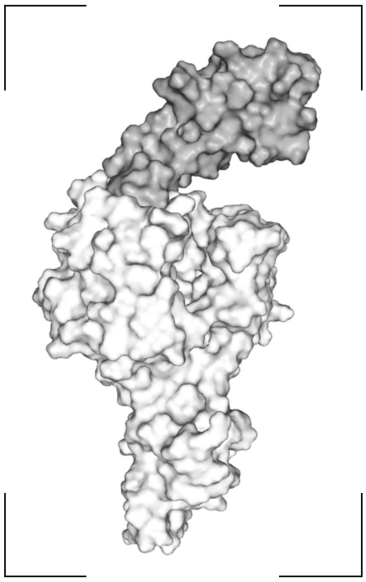
Nairovirus nucleoproteins: development of diagnostic tools, structural characterisation and functional studies

Beatriz Álvarez Rodríguez

Submitted in accordance with the requirements
for the degree of Doctor of Philosophy



Supervisor: Dr. John N. Barr

University of Leeds

The School of Molecular and Cellular Biology
The Astbury Centre for Structural Molecular Biology

August 2020

Intellectual property and publication statement

The candidate confirms that the work submitted is her own and that appropriate credit has been given where reference has been made to the work of others.

To Silvia, Clara, Jacobo and Mamá.

To the memory of my father.

Sometimes I've believed as many as
six impossible things before breakfast!
Lewis Carrol, Alice in Wonderland

Acknowledgements

Chapter 3: Dr. Christian Tiede (BSTG, University of Leeds) performed the CCHFV NP-specific Affimer screening.

Chapter 4: Dr. Chi Trinh helped with X-ray crystallography data collection and first stages of data processing. Dr. Thomas Edwards solved the crystal structure and performed data refinement.

Chapter 5: Alexis Hoste produced recombinant RVFV and SBV NPs, used as negative controls.

Foremost, I would like to thank John for his support since my first day in Leeds and his excellent guidance throughout this project. Ed, Paul, Jamel and Juan, thanks for getting involved in enthusiastic scientific discussions with me. A special thanks goes to everyone in 8.61, particularly to past and present members of the Barr group, for your help and support during these 3 years. I would like to acknowledge the assistance of the SPR, CD, crystallography and EM facilities, particularly to the tireless patience of Ian Manfield, Chi Trihn and Martin Fuller, who have invested a good amount of time in my training. A specific mention goes to the support staff of the University, without whom the science of this project would have not been possible, and the great team of Eurofins Ingenasa, that warmly hosted my PhD secondment.

A huge thanks goes to the EU Marie-Curie funding scheme, the HONOURs network and all the people involved in it. Being part of this network is a once in a lifetime experience that few people can have the pleasure to enjoy. Thanks for all the trips, science and beers we have shared.

I would like to thank Sergio for his -much needed- support. For all the science you have patiently listened to, all the drives to and back from the lab, and all the climbs and outdoor trips you have shared with me. Thanks to the climbing crew for all those evenings in 'the other' Lab and for enjoying with me the outdoor beauties of Yorkshire. Thanks to my family and friends from Spain. Silvia, Clara, Jacobo and Mamá, thanks for always receiving me as if I had never left.

Abbreviations

A	absorbance
Ab	antibody
AF	Affimer
amp	ampicillin
AP	affinity precipitation
ATP	adenosine triphosphate
Au	arbitrary units
AU	asymmetric unit
BHK	baby hamster kidney
BSA	bovine serum albumin
BSL	biosafety level
BSTG	BioScreening Technology Group
BUNV	Bunyamwera virus
C3	caspase 3
carb	carbenicillin
CCD	charge coupled device
CCHF	Crimean–Congo hemorrhagic fever
CCHFV	Crimean–Congo hemorrhagic fever virus
CD	circular dichroism
cDNA	complementary DNA
CDK-2	cyclin-dependent protein kinase 2
CLEM	correlative light electron microscopy
CMC	carboxy-methyl cellulose
CM-dextran	carboxymethyl-dextran
CMV	human cytomegalovirus
cRNA	complementary RNA
cRNP	complementary RNP
cryo-TEM	cryogenic transmission electron microscopy
dAb	domain antibodies
DBPA	2,4-dioxo-4-phenylbutanoic acid
DCs	dendritic cells
ddH_2O	double-distilled dihydrogen monoxide
DLS	dynamic light scattering
DMSO	dimethyl sulfoxide
DNase	deoxyribonuclease
ds	double stranded
DsbA	disulfide bond formation protein A
DTT	dithiothreitol
EBOV	Ebola virus
EMCV	encephalomyocarditis virus
ECs	endothelial cells
EDC	1-ethyl-3-(3-dimethylaminopropyl) carbodiimide
EDTA	ethylenediaminetetraacetic acid
eGFP	enhanced green fluorescence protein
eGFPcomp	complemented enhanced green fluorescence protein
ELISA	enzyme-linked immunosorbent assay
EM	electron microscopy
ER	endoplasmic reticulum

ESCRT	endosomal sorting complex required for transport
Fab	antigen-binding fragments
FBS	fetal bovine serum
Fc	fragment crystallizable
FcγRs	Fcγ receptors
FISH	fluorescence *in situ* hybridisation
Fl	fluorescein
FTCD	formiminotransferase cyclodeaminase
Fv	variable fragment
g3p	gene 3 protein
GAPDH	glyceraldehyde-3-phosphate dehydrogenase
GPC	glycoprotein precursor
GST	glutathione-S-transferase
GSH	glutathione
H	heavy chain
HAZV	Hazara virus
HDV RZ	hepatitis delta virus ribozyme
HLA	human leukocyte antigen
hpi	hours post-infection
HPLC	high performance liquid chromatography
hpt	hours post-transfection
HRP	horseradish peroxidase
HSP70	70 kilodalton heat shock proteins
I-CLiPs	intramembrane cleaving proteases
IC_{50}	half maximal inhibitory concentration
ICAM-1	intercellular adhesion molecule 1
ICTV	International Committee on Taxonomy of Viruses
IL	interleukin
IF	immunofluorescence
IFN	interferon
IMAC	immobilised metal affinity chromatography
IRES	internal ribosome entry site
IRF-3	interferon regulatory factor-3
Ig	immunoglobulin
ISG15	interferon-stimulated gene 15
J chain	joining chain
kan	kanamycin
kbp	kilobasepairs
K_D	dissociation constant
kDa	kiloDalton
KLD	kinase, ligase and *DpnI*
k_{off}	dissociation rate constant
k_{on}	association rate constant
KUPV	Kupe virus
L	light chain
L segment	large segment
LACV	La Crosse virus
LASV	Lassa virus
LB	Luria-Bertani
LDS	lithium dodecyl sulfate
LFA	lateral flow assay
LOD	limit of detection
M segment	medium segment
m^7G	7-methylguanylate
Mb	myoglobin
MLD	mucin-like domain
M.O.I.	multiplicity of infection
MR	molecular replacement
MRI	magnetic resonance imaging
mRNA	messenger RNA
MSCA-ITN	Marie Skłodowska-Curie Actions Innovative Training Network

