magellan software.

3.2.1.10. qPCR-based product-enhanced reverse transcriptase (PERT) – Assay

PERT (F-PERT) assay is a retroviral titration method that depends on quantifying reverse transcriptase (RT) activity, the associated enzyme with all retroviral particles. Quantification is done by qPCR using cDNA-specific fluorescent-labeled probes (Taqman chemistry). During the reaction, the retroviral RT transcribes the template RNA into cDNA, which is then amplified and quantified in each cycle of PCR. RT activity is estimated based on the amount of the converted cDNA from an RNA template by the retroviral RT. The amount of synthesized cDNA

is proportional to the RT activity of the examined virus (Horie *et al.*, 2010; Pyra *et al.*, 1994; Vermeire *et al.*, 2012).

The assay was used to estimate the titration of library viral supernatants collected from 96-well plates. HIV-1 rRT Standard Lenti (Cavidi) was recruited as a standard in our assay. Twelve standard serial dilutions were prepared according to the Cavidi protocol but with DMEM complete medium (DMEM c) instead of the base buffer.

Table 3-25 Master mix preparation for qPCR of PERT assay

Component	***final conc./vol.***
Bioline Master mix (2x)	1x
RNase-OUT, 1:10 Fresh 4U/µl	0,02 U/µl
Horie2010_PERT-F	500 nM
Horie2010_PERT-R	500 nM
MS2-PERT_Horie-Probe (FAM)	250 nM
MS2 RNA Template	0,1 µl
Water	to 20 µl
Virus Lysate	9 µl

Assessed viruses and standards were lysed by mixing 5ul of each with 5ul of 2x lysis buffer contained RNase-OUT (Thermo Fisher Scientific) at concentration 1:50 in a pre-cooled 96-well plate, 5 µl of DMEM c was applied as a negative control.

Table 3-26 Thermal cycling profile of qPCR PERT assay.

Step		*Temp.*	*Time*
Reverse transcription		42°C	20 min.
Initial denaturation		95°C	10 min.
45 cycles	Denature	95°C	30 sec.
	Anneal	60°C	30 sec.
		Fluorescent detection	
	Extend	72°C	15 sec.

The plate was then incubated for 10 min at room temperature before adding 90 µl/well of ddH2O. Meanwhile, the reaction plate with master mix (Bioline), MS2 RNA template (Sigma), as well as specific MS2 primers and TaqMan probes (Table 3-16) (Horie *et al.*, 2010), was set up according to (Table 3-25), and 9 µl/well of the lysed viruses or standards were then added to the reaction plate. The qPCR test was conducted using the CFX96 Touch Real-Time System

(BioRad) using the parameters in Table 3-26, and the results were evaluated using the CFX Maestro software (Bio-Rad).

3.2.2. Cellular methods

3.2.2.1. Cell culture

All cell lines utilized in the study were adherent, maintained at 37°C and 5% CO2 in DMEMc or RPMI-1640c media, depending on the cell line (Table 3-10). The complete medium was supplied with 10% FBS, 2mM L-Glutamine, and 1% Penicillin/Streptomycin. The cells were passaged twice a week according to the cell line and its confluency in the used culture flask. To detach the cells, they were washed gently in PBS buffer and then treated with trypsin/EDTA for 1-5 minutes at 37°C. Trypsin neutralization was achieved by treating cells with complete media equal to the trypsin volume. The cells were then pipetted several times to prevent any aggregations. After that, some of the cells were further cultured according to the cell division rates, and the rest were discarded. However, when cells were required for many experiments, they were centrifuged for 5 minutes at 1400 rpm, resuspended in 10 ml of medium, and counted using Coulter Counter Z2 (Beckmann Coulter) or in a Neubauer counting chamber (Carl Roth) in which the following formula was used to quantify the cells:
Total cells/ml = (total cells counted/4) x Dilution factor x 10^4
They were sown in various culture vessels following the quantification based on their growth rates.

3.2.2.2. Cryopreservation of cell lines

The cells' working stock was prepared from the suspended cells in FBS at a 2 x 10^6–2 x10^7 cells/mL mixed 1:1 with a freezing medium at a final concentration of 10% DMSO in FBS. The cryotubes were immediately stored at -20°C for 1h before moving them to -80°C. For long-term storage, tubes were stored in liquid nitrogen.

3.2.2.3. Transfection

The transfection of HEK293T cells for protein overexpression or viral production was performed in several culture vessels, depending on the amount of protein or the required virus. PEI (polyethyleneimine) was used as a transfection reagent to introduce plasmid DNA into cell lines. Table 3-27 lists the various transfection formats with the number of seeded cells and the total DNA needed for transfection.

Table 3-27 HEK293T Cells seeding number and amount of transfected DNA according to the culture vessel

Culture Vessel	Seeding Density x10^5	Total DNA µg	Basal medium (ml)
10 cm^2	30	8	1
150 cm^2	75	25	5
6-well	5	2.5-3	0.2
96-well (F-Bottom)	0.4	0.300/0.563*	0.5
T-175	85	25	5
T-300	220	50	18

* ***Amount of total DNA needed for virus production in 96-well format***

The cells were transfected when the cells' confluency reached about 60-80%. The required plasmids were diluted in the basal medium at the corresponding amount mentioned in Table 3-27. For each transfection process, pMax- GFP plasmid was used to transfect HEK293T cells in 6 well plates to monitor the transfection efficiency. PEI was added as the last step at a ratio of 3:1 of the total DNA. Gently vortex was applied before incubation at room temperature for 15 min. After that, the transfection mixture was carefully given to the cells. Due to the toxicity of PEI to the cells, the medium was changed 16-18h post-transfection.

48h post-transfection, GFP fluorescence was observed under the fluorescent microscope (Nikon), and the expressed target protein was detected either in cell lysate (GP73 and library membrane proteins) or in cell lysates and supernatant (fusion proteins). In the case of producing fusion proteins or viral particles, the 48h-supernatants were collected and stored at 4°C, and the cells were treated with a fresh medium for an additional 24h. Later, both supernatants were pooled together and processed accordingly.

3.2.2.4. <u>Pseudotyped HIV viral particles production</u>

a) In order to produce pseudotyped-HIV luciferase reporter viruses, HEK293T cells were seeded in 150 cm^2 dishes or T-175 Flasks and transfected with total DNA as listed in Table 3-27. The transfection mixture contained a combination of the viral packaging plasmid psPAX2, the transgene pWPXL-Luc, and the viral envelope plasmid (Table 3-13) at a ratio of 2:4:1, respectively. PEI was used as in 3.2.2.3 at a ratio of 3:1 to the total DNA.

   The 48h- and 72h- post-transfection supernatants were collected, pooled, and centrifuged at 1400rpm to eliminate cell debris, then passed through a filter with a pore size of 0.45 µm (Sartorius). The clear supernatant was then added to clear tubes (Beckman Coulter) prefilled with 5 ml of 20% sucrose cushion and ultracentrifuged at

4°C on 30x10^3 rpm for 2:30h using SW32 Ti rotor (Beckman Coulter). After that, the tubes were decanted, and the virus pellets were resuspended in 100-150 µl/tube of basal RPMI-1640. After incubation on ice for 20-60 min, virus suspensions were collected and stored at -80°C for later use.

b) In order to produce HML-2 Env pseudovirus on a large scale for library screening, HEK293T cells were seeded in 6-12 x T-300 flasks. The transfection was carried out as described above and according to Table 3-27. The collected supernatants were processed similarly and subjected to ultracentrifuge without sucrose cushion at 18000 rpm at 4°C for 2h using a type-19 rotor (Beckman Coulter). After that, the viral pellets were resuspended in basal RPMI-1640 (5 ml/container) and incubated at 4°C overnight. Later, the viral suspensions were pooled and ultracentrifuged again in the presence of sucrose cushion using SW32Ti rotor as described in (3.2.2.4;a). The virus preparation was stored at -80°C.

c) For library screening, VSV-G pseudotyped HIV viruses carrying the library genes were produced in 96-well format by transient transfection according to Table 3-27. In a Costar 3896 plate, the virus production plasmids, including the packaging plasmid psPAX2 (175 ng) and the expression library clone DNA in pXL304 (300 ng), in addition to the envelope plasmid pCMV-VSV-G (87.5 ng), were mixed with the basal DMEM in a total 25 µl/well. PEI was diluted in a basal DMEM before 25 µl was given to each well. The plate was gently tapped and incubated at room temperature for 15 min. Meanwhile, 50 µl of the 150 µl medium/well from the cell plate was aspirated, followed by adding 50 µl of the transfection mix per well and incubation overnight at 37°C. The medium was changed 16-18h post-transfection with 110 µl fresh complete medium per well. The plates were centrifuged at 1400 rpm for 5 min at 4°C after 48h, and the viral supernatants were collected into a V-96-well plate (Thermo Fisher Scientific). After that, the plates were sealed with an adhesive foil and stored at -80°C until use.

3.2.2.5. Cells infection and transduction

a) All luciferase reporter viruses' infections were conducted in 96-well plates with 8 µg/ml of polybrene (Sigma). Before being used to infect the cells, The ultraconcentrated viruses were normalized using ELISA p24. The required virus titers were diluted to a final volume of 50 µl/well in a basal medium containing polybrene. The media was

removed from the cells before applying the infection mixture for 2h at 37°C. The infection mixture was then aspirated, and the cells were washed in PBS before adding 150-200 µl fresh complete medium per well. 72h post-infection, luciferase activity was measured.

b) For the infections in the presence of the soluble fusion proteins XSU-Fc or SU-FT-Fc, cells were first pretreated with the indicated concentrations of the soluble proteins for 1h at 23°C. Then, they were infected with the pseudotyped viruses for another 1h at 37°C. Subsequently, the cells were washed with PBS and maintained in a complete medium for 72h. In the case of XMRV infection, luciferase assay was performed after 24h -48h.

c) Infection of transduced cells with cDNA library pseudoviruses: cells were seeded in a 96-well plate at 15000 cells /well. The next day, 35 µl of the lentiviral cDNA library supernatants were mixed with 35 µl of basal RPMI-1640 with polybrene (11 ug/ml) in a V-96-well plate (Thermo Fisher Scientific). Then the medium was removed from the cells' wells and treated with 70 µl of the infection mixture for 1-2h. Later, 150 µl of complete medium was added per well. The plates were then incubated for 48h before being infected with HML-2 Env, VSV-G, or JSRV Env pseudotyped viruses according to the method described in (3.2.2.5; a).

3.2.2.6. Cell surface staining for flow cytometry analysis (FACS)

Cell surface staining was performed to detect the expressed transmembrane protein GP73 at the cell surface and the binding of the fusion proteins to the cell surface.

First, cells were seeded in 2 x 10 cm^2 plates for each cell line. The following day, one plate for each cell line was transfected with GP73 expression plasmid according to the transfection method (3.2.2.3), and the other plate was left as a control. 48h post-transfection, the medium was aspirated from the plates, washed with PBS, and harvested by enzymatic release using 1x Accutase solution (Thermo Fisher Scientific) at 37°C for 2-10 min depending on the cell line. The harvested cells were suspended in a complete medium, passed through a 40µm strainer (Corning), and counted.

A single-cell suspension at a concentration of 1-10 x 10^6 was prepared in a cold PBS. Fixable Viability Dye eFluor 780 (FVD) (Thermo Fisher Scientific) was incubated with cells at a concentration of 1:1000 for 30 min on ice and in the dark to stain dead cells and exclude them later from the flow cytometry analysis. The cells were then rinsed and resuspended in FACS

staining buffer (Table 3-7) before aliquoting 50µl per round-bottom polystyrene FACS tube (Carl Roth). The primary antibodies, including anti-Gp73 (Thermo Fisher Scientific) at a concentration of 1:100 and the fusion proteins XSU-hFc and SU-FT-hFc at concentrations of 13 µg and 30 µg per ml, were incubated with the cells on ice for 1h. Three rounds of centrifugation at 1400 rpm for 5 min/washing with FACS staining buffer were performed. Then cells were treated for 1h on ice with 1 µg/ml of the Alexa Fluor 488 labeled secondary antibodies, anti-h-IgG, and anti-m-IgG (Thermo Fisher Scientific). The control cell tubes were incubated with 1 µg/ml of the secondary anti-h-IgG alone or with mouse IgG 1 Kappa isotype (Thermo Fisher Scientific). Cells were rewashed three times and then fixed with the FACS Fixation buffer (Table 3-7) at a final concentration of 0.5% paraformaldehyde (PFA). Finally, the samples were measured and analyzed using FACSCalibur and FlowJo software (Becton-Dickinson).

3.2.2.7. Cell treatment with heparinase I/III

Cells were grown in 10 cm^2 plates at standard cultivation conditions till they reached 80-90% confluency. They were then harvested by enzymatic release with Accutase (Thermo Fisher Scientific) followed by two washing rounds with PBS. Heparinase I and III -blend (Sigma) was diluted with Heparinase reaction buffer (Table 3-7) at a final concentration of 5 U/ml, in which the cells were resuspended and incubated for 1h at 37°C. After that, the solution was removed, and the cells were washed with PBS, then counted and subjected to the FACS staining protocol (3.2.2.6).

3.2.3. Protein analysis methods

3.2.3.1. Cell lysis

We prepared lysates of transiently transfected cells to detect the expressed proteins by western blot. The medium was aspirated 48h after transfection, and the cells were washed with cold PBS. Then, NP-40 lysis buffer with a protease inhibitor (Roche) (Table 3-7) was added to the cells and incubated for 10 min on ice. The cell lysates were then collected into a 1.5 ml tube and centrifuged at 4000 rpm for 10 min at 4°C. The supernatant was carefully transferred into a new 1.5 ml tube and stored at -20°C.

3.2.3.2. Transmembrane protein extraction

Membrane proteins were extracted from cells as described by Mei and coworkers (Mei *et al.*, 2015). Different cell lines were grown in 10cm^2 dishes till 90-100% confluency. After removing

the medium, the cells were washed with a cold PBS before being treated with 1 ml of homogenizing buffer (Table 3-7) for 10 min on ice. Cells were then collected using scrapers into 1.5 ml tubes and centrifuged at 6000g and 4°C for 5 min. The supernatants that included the membrane proteins were collected into new tubes and re-centrifuged at 13200g for 60 min at 4°C. The membrane protein pellets were resuspended in the homogenizing buffer and stored at -80°C.

3.2.3.3. Fusion protein purification

The fusion proteins were produced by transient transfection of HEK293T cells in T-175 Flasks. For Fc-tagged proteins, the cells were maintained in a complete DMEM including IgG Stripped-FBS, starting from 18h post-transfection. 48h post-transfection, the supernatant was collected, centrifuged at 1200 rpm for 10 min at 4°C, passed through a 0.45 μm filter, and diluted in PBS 1:3 or 1:2 (v:v of PBS). Meanwhile, Pump–P1 (Amersham biosciences) was prepared by rinsing the tubing system with PBS, then attaching the HiTrap column (GE Healthcare) to the pump, setting the dropping system (drop to drop), and washing the column with PBS before applying the sample. After that, the column was rewashed until no more material appeared in the effluent. The protein was eluted by 2 to 5 column volumes of elution buffer (Table 3-7) in a 50 ml tube prefilled with 2 ml of the collecting buffer (Table 3-7). Next, the buffer was exchanged with PBS using Amicon Ultra-15 Centrifugal Filter Device (Millipore). The filter tube was rinsed first with PBS and centrifuged for 30 min at 4000 rpm using a swing rotor to eliminate any trace of glycerin before adding 12.5 ml of the purified protein sample to the filter device drop by drop and centrifuging it at 4000 rpm for 30 min. The last step included adding 12.5 ml of PBS and then centrifuging again. The concentrated protein sample (200-400μl) was collected into a 1.5 ml tube, stabilized with sucrose (5%), aliquoted, and stored at -20°C.

For the V5-tagged proteins, the V5-tagged protein purification kit Ver.2 (MBL) was used according to the manufacturer's instructions.

Using the BCA assay, the protein concentration was measured.

3.2.3.4. Measurement of protein concentration

Pierce BCA Protein Assay Kit (Thermo Fisher Scientific) measures the protein concentration of the whole cell lysate and the purified protein samples. The manufacturer's microplate protocol was followed, and measurement was conducted using the microplate spectrophotometer Multiskan GO (Thermo Fisher Scientific).

3.2.3.5. Protein deglycosylation

In order to remove all N-linked glycans from the glycoproteins, the PNGase F enzyme (N-glycosidase F, from *Flavobacterium meningosepticum*) was recruited. The enzyme specifically removes glycosylations on asparagine side chains. We conducted the deglycosylation reaction under non-denaturing conditions on the purified fusion proteins according to the manufacturer's instructions (NEB). The deglycosylated proteins were then subjected to SDS-PAGE.

3.2.3.6. Polyacrylamide gel electrophoresis (PAGE)

Proteins were separated by molecular weight using one-dimensional polyacrylamide gel electrophoresis (PAGE), either under denaturing conditions with Sodium Dodecyl Sulfate (SDS-PAGE) or non-denaturing conditions (Native), in which neither the sample nor the electrophoretic buffer contained any SDS. For SDS-PAGE, the sample was treated with an equal volume of 2x SDS sample loading buffer (Sigma), containing 2% SDS and 5% β-mercaptoethanol, then heated at 70°C for 10 min using a heated block (Eppendorf). Under non-denaturing conditions (Native-PAGE), the samples were loaded directly into the gel after mixing with 2x Tris-Glycine Native Sample Buffer (Table 3-7), and no heating was applied.
The concentration of the used polyacrylamide gels depended on the protein size. Gels for fusion protein separation were either hand-casted at 7.5% for separation gel and 5% for stacking gel (Table 3-28) or ready pre-casted 4–15% gradient gels (Bio-Rad). However, 4-20% gradient gels (Bio-Rad) were used to separate other types of protein samples.
The Samples were separated in the stacking gel at 100 V for 10 min, then the voltage was increased to 160-180 V for 45-60 min. PageRuler Prestained Protein Ladder (Thermo Fisher Scientific), and Precision Plus Protein Dual Color Standards (Bio-Rad), in addition to Native Marker (SERVA), were used. After electrophoresis, the gels were either stained with silver (sigma) or Coomassie stain (SERVA) to visualize the protein bands or subjected to Western blot.

Table 3-28 Recipe for preparing hand-casted gels.

Components	2x Separation gel 7.5 %	2x Stacking gel 5%
ddH2O	9.8 ml	5,7 ml
30% acrylamide	5 ml	1,7 ml
Resolving Buffer, 1.5M Tris-HCl pH 8.8	5 ml	-
Stacking Buffer, 0.5 M Tris-HCl pH 6.8	-	2,5 ml
10% SDS	100 µl	100 µl
10% APS	50 µl	50 µl
TEMED	10 µl	10 µl

3.2.3.7. Western blot

The separated proteins by PAGE (SDS or native) were electrophoretically transferred to a 0.45 mm nitrocellulose membrane (Bio-Rad) using the wet transfer method (Bio-Rad) at 90 V for 30-50 min at 4°C. Then, the membrane was incubated overnight on an agitation (GFL) with WB blocking buffer (Table 3-7) at 4°C. Before adding the antibodies, the membrane was washed thrice with WB washing buffer (Table 3-7) for 5min each. The primary antibody was diluted in the washing buffer and incubated with the membrane overnight at 4°C. Table 3-9 shows all antibodies utilized in this study and their working dilution. Three washing rounds were required to remove the excess antibodies before incubation with peroxidase-conjugated secondary antibodies for 1h at room temperature. Following the manufacturer's directions, three rounds of washing were performed before detecting the blotted proteins using an Enhanced Chemiluminescence (ECL) reaction using SuperSignal West Pico PLUS Chemiluminescent Substrate (Thermo Fisher Scientific). The emitted light resulting from this reaction was captured and visualized on an X-ray film (CL-XPosure Film, Thermo Fisher Scientific) after developing it using the Curix 60 Film processor (AGFA).

Stripping membrane, when needed, was done using stripping buffer (Table 3-7) according to the manufacturer's instruction, followed by repeating the steps mentioned above, starting from the blocking step.

Dot blot was performed by directly applying 5-7µl of lysates on a 0.45 mm nitrocellulose membrane. The membrane was then let dry at RT for 5-10 min. It was then blocked, treated with antibodies, and processed as described above.

3.2.3.8. Silver staining and mass spectrometry analysis

In order to visualize the separated proteins on the gels, silver staining using the ProteoSilver Plus kit (Sigma) was performed after the SDS-PAGE and according to the manufacturer's protocol. The silver-stained protein bands were identified using the mass spectrometry analysis by (Proteome Factory AG).

3.2.3.9. GAG-Binding assay

The protein-GAG interaction was assessed by using heparin agarose beads (Sigma) as described by the reference (Robinson-McCarthy *et al.*, 2018). 50 µl of beads (heparin-beads, protein G-beads, streptavidin-beads or anti-V5-agarose beads (Sigma)) were washed before incubation with either 10 ng of concentrated pseudotyped viruses in TNE (Table 3-7), 10 µg of cell lysate in TNE buffer, or 3 µg of purified fusion proteins in TN buffer (Table 3-7). The incubation was

done at 4°C for 2h with end-to-end mixing in the presence and absence of 50 µg/ml of soluble heparin (Sigma) or heparan sulfate (AMSBIO). The beads were washed in the incubation buffer five times, and the bound proteins were eluted at 70°C for 5 min with 2x SDS sample buffer. Samples were analyzed by SDS-PAGE and Western blot using specific antibodies.

3.2.3.10. Affinity isolation and receptor pull-down assay

A modified protocol from (Mei *et al.*, 2015; Raj *et al.*, 2013) was applied to pull down the potential receptor of HML-2 using its fusion protein. 15 µg/ml of the purified HML-2 fusion protein (SU-FT-hFc) or XMRV (XSU-hFc) were incubated with 50µl of membrane proteins extracts (3.2.3.2) from different cell lines in homogenizing buffer (Table 3-7) at 4°C on end-to-end mixing overnight. Next, protein G Sepharose (Sigma) was added for 2h at 4°C. After thrice washes using homogenizing buffer, the precipitated proteins were treated with a 2x SDS sample loading buffer (Sigma) at 50°C for 5 min and subjected to SDS-PAGE. After that, the gels were stained using a Silver Stain Kit (Sigma) and analyzed with mass spectrometry. Alternatively, they were analyzed by blotting on a membrane using specific antibodies, anti-h-IgG and anti-XPR1.

3.2.3.11. Luciferase assay

In order to assess luciferase pseudoviral entries, luciferase reporter activity was measured after 72h of infection with the pseudoviruses carrying the luciferase reporter gene. The assay was performed using the firefly luciferase assay system (Promega). The medium was first aspirated, and the cells were rinsed with PBS before incubating them with 20 µl/well of 1x Luciferase Cell Culture Lysis Reagent (CCRL) (Promega) for 10 min at room temperature. The plate was tapped from the sides during the incubation to ensure that all cells were detached. Next, the plate was centrifuged at 4000 rpm for 10 min at 4°C.

Meanwhile, the Luciferase Assay Reagent (Promega) was mixed with the Substrate (Promega) and kept at RT till measurement. 10 µl of cell lysate was transferred from each well to the measurement plate (Thermo nunclon white plate). The luminometer Lumistar Omega (BMG LABTECH) was recruited for the measurement using the inject-then-read mode. The injection volume was 50ul/well, the delay time was 2 sec, and the read time was 10 sec. MARS Data Analysis Software (BMG LABTECH) analyzed the data.

3.2.3.12. p24 ELISA

The p24 enzyme-linked immunosorbent assay (ELISA) is a method to determine the lentiviral titer using p24 capture antibodies, which bind the HIV-1 capsid protein 24. In this assay, a 96-

well microtiter plate Nunc MaxiSorp (Thermo Fisher Scientific) was incubated overnight at 4°C with 50 µl/well of the monoclonal AG3.0 antibody diluted 1:800 in the carbonate bicarbonate buffer (Sigma). The wells were then washed thrice with PT buffer (Table 3-7) before being blocked for 1h at 37°C with PM buffer (Table 3-7).

Meanwhile, the virus samples were inactivated for 10-30 min at room temperature with 0.02% Tween-20. Next, the wells were rinsed with PT buffer before the inactivated virus samples were applied to the first well of each row. Then, in the PMT buffer (Table 3-7), eleven serial dilutions in each row were produced at a 1:2 ratio in 50 µl/well as the total volume and incubated for 1h at 37°C. The wells were washed three times and incubated for 1h at 37°C with 50 µl/well of HIV-1 plasma pool diluted 1: 10000 in PMT buffer.

Similarly, the secondary antibody anti-h-IgG conjugated with HRP diluted at 1:1000 in PMT buffer was given to the wells. Next, the wells were washed five times before adding 50 µl/well of the substrate for 10–15 min at room temperature. Then, 25 µl/well of the sulfuric acid 30% was added to stop the reaction. The plate was then read at 492 nm / 620 nm wavelength using the microplate reader Sunrise (Tecan).

3.2.4. Statistics

The statistical analysis was conducted in GraphPad Prism software version 9 (Table 3-3), using unpaired two-tailed Student's t-tests. The Spearman test was used for correlation analysis. The statistical significance was set as $p < 0.05$.

4. Results

4.1. Mutational analysis of a putative HS-binding domain in HML-2 Env

One good strategy for most viruses to increase their chances of attachment to the cell surface is binding to GAGs. GAGs are abundant components of both the cellular matrix and plasma membrane. Interactions with them would increase the concentration of viral particles at the cellular membrane and subsequently increase the probability of encountering a specific entry receptor to infect the cell (Cagno *et al.*, 2019).
The ability for HML-2 Env to interact with GAGs, particularly HS, and there by, to enter cells had been demonstrated without identifying the involved domain on HML-2 Env in this interaction (Geppert, 2019; Robinson-McCarthy *et al.*, 2018).

4.1.1. Binding K113 Envs mutated in the presumed RBS to the HP-beads

As we mentioned earlier, the HS role is indispensable in HML-2 Env entry and acts as an attachment factor to it (Geppert, 2019; Robinson-McCarthy *et al.*, 2018). Since the RBS of HML-2 has been identified in recent years (Bannert, 2017; Wamara, 2020), we investigated whether HML-2 uses the RBS for binding HS or involves another region of its Env. The RBS indicated in green (Figure 4-4) has been determined by mutational analysis at the N-terminal region of the Env extended between residues N_{128} to N_{153}. The mutations (Figure 1-9) were introduced individually in the K113 oricoEnvΔC1 expression plasmids: (R140C, R140A, N128A, N153A, EEE146-148 AAA, S138A, S155A, G135A, P145, and D138A. In addition to a deletion Δ144-152), they were shown to prevent HML-2 Env-mediated entry (Priesnitz, 2019; Wamara, 2020).

First, we produced HIV pseudotyped luciferase reporter viruses with RBS-mutated envelopes and ultraconcentrated them according to the methods presented earlier (3.2.2.4). Then, we analyzed them by Western blot to detect the Env and Gag. The processed viruses showed the typical bands of the Env and p24 Gag proteins. Mutations that affected the N-linked glycosylation sites showed a weak signal on the precursor bands (Figure 4-1). However, we could not detect the Env on the blot of the virus with Δ144-152 Env, even though the Gag proteins were detected. The results indicated that all the mutated Envs were incorporated into viral particles except Δ144-152 Env, which was excluded from the further analysis.

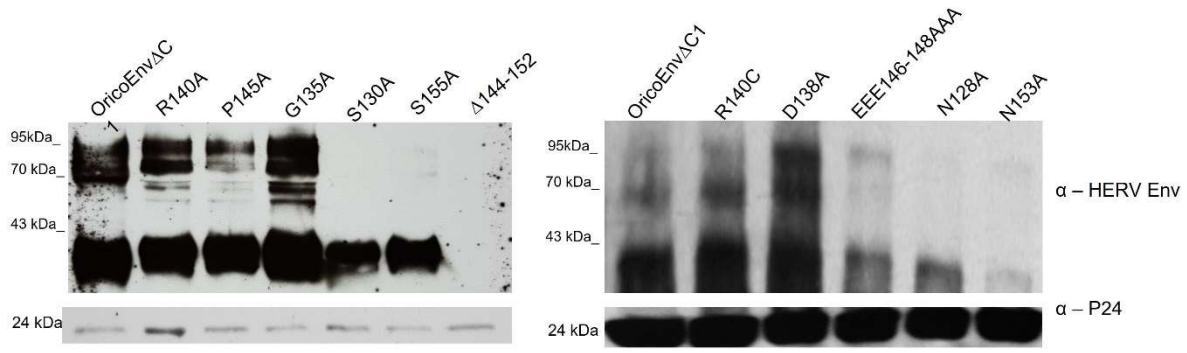

Figure 4-1 The incorporation of the K113 oricoEnvΔC1 RBS mutated Envs into viral particles.
Western blot analysis of HIV pseudotyped reporter viruses carrying mutated RBS envelopes. The viruses were ultraconcentrated through a sucrose cushion and normalized by p24 ELISA. 10 ng of p24/each virus was subject to Western blotting using specific antibodies: anti-HERV-Env, and anti- p24.

In the next step, we tested the mutated Envs for HP-binding. An equal concentration of normalized pseudotyped viruses with each of the mutated Envs was incubated with HP-coated beads in the presence and absence of soluble HP as a pre-incubation step. The wild-type Env (K113 oricoEnvΔC1) and Vesicular Stomatitis Virus envelope glycoprotein (VSV-G) pseudoviruses were used as controls. Additionally, we incubated viruses with protein G-beads as a binding negative control. The unbound viruses were washed off, and the bound viruses were then eluted and analyzed by Western blot using anti-HERV-Env.

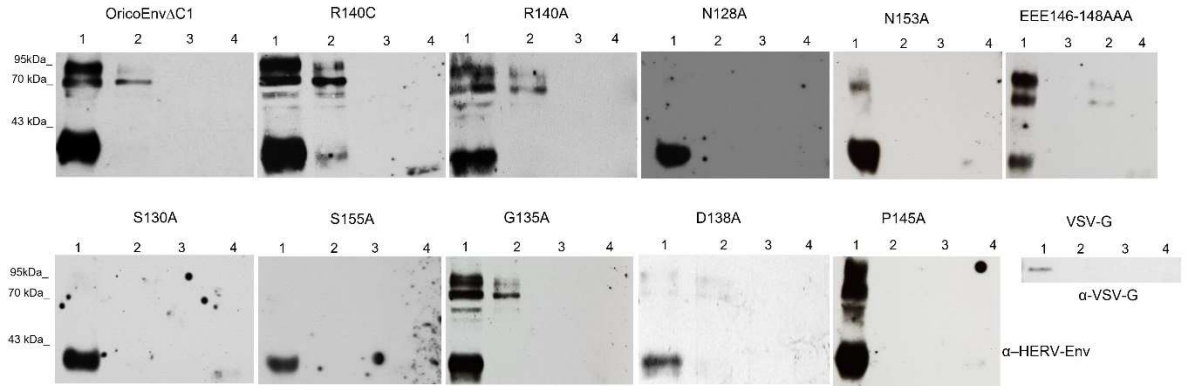

Figure 4-2 Western blot analysis of the pulled-down K113 RBS mutated Env viruses by HP-beads.
Normalized HIV pseudotyped viruses with K113 RBS mutated envelopes, wild-type Env and VSV-G were incubated with HP-beads only (2), HP-beads with soluble HP (3) added as a competitor, or with protein G (4); (1) virus only: 10% of total input viruses. The bound viruses were eluted from the beads and subjected to SDS-PAGE, followed by Western blotting using specific antibodies against HERV-Env and VSV-G. Data is a representative experiment from two independent experiments.

Results showed that Envs with N-linked glycosylation site mutations (N128A, N153A S138A, and S155A) had not been detected when incubated with HP-beads (Figure 4-2). In contrast, the

other mutations appeared to bind HP-beads similarly to the control. The binding (when found) was abolished upon incubation with soluble HP, and in all cases, protein G-beads showed no unspecific binding.

These results disproved the implication of the RBS in binding HP since all RBS mutated Envs bound HP-beads. The Envs with the mutations N128A, N153A S138A, and S155A demonstrated full envelope cleavage enabling us to detect the TM proteins in the first lanes of their blots (Figure 4-2). This suggests that HP-beads bound the SU, which was undetectable by the TM-specific antibody used. In order to prove this assumption, a version of K113 oricoEnv mutated in all its predicted glycosylation sites (K113 oricoEnvGly$^-$) was used to transfect HEK293T cells. Since this mutated Env does not incorporate into viral particles (Hanke *et al.*, 2009), the cell lysates were incubated instead with the HP-beads, protein A-beads as a negative control, and V5-beads as a binding positive control since the envelope proteins used are V5-tagged.

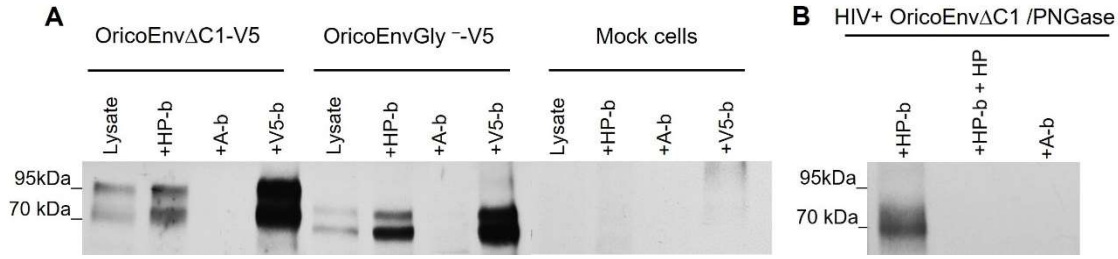

Figure 4-3 pull-down assay for the non-glycosylated Envs.

***A.** Western blot analysis of the HEK293Tcell lysates expressing K113 oricoEnvΔC1-V5 and K113 oricoEnvGly$^-$-V5, or mock subjected to pull-down assay by HP-beads, protein A-beads, and V5-beads. Anti-V5-tag was used; lysate lane: 5% of the total input; +HP-b= with HP-beads; +A-b= with protein A-beads, +V5-b= with V5-beads. Mock cells= non-transfected cells. **B.** Western blot analysis of the pulled-down K113 oricoEnvΔC1 pseudotyped viruses treated with PNGase F; +HP-b+ HP= with HP-beads and soluble HP. A representative experiment of two independent experiments is shown.*

Similar steps were followed as previously described, and Western blot analysis against HERV-Env detected the non-glycosylated Env by HP-beads at the expected size (Figure 4-3; A). Pull-down K113 oricoEnvΔC1 pseudotyped viruses treated with PNGase F enzyme to remove N-linked glycans yielded similar findings (Figure 4-3; B). This means that K113 RBS was not used to bind HP, and there is another site involved in HS binding.

The results demonstrate that neither the HML-2 RBS nor the glycans on the proteins are used to bind HP. There has to be another site in the SU subunit involved in HS binding.

4.1.2. Mutational analysis of the putative HML-2 HBD

To identify the domain of HML-2 Env that interacts with HS, we used an approach similar to one previously used to identify HML-2 RBS (Wamara, 2020). We aligned the protein sequence of K113 Env as a reference for HML-2 in a codon-optimized version (Hanke *et al.*, 2009) to the sequence of the MMTV Env (UniProt accession number: Q85646), which has strong homology with K113 (Figure 4-4). The MMTV Env interacts with HS through a known HBD (Willer *et al.*, 1997; Zhang *et al.*, 2003). The latter shows the mammalian consensus sequences of an HBD motif "XBBBXXBX" where X represents any aa and B a basic aa (Zhang *et al.*, 2003). The linear alignment allowed us to identify a region on K113 Env corresponding to the MMTV HBD rich in positively charged residues, suggesting that it might also be involved in binding HS. However, unlike the MMTV HBD, the sequence of the corresponding region "Q_{215}RSLKFRPKGKPCPKEIPKESK$_{236}$" showed the following pattern "XBXXBXBXBXBXXXBXXXBXXB" that did not resemble mammalian consensus sequences of HBDs. Nonetheless, it showed frequent clusters of basic amino acids (B) spaced with non-basic residues (X), which could still be considered a common pattern in HBDs according to (Fromm *et al.*, 1997). The putative domain, however, included 8 basic aa, mostly Lys (K= 6) and less Arg (R=2), which may suggest lower affinity in binding HS since R residues in HBDs bind more strongly to sulfate groups than do K residues (Fromm *et al.*, 1997). When comparing the predicted three-dimensional model of both Envs, we found that the putative HBD of HML-2 Env mapped mostly to a predicted β-strand motif. Only the two basic residues of the domain R_{216} and K_{219} were mapped to an α-helix motif. However, the modeling confidence in this region is low (Figure 4-5) (Jumper *et al.*, 2021; Varadi *et al.*, 2022). In contrast, the MMTV HBD forms an α-helix motif (Zhang *et al.*, 2003).

To determine whether the positively charged residues in the putative HML-2 HBD had effects on both HP-binding and infectivity, we performed site-directed mutagenesis on the N-terminal region of SU of the K113 oricoEnvΔC1. The altered tandemly positioned positively charged amino acids (arginine and lysine) were mutated to alanine neutrally charged residue so that the whole region was mutated by double overlapped mutations (Figure 4-6). The double mutations were generated using one primer when the two targeted sites were adjacent. However, the mutagenesis was performed in two sequential steps using two primers when the two targeted sites were distant.