MVB	multivesicular bodies
MW	molecular weight
MWCO	molecular weight cut-off
MxA	myxovirus resistance protein 1
NHS	N-hydroxysuccinimide
Ni^{2+}-NTA	nickel-nitrilotriacetic acid
NMR	nuclear magnetic resonance
NP	nucleoprotein
NSDV	Nairobi sheep disease virus
NS_S	non-structural S protein
NTA	nitrilotriacetic acid
NTPs	nucleoside triphosphates
OD	optical density
ORD	optical rotary dispersion
ORF	open reading frame
OTU	ovarian tumour
p-polarised	parallel polarised
PABP	poly-A binding protein
PCR	polymerase chain reaction
PD	pull down
PDB	protein data bank
PE	promoter element
pI	isoelectric point
pfu	plaque forming units
PMT	photomultiplier tubes
pol I	polymerase I
PVDF	polyvinylidene fluoride
RdRp	RNA dependent RNA polymerase
rHAZV	recombinant Hazara virus
RIPA	radioimmunoprecipitation assay buffer
RFP	red fluorescent protein
RNase	ribonuclease
RNP	ribonucleoprotein
RT	room temperature
RT-PCR	reverse transcription polymerase chain reaction
RT-qPCR	reverse transcription quantitative polymerase chain reaction
RU	response units
RVFV	Rift Valley fever virus
S segment	small segment
scFv	single chain variable fragments
SBV	Schmallenberg virus
SD	standard deviation
SDM	site-directed mutagenesis
SDS	sodium dodecyl sulfate
SDS-PAGE	SDS-polyacrylamide gel electrophoresis
SEC	size exclusion chromatography
SELEX	systematic evolution of ligands by exponential enrichment
SHG	second-harmonic generation
SKI-1/S1P	subtilisin kexin isozyme-1/site-1 protease
sNSVs	segmented negative-strand RNA viruses
SNV	sin nombre virus
SOC	super optimal broth with catabolite repression
SP	signal peptidases
spcr	spectinomycin
SPR	surface plasmon resonance
ss	single stranded
SUMO	small ubiquitin-like modifier
T7 pol	bacteriophage T7 DNA dependent RNA polymerase
TAE	Tris-acetate-EDTA
TCEP	Tris(2-carboxyethyl)phosphine

TEM transmission electron microscopy

TFLV Tofla virus

TGN trans-Golgi network

Tm melting temperature

TMB 3,3',5,5'-tetramethylbenzidine

TNE Tris-NaCl-EDTA

TNF-α tumour necrosis factor-α

TrxA thioredoxin 1

UB unbound

Ub ubiquitin

Ulp1 ubiquitin-like specific protease 1

UTR untranslated region

UUKV Uukuniemi phlebovirus

UV ultraviolet

UV-TPEF two-photon excited ultraviolet fluorescence

v volume

V variable region

VHH variable heavy homodimers

VLP virus-like particle

VNAR variable new antigen receptor

vRNA viral ribonucleic acid

vRNP viral ribonucleoprotein

w weight

WHO World Health Organisation

WT wild type

YARU Yale arbovirus research unit

ZIKV Zika virus

Abstract

Crimean-Congo hemorrhagic fever virus (CCHFV) is one of the most widespread medically important arboviruses, causing human infections that result in case fatality rates of up to 60%, and for which no effective preventative or therapeutic measures are available. Here, the selection and biophysical characterisation of a high-affinity small protein (Affimer-NP) that binds specifically to the nucleoprotein (NP) of CCHFV and possesses potential antiviral properties are described. Affimer-NP interferes in the RNA binding function of CCHFV NP, and inhibits CCHFV gene expression, reconstituted using a mini-genome system in mammalian cells. Solution of the crystallographic structure of the complex formed by these two molecules at 2.84 Å resolution revealed the structural basis for these interferences, with the Affimer-NP binding site positioned at a critical region of the NP involved in RNA binding and oligomerization. Additionally, the *in vitro* application of this novel molecule for the development of enzyme-linked immunosorbent and lateral flow assays is validated, presenting the first published point-of-care test able to detect CCHFV NP in spiked human and animal sera.

The study of nairoviral NPs is further deepened using a reverse genetics system for Hazara virus (HAZV), a closely related homologue of CCHFV, for the incorporation of clonable tags to study the morphology and intracellular trafficking of nairoviral ribonucleoproteins (RNPs). A C-terminal 6xHis tag was engineered in the NP of HAZV for the purification of native viral RNPs in the context of infectious virus using affinity chromatography. Structural analyses of purified RNPs using electron microscopy revealed the helical structure of native nairoviral RNPs. A split-enhanced green fluorescent protein (split-eGFP) strategy was used to tag and visualise the NP of HAZV in infected cells. This system, combined with other fluorescence-based labelling techniques, allowed the characterisation of the intracellular dynamics of HAZV components and revealed the localisation of nairoviral replication factories using confocal microscopy. These results validate a novel approach to design and produce tagged-NP replication-competent recombinant HAZV that represents a valuable tool for the study of nairoviruses.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Viruses are infectious molecular machines lying on the borderline between inanimate and living entities. The term virus, borrowed from Latin *vīrus*, referring to 'poison, slime or venom', was used for the first time in 1728 to describe an agent that causes infectious disease. But it was not until 1892 when Dimitri Ivanovski and Martinus Beijerinck identified the first virus, tobacco mosaic virus, using a porcelain filter previously designed by Charles Chamberland to isolate it [1]. Despite disagreements on the origin of viruses, and the lack of tools to investigate the nanoscopic world until the 20th century, there is scientific evidence that these entities have evolved and co-existed with life for millions of years [2]. Viruses are obligate intracellular parasites which require the machinery and metabolism of a host cell for replication. Only a limited range of hosts can be infected by each particular virus, but within the vast diversity of existing viruses the host range includes members from the six kingdoms of life: *Archaebacteria, Eubacteria, Protista, Fungi, Plantae* and *Animalia*.

Humans are constantly exposed to this large diversity of viruses, which are very genetically heterogeneous and may have the potential to infect humans or to evolve to do so. There are more than 200 viral species known to be able to infect humans, and this number increases with the finding of new species every year [3]. A substantial proportion of mammalian viruses may be capable of crossing the species barrier into humans, a phenomenon known as zoonosis [4]. It is predicted that only around half of these would be capable of being transmitted by humans, and around half again of causing major outbreaks [3]. With globalisation of human movement, increased density of human populations, climate change and loss of natural habitats, there is increased risk of interaction between species, and within species. This reinforces the emergence of infectious diseases by increased contact and cross-species transmission, as well as

movement of infectious agents into geographic regions where they have not been found before [5].

Emerging infectious diseases are a major contemporary concern due to their impact on both the public health and economy of developing and developed nations. A recognisable example is the ongoing pandemic outbreak of pneumonia caused by a novel coronavirus, severe acute respiratory syndrome coronavirus-2, which has been responsible for more than 800,000 deaths as of August 2020, and is causing the worst worldwide economic recession since the Great Depression of the 1930s [6]. Viral outbreaks are generally unexpected, and therefore preparedness for their prevention and control is crucial. Tools and abilities for immediate action in outbreak situations include the identification and characterisation of the causative agent, the evaluation of its pathogenic potential, and the fast development of diagnostic assays. The work presented in this thesis forms part of the HONOURs project, a Marie Skłodowska-Curie Actions Innovative Training Network (MSCA-ITN) teaching 15 early stage researchers to become preparedness-experts on host switching pathogens, infectious outbreaks and zoonosis [7].

Nearly two thirds of emerging infectious diseases are caused by zoonotic viruses [8], most of them having wild mammal origins such as human immunodeficiency virus or Ebola virus (EBOV). There are several routes of transmission for zoonotic diseases, which can be divided into either contact with the infected animal host or contact with the infected arthropod host. In the case of the animal, virus contact can be through virus contaminated objects known as fomites, inhalation of aerosolised virus from infected body fluids, or direct contact with the virus through an animal bite. In the case of the arthropod vector, transmission can occur by biting or mechanical transfer of the virus, which is predicted to be the most common infection route for viral zoonosis [8]. Arboviruses are defined as arthropod-borne viruses, and commonly cause unprecedented outbreaks such as the recent Chikungunya and Zika virus (ZIKV) epidemics [9]. They are becoming a threat worldwide due to their ability to spillover from their natural niches, mainly occurring in the regions of Asia, America and Africa, to susceptible urban regions [9, 10, 11].

Arboviruses affecting humans involve three main viral groups: family *Flaviviridae*, family *Togaviridae* and order *Bunyavirales*; and a small number of members belong to the *Reoviridae*, and *Orthomyxoviridae* families [10, 11, 12]. In this project, the focus has been placed on two arboviruses of the order *Bunyavirales*, family *Nairoviridae*: the zoonotic and highly-pathogenic Crimean-Congo hemorrhagic fever virus (CCHFV), and its homologue and non-pathogenic in humans Hazara virus (HAZV).

1.1 CCHFV and HAZV

1.1.1 Discovery

Crimean-Congo hemorrhagic fever virus was first recognised in the Crimean peninsula in 1944, when more than 200 cases of a severe febrile illness accompanied by bleeding and shock occurred in Soviet soldiers working to restore agricultural production in farmland abandoned during the German occupation in World War II [13, 14]. Mikhail Chumakov led the team of scientists sent from Moscow to investigate the outbreak. An association of what was then called 'Crimean disease with tick bites' was quickly recognised and linked to the increase in the population of hares and other wild hosts of *Hyalomma* ticks as consequence of the abundance of abandoned farm land during the German occupation. Etiology of the disease was established through experimental infection studies via inoculation of filtered tick extracts in military volunteers and psychiatric patients [15, 16].

In 1967, Chumakov and his colleagues at the Institute of Poliomyelitis and Viral Encephalitis in Moscow, first used new-born white mice and rats for propagation and isolation of CCHFV through inoculation of samples from patients, resulting in the isolation of nine CCHFV strains. The Drozdov strain, named after the patient it was isolated from, was used as the prototype to develop antigens and antibodies for consecutive serological surveys and investigations of virus isolates from different geographical locations [17]. In 1968, the Drozdov strain was transferred to the Yale arbovirus research unit (YARU), where Jordi Casals associated the Crimean virus with three different strains of Congo, at the time a poorly characterised virus isolated from human patients from the Congo and Uganda, and from *Hyalomma* ticks from Pakistan [18]. The Crimean and the Congo viruses seemed to be antigenically indistinguishable using complement fixation and neutralisation test experiments [19].

Less than 2 years after the first isolation of CCHFV by Chumakov and his colleagues, it was demonstrated that the Crimean-Congo hemorrhagic fever (CCHF) causative agent occurs in nature from Pakistan to Bulgaria and several regions of Africa, leading to the renaming of the two virus isolates to Crimean-Congo hemorrhagic fever virus [20]. Since the first documented appearance of CCHFV in 1944, more than 140 CCHFV outbreaks and over 5000 cases have been reported worldwide [21]. CCHFV is now considered one of the most widespread medically important emerging infectious diseases, and its geographical range is drastically expanding to new areas where the CCHFV agent had never been described before [22].

HAZV was first isolated from a pool of six adult *Ixodes redikorzevi* ticks collected from subarctic terrain, 12,000 ft above sea level, in the vole *Alticola roylei*, in the Kaghan Valley, West Pakistan [23, 24]. The virus, firstly named JC 280, was adapted to suckling mice and the preparation of antigens and mouse immune serum were described together with eight other viral agents isolated from hard ticks collected in Pakistan [25]. JC 280 showed cross-reactivity in neutralisation tests with CCHFV, indicating a serological relationship between them, but no cross-reactivity was shown in complement fixation tests with CCHFV antigen. These findings suggested that JC 280 was a distinct virus from CCHFV, and thus JC 280 gained its current nomenclature, HAZV, from the place it was isolated from, the Hazara District [23]. HAZV has been shown to be lethal in adult mice deficient in the type I interferon receptor (IFNAR), but infectivity and pathogenicity have not been reported in humans to date. Therefore, infectious HAZV can be handled at containment level 2 and has been used as a model virus for CCHFV.

In 2016, a novel orthonairovirus named Tofla virus (TFLV) was described in Japan [26]. Two different strains were isolated from *Heamaphysalis flava* and *Heamaphysalis formsensis* ticks collected in Tokushima and Nagasaki, respectively. Sequence analyses revealed similarities between these viruses and HAZV with 75.7%, 61.3% and 21.4% amino acid shared identity of the small (S), medium (M) and large (L) segments, respectively. Cross-neutralisation between HAZV and TFLV was also described, suggesting that TFLV forms part of the CCHFV serogroup. Furthermore, TFLV exhibited lethal infection in IFNAR-knockout mice, as HAZV did, suggesting that TFLV could also serve as modelling virus for CCHFV. TFLV infections in humans and animals are currently unknown, but this virus is normally handled under BSL-3 facilities because its pathogenic potential in humans is not clear [26]. Due to the similarities between HAZV and TFLV, both of them have been classified as members of the *Hazara orthonairovirus* species [27].

1.1.2 Classification

The realm *Riboviria* comprises all RNA viruses and viroids. It includes the phylum *Negarnaviricota*, containing all negative sense single-stranded RNA (ssRNA) viruses (group V in Baltimore classification), and multiple unclassified orders, families, and genera (Figure 1.1). The phylum *Negarnaviricota* is, in turn, divided into the subphyla *Haploviricotina* and *Polyploviricotina*. As their nomenclatures suggest, the *Haploviricotina* subphylum (from the Ancient Greek haplo, 'simple') is formed by non-segmented negative sense RNA viruses, and the *Polyploviricotina* subphylum (from the

Ancient Greek polyplo, 'complex') is formed by segmented negative sense RNA viruses.