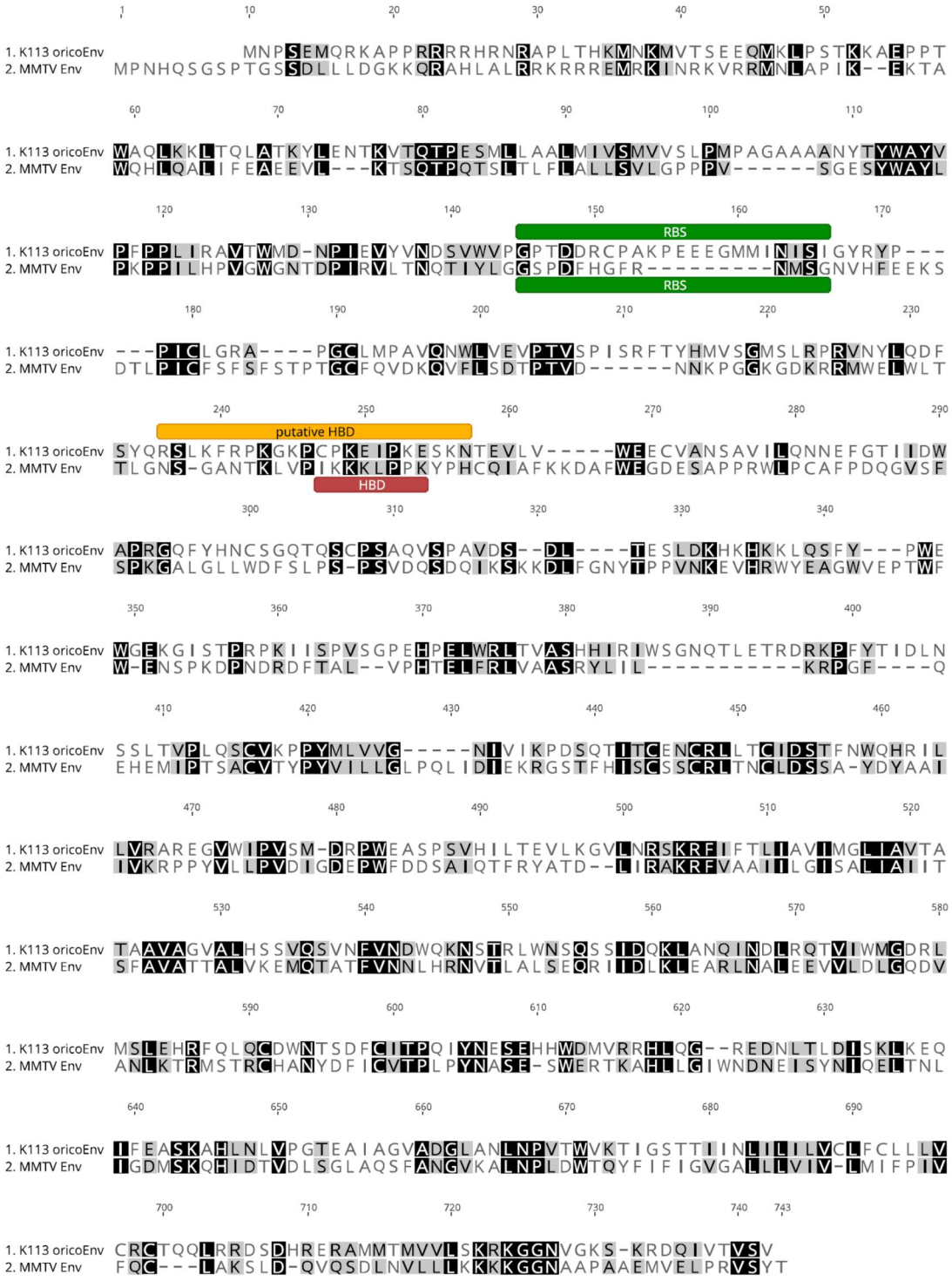

Figure 4-4 Alignment of K113 and MMTV Envs sequences.

The black shaded residues are identical; the gray shaded are similar. The RBSs for both viruses are marked by green annotation above and under the sequences. The MMTV HBD (IKKKLPPK) is marked by red annotation under the sequence, and the corresponding region to MMTV HBD at K113 Env is marked with yellow as the putative HBD. The RBS and HBD of MMTV are positioned according to (Zhang, 2003). (Geneious version 2021.2 created by Biomatters).

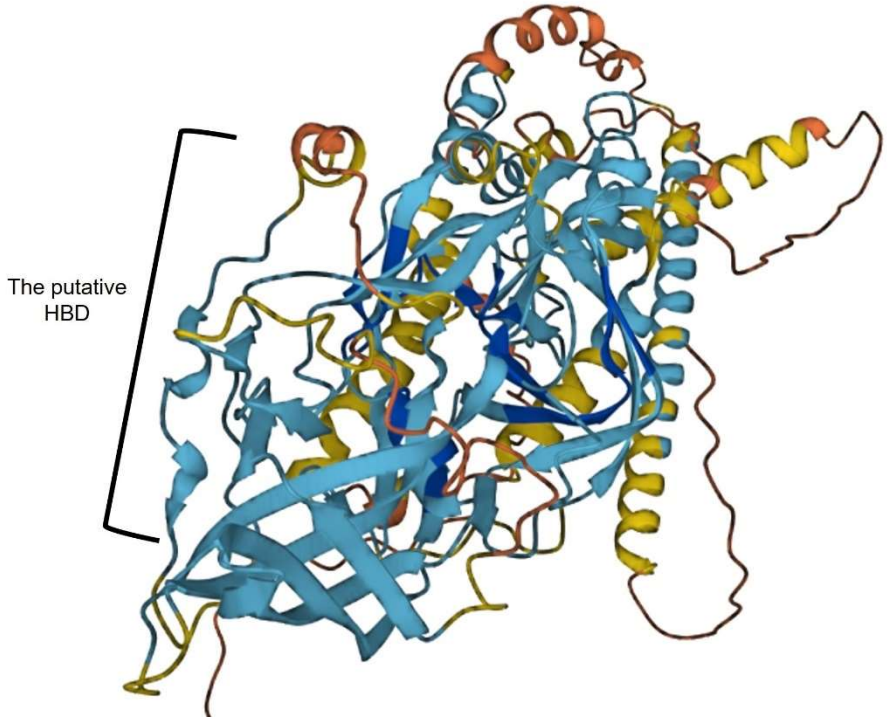

Figure 4-5 Three-dimensional model of K113 Env.
The putative HBD is mapped mostly to a β-strand motif. The 3D structure is predicted by AlphaFold Protein Structure Database. A model confidence score (pLDDT) between 0 and 100 was produced by AlphaFold. The dark blue=(pLDDT > 90) Very high; the blue= (90 > pLDDT > 70); the yellow= low confidence (70 > pLDDT > 50); the orange= Very low (pLDDT < 50) the regions below 50 pLDDT may be unstructured in isolation (Jumper et al., 2021; Varadi et al., 2022). The model is adapted from (AlphaFold Protein Structure Database).

The first mutagenesis steps resulted in four Envs with a single site mutation, and we included them in this study. In addition, we mutagenized all arginine and lysine sites of HBD to alanine in a one Env expression plasmid, which was mutated using different primers in sequential steps (Table 3-15), making the total number of HBD mutated Env expression plasmids eleven. The eleven mutant plasmids were then sequenced using the primers listed in Table 3-17 to confirm the presence of the designated mutations. Sequencing data for them were as expected in all clones.

To check the expression of the mutated Envs, we transfected HEK293T cells with the mutated envelope expression plasmids, K113 oricoEnvΔC1, and the empty vector pcDNA as controls. Lysates were then collected 48h post-transfection and analyzed by Western blot using a specific HERV-Env antibody. All the clones that carry the single site or the double-site mutations were expressed at similar levels to that of K113 oricoEnvΔC1, and all showed both the glycosylated precursor at 95kDa and TM at 38kDa. The fully mutated HBD Env showed the precursor at the

expected size. However, the TM band was faint, suggesting that the mutations prevented the efficient cleavage of the precursor into the two subunits SU and TM (Figure 4-7; A).

Next, we investigated whether the eleven mutated HBD Envs were able to incorporate into lentiviral particles. We transfected HEK293T cells to produce HIV pseudotyped luciferase reporter viruses with the mutated K113 HBD Envs, K113 oricoEnvΔC1, and VSV-G.

Figure 4-6 Generation of K113-HBD mutations.
K113 Env is represented on the top with positions of the potential glycosylation sites (NXS/T) according to (Hanke et al., 2009). The SU glycoprotein comprises both the receptor-binding site (RBS) and the putative HP-binding domain (HBD) downstream of the RBS. The positively charged residues in the HBD are indicated in red. The mutations shown in blue were introduced as single or double sites in ten clones out of eleven. All arginine and lysine residues were mutated to alanine in the last clone.

The virus supernatants were harvested, ultraconcentrated through a sucrose cushion, and normalized by p24 ELISA. Western blot analysis was performed on virus pellets, and viral incorporation was assessed by detecting the Env and Gag proteins using anti- HERV-Env and p24 antibodies (Figure 4-7; B). Almost all mutant Envs were incorporated into viral particles with similar efficiency as wild-type Envs, except for viruses bearing the completely mutated

HBD Envs, which had a low degree of incorporation. Gag proteins were, however, comparable to those of the wild-type Env virus.

Together, these results demonstrate that the HBD single-site and double-site mutations did not affect Env expression or incorporation into viral particles. However, when gathered in one fully mutated domain, these same mutations influenced the Env cleavage efficiency, leading to a lower level of viral incorporation.

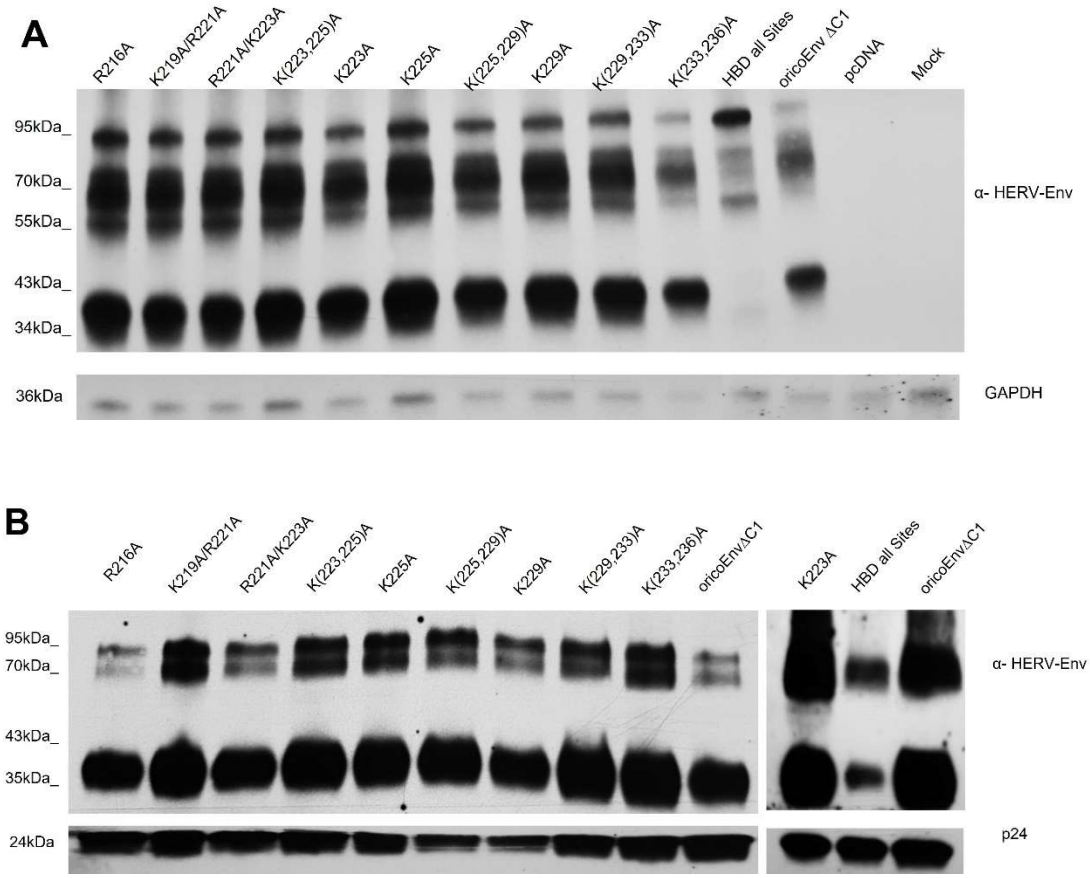

Figure 4-7 Expression of the mutated HBD Envs in HEK293T cells and incorporation into lentiviral particles. ***A.*** *Western blot analysis of the lysates of transfected HEK293T cells with the expression plasmids of mutated HBD Envs, the wild-type Env, and pcDNA as a negative control.* ***B.*** *Western blot analysis of HIV pseudotyped luciferase reporter viruses carry the mutated HBD Envs. We subjected 10 ng of p24/each virus to Western blot analysis using specific antibodies to detect the Envs (by anti-HERV-Env) and the Gag (by anti-p24).*

4.1.3. Binding of mutated HBD envelopes to HP-beads

To determine whether the corresponding MMTV HBD on K113 Env functions as the virus HS-binding domain, we recruited the HP-coated beads to study this interaction, similarly to what had been done previously (section 4.1.1). Equal concentrations of normalized pseudotyped

viruses with the mutated Envs were incubated with HP-beads in the presence and the absence of soluble HP as a pre-incubation step. The wild-type Env and VSV-G pseudoviruses were used as controls. The protein G-beads were used as a binding negative control. The unbound viruses were washed off. Then viruses were eluted and analyzed by Western blot using anti-HERV-Env (Figure 4-8).

Most of the analyzed mutations in the putative HBD reacted with the same degree of affinity to HP. However, HBD mutant Env (HBD all sites) interacted with less affinity to HP-beads, indicating that this domain is the main site to bind HS on the cell surface. Since we could detect a trace of interaction with HP, we hypothesized the implication of another secondary domain in the interaction. Nevertheless, the double mutation K(223,225)A failed to bind HP-beads. In contrast, the Env with single-site mutation K225A did bind to HS. It presumably retains the ability to bind because the "K" is at the nearby position 223. However, the Env with the single mutation K223A showed a reduction in affinity to HP compared to the control and other mutations, suggesting that this residue is essential in binding HS on the cell surface. The two sites together affect the Env-HS interaction.

On the other hand, other analyzed mutations in K229, K233, and K236 reacted oppositely. They showed stronger binding to HP compared to the wild-type Env.

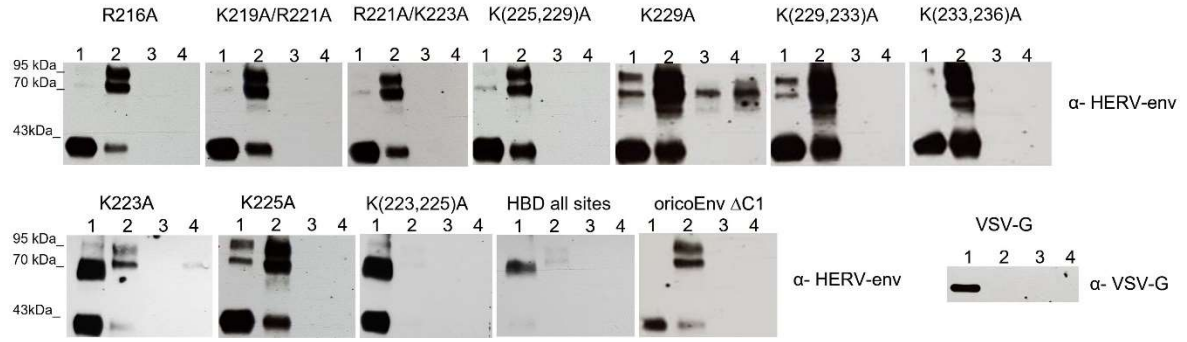

Figure 4-8 Western blot analysis of the pulled-down K113 HBD mutated Env viruses by HP-beads.
The normalized HIV pseudotyped viruses with HBD mutated Envs, wild-type Env, or VSV-G were incubated with HP-beads only (2), HP-beads with soluble HP (3) added as a competitor, or with protein G (4); (1) virus only:10% of the total input viruses in the pull-down assay. The bound viruses were analyzed by Western blot using specific antibodies for HERV-Env and VSV-G. Data show a representative experiment from two independent experiments.

Furthermore, the binding to HP-beads (when detected) was abolished in the presence of soluble HP, which functioned as a competitor to the beads. The binding was specific to HP-beads but not protein G-beads in all tested viruses, except for the virus with K229A Env, which bound to

protein G- in addition to HP-beads. Other double mutations with the same site K(225,229)A or K(229,233)A had different behavior, suggesting that the single-site mutation K229A may allow the virus to bind to the cell surface more robustly.

4.1.4. Entry assay of the pseudotyped viruses with HBD mutated Envs:

To further investigate the influence of the HBD mutations on the virus infectivity (the level of the virus entry), we infected HEK293T cells with equal concentrations of normalized HIV luciferase reporter viruses carry HBD mutated Envs, the wild-type Env, the non-functional Env R140C as a negative control and VSV-G as positive infection control. 72h post-infection, we collected cell lysates and measured the luciferase activity.

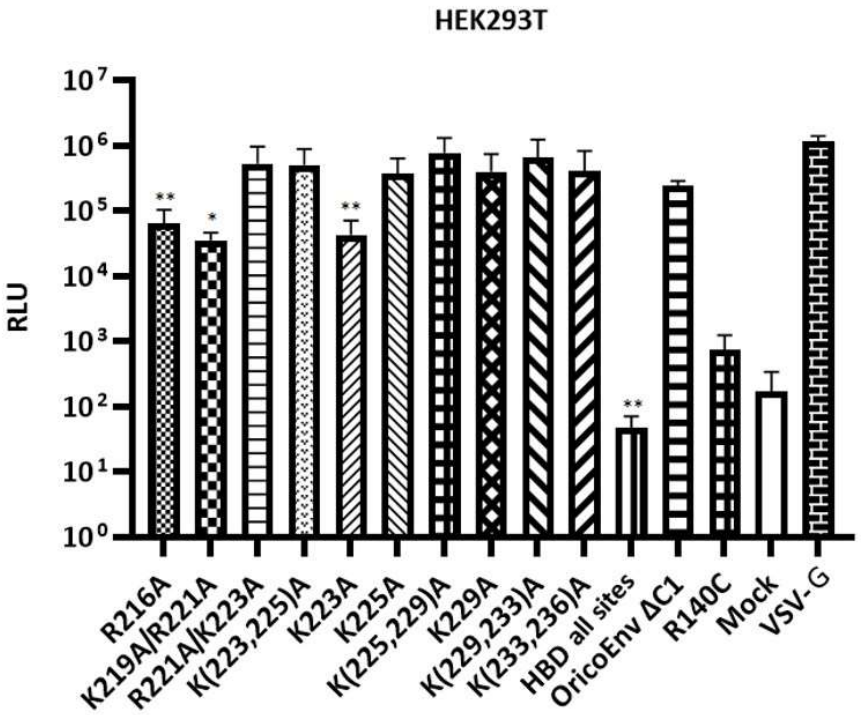

Figure 4-9 Entry of lentiviral particles pseudotyped with HBD mutated Envs.
HEK293T cells were infected with 100ng p24 of the normalized viral particles carrying HBD mutated Envs, the wild-type Env, or R140C Env as a negative control. 1ng of VSV-G virus was used to infect the cells. Luciferase activities were measured. The background level was set according to the negative control entry.
*Mock: non-infected cells. Each bar is the mean of at least three replicates from two independent experiments.(error bars=SD). *p < 0.05, ** p < 0.01*

Results showed that the viral particles with HBD all sites mutated Env failed to enter the cell completely ($p = 0.009$) due to lacking the ability to bind HP. However, the incorporation aspect might contribute to this inhibition (Figure 4-9). The Env with the double mutations

K(223,225)A, which failed to bind HP-beads earlier, allowed an efficient entry similar to the wild-type Env, suggesting that other cell surface glycans compensated for virus binding. Envs with the single mutation R216A and the double sites mutations K219/R221A showed a significant reduction in K113 entry-level (p = 0.004 and p = 0.01). We also observed a significant reduction in the K223A Env entry-level (p = 0.003). This mutated Env showed earlier a lower affinity to HP than other mutated Envs. The other mutations in the putative HBD domain allowed efficient viral entry and were at comparable levels to the control Env.

These results confirmed the implication of the putative HBD corresponding MMTV HBD in binding HS and showed that arginine residues, especially R_{216} and lysine residues K_{219} and K_{223}, were influential in HS binding, and alteration in these residues affected the viral entry. However, mutations of all HBD (HBD all sites) or most of the basic residues in this domain are required to prevent viral entry.

4.2. Affinity isolation approach using fusion glycoprotein

In order to find the cell surface molecule or protein that serves as a receptor for K113 (HML-2), we utilized the typical approach of affinity isolation that is widely used to determine the receptors of different viruses, including the Middle Eastern respiratory syndrome coronavirus, avian leukosis virus subgroup J, and Japanese encephalitis virus (Mei *et al.*, 2015; Mukherjee *et al.*, 2018; Raj *et al.*, 2013).

The principle of this method is to design a fusion protein that consists of the viral spike protein-coding gene fused with a tag. It is then cloned into an expression vector to express it in cells. After purification, the fusion protein is incubated with lysates of cells known to be susceptible to the virus. A complex of receptor-fusion proteins will be formed due to the affinity between them. The complex can be isolated using PAGE, then identified by mass-spectrometry (Flint *et al.*, 2015; Raj *et al.*, 2013).

4.2.1. Generation of fusion glycoproteins of K113

HML-2, as a member of betaretroviruses, shows a noncovalent association between the two subunits of the Env (Henzy & Coffin, 2013; Jern *et al.*, 2005). However, this type of bounds is labile, causing some SUs to shade off (Robey *et al.*, 1987), which results in a nonfunctional Env (Yang *et al.*, 2000) That is mainly because the SU protein functions as the key step in the entry process. It has specific domains that interact with cellular receptors on the cell surface, leading to conformational changes that expose the fusion peptide at the N- terminus of TM.

Such a peptide is crucial for insertion into the membrane of the target cell and triggering the fusion process between viral and cellular membranes (Hunter & Swanstrom, 1990).

We wanted to produce a soluble form of the viral Env trimer that resembles the native state but is more stable to empower binding to its receptor. We adapted some stabilization strategies formerly used to stabilize the trimer of HIV Env (Yang *et al.*, 2002). The first used strategy included employing a trimerization domain from the C-terminus of bacteriophage T4, the fibritin (FT) domain, a motif consisting of 27aa (GYIPEAPRDGQAYVRKDGEWVLLSTFL) that promotes trimerization of proteins (Figure 4-10, C) (Frank *et al.*, 2001; Guthe *et al.*, 2004; Tao *et al.*, 1997). The second strategy was to stabilize the trimer by changing the proteolytic cleavage site between SU and TM Envs by a mutation to prevent cleavage (Yang *et al.*, 2000; Yang *et al.*, 2002).

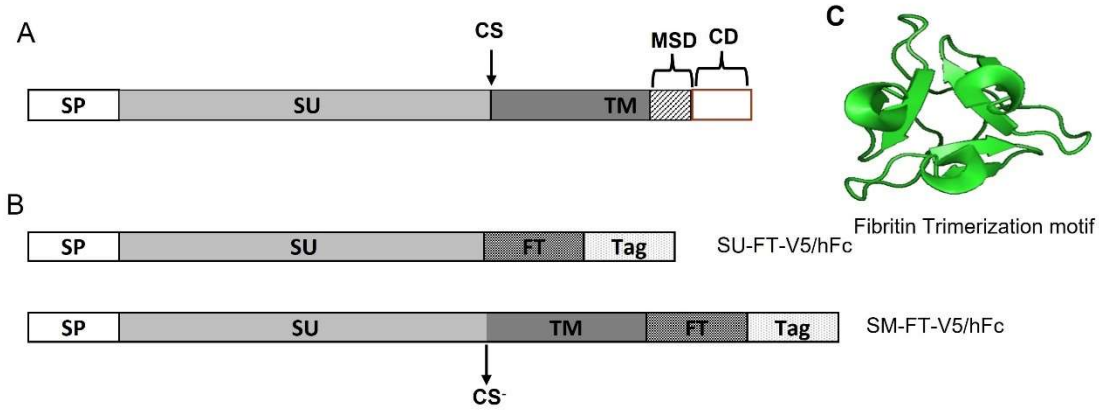

Figure 4-10 Schematic diagram of the designated trimeric fusion glycoproteins of the K113 Env.
A. *Structural regions of the K113 Env according to(Hanke et al., 2009). SP: signal peptide; SU: surface protein; CS: cleavage site; TM: Transmembrane protein; MSD: membrane-spanning domain; CD: cytoplasmic domain.* ***B.*** *Fibritin stabilized trimers of the two described fusion glycoproteins of the reconstituted oricoEnv, SU-FT-Tag, and the uncleaved ectodomains SM-FT-Tag. CS-: the mutated cleavage site.* ***C.*** *Cartoon representation of the T4 phage fibritin. Adapted from PDB ID 1RFO, DOI: 10.2210/pdb1rfo/pdb (Guthe et al., 2004), the picture was taken using PyMol.*

Following these two strategies together, we designed putative soluble forms of the trimer protein using the reconstituted K113 Env sequence (Hanke *et al.*, 2009). In the beginning, we cloned the complete SU and TM ectodomains with mutations in the SU/TM furin cleavage site ($_{461}$NRSKR was changed to $_{461}$NASAA), obtained from K113 oricoEnvCS$^-$-V5 plasmid (Hanke *et al.*, 2009), that prevents the cleavage process. The cloned cDNA was inserted into the pTH vector between an introduced EcoRI site and HindIII site. The fibritin trimerization domain was fused on the Env C-terminus to increase the trimer stability. We tagged the recombinant protein

with either V5-tag (SM-FT-V5) or the Fc domain of human IgG heavy-chain molecule (SM-FT-hFc) by cloning them into NheI and NotI sites of the pTH downstream of the fibritin motif, as shown in (Figure 4-10, B).

To avoid any complications caused by cleavage prevention, the second version of the trimer included only the SU protein since it is the direct interacting protein with the receptor (Hunter, 1997; Hunter & Swanstrom, 1990). However, expressing the SU protein without the TM protein results in a monomeric SU protein product (Earl & Moss, 1993; Wamara, 2020), which might have a negative impact on its interaction with the receptor, as in the case of HIV, where the oligomeric form (the trimer form) of the gp120 is required for efficient binding to CD4 (Moore *et al.*, 1991). Taking this into account, we designed the second form of the soluble glycoprotein from the SU portion only but in an oligomeric form by employing the fibritin motif again. Similarly, we cloned the SU cDNA from the K113 oricoEnvΔC1 plasmid into the pTH vector between an introduced EcoRI site and HindIII site. We linked it with the fibritin motif, followed by either a V5-tag (SU-FT-V5) or the Fc domain of a human IgG heavy-chain molecule (SU-FT-hFc). Figure 4-10; B) shows the structures of the two designed fusion proteins.

We transiently transfected HEK293T cells with the expression plasmids of the fusion proteins SM-FT-V5, SM-FT-hFc, SU-FT-hFc, SU-FT-V5, and XSU-hFc (Battini *et al.*, 1999). The latter is the plasmid of the XMRV fusion protein. It was a gift from Dr. Dusty Miller and used as a control in addition to K113 oricoEnvΔC1 and pcDNA plasmids. Then, we collected the cell lysates, subjected them to denaturing SDS-PAGE, and analyzed them by Western blot using specific antibodies against the tags (Figure 4-11; A-B). We could also detect the TM portion in the SM fusion proteins using anti-HERV-Env, a TM-specific antibody (Figure 4-11; C). Since both types of the tagged proteins (-V5/-hFc) were expressed similarly and due to the availability of a large-scale purifying method for the Fc-tagged proteins, we decided to use the Fc-tagged protein in our further investigation to find the receptor.

We purified the soluble fusion proteins (SM-FT-hFc, SU-FT-hFc, and XSU-hFc) from the supernatant of the transfected HEK293Tcells (3.2.3.3) and concentrated them using the BSA assay (3.2.3.4). The purified and concentrated fusion proteins were then analyzed by Western blot using anti-h-IgG. We experienced (especially in SM-FT-hFc) protein cleavage at the tag level when high temperatures were applied during sample preparation for the denaturing SDS-PAGE, which prevented the detection of the proteins using anti-h-IgG.

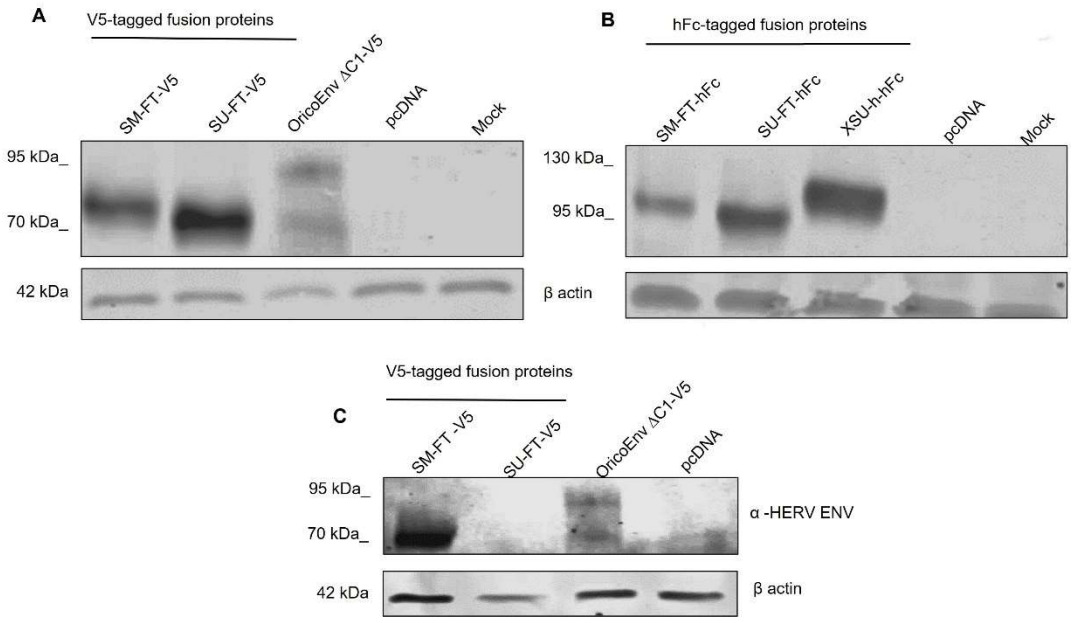

Figure 4-11 Expression of the designed K113 fusion glycoproteins.

HEK293T cells were transfected with the pTH-SM-FT-hFc/V5, pTH-SU-FT-hFc/V5, pcDNA, K113 oricoEnvΔC1-V5, or XSU-hFc; the latter is the fusion protein of XMRV and was used as a control. Then, the cell lysates were subjected to SDS-PAGE under denaturing conditions, followed by Western blot.

A. *Western blot analysis of the V5-tagged fusion proteins using an anti-V5 antibody.* ***B.*** *Western blot analysis of the hFc-tagged fusion proteins using anti-human IgG. Mock: non-transfected cells.* ***C.*** *Detection of the TM protein in the designed SM-FT fusion protein but not in the SU-FT that lacks TM portion, compared to the wild type K113 oricoEnvΔC1, using anti-HERV-Env (TM specific antibody). β-actin was detected in all lysates using a specific antibody.*

We heated samples to 70°C for either 1min or 5min. The proteins' electrophoretic mobilities were consistent with those detected in the cell lysates, and 5min heating was enough for reduction without losing the Fc-tag. Heating the samples for 1min allowed us to detect two forms of SU-FT (the dimer and the monomer) as well as the control XSU, which is expressed as a dimer (Battini *et al.*, 1999) (Figure 4-12; A). However, for the purified SM-FT proteins, we experienced aggregation, which made it challenging to migrate through the gel and to detect by Western blot. The V5-tagged version SM-FT-V5 ran into a similar issue. The protein aggregation caused different atypical structures that do not resemble the native Env (data are not shown).

Therefore, we continued our investigation with the SU-FT-hFc protein, which showed a molecular weight of ~92kDa higher than the predicted 78 kDa, indicating glycosylation of the protein. We know that the SU subunit has five N-linked glycosylation sites, as reported by Hanke et al.(Hanke *et al.*, 2009). Four sites should be found in our fusion protein since the last

one is located in the cleavage site between SU-TM. It was disrupted in our protein due to the introduced EcoRI site. However, Fc-tag has one N-linked glycosylation site. This makes the total glycosylation sites on the fusion protein five. These five sites correspond to a 14 kDa shift, considering a reduction of 2.8 for each oligosaccharide (Cheynet *et al.*, 2005). Treating the fusion protein SU-FT-hFc with PNGase F under reduced conditions and analyzing it by Western blot confirmed the presumption (Figure 4-12; C).

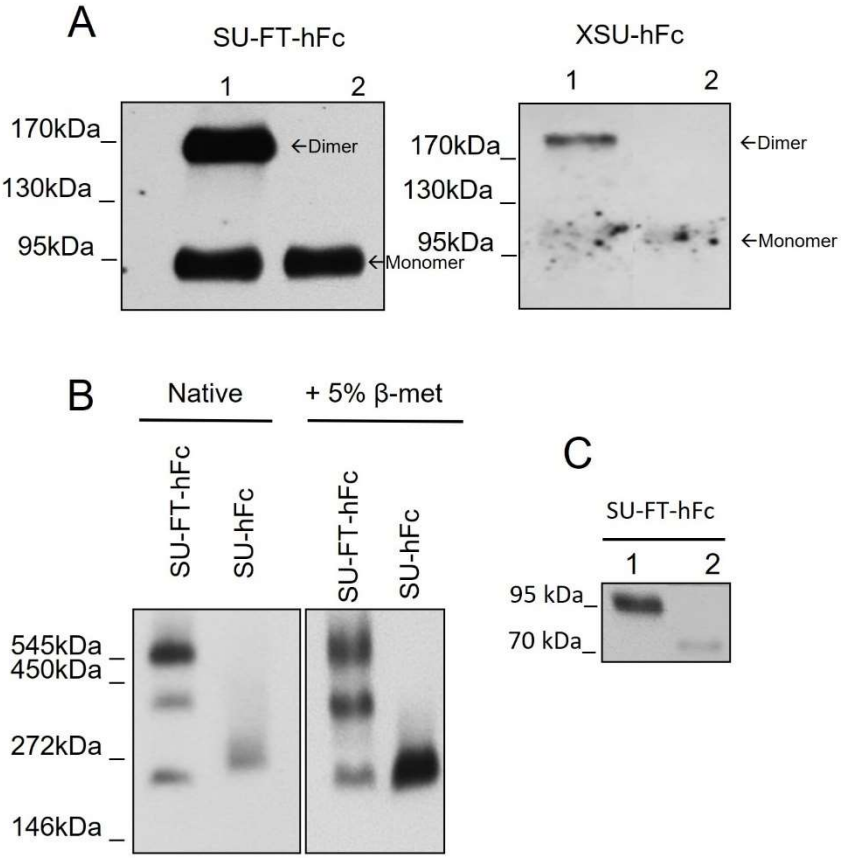

Figure 4-12 Analysis of the purified fusion glycoproteins under native and denaturing conditions.
***A**. The fusion proteins were separated by denaturing SDS-PAGE using Laemmli buffer and heating at 70°C for either 1min (1) or 5min (2), then analyzed by Western blot.* ***B.** The two forms of K113 fusion proteins, the trimer SU-FT- hFc and the monomer SU-hFc, were separated by native-PAGE using native loading buffer or Laemmli buffer containing 5% β-Mercaptoethanol at 37°C for 30 min, then analyzed by Western blot.*
***C.** The purified fusion protein SU-FT was treated with PNGase F at 37°C for 1h, then analyzed by SDS-PAGE and Western blot. The treated protein (2) showed shifted band compared to the untreated (1). The proteins were detected using anti-h-IgG*

Next, we examined the SU-FT oligomeric (trimer) structure and its stability by Western blot analysis. We employed the monomeric fusion protein of the K113 Env SU-hFc as a control. It was designed similarly to ours but without the FT domain (Wamara, 2020). We examined the trimeric structure at native conditions (non-reduced) and checked the trimer stability by treating

the protein with 5% β-mercaptoethanol at 37°C. SU-FT showed a trimeric structure in the native conditions. The trimeric structure was maintained even in 5% β-mercaptoethanol (Figure 4-12; B), though the monomeric fusion protein SU-hFc showed only the monomeric form in both tested conditions.

Taking these results together, we could produce a soluble fusion protein for K113 (SU-FT) that is expressed, glycosylated, and oligomerized as a trimer stabilized by FT motif, similar to the wild-type Env SU trimer. Accordingly, we can use it in the next steps for binding to the receptor.

4.2.2. Interaction of the soluble trimer protein SU-FT with HP

We further investigated whether our SU-FT trimer could bind glycans with similar efficiency to the wild-type Env.

The purified trimer protein SU-FT-hFc and the monomeric form SU-Fc were incubated with HP-coated beads in the presence and the absence of soluble HP or HS as a pre-incubation step. XSU-hFc was included in this investigation as a negative control, given that XMRV Env binds cell surface in an HP-independent manner (Xu & Eiden, 2011). Additionally, we incubated the soluble fusion proteins with protein G-beads and Streptavidin-beads as binding positive and negative controls, respectively. The concentrations of the tested soluble proteins were equal in all tested tubes. Similar test conditions were applied to capture and precipitate the wild-type Env using ultraconcentrated and normalized K113 oricoEnvΔC1 pseudotyped HIV viruses. The viral particles were treated similarly and incubated with HP- and streptavidin-beads. After incubation, all beads were washed. The bound proteins were eluted and analyzed by Western blot using anti-h-IgG to detect the Fc-tagged proteins and anti-HERV-Env to detect the wild-type Env. No virus was captured when we used the soluble HP or the streptavidin-beads (Figure 4-13), indicating that the soluble HP saturated the interaction sites on the Env. However, we could detect a low amount of the Env in the presence of the soluble HS compared to the HP-beads only, suggesting that the interacting sites were not completely saturated by the concentration of HS used since it is less sulfated than HP.

Results of the fusion proteins showed that all proteins were detected using protein G-beads. However, only SU-FT-hFc was detected using HP-beads, and a smaller amount in the presence of both soluble competitors HP and HS (Figure 4-13), suggesting that a higher concentration of the competitors was needed than the one used to achieve the saturation. No proteins were detected using streptavidin-beads.

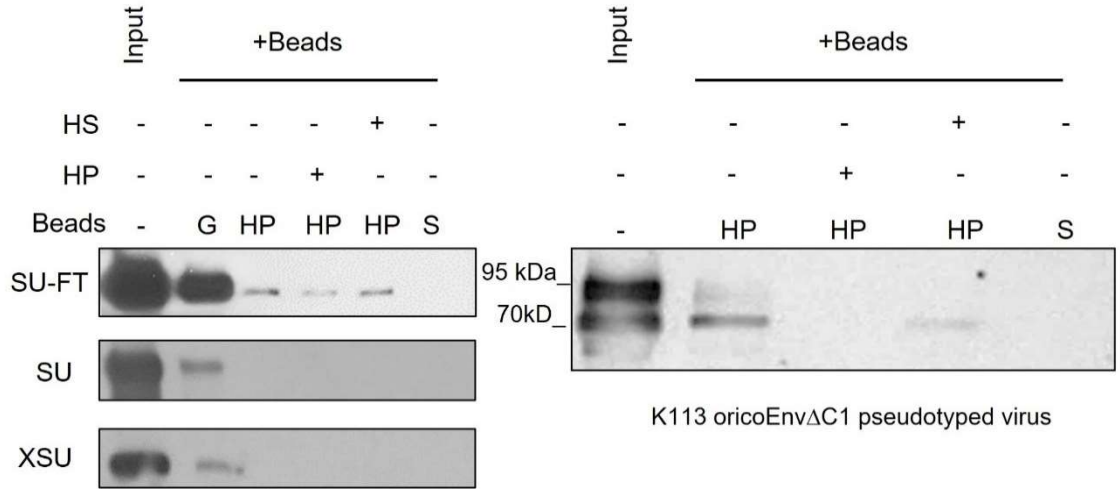

Figure 4-13 Binding of the fusion proteins and K113 Env pseudotyped HIV to HP-coated beads.
10 ng p24 K113 Env pseudotyped HIV and 3ug of each purified fusion protein were incubated with protein G-beads (G), streptavidin-beads (S), or heparin-beads (HP) in the presence or absence of the excess soluble HP or HS at a concentration of 50 µg/ml. The captured virus and proteins were analyzed by Western blot using anti-h-IgG to detect the fusion proteins and anti-HERV-Env to detect the viral Env. Input= the total input. The data shown is a representative experiment of two independent experiments.