The *Polyploviricotina* subphylum is constituted by two classes: *Ellioviricetes* and *Insthoviricetes*, which encompass the orders *Bunyavirales* and *Articulavirales*, respectively. According to the last taxonomic update accepted by the International Committee on Taxonomy of Viruses (ICTV) [27], the order *Bunyavirales* comprises twelve families: *Arenaviridae, Cruliviridae, Fimoviridae, Hantaviridae, Leishbuviridae, Mypoviridae, Nairoviridae, Peribunyaviridae, Phasmaviridae, Phenuiviridae, Tospoviridae* and *Wupedeviridae*. Most members from the *Bunyavirales* order are transmitted by hematophagous arthropods including mosquitoes, midges, flies, and ticks. Exceptions are hantaviruses and arenaviruses, which are transmitted by rodents, and tospoviruses and fimoviruses, which are plant-specific and are transmitted via non-hematophagous vectors, namely thrips and eriophyid mite vectors, respectively [28].

Bunyavirus families present very different host ranges. Members of the *Tospoviridae* and *Fimoviridae* families infect plants, whereas the *Phasmaviridae* family is insect-specific. *Leishbuviridae* family members are protozoa-specific and *Cruliviridae, Mypoviridae* and *Wupedeviridae* family members infect invertebrate hosts. Members belonging to other bunyavirus families infect vertebrates and cause serious human diseases, such as Lassa virus (LASV; family *Arenaviridae*), Rift Valley fever virus (RVFV; family *Phenuiviridae*), sin nombre virus (SNV; family *Hantavirus*), La Crosse virus (LACV; family *Peribunyaviridae*) or CCHFV (family *Nairoviridae*).

Both CCHFV and HAZV are members of the family *Nairoviridae*, genus *Orthonairovirus*. CCHFV is the only member of the *Crimean-Congo hemorrhagic fever orthonairovirus* species, and HAZV, together with TFLV, forms the *Hazara orthonairovirus* species. These three members (CCHFV, HAZV and TFLV) form the CCHFV serogroup, one of the nine serogroups of the *Orthonairovirus* genus. Another remarkable serogroup of this genus is the Nairobi sheep disease serogroup, formed by Nairobi sheep disease virus (NSDV), Dugbe virus (DUGV) and Kupe virus (KUPV). NSDV causes severe disease in sheep and goats in East Africa and India with a case fatality rate of around 70% [29]. DUGV affects cattle in Nigeria, but the pathogenesis of KUPV in mammals is unknown [30].

1.1.3 Transmission

CCHFV circulates in nature between tick and vertebrate hosts (Figure 1.2). Several tick species of the family *Ixodidae*, characterised by having a hard shield, play a role in

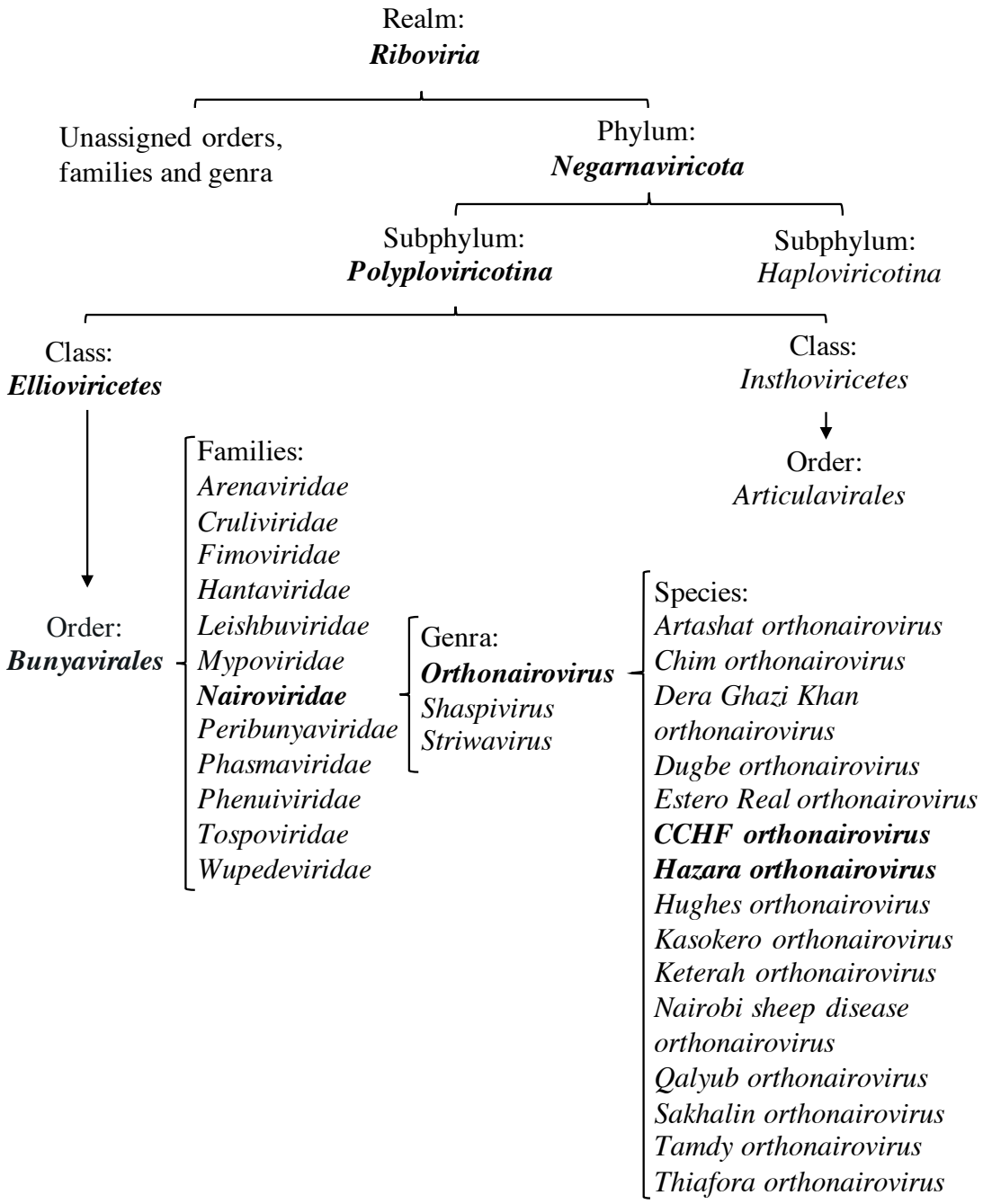


Figure 1.1: Taxonomy of CCHFV and HAZV. CCHFV and HAZV belong to the realm *Riboviria*, phylum *Negarnaviricota*, subphylum *Polyploviricotina*, class *Ellioviricetes*, order *Bunyavirales*, family *Nairoviridae*, genus *Orthonairovirus*, species *CCHFV orthonairovirus* and *Hazara orthonairovirus*, respectively.

the transmission of CCHFV. Members of the *Hyalomma, Rhipicephalus and Dermacentor* ticks species are known to be capable of transmitting the virus, but ticks of the *Hyalomma* genus are the main source of human infection [31], and their global distribution closely matches that of CCHFV incidence. Vertical viremic transmission of CCHFV within tick

populations can occur through trans-stadial, trans-ovarial and venereal transmission [15]. Bursts of amplification normally occur during spring and summer, when ticks take blood meals from small and large vertebrates, and infect them with the virus. Ticks take blood meals once during each developmental stage (larva, nymph and adult), which are required to acquire the nutrients necessary for their maturation and egg production. Non-viremic horizontal way of transmission can also occur between ticks, without the presence of an infected animal host. Tick saliva contains substances that enhance the transmission of the virus to other ticks feeding nearby [32].