These results demonstrated that the soluble fusion protein SU-FT-hFc can interact with glycans through GAG binding sites on the SUs. Also, the trimer protein but not the monomeric form can interact with the glycans, in view of the inability of the monomer to interact with the HP-beads in presence or absence of the competitors. This means that our protein SU-FT-hFc behaves similarly to the native state envelope trimer.

4.2.3. Binding of the trimer fusion protein SU-FT-hFc to the cell surface

Since the receptor-binding domain is included in the SU-FT fusion protein, we investigated the capability of the trimer SU-FT to bind to the surface of cells that allow HML-2 Env mediated entry. These cells should express the elusive receptor protein of HML-2. Moreover, the soluble SU-FT proteins should block the entry of HML-2 Env pseudotyped viruses.

The binding of the soluble trimer SU-FT to the unknown cell surface receptor was assessed by flow cytometry. More than one receptor protein potentially allows entry of HML-2 (Robinson-McCarthy *et al.*, 2018; Robinson & Whelan, 2016). Since the SK-MEL-28 cell line, one of the cell lines included in the NCI-60 cell line panel (Shoemaker, 2006), had been experimentally proven to be permissive to HML-2 Env infection (N. Bannert, unpublished data), we incubated cells from this cell line with the purified soluble proteins SU-FT-hFc and XSU-hFc as a control. We expected no binding for XSU since the used cell line does not express XPR1, the receptor

of XMRV (Battini *et al.*, 1999), based on its Z-Score (Z-score = -1.033 in the used cells). The Z-Score was obtained using the NCI-60 CellMinor tool (Reinhold *et al.*, 2012).

Removing the cellular GAGs decreases viral particles attached to the cell surface since the binding to the HS is essential for HML-2 Env-mediated entry (Geppert, 2019). On the other hand, the previous results in Figure 4-13 proved the ability of the trimer SU-FT protein to bind HS. Accordingly, removing HS from the cell surface would decrease the trimer attachment as same as the viral particles. Therefore, we also examined the binding of SU-FT trimer protein to the cell surface in the absence of cellular HS. We incubated the fusion proteins with SK-MEL-28 cells after enzymatic removal of the surface GAGs according to the method described in section (3.2.2.7).

We assessed the binding of the fusion proteins to the cell surface using a fluorescent-conjugated secondary antibody directed against the Fc-tag. We stained the dead cells with the FVD according to the described method in section (3.2.2.6) to gate them on a different channel and exclude them from the flow cytometry analysis, as shown in Figure 4-14. The SU-FT trimer fusion proteins bound the cells and were observed as a fluorescence shift compared to the incubated cells with the secondary antibody alone (Figure 4-14; A). This binding was specific since XSU-hFc was unable to bind the cells. However, the binding of SU-FT to the cells was abolished when it was incubated with pretreated cells with heparinase I/III (Figure 4-14; B), which affirmed the direct role of GAGs in HML-2 Env attachment. No changes in the binding of XSU to the cells were observed after the enzymatic removal of the cellular GAGs.

In the next step, we investigated whether the interaction between the fusion protein SU-FT and cell surface receptors could inhibit HML-2 Env pseudovirus infection.

First, we performed a control inhibition assay using the XSU fusion protein to inhibit XMRV Env pseudoviral infections. The XSU fusion protein was reported to inhibit XMRV infections (Côté *et al.*, 2012).

We produced HIV pseudotyped luciferase reporter viruses carrying XMRV Env, K113 oricoEnvΔC1, VSV-G as positive infection control, and K113 R140C Env as negative control by transfecting HEK293T cells according to the method described in section (3.2.2.4). The viruses were ultraconcentrated and normalized by p24 ELISA before being used to infect the cells.

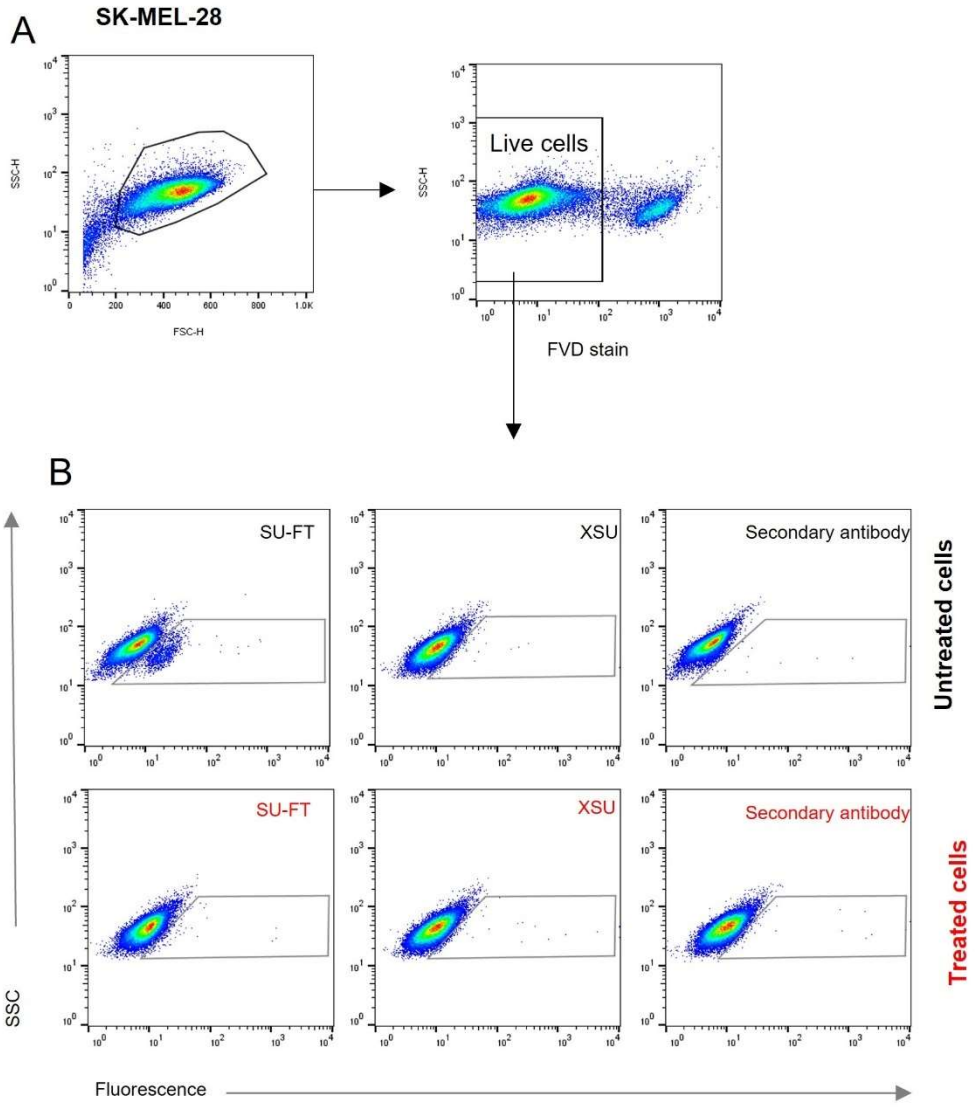

Figure 4-14 Binding of the soluble trimeric fusion protein SU-FT to an unknown receptor on the cell surface of permissive SK-MEL-28 cells.

***A.** Gating strategy. In the gating strategy, dead cells were excluded by gating on the FL4 channel. Then, the living cell population was gated on the FL1 channel, as illustrated. **B.** The bound fusion proteins to heparinase-treated and untreated SK-MEL-28 cells. The fusion proteins were detected using Alexa Fluor 488-labeled anti-human IgG. The background is represented by cells incubated with the secondary antibody alone.*

We exposed DU145 cells, a permissive cell line for XMRV (XPR1 Z-score = 0,569 (Reinhold *et al.*, 2012)) to different concentrations of the soluble fusion protein XSU-hFc (0, 1, 2, 4, and 8ug/well) before being infected with XMRV Env, VSV-G, and K113 R140C Env pseudotyped viruses. We lysed the cells 72h post-infection and measured the luciferase activity. We sat the background signal according to K113 R140C Env entry-level and minimized it from the infection levels of both viruses XMRV Env and VSV-G. Results showed that XSU soluble fusion protein blocked the infection of XMRV Env pseudovirus in a dose-dependent manner,

resulting in about 50% inhibition in XMRV Env entry at a concentration of 8 µg/well ($p < 0.05$). The entry of the VSV-G pseudovirus was not blocked at this concentration of XSU-hFc (Figure 4-15, A).

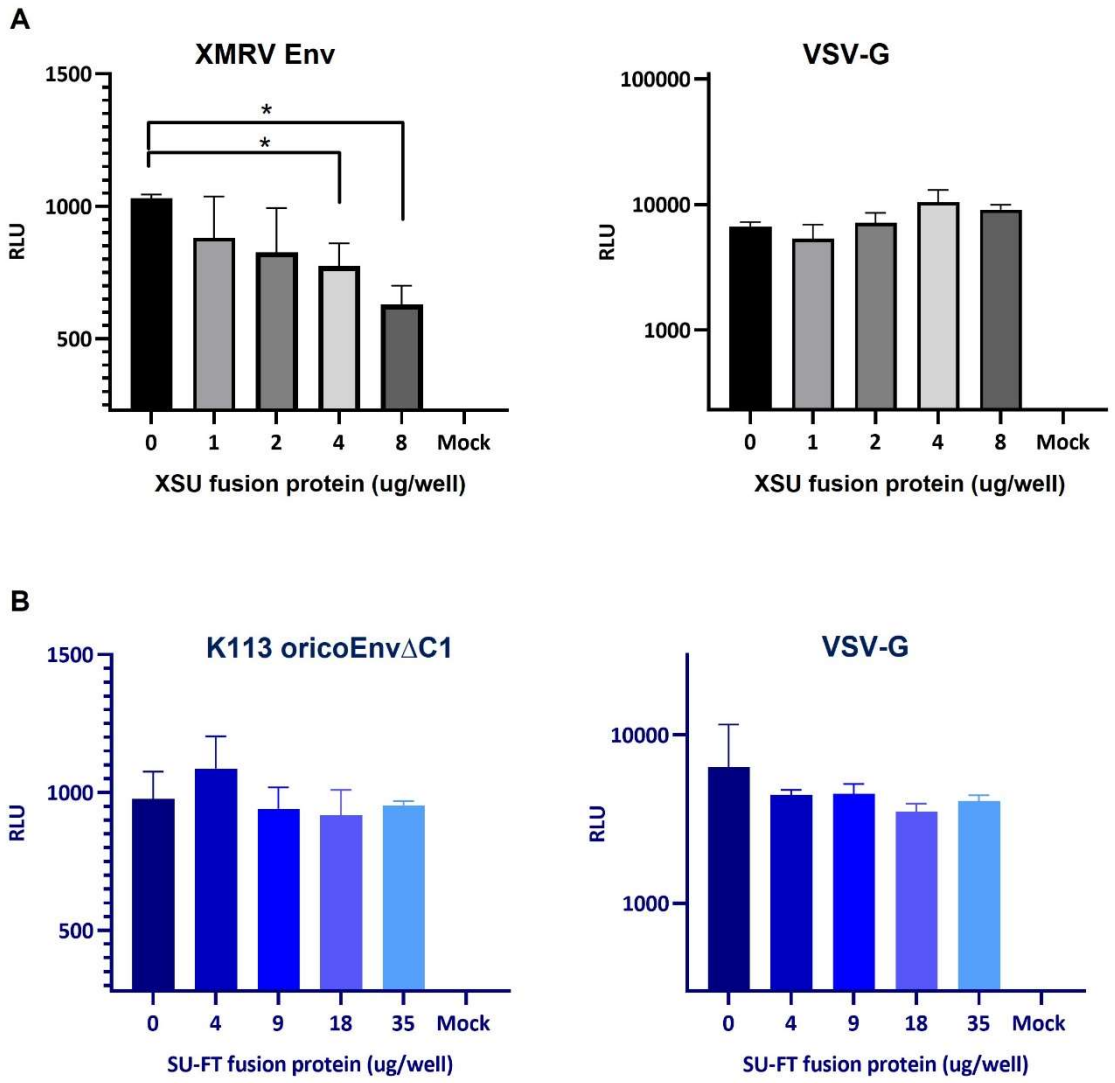

Figure 4-15 inhibition of pseudoviral infection using soluble fusion proteins.

A. *Inhibition of XMRV Env pseudotyped lentiviral infection of DU145 cells by using XSU soluble fusion protein. 10ng p24 of XMRV Env, K113 R140C Env, or 1ng of VSV-G pseudoviruses were used to infect cells in the presence of increasing concentrations of the XSU-hFc protein.* ***B.*** *Infection of the HML-2 permissive cell line SK-MEL-28 in the presence of increasing concentrations of the soluble fusion protein SU-FT-hFc with the lentiviral pseudotypes carrying the K113 oricoEnvΔC1 (75 ng p24), K113 R140C Env (75 ng p24), or VSV-G (1 ng p24). No inhibition of any of the pseudotypes was observed ($p > 0.05$). The background was set to the negative control K113 R140C Env pseudovirus entry-level. The means of 3 wells ± SD are shown. Data is representative of three independent experiments. * $p < 0.05$*

In the same way, we incubated SK-MEL-28 cells with different concentrations of the HML-2 soluble fusion protein SU-FT-hFc at (0, 4, 9, 18, and 35ug/well) to inhibit K113 oricoEnvΔC1 pseudovirus infection. VSV-G and K113 R140C Env pseudoviruses were used as well. No inhibition of K113 oricoEnvΔC1 infection was observed in any of the used concentrations ($p > 0.05$). The infection levels in all wells were comparable to the control cells (0µg/well), similar to that of the infection control VSV-G (Figure 4-15; B). Comparable results were obtained when we infected other cell lines, including SK-MEL-5 and DU145 cells, with less permissivity for HML-2 Env (medium and low, respectively) (data not shown).

Together, these results demonstrate the ability of the soluble HML-2 trimer protein SU-FT to bind to cell surface glycoproteins. The interaction is mediated through cellular GAGs that can be removed by heparinase, which prevents the binding of SU-FT to the cell surface. Nevertheless, unlike XMRV, which does not bind to GAGs, HML-2 fusion protein SU-FT at the used concentrations cannot inhibit the infection by the HML-2 Env pseudovirus.

4.2.4. Identification of GP73 transmembrane protein by the pull-down assay

As we were able to detect the binding of the HML-2 soluble fusion protein SU-FT to GAGs on the surface of the infection-permissive SK-MEL-28 cells, we would be able, in principle, to pull down such glycoproteins that serve as an attachment factor or even as a receptor for the virus. To verify this strategy, we conducted a control pull-down assay for the XPR1 protein using the fusion protein XSU-hFc. We chose two cell lines from the NCI-60 panel based on their XPR1 expression, the first was NCI-H522, which expresses the XRP1 (Z-score = 0.473), and the other was NCI-H226 which does not express XPR1 (Z-score = -0.248) (Reinhold *et al.*, 2012). We extracted membrane proteins from the two cell lines according to the method described in section (3.2.3.2). The extracts were incubated with the XSU fusion protein overnight at 4 °C. At this step, the fusion protein would bind the receptor specifically, forming a protein-receptor complex, which was isolated in a further step using protein G-beads that have an affinity to the hFc-tag.

Additionally, we incubated the extracts and fusion proteins individually with protein G-beads as controls for the cross-reactivity. We then eluted the co-immunoprecipitated (Co-IP) proteins, processed them by PAGE, and subjected them to silver staining. The staining revealed a band at ~65 kDa in the extracts of H522 cells incubated with XSU but not in the extracts of H226 cells (Figure 4-16; A), indicating that XSU had pulled down its receptor since XPR1 protein

migrates at a molecular weight of ~65 kDa (Ansermet *et al.*, 2017). Western blot analysis of a similar PAGE using a polyclonal antibody against XPR1 showed the 65 kDa bands in both the extracted membrane protein lysate of NCI-H522 alone and the one Co-IP by XSU (Figure 4-16; B). This band was absent in homologs samples prepared from the NCI-H226 cell line and on other lanes. These results indicate the validity of the method used to pull down the specific receptor of XMRV and justify using it in an approach to isolate a glycoprotein that functions as an HML-2 attachment factor or entry receptor.

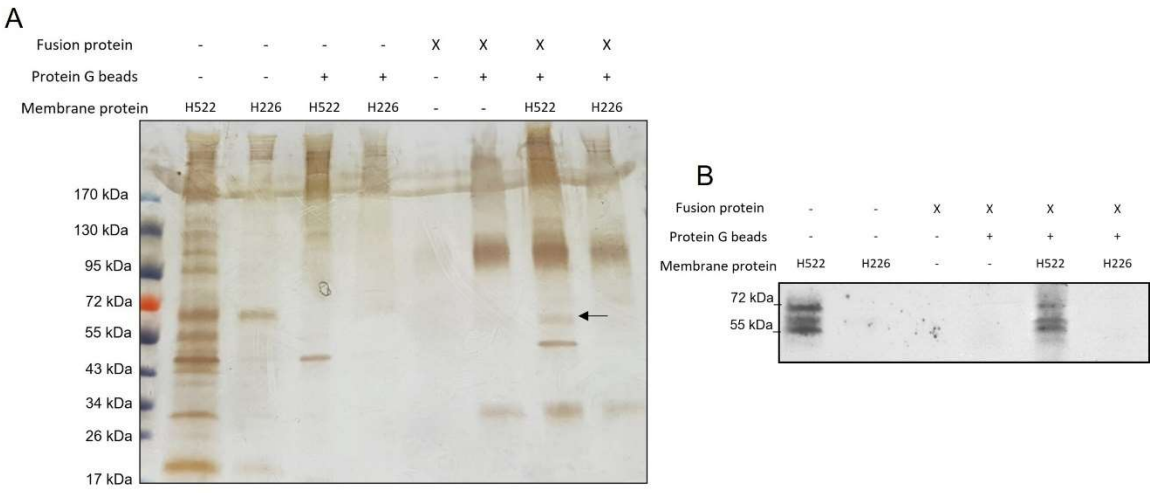

Figure 4-16 Pull-down of XPR1 receptor by XSU fusion protein and protein G-beads.
Lysates of the extracted membrane proteins of the indicated cell lines were incubated with XSU-hFc fusion protein. Co-IP using protein G-beads was performed, followed by PAGE to be either silver stained or analyzed by Western blot. ***A.*** *Pull-down of the XPR1 receptor by the XSU fusion protein (X), the arrow on the silver-stained gel (on the left) indicates the ~60 kDa band, which corresponds to the XPR1;* ***B.*** *Western blot analysis of the same sample using XPR1 polyclonal antibody. Three bands of the XPR1 were detected of around 70-60 kDa.*

To pull down the glycoprotein that functions as an attachment factor or a receptor for HML-2 Env, we chose two experimentally proven cell lines that K113 Env is able to infect, SK-MEL-28 and U125, and the non-permissive cell line NCI-H226 (N Bannert, unpublished data). Membrane proteins of these cells were extracted and then incubated with SU-FT-hFc or XSU-hFc as a negative control since the mentioned cell lines do not express XPR1 (Reinhold *et al.*, 2012). After incubation with the soluble fusion proteins, we added protein G-beads as described above. Precipitants were then subjected to PAGE, followed by silver staining. The latter revealed one different band at a high molecular weight in the extracts of both permissive cells Co-IP by SU-FT but not with the control XSU fusion protein or in the precipitants of the non-permissive cell line with SU-FT (Figure 4-17; A). Due to the high molecular weight of the

observed band, we assumed that it contained the trimer fusion protein SU-FT bound to the putative receptor and formed a complex. The assumed complex migrated at a higher molecular weight than the trimer SU-FT alone did. Accordingly, employing an antibody against the Fc-tag would enable us to detect, at this high molecular weight, the assumed complex in the lane of the Co-IP membranes by the SU-FT, but not in the lane of the SU-FT trimer alone. By analyzing the same samples with Western blot, we were able to detect, at the high molecular weight, the assumed complex (fusion protein-putative receptor) in the lanes of the two cell line extracts Co-IP by SU-FT but not in the SU-FT alone or other lanes (Figure 4-17; B).

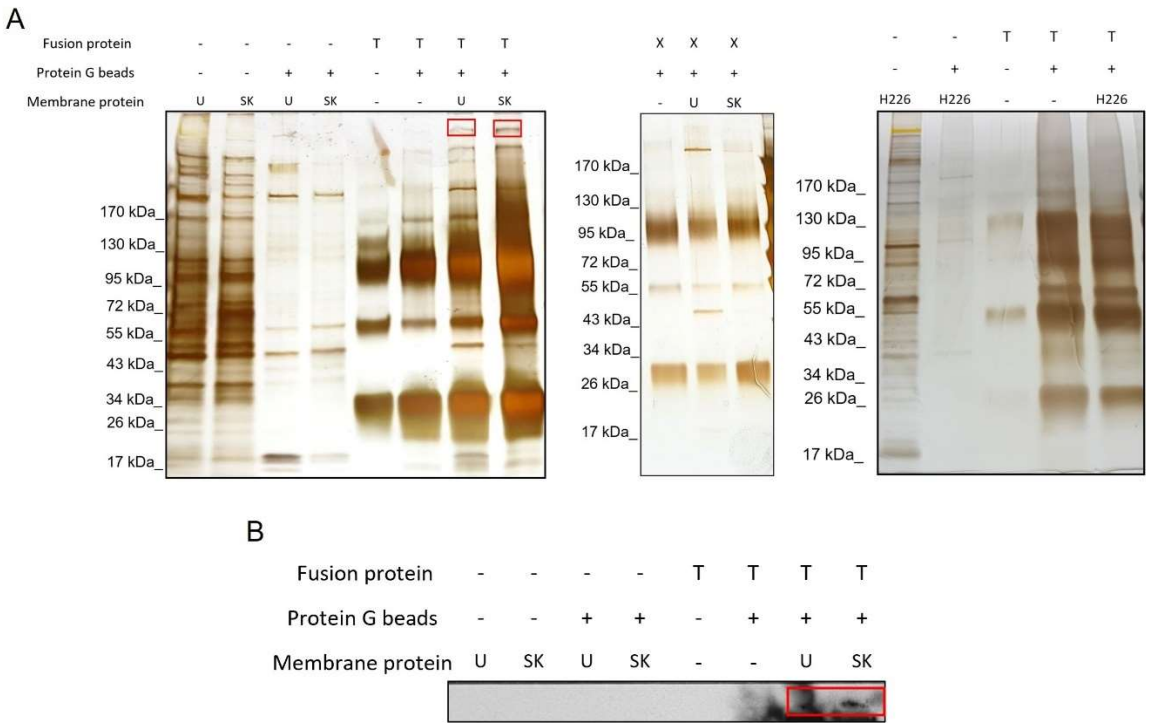

Figure 4-17 Pull-downs of the potential receptor or attachment factor by HML-2 fusion protein and protein G-beads.

Lysates of extracted membrane proteins of the indicated cell lines were incubated with the fusion proteins. Co-IP using protein G-beads was performed, followed by PAGE to be either silver stained or subject to Western blot. ***A.*** *silver-stained gel for the results of pull-down assay of HML-2 potential receptor/attachment factor by the trimeric fusion protein SU-FT-hFc (T) and the control XSU (X) in the cell lines SK-MEL-28 (SK), U251 (U) and NCI-H226 (H226). The red boxes on the gel (on the left) indicate the potential bands of SU-FT-hFc with a cellular membrane protein.* ***B.*** *Western blot analysis of the same samples using a specific antibody against the Fc-tag detects the expected bands (indicated in red boxes).*

Mass spectrometric analysis results revealed specific peptides that confirmed the presence of the fusion protein-receptor complex and identified the Co-IP potential receptor protein as Golgi Membrane Protein-1 (GOLM1, also called GP73 or GOLPH2), a highly conserved type II

transmembrane protein with a molecular weight of 73 kDa, normally expressed in epithelial cells. It is translated, oligomerized as a dimer, and glycosylated in ER, then translocated to Golgi, from where it cycles to the cell surface from TGN and backs to it through late endosome bypass (Kladney *et al.*, 2000; Li *et al.*, 2012; Munro, 1998; Puri *et al.*, 2002). Its function is not fully understood; however, since its expression is controlled by a housekeeping gene-like basal promoter structure, it may be involved in certain housekeeping functions (Gong *et al.*, 2012), including protein sorting, modification, and cargo transport (Ye *et al.*, 2016).

4.2.5. GP73 as a potential attachment factor for HML-2 Env

The GP73 protein sequence is divided into 5 structural regions (Figure 4-18; A): an N-terminal cytoplasmic tail and transmembrane domain, a coiled-coil domain, a flexible loop, and an acidic tail at the C-terminus. It is expressed as a dimer stabilized by disulfate bonds found in the coiled-coil region (Hu *et al.*, 2011).

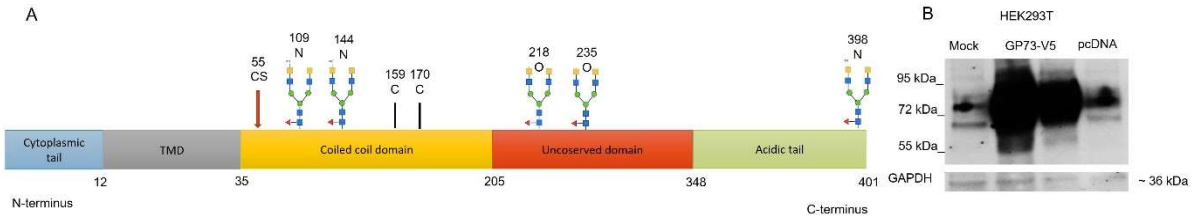

Figure 4-18 GP73 protein structure and expression.
A. *schematic structural regions of the GP73 protein based on (Hu et al., 2011) with the predicted N- and O-linked glycosylation sites in addition to cleavage site (CS) and cysteine residues in the coiled-coil region that are involved in disulfate bonds;* ***B.*** *Western blot analysis of GP73 expression in the cell lysates of transiently transfected HEK293T cells with the cloned GP73-V5 expression plasmid or the empty vector pcDNA, mock= non-transfected cells. Anti-GP73 was used for detection. GAPDH was detected in all lysates using a specific antibody.*

In addition, a conserved cleavage site is located in the coiled-coil region ($R_{52}VRR_{55}$) that is recognized by the proprotein convertases, resulting in the N-terminal cleavage and secretion of the C-terminal ectodomain (Gong *et al.*, 2012). The cleavage correlates with abnormal overexpression levels (Hu *et al.*, 2011). GP73 is a sulfated glycoprotein. It has five glycosylation sites in the ectodomain, three N-glycosylation sites, and two *O*-glycosylation sites; it carries a complex-type N-glycan with core fucose and 4′-*O*-sulfated LacdiNAc glycan (Toyoda *et al.*, 2016).

Since the GP73 protein exists in the cytoplasm and at the cell membrane (Puri *et al.*, 2002), we would not benefit much from its Z-score in the NCI-60 cells panel due to our interest in its

expression at the cell membrane only where the interaction with HML-2 Env occurs. On the other hand, overexpression of GP73 in any cell might increase its secretion, as reported by (Hu *et al.*, 2011).

We cloned the ORF of GP73 that consists of 1200bp from its full-length cDNA of 3042bp (pCMV6-XL5, Origene) into the pTH vector with a V5-tag fused at its C-terminus. Then, transfected HEK293T cells with the newly cloned expression plasmid. The expression was detected using Western blot and a specific antibody against GP73. We detected a high expression level of GP73 in the lysates of GP73-V5-transfected cells compared to the controls, the pcDNA-transfected, and the non-transfected cells (mock) (Figure 4-18; B). After that, we examined the surface expression of GP73 on permissive and non-permissive cells and investigated whether GP73 overexpression might increase the cycled protein to the cell membrane and whether binding of the HML-2 Env might increase upon increasing the plasma membrane-GP73.

We chose 7 cell lines with different experimentally proven permissivity to HML-2 Env from high to none (SK-MEL-28, MALME-3M, NCI-H522, SK-MEL-5, SK-MEL-2, NCI-H226, and HEp-2), respectively (N. Bannert, unpublished data). We transfected cells from these lines with the GP73-V5 expression plasmid. We detected the overexpressed GP73 proteins in the transfected cells using Western blot analysis. Then, we performed cell surface staining on non-transfected cells (naive) and the GP73-transfected cells to detect GP73 at the cell surface using a specific antibody and fluorescent-labeled secondary antibody (Figure 4-19; B). We used the fusion protein SU-FT-hFc for the envelope binding study. The binding of SU-FT-hFc was detected using a fluorescent-labeled secondary antibody directed to its tag (Figure 4-19; C). FVD was used to stain dead cells to exclude them from the flow cytometry (Figure 4-19; A). Finally, we compared the surface staining results obtained from the GP73-transfected cells with those of non-transfected cells (Figure 4-19; E and F).

Western blot analysis showed that GP73 was expressed and detected in the lysate of the transfected cells compared to the control (the naïve) (Figure 4-19; D). The expression was varied since the transfection efficiency is cell line-dependent.

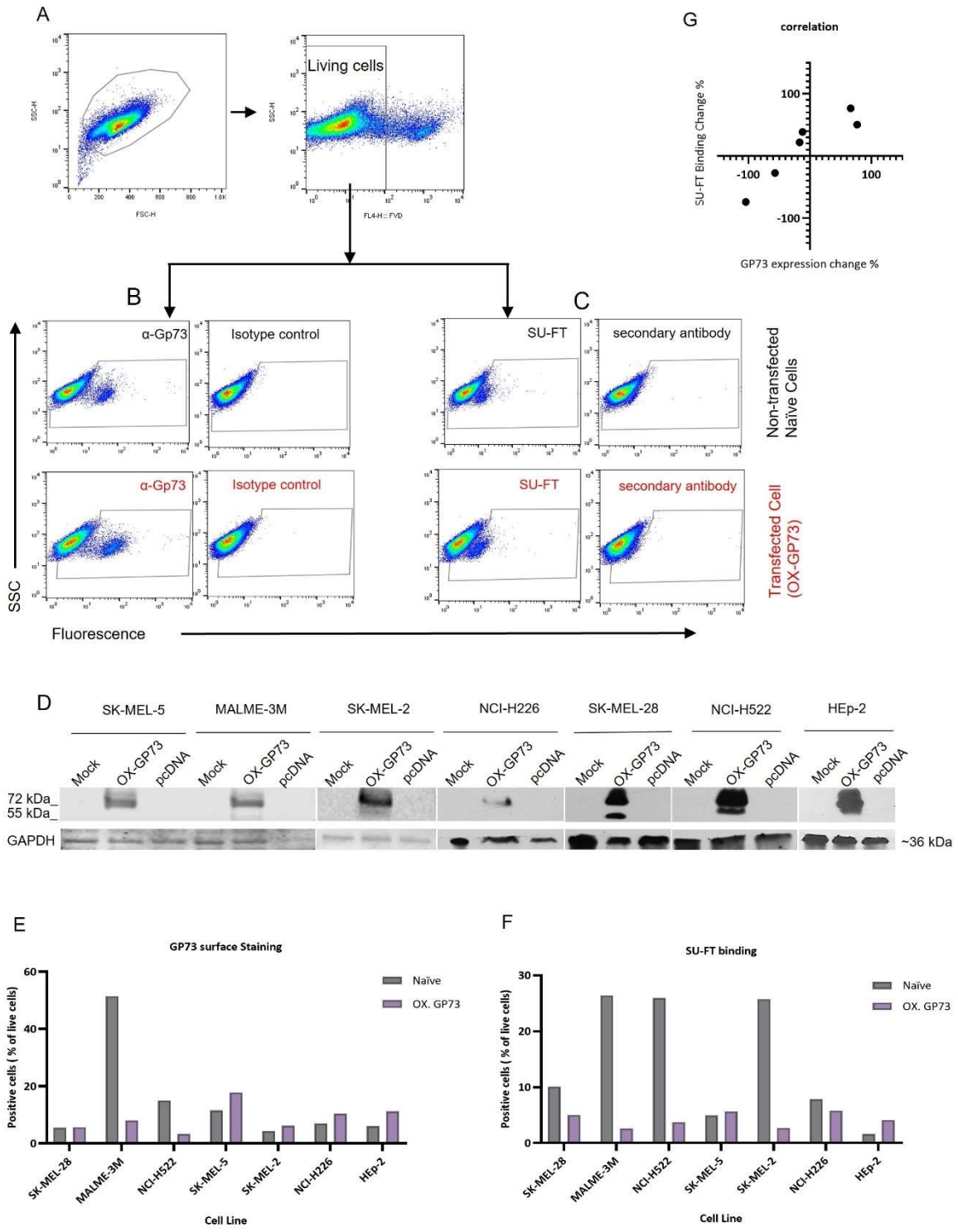

Figure 4-19 Cell surface staining to detect GP73 membrane expression and SU-FT binding.

A. *Gating strategy: in the gating strategy, the transfected cells with GP73-V5(OX-GP73) and non-transfected cells (Naive) from the seven cell lines were gated on the FL4 channel to exclude the dead cell populations. The living cell populations were gated on the FL1 channel, as illustrated.* ***B.*** *The surface expression of GP73 proteins in OX-GP73 and naive cells were detected with mouse anti-GP73 and Alexa Fluor 488-conjugated anti-mouse IgG. The background is represented by cells incubated with anti-IgG mouse isotype control.*

C. *The bound fusion proteins to the OX-GP73 and naive cells. The fusion proteins were detected using Alexa Fluor 488-labeled anti-human IgG. The background is represented by cells incubated with the secondary antibody alone. (Results of HEp-2 cells surface staining are presented in B and C)* ***D.*** *Western blot analysis of the transfected cells (from the seven used cell lines) with GP73-V5 expression plasmid using a specific antibody against V5-tag and GAPDH.* ***E.*** *The surface expressions of GP73 protein in OX-GP73 and naive cells of the seven tested cell lines. Each bar represents the percentage of positively stained cells.* ***F****. The bound fusion proteins to the OX-GP73 and naive cells of the seven tested cell lines. Each bar represents the percentage of positively stained cells.* ***G.*** *The correlation between the percentage of change ratio of the SU protein binding and the change ratio of GP73 surface protein expression in the seven cell lines, calculated using Spearman correlation test Spearman r = 0.9, p = 0.006.*

Nevertheless, we noticed a decrement in plasma membrane-GP73 amount in the high permissive cell lines (MALME-3M and NCI-H522) upon transfection (Figure 4-19; E), suggesting that the protein was secreted instead of cycling to the cell membrane. This decrement coincided with a decrement in SU-FT protein binding (Figure 4-19; F). In contrast, the low/non-permissive cell lines (SK-MEL-5 and HEp-2) showed an increment in the plasma membrane-GP73 upon transfection and coincided with an increment in the SU-FT binding. On the other hand, the two cell lines SK-MEL-2 and NCI-H226 increased the plasma membrane-GP73 upon transfection. However, this coincided with a decrease in SU-FT binding in the two cell lines, suggesting the involvement of other binding factors or glycans, which were affected by the GP73 up-regulation. For the cell-line SK-MEL-28, we observed no changes in the plasma membrane-GP73. However, a decrement in SU-FT protein binding was observed, suggesting that the cells maintained the plasma membrane-GP73, and the overexpressed proteins were cleaved in the cytoplasm since we detected the cleaved form 55kDa band in lysates of GP73-transfected SK-MEL-28 cells. However, it impacted the bound SU-FT to the cell surface, suggesting the involvement of other binding factors or glycans in binding SU-FT to the cell surface.

We found that the fold changes in the cell surface-bound SU-FT and the fold changes of the plasma membrane-GP73 in the seven tested cell lines were significantly strongly correlated (Spearman $r = 0.9$, $p = 0.006$) (Figure 4-19; G).

In the next step, we examined whether the increment in plasma membrane-GP73 contributes to or confers the susceptibility to HML-2 Env pseudoviral infection. We chose the two cell lines SK-MEL-5 and HEp-2 since both showed increments in the plasma membrane-GP73 coincided with increments in binding SU-FT. Then, we transiently transfected the cells with the GP73 expression plasmid and pcDNA as negative control (in HEP-2 non-transfected cells (naïve) were used as negative control). 48h post-transfection, we infected the cells with normalized

luciferase reporter pseudoviruses with K113 oricoEnvΔC1, VSV-G as infection positive control, and K113 R140C Env as a negative control.

Although the plasma membrane-GP73 increased in the low/non-permissive cell lines SK-MEL-5 and HEp-2 upon transfection, we observed no changes in HML-2 Env entry (Figure 4-20), indicating that GP73 is neither the receptor nor a significant attachment factor in these cells. Comparable results were obtained from the HEK293T cell line (data are not shown).

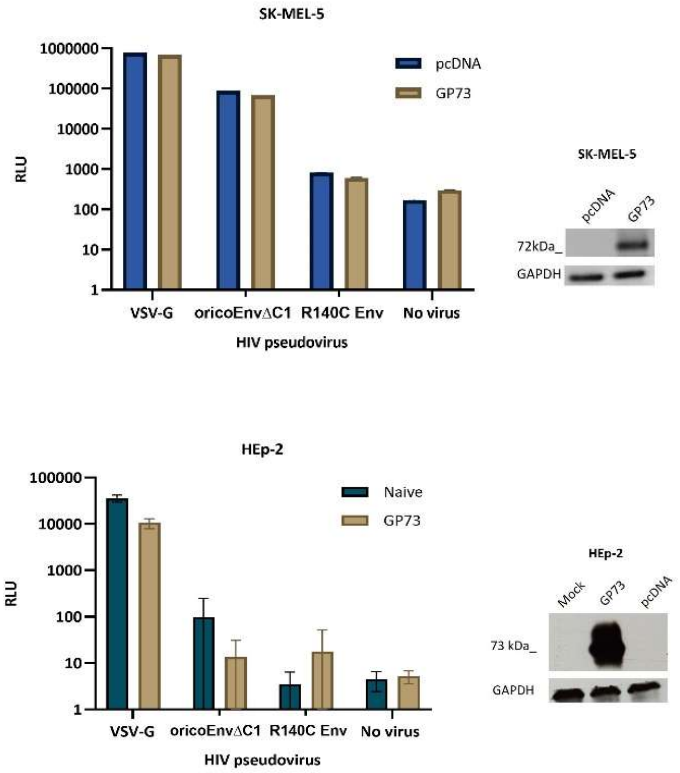

Figure 4-20 HML-2 Env and the control Envs pseudoviral entries in the transfected cells with GP73 expression plasmid.

Cells from SK-MEL-5 and HEp-2 cell lines were transfected with the GP73-v5 expression vector and pcDNA in a 96-well plate. After 48h, GP73 expression was checked by Western blot analysis on lysates obtained from some wells using an anti-V5 tag-specific antibody. The rest of the cells were infected with 500 ng/well of K113 oricoEnvΔC1 and K113 R140C Env and 1 ng/well of VSV-G pseudoviruses. 72h post-infections, cells were lysed, and the luciferase assay was conducted; Naïve: non-transfected cells. The mean of three wells from at least two independent experiments is shown with SD error bars.

Together, the surface staining results demonstrated a direct relationship between the expressed GP73 on the cell membrane and the binding of HML-2 Env, represented by the trimer fusion protein, to the cell surface. The changes in plasma membrane-GP73 affected the Env binding. However, the results in SK-MEL-2, NCI-H226, and SK-MEL-28 cell lines suggested an

implication of other factors or glycans in binding the Env (which is HS for HML-2.) that were impacted by GP73-overexpression.

On the other hand, the increment of the plasma membrane-GP73 did not confer the cell susceptibility to HML-2 Env pseudoviral infections.

Given that GP73 has a sulfated glycan, and the HML-2 Env binds the negatively charged sulfate groups in HS, we proposed that HML-2 Env might bind GP73 as an attachment factor that does not promote entry.

4.3. Screening of human integral membrane proteins cDNA library

To identify the cellular proteins or factors, including the receptor, involved in the early entry pathway of HML-2 Env, we utilized a gain-of-function cDNA library screen, a subset library of mammalian gene ORF clone library for integral membrane proteins in a lentiviral vector (Transomic Technologies). It was created using the Gene Ontology (GO:0016021) project (http://www.geneontology.org/) terms and the Gene Finder bioinformatics tool from the National Cancer Institute (NCI) Cancer Genome Anatomy Project (Yang *et al.*, 2011). The subset included 2440 clones of 2247 integral membrane genes, the most common transmembrane genes among cells (*The Human Protein Atlas*, 2022; Pontén *et al.*, 2008). Each gene's ORF was inserted in the lentiviral vector pXL304 and V5-tagged at its C-terminus (Yang *et al.*, 2011).

4.3.1. Protocol setup and optimization

First, we assessed the incorporation and titers of the produced viruses that carry the library genes. We chose the VSV envelope to produce pseudoviruses carrying the genes due to its high infectivity in most human cell lines (Finkelshtein *et al.*, 2013). Production was done by transient transfection of HEK293T cells in 15 cm^2 plates with two randomly chosen library gene expression plasmids (E6 plate 16 and G12-plate 6) combined with lentiviral packaging and the VSV-G plasmids. Virus supernatants were collected 48h and 72h post-transfection, pooled together and concentrated by ultracentrifuge through a sucrose cushion, then normalized by p24 ELISA to determine the virus titer according to the method described in section (3.2.3.12). Incorporation into viral particles was confirmed by Western blot analysis. We detected the envelope and p24 proteins using specific antibodies (Figure 4-21; A), and both viruses' titers were at comparable levels (Figure 4-21; B).

Then the previous protocol was scaled down to a 96-well plate format for virus production, according to the described method in section (3.2.2.4; C)

After producing the VSV-G pseudovirus with library genes in 96-well plates, we collected the supernatants of 8 random wells from one plate and measured their titers. Since the p24 ELISA assay measures both functional and non-functional viral particles (Geraerts *et al.*, 2006), we measured the functional titer based on their RT activity using the PERT assay. All wells showed RT activity >$10x10^3$ pg/ml, and only two virus supernatants showed RT activity < $10x10^3$ pg/ml (Figure 4-21; C), presuming that the library might include high and low titer-producing genes, which would cause variation in transduction efficiency during screening.

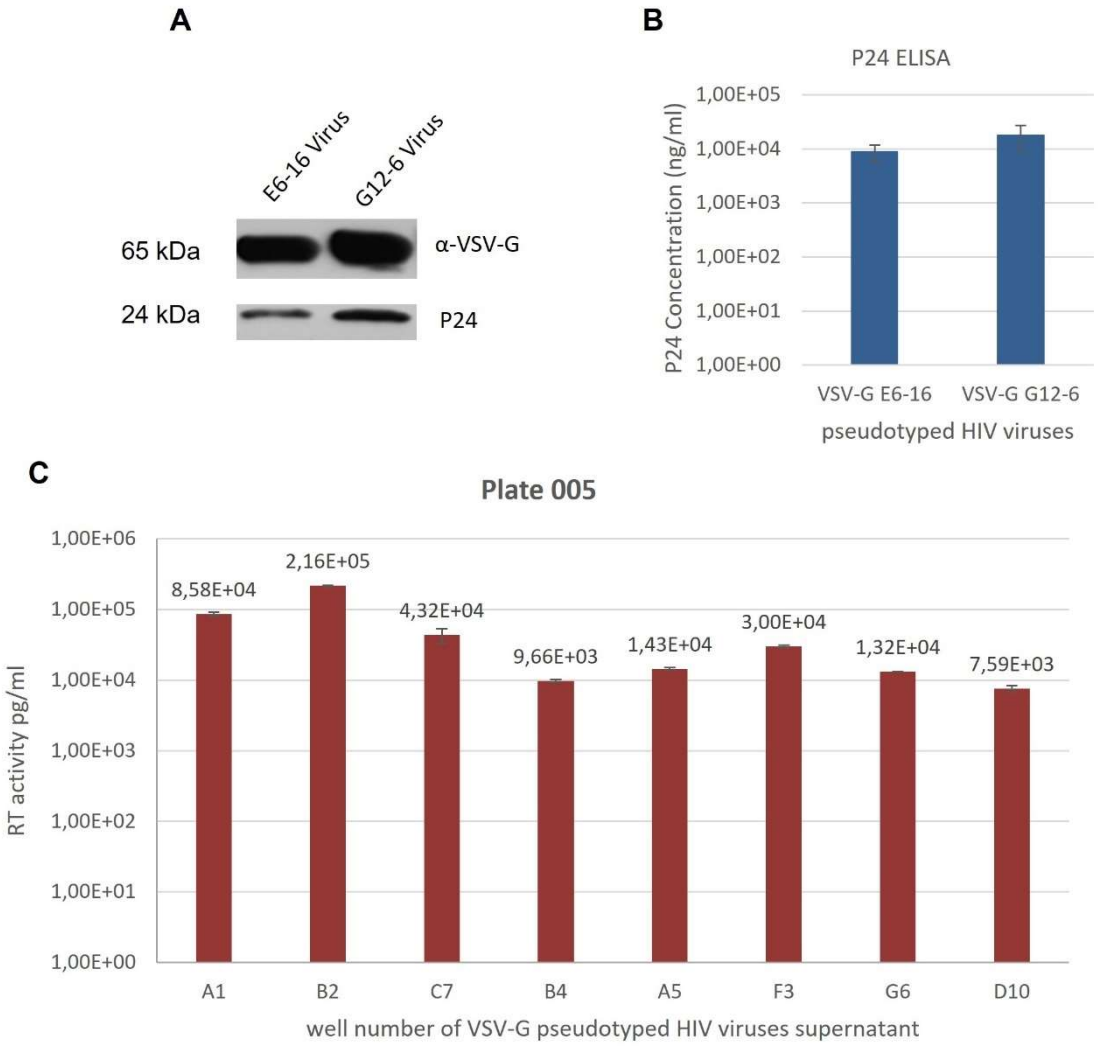

Figure 4-21 Evaluation of the produced VSV-G pseudotyped HIV carrying library genes.
A. *Western blot analysis of VSV-G pseudotyped HIV viruses carrying the library genes from two wells (E6-plate 16 and G12- plate 6) using VSV-G -specific monoclonal antibody and HIV p24 specific antibody which was used to determine the particle load.* ***B.*** *p24 concentration of the ultraconcentrated viruses. Each bar represents the mean of two replicates.* ***C.*** *RT activity of viruses' supernatants collected from 8 random wells of plate 005. Each bar represents the mean of two replicates. Error bars = SD.*

We chose the HCT116 cell line from the NCI-60 cell panel (Shoemaker, 2006) for the transduction step. The candidate cell line has a low entry profile for HML-2 Env but not for VSV-G pseudotyped HIV viruses (N. Bannert, unpublished data), implying that the low entry for HML-2 Env is due to the absence or low expression of the unknown receptor/s rather than other factors including post entry restrictions. This can be distinguished by using the same viral core but a different Env, in this case, VSV-G.

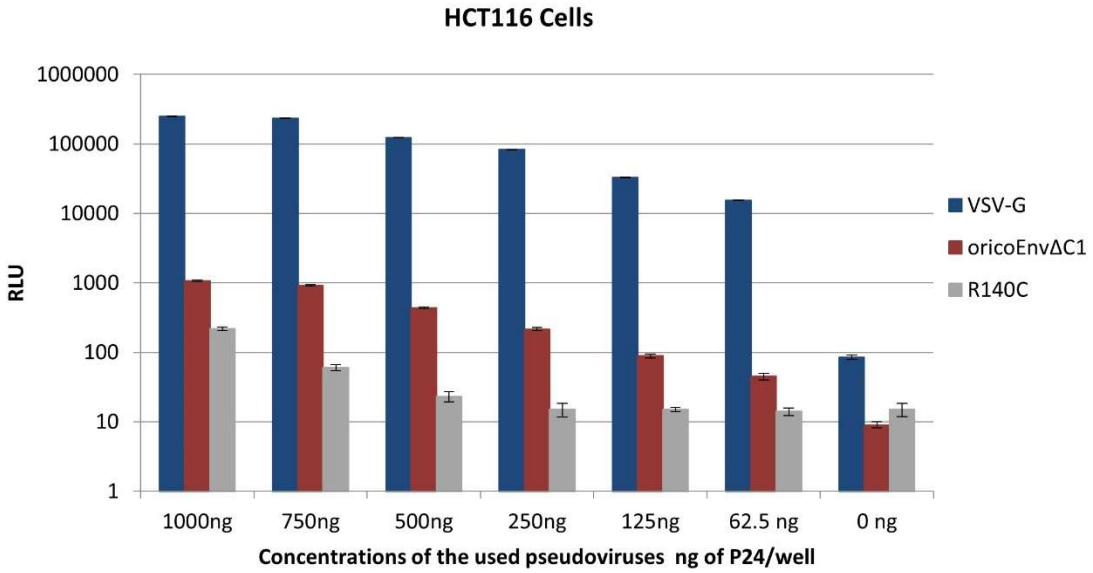

Figure 4-22 Titration assay in HCT116 cell line.
Cells were infected with 7 concentrations ranging from 0-1000 ng of p24 of HIV luciferase reporter viruses pseudotyped with the three Envs. Viruses with the VSV-G envelope were used as a positive control. The non-functional envelope K113 R140C Env served as a negative control. 72h post-infection luciferase assay was performed to measure the entry. Each bar represents the mean of three replicates (error bars= SD).

We performed a titration experiment to determine the titer of HML-2 Env pseudoviruses needed to infect HCT116 cells to assess the HML-2 Env viral entry during screening. HCT116 cells were infected with different titers ranging between (0 - 1000 ng/well of p24) of normalized pseudotyped HIV luciferase reporter viruses with K113 oricoEnvΔC1, VSV-G, and K113 R140C Env. The last two Env viruses served as positive and negative controls, respectively. Results in Figure 4-22) showed that the proper concentration of HML-2 Env (K113 oricoEnvΔC1) for the screening was between 250-500 ng of p24 (at a MOI=200), where the viral entry was specific and yet distinct from the background represented by the entry of the K113 R140C Env pseudovirus at the same MOI. Accordingly, 250 ng p24 of HML-2 Env pseudovirus was chosen to infect HCT116 cells expressing library genes.

4.3.2. Library Screening

For each plate of the library, the previously optimized protocol was performed. Briefly, HEK293T cells were seeded in 96-well plates and transfected with expression clones and lentiviral packaging plasmids (methods section 3.2.2.3) to produce viral supernatant. pMax-GFP expression plasmid was transfected into HEK293T cells in 6-well plates as a transfection control for each transfection process. 48h-72h post-transfection GFP expression was monitored by fluorescence microscopy, and viral supernatants were collected from all wells and stored at -80˚C till screening.

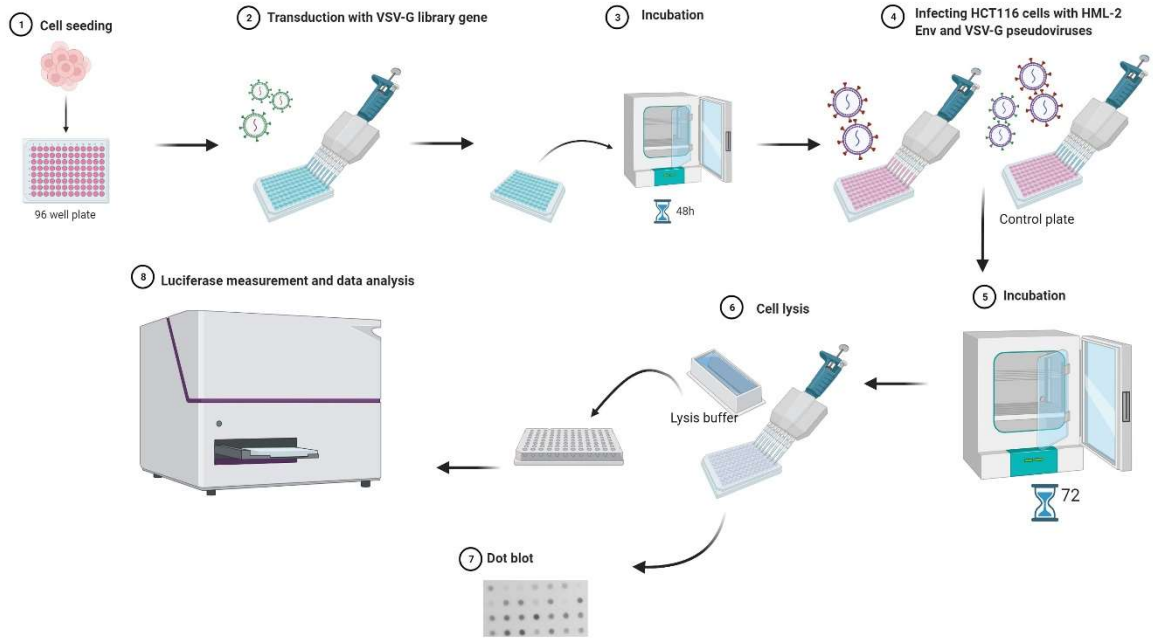

Figure 4-23 Diagram summarising the working flow of the library screening.
Plasmid DNA Extraction was performed in 96-microwell plate format and concentrated by Pico green assay. The extracted library plasmids were combined with the viral packaging plasmids to transfect the HEK293T cells for virus production. Viral supernatant was collected in fresh plates and used to transduce HCT116 cells and allowed to express the transmembrane proteins for 48h. Cells were then infected with HML-2 Env and the control VSV-G pseudoviruses with luciferase reporter as the readout for infection efficiency. 72h post-infection, cells were analyzed for luciferase activity, and dot blot was used to detect the expressed proteins. Created with BioRender.com.

Next, we performed the library screening (Figure 4-23) by transducing HCT116 cells with the VSV-G pseudoviruses carrying the cDNA library genes. Cells in the control wells were infected with VSV-G pseudoviruses carrying ovalbumin instead of the cDNA library gene. All wells were allowed to express the transduced genes for 48h before being infected with the luciferase reporter pseudoviruses with HML-2 Env represented by K113 oricoEnvΔC1. The control

ovalbumin-expressing cells were infected with luciferase reporter pseudoviruses with HML-2 Env and VSV-G.

The cells were lysed 72h post-infection, and luciferase activity was measured. Dot blot was also used to analyze the lysates as a confirmatory check using an anti-V5 antibody since all proteins were tagged with the V5-tag. The method detected most expressed proteins in each plate, but not all in some cases due to the high background and the antibody used.

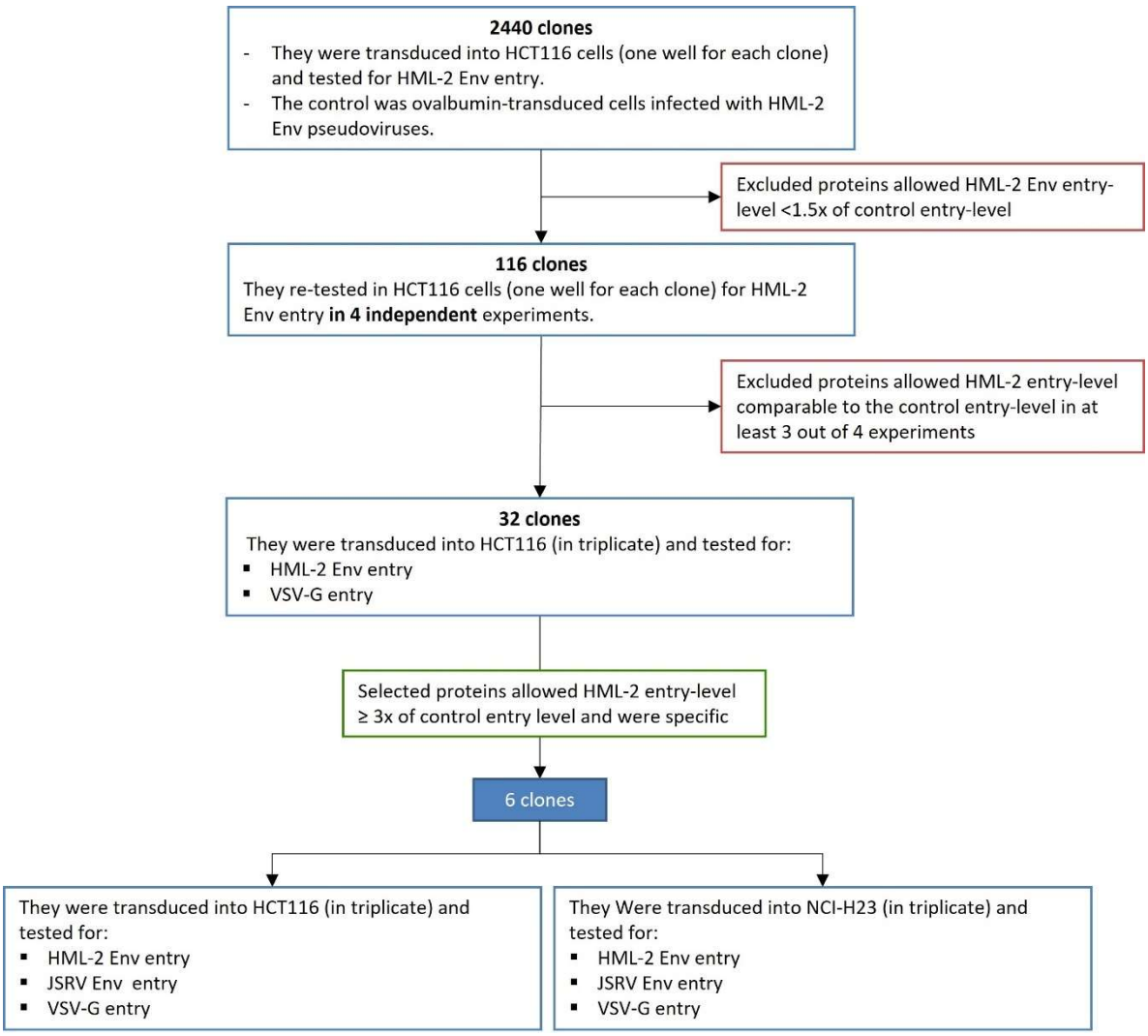

Figure 4-24 The cDNA library screening strategy and hits selection criteria

Initial screening results showed that HML-2 Env entry-levels over the entire library ranged between 1x to 7x the control entry-level. The cutoff was set on 3x of the HML-2 Env control entry-level. However, we observed variations in the HML-2 Env control entry-level in different experiments, most likely resulted from batch-to-batch variations in the ratio of viral particles to infectious ones (Tatsis *et al.*, 2006). We also considered the variations in the transduction efficiency of the library pseudoviruses caused by the high and low titer-producing clones.

Therefore, we decided to re-test all proteins that allowed HML-2 Env entry ≥1.5x control entry-level. They were 116 proteins. Figure 4-24 shows the screening strategy.

We transduced HCT116 cells with relatively similar concentrations of the normalized virus supernatants of the 116 selected clones. Then, we re-evaluated them for HML-2 Env entry in four independent experiments (one well for each clone). The excluded proteins allowed HML-2 Env entries comparable to the control in at least three experiments.

We assessed the remaining proteins that allowed entry-level ranging between (2x- 7x the control) in triplicates for HML-2 Env and VSV-G pseudoviruses entries. The short-listed proteins were only those whose HML-2 Env entry-level was above the cutoff (= 3x of control entry-level) and specific for HML-2 Env but not VSV-G.

4.3.3. Testing short-listed hits

The short-listed hits included six transmembrane proteins encoding genes. They showed a high entry-level of HML-2 Env among all tested clones compared to the control. Confirmatory sequence analysis was performed on their expression plasmids. The six transmembrane proteins were: the transmembrane protein 9 (TMEM9), chromosome 12 open reading frame 59 (C12ORF59), also known as Transmembrane protein 52B (TMEM52B), interleukin 1 receptor accessory protein (IL1RAP), prostate stem cell antigen (PSCA), LETM1 domain-containing protein 1(LETMD1) and Macrophage-expressed gene 1 (MPEG1) or Perforin-2 (PRF2).

The proteins C12ORF59, IL1RAP, PSCA, and MPEG1 are plasma membrane proteins (Gerhard *et al.*, 2004; Greenfeder *et al.*, 1995; McCormack *et al.*, 2015; Reiter *et al.*, 1998). MPEG1 protein is also known to localize to the endosomes /lysosomes (McCormack *et al.*, 2015). TMEM9 is an intracellular transmembrane protein localized on lysosomes (Kveine *et al.*, 2002). Finally, LETMD1 is an outer mitochondrial transmembrane protein (*GeneCards (RRID:SCR_002773) the human gene database*; Snyder *et al.*, 2021; Stelzer *et al.*, 2016). Table 4-1 lists their functions.

We recruited the JSRV Env to investigate whether the enhancements of HML-2 Env entry in the cells expressing these proteins were specific, and the expressed proteins do not influence the entry of other betaretroviruses.

We transduced the HCT116 cells with an equal concentration of the normalized VSV-G pseudoviruses carrying the six candidate clones. We included the control VSV-G ovalbumin as well. 48h post-transduction, we detected the expressed proteins by Western blot analysis

(Figure 4-25; A) and infected the cells with normalized luciferase reporter pseudoviruses carrying HML-2 Env, JSRV Env, and VSV-G. Then, we measured the luciferase activity in the 72h post-infection lysates. Entry of HML-2 Env in cells expressing the hit-proteins was significantly enhanced compared to the control entry-level (the ovalbumin expressing cells infected with HML-2 Env).

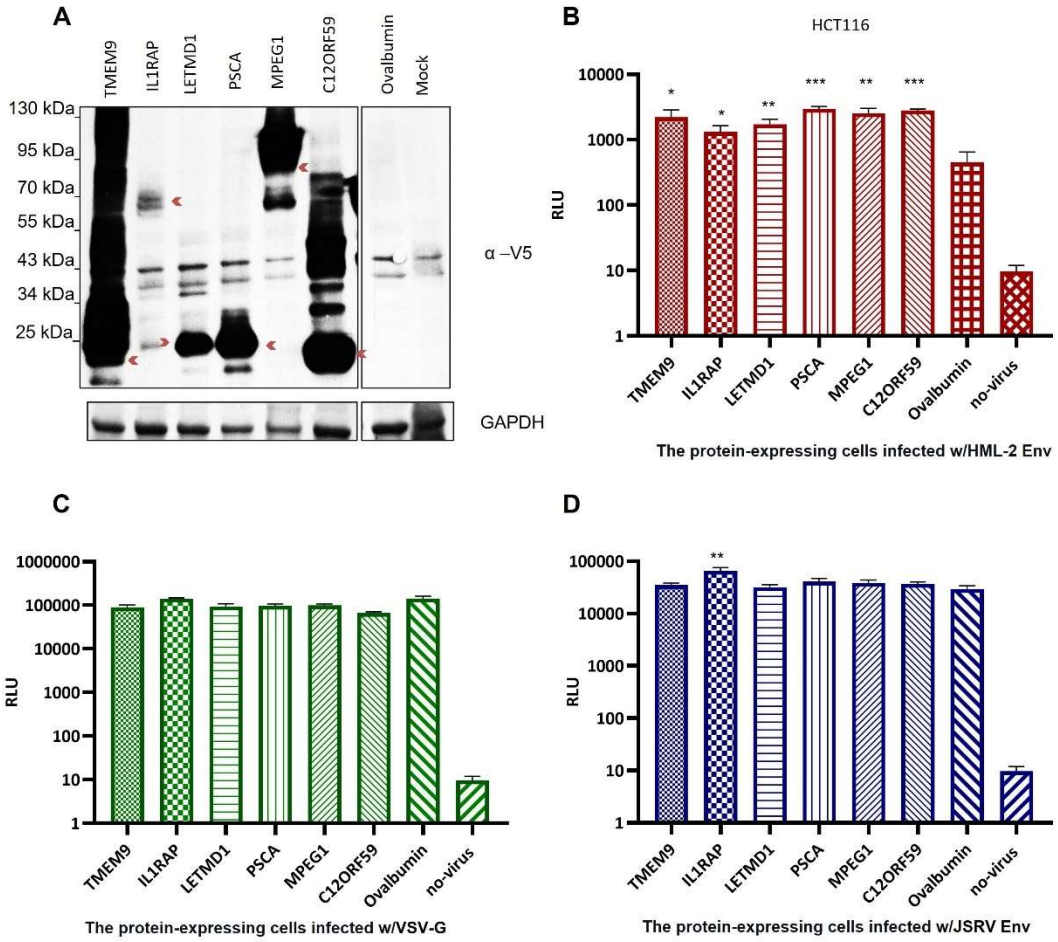

Figure 4-25 Entry of HML-2 Env, JSRV Env, and VSV-G pseudoviruses into HCT116 cells expressing the hit-proteins.

HCT116 cells were transduced with 45ng P24/well of VSV-G pseudoviruses carrying the hits' genes. VSV-G-ovalbumin was used as a control. Cells were then infected with the luciferase reporter pseudoviruses. ***A****. Expressions of proteins were detected 48h post-transduction by Western blot using an anti-V5 antibody. All proteins were expressed and showed the predicted molecular weight; Mock: non-transduced cells.* ***B****. (red) HCT116 cells expressing the hits proteins and ovalbumin infected with HML-2 Env pseudovirus.* ***C****. (green) HCT116 cells infected with VSV-G pseudovirus* ***D****. (blue) HCT116 cells infected with JSRV Env pseudovirus. Each bar represents the mean of three replicates (error bars=SD). *p < 0.05, ** p < 0.01, ***p < 0.001*

Table 4-1 The short-listed hits of the transmembrane proteins that confer the susceptibility of HML-2 Env pseudoviruses in HCT116 cells

Gene	Membrane type	Location	MW kDa.	Function
TMEM9	Single-pass type I	Lysosome membrane	21	Transmembrane protein that binds to and facilitates the assembly of lysosomal proton-transporting V-type ATPase (v-ATPase), resulting in enhanced lysosomal acidification and trafficking (Jung *et al.*, 2018)[1].
IL1RAP	Single-pass type I	Cell membrane	65	IL1RAP functions as a coreceptor with IL1R1 in the IL-1 signaling system (Greenfeder *et al.*, 1995), and a coreceptor for tyrosine kinase receptors FLT3 and c-KIT (Mitchell *et al.*, 2018). It is also a coreceptor for IL1RL2 in the IL-36 signaling system (by similarity) [1].
LETMD1	Single-pass type I	Mitochondrial outer membrane	41	LETMD1 is essential for mitochondrial structure and function and the thermogenesis of brown adipocytes (Snyder *et al.*, 2021).
PSCA	Lipid-anchor, GPI-anchor	Cell membrane	12	It is a glycosylphosphatidylinositol-anchored cell membrane glycoprotein, which may function as a modulator of the nicotinic acetylcholine receptors (nAChRs) activity (Jensen *et al.*, 2015)[1].
MPEG1	Single-pass type I	Cell membrane and cytoplasmic vesicle membrane	78	Macrophage-expressed gene 1 (MPEG1) /Perforin-2 (PRF2). Common in the cell membrane of macrophages, phagocytes, and others. It localizes to early endosomes, and phagosomes/lysosomes and forms pore in the membrane of engulfed bacteria (McCormack *et al.*, 2015).
C12ORF59	Single-pass type I	Cell membrane	20	unknown function[1].

[1] Obtained data from (*GeneCards (RRID:SCR_002773) the human gene database*; Stelzer *et al.*, 2016)

The highest entry was in the cells that expressed the GPI-anchored protein PSCA (7x) ($p < 0.001$), followed by (6x) MEGP1 ($p = 0.009$) and C12ORF59 ($p < 0.001$), then TMEM9 (5x) ($p = 0.03$), and lastly the mitochondrial outer membrane protein LETMD1 (4x) ($p = 0.008$) and IL1RAP (3x) ($p = 0.02$) (Figure 4-25; B).

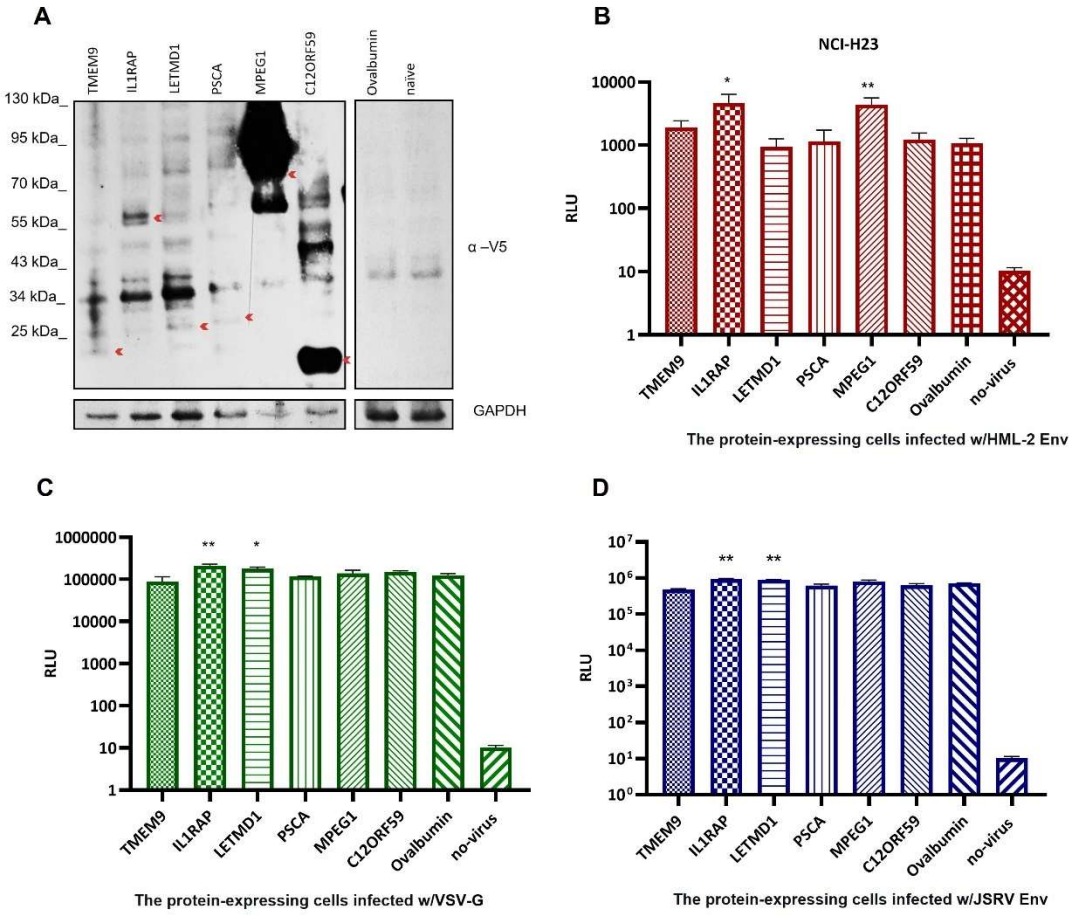

Figure 4-26 Entry of HML-2 Env, JSRV Env, and VSV-G pseudoviruses into NCI-H23 cells expressing the hit-proteins.

NCI-H23 cells were transduced with 45ng P24/well of VSV-G pseudoviruses carrying the hits' genes. VSV-G-ovalbumin was used as a control. Cells were then infected with the luciferase reporter pseudoviruses. ***A****. Expressions of proteins were detected 48h post-transduction by Western blot using an anti-V5 antibody. All proteins were expressed and showed the predicted molecular weight; Mock: non-transduced cells.* ***B****. (red) NCI-H23 cells expressing the hits proteins and ovalbumin infected with HML-2 Env pseudovirus.* ***C****. (green) NCI-H23 cells infected with VSV-G pseudovirus.* ***D****. (blue) NCI-H23 cells infected with JSRV Env pseudovirus. Each bar represents the mean of three replicates (error bars=SD). *p < 0.05, ** p < 0.01.*

No changes in VSV-G or JSRV Env pseudoviral entry were observed with cells expressing the hits-proteins except for IL1RAP (Figure 4-25; C, D), in which a significant increase in JSRV

Env entry was observed ($p = 0.003$) (Figure 4-25;D). This suggested that the entry of HML-2 Env pseudovirus was specific in the presence of at least five expressed proteins.

To confirm the results above and rule out the cell line-specific effect, we evaluated the entry of the HML-2 Env pseudovirus in NCI-H23 cells, a poorly permissive cell line for HML-2 Env (N. Bannert, unpublished data) but permissive for both VSV-G and JSRV Env pseudoviruses (data are not shown). As in the previous experiment, we transduced the NCI-H23 with an equal concentration of the normalized VSV-G pseudoviruses carrying the library candidate genes. The expression of the proteins was confirmed after 48h by Western blot analysis using an anti-V5-tag (Figure 4-26; A). Then, cells were infected with luciferase reporter pseudoviruses with HML-2 Env, JSRV Env, and VSV-G. Only the cells expressing IL1RAP or MPEG1 proteins were significantly highly susceptible to HML-2 Env 8x ($p = 0.02$) and 6x ($p = 0.009$) control entry-level, respectively. Furthermore, TMEM9 protein allowed HML-2 Env pseudoviral entry ~3-fold of the control entry ($p > 0.05$) (Figure 4-26; B). In contrast, other cells expressing proteins did not confer susceptibility and were at comparable levels to the control ($p > 0.05$).

On the other hand, VSV-G and JSRV Env pseudoviruses showed comparable levels to their controls. However, significant enhancements in both of their entry levels were observed in cells expressing LETMD1 (VSV-G $p = 0.01$, JSRV Env $p = 0.001$) and IL1RAP (VSV-G $p = 0.006$, JSRV Env $p = 0.001$) proteins (Figure 4-26; C, D), but not in cells expressing the other proteins, suggesting that these proteins may contribute specifically to the virus infection cycle. Therefore, HML-2 Env entry was specific, at least in the cells expressing the MPEG1 protein. This protein might function in the entry process of HML-2 Env. Although the other proteins enhanced the entry of other viruses, this does not exclude them as playing a role in the HML-2 Env entry.

5. Discussion

Early events of the retroviral infection cycle include binding the viral envelope glycoprotein to the cell surface through receptors and coreceptors, which usually define the features of species and cellular tropism (Schneider-Schaulies, 2000). SU protein on the viral envelope mediates this interaction with at least one specific cellular receptor on the cell membrane. However, many viruses require more than one cell surface molecule to permit viral entry. While some of the required molecules provide attachment only, others promote signaling, induce plasma membrane ruffling, trigger changes in the viral particles, and activate endocytosis (Yamauchi & Greber, 2016). Thus, SU may hold many binding sites, as in MMTV, HIV-1, and HTLV (Clapham & McKnight, 2001; Lambert *et al.*, 2009; Zhang *et al.*, 2003).

In recent years, HML-2 has increased interest. The K108 Env was recognized to have a functional activity (Dewannieux *et al.*, 2006). K113 infectivity was restored by reversing some mutations in its Env (Hanke *et al.*, 2009). Moreover, the HML-2 Env is shown to mediate fusion under acidic pH, and it enters the cells by endocytosis through a dynamin-dependent pathway (Robinson & Whelan, 2016). Furthermore, it imparts broad species and tissue tropism, including human, rodent, dog, feline, and even reptile cells (Kramer *et al.*, 2016; Robinson & Whelan, 2016). Recently, studies demonstrated the indispensable role of GAGs in the virus entry and the ability of the HML-2 Env to bind these GAGs, especially HS (Geppert, 2019; Robinson-McCarthy *et al.*, 2018). Finally, the receptor-binding site was identified by mutational analysis (Wamara, 2020).
However, the HBD on HML-2 Env is still not identified, nor is the one or more cellular proteins that serve as entry receptor(s). This study aimed to identify them using different approaches.

5.1. The HBD of HML-2:

Viral attachment to the surface of a target cell is the first step of the virus-cell complex interactions. Some of these interactions are weak and electrostatic, aimed to increase virus concentration at the cell surface. Consequently, increasing their chances of binding the specific entry receptor by establishing the other type, the specific interactions (Rusnati *et al.*, 2009). Engagement of viruses with cellular GAGs falls into the first type of interactions, and it is common among viruses, especially with HS, the most prevalent cell-surface GAG (Ihrcke *et al.*, 1993; Li *et al.*, 2021).

HML-2 and other retroviruses like HTLV, HIV-1, and MMTV exploit the cellular HS as an attachment factor. HIV-1 and MMTV bind HS through specific HBDs rich with basic residues

(Crublet *et al.*, 2008; Jones *et al.*, 2005; Patel *et al.*, 1993; Robinson-McCarthy *et al.*, 2018; Zhang *et al.*, 2003). However, binding HS does not necessarily require specific linear patterns of basic residues. It could be achieved through a spatial structural motif, in which the basic residues could be close to each other in the space but not in the linear protein sequence (Margalit *et al.*, 1993).

Using HP-coated beads and HML-2 Env pseudoviruses with mutations in the RBS, we confirmed that HML-2 Env uses a different region of its envelope to bind HS other than the RBS, which is not required for HS-interaction. Additionally, the Env could still bind HP-beads even after removing N-glycosylations, which are very important to binding the receptor (Li *et al.*, 2021).