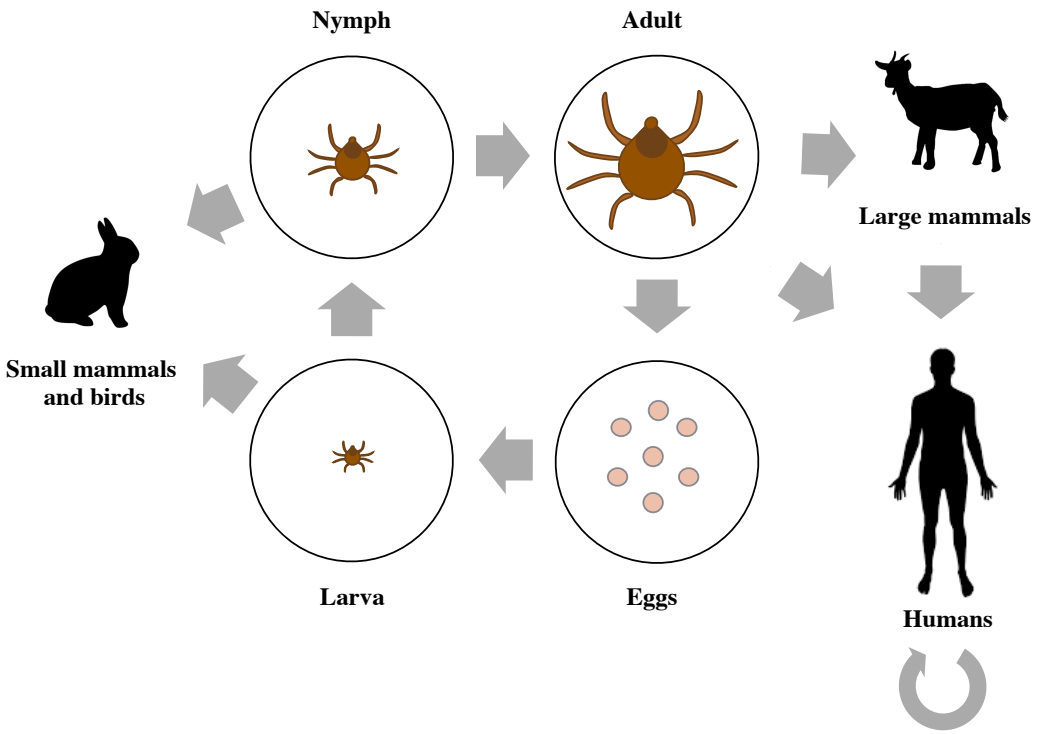


Figure 1.2: Transmission cycle of CCHFV. *Hyalomma* tick species is the natural reservoir of CCHFV. Vertical transmission (trans-stadial, transovarial and venereal) occurs during *Hyalomma* ticks life-cycle. Small and large vertebrates are normally infected by ticks during their blood meals acting as transient hosts of the virus. Humans are considered 'dead-end' hosts of the virus. Human infections can occur by bite of infected ticks or by contact with blood or tissues of viremic patients or animals.

Members of the *Hyalomma* genus normally feed on two or three animal hosts per life cycle. Commonly, larvae and nymphs feed on rodents, hares, ground-feeding birds or other small animals and adults feed on larger mammals such as sheep or cattle [32]. CCHFV can replicate in a wide range of vertebrates including horses, donkeys, sheep, cattle, scrub hares, mice, rats and ostriches [33]. Most vertebrates infected with CCHFV develop only a transient viremia without symptomatic illness which is resolved in 1-2

weeks. Viremic animals serve as transient reservoirs of the virus, infecting naive ticks when they feed on them. But the only true natural reservoirs of the virus are ticks, which remain infected through their entire lives[15].

Humans are infected by bite of infected ticks or by contact with blood or tissues of viremic patients or animals [34]. CCHFV infections are particularly common among professional workers that come into contact with potentially infected animals or human blood on a daily basis, such as farmers, shepherds, veterinarians, abattoir workers, laboratory workers and health workers. High-risk populations also include people exposed to ticks during recreational activities such as hiking or camping [35]. Humans are considered only accidental or 'dead-end' hosts because they are not a source of infection for ticks, the natural reservoir of the virus.

1.1.4 Epidemiology and geographical distribution

The geographic distribution of CCHFV closely matches the range of its *Hyalomma* tick host, which is widespread throughout Africa, Asia, the Middle East, and South-eastern Europe [36]. As such, CCHFV is one of the most widespread tick-borne viruses on earth [21].

After the first description of CCHFV in 1967, initial cases were reported in the former Soviet Union (Crimea, Astrakhan, Rostov, Uzbekistan, Kazakhstan, Tajikistan) and Bulgaria. These were followed by outbreaks in several African countries that included Democratic Republic of the Congo, Uganda, and Mauritania. Presence of CCHFV in Middle East countries was first reported in the 1970s in animal and human serosurveys in Iran. Since then, cases have been reported in Iraq, the United Arab Emirates and Saudi Arabia. CCHFV has also been found in central and eastern Asia where outbreaks have been described in China and India. In the recent years, most CCHFV cases have been reported in Pakistan, Iran, Bulgaria, Turkey and India [21, 15]. The first autochthonous case of CCHFV in Western Europe occurred in Spain, in 2016. A 62-year-old man who had been bitten by a tick died, and nosocomial spread to a nurse was detected [22].

CCHFV is considered an emerging virus due to its recent geographic expansion and increase in the number of countries where it is considered to be endemic. Currently 52 countries are recognised as endemic or potentially endemic regions. In most of these regions, CCHFV human infections have been described. But evidence of circulating CCHFV has also been reported based on the recovery of viral sequences or live virus from ticks or animals, and the detection of CCHFV-specific antibodies in serosurveys [21, 15].

Countries with at least one case of virological or serological evidence of CCHFV are highlighted in Figure 1.3.