Then, by aligning the envelope protein sequence of HML-2 to the relative MMTV, we identified a predicted domain to bind HS within the SU protein sequence. The HML-2 HBD did not exhibit one of the mammalian consensus sequences like MMTV HBD (Zhang *et al.*, 2003). However, the pattern of the predicted domain falls into HBDs and is akin to that of HSV-1 glycoprotein C (Mårdberg *et al.*, 2001) and respiratory syncytial virus (RSV) glycoprotein G (Feldman *et al.*, 1999). The predicted domain included a cluster of basic residues with a single isolated basic amino acid XBX, a common pattern in HBDs (Fromm *et al.*, 1997). The spacing between the basic residues was between (1, 2, and 3 non-basic residues), which is still relatively common. That and spacers with a length of up to 3 non-basic residues have the highest affinity to HS (Fromm *et al.*, 1997). Through the three-dimensional protein structure of HML-2 Env predicted by the Alpha fold (Jumper *et al.*, 2021; Varadi *et al.*, 2022), we found that the predicted domain is folded in a predicted β-strand (Figure 4-5), in contrast to MMTV HBD (Zhang *et al.*, 2003). HBD's molecular modeling studies show that the common sequences XBBXBX and XBBBXXBX are folded into a β-strand and α-helix conformations, orienting the basic amino acids in them on one face or one side and pointing the hydropathic residues back into the protein core or the other side (Cardin & Weintraub, 1989; Margalit *et al.*, 1993).

Involvement of this domain in HS-interaction was investigated by inducing mutations in the basic amino acid clusters and producing Envs that have these mutations incorporated into viral particles. These viral particles were then tested for HP binding and viral entry. The entirely mutant domain failed to bind HP-beads and thereby enter the cells. This proved that the predicted HBD is the primary domain to bind HS. On the other hand, the entry inhibition might result from folding changes in the Env affecting its stability since we detected slight differences in processed Env on WB (Figure 4-7).

Nevertheless, the double mutation in two lysine residues K(223,225)A separated by one non-basic residue spacer glycine showed an extreme reduction in binding HP than others. However, it did not influence Env-mediated viral entry, suggesting compensations by binding other types of cellular GAGs, consistent with the findings of Robinson-McCarthy and coworker (Robinson-McCarthy *et al.*, 2018). Moreover, Envs with single mutations in the same lysine residues K225A and K223A bound HP-beads efficiently. However, K223A mutated Env showed less affinity to HP-beads, which appeared as a reduction in viral entry.

HBDs can bind other GAGs, especially those that show structural similarity to HS, like chondroitin sulfate and DS (Radek *et al.*, 2009; Trowbridge *et al.*, 2002; Xu & Esko, 2014). For HML-2, chondroitin sulfate and DS do not affect its Env-mediated entry (Robinson-McCarthy *et al.*, 2018). However, changing the arrangement of positively charged residues enables the protein to tolerate different negative-charge patterns in HS or other GAGs (Carter *et al.*, 2005).

On the other hand, these results also suggest an involvement of other residues or another domain on HML-2 Env, which also explains the trace bands on the Western blot analysis when the entirely mutant domain or the double mutant K(223,225)A envelope was used (Figure 4-8).

The non-basic residues like serine and glycine in HBDs are frequent and found explicitly in domains interacting with HS. They have small side chains and provide good flexibility for protein interaction with GAGs (Caldwell *et al.*, 1996; Pearson, 1963). In HML-2 Env HBD, there are two serine residues and one glycine residue that we did not mutate in this study and might play an important role in HS binding. Moreover, the glycine residue is the non-basic spacer found between the two mutated lysine residues K_{223} and K_{225}. Since the positive charge density and spacing play a critical role in binding HS (Caldwell *et al.*, 1996), increasing spacing would decrease the affinity to HS (Fromm *et al.*, 1997). Consequently, these three residues, K_{223}, G_{224}, and K_{225}, may play a significant role in binding HS specifically.

Arginine residues in HBDs are frequent and known to bind sulfate groups about 2.5x more tightly by forming stable hydrogen bonds and stronger electrostatic interactions (Fromm *et al.*, 1997; Pearson, 1963). The ratio of arginine to lysine residues defines the affinity to HS (Fromm *et al.*, 1997; Hileman *et al.*, 1998). Accordingly, mutations in any arginine residue would be reflected at the affinity to HS. So, the Envs with mutations R216A and double sites mutation K219A/R221A were still able to bind HP-beads but showed a reduction in viral entry due to a reduction in HS affinity.

Other mutations in the domain, including K(225,229)A, K229A, K(229,233)A, and K(233,236)A, did not influence Env-mediated viral entries significantly. They showed comparable levels to the wild-type Env. The K229 residue is conserved with MMTV HBD and seemed to be involved in increasing the viral entry by altering the specificity to HS. However, its effect in the tested cell line HEK293T was not significant.

5.2. GP73 protein as a potential attachment factor

A traditional method to identify viral receptors is using the SUs as ligands to pull down their receptors based on their affinity. This method was used to discover virus receptors within a short time, especially emerging viruses like the middle east respiratory syndrome-related coronavirus (HCoV-EMC) (Mei *et al.*, 2015; Raj *et al.*, 2013). However, monomeric forms of the SU fusion proteins may show lower affinity to the receptors than the native state oligomeric spike proteins. The trimeric ligands have a higher affinity to the receptors than the monomeric ones, especially if the receptor molecule is a dimer (Lauffer *et al.*, 1995).

In this part of our study, we utilized stabilization strategies based on the synthetic trimerization T4-fibritin motif to produce a recombinant trimer that mimics the native trimer on the viral Env and provides a native-like binding epitope with the highest affinity to the specific receptor in order to identify it.

The two designed versions of HML-2 soluble fusion proteins SU-FT and SM-FT were efficiently expressed and secreted from the cells. However, the purified SM-FT fusion protein, which has its CS eliminated, tended to aggregate, adopting aberrant structures that are not similar to the native trimer. This was experienced in both SM-FT tagged proteins with hFc- tag and V5-tag. One research group reported a similar issue on HIV 140gp trimer, which has its cleavage site eliminated and stabilized with FT. The problem was linked mainly to the prevention of the cleavage (Klasse *et al.*, 2013).

The aggregation problem might also result from the elution step in the purification procedure of Fc-tagged proteins, which depends on brief exposure to low pH (Arosio *et al.*, 2011). However, some elution buffers, such as acetate and glycine buffer, as well as Tris–HCl buffer at pH 7.2 - 8, can reduce the aggregation to the minimum (Joshi *et al.*, 2014). Both of these were utilized in the purification of our Fc-tagged proteins.

Furthermore, the aggregation might have been caused by conformational changes in the TM subunit caused by brief exposure to low pH (Robinson & Whelan, 2016; Zhao *et al.*, 1998), which could influence the SM-FT-Fc but not the SM-FT-V5 since the latter was purified using a different method that did not include low pH-dependent elution.

However, the experienced reduction in both the melting temperature of the purified fusion proteins and the stability of the Fc domain can be linked to the purification method (Ejima *et al.*, 2007; Welfle *et al.*, 1999).

On the other hand, we did not experience aggregation in the fusion protein SU-FT-hFc, which was oligomerized in a trimer by the FT-motif (Figure 4-12). Without the synthetic trimerization domain FT, SU protein was expressed and secreted as a monomer (Wamara, 2020) similar to the soluble fusion protein of JSRV (Vigdorovich *et al.*, 2005). The trimer SU-FT fusion protein showed chemical stability in the presence of β-mercaptoethanol as previously reported for the stabilized HIV trimer fusion proteins with the FT-motif (Yang *et al.*, 2000; Yang *et al.*, 2002). The proper SU glycosylation is essential for proper folding to expose the RBS (Kayman *et al.*, 1991; Li *et al.*, 1993; Pinter *et al.*, 1984). The deglycosylation experiment shifted the SU-FT trimer protein by ~14 kDa. This difference between the glycosylated and non-glycosylated protein indicated that the SU-FT trimer was glycosylated similarly to the native trimer (Hanke *et al.*, 2009).

In the GAGs binding experiment, the SU-FT trimer was the only protein that bound HP-beads in a similar affinity to that of the Env trimer on the pseudoviruses, in keeping with the previous results on the viral particles (Geppert, 2019; Robinson-McCarthy *et al.*, 2018).

However, no interaction occurred with the monomeric SU form. In contrast to previous study findings on K108 fusion protein (Robinson-McCarthy *et al.*, 2018). Our results suggest that the oligomer is required to form the HBD that binds the cellular HSPG. Similarly, differences in affinity to α4β7 integrin between HIV Env monomers and trimers were reported (Chand *et al.*, 2017). Moreover, HS binding to oligomers is well-known among HS ligands, especially chemokines. The CXCL and CCL chemokines exist as dimers. However, both are required to bind HS. Thus, HBD is, in this case, formed from a tetramer (Lau *et al.*, 2004). The high affinity between oligomers and HS involves better stability and protection of the trimeric structure (Salanga & Handel, 2011). Most importantly, it guarantees interaction with the entry receptor in the oligomeric state (Gómez Toledo *et al.*, 2021).

Since the trimer fusion protein SU-FT was able to bind HS, binding it to the cell surface should involve interactions with HSPG. This binding almost vanished upon enzymatic removal of the cellular HSPG, affirming the role of HSPG as the primary attachment factor of HML-2 Env (Geppert, 2019; Robinson-McCarthy *et al.*, 2018).

In the viral inhibition assay, binding of SU-FT fusion protein to the cell surface was not enough to block or inhibit the infection of pseudoviruses with HML-2 Env. Unlike XMRV Env

infection, which was blocked by its fusion protein XSU. The fusion protein XSU competed with XMRV viral particles to bind the specific entry receptor XRP1 leading to a dose-dependent inhibition. The fusion proteins of XMRV, JSRV, enzootic nasal tumor virus (ENTV), and other retroviruses were reported to bind the specific receptor and block the Env-mediated entry since the SUs possess the RBS. However, these viruses used one receptor molecule for their HP-independent entries (Battini *et al.*, 1999; Côté *et al.*, 2011; Côté *et al.*, 2012; Xu & Eiden, 2011).

In the case of HIV infection, many receptors are required for its entry (Bleul *et al.*, 1997; Feng *et al.*, 1996; McDougal *et al.*, 1986). It is well established that both of its envelope subunits are required with a covalent link to inhibit HIV entry (Kagiampakis *et al.*, 2011), which might be required in the case of HML-2 Env to block its viral entry. Consequently, this emphasizes the roles of multiple receptors in HML-2 Env entry.

After all, using the trimer fusion protein SU-FT, we reported the isolation and identification of GP73 protein as a potential binding factor for HML-2 Env. The functionality of the utilized method was confirmed by the control XSU, in which the XPR1 receptor was detected by Western blot (Figure 4-16; B).

In the cell surface staining experiment, the FACS analysis for the GP73-transfected and non-transfected cells showed a significantly strong correlation between fold changes of the cell surface-bound SU-FT and fold changes of the expressed plasma membrane-GP73 (Figure 4-19). Nevertheless, increasing the plasma membrane-GP73 did not boost HML-2 Env pseudoviral entry (Figure 4-20).

The overexpression of GP73 in MALME-3M and NCI-H522 cell lines led to a decrement in the cycled GP73 to the plasma membrane, which coincided with a decrement in the cell surface-bound SU-FT. The decrement in plasma membrane-GP73 might result from the protein secretion induced by abnormal expression levels of GP73 (Bachert *et al.*, 2007; Hu *et al.*, 2011). The expression pattern of GP73 differs among different cells and tissues (Donizy *et al.*, 2016; Fritzsche *et al.*, 2008; Liu *et al.*, 2019). The overexpression of GP73 in abundant plasma membranes and endosomes causes more amounts to leave the early Golgi compartment and come into contact with proteases, including furin, leading to increased cleavage and ectodomain release (Bachert *et al.*, 2007). However, it is unclear why some cells increased the cycled GP73 to the cell surface. It might be increased in plasma membranes with sparse cell membrane-GP73. Regardless, this was noticed in the SK-MEL-5, NCI-H226, SK-MEL-2, and HEp-2 cell lines.

The increase in the amount of bound SU-FT to the cell surface upon GP73 overexpression suggests the involvement of GP73 in binding SU-FT at the cell surface.

However, NCI-H226 and SK-MEL-2 cell lines with increased plasma membrane-GP73 showed decreased amounts of the cell surface-bound SU-FT, suggesting the involvement of other factors or glycans, which seemed to be impacted by the overexpressed-GP73. We already know that HML-2 Env attaches to the cell surface through binding HSPG (Geppert, 2019; Robinson-McCarthy *et al.*, 2018). Over expression of GP73 can enhance the epidermal growth factor receptor (EGFR) levels (Yan *et al.*, 2020; Ye *et al.*, 2016). EGFR and Golgi membrane phosphoprotein 3 (GOLPH3), another transmembrane protein of Golgi apparatus, are negatively correlated (Arriagada *et al.*, 2020; Blagoveshchenskaya *et al.*, 2008). GOLPH3 modulates the activities of many glycosylation enzymes, especially glycosyltransferase enzymes encoded by the exostosin (EXT1/2) genes (Chang *et al.*, 2013; Welch *et al.*, 2021). These enzymes are necessary for the biosynthetic process of HSPG (Esko & Selleck, 2002). Up- or down-regulation of GOLPH3 would disrupt this process resulting in incomplete HSPGs (Chang *et al.*, 2013). However, different cancer cell lines show different levels of GOPLH3 expression and functional heterogeneity in their effects (Arriagada *et al.*, 2020). Based on the above, The SKMEL-2 cell line that expresses a high amount of HS on its surface (Baljinnyam *et al.*, 2011) allows a high trimer SU-FT binding rate the surface of the non-transfected cells. The overexpression of GP73, however, resulted in a reduction in the bound SU-FT, most likely caused by losing HS.

Although this cell line bound the trimer protein SU-FT at a high level, it is still low-permissive for HML-2 Env entry. Similarly, non-permissive cell lines for the HTLV virus can bind its SU fusion protein. These cells might be deficient in expressing the entry receptor or have some factors that prevent the viral entry at the late post-receptor-binding stage (Jassal *et al.*, 2001; Jones *et al.*, 2002; Li *et al.*, 1996; Manel, Kinet*, et al.*, 2003; Okuma *et al.*, 1999; Sutton & Littman, 1996).

Although GP73 did not render HML-2 Env entry, these findings do not preclude the involvement of GP73 in the attachment of HML-2 Env to the cell surface. HML-2 Env might bind it as a secondary attachment factor to increase the attachment to the cell surface, or to compensate for the lack of HS, especially since GP73 is a sulfated glycoprotein that carries 4′-*O*-sulfated LacdiNAc (Toyoda *et al.*, 2016), and HML-2 Env binds the sulfate groups in HS (Robinson-McCarthy *et al.*, 2018). This is consistent with previous observations that suggested HML-2 Env can bind other sulfate groups in the absence of HS (Robinson-McCarthy *et al.*, 2018). Not to mention that the biosynthetic pathway of the LacdiNAc group involved

glycosylation enzymes that show high homology to chondroitin sulfate synthase (Gotoh *et al.*, 2004; Sato *et al.*, 2003). Furthermore, chondroitin sulfate, which shows a relative structure to HS (Munoz & Linhardt, 2004), can bear 4'-*O*-sulfation or 6-*O*-sulfation (Plaas *et al.*, 1997). However, further investigation is still required to prove this.

No previous study indicated whether HML-2 Env binds one or several glycans, but Robinson-McCarthy *et al.*, 2018 reported that HML-2 Env binds HS via 2-*O* sulfation (Robinson-McCarthy *et al.*, 2018). However, for HIV-1, it has been reported that all syndecans can bind it (Gallay, 2004; Saphire *et al.*, 2001) via 6-*O*-sulfation (De Parseval *et al.*, 2005).

The LacdiNAc glycan is a minor disaccharide (GalNAcβ1–4GlcNAc-R) on N-glycans in mammalian cells. It can occur within 4-*O*-sulfated, α1,3-fucosylated, or α2,6-sialylated forms (Kawar *et al.*, 2005). It is protein-specific (Benicky *et al.*, 2019), recognized by various receptors and known to play a vital role in signaling pathways, including regulating the circulatory half-life of pituitary glycoprotein hormones, cell identification, and others (Baenziger *et al.*, 1992; Hirano & Furukawa, 2022; Manzella *et al.*, 1996; Mi *et al.*, 2002). Furthermore, sialylated LacdiNAc can be recognized and bound by most Influenza viruses as their receptor, including the H3N2 virus (Stevens *et al.*, 2010), the seasonal H1N1 influenza virus (Xu *et al.*, 2013), and the droplet-transmissible H5N1 strain (De Vries *et al.*, 2014; Jia *et al.*, 2014). Helicobacter pylori bacteria can also bind LacdiNAc glycan in the mucin layer of the human stomach mucosa to establish their colonies (Kenny *et al.*, 2012; Matos *et al.*, 2021; Paraskevopoulou *et al.*, 2021).

Although the viral Envs and glycans interactions are classified as electrostatic low-affinity interactions, they still imply specificity. However, they are not always essential for viral infectivity (Maginnis, 2018; Marsh & Helenius, 2006; Stencel-Baerenwald *et al.*, 2014; Yamauchi & Helenius, 2013).

On the other hand, GP73 is a heterotypic trafficking protein. It traffics in two pathways, from ER to Golgi and bi-directional traffic between endosomes and the TGN (Puri *et al.*, 2002). It plays a role in protein sorting, modification, and cargo transport. GP73 was reported to interact with EGFR/RTK as a specific cargo adapter and assist in cycling it to the plasma membrane (Ye *et al.*, 2016). Accordingly, the interaction between HML-2 Env and GP73 might also suggest the involvement of GP73 in the late HML-2 replication cycle by serving as a specific cargo adapter for HML-2 Env in trafficking from TGN to the plasma membrane during assembly. This would be supported by the fact that GP73 expression is up-regulated upon viral infections (Kladney *et al.*, 2000). In the infected cells, viruses hijack the cellular gene-

expression machinery to express their genes in addition to other cellular genes essential for their life cycles (Hassan *et al.*, 2021; Rabson & Graves, 2011; Risco & Fernández de Castro, 2013). They create a cellular environment that facilitates replication by up-regulating some genes and down-regulating others (Mattoscio *et al.*, 2013; Reza Etemadi *et al.*, 2017; Wie *et al.*, 2013; Yang *et al.*, 2017). They also hijack the membrane trafficking system to promote their replication. HIV and other enveloped RNA viruses that assembled their viral particles at the plasma membrane hijack the post-Golgi trafficking pathway to transport their Envs using the cycled endosomes or directly from TGN to the plasma membrane (Coller *et al.*, 2012; Hassan *et al.*, 2021; Robinson *et al.*, 2018; Taylor *et al.*, 2011).

5.3. Identification of potential HML-2 receptors

In the last part of this study, we recruited another strategy to identify HML-2 entry receptors and any other proteins required in the early entry pathway. Using a subset library that includes human cDNA transmembrane ORF collection, we conducted the screening by exposing the low-permissive cell line HCT116 to VSV-G pseudoviral vectors carrying the coding genes of transmembrane proteins.

We identified, for the first time to our knowledge, six human transmembrane proteins (TMEM9, C12ORF59, IL1RAP, PSCA, LETMD1, and MPEG1) that confer susceptibility to HML-2 Env pseudoviruses. Two proteins (TMEM9 and LETMD1) are intracellular membrane proteins (Kveine *et al.*, 2002; Snyder *et al.*, 2021), and the remaining (C12ORF59, IL1RAP, MPEG1, and PSCA) are plasma membrane proteins. The latter is a GPI anchor protein. (Gerhard *et al.*, 2004; Greenfeder *et al.*, 1995; McCormack *et al.*, 2015; Reiter *et al.*, 1998). Elevated levels of HML-2 Env pseudoviral infection in HCT116 cells were significantly associated with the individual expression of each identified transmembrane protein.

Nevertheless, only IL1RAP and MPEG1 proteins were shown to render the entry process in the poor-permissive NCI-H23 cells, suggesting that the potential receptor and/or co-receptor might be one or both of them. However, that does not disqualify the other proteins from being involved in the entry process since they also confer HML-2 Env susceptibility in the other low-permissive cell line (HCT116). These results also suggest that the six proteins may serve as receptors or contribute directly or indirectly to the HML-2 Env entry process in a cell-dependent manner. This agrees with previous findings that suggest different entry requirements for HML-2 Env in different cell types and a flexible entry route utilizing multiple cellular uptake pathways (Robinson-McCarthy *et al.*, 2018; Robinson & Whelan, 2016). Regardless, none of the identified proteins were reported before to function as a receptor for other viruses.

TMEM9 is a type I transmembrane protein. It is a member of a phylogenetically conserved protein family. It localizes to the late endosomes and lysosomes (Kveine *et al.*, 2002) and functions as a vesicular acidification regulator through binding to the vacuolar proton pump, the vacuolar-ATPase (V-ATPase), and facilitating its assembly. As a result, the vesicular acidification and trafficking are enhanced (Jung *et al.*, 2018). HML-2 Env-mediated entry is dependent on the endosomal uptake and acidic pH, which is implied by the inhibition of HML-2 Env pseudoviral infections up on treatment with V-ATPase inhibitors such as Bafilomycin (Robinson & Whelan, 2016) (N. Bannert, unpublished data). TMEM9 promoted HML-2 Env pseudoviral infection ~ 5 folds in HCT116 and 3 folds in NCI-H23. Therefore, TMEM9 might contribute indirectly to HML-2 Env pseudoviral infection via increasing the endosomal uptake and the acidification in the endocytic pathway. Similarly, increasing the caveolar vesicle trafficking upon expression of the EH-domain–containing protein 2, a dynamin-related ATPase, leads to an increase in simian virus 40 (SV40) infection efficiency (Cossart & Helenius, 2014; Stoeber *et al.*, 2012). Accordingly, the TMEM9 protein might be involved in the HML-2 late entry pathway rather than being a receptor for HML-2 Env.

PSCA is a GPI-anchored membrane protein located in a lipid raft of the cell membrane. It belongs to the urokinase-type plasminogen activator receptor (uPAR) proteins. The family includes GPI-anchored membrane proteins that share the presence of the LU domain resulting from a distinct pattern of disulfates bridges between the cysteine residues (Behrendt, 2004; Gumley *et al.*, 1995; Reiter *et al.*, 1998; Sharom & Radeva, 2004; Tsetlin, 2015). Proteins of this family were identified first in mouse lymphocytes before the proteins homology was isolated from humans (McKenzie *et al.*, 1977). PSCA is expressed in different normal human tissues. However, it shows an organ-dependent expression pattern in cancer cells (Ono *et al.*, 2012). It is involved in cell proliferation and differentiation (Li *et al.*, 2017; Ono *et al.*, 2018), as well as immunity through down-regulating IL-8, IL-1 receptor antagonist, and S100 calcium-binding proteins A8 and A9 (Saeki *et al.*, 2015). In addition, it modulates nicotinic acetylcholine receptors in the brain (Miwa *et al.*, 2019).

The lymphocyte antigen-6 (LY6E) protein, a member of the uPAR family, was reported to enhance the viral infectivity of many viruses, including HIV-1, Zeka virus, West Nile virus, and yellow fever virus (Loeuillet *et al.*, 2008; Mar *et al.*, 2018; Yu *et al.*, 2017). PSCA, the only protein from the family that shares a common sequence and structures with Ly6E protein, was also found to promote the West Nile and yellow fever viral infections (Mar *et al.*, 2018; Yu *et al.*, 2017). The promotion of viral infection was linked to the involvement of the protein in enhancing the uncoating after endosomal escape (Mar *et al.*, 2018).

However, the mouse LY6E, the homologs protein to human LY6E, functions as a receptor for the syncytin A protein (Bacquin *et al.*, 2017), which is an endogenous retroviral envelope protein expressed in mice and homologs to syncytin 1 (HERV-W) in humans. Syncytin A and syncytin B proteins form the placenta in the syncytiotrophoblast mice cells, where they specifically express (Dupressoir *et al.*, 2009; Dupressoir *et al.*, 2011; Hughes *et al.*, 2013). Furthermore, they are homologs to syncytin 1 and syncytin 2 in humans (Dupressoir *et al.*, 2005). Accordingly, the PSCA protein might be involved in the HML-2 entry pathway, but it might also serve as an entry receptor.

LETMD1 is an integral membrane protein in the outer membrane of the mitochondria. It is involved in cellular proliferation and survival, including energy metabolism, ROS generation, Ca^{2+} homeostasis, functional polarization, and cellular respiration (Morciano *et al.*, 2016; Van den Bossche *et al.*, 2016). It is also known as the human cervical cancer gene (HCC-1) since increased levels of its expression were observed in the serum of cancer patients (Cho *et al.*, 2007). A recent study revealed an immunomodulatory role of LETMD1 protein in suppressing phagocytosis and migration of macrophages and promoting the activation of the nuclear factor "kappa-light-chain-enhancer" of activated B-cells (NF-κB), MAP kinases, p53, and JAK-STAT. Furthermore, it has been shown that LETMD1 specifically regulates Lipopolysaccharide (LPS)-induced activation of NF-κB via myeloid differentiation primary response 88 (MyD88) but not via TIR-domain–containing adapter-inducing IFN-b (TRIF) (Lim *et al.*, 2020). MyD88 protein is known to be recruited as a signaling adapter by Toll-like receptors (TLRs), except TLR-3, to activate NF-κB and MAP kinases (Bovijn *et al.*, 2013; Hennessy *et al.*, 2010; Kawai & Akira, 2010; O'Neill & Bowie, 2007). However, the NF-κB pathway is an attractive target to viral pathogens. Once activated, viral replication is enhanced (Hiscott *et al.*, 2001). Envs of JSRV, MMTV, and MMLV (Burzyn *et al.*, 2004; Hofacre & Fan, 2010; Rassa *et al.*, 2002) as well as HIV-1 (Zhou *et al.*, 2018) activate this pathway through binding the TLR 4 receptor, which uniquely recruits both MyD88 and TRIF as signaling adapters (Fitzgerald *et al.*, 2003; Kagan *et al.*, 2008). This leads to a signaling cascade that activates the NF-κB transcription factor and other signaling molecules (Hofacre & Fan, 2010).

Thus, overexpression of the LETMD1 protein enhanced the infectivity of both control pseudoviruses JSRV Env and VSV-G in the NCI-H23 cell line but not HCT116 which does not express TLR4 protein on the cell surface (Davoodi *et al.*, 2012). Nevertheless, VSV-G induces a TLR4-dependent response pathway, which recruits an adapter other than MyD88 (Georgel *et al.*, 2007). LETMD1 expressing cells showed an enhanced HML-2 Env infection in HCT116

but not in NCI-H23, suggesting cell-dependent involvement of LETMD1 in the HML-2 Env infection pathway and emphasizing the different entry requirements for HML-2 Env in different cell types (Robinson & Whelan, 2016).

IL1RAP protein functions as a coreceptor for IL-1 receptor 1 (IL-1R1) and receptor-like 2 (IL1RL2) in the IL-1, IL-33 and IL-36 signaling pathways (Ali *et al.*, 2007; Dinarello, 2009; Huang *et al.*, 1997; Yamagata *et al.*, 2015). It is essential to form the IL-1R complex that mediates the activation of the NF-κB pathway and other pathways (Cullinan *et al.*, 1998; Huang *et al.*, 1997; Yamagata *et al.*, 2015). The membrane form of IL1RAP is implicated in activating and promoting NF-κB transcription factors (Liu *et al.*, 2018).
Expression of the transmembrane protein IL1RAP enhanced the infections of JSRV Env, VSV-G, and HML-2 Env pseudoviruses in the two tested cell lines (Figure 4-25) and (Figure 4-26). This suggests that the protein is involved in post-entry events.
IL1RAP is one of many proteins that are upregulated during HIV infection, especially in long-term non-progressor patients, according to one study (Salgado *et al.*, 2011). Furthermore, signaling through IL-1 induces HIV-1 expression (Granowitz *et al.*, 1995). IL1RAP might function similarly to LETMD1 in the post-entry events of the HML-2 infection pathway.

C12ORF59 is a novel protein whose biological function in the cell is not entirely understood. However, it is differentially expressed at the cellular membrane of distal convoluted tubule cells of the kidney (Park *et al.*, 2018). It has been shown that kidney resident macrophages implicated in renal artery stenosis repair express C12ORF59, unlike tissue-resident macrophages, which do not express this gene at their steady-state (Puranik *et al.*, 2018; Stamatiades *et al.*, 2016) At the subcellular level, it is localized to the focal adhesion sites (Thul *et al.*, 2017). Recently, reports revealed part of its role in the cell as a modulator to E-cadherin and EGFR activity, in which C12ORF59 suppression reduces E-cadherin stability and expression. On the other hand, it enhances the activation and internalization of EGFR (Lee *et al.*, 2021). E-cadherin modulates the entry of some viruses, including HBV (Hu *et al.*, 2020), HCV (Li *et al.*, 2016), and HSV-1 (Drees *et al.*, 2005). It contributes to the redistribution of the functional receptor of HBV and HCV, leading to an increase in viral entry (Hu *et al.*, 2020; Li *et al.*, 2016). In contrast, HIV infection is correlated with a reduction in E-cadherin (Lien *et al.*, 2019).

In our results, expression of C12ORF59 was associated with elevated levels of HML-2 Env entry in the HCT116 cell line (Figure 4-25) but not in NCI-H23 cells, where no change in HML-2 Env entry-level was observed. At the same time, JSRV Env and VSV-G pseudoviral entry levels were at the range of their controls. E-cadherin neither influences VSV-G (Li *et al.*, 2016)

nor JSRV Env entries (Martineau *et al.*, 2011). These results suggest an indirect role of the C12ORF59 in modulating the HML-2 Env entry through increasing E-cadherin. It also suggests that HCT116 cells but not NCI-H23 might express the functional receptor of HML-2 Env since E-cadherin contributes to the redistribution of the functional receptor in order to promote viral infections.

However, that does not preclude C12ORF59 of being involved directly in the HML-2 Env entry pathway as a receptor or coreceptor since its function in the cell is still not completely understood. To that, further investigation is recommended.

The last membrane protein that was observed to enhance the entry-level of HML-2 Env pseudovirus is MPEG1. It enhanced the viral entry in both tested non-permissive cell lines. However, no effect was observed in the entry of the control JSRV Env and VSV-G pseudoviruses, raising the possibility of being the putative receptor for HML-2 Env, especially since MPEG1 is a highly conserved protein from sponges to *Homo sapiens* (Bayly-Jones *et al.*, 2020; McCormack & Podack, 2015; Wiens *et al.*, 2005).

MPEG1 is the most ancient metazoan intracellular immune effector and defense protein belonging to the Membrane Attack Complex/Perforin MACPF/CDCs superfamily. It consists of immune pore-forming proteins called perforins (D'Angelo *et al.*, 2012; McCormack & Podack, 2015; Spilsbury *et al.*, 1995). Furthermore, it is believed that the Mammalian perforin immune proteins and the complement components (MAC) evolved from MPEG1 gene duplications (D'Angelo *et al.*, 2012).

It has been postulated that MPEG1 is the closest representative paralog to the metazoan common MACPF ancestor gene. Thus, its homologs can be found throughout the *Animalia* kingdom including *Porifera*, *Vertebrata*, and *Invertebrata* (D'Angelo *et al.*, 2012; McCormack & Podack, 2015; Spilsbury *et al.*, 1995; Wiens *et al.*, 2005). This might explain the broad species tropism of HML-2 Env (Kramer *et al.*, 2016; Robinson & Whelan, 2016). However, the THP-1 cell line prevented HML-2 Env entry, as reported by (Kramer *et al.*, 2016).

MPEG1 protein is commonly expressed in macrophages and phagocytes. However, it is also expressed by induction in cellular barriers such as keratinocytes, epithelial cells, and fibroblasts (McCormack *et al.*, 2015; Strbo *et al.*, 2019). Several signaling factors have been shown to drive MPEG1 expression independently, such as tumor necrosis factor-α and LPS. Therefore, MyD88 and NF-κB pathways play an indispensable role in its expression (Benard *et al.*, 2015; Wiens *et al.*, 2005). In this way, neither polarization of THP-1 macrophages into activated M2 macrophages using cytokines IL-10 and IL-4, nor stimulation of THP-1 cells with M-CSF, GM-

CSF, or PMA, have any effects on MPEG1 expression (Xiong *et al.*, 2017). This might be why the THP-1 cell line is non-permissive for HML-2 entry.

Different studies revealed the mechanism and functionality behind MPEG1 protein. However, some important mechanisms of pre-pores to pore-formation, in addition to triggering factors, are not fully understood (Bayly-Jones *et al.*, 2020; Law *et al.*, 2010). Well established is that MPEG1 localizes to the early endosomes and phagosomes/lysosomes (McCormack *et al.*, 2015; Xiong *et al.*, 2017), where it forms an oligomeric structure that resembles the Perforin and MAC proteins. It seems to function similarly to them as well, through forming membrane-spanning pores in the target's membrane (Leung *et al.*, 2017; Lukoyanova *et al.*, 2015). Its function and activity depend entirely on acidification, which can be maintained in the endosomes. The oligomerization of murine MPEG1 and pore-forming processes are shown to be triggered by low pH (Bayly-Jones *et al.*, 2020; Ni *et al.*, 2020; Pang *et al.*, 2019). However, gradual acidification is a specific feature for endosomes/ phagosomal vesicles in the presence of MPEG1 protein (Bayly-Jones *et al.*, 2020). Generally, Perforin proteins trigger endocytosis through clathrin- and dynamin-dependent pathways (Thiery *et al.*, 2010). The latter is the pathway used by HML-2 Env endocytosis (Robinson & Whelan, 2016).

Altogether, the short-listed proteins that enhanced HML-2 Env entry were involved in the entry process directly or indirectly. Most of them are expressed proteins in macrophages/lymphocytes and play an immune-modulatory role or activate the NF-κB pathway, raising the question of whether HML-2 Env initially infected the immune cells. Furthermore, the question of whether HML-2 Env utilizes multiple receptors in its entry pathway is still open.

These are preliminary results that require further investigation. Conducting a loss of function research for these proteins in the future is highly required.

6. References

Aaronson, S. A., Todaro, G. J., & Scolnick, E. M. (1971). Induction of murine C-type viruses from clonal lines of virus-free BALB/3T3 cells. *Science*, *174*(4005), 157-159.

Acton, O., Grant, T., Nicastro, G., Ball, N. J., Goldstone, D. C., Robertson, L. E., Sader, K., Nans, A., Ramos, A., Stoye, J. P., Taylor, I. A., & Rosenthal, P. B. (2019). Structural basis for Fullerene geometry in a human endogenous retrovirus capsid. *Nat Commun*, *10*(1), 5822.

Ahmed, Y. F., Hanly, S., Malim, M., Cullen, B., & Greene, W. (1990). Structure-function analyses of the HTLV-I Rex and HIV-1 Rev RNA response elements: insights into the mechanism of Rex and Rev action. *Genes & development*, *4*(6), 1014-1022.

Alexopoulou, A. N., Multhaupt, H. A., & Couchman, J. R. (2007). Syndecans in wound healing, inflammation and vascular biology. *The international journal of biochemistry & cell biology*, *39*(3), 505-528.

Alfahad, T., & Nath, A. (2013). Retroviruses and amyotrophic lateral sclerosis. *Antiviral research*, *99*(2), 180-187.

Ali, S., Huber, M., Kollewe, C., Bischoff, S. C., Falk, W., & Martin, M. U. (2007). IL-1 receptor accessory protein is essential for IL-33-induced activation of T lymphocytes and mast cells. *Proceedings of the National Academy of Sciences*, *104*(47), 18660-18665.

AlphaFold Protein Structure Database. EMBL-EBI. Retrieved 25.05.2022 from https://alphafold.ebi.ac.uk/entry/Q902F9

Ansermet, C., Moor, M. B., Centeno, G., Auberson, M., Hu, D. Z., Baron, R., Nikolaeva, S., Haenzi, B., Katanaeva, N., & Gautschi, I. (2017). Renal fanconi syndrome and hypophosphatemic rickets in the absence of xenotropic and polytropic retroviral receptor in the nephron. *Journal of the American Society of Nephrology*, *28*(4), 1073-1078.

Aquino, R. S., & Park, P. W. (2016). Glycosaminoglycans and infection. *Frontiers in bioscience (Landmark edition)*, *21*, 1260.

Armbruester, V., Sauter, M., Krautkraemer, E., Meese, E., Kleiman, A., Best, B., Roemer, K., & Mueller-Lantzsch, N. (2002). A novel gene from the human endogenous retrovirus K expressed in transformed cells. *Clinical Cancer Research*, *8*(6), 1800-1807.

Arosio, P., Barolo, G., Müller-Späth, T., Wu, H., & Morbidelli, M. (2011). Aggregation stability of a monoclonal antibody during downstream processing. *Pharmaceutical research*, *28*(8), 1884-1894.

Arriagada, C., Cavieres, V. A., Luchsinger, C., González, A. E., Muñoz, V. C., Cancino, J., Burgos, P. V., & Mardones, G. A. (2020). GOLPH3 regulates EGFR in T98G glioblastoma cells by modulating its glycosylation and ubiquitylation. *International Journal of Molecular Sciences*, *21*(22), 8880.

Arru, G., Mameli, G., Deiana, G., Rassu, A., Piredda, R., Sechi, E., Caggiu, E., Bo, M., Nako, E., & Urso, D. (2018). Humoral immunity response to human endogenous retroviruses K/W differentiates between amyotrophic lateral sclerosis and other neurological diseases. *European journal of neurology*, *25*(8), 1076-e1084.

Arthos, J., Cicala, C., Martinelli, E., Macleod, K., Van Ryk, D., Wei, D., Xiao, Z., Veenstra, T. D., Conrad, T. P., & Lempicki, R. A. (2008). HIV-1 envelope protein binds to and signals through integrin α4β7, the gut mucosal homing receptor for peripheral T cells. *Nature immunology*, *9*(3), 301-309.

Bachert, C., Fimmel, C., & Linstedt, A. D. (2007). Endosomal trafficking and proprotein convertase cleavage of cis Golgi protein GP73 produces marker for hepatocellular carcinoma. *Traffic*, *8*(10), 1415-1423.

Bacquin, A., Bireau, C., Tanguy, M., Romanet, C., Vernochet, C., Dupressoir, A., & Heidmann, T. (2017). A cell fusion-based screening method identifies glycosylphosphatidylinositol-anchored protein Ly6e as the receptor for mouse endogenous retroviral envelope syncytin-A. *Journal of virology*, *91*(18), e00832-00817.

Baenziger, J. U., Kumar, S., Brodbeck, R. M., Smith, P. L., & Beranek, M. C. (1992). Circulatory half-life but not interaction with the lutropin/chorionic gonadotropin receptor is modulated by sulfation of bovine lutropin oligosaccharides. *Proceedings of the National Academy of Sciences*, *89*(1), 334-338.

Baljinnyam, E., Umemura, M., De Lorenzo, M. S., Iwatsubo, M., Chen, S., Goydos, J. S., & Iwatsubo, K. (2011). Epac1 promotes melanoma metastasis via modification of heparan sulfate. *Pigment cell & melanoma research*, *24*(4), 680-687.

Bannert, N. (2017, Aug 27 – 31). Identification of a receptor binding site of the human endogenous retrovirus HERV-K(HML-2) [Oral presentation]. RETROPATH - International Workshop on Retroviral Pathogenesis, Chateau Liblice, Czech Republic, https://retropath.img.cas.cz/.