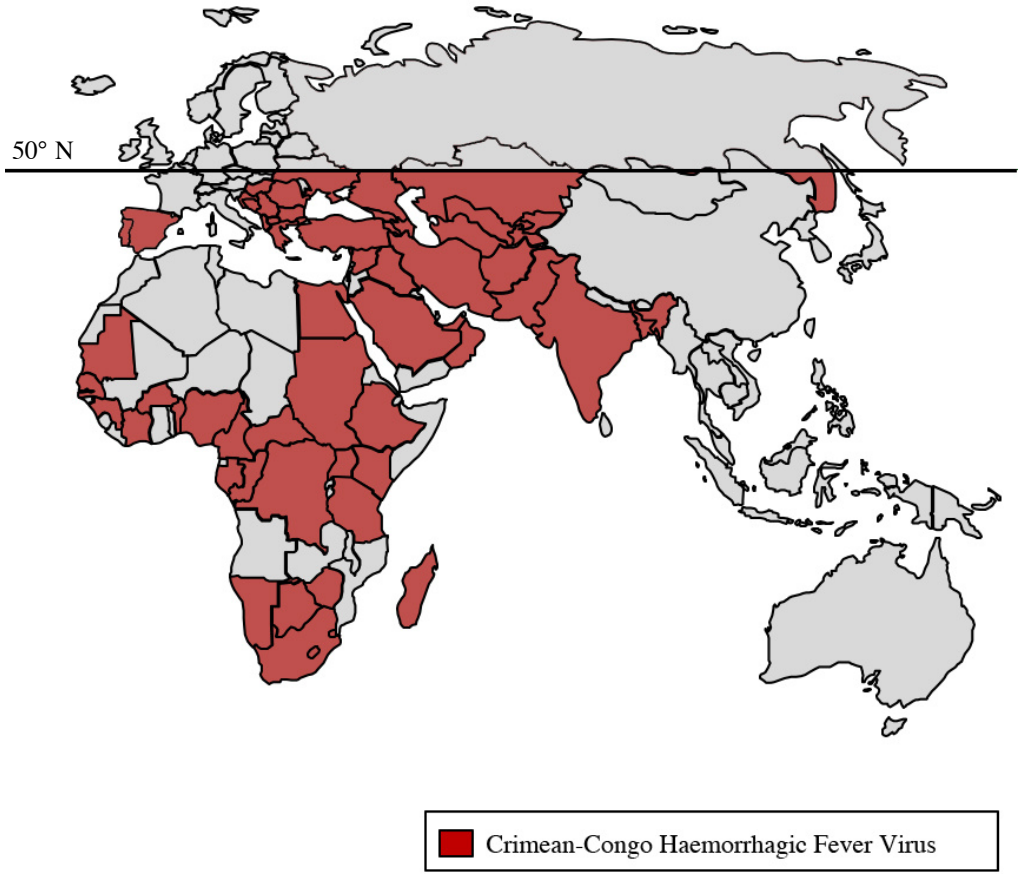


Figure 1.3: Geographical distribution of CCHFV. Countries with at least one case of virological or serological evidence of CCHFV are highlighted in the map. Data obtained from reference [37]. North limit for the geographic distribution of *Hyalomma* species ticks (50 °N) is represented with a line.

Expansion of CCHFV has been attributed to several causes that include human case introduction, infected tick introduction and movement of animals [38]. Person-to-person transmission due to contact with infected body fluids (especially blood, semen and saliva) has been reported in several cases. Nosocomial spread to nurses is often seen when standard barrier nursing methods are not properly executed and cases of sexual transmission have also been documented [39]. Introduction of CCHFV to non-endemic areas has also been reported by travel of infected patients [38].

Importation of ticks into new areas is concerning due to the possibility of establishment of new populations able to sustain the introduction of viruses in a non-endemic region. Movement of infected ticks may also result in transmission of the virus to humans, starting a wave of human-to-human transmission. Dispersion of ticks

is associated with expansion of their host populations and bird migrations. Although experimental infection of birds does not support efficient viral replication (with the exception of ostriches), avian populations seem to play an important role for long-distance movement of ticks [38]. Movement of infected animals is also a promoter for expansion of viral diseases. Lack of clinical signs in CCHFV infected livestock makes it difficult to prevent movement of infected animals between regions, resulting in international movement of the virus through livestock trade [15].

1.1.5 Symptoms and pathogenesis

Most CCHFV human infections are asymptomatic or cause a non-specific mild febrile illness. In some cases, severe progression of the disease is observed, but factors contributing to development of fatal disease are unknown [40]. After brief incubation periods (between 3 and 7 days), CCHFV causes an acute febrile illness accompanied by non-specific symptoms that include myalgia, vomiting and diarrhea. A second phase of the disease is commonly characterised by symptoms that range from petechiae to large areas of ecchymosis and hemorrhage of the gums, nose, internal organs, and gastrointestinal system [41]. This stage is generally short and frequently results in high fatality records with 30-50% mortality rates [15]. Higher fatality rates have been reported in various outbreaks, such as in the United Arab Emirates or China with up to 72% and 80% mortality records, respectively [42, 43].

In severe cases, multi-organ failure, disseminated intravascular coagulation, and circulatory shock result in death of patients. Fatal outcome is associated with high amount of virus in blood, decreased antibody response and platelet counts, elevated liver enzymes, prolonged bleeding times, decreased fibrinogen levels, somnolence, and gastrointestinal bleeding [44, 45, 46, 47]. Certain human leukocyte antigen (HLA) alleles (HLA-A*23) have also been correlated to severity of the disease [47]. In survivor patients, a convalescent period normally begins 10-20 days after onset of illness. Some characteristic symptoms of this stage include feeble pulse, tachycardia, loss of hearing, memory and hair [21].

Very little is known about the pathogenesis of CCHFV, due to the high biosafety required to handle the virus, lack of animal models and occurrence of infections in areas where limited research facilities are available. The main responsible factor for CCHFV pathogenesis is believed to be the interaction between virus and host cells, mainly endothelial cells (ECs) and immune cells [47] (Figure 1.4). CCHFV overcomes the epithelial barrier with help of the tick bite, and is released in ECs inducing endothelial

damage, characterised by up-regulation of intercellular adhesion molecule 1 (ICAM-1), release of interleukin (IL)-6 and IL-8 and increased adhesion of leukocytes [48]. The virus is then amplified by tissue resident macrophages and dendritic cells (DCs), facilitating its transmission to local lymph nodes, spleen, and finally to systemic circulation of the host. CCHFV easily enters the liver due to the fenestrated liver sinusoids and the lack of a basement membrane. Pro-inflamatory cytokines IL-1, IL- 6, IL-8, IL-10 and tumour necrosis factor-α (TNF-α) are released by ECs, macrophages and DCs inducing a strong immune response. Excessive release of these cytokines causes systemic vascular collapse and has been demonstrated to be a prognostic factor for disease severity [49, 47].

Other factors contributing to the pathogenicity of CCHFV include mechanisms for impairment of the innate immune system and delay in the adaptive immune response. CCHFV infection is characterised by partial activation of macrophages and DCs, resulting in a lack of up-regulation of major histocompatibility complex II (MHC II), which is an important partner for the activation of naive T cells [50]. Hyper-activation of monocytes and macrophages is also a common feature upon CCHFV infection, causing cytopenias as a result of an excessive phagocytosis of blood cells (hemophagocytosis) [51]. CCHFV infection delays production of interferons (IFN) and prevents the anti-viral effects of these molecules. Some of these mechanisms include processing the viral 5' RNA triphosphate termini, avoiding RNA helicase (retinoic acid-inducible gene I) RIG-I stimulation and IFN transcription [52], delayed stimulation of IFN regulatory factor-3 (IRF-3) nuclear translocation during late infection causing a delay in the IFN response [53] and the lack of production of significant amounts of the IFN inducer molecule double-stranded RNA [54].

Disregulation of immunological and inflammatory pathways and cells upon CCHFV infection results in disruption of hemostasis [47]. Endothelial damage stimulates platelet aggregation and degranulation, followed by activation of the intrinsic coagulation cascade and disseminated intravascular coagulation. This results in thrombocytopenia, characterised by low platelet counts, and low levels of plasma coagulation factors, which are also impaired due to liver disfunction. Vascular endothelial injury, disseminated intravascular coagulation, thrombocytopenia, liver dysfunction, and diminished levels of coagulation factors are responsible for hemorrhage, hypotension, multiple organ failure, and shock, leading to fatal outcomes in CCHFV infected patients [47].

HAZV has only been reported to replicate in one tick species, *Ixodes redikorzevi*, where it was first isolated [23, 24]. Mammalian natural hosts for HAZV are widely unknown, but antibodies against the virus have been detected in wild rodent sera [55]. Experimental infections with HAZV have demonstrated its ability to replicate in a wide

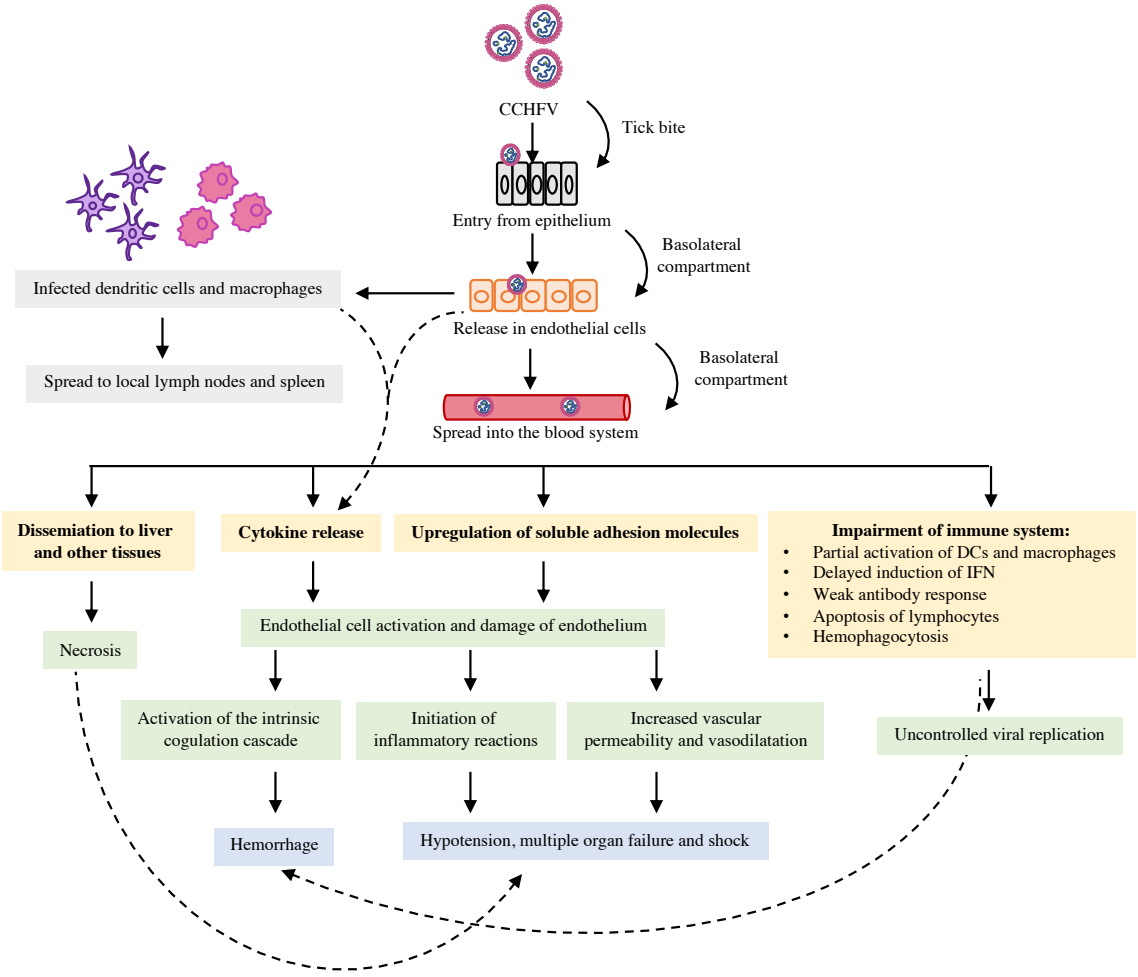


Figure 1.4: Pathogenesis of CCHFV. CCHFV overcomes the epithelial barrier with the help of the tick bite, and is released in endothelial cells inducing endothelial damage. The virus is then amplified by tissue resident macrophages and dendritic cells, facilitating its transmission to local lymph nodes, spleen, and finally to systemic circulation of the host. Dissemination of virus, excessive cytokine release, up-regulation of soluble adhesion molecules and impairment of immune system result in hypotension, multiple organ failure, shock and hemorrhage.

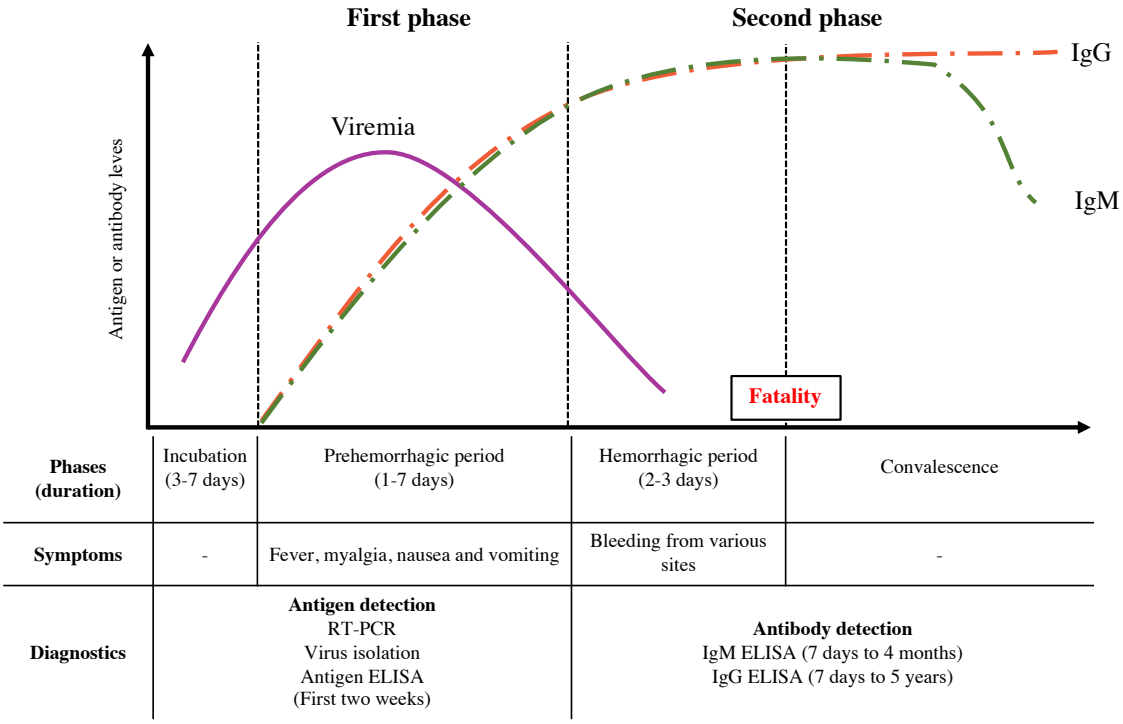
range of vertebrate hosts including mice, rats, hamsters, guinea pigs, rabbits, donkeys, sheep and calves [56]. As for CCHFV, HAZV infected animals are asymptomatic.