Bannert, N., Fiebig, U., & Hohn, O. (2010). Retroviral Particles, Proteins and Genomes. In R. Kurth & N. Bannert (Eds.), *Retroviruses: molecular biology, genomics and pathogenesis* (pp. 71-106). Caister Academic Press.

Bannert, N., & Kurth, R. (2004). Retroelements and the human genome: new perspectives on an old relation. *Proc Natl Acad Sci U S A*, *101 Suppl 2*, 14572-14579.

Bannert, N., & Kurth, R. (2006). The evolutionary dynamics of human endogenous retroviral families. *Annu Rev Genomics Hum Genet*, *7*, 149-173.

Bannert, N., Schenten, D., Craig, S., & Sodroski, J. (2000). The level of CD4 expression limits infection of primary rhesus monkey macrophages by a T-tropic simian immunodeficiency virus and macrophagetropic human immunodeficiency viruses. *Journal of virology*, *74*(23), 10984-10993.

Barabás, O., Rumlová, M., Erdei, A., Pongrácz, V., Pichová, I., & Vértessy, B. G. (2003). dUTPase and nucleocapsid polypeptides of the Mason-Pfizer monkey virus form a fusion protein in the virion with homotrimeric organization and low catalytic efficiency. *Journal of Biological Chemistry*, *278*(40), 38803-38812.

Bartlett, A. H., & Park, P. W. (2010). Proteoglycans in host–pathogen interactions: molecular mechanisms and therapeutic implications. *Expert reviews in molecular medicine*, *12*.

Battini, J.-L., Rasko, J. E., & Miller, A. D. (1999). A human cell-surface receptor for xenotropic and polytropic murine leukemia viruses: possible role in G protein-coupled signal transduction. *Proceedings of the National Academy of Sciences*, *96*(4), 1385-1390.

Bayly-Jones, C., Pang, S. S., Spicer, B. A., Whisstock, J. C., & Dunstone, M. A. (2020). Ancient but Not Forgotten: New Insights Into MPEG1, a Macrophage Perforin-Like Immune Effector. *Frontiers in Immunology*, *11*, 2690.

Behrendt, N. (2004). The urokinase receptor (uPAR) and the uPAR-associated protein (uPARAP/Endo180): membrane proteins engaged in matrix turnover during tissue remodeling.

Beimforde, N., Hanke, K., Ammar, I., Kurth, R., & Bannert, N. (2008). Molecular cloning and functional characterization of the human endogenous retrovirus K113. *Virology*, *371*(1), 216-225.

Belshaw, R., Dawson, A. L., Woolven-Allen, J., Redding, J., Burt, A., & Tristem, M. (2005). Genomewide screening reveals high levels of insertional polymorphism in the human endogenous retrovirus family HERV-K (HML2): implications for present-day activity. *Journal of virology*, *79*(19), 12507-12514.

Belshaw, R., Pereira, V., Katzourakis, A., Talbot, G., Pačes, J., Burt, A., & Tristem, M. (2004). Long-term reinfection of the human genome by endogenous retroviruses. *Proceedings of the National Academy of Sciences*, *101*(14), 4894-4899.

Benachenhou, F., Blikstad, V., & Blomberg, J. (2009). The phylogeny of orthoretroviral long terminal repeats (LTRs). *Gene*, *448*(2), 134-138.

Benard, E. L., Racz, P. I., Rougeot, J., Nezhinsky, A. E., Verbeek, F. J., Spaink, H. P., & Meijer, A. H. (2015). Macrophage-expressed perforins mpeg1 and mpeg1. 2 have an anti-bacterial function in zebrafish. *Journal of innate immunity*, *7*(2), 136-152.

Benicky, J., Sanda, M., Kennedy, Z. B., & Goldman, R. (2019). N-Glycosylation is required for secretion of the precursor to brain-derived neurotrophic factor (proBDNF) carrying sulfated LacdiNAc structures. *Journal of Biological Chemistry*, *294*(45), 16816-16830.

Bentvelzen, P., Daams, J., Hageman, P., & Calafat, J. (1970). Genetic transmission of viruses that incite mammary tumor in mice. In: National Acad Sciences.

Bertrand, P., Côté, M., Zheng, Y.-M., Albritton, L. M., & Liu, S.-L. (2008). Jaagsiekte sheep retrovirus utilizes a pH-dependent endocytosis pathway for entry. *Journal of virology*, *82*(5), 2555-2559.

Bhardwaj, N., Maldarelli, F., Mellors, J., & Coffin, J. M. (2014). HIV-1 infection leads to increased transcription of human endogenous retrovirus HERV-K (HML-2) proviruses in vivo but not to increased virion production. *Journal of virology*, *88*(19), 11108-11120.

Bishop, J. R., Schuksz, M., & Esko, J. D. (2007). Heparan sulphate proteoglycans fine-tune mammalian physiology. *Nature*, *446*(7139), 1030-1037.

Blagoveshchenskaya, A., Cheong, F. Y., Rohde, H. M., Glover, G., Knodler, A., Nicolson, T., Boehmelt, G., & Mayinger, P. (2008). Integration of Golgi trafficking and growth factor signaling by the lipid phosphatase SAC1. *The Journal of cell biology*, *180*(4), 803-812.

Blaise, S., de Parseval, N., Bénit, L., & Heidmann, T. (2003). Genomewide screening for fusogenic human endogenous retrovirus envelopes identifies syncytin 2, a gene conserved on primate evolution. *Proceedings of the National Academy of Sciences*, *100*(22), 13013-13018.

Bleul, C. C., Wu, L., Hoxie, J. A., Springer, T. A., & Mackay, C. R. (1997). The HIV coreceptors CXCR4 and CCR5 are differentially expressed and regulated on human T lymphocytes. *Proceedings of the National Academy of Sciences*, *94*(5), 1925-1930.

Blond, J.-L., Lavillette, D., Cheynet, V., Bouton, O., Oriol, G., Chapel-Fernandes, S., Mandrand, B., Mallet, F., & Cosset, F.-L. (2000). An envelope glycoprotein of the

human endogenous retrovirus HERV-W is expressed in the human placenta and fuses cells expressing the type D mammalian retrovirus receptor. *Journal of virology*, *74*(7), 3321-3329.

Bobardt, M. D., Saphire, A. C., Hung, H.-C., Yu, X., Van der Schueren, B., Zhang, Z., David, G., & Gallay, P. A. (2003). Syndecan captures, protects, and transmits HIV to T lymphocytes. *Immunity*, *18*(1), 27-39.

Boller, K., König, H., Sauter, M., Mueller-Lantzsch, N., Löwer, R., Löwer, J., & Kurth, R. (1993). Evidence that HERV-K is the endogenous retrovirus sequence that codes for the human teratocarcinoma-derived retrovirus HTDV. *Virology*, *196*(1), 349-353.

Boller, K., Schonfeld, K., Lischer, S., Fischer, N., Hoffmann, A., Kurth, R., & Tonjes, R. R. (2008). Human endogenous retrovirus HERV-K113 is capable of producing intact viral particles. *J Gen Virol*, *89*(Pt 2), 567-572.

Bovijn, C., Desmet, A.-S., Uyttendaele, I., Van Acker, T., Tavernier, J., & Peelman, F. (2013). Identification of binding sites for myeloid differentiation primary response gene 88 (MyD88) and Toll-like receptor 4 in MyD88 adapter-like (Mal). *Journal of Biological Chemistry*, *288*(17), 12054-12066.

Brady, T., Lee, Y. N., Ronen, K., Malani, N., Berry, C. C., Bieniasz, P. D., & Bushman, F. D. (2009). Integration target site selection by a resurrected human endogenous retrovirus. *Genes & development*, *23*(5), 633-642.

Brinzevich, D., Young, G. R., Sebra, R., Ayllon, J., Maio, S. M., Deikus, G., Chen, B. K., Fernandez-Sesma, A., Simon, V., & Mulder, L. C. (2014). HIV-1 interacts with human endogenous retrovirus K (HML-2) envelopes derived from human primary lymphocytes. *Journal of virology*, *88*(11), 6213-6223.

Bronson, D. L., Ritzi, D. M., Fraley, E. E., & Dalton, A. J. (1978). Morphologic evidence for retrovirus production by epithelial cells derived from a human testicular tumor metastasis: brief communication. *Journal of the National Cancer Institute*, *60*(6), 1305-1308.

Bukrinsky, M. I., Sharova, N., McDonald, T. L., Pushkarskaya, T., Tarpley, W. G., & Stevenson, M. (1993). Association of integrase, matrix, and reverse transcriptase antigens of human immunodeficiency virus type 1 with viral nucleic acids following acute infection. *Proceedings of the National Academy of Sciences*, *90*(13), 6125-6129.

Burnett, J. C., Miller-Jensen, K., Shah, P. S., Arkin, A. P., & Schaffer, D. V. (2009). Control of stochastic gene expression by host factors at the HIV promoter. *PLoS pathogens*, *5*(1), e1000260.

Burzyn, D., Rassa, J. C., Kim, D., Nepomnaschy, I., Ross, S. R., & Piazzon, I. (2004). Toll-like receptor 4-dependent activation of dendritic cells by a retrovirus. *Journal of virology*, *78*(2), 576-584.

Büscher, K., Trefzer, U., Hofmann, M., Sterry, W., Kurth, R., & Denner, J. (2005). Expression of human endogenous retrovirus K in melanomas and melanoma cell lines. *Cancer research*, *65*(10), 4172-4180.

Cagno, V., Tseligka, E. D., Jones, S. T., & Tapparel, C. (2019). Heparan Sulfate Proteoglycans and Viral Attachment: True Receptors or Adaptation Bias? *Viruses*, *11*(7).

Caldwell, E. E., Nadkarni, V. D., Fromm, J. R., Linhardt, R. J., & Weiler, J. M. (1996). Importance of specific amino acids in protein binding sites for heparin and heparan sulfate. *The international journal of biochemistry & cell biology*, *28*(2), 203-216.

Callahan, R., Drohan, W., Tronick, S., & Schlom, J. (1982). Detection and cloning of human DNA sequences related to the mouse mammary tumor virus genome. *Proceedings of the National Academy of Sciences*, *79*(18), 5503-5507.

Cameron, P. U., Freudenthal, P. S., Barker, J. M., Gezelter, S., Inaba, K., & Steinman, R. M. (1992). Dendritic cells exposed to human immunodeficiency virus type-1 transmit a vigorous cytopathic infection to CD4+ T cells. *Science*, *257*(5068), 383-387.

Capila, I., & Linhardt, R. J. (2002). Heparin–protein interactions. *Angewandte Chemie International Edition*, *41*(3), 390-412.

Cardin, A. D., & Weintraub, H. (1989). Molecular modeling of protein-glycosaminoglycan interactions. *Arteriosclerosis: An Official Journal of the American Heart Association, Inc.*, *9*(1), 21-32.

Carter, W. J., Cama, E., & Huntington, J. A. (2005). Crystal structure of thrombin bound to heparin. *Journal of Biological Chemistry*, *280*(4), 2745-2749.

Casale, J., & Crane, J. S. (2019). *Biochemistry, Glycosaminoglycans*. StatPearls Publishing LLC. .

Celma, C. C., Paladino, M. G., González, S. A., & Affranchino, J. L. (2007). Importance of the short cytoplasmic domain of the feline immunodeficiency virus transmembrane glycoprotein for fusion activity and envelope glycoprotein incorporation into virions. *Virology*, *366*(2), 405-414.

Chaitanya, K. (2019). Structure and organization of virus genomes. In *Genome and Genomics* (pp. 1-30). Springer.

Chand, S., Messina, E. L., AlSalmi, W., Ananthaswamy, N., Gao, G., Uritskiy, G., Padilla-Sanchez, V., Mahalingam, M., Peachman, K. K., & Robb, M. L. (2017). Glycosylation and oligomeric state of envelope protein might influence HIV-1 virion capture by α4β7 integrin. *Virology*, *508*, 199-212.

Chang, W.-L., Chang, C.-W., Chang, Y.-Y., Sung, H.-H., Lin, M.-D., Chang, S.-C., Chen, C.-H., Huang, C.-W., Tung, K.-S., & Chou, T.-B. (2013). The Drosophila GOLPH3 homolog regulates the biosynthesis of heparan sulfate proteoglycans by modulating the retrograde trafficking of exostosins. *Development*, *140*(13), 2798-2807.

Chappell, J. D., & Dermody, T. S. (2015). Biology of viruses and viral diseases. In J. E. Bennett, R. Dolin, & M. J. Blaser (Eds.), *Mandell, Douglas, and Bennett's Principles and Practice of Infectious Diseases* (8th ed., Vol. 2, pp. 1681-1693.). Elsevier.

Cheynet, V., Ruggieri, A., Oriol, G., Blond, J.-L., Boson, B., Vachot, L., Verrier, B., Cosset, F.-L., & Mallet, F. (2005). Synthesis, assembly, and processing of the Env ERVWE1/syncytin human endogenous retroviral envelope. *Journal of virology*, *79*(9), 5585-5593.

Cho, G.-W., Shin, S. M., Kim, H. K., Ha, S.-A., Kim, S., Yoon, J.-H., Hur, S. Y., Kim, T. E., & Kim, J. W. (2007). HCCR-1, a novel oncogene, encodes a mitochondrial outer membrane protein and suppresses the UVC-induced apoptosis. *BMC Cell Biology*, *8*(1), 1-12.

Clapham, P. R., & McKnight, Á. (2001). HIV-1 receptors and cell tropism. *British medical bulletin*, *58*(1), 43-59.

Clausen, T. M., Sandoval, D. R., Spliid, C. B., Pihl, J., Perrett, H. R., Painter, C. D., Narayanan, A., Majowicz, S. A., Kwong, E. M., & McVicar, R. N. (2020). SARS-CoV-2 infection depends on cellular heparan sulfate and ACE2. *Cell*, *183*(4), 1043-1057.

Coffin, J. M. (1992a). Genetic diversity and evolution of retroviruses. *Genetic diversity of RNA viruses*, 143-164.

Coffin, J. M. (1992b). Structure and classification of retroviruses. In *The retroviridae* (pp. 19-49). Springer.

Coffin, J. M. (1996). Retroviridae and their replication In Virology. In (ed. B.N. Fields et al. ed., pp. 1767-1848): Raven Press: New York.

Coffin, J. M., Hughes, S. H., & Varmus, H. E. (1997). *Retroviruses*. Cold Spring Harbor Laboratory Press.

Cohen, M., & Larsson, E. (1988). Human endogenous retroviruses. *Bioessays*, *9*(6), 191-196.

Coller, K. E., Heaton, N. S., Berger, K. L., Cooper, J. D., Saunders, J. L., & Randall, G. (2012). Molecular determinants and dynamics of hepatitis C virus secretion. *PLoS pathogens*, *8*(1), e1002466.

Contreras-Galindo, R., Almodóvar-Camacho, S., González-Ramírez, S., Lorenzo, E., & Yamamura, Y. (2007). Comparative longitudinal studies of HERV-K and HIV-1 RNA titers in HIV-1-infected patients receiving successful versus unsuccessful highly active antiretroviral therapy. *AIDS research and human retroviruses*, *23*(9), 1083-1086.

Contreras-Galindo, R., Contreras-Galindo, A., Lorenzo, E., & Yamamura, Y. (2006). Evidence for replication of human endogenous retroviruses type-K (HERV-K) in HIV-1 positive patients. *Retrovirology*, *3*(1), 1-1.

Contreras-Galindo, R., Kaplan, M. H., Contreras-Galindo, A. C., Gonzalez-Hernandez, M. J., Ferlenghi, I., Giusti, F., Lorenzo, E., Gitlin, S. D., Dosik, M. H., & Yamamura, Y. (2012). Characterization of human endogenous retroviral elements in the blood of HIV-1-infected individuals. *Journal of virology*, *86*(1), 262-276.

Contreras-Galindo, R., López, P., Vélez, R., & Yamamura, Y. (2007). HIV-1 infection increases the expression of human endogenous retroviruses type K (HERV-K) in vitro. *AIDS research and human retroviruses*, *23*(1), 116-122.

Cossart, P., & Helenius, A. (2014). Endocytosis of viruses and bacteria. *Cold Spring Harbor perspectives in biology*, *6*(8), a016972.

Côté, M., Zheng, Y.-M., Albritton, L. M., & Liu, S.-L. (2008). Fusogenicity of Jaagsiekte sheep retrovirus envelope protein is dependent on low pH and is enhanced by cytoplasmic tail truncations. *Journal of virology*, *82*(5), 2543-2554.

Côté, M., Zheng, Y.-M., Albritton, L. M., & Liu, S.-L. (2011). Single residues in the surface subunits of oncogenic sheep retrovirus envelopes distinguish receptor-mediated triggering for fusion at low pH and infection. *Virology*, *421*(2), 173-183.

Côté, M., Zheng, Y.-M., & Liu, S.-L. (2012). Membrane fusion and cell entry of XMRV are pH-independent and modulated by the envelope glycoprotein's cytoplasmic tail. *PLoS One*, *7*(3), e33734.

Crublet, E., Andrieu, J. P., Vives, R. R., & Lortat-Jacob, H. (2008). The HIV-1 envelope glycoprotein gp120 features four heparan sulfate binding domains, including the co-receptor binding site. *J Biol Chem*, *283*(22), 15193-15200.

Csibra, E., Brierley, I., & Irigoyen, N. (2014). Modulation of stop codon read-through efficiency and its effect on the replication of murine leukemia virus. *Journal of virology*, *88*(18), 10364-10376.

Cullinan, E. B., Kwee, L., Nunes, P., Shuster, D. J., Ju, G., McIntyre, K. W., Chizzonite, R. A., & Labow, M. A. (1998). IL-1 receptor accessory protein is an essential component of the IL-1 receptor. *The Journal of Immunology*, *161*(10), 5614-5620.

D'Angelo, M. E., Dunstone, M. A., Whisstock, J. C., Trapani, J. A., & Bird, P. I. (2012). Perforin evolved from a gene duplication of MPEG1, followed by a complex pattern of gene gain and loss within Euteleostomi. *BMC evolutionary biology*, *12*(1), 1-12.

Dar, R. D., Razooky, B. S., Singh, A., Trimeloni, T. V., McCollum, J. M., Cox, C. D., Simpson, M. L., & Weinberger, L. S. (2012). Transcriptional burst frequency and burst size are equally modulated across the human genome. *Proceedings of the National Academy of Sciences*, *109*(43), 17454-17459.

Davoodi, H., Hashemi, S. R., & Seow, H. F. (2012). Increased NFκ-B activity in HCT116 colorectal cancer cell line harboring TLR4 Asp299Gly variant. *Iranian Journal of Allergy, Asthma and Immunology*, *11*(2), 121-132.

De Parseval, A., Bobardt, M. D., Chatterji, A., Chatterji, U., Elder, J. H., David, G., Zolla-Pazner, S., Farzan, M., Lee, T.-H., & Gallay, P. A. (2005). A highly conserved arginine in gp120 governs HIV-1 binding to both syndecans and CCR5 via sulfated motifs. *Journal of Biological Chemistry*, *280*(47), 39493-39504.

De Parseval, N., & Heidmann, T. (2005). Human endogenous retroviruses: from infectious elements to human genes. *Cytogenetic and genome research*, *110*(1-4), 318-332.

De Vries, R. P., Zhu, X., McBride, R., Rigter, A., Hanson, A., Zhong, G., Hatta, M., Xu, R., Yu, W., & Kawaoka, Y. (2014). Hemagglutinin receptor specificity and structural analyses of respiratory droplet-transmissible H5N1 viruses. *Journal of virology*, *88*(1), 768-773.

De Witte, L., Bobardt, M., Chatterji, U., Degeest, G., David, G., Geijtenbeek, T. B., & Gallay, P. (2007). Syndecan-3 is a dendritic cell-specific attachment receptor for HIV-1. *Proceedings of the National Academy of Sciences*, *104*(49), 19464-19469.

Debaisieux, S., Rayne, F., Yezid, H., & Beaumelle, B. (2012). The ins and outs of HIV-1 Tat. *Traffic*, *13*(3), 355-363.

Deininger, P. L., & Batzer, M. A. (2002). Mammalian retroelements. *Genome research*, *12*(10), 1455-1465.

Delwart, E. L., Mosialos, G., & Gilmore, T. (1990). Retroviral envelope glycoproteins contain a "leucine zipper"-like repeat. *AIDS Res Hum Retroviruses*, *6*(6), 703-706.

Delwart, E. L., & Panganiban, A. T. (1990). N-linked glycosylation and reticuloendotheliosis retrovirus envelope glycoprotein function. *Virology*, *179*(2), 648-657.

Derse, D., Crise, B., Li, Y., Princler, G., Lum, N., Stewart, C., McGrath, C. F., Hughes, S. H., Munroe, D. J., & Wu, X. (2007). Human T-cell leukemia virus type 1 integration target sites in the human genome: comparison with those of other retroviruses. *Journal of virology*, *81*(12), 6731-6741.

Desfarges, S., & Ciuffi, A. (2012). Viral integration and consequences on host gene expression. In *Viruses: essential agents of life* (pp. 147-175). Springer.

Dewannieux, M., Blaise, S., & Heidmann, T. (2005). Identification of a functional envelope protein from the HERV-K family of human endogenous retroviruses. *J Virol*, *79*(24), 15573-15577.

Dewannieux, M., Harper, F., Richaud, A., Letzelter, C., Ribet, D., Pierron, G., & Heidmann, T. (2006). Identification of an infectious progenitor for the multiple-copy HERV-K human endogenous retroelements. *Genome Res*, *16*(12), 1548-1556.

Dinarello, C. A. (2009). Immunological and inflammatory functions of the interleukin-1 family. *Annual review of immunology*, *27*, 519-550.

Djerbal, L., Lortat-Jacob, H., & Kwok, J. (2017). Chondroitin sulfates and their binding molecules in the central nervous system. *Glycoconjugate journal*, *34*(3), 363-376.

Donizy, P., Kaczorowski, M., Biecek, P., Halon, A., Szkudlarek, T., & Matkowski, R. (2016). Golgi-related proteins GOLPH2 (GP73/GOLM1) and GOLPH3 (GOPP1/MIDAS) in cutaneous melanoma: patterns of expression and prognostic significance. *International Journal of Molecular Sciences*, *17*(10), 1619.

Douville, R., Liu, J., Rothstein, J., & Nath, A. (2011). Identification of active loci of a human endogenous retrovirus in neurons of patients with amyotrophic lateral sclerosis. *Annals of neurology*, *69*(1), 141-151.

Dragan, A., Casas-Finet, J., Bishop, E., Strouse, R., Schenerman, M., & Geddes, C. (2010). Characterization of PicoGreen interaction with dsDNA and the origin of its fluorescence enhancement upon binding. *Biophysical journal*, *99*(9), 3010-3019.

Drees, F., Pokutta, S., Yamada, S., Nelson, W. J., & Weis, W. I. (2005). α-catenin is a molecular switch that binds E-cadherin-β-catenin and regulates actin-filament assembly. *Cell*, *123*(5), 903-915.

Dunn, C. A., Romanish, M. T., Gutierrez, L. E., van de Lagemaat, L. N., & Mager, D. L. (2006). Transcription of two human genes from a bidirectional endogenous retrovirus promoter. *Gene*, *366*(2), 335-342.

Dupressoir, A., Marceau, G., Vernochet, C., Bénit, L., Kanellopoulos, C., Sapin, V., & Heidmann, T. (2005). Syncytin-A and syncytin-B, two fusogenic placenta-specific murine envelope genes of retroviral origin conserved in Muridae. *Proceedings of the National Academy of Sciences*, *102*(3), 725-730.

Dupressoir, A., Vernochet, C., Bawa, O., Harper, F., Pierron, G., Opolon, P., & Heidmann, T. (2009). Syncytin-A knockout mice demonstrate the critical role in placentation of a fusogenic, endogenous retrovirus-derived, envelope gene. *Proceedings of the National Academy of Sciences*, *106*(29), 12127-12132.

Dupressoir, A., Vernochet, C., Harper, F., Pierron, G., Guégan, J., Dessen, P., & Heidmann, T. (2011). Contribution of captured retroviral envelope genes, the" synctins" to the formation of the mouse placenta. *Retrovirology*, *8*(2), 1-2.

Earl, P. L., & Moss, B. (1993). Mutational analysis of the assembly domain of the HIV-1 envelope glycoprotein. *AIDS Res Hum Retroviruses*, *9*(7), 589-594.

Ejima, D., Tsumoto, K., Fukada, H., Yumioka, R., Nagase, K., Arakawa, T., & Philo, J. S. (2007). Effects of acid exposure on the conformation, stability, and aggregation of monoclonal antibodies. *Proteins*, *66*(4), 954-962.

Engelman, A. (2010). Reverse Transcription and Integration. In R. Kurth & N. Bannert (Eds.), *Retroviruses: molecular biology, genomics and pathogenesis* (pp. 129-159). Caister Academic Press.

Esko, J. D. (1999). Proteoglycans and Glycosaminoglycans. In A. Varki, R. D. Cummings, J. D. Esko, H. Freeze, G. W. Hart, & J. Marth (Eds.), *Essentials of Glycobiology* (3rd ed.). Cold Spring Harbor Laboratory Press.

Esko, J. D., & Selleck, S. B. (2002). Order out of chaos: assembly of ligand binding sites in heparan sulfate. *Annual review of biochemistry*, *71*(1), 435-471.

Esnault, C., Priet, S., Ribet, D., Vernochet, C., Bruls, T., Lavialle, C., Weissenbach, J., & Heidmann, T. (2008). A placenta-specific receptor for the fusogenic, endogenous retrovirus-derived, human syncytin-2. *Proceedings of the National Academy of Sciences*, *105*(45), 17532-17537.

Esqueda, D., Xu, F., Moore, Y., Yang, Z., Huang, G., Lennon, P. A., Hu, P. C., & Dong, J. (2013). Lack of correlation between HERV-K expression and HIV-1 viral load in plasma specimens. *Annals of Clinical & Laboratory Science*, *43*(2), 122-125.

Faschinger, A., Rouault, F., Sollner, J., Lukas, A., Salmons, B., Günzburg, W. H., & Indik, S. (2008). Mouse mammary tumor virus integration site selection in human and mouse genomes. *Journal of virology*, *82*(3), 1360-1367.

Feldman, S. A., Hendry, R. M., & Beeler, J. A. (1999). Identification of a linear heparin binding domain for human respiratory syncytial virus attachment glycoprotein G. *Journal of virology*, *73*(8), 6610-6617.

Feng, Y., Broder, C. C., Kennedy, P. E., & Berger, E. A. (1996). HIV-1 entry cofactor: functional cDNA cloning of a seven-transmembrane, G protein-coupled receptor. *Science*, *272*(5263), 872-877.

Fenner, F., Bachmann, P. A., Gibbs, E. P. J., Murphy, F. A., STUDDERT, M. J., & WHITE, D. O. (1987). Structure and Composition of Viruses. In F. Fenner, P. A. Bachmann, E. P. J. Gibbs, F. A. Murphy, M. J. STUDDERT, & D. O. WHITE (Eds.), *Veterinary Virology* (pp. 3-19). Academic Press.

Feuchter, A., & Mager, D. (1990). Functional heterogeneity of a large family of human LTR-like promoters and enhancers. *Nucleic acids research*, *18*(5), 1261-1270.

Finkelshtein, D., Werman, A., Novick, D., Barak, S., & Rubinstein, M. (2013). LDL receptor and its family members serve as the cellular receptors for vesicular stomatitis virus. *Proceedings of the National Academy of Sciences*, *110*(18), 7306-7311.

Fitzgerald, K. A., Rowe, D. C., Barnes, B. J., Caffrey, D. R., Visintin, A., Latz, E., Monks, B., Pitha, P. M., & Golenbock, D. T. (2003). LPS-TLR4 signaling to IRF-3/7 and NF-κB involves the toll adapters TRAM and TRIF. *The Journal of experimental medicine*, *198*(7), 1043-1055.

Flint, S. J., Racaniello, V. R., Rall, G. F., & Skalka, A. M. (2015). *Principles of Virology: Molecular Biology* (Vol. 1). John Wiley & Sons.

Frank, S., Kammerer, R. A., Mechling, D., Schulthess, T., Landwehr, R., Bann, J., Guo, Y., Lustig, A., Bachinger, H. P., & Engel, J. (2001). Stabilization of short collagen-like triple helices by protein engineering. *J Mol Biol*, *308*(5), 1081-1089.

Franklin, G. C., Chretien, S., Hanson, I. M., Rochefort, H., May, F., & Westley, B. (1988). Expression of human sequences related to those of mouse mammary tumor virus. *Journal of virology*, *62*(4), 1203-1210.

Frevert, C. W., & Wight, T. N. (2006). EXTRACELLULAR MATRIX | Matrix Proteoglycans. In G. J. Laurent & S. D. Shapiro (Eds.), *Encyclopedia of Respiratory Medicine* (pp. 184-188). Academic Press.

Fritzsche, F. R., Riener, M.-O., Dietel, M., Moch, H., Jung, K., & Kristiansen, G. (2008). GOLPH2 expression in renal cell cancer. *BMC urology*, *8*(1), 1-6.

Fromm, J., Hileman, R., Caldwell, E., Weiler, J., & Linhardt, R. (1995). Differences in the interaction of heparin with arginine and lysine and the importance of these basic amino acids in the binding of heparin to acidic fibroblast growth factor. *Archives of biochemistry and biophysics*, *323*(2), 279-287.

Fromm, J., Hileman, R., Caldwell, E., Weiler, J., & Linhardt, R. (1997). Pattern and spacing of basic amino acids in heparin binding sites. *Archives of biochemistry and biophysics*, *343*(1), 92-100.

Fuchs, N. V., Kraft, M., Tondera, C., Hanschmann, K.-M., Löwer, J., & Löwer, R. (2011). Expression of the human endogenous retrovirus (HERV) group HML-2/HERV-K does not depend on canonical promoter elements but is regulated by transcription factors Sp1 and Sp3. *Journal of virology*, *85*(7), 3436-3448.

Fuchs, N. V., Loewer, S., Daley, G. Q., Izsvák, Z., Löwer, J., & Löwer, R. (2013). Human endogenous retrovirus K (HML-2) RNA and protein expression is a marker for human embryonic and induced pluripotent stem cells. *Retrovirology*, *10*(1), 1-6.

Gallaher, W. R., Ball, J. M., Garry, R. F., Griffin, M. C., & MONTELARO, R. C. (1989). A general model for the transmembrane proteins of HIV and other retroviruses. *AIDS research and human retroviruses*, *5*(4), 431-440.

Gallay, P. (2004). Syndecans and HIV-1 pathogenesis. *Microbes Infect*, *6*(6), 617-622.

Garcia-Montojo, M., Doucet-O'Hare, T., Henderson, L., & Nath, A. (2018). Human endogenous retrovirus-K (HML-2): a comprehensive review. *Crit Rev Microbiol*, *44*(6), 715-738.

Garcia-Montojo, M., Fathi, S., Norato, G., Smith, B., Rowe, D., Kiernan, M., Vucic, S., Mathers, S., van Eijk, R., & Santamaria, U. (2021). Inhibition of HERV-K (HML-2) in amyotrophic lateral sclerosis patients on antiretroviral therapy. *Journal of the neurological sciences*, *423*, 117358.

Garcia-Montojo, M., Li, W., & Nath, A. (2019). Technical considerations in detection of HERV-K in amyotrophic lateral sclerosis: Selection of controls and the perils of qPCR. *Acta Neuropathologica Communications*, *7*(1), 1-3.

Garrison, K. E., Jones, R. B., Meiklejohn, D. A., Anwar, N., Ndhlovu, L. C., Chapman, J. M., Erickson, A. L., Agrawal, A., Spotts, G., & Hecht, F. M. (2007). T cell responses to human endogenous retroviruses in HIV-1 infection. *PLoS pathogens*, *3*(11), e165.

GeneCards (RRID:SCR_002773) the human gene database. Retrieved 01.12.2021 from http://genecards.org

George, M., Schwecke, T., Beimforde, N., Hohn, O., Chudak, C., Zimmermann, A., Kurth, R., Naumann, D., & Bannert, N. (2011). Identification of the protease cleavage sites in a

reconstituted Gag polyprotein of an HERV-K (HML-2) element. *Retrovirology*, *8*(1), 1-15.

Georgel, P., Jiang, Z., Kunz, S., Janssen, E., Mols, J., Hoebe, K., Bahram, S., Oldstone, M. B., & Beutler, B. (2007). Vesicular stomatitis virus glycoprotein G activates a specific antiviral Toll-like receptor 4-dependent pathway. *Virology*, *362*(2), 304-313.

Geppert, P. (2019). *The specific Cell Entry Pathway of the Human Endogenous Retrovirus-K (HML-2) and the Role of Heparan Sulfate Proteoglycans* [Master's thesis, Brandenburgische Technische Universität]. Robert Koch-Institut, edoc server http://edoc.rki.de/176904/5899.

Geraerts, M., Willems, S., Baekelandt, V., Debyser, Z., & Gijsbers, R. (2006). Comparison of lentiviral vector titration methods. *BMC biotechnology*, *6*(1), 1-10.

Gerhard, D. S., Wagner, L., Feingold, E. A., Shenmen, C. M., Grouse, L. H., Schuler, G., Klein, S. L., Old, S., Rasooly, R., & Good, P. (2004). The status, quality, and expansion of the NIH full-length cDNA project: the Mammalian Gene Collection (MGC). *Genome research*, *14*(10 B), 2121-2127.

Ghez, D., Lepelletier, Y., Lambert, S., Fourneau, J. M., Blot, V., Janvier, S., Arnulf, B., van Endert, P. M., Heveker, N., Pique, C., & Hermine, O. (2006). Neuropilin-1 is involved in human T-cell lymphotropic virus type 1 entry. *J Virol*, *80*(14), 6844-6854.

Gifford, R., Kabat, P., Martin, J., Lynch, C., & Tristem, M. (2005). Evolution and distribution of class II-related endogenous retroviruses. *Journal of virology*, *79*(10), 6478-6486.

Gifford, R., & Tristem, M. (2003). The evolution, distribution and diversity of endogenous retroviruses. *Virus genes*, *26*(3), 291-315.

Gifford, R. J., Katzourakis, A., Tristem, M., Pybus, O. G., Winters, M., & Shafer, R. W. (2008). A transitional endogenous lentivirus from the genome of a basal primate and implications for lentivirus evolution. *Proceedings of the National Academy of Sciences*, *105*(51), 20362-20367.

Gómez Toledo, A., Sorrentino, J. T., Sandoval, D. R., Malmström, J., Lewis, N. E., & Esko, J. D. (2021). A systems view of the heparan sulfate interactome. *Journal of Histochemistry & Cytochemistry*, *69*(2), 105-119.

Gong, Y., Long, Q., Xie, H., Zhang, T., & Peng, T. (2012). Cloning and characterization of human Golgi phosphoprotein 2 gene (GOLPH2/GP73/GOLM1) promoter. *Biochem Biophys Res Commun*, *421*(4), 713-720.

Gotoh, M., Sato, T., Kiyohara, K., Kameyama, A., Kikuchi, N., Kwon, Y.-D., Ishizuka, Y., Iwai, T., Nakanishi, H., & Narimatsu, H. (2004). Molecular cloning and characterization of β1, 4-N-acetylgalactosaminyltransferases IV synthesizing N, N′-diacetyllactosediamine1. *FEBS letters*, *562*(1-3), 134-140.

Göttlinger, H. G., & Weissenhorn, W. (2010). Assembly and Release. In R. Kurth & N. Bannert (Eds.), *Retroviruses: molecular biology, genomics and pathogenesis* (pp. 187-215). Caister Academic Press.

Graff, S., Moore, D. H., Stanley, W. M., Randall, H. T., & Haagensen, C. (1949). Isolation of mouse mammary carcinoma virus. *Cancer*, *2*(5), 755-762.

Granowitz, E. V., Saget, B. M., Wang, M. Z., Dinarello, C. A., Skolnik, P. R., & Fauci, A. (1995). Interleukin 1 induces HIV-1 expression in chronically infected U1 cells:

blockade by interleukin 1 receptor antagonist and tumor necrosis factor binding protein type 1. *Molecular medicine*, *1*(6), 667-677.

Gray, L. R., Jackson, R. E., Jackson, P. E. H., Bekiranov, S., Rekosh, D., & Hammarskjold, M. L. (2019). HIV-1 Rev interacts with HERV-K RcREs present in the human genome and promotes export of unspliced HERV-K proviral RNA. *Retrovirology*, *16*(1), 40.

Greenfeder, S. A., Nunes, P., Kwee, L., Labow, M., Chizzonite, R. A., & Ju, G. (1995). Molecular Cloning and Characterization of a Second Subunit of the Interleukin 1 Receptor Complex∗. *Journal of Biological Chemistry*, *270*(23), 13757-13765.

Grove, J., & Marsh, M. (2011). The cell biology of receptor-mediated virus entry. *Journal of Cell Biology*, *195*(7), 1071-1082.

Gumley, T. P., McKenzie, I. F., & Sandrin, M. S. (1995). Tissue expression, structure and function of the murine Ly-6 family of molecules. *Immunology and cell biology*, *73*(4), 277-296.

Guo, Y., Yang, C., Liu, Y., Li, T., Li, H., Han, J., Jia, L., Wang, X., Zhang, B., & Li, J. (2022). High Expression of HERV-K (HML-2) Might Stimulate Interferon in COVID-19 Patients. *Viruses*, *14*(5), 996.

Guthe, S., Kapinos, L., Moglich, A., Meier, S., Grzesiek, S., & Kiefhaber, T. (2004). Very fast folding and association of a trimerization domain from bacteriophage T4 fibritin. *J Mol Biol*, *337*(4), 905-915.

Hanke, K., Kramer, P., Seeher, S., Beimforde, N., Kurth, R., & Bannert, N. (2009). Reconstitution of the ancestral glycoprotein of human endogenous retrovirus k and modulation of its functional activity by truncation of the cytoplasmic domain. *J Virol*, *83*(24), 12790-12800.

Hassan, Z., Kumar, N. D., Reggiori, F., & Khan, G. (2021). How Viruses Hijack and Modify the Secretory Transport Pathway. *Cells*, *10*(10), 2535.

Hayward, A. (2017). Origin of the retroviruses: when, where, and how? *Current opinion in virology*, *25*, 23-27.

Hayward, A., Cornwallis, C. K., & Jern, P. (2015). Pan-vertebrate comparative genomics unmasks retrovirus macroevolution. *Proceedings of the National Academy of Sciences*, *112*(2), 464-469.

Heise, M. (2014). Viral pathogenesis. *Reference Module in Biomedical Sciences*.

Helland, D., Welles, J. L., Caputo, A., & Haseltine, W. A. (1991). Transcellular transactivation by the human immunodeficiency virus type 1 tat protein. *Journal of virology*, *65*(8), 4547-4549.

Hennessy, E. J., Parker, A. E., & O'neill, L. A. (2010). Targeting Toll-like receptors: emerging therapeutics? *Nature reviews Drug discovery*, *9*(4), 293-307.

Henzy, J. E., & Coffin, J. M. (2013). Betaretroviral envelope subunits are noncovalently associated and restricted to the mammalian class. *J Virol*, *87*(4), 1937-1946.

Herniou, E., Martin, J., Miller, K., Cook, J., Wilkinson, M., & Tristem, M. (1998). Retroviral diversity and distribution in vertebrates. *Journal of virology*, *72*(7), 5955-5966.

Herve, C., Lugli, E., Brand, A., Griffiths, D., & Venables, P. (2002). Autoantibodies to human endogenous retrovirus-K are frequently detected in health and disease and react with multiple epitopes. *Clinical & Experimental Immunology*, *128*(1), 75-82.

Hileman, R. E., Fromm, J. R., Weiler, J. M., & Linhardt, R. J. (1998). Glycosaminoglycan-protein interactions: definition of consensus sites in glycosaminoglycan binding proteins. *Bioessays*, *20*(2), 156-167.

Hirano, K., & Furukawa, K. (2022). Biosynthesis and Biological Significances of LacdiNAc Group on N-and O-Glycans in Human Cancer Cells. *Biomolecules*, *12*(2), 195.

Hiscott, J., Kwon, H., & Génin, P. (2001). Hostile takeovers: viral appropriation of the NF-kB pathway. *The Journal of clinical investigation*, *107*(2), 143-151.

Hizi, A., & Herzig, E. (2015). dUTPase: the frequently overlooked enzyme encoded by many retroviruses. *Retrovirology*, *12*(1), 1-15.

Ho, D. D., Rota, T. R., & Hirsch, M. S. (1986). Infection of monocyte/macrophages by human T lymphotropic virus type III. *The Journal of clinical investigation*, *77*(5), 1712-1715.

Hofacre, A., & Fan, H. (2010). Jaagsiekte sheep retrovirus biology and oncogenesis. *Viruses*, *2*(12), 2618-2648.

Hohn, O., Hanke, K., & Bannert, N. (2013). HERV-K(HML-2), the Best Preserved Family of HERVs: Endogenization, Expression, and Implications in Health and Disease. *Front Oncol*, *3*, 246.

Hohn, O., Hanke, K., Lausch, V., Zimmermann, A., Mostafa, S., & Bannert, N. (2014). CMV-promoter driven codon-optimized expression alters the assembly type and morphology of a reconstituted HERV-K(HML-2). *Viruses*, *6*(11), 4332-4345.

Hong, S., Klein, E. A., Das Gupta, J., Hanke, K., Weight, C. J., Nguyen, C., Gaughan, C., Kim, K. A., Bannert, N., Kirchhoff, F., Munch, J., & Silverman, R. H. (2009). Fibrils of prostatic acid phosphatase fragments boost infections with XMRV (xenotropic murine leukemia virus-related virus), a human retrovirus associated with prostate cancer. *J Virol*, *83*(14), 6995-7003.

Horie, M., Honda, T., Suzuki, Y., Kobayashi, Y., Daito, T., Oshida, T., Ikuta, K., Jern, P., Gojobori, T., & Coffin, J. M. (2010). Endogenous non-retroviral RNA virus elements in mammalian genomes. *Nature*, *463*(7277), 84-87.

Hu, L., Li, L., Xie, H., Gu, Y., & Peng, T. (2011). The Golgi localization of GOLPH2 (GP73/GOLM1) is determined by the transmembrane and cytoplamic sequences. *PLoS One*, *6*(11), e28207.

Hu, Q., Zhang, F., Duan, L., Wang, B., Ye, Y., Li, P., Li, D., Yang, S., Zhou, L., & Chen, W. (2020). E-cadherin plays a role in hepatitis B virus entry through affecting glycosylated sodium-taurocholate cotransporting polypeptide distribution. *Frontiers in Cellular and Infection Microbiology*, *10*, 74.

Huang, J., Gao, X., Li, S., & Cao, Z. (1997). Recruitment of IRAK to the interleukin 1 receptor complex requires interleukin 1 receptor accessory protein. *Proceedings of the National Academy of Sciences*, *94*(24), 12829-12832.

Hughes, M., Natale, B., Simmons, D., & Natale, D. (2013). Ly6e expression is restricted to syncytiotrophoblast cells of the mouse placenta. *Placenta*, *34*(9), 831-835.

Huigen, M., Kamp, W., & Nottet, H. (2004). Multiple effects of HIV-1 trans-activator protein on the pathogenesis of HIV-1 infection. *European journal of clinical investigation*, *34*(1), 57-66.

The Human Protein Atlas. (2022). Retrieved 04.06.2022 from https://www.proteinatlas.org/humanproteome/subcellular/plasma+membrane

Hunter, E. (1997). Viral Entry and Receptors. In: Cold Spring Harbor Laboratory Press, Cold Spring Harbor (NY).

Hunter, E., & Swanstrom, R. (1990). Retrovirus envelope glycoproteins. *Curr Top Microbiol Immunol*, *157*, 187-253.

ICTV. (2021). *Virus Taxonomy: 2021 Release*. Retrieved 04.05.2022 from https://talk.ictvonline.org/taxonomy/

Ihrcke, N. S., Wrenshall, L. E., Lindman, B. J., & Platt, J. L. (1993). Role of heparan sulfate in immune system-blood vessel interactions. *Immunology today*, *14*(10), 500-505.

Jassal, S. R., Pöhler, R. G., & Brighty, D. W. (2001). Human T-cell leukemia virus type 1 receptor expression among syncytium-resistant cell lines revealed by a novel surface glycoprotein-immunoadhesin. *Journal of virology*, *75*(17), 8317-8328.

Jensen, M. M., Arvaniti, M., Mikkelsen, J. D., Michalski, D., Pinborg, L. H., Hartig, W., & Thomsen, M. S. (2015). Prostate stem cell antigen interacts with nicotinic acetylcholine receptors and is affected in Alzheimer's disease. *Neurobiol Aging*, *36*(4), 1629-1638.

Jern, P., & Coffin, J. M. (2008). Effects of retroviruses on host genome function. *Annual review of genetics*, *42*, 709-732.

Jern, P., Sperber, G. O., & Blomberg, J. (2005). Use of endogenous retroviral sequences (ERVs) and structural markers for retroviral phylogenetic inference and taxonomy. *Retrovirology*, *2*(1), 1-12.

Jha, A. R., Pillai, S. K., York, V. A., Sharp, E. R., Storm, E. C., Wachter, D. J., Martin, J. N., Deeks, S. G., Rosenberg, M. G., & Nixon, D. F. (2009). Cross-sectional dating of novel haplotypes of HERV-K 113 and HERV-K 115 indicate these proviruses originated in Africa before Homo sapiens. *Molecular biology and evolution*, *26*(11), 2617-2626.

Jia, N., Barclay, W. S., Roberts, K., Yen, H.-L., Chan, R. W., Lam, A. K., Air, G., Peiris, J. M., Dell, A., & Nicholls, J. M. (2014). Glycomic characterization of respiratory tract tissues of ferrets: implications for its use in influenza virus infection studies. *Journal of Biological Chemistry*, *289*(41), 28489-28504.

Jinno, A., & Park, P. W. (2015). Role of glycosaminoglycans in infectious disease. In *Glycosaminoglycans* (pp. 567-585). Springer.

Johnson, K. M., Kines, R. C., Roberts, J. N., Lowy, D. R., Schiller, J. T., & Day, P. M. (2009). Role of heparan sulfate in attachment to and infection of the murine female genital tract by human papillomavirus. *Journal of virology*, *83*(5), 2067-2074.

Johnson, W. E. (2015). Endogenous retroviruses in the genomics era. *Annual review of virology*, *2*, 135-159.

Jones, K. S., Nath, M., Petrow-Sadowski, C., Baines, A. C., Dambach, M., Huang, Y., & Ruscetti, F. W. (2002). Similar regulation of cell surface human T-cell leukemia virus type 1 (HTLV-1) surface binding proteins in cells highly and poorly transduced by HTLV-1-pseudotyped virions. *Journal of virology*, *76*(24), 12723-12734.

Jones, K. S., Petrow-Sadowski, C., Bertolette, D. C., Huang, Y., & Ruscetti, F. W. (2005). Heparan sulfate proteoglycans mediate attachment and entry of human T-cell leukemia virus type 1 virions into CD4+ T cells. *Journal of virology*, *79*(20), 12692-12702.

Jones, R. B., Garrison, K. E., Mujib, S., Mihajlovic, V., Aidarus, N., Hunter, D. V., Martin, E., John, V. M., Zhan, W., & Faruk, N. F. (2012). HERV-K–specific T cells eliminate

diverse HIV-1/2 and SIV primary isolates. *The Journal of clinical investigation*, *122*(12).

Joshi, V., Shivach, T., Kumar, V., Yadav, N., & Rathore, A. (2014). Avoiding antibody aggregation during processing: establishing hold times. *Biotechnology Journal*, *9*(9), 1195-1205.

Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., & Potapenko, A. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature*, 1-11.

Jung, Y.-S., Jun, S., Kim, M. J., Lee, S. H., Suh, H. N., Lien, E. M., Jung, H.-Y., Lee, S., Zhang, J., & Yang, J.-I. (2018). TMEM9 promotes intestinal tumorigenesis through vacuolar-ATPase-activated Wnt/β-catenin signalling. *Nature cell biology*, *20*(12), 1421-1433.

Kagan, J. C., Su, T., Horng, T., Chow, A., Akira, S., & Medzhitov, R. (2008). TRAM couples endocytosis of Toll-like receptor 4 to the induction of interferon-β. *Nature immunology*, *9*(4), 361-368.

Kagiampakis, I., Gharibi, A., Mankowski, M. K., Snyder, B. A., Ptak, R. G., Alatas, K., & LiWang, P. J. (2011). Potent strategy to inhibit HIV-1 by binding both gp120 and gp41. *Antimicrobial agents and chemotherapy*, *55*(1), 264-275.

Kameoka, J., Tanaka, T., Nojima, Y., Schlossman, S. F., & Morimoto, C. (1993). Direct association of adenosine deaminase with a T cell activation antigen, CD26. *Science*, *261*(5120), 466-469.

Kämmerer, U., Germeyer, A., Stengel, S., Kapp, M., & Denner, J. (2011). Human endogenous retrovirus K (HERV-K) is expressed in villous and extravillous cytotrophoblast cells of the human placenta. *Journal of reproductive immunology*, *91*(1-2), 1-8.

Karamitros, T., Paraskevis, D., Hatzakis, A., Psichogiou, M., Elefsiniotis, I., Hurst, T., Geretti, A.-M., Beloukas, A., Frater, J., & Klenerman, P. (2016). A contaminant-free assessment of Endogenous Retroviral RNA in human plasma. *Scientific reports*, *6*(1), 1-12.

Katzourakis, A., Tristem, M., Pybus, O. G., & Gifford, R. J. (2007). Discovery and analysis of the first endogenous lentivirus. *Proc Natl Acad Sci U S A*, *104*(15), 6261-6265.

Kawai, T., & Akira, S. (2010). The role of pattern-recognition receptors in innate immunity: update on Toll-like receptors. *Nature immunology*, *11*(5), 373-384.

Kawar, Z. S., Haslam, S. M., Morris, H. R., Dell, A., & Cummings, R. D. (2005). Novel poly-GalNAcβ1–4GlcNAc (LacdiNAc) and fucosylated poly-LacdiNAc N-glycans from mammalian cells expressing β1, 4-N-acetylgalactosaminyltransferase and α1, 3-fucosyltransferase. *Journal of Biological Chemistry*, *280*(13), 12810-12819.

Kayman, S. C., Kopelman, R., Projan, S., Kinney, D. M., & Pinter, A. (1991). Mutational analysis of N-linked glycosylation sites of Friend murine leukemia virus envelope protein. *Journal of virology*, *65*(10), 5323-5332.

Kenny, D. T., Skoog, E. C., Lindén, S. K., Struwe, W. B., Rudd, P. M., & Karlsson, N. G. (2012). Presence of terminal N-acetylgalactosamineβ1-4 N-acetylglucosamine residues on O-linked oligosaccharides from gastric MUC5AC: involvement in helicobacter pylori colonization? *Glycobiology*, *22*(8), 1077-1085.

Kiernan, M. C., Vucic, S., Cheah, B. C., Turner, M. R., Eisen, A., Hardiman, O., Burrell, J. R., & Zoing, M. C. (2011). Amyotrophic lateral sclerosis. *The lancet*, *377*(9769), 942-955.

Kitamura, Y., Ayukawa, T., Ishikawa, T., Kanda, T., & Yoshiike, K. (1996). Human endogenous retrovirus K10 encodes a functional integrase. *Journal of virology*, *70*(5), 3302-3306.

Kjellén, L., & Lindahl, U. (1991). Proteoglycans: structures and interactions. *Annual review of biochemistry*, *60*(1), 443-475.

Kladney, R. D., Bulla, G. A., Guo, L., Mason, A. L., Tollefson, A. E., Simon, D. J., Koutoubi, Z., & Fimmel, C. J. (2000). GP73, a novel Golgi-localized protein upregulated by viral infection. *Gene*, *249*(1-2), 53-65.

Klasse, P. J., Depetris, R. S., Pejchal, R., Julien, J.-P., Khayat, R., Lee, J. H., Marozsan, A. J., Cupo, A., Cocco, N., & Korzun, J. (2013). Influences on trimerization and aggregation of soluble, cleaved HIV-1 SOSIP envelope glycoprotein. *Journal of virology*, *87*(17), 9873-9885.

Konstantoulas, C. J., & Indik, S. (2014). Mouse mammary tumor virus-based vector transduces non-dividing cells, enters the nucleus via a TNPO3-independent pathway and integrates in a less biased fashion than other retroviruses. *Retrovirology*, *11*(1), 1-15.

Köppe, B., Menendez-Arias, L., & Oroszlan, S. (1994). Expression and purification of the mouse mammary tumor virus gag-pro transframe protein p30 and characterization of its dUTPase activity. *Journal of virology*, *68*(4), 2313-2319.

Kramer, P., Lausch, V., Volkwein, A., Hanke, K., Hohn, O., & Bannert, N. (2016). The human endogenous retrovirus K(HML-2) has a broad envelope-mediated cellular tropism and is prone to inhibition at a post-entry, pre-integration step. *Virology*, *487*, 121-128.

Krilleke, D., DeErkenez, A., Schubert, W., Giri, I., Robinson, G. S., Ng, Y.-S., & Shima, D. T. (2007). Molecular mapping and functional characterization of the VEGF164 heparin-binding domain. *Journal of Biological Chemistry*, *282*(38), 28045-28056.

Krzysztalowska-Wawrzyniak, M., Ostanek, M., Clark, J., Binczak-Kuleta, A., Ostanek, L., Kaczmarczyk, M., Loniewska, B., Wyrwicz, L. S., Brzosko, M., & Ciechanowicz, A. (2011). The distribution of human endogenous retrovirus K-113 in health and autoimmune diseases in Poland. *Rheumatology*, *50*(7), 1310-1314.

Kurth, R., & Bannert, N. (2010). Beneficial and detrimental effects of human endogenous retroviruses. *International journal of cancer*, *126*(2), 306-314.

Kveine, M., Tenstad, E., Døsen, G., Funderud, S., & Rian, E. (2002). Characterization of the novel human transmembrane protein 9 (TMEM9) that localizes to lysosomes and late endosomes. *Biochemical and biophysical research communications*, *297*(4), 912-917.

Lambert, S., Bouttier, M., Vassy, R., Seigneuret, M., Petrow-Sadowski, C., Janvier, S., Heveker, N., Ruscetti, F. W., Perret, G., Jones, K. S., & Pique, C. (2009). HTLV-1 uses HSPG and neuropilin-1 for entry by molecular mimicry of VEGF165. *Blood*, *113*(21), 5176-5185.

Larsson, E., Kato, N., & Cohen, M. (1989). Human endogenous proviruses. *Oncogenes and Retroviruses*, 115-132.

Lau, E. K., Paavola, C. D., Johnson, Z., Gaudry, J.-P., Geretti, E., Borlat, F., Kungl, A. J., Proudfoot, A. E., & Handel, T. M. (2004). Identification of the glycosaminoglycan binding site of the CC chemokine, MCP-1: implications for structure and function in vivo. *Journal of Biological Chemistry*, *279*(21), 22294-22305.

Lauffer, L., Kanzy, E. J., Kohler, R., Kurrle, R., Enssle, K., & Seiler, F. R. (1995). Monomeric and dimeric forms of soluble receptors can differ in their neutralization potential. *Behring Inst Mitt*(96), 21-31.

Law, R. H., Lukoyanova, N., Voskoboinik, I., Caradoc-Davies, T. T., Baran, K., Dunstone, M. A., D'Angelo, M. E., Orlova, E. V., Coulibaly, F., & Verschoor, S. (2010). The structural basis for membrane binding and pore formation by lymphocyte perforin. *Nature*, *468*(7322), 447-451.

Lee, Y., Ko, D., Yoon, J., Lee, Y., & Kim, S. (2021). TMEM52B suppression promotes cancer cell survival and invasion through modulating E-cadherin stability and EGFR activity. *Journal of Experimental & Clinical Cancer Research*, *40*(1), 1-20.

Lee, Y. N., & Bieniasz, P. D. (2007). Reconstitution of an infectious human endogenous retrovirus. *PLoS Pathog*, *3*(1), e10.

Lefkowitz, E. J., Dempsey, D. M., Hendrickson, R. C., Orton, R. J., Siddell, S. G., & Smith, D. B. (2018). Virus taxonomy: the database of the International Committee on Taxonomy of Viruses (ICTV). *Nucleic acids research*, *46*(D1), D708-D717.

Leistner, C. M., Gruen-Bernhard, S., & Glebe, D. (2008). Role of glycosaminoglycans for binding and infection of hepatitis B virus. *Cellular microbiology*, *10*(1), 122-133.

Leung, C., Hodel, A. W., Brennan, A. J., Lukoyanova, N., Tran, S., House, C. M., Kondos, S. C., Whisstock, J. C., Dunstone, M. A., & Trapani, J. A. (2017). Real-time visualization of perforin nanopore assembly. *Nature nanotechnology*, *12*(5), 467-473.

Lewis, P. F., & Emerman, M. (1994). Passage through mitosis is required for oncoretroviruses but not for the human immunodeficiency virus. *Journal of virology*, *68*(1), 510-516.

Li, E., Liu, L., Li, F., Luo, L., Zhao, S., Wang, J., Kang, R., Luo, J., & Zhao, Z. (2017). PSCA promotes prostate cancer proliferation and cell-cycle progression by up-regulating c-Myc. *The Prostate*, *77*(16), 1563-1572.

Li, L., Wen, L., Gong, Y., Mei, G., Liu, J., Chen, Y., & Peng, T. (2012). Xenopus as a model system for the study of GOLPH2/GP73 function: Xenopus GOLPH2 is required for pronephros development. *PLoS One*, *7*(6), e38939.

Li, Q.-X., Camerini, D., Xie, Y., Greenwald, M., Kuritzkes, D. R., & Chen, I. S. (1996). Syncytium formation by recombinant HTLV-II envelope glycoprotein. *Virology*, *218*(1), 279-284.

Li, Q., Sodroski, C., Lowey, B., Schweitzer, C. J., Cha, H., Zhang, F., & Liang, T. J. (2016). Hepatitis C virus depends on E-cadherin as an entry factor and regulates its expression in epithelial-to-mesenchymal transition. *Proceedings of the National Academy of Sciences*, *113*(27), 7620-7625.

Li, W., Lee, M.-H., Henderson, L., Tyagi, R., Bachani, M., Steiner, J., Campanac, E., Hoffman, D. A., Von Geldern, G., & Johnson, K. (2015). Human endogenous retrovirus-K contributes to motor neuron disease. *Science translational medicine*, *7*(307), 307ra153-307ra153.

Li, Y., Liu, D., Wang, Y., Su, W., Liu, G., & Dong, W. (2021). The importance of glycans of viral and host proteins in enveloped virus infection. *Frontiers in Immunology*, *12*, 1544.

Li, Y., Luo, L., Rasool, N., & Kang, C. Y. (1993). Glycosylation is necessary for the correct folding of human immunodeficiency virus gp120 in CD4 binding. *Journal of virology*, *67*(1), 584-588.

Lien, K., Mayer, W., Herrera, R., Rosbe, K., & Tugizov, S. M. (2019). HIV-1 proteins gp120 and tat induce the epithelial–mesenchymal transition in oral and genital mucosal epithelial cells. *PLoS One*, *14*(12), e0226343.

Lim, S.-G., Suk, K., & Lee, W.-H. (2020). Letmd1 regulates phagocytosis and inflammatory responses to lipopolysaccharide via reactive oxygen species generation and nf-κb activation in macrophages. *The Journal of Immunology*, *204*(5), 1299-1309.

Lindahl, U., Kusche-Gullberg, M., & Kjellén, L. (1998). Regulated diversity of heparan sulfate. *Journal of Biological Chemistry*, *273*(39), 24979-24982.

Liu, C., Wen, C., Wang, X., Wei, Y., Xu, C., Mu, X., Zhang, L., Wang, X., Tian, J., & Ma, P. (2019). Golgi membrane protein GP73 modified-liposome mediates the antitumor effect of survivin promoter-driven HSVtk in hepatocellular carcinoma. *Experimental cell research*, *383*(1), 111496.

Liu, F., Dai, M., Xu, Q., Zhu, X., Zhou, Y., Jiang, S., Wang, Y., Ai, Z., Ma, L., & Zhang, Y. (2018). SRSF10-mediated IL1RAP alternative splicing regulates cervical cancer oncogenesis via mIL1RAP-NF-κB-CD47 axis. *Oncogene*, *37*(18), 2394-2409.

Loeuillet, C., Deutsch, S., Ciuffi, A., Robyr, D., Taffé, P., Muñoz, M., Beckmann, J. S., Antonarakis, S. E., & Telenti, A. (2008). In vitro whole-genome analysis identifies a susceptibility locus for HIV-1. *PLoS biology*, *6*(2), e32.

Löwer, R., Löwer, J., Frank, H., Harzmann, R., & Kurth, R. (1984). Human teratocarcinomas cultured in vitro produce unique retrovirus-like viruses. *Journal of General Virology*, *65*(5), 887-898.

Löwer, R., Löwer, J., & Kurth, R. (1996). The viruses in all of us: characteristics and biological significance of human endogenous retrovirus sequences. *Proceedings of the National Academy of Sciences*, *93*(11), 5177-5184.

Löwer, R., Tönjes, R., Korbmacher, C., Kurth, R., & Löwer, J. (1995). Identification of a Rev-related protein by analysis of spliced transcripts of the human endogenous retroviruses HTDV/HERV-K. *Journal of virology*, *69*(1), 141-149.

Lu, M., Blacklow, S. C., & Kim, P. S. (1995). A trimeric structural domain of the HIV-1 transmembrane glycoprotein. *Nature structural biology*, *2*(12), 1075-1082.

Lukoyanova, N., Kondos, S. C., Farabella, I., Law, R. H., Reboul, C. F., Caradoc-Davies, T. T., Spicer, B. A., Kleifeld, O., Traore, D. A., & Ekkel, S. M. (2015). Conformational changes during pore formation by the perforin-related protein pleurotolysin. *PLoS biology*, *13*(2), e1002049.

Macfarlane, C., & Simmonds, P. (2004). Allelic variation of HERV-K (HML-2) endogenous retroviral elements in human populations. *Journal of molecular evolution*, *59*(5), 642-656.

Mager, D. L., & Medstrand, P. (2003). Retroviral repeat sequences. *Nature encyclopedia of the human genome*, *5*, 57-63.

Maginnis, M. S. (2018). Virus-Receptor Interactions: The Key to Cellular Invasion. *J Mol Biol*, *430*(17), 2590-2611.

Malassine, A., Handschuh, K., Tsatsaris, V., Gerbaud, P., Cheynet, V., Oriol, G., Mallet, F., & Evain-Brion, D. (2005). Expression of HERV-W Env glycoprotein (syncytin) in the extravillous trophoblast of first trimester human placenta. *Placenta*, *26*(7), 556-562.

Mammano, F., Salvatori, F., Indraccolo, S., De Rossi, A., Chieco-Bianchi, L., & Göttlinger, H. (1997). Truncation of the human immunodeficiency virus type 1 envelope glycoprotein allows efficient pseudotyping of Moloney murine leukemia virus particles and gene transfer into CD4+ cells. *Journal of virology*, *71*(4), 3341-3345.

Manel, N., Kim, F. J., Kinet, S., Taylor, N., Sitbon, M., & Battini, J.-L. (2003). The ubiquitous glucose transporter GLUT-1 is a receptor for HTLV. *Cell*, *115*(4), 449-459.

Manel, N., Kinet, S., Battini, J.-L., Kim, F. J., Taylor, N., & Sitbon, M. (2003). The HTLV receptor is an early T-cell activation marker whose expression requires de novo protein synthesis. *Blood, The Journal of the American Society of Hematology*, *101*(5), 1913-1918.

Manzella, S. M., Hooper, L. V., & Baenziger, J. U. (1996). Oligosaccharides Containing β1, 4-Linked N-Acetylgalactosamine, a Paradigm for Protein-specific Glycosylation (∗). *Journal of Biological Chemistry*, *271*(21), 12117-12120.

Mar, K. B., Rinkenberger, N. R., Boys, I. N., Eitson, J. L., McDougal, M. B., Richardson, R. B., & Schoggins, J. W. (2018). LY6E mediates an evolutionarily conserved enhancement of virus infection by targeting a late entry step. *Nature communications*, *9*(1), 1-14.

Mårdberg, K., Trybala, E., Glorioso, J. C., & Bergström, T. (2001). Mutational analysis of the major heparan sulfate-binding domain of herpes simplex virus type 1 glycoprotein C. *Journal of General Virology*, *82*(8), 1941-1950.

Margalit, H., Fischer, N., & Ben-Sasson, S. A. (1993). Comparative analysis of structurally defined heparin binding sequences reveals a distinct spatial distribution of basic residues. *Journal of Biological Chemistry*, *268*(26), 19228-19231.

Mariani-Costantini, R., Horn, T., & Callahan, R. (1989). Ancestry of a human endogenous retrovirus family. *Journal of virology*, *63*(11), 4982-4985.

Marsh, M., & Helenius, A. (2006). Virus entry: open sesame. *Cell*, *124*(4), 729-740.

Martin, M. A., Bryan, T., Rasheed, S., & Khan, A. S. (1981). Identification and cloning of endogenous retroviral sequences present in human DNA. *Proceedings of the National Academy of Sciences*, *78*(8), 4892-4896.

Martineau, H. M., Cousens, C., Imlach, S., Dagleish, M. P., & Griffiths, D. J. (2011). Jaagsiekte sheep retrovirus infects multiple cell types in the ovine lung. *Journal of virology*, *85*(7), 3341-3355.

Masrori, P., & Van Damme, P. (2020). Amyotrophic lateral sclerosis: a clinical review. *European journal of neurology*, *27*(10), 1918-1929.

Matos, R., Amorim, I., Magalhães, A., Haesebrouck, F., Gärtner, F., & Reis, C. A. (2021). Adhesion of helicobacter species to the human gastric mucosa: a deep look into glycans role. *Frontiers in molecular biosciences*, *8*, 386.

Mattoscio, D., Segré, C. V., & Chiocca, S. (2013). Viral manipulation of cellular protein conjugation pathways: The SUMO lesson. *World journal of virology*, *2*(2), 79.

Mayer, J., & Meese, E. U. (2003). Presence of dUTPase in the various human endogenous retrovirus K (HERV-K) families. *Journal of molecular evolution*, *57*(6), 642-649.

McCormack, R., & Podack, E. R. (2015). Perforin-2/Mpeg1 and other pore-forming proteins throughout evolution. *Journal of leukocyte biology*, *98*(5), 761-768.

McCormack, R. M., de Armas, L. R., Shiratsuchi, M., Fiorentino, D. G., Olsson, M. L., Lichtenheld, M. G., Morales, A., Lyapichev, K., Gonzalez, L. E., & Strbo, N. (2015). Perforin-2 is essential for intracellular defense of parenchymal cells and phagocytes against pathogenic bacteria. *Elife*, *4*, e06508.

McDougal, J., Kennedy, M., Sligh, J., Cort, S., Mawle, A., & Nicholson, J. (1986). Binding of HTLV-III/LAV to T4+ T cells by a complex of the 110K viral protein and the T4 molecule. *Science*, *231*(4736), 382-385.

McKenzie, I., Gardiner, J., Cherry, M., & Snell, G. (1977). Lymphocyte antigens: Ly-4, Ly-6, and Ly-7.

Medstrand, P., & Blomberg, J. (1993). Characterization of novel reverse transcriptase encoding human endogenous retroviral sequences similar to type A and type B retroviruses: differential transcription in normal human tissues. *Journal of virology*, *67*(11), 6778-6787.

Medstrand, P., Van De Lagemaat, L. N., & Mager, D. L. (2002). Retroelement distributions in the human genome: variations associated with age and proximity to genes. *Genome research*, *12*(10), 1483-1495.

Mei, M., Ye, J., Qin, A., Wang, L., Hu, X., Qian, K., & Shao, H. (2015). Identification of novel viral receptors with cell line expressing viral receptor-binding protein. *Sci Rep*, *5*, 7935.

Mi, Y., Shapiro, S. D., & Baenziger, J. U. (2002). Regulation of lutropin circulatory half-life by the mannose/N-acetylgalactosamine-4-SO 4 receptor is critical for implantation in vivo. *The Journal of clinical investigation*, *109*(2), 269-276.

Michaud, H.-A., de Mulder, M., SenGupta, D., Deeks, S. G., Martin, J. N., Pilcher, C. D., Hecht, F. M., Sacha, J. B., & Nixon, D. F. (2014). Trans-activation, post-transcriptional maturation, and induction of antibodies to HERV-K (HML-2) envelope transmembrane protein in HIV-1 infection. *Retrovirology*, *11*(1), 1-15.

Mitchell, K., Barreyro, L., Todorova, T. I., Taylor, S. J., Antony-Debré, I., Narayanagari, S.-R., Carvajal, L. A., Leite, J., Piperdi, Z., & Pendurti, G. (2018). IL1RAP potentiates multiple oncogenic signaling pathways in AML. *Journal of Experimental Medicine*, *215*(6), 1709-1727.

Mitchell, R. S., Beitzel, B. F., Schroder, A. R. W., Shinn, P., Chen, H., Berry, C. C., Ecker, J. R., Bushman, F. D., & Emerman, M. (2004). Retroviral DNA integration: ASLV, HIV, and MLV show distinct target site preferences. *PLoS biology*, *2*(8), e234.

Miwa, J. M., Anderson, K. R., & Hoffman, K. M. (2019). Lynx prototoxins: Roles of endogenous mammalian neurotoxin-like proteins in modulating nicotinic acetylcholine receptor function to influence complex biological processes. *Frontiers in pharmacology*, *10*, 343.

Moore, J. P., McKeating, J. A., Norton, W. A., & Sattentau, Q. J. (1991). Direct measurement of soluble CD4 binding to human immunodeficiency virus type 1 virions: gp120 dissociation and its implications for virus-cell binding and fusion reactions and their neutralization by soluble CD4. *J Virol*, *65*(3), 1133-1140.

Morciano, G., Giorgi, C., Balestra, D., Marchi, S., Perrone, D., Pinotti, M., & Pinton, P. (2016). Mcl-1 involvement in mitochondrial dynamics is associated with apoptotic cell death. *Molecular biology of the cell*, *27*(1), 20-34.

Morozov, V. A., Dao Thi, V. L., & Denner, J. (2013). The transmembrane protein of the human endogenous retrovirus--K (HERV-K) modulates cytokine release and gene expression. *PLoS One*, *8*(8), e70399.

Morozov, V. A., & Morozov, A. V. (2021). A Comprehensive Analysis of Human Endogenous Retroviruses HERV-K (HML. 2) from Teratocarcinoma Cell Lines and Detection of Viral Cargo in Microvesicles. *International Journal of Molecular Sciences*, *22*(22), 12398.

Mothes, W., & Uchil, P. D. (2010). Retroviral Entry and Uncoating. In R. Kurth & N. Bannert (Eds.), *Retroviruses: molecular biology, genomics and pathogenesis* (pp. 107-128). Caister Academic Press.

Mukherjee, S., Sengupta, N., Chaudhuri, A., Akbar, I., Singh, N., Chakraborty, S., Suryawanshi, A. R., Bhattacharyya, A., & Basu, A. (2018). PLVAP and GKN3 are two critical host cell receptors which facilitate Japanese encephalitis virus entry into neurons. *Scientific reports*, *8*(1), 1-16.

Munoz, E. M., & Linhardt, R. J. (2004). Heparin-binding domains in vascular biology. *Arterioscler Thromb Vasc Biol*, *24*(9), 1549-1557.

Munro, S. (1998). Localization of proteins to the Golgi apparatus. *Trends in cell biology*, *8*(1), 11-15.

Muradrasoli, S., Forsman, A., Hu, L., Blikstad, V., & Blomberg, J. (2006). Development of real-time PCRs for detection and quantitation of human MMTV-like (HML) sequences: HML expression in human tissues. *Journal of virological methods*, *136*(1-2), 83-92.

Murakami, T., & Freed, E. O. (2000). The long cytoplasmic tail of gp41 is required in a cell type-dependent manner for HIV-1 envelope glycoprotein incorporation into virions. *Proceedings of the National Academy of Sciences*, *97*(1), 343-348.

Muranyi, W., Malkusch, S., Müller, B., Heilemann, M., & Kräusslich, H.-G. (2013). Super-resolution microscopy reveals specific recruitment of HIV-1 envelope proteins to viral assembly sites dependent on the envelope C-terminal tail. *PLoS pathogens*, *9*(2), e1003198.

Murphy, F. A., Fauquet, C. M., Bishop, D. H., Ghabrial, S. A., Jarvis, A. W., Martelli, G. P., Mayo, M. A., & Summers, M. D. (2012). *Virus taxonomy: classification and nomenclature of viruses* (Vol. 10). Springer Science & Business Media.

Muster, T., Waltenberger, A., Grassauer, A., Hirschl, S., Caucig, P., Romirer, I., Födinger, D., Seppele, H., Schanab, O., & Magin-Lachmann, C. (2003). An endogenous retrovirus derived from human melanoma cells. *Cancer research*, *63*(24), 8735-8741.

Ni, T., Jiao, F., Yu, X., Aden, S., Ginger, L., Williams, S. I., Bai, F., Pražák, V., Karia, D., & Stansfeld, P. (2020). Structure and mechanism of bactericidal mammalian perforin-2, an ancient agent of innate immunity. *Science advances*, *6*(5), eaax8286.

Nugent, M. A., Zaia, J., & Spencer, J. L. (2013). Heparan sulfate-protein binding specificity. *Biochemistry (Moscow)*, *78*(7), 726-735.

O'Neill, L. A., & Bowie, A. G. (2007). The family of five: TIR-domain-containing adaptors in Toll-like receptor signalling. *Nature Reviews Immunology*, *7*(5), 353-364.

Oja, M., Peltonen, J., Blomberg, J., & Kaski, S. (2007). Methods for estimating human endogenous retrovirus activities from EST databases. *BMC Bioinformatics*, *8*(2), 1-10.

Okuma, K., Dalton, K. P., Buonocore, L., Ramsburg, E., & Rose, J. K. (2003). Development of a novel surrogate virus for human T-cell leukemia virus type 1: inhibition of infection by osteoprotegerin. *Journal of virology*, *77*(15), 8562-8569.

Okuma, K., Nakamura, M., Nakano, S., Niho, Y., & Matsuura, Y. (1999). Host range of human T-cell leukemia virus type I analyzed by a cell fusion-dependent reporter gene activation assay. *Virology*, *254*(2), 235-244.

Olson, S. K., & Esko, J. D. (2013). Proteoglycans,. In W. J. Lennarz & M. Daniel Lane (Eds.), *Encyclopedia of Biological Chemistry (Second Edition)* (pp. 654-660). Academic Press.

Ono, A. (2010). HIV-1 assembly at the plasma membrane. *Vaccine*, *28*, B55-B59.

Ono, H., Hiraoka, N., Lee, Y. S., Woo, S. M., Lee, W. J., Choi, I. J., Saito, A., Yanagihara, K., Kanai, Y., & Ohnami, S. (2012). Prostate stem cell antigen, a presumable organ-dependent tumor suppressor gene, is down-regulated in gallbladder carcinogenesis. *Genes, Chromosomes and Cancer*, *51*(1), 30-41.

Ono, H., Sakamoto, H., Yoshida, T., & Saeki, N. (2018). Prostate stem cell antigen is expressed in normal and malignant human brain tissues. *Oncology letters*, *15*(3), 3081-3084.

Ori, A., Wilkinson, M. C., & Fernig, D. G. (2011). A systems biology approach for the investigation of the heparin/heparan sulfate interactome. *Journal of Biological Chemistry*, *286*(22), 19892-19904.

Overbaugh, J., Miller, A. D., & Eiden, M. V. (2001). Receptors and entry cofactors for retroviruses include single and multiple transmembrane-spanning proteins as well as newly described glycophosphatidylinositol-anchored and secreted proteins. *Microbiology and Molecular Biology Reviews*, *65*(3), 371-389.

Pang, S. S., Bayly-Jones, C., Radjainia, M., Spicer, B. A., Law, R. H., Hodel, A. W., Parsons, E. S., Ekkel, S. M., Conroy, P. J., & Ramm, G. (2019). The cryo-EM structure of the acid activatable pore-forming immune effector Macrophage-expressed gene 1. *Nature communications*, *10*(1), 1-9.

Paprotka, T., Delviks-Frankenberry, K. A., Cingöz, O., Martinez, A., Kung, H.-J., Tepper, C. G., Hu, W.-S., Fivash Jr, M. J., Coffin, J. M., & Pathak, V. K. (2011). Recombinant origin of the retrovirus XMRV. *Science*, *333*(6038), 97-101.

Paraskevopoulou, V., Schimpl, M., Overman, R. C., Stolnik, S., Chen, Y., Nguyen, L., Winkler, G. S., Gellert, P., Klassen, J. S., & Falcone, F. H. (2021). Structural and binding characterization of the LacdiNAc-specific adhesin (LabA; HopD) exodomain from Helicobacter pylori. *Current Research in Structural Biology*, *3*, 19-29.

Park, J., Shrestha, R., Qiu, C., Kondo, A., Huang, S., Werth, M., Li, M., Barasch, J., & Suszták, K. (2018). Single-cell transcriptomics of the mouse kidney reveals potential cellular targets of kidney disease. *Science*, *360*(6390), 758-763.

Patel, M., Yanagishita, M., Roderiquez, G., Bou-Habib, D. C., Oravecz, T., Hascall, V. C., & Norcross, M. (1993). Cell-surface heparan sulfate proteoglycan mediates HIV-1 infection of T-cell lines. *AIDS research and human retroviruses*, *9*(2), 167-174.

Pearson, R. G. (1963). Hard and soft acids and bases. *Journal of the American Chemical society*, *85*(22), 3533-3539.

Pecheur, E., Sainte-Marie, J., Bienvenüe, A., & Hoekstra, D. (1999). Peptides and membrane fusion: towards an understanding of the molecular mechanism of protein-induced fusion. *The Journal of membrane biology*, *167*(1), 1-17.

Pérot, P., Mugnier, N., Montgiraud, C., Gimenez, J., Jaillard, M., Bonnaud, B., & Mallet, F. (2012). Microarray-based sketches of the HERV transcriptome landscape. *PLoS One*, *7*(6), e40194.

Peters, G., & Glover, C. (1980). tRNA's and priming of RNA-directed DNA synthesis in mouse mammary tumor virus. *Journal of virology*, *35*(1), 31-40.

Pinon, J. D., Klasse, P., Jassal, S. R., Welson, S., Weber, J., Brighty, D. W., & Sattentau, Q. J. (2003). Human T-cell leukemia virus type 1 envelope glycoprotein gp46 interacts with cell surface heparan sulfate proteoglycans. *Journal of virology*, *77*(18), 9922-9930.

Pinter, A., Honnen, W. J., & Li, J. S. (1984). Studies with inhibitors of oligosaccharide processing indicate a functional role for complex sugars in the transport and proteolysis of Friend mink cell focus-inducing murine leukemia virus envelope proteins. *Virology*, *136*(1), 196-210.

Plaas, A. H., Wong-Palms, S., Roughley, P. J., Midura, R. J., & Hascall, V. C. (1997). Chemical and immunological assay of the nonreducing terminal residues of chondroitin sulfate from human aggrecan. *Journal of Biological Chemistry*, *272*(33), 20603-20610.

Ploubidou, A., & Way, M. (2001). Viral transport and the cytoskeleton. *Current opinion in cell biology*, *13*(1), 97-105.

Pontén, F., Jirström, K., & Uhlen, M. (2008). The Human Protein Atlas—a tool for pathology. *The Journal of Pathology: A Journal of the Pathological Society of Great Britain and Ireland*, *216*(4), 387-393.

Priesnitz, M. (2019). *Funktionelle Analyse der Putativen Rezeptorbindungsregion des Hüllproteins von HERV-K (HML-2)* [Bachelor's thesis, Beuth-Hochschule für Technik Berlin]. Robert-Koch Institute, edoc server http://edoc.rki.de/176904/5921.

Puranik, A. S., Leaf, I. A., Jensen, M. A., Hedayat, A. F., Saad, A., Kim, K.-W., Saadalla, A. M., Woollard, J. R., Kashyap, S., & Textor, S. C. (2018). Kidney-resident macrophages promote a proangiogenic environment in the normal and chronically ischemic mouse kidney. *Scientific reports*, *8*(1), 1-15.

Puri, S., Bachert, C., Fimmel, C. J., & Linstedt, A. D. (2002). Cycling of early Golgi proteins via the cell surface and endosomes upon lumenal pH disruption. *Traffic*, *3*(9), 641-653.

Pyra, H., Böni, J., & Schüpbach, J. (1994). Ultrasensitive retrovirus detection by a reverse transcriptase assay based on product enhancement. *Proceedings of the National Academy of Sciences*, *91*(4), 1544-1548.

Qi, M., Williams, J. A., Chu, H., Chen, X., Wang, J.-J., Ding, L., Akhirome, E., Wen, X., Lapierre, L. A., & Goldenring, J. R. (2013). Rab11-FIP1C and Rab14 direct plasma membrane sorting and particle incorporation of the HIV-1 envelope glycoprotein complex. *PLoS pathogens*, *9*(4), e1003278.

Rabson, A., & Graves, B. (2011). Synthesis and processing of viral RNA.

Radek, K. A., Taylor, K. R., & Gallo, R. L. (2009). FGF-10 and specific structural elements of dermatan sulfate size and sulfation promote maximal keratinocyte migration and cellular proliferation. *Wound Repair and Regeneration*, *17*(1), 118-126.

Rai, S. K., DeMartini, J. C., & Miller, A. D. (2000). Retrovirus vectors bearing jaagsiekte sheep retrovirus Env transduce human cells by using a new receptor localized to chromosome 3p21. 3. *Journal of virology*, *74*(10), 4698-4704.

Raj, V. S., Mou, H., Smits, S. L., Dekkers, D. H., Muller, M. A., Dijkman, R., Muth, D., Demmers, J. A., Zaki, A., Fouchier, R. A., Thiel, V., Drosten, C., Rottier, P. J., Osterhaus, A. D., Bosch, B. J., & Haagmans, B. L. (2013). Dipeptidyl peptidase 4 is a functional receptor for the emerging human coronavirus-EMC. *Nature*, *495*(7440), 251-254.

Raman, R., Sasisekharan, V., & Sasisekharan, R. (2005). Structural insights into biological roles of protein-glycosaminoglycan interactions. *Chemistry & biology*, *12*(3), 267-277.

Ramdas, P., Sahu, A. K., Mishra, T., Bhardwaj, V., & Chande, A. (2020). From entry to egress: Strategic exploitation of the cellular processes by HIV-1. *Frontiers in Microbiology*, 3021.

Rassa, J. C., Meyers, J. L., Zhang, Y., Kudaravalli, R., & Ross, S. R. (2002). Murine retroviruses activate B cells via interaction with toll-like receptor 4. *Proceedings of the National Academy of Sciences*, *99*(4), 2281-2286.

Redmond, S., Peters, G., & Dickson, C. (1984). Mouse mammary tumor virus can mediate cell fusion at reduced pH. *Virology*, *133*(2), 393-402.

Reinhold, W. C., Sunshine, M., Liu, H., Varma, S., Kohn, K. W., Morris, J., Doroshow, J., & Pommier, Y. (2012). CellMiner: a web-based suite of genomic and pharmacologic tools to explore transcript and drug patterns in the NCI-60 cell line set. *Cancer Res*, *72*(14), 3499-3511.

Reiter, R. E., Gu, Z., Watabe, T., Thomas, G., Szigeti, K., Davis, E., Wahl, M., Nisitani, S., Yamashiro, J., & Le Beau, M. M. (1998). Prostate stem cell antigen: a cell surface marker overexpressed in prostate cancer. *Proceedings of the National Academy of Sciences*, *95*(4), 1735-1740.

Reza Etemadi, M., Ling, K.-H., Zainal Abidin, S., Chee, H.-Y., & Sekawi, Z. (2017). Gene expression patterns induced at different stages of rhinovirus infection in human alveolar epithelial cells. *PLoS One*, *12*(5), e0176947.

Risco, C., & Fernández de Castro, I. (2013). Virus morphogenesis in the cell: Methods and observations. *Structure and Physics of Viruses*, 417-440.

Robey, W. G., Nara, P. L., Poore, C. M., Popovic, M., McLane, M. F., Barin, F., Essex, M., & Fischinger, P. J. (1987). Rapid assessment of relationships among HIV isolates by oligopeptide analyses of external envelope glycoproteins. *AIDS research and human retroviruses*, *3*(4), 401-408.

Robinson-McCarthy, L. R., McCarthy, K. R., Raaben, M., Piccinotti, S., Nieuwenhuis, J., Stubbs, S. H., Bakkers, M. J. G., & Whelan, S. P. J. (2018). Reconstruction of the cell entry pathway of an extinct virus. *PLoS Pathog*, *14*(8), e1007123.

Robinson, L. R., & Whelan, S. P. (2016). Infectious Entry Pathway Mediated by the Human Endogenous Retrovirus K Envelope Protein. *J Virol*, *90*(7), 3640-3649.

Robinson, M., Schor, S., Barouch-Bentov, R., & Einav, S. (2018). Viral journeys on the intracellular highways. *Cellular and Molecular Life Sciences*, *75*(20), 3693-3714.

Roe, T., Reynolds, T. C., Yu, G., & Brown, P. (1993). Integration of murine leukemia virus DNA depends on mitosis. *The EMBO journal*, *12*(5), 2099-2108.

Ross, S. R., Schofield, J. J., Farr, C. J., & Bucan, M. (2002). Mouse transferrin receptor 1 is the cell entry receptor for mouse mammary tumor virus. *Proceedings of the National Academy of Sciences*, *99*(19), 12386-12390.

Rostand, K. S., & Esko, J. D. (1997). Microbial adherence to and invasion through proteoglycans. *Infection and immunity*, *65*(1), 1-8.

Ruggieri, A., Maldener, E., Sauter, M., Mueller-Lantzsch, N., Meese, E., Fackler, O. T., & Mayer, J. (2009). Human endogenous retrovirus HERV-K (HML-2) encodes a stable signal peptide with biological properties distinct from Rec. *Retrovirology*, *6*(1), 1-20.

Rusnati, M., Coltrini, D., Oreste, P., Zoppetti, G., Albini, A., Noonan, D., Di Fagagna, F. D. A., Giacca, M., & Presta, M. (1997). Interaction of HIV-1 Tat protein with heparin: role of the backbone structure, sulfation, and size. *Journal of Biological Chemistry*, *272*(17), 11313-11320.

Rusnati, M., & Presta, M. (2002). HIV-1 Tat protein: a target for the development of anti-AIDS therapies. *Drug Fut*, *27*, 481-493.

Rusnati, M., Tulipano, G., Spillmann, D., Tanghetti, E., Oreste, P., Zoppetti, G., Giacca, M., & Presta, M. (1999). Multiple interactions of HIV-I Tat protein with size-defined heparin oligosaccharides. *Journal of Biological Chemistry*, *274*(40), 28198-28205.

Rusnati, M., & Urbinati, C. (2009). Polysulfated/sulfonated compounds for the development of drugs at the crossroad of viral infection and oncogenesis. *Current pharmaceutical design*, *15*(25), 2946-2957.

Rusnati, M., Vicenzi, E., Donalisio, M., Oreste, P., Landolfo, S., & Lembo, D. (2009). Sulfated K5 Escherichia coli polysaccharide derivatives: a novel class of candidate antiviral microbicides. *Pharmacology & therapeutics*, *123*(3), 310-322.

Saeki, N., Ono, H., Sakamoto, H., & Yoshida, T. (2015). Down-regulation of immune-related genes by PSCA in gallbladder cancer cells implanted into mice. *Anticancer research*, *35*(5), 2619-2625.

Salanga, C., & Handel, T. (2011). Chemokine oligomerization and interactions with receptors and glycosaminoglycans: the role of structural dynamics in function. *Experimental cell research*, *317*(5), 590-601.

Salgado, M., López-Romero, P., Callejas, S., López, M., Labarga, P., Dopazo, A., Soriano, V., & Rodés, B. (2011). Characterization of host genetic expression patterns in HIV-infected individuals with divergent disease progression. *Virology*, *411*(1), 103-112.

Salmivirta, M., Lidholt, K., & Lindahl, U. (1996). Heparan sulfate: a piece of information. *The FASEB Journal*, *10*(11), 1270-1279.

Saphire, A. C., Bobardt, M. D., Zhang, Z., David, G., & Gallay, P. A. (2001). Syndecans serve as attachment receptors for human immunodeficiency virus type 1 on macrophages. *Journal of virology*, *75*(19), 9187-9200.

Sato, T., Gotoh, M., Kiyohara, K., Kameyama, A., Kubota, T., Kikuchi, N., Ishizuka, Y., Iwasaki, H., Togayachi, A., & Kudo, T. (2003). Molecular cloning and characterization of a novel human β1, 4-N-acetylgalactosaminyltransferase, β4GalNAc-T3, responsible for the synthesis of N, N′-diacetyllactosediamine, GalNAcβ1–4GlcNAc. *Journal of Biological Chemistry*, *278*(48), 47534-47544.

Saunders, S., Paine-Saunders, S., & Lander, A. D. (1997). Expression of the cell surface proteoglycan glypican-5 is developmentally regulated in kidney, limb, and brain. *Developmental biology*, *190*(1), 78-93.

Schmitt, K., Reichrath, J., Roesch, A., Meese, E., & Mayer, J. (2013). Transcriptional profiling of human endogenous retrovirus group HERV-K (HML-2) loci in melanoma. *Genome biology and evolution*, *5*(2), 307-328.

Schneider-Schaulies, J. (2000). Cellular receptors for viruses: links to tropism and pathogenesis. *Microbiology*, *81*(6), 1413-1429.

Schnierle, B. S., Stitz, J., Bosch, V., Nocken, F., Merget-Millitzer, H., Engelstädter, M., Kurth, R., Groner, B., & Cichutek, K. (1997). Pseudotyping of murine leukemia virus with the envelope glycoproteins of HIV generates a retroviral vector with specificity of infection for CD4-expressing cells. *Proceedings of the National Academy of Sciences*, *94*(16), 8640-8645.

Schröder, A. R., Shinn, P., Chen, H., Berry, C., Ecker, J. R., & Bushman, F. (2002). HIV-1 integration in the human genome favors active genes and local hotspots. *Cell*, *110*(4), 521-529.

Seifarth, W., Frank, O., Zeilfelder, U., Spiess, B., Greenwood, A. D., Hehlmann, R. d., & Leib-Mösch, C. (2005). Comprehensive analysis of human endogenous retrovirus transcriptional activity in human tissues with a retrovirus-specific microarray. *Journal of virology*, *79*(1), 341-352.

SenGupta, D., Tandon, R., Vieira, R. G., Ndhlovu, L. C., Lown-Hecht, R., Ormsby, C. E., Loh, L., Jones, R. B., Garrison, K. E., & Martin, J. N. (2011). Strong human endogenous retrovirus-specific T cell responses are associated with control of HIV-1 in chronic infection. *Journal of virology*, *85*(14), 6977-6985.

Sharom, F. J., & Radeva, G. (2004). GPI-anchored protein cleavage in the regulation of transmembrane signals. *Membrane Dynamics and Domains*, 285-315.

Sherman, M. P., & Greene, W. C. (2002). Slipping through the door: HIV entry into the nucleus. *Microbes and infection*, *4*(1), 67-73.

Shi, Q., Jiang, J., & Luo, G. (2013). Syndecan-1 serves as the major receptor for attachment of hepatitis C virus to the surfaces of hepatocytes. *Journal of virology*, *87*(12), 6866-6875.

Shoemaker, R. H. (2006). The NCI60 human tumour cell line anticancer drug screen. *Nature Reviews Cancer*, *6*(10), 813-823.

Shriver, Z., Capila, I., Venkataraman, G., & Sasisekharan, R. (2012). Heparin and heparan sulfate: analyzing structure and microheterogeneity. *Heparin-A Century of Progress*, 159-176.

Shukla, D., Liu, J., Blaiklock, P., Shworak, N. W., Bai, X., Esko, J. D., Cohen, G. H., Eisenberg, R. J., Rosenberg, R. D., & Spear, P. G. (1999). A novel role for 3-O-sulfated heparan sulfate in herpes simplex virus 1 entry. *Cell*, *99*(1), 13-22.

Singer, V. L., Jones, L. J., Yue, S. T., & Haugland, R. P. (1997). Characterization of PicoGreen reagent and development of a fluorescence-based solution assay for double-stranded DNA quantitation. *Analytical biochemistry*, *249*(2), 228-238.

Singh, A., Razooky, B., Cox, C. D., Simpson, M. L., & Weinberger, L. S. (2010). Transcriptional bursting from the HIV-1 promoter is a significant source of stochastic noise in HIV-1 gene expression. *Biophysical journal*, *98*(8), L32-L34.

Snyder, M. M., Yue, F., Zhang, L., Shang, R., Qiu, J., Chen, J., Kim, K. H., Peng, Y., Oprescu, S. N., & Donkin, S. S. (2021). LETMD1 is required for mitochondrial structure and thermogenic function of brown adipocytes. *The FASEB Journal*, *35*(11), e21965.

Spillmann, D. (2001). Heparan sulfate: anchor for viral intruders? *Biochimie*, *83*(8), 811-817.

Spilsbury, K., O'Mara, M.-A., Wu, W. M., Rowe, P. B., Symonds, G., & Takayama, Y. (1995). Isolation of a novel macrophage-specific gene by differential cDNA analysis.

Stamatiades, E. G., Tremblay, M.-E., Bohm, M., Crozet, L., Bisht, K., Kao, D., Coelho, C., Fan, X., Yewdell, W. T., & Davidson, A. (2016). Immune monitoring of trans-endothelial transport by kidney-resident macrophages. *Cell*, *166*(4), 991-1003.

Stelzer, G., Rosen, N., Plaschkes, I., Zimmerman, S., Twik, M., Fishilevich, S., Stein, T. I., Nudel, R., Lieder, I., & Mazor, Y. (2016). The GeneCards suite: from gene data mining to disease genome sequence analyses. *Current protocols in bioinformatics*, *54*(1), 1.30. 31-31.30. 33.

Stencel-Baerenwald, J. E., Reiss, K., Reiter, D. M., Stehle, T., & Dermody, T. S. (2014). The sweet spot: defining virus–sialic acid interactions. *Nature Reviews Microbiology*, *12*(11), 739-749.

Stevens, J., Chen, L.-M., Carney, P. J., Garten, R., Foust, A., Le, J., Pokorny, B. A., Manojkumar, R., Silverman, J., & Devis, R. (2010). Receptor specificity of influenza A H3N2 viruses isolated in mammalian cells and embryonated chicken eggs. *Journal of virology*, *84*(16), 8287-8299.

Stick, R. V., & Williams, S. J. (2009). Glycoproteins and Proteoglycans. In R. V. Stick & S. J. Williams (Eds.), *Carbohydrates: The Essential Molecules of Life* (2nd ed., pp. 369-412). Elsevier.

Stoeber, M., Stoeck, I. K., Hänni, C., Bleck, C. K. E., Balistreri, G., & Helenius, A. (2012). Oligomers of the ATPase EHD2 confine caveolae to the plasma membrane through association with actin. *The EMBO journal*, *31*(10), 2350-2364.

Stoye, J. P. (2012). Studies of endogenous retroviruses reveal a continuing evolutionary saga. *Nature Reviews Microbiology*, *10*(6), 395-406.

Strbo, N., Pastar, I., Romero, L., Chen, V., Vujanac, M., Sawaya, A. P., Jozic, I., Ferreira, A. D., Wong, L. L., & Head, C. (2019). Single cell analyses reveal specific distribution of anti-bacterial molecule Perforin-2 in human skin and its modulation by wounding and Staphylococcus aureus infection. *Experimental dermatology*, *28*(3), 225-232.

Subramanian, R. P., Wildschutte, J. H., Russo, C., & Coffin, J. M. (2011). Identification, characterization, and comparative genomic distribution of the HERV-K (HML-2) group of human endogenous retroviruses. *Retrovirology*, *8*(1), 1-22.

Sugahara, K., & Kitagawa, H. (2002). Heparin and heparan sulfate biosynthesis. *IUBMB life*, *54*(4), 163-175.

Suntsova, M., Gogvadze, E. V., Salozhin, S., Gaifullin, N., Eroshkin, F., Dmitriev, S. E., Martynova, N., Kulikov, K., Malakhova, G., & Tukhbatova, G. (2013). Human-specific endogenous retroviral insert serves as an enhancer for the schizophrenia-linked gene PRODH. *Proceedings of the National Academy of Sciences*, *110*(48), 19472-19477.

Sutton, R. E., & Littman, D. R. (1996). Broad host range of human T-cell leukemia virus type 1 demonstrated with an improved pseudotyping system. *Journal of virology*, *70*(10), 7322-7326.

Tailor, C. S., Nouri, A., Lee, C. G., Kozak, C., & Kabat, D. (1999). Cloning and characterization of a cell surface receptor for xenotropic and polytropic murine leukemia viruses. *Proceedings of the National Academy of Sciences*, *96*(3), 927-932.

Tailor, C. S., Nouri, A., Zhao, Y., Takeuchi, Y., & Kabat, D. (1999). A sodium-dependent neutral-amino-acid transporter mediates infections of feline and baboon endogenous retroviruses and simian type D retroviruses. *Journal of virology*, *73*(5), 4470-4474.

Tam, O. H., Rozhkov, N. V., Shaw, R., Kim, D., Hubbard, I., Fennessey, S., Propp, N., Phatnani, H., Kwan, J., & Sareen, D. (2019). Postmortem cortex samples identify distinct molecular subtypes of ALS: retrotransposon activation, oxidative stress, and activated glia. *Cell reports*, *29*(5), 1164-1177. e1165.

Tandon, R., SenGupta, D., Ndhlovu, L. C., Vieira, R. G., Jones, R. B., York, V. A., Vieira, V. A., Sharp, E. R., Wiznia, A. A., & Ostrowski, M. A. (2011). Identification of human endogenous retrovirus-specific T cell responses in vertically HIV-1-infected subjects. *Journal of virology*, *85*(21), 11526-11531.

Tao, Y., Strelkov, S. V., Mesyanzhinov, V. V., & Rossmann, M. G. (1997). Structure of bacteriophage T4 fibritin: a segmented coiled coil and the role of the C-terminal domain. *Structure*, *5*(6), 789-798.

Tatsis, N., Tesema, L., Robinson, E., Giles-Davis, W., McCoy, K., Gao, G., Wilson, J., & Ertl, H. (2006). Chimpanzee-origin adenovirus vectors as vaccine carriers. *Gene therapy*, *13*(5), 421-429.

Taylor, M. P., Koyuncu, O. O., & Enquist, L. W. (2011). Subversion of the actin cytoskeleton during viral infection. *Nature Reviews Microbiology*, *9*(6), 427-439.

Tedbury, P. R., & Freed, E. O. (2015). The cytoplasmic tail of retroviral envelope glycoproteins. *Progress in molecular biology and translational science*, *129*, 253-284.

Thiery, J., Keefe, D., Saffarian, S., Martinvalet, D., Walch, M., Boucrot, E., Kirchhausen, T., & Lieberman, J. (2010). Perforin activates clathrin-and dynamin-dependent endocytosis, which is required for plasma membrane repair and delivery of granzyme B for granzyme-mediated apoptosis. *Blood, The Journal of the American Society of Hematology*, *115*(8), 1582-1593.

Thul, P. J., Åkesson, L., Wiking, M., Mahdessian, D., Geladaki, A., Ait Blal, H., Alm, T., Asplund, A., Björk, L., & Breckels, L. M. (2017). A subcellular map of the human proteome. *Science*, *356*(6340), eaal3321.

Toida, T., Yoshida, H., Toyoda, H., Koshiishi, I., Imanari, T., Hileman, R. E., Fromm, J. R., & Linhardt, R. J. (1997). Structural differences and the presence of unsubstituted amino groups in heparan sulphates from different tissues and species. *Biochemical Journal*, *322*(2), 499-506.

Tönjes, R. R., Boller, K., Limbach, C., Lugert, R., & Kurth, R. (1997). Characterization of human endogenous retrovirus type K virus-like particles generated from recombinant baculoviruses. *Virology*, *233*(2), 280-291.

Tönjes, R. R., Löwer, R., Boller, K., Denner, J., Hasenmaier, B., Kirsch, H., König, H., Korbmacher, C., Limbach, C., & Lugert, R. (1996). HERV-K: the biologically most active human endogenous retrovirus family. *JAIDS Journal of Acquired Immune Deficiency Syndromes*, *13*, S261-S267.

Toyoda, M., Kaji, H., Sawaki, H., Togayachi, A., Angata, T., Narimatsu, H., & Kameyama, A. (2016). Identification and characterization of sulfated glycoproteins from small cell lung carcinoma cells assisted by management of molecular charges. *Glycoconj J*, *33*(6), 917-926.

Trowbridge, J. M., Rudisill, J. A., Ron, D., & Gallo, R. L. (2002). Dermatan sulfate binds and potentiates activity of keratinocyte growth factor (FGF-7). *Journal of Biological Chemistry*, *277*(45), 42815-42820.

Tsai, B., Ye, Y., & Rapoport, T. A. (2002). Retro-translocation of proteins from the endoplasmic reticulum into the cytosol. *Nature reviews Molecular cell biology*, *3*(4), 246-255.

Tsetlin, V. I. (2015). Three-finger snake neurotoxins and Ly6 proteins targeting nicotinic acetylcholine receptors: pharmacological tools and endogenous modulators. *Trends in pharmacological sciences*, *36*(2), 109-123.

Turner, G., Barbulescu, M., Su, M., Jensen-Seaman, M. I., Kidd, K. K., & Lenz, J. (2001). Insertional polymorphisms of full-length endogenous retroviruses in humans. *Current Biology*, *11*(19), 1531-1535.

Tyagi, M., Rusnati, M., Presta, M., & Giacca, M. (2001). Internalization of HIV-1 tat requires cell surface heparan sulfate proteoglycans. *Journal of Biological Chemistry*, *276*(5), 3254-3261.

Tyrrell, D. J., Kilfeather, S., & Page, C. P. (1995). Therapeutic uses of heparin beyond its traditional role as an anticoagulant. *Trends in pharmacological sciences*, *16*(6), 198-204.

Urisman, A., Molinaro, R. J., Fischer, N., Plummer, S. J., Casey, G., Klein, E. A., Malathi, K., Magi-Galluzzi, C., Tubbs, R. R., & Ganem, D. (2006). Identification of a novel Gammaretrovirus in prostate tumors of patients homozygous for R462Q RNASEL variant. *PLoS pathogens*, *2*(3), e25.

Van de Lagemaat, L. N., Landry, J.-R., Mager, D. L., & Medstrand, P. (2003). Transposable elements in mammals promote regulatory variation and diversification of genes with specialized functions. *TRENDS in Genetics*, *19*(10), 530-536.

Van den Bossche, J., Baardman, J., Otto, N. A., van der Velden, S., Neele, A. E., van den Berg, S. M., Luque-Martin, R., Chen, H.-J., Boshuizen, M. C., & Ahmed, M. (2016). Mitochondrial dysfunction prevents repolarization of inflammatory macrophages. *Cell reports*, *17*(3), 684-696.

Van Rheenen, W., Shatunov, A., Dekker, A. M., McLaughlin, R. L., Diekstra, F. P., Pulit, S. L., Van Der Spek, R. A., Võsa, U., De Jong, S., & Robinson, M. R. (2016). Genome-wide association analyses identify new risk variants and the genetic architecture of amyotrophic lateral sclerosis. *Nature genetics*, *48*(9), 1043-1048.

Varadi, M., Anyango, S., Deshpande, M., Nair, S., Natassia, C., Yordanova, G., Yuan, D., Stroe, O., Wood, G., & Laydon, A. (2022). AlphaFold Protein Structure Database: Massively expanding the structural coverage of protein-sequence space with high-accuracy models. *Nucleic acids research*, *50*(D1), D439-D444.

Vargiu, L., Rodriguez-Tomé, P., Sperber, G. O., Cadeddu, M., Grandi, N., Blikstad, V., Tramontano, E., & Blomberg, J. (2016). Classification and characterization of human endogenous retroviruses; mosaic forms are common. *Retrovirology*, *13*(1), 1-29.

Vergara Bermejo, A., Ragonnaud, E., Daradoumis, J., & Holst, P. (2020). Cancer associated endogenous retroviruses: ideal immune targets for adenovirus-based immunotherapy. *International Journal of Molecular Sciences*, *21*(14), 4843.

Vermeire, J., Naessens, E., Vanderstraeten, H., Landi, A., Iannucci, V., Van Nuffel, A., Taghon, T., Pizzato, M., & Verhasselt, B. (2012). Quantification of reverse transcriptase activity

by real-time PCR as a fast and accurate method for titration of HIV, lenti-and retroviral vectors. *PLoS One*, *7*(12), e50859.

Vigdorovich, V., Strong, R. K., & Miller, A. D. (2005). Expression and characterization of a soluble, active form of the jaagsiekte sheep retrovirus receptor, Hyal2. *Journal of virology*, *79*(1), 79-86.

Vincendeau, M., Göttesdorfer, I., Schreml, J. M., Wetie, A. G. N., Mayer, J., Greenwood, A. D., Helfer, M., Kramer, S., Seifarth, W., & Hadian, K. (2015). Modulation of human endogenous retrovirus (HERV) transcription during persistent and de novo HIV-1 infection. *Retrovirology*, *12*(1), 1-17.

ViralZone (RRID:SCR_006563). Swiss Institute of Bioinformatics. Retrieved 06.05.2022 from https://viralzone.expasy.org

Vogt, V. (1996). Proteolytic processing and particle maturation. In H. Kräusslich (Ed.), *Morphogenesis and maturation of retroviruses* (pp. 95-131). Springer Science & Business Media.

Walker, P. J., Siddell, S. G., Lefkowitz, E. J., Mushegian, A. R., Adriaenssens, E. M., Dempsey, D. M., Dutilh, B. E., Harrach, B., Harrison, R. L., & Hendrickson, R. C. (2020). Changes to virus taxonomy and the Statutes ratified by the International Committee on Taxonomy of Viruses (2020). In: Springer.

Wamara, J. (2020). *Identifizierung einer Rezeptorbindestelle im Hüllprotein des Humanen Endogenen Retrovirus K(HML-2) und weitergehende Analysen der Interaktion des Hüllproteins mit Zellulären Rezeptoren* [Doctoral dissertation, Technische Universität Berlin]. datenbank http://dx.doi.org/10.14279/depositonce-9983.

Wang-Johanning, F., Frost, A. R., Johanning, G. L., Khazaeli, M., LoBuglio, A. F., Shaw, D. R., & Strong, T. V. (2001). Expression of human endogenous retrovirus k envelope transcripts in human breast cancer. *Clinical Cancer Research*, *7*(6), 1553-1560.

Wang-Johanning, F., Liu, J., Rycaj, K., Huang, M., Tsai, K., Rosen, D. G., Chen, D. T., Lu, D. W., Barnhart, K. F., & Johanning, G. L. (2007). Expression of multiple human endogenous retrovirus surface envelope proteins in ovarian cancer. *International journal of cancer*, *120*(1), 81-90.

Weiss, R. (1969). The host range of Bryan strain Rous sarcoma virus synthesized in the absence of helper virus. *Journal of General Virology*, *5*(4), 511-528.

Welch, L. G., Peak-Chew, S.-Y., Begum, F., Stevens, T. J., & Munro, S. (2021). GOLPH3 and GOLPH3L are broad-spectrum COPI adaptors for sorting into intra-Golgi transport vesicles. *Journal of Cell Biology*, *220*(10).

Welfle, K., Misselwitz, R., Hausdorf, G., Höhne, W., & Welfle, H. (1999). Conformation, pH-induced conformational changes, and thermal unfolding of anti-p24 (HIV-1) monoclonal antibody CB4-1 and its Fab and Fc fragments. *Biochimica et Biophysica Acta (BBA)-Protein Structure and Molecular Enzymology*, *1431*(1), 120-131.

Wie, S.-H., Du, P., Luong, T. Q., Rought, S. E., Beliakova-Bethell, N., Lozach, J., Corbeil, J., Kornbluth, R. S., Richman, D. D., & Woelk, C. H. (2013). HIV downregulates interferon-stimulated genes in primary macrophages. *Journal of Interferon & Cytokine Research*, *33*(2), 90-95.

Wiens, M., Korzhev, M., Krasko, A., Thakur, N. L., Perovic-Ottstadt, S., Breter, H. J., Ushijima, H., Diehl-Seifert, B. r., Müller, I. M., & Müller, W. E. (2005). Innate immune defense of the sponge Suberites domuncula against bacteria involves a MyD88-

dependent signaling pathway: induction of a perforin-like molecule. *Journal of Biological Chemistry*, *280*(30), 27949-27959.

Wildschutte, J. H., Williams, Z. H., Montesion, M., Subramanian, R. P., Kidd, J. M., & Coffin, J. M. (2016). Discovery of unfixed endogenous retrovirus insertions in diverse human populations. *Proceedings of the National Academy of Sciences*, *113*(16), E2326-E2334.

Willer, A., Saußele, S., Gimbel, W., Zeifarth, W., Kister, P., & Hehlmann, R. (1997). Two groups of endogenous MMTV related retroviral env transcripts expressed in human tissues. *Virus genes*, *15*(2), 123-133.

Xiong, P., Shiratsuchi, M., Matsushima, T., Liao, J., Tanaka, E., Nakashima, Y., Takayanagi, R., & Ogawa, Y. (2017). Regulation of expression and trafficking of perforin-2 by LPS and TNF-α. *Cellular Immunology*, *320*, 1-10.

Xu, D., & Esko, J. D. (2014). Demystifying heparan sulfate–protein interactions. *Annual review of biochemistry*, *83*, 129-157.

Xu, R., de Vries, R. P., Zhu, X., Nycholat, C. M., McBride, R., Yu, W., Paulson, J. C., & Wilson, I. A. (2013). Preferential recognition of avian-like receptors in human influenza A H7N9 viruses. *Science*, *342*(6163), 1230-1235.

Xu, W., & Eiden, M. V. (2011). Primate gammaretroviruses require an ancillary factor not required for murine gammaretroviruses to infect BHK cells. *J Virol*, *85*(7), 3498-3506.

Xu, X., Zhao, H., Gong, Z., & Han, G.-Z. (2018). Endogenous retroviruses of non-avian/mammalian vertebrates illuminate diversity and deep history of retroviruses. *PLoS pathogens*, *14*(6), e1007072.

Yamagata, A., Yoshida, T., Sato, Y., Goto-Ito, S., Uemura, T., Maeda, A., Shiroshima, T., Iwasawa-Okamoto, S., Mori, H., & Mishina, M. (2015). Mechanisms of splicing-dependent trans-synaptic adhesion by PTPδ–IL1RAPL1/IL-1RAcP for synaptic differentiation. *Nature communications*, *6*(1), 1-11.

Yamauchi, Y., & Greber, U. F. (2016). Principles of virus uncoating: cues and the snooker ball. *Traffic*, *17*(6), 569-592.

Yamauchi, Y., & Helenius, A. (2013). Virus entry at a glance. *Journal of cell science*, *126*(6), 1289-1295.

Yan, J., Zhou, B., Guo, L., Chen, Z., Zhang, B., Liu, S., Zhang, W., Yu, M., Xu, Y., & Xiao, Y. (2020). GOLM1 upregulates expression of PD-L1 through EGFR/STAT3 pathway in hepatocellular carcinoma. *American journal of cancer research*, *10*(11), 3705.

Yang, C.-H., Li, H.-C., Shiu, Y.-L., Ku, T.-S., Wang, C.-W., Tu, Y.-S., Chen, H.-L., Wu, C.-H., & Lo, S.-Y. (2017). Influenza A virus upregulates PRPF8 gene expression to increase virus production. *Archives of virology*, *162*(5), 1223-1235.

Yang, X., Boehm, J. S., Yang, X., Salehi-Ashtiani, K., Hao, T., Shen, Y., Lubonja, R., Thomas, S. R., Alkan, O., Bhimdi, T., Green, T. M., Johannessen, C. M., Silver, S. J., Nguyen, C., Murray, R. R., Hieronymus, H., Balcha, D., Fan, C., Lin, C., Ghamsari, L., Vidal, M., Hahn, W. C., Hill, D. E., & Root, D. E. (2011). A public genome-scale lentiviral expression library of human ORFs. *Nat Methods*, *8*(8), 659-661.

Yang, X., Florin, L., Farzan, M., Kolchinsky, P., Kwong, P. D., Sodroski, J., & Wyatt, R. (2000). Modifications that stabilize human immunodeficiency virus envelope glycoprotein trimers in solution. *J Virol*, *74*(10), 4746-4754.

Yang, X., Lee, J., Mahony, E. M., Kwong, P. D., Wyatt, R., & Sodroski, J. (2002). Highly stable trimers formed by human immunodeficiency virus type 1 envelope glycoproteins fused with the trimeric motif of T4 bacteriophage fibritin. *J Virol*, *76*(9), 4634-4642.

Yang, Y.-L., Guo, L., Xu, S., Holland, C. A., Kitamura, T., Hunter, K., & Cunningham, J. M. (1999). Receptors for polytropic and xenotropic mouse leukaemia viruses encoded by a single gene at Rmc1. *Nature genetics*, *21*(2), 216-219.

Ye, Q. H., Zhu, W. W., Zhang, J. B., Qin, Y., Lu, M., Lin, G. L., Guo, L., Zhang, B., Lin, Z. H., Roessler, S., Forgues, M., Jia, H. L., Lu, L., Zhang, X. F., Lian, B. F., Xie, L., Dong, Q. Z., Tang, Z. Y., Wang, X. W., & Qin, L. X. (2016). GOLM1 Modulates EGFR/RTK Cell-Surface Recycling to Drive Hepatocellular Carcinoma Metastasis. *Cancer Cell*, *30*(3), 444-458.

Young, J. (2001). Virus entry and uncoating. In D. M. Knipe & P. M. Howley (Eds.), *Fields virology* (Vol. 1, pp. 87–103.). Lippincott Williams & Wilkins, Philadelphia.

Yu, J., Liang, C., & Liu, S.-L. (2017). Interferon-inducible LY6E protein promotes HIV-1 infection. *Journal of Biological Chemistry*, *292*(11), 4674-4685.

Zhang, H., Dornadula, G., Orenstein, J., & Pomerantz, R. (2000). Morphologic changes in human immunodeficiency virus type 1 virions secondary to intravirion reverse transcription: evidence indicating that reverse transcription may not take place within the intact viral core. *Journal of human virology*, *3*(3), 165-172.

Zhang, L.-l., Wei, J.-y., Wang, L., Huang, S.-l., & Chen, J.-l. (2017). Human T-cell lymphotropic virus type 1 and its oncogenesis. *Acta Pharmacologica Sinica*, *38*(8), 1093-1103.

Zhang, Y., Rassa, J. C., deObaldia, M. E., Albritton, L. M., & Ross, S. R. (2003). Identification of the receptor binding domain of the mouse mammary tumor virus envelope protein. *J Virol*, *77*(19), 10468-10478.

Zhao, Y., Zhu, L., Benedict, C. A., Chen, D., Anderson, W. F., & Cannon, P. M. (1998). Functional domains in the retroviral transmembrane protein. *Journal of virology*, *72*(7), 5392-5398.

Zhou, Z., Xu, L., Sennepin, A., Federici, C., Ganor, Y., Tudor, D., Damotte, D., Delongchamps, N. B., Zerbib, M., & Bomsel, M. (2018). The HIV-1 viral synapse signals human foreskin keratinocytes to secrete thymic stromal lymphopoietin facilitating HIV-1 foreskin entry. *Mucosal immunology*, *11*(1), 158-171.

Zwolińska, K., Knysz, B., Gąsiorowski, J., Pazgan-Simon, M., Gładysz, A., Sobczyński, M., & Piasecki, E. (2013). Frequency of human endogenous retroviral sequences (HERV) K113 and K115 in the Polish population, and their effect on HIV infection. *PLoS One*, *8*(10), e77820.

7. List of publication

Conference poster

Ramadan, A., Andus, D., Hohn, O., Bannert, N. (2021) Mutational analysis of a putative heparan sulfate-binding domain in the HERV-K(HML-2) envelope. [Poster presentation]. The 30th Annual Meeting of the Society for Virology: GFV 24-26 March. Digital.