1.1.6 Diagnostics

CCHFV disease in humans is characterised by an early febrile phase with potential to proceed further, causing severe hemorrhagic fever. Differential diagnosis during the early disease phase is difficult due to associated non-specific febrile symptoms. After the incubation phase, a first viremic phase is characterised by low CCHFV-specific antibody levels and high virus titres (Figure 1.5). Current assays used for detection of CCHFV during the viremic phase involve virus isolation, which must be performed under biosafety level 4 (BSL-4) laboratory conditions, with reverse transcription polymerase chain reaction (RT-PCR) detection representing the current gold standard, offering sensitive and highly specific results [57]. Different RT-PCRs have been designed for CCHFV detection, most of them targeting the small (S) segment of the virus. Some RT-PCR assays are able to detect strains from all the different clades of the virus [58], and quantitative RT-PCR (RT-qPCR) assays enable the measurement of viral load, and thus the prediction of disease severity and likelihood of death [59]. An antigen-detection enzyme-linked immunosorbent assay (ELISA) test (VectoCrimea-CHF-antigen, VectorBest) is also commercially available to detect CCHFV antigens during the viremic phase.

During latter stages of the disease, hemorrhagic period and convalescence, virus titres drop and high levels of immunoglobulin G (IgG) and immunoglobulin M (IgM) antibodies are developed (Figure 1.5). Most available serologic methods for detection of CCHFV during the later phase of the disease are ELISA tests based on the detection of CCHFV-specific IgG or IgM [60, 61, 62, 63, 64, 65], which are also commonly used in animal serosurveys [64, 66]. But in patients exhibiting severe disease, production of antibodies is normally delayed, low, or even absent [67], making viral RNA (vRNA) detection by RT-PCR the most reliable method of diagnosis in human cases.

Fatal outcome is associated with higher viral loads [44, 45, 46] and late diagnosis of patients decreases treatment efficacy and increases the risk of fatal outcome and nosocomial spread [68], making quick and accurate detection of antigen important for disease management. In addition, rapid detection of CCHFV antigen in remote areas or in low-resource settings is an urgent and unmet need. For these reasons, the development of a commercially available, bio-safe, rapid point-of-care test is considered as urgent need for research by the World Health Organisation (WHO) [35] and this objective has been addressed in this project.



Phases (duration)	Incubation (3-7 days)	Prehemorrhagic period (1-7 days)	Hemorrhagic period (2-3 days)	Convalescence
Symptoms	-	Fever, myalgia, nausea and vomiting	Bleeding from various sites	-
Diagnostics		**Antigen detection** RT-PCR Virus isolation Antigen ELISA (First two weeks)	**Antibody detection** IgM ELISA (7 days to 4 months) IgG ELISA (7 days to 5 years)	

Figure 1.5: Current diagnostic methods for CCHFV infected patients. Schematic representation of phases and symptoms of CCHFV patients. Standard diagnostic methods for each phase are described.

1.1.7 Treatment

CCHF disease management has proven difficult and there are no broadly licensed therapeutics or vaccines. The current approach used to treat CCHF is based on general supportive measures. Monitoring the hematological and coagulation status of patients is crucial, including potential bleeding foci, fluid and electrolyte balance and blood count. Replacement of cells and factors can be carried out if needed, including the administration of thrombocytes, fresh frozen plasma, and sometimes erythrocyte preparations [69].

The drug ribavirin, which is a synthetic purine nucleoside analogue, has been employed to treat CCHF. It inhibits the replication of a wide range of RNA and DNA viruses *in vitro*, but its mechanism of action remains unclear. The efficacy of ribavirin for the treatment of CCHF is controversial and placebo-controlled studies are difficult due to ethical considerations, preventing definitive conclusions on the efficacy of ribavirin in patients with CCHF [40]. Nonetheless, ribavirin treatment is considered to be beneficial when administered during the early phase of the disease and is recommended as post-exposure prophylaxis for healthcare workers dealing with CCHFV patients [70]. When administered to CCHFV infected IFNAR knock-out mice, ribavirin prolonged the time until death, but did not prevent it [71]. Another drug that has been explored for CCHFV

treatment is favipiravir, an RNA polymerase inhibitor, which has shown efficacy in suppressing viral replication and preventing mortality following CCHFV infection in mouse models [71]. The use of monoclonal antibodies capable of neutralising CCHFV infection is also considered as a potential alternative treatment for CCHF [40].

1.1.8 Vaccine development

In the 1970s the first CCHFV vaccine was developed in the Soviet Union, a chloroform-inactivated mouse brain derived vaccine. It was licensed in Bulgaria and has been administered to military personnel including medical and agricultural workers since 1974. The current vaccine consists of CCHFV strain V42/81, and was isolated from a patient in 1981 [72]. Evidence of its protective efficacy is based on the reduction in CCHFV cases since 1974. CCHF has never been observed in vaccinated individuals but the vaccine is unlikely to gain widespread international regulatory approval due to its crude preparation [73].

In recent years, several approaches have been studied for CCHFV vaccine development. Most vaccines are based on the CCHFV nucleoprotein (NP) and glycoproteins: the NP is highly conserved in different strains of the virus and plays an essential role in viral replication, whereas the glycoproteins (G_N and G_C) are more diverse but contain neutralising epitopes that have the potential to generate immune protection against the virus.

Different approaches used for vaccine development include CCHFV inactivated virus [74], virus vectors (modified vaccinia Ankara vaccines [75], herpesvirus [76], adenovirus [77, 78] and vesicular stomatitis virus-based vaccines [79]), DNA and messenger RNA (mRNA) vaccines [80, 81, 82, 83], transgenic plants [84], protein-based vaccines [85, 86] and virus-like particles [87]. Many of these vaccines have been proven successful in the induction of an immune response and protection from lethal disease in mouse models. But the absence of larger animal models of CCHF has hindered further evaluation of the efficacy of these vaccines, and controlled human studies have not been reported [73]. An alternative animal anti-tick vaccine might also reduce the risk of zoonotic transmission, and the pool of ticks carrying CCHFV [73]. Animal anti-tick vaccines are thus potentially valuable, and may also be effective against CCHFV spread [35].

1.2 Molecular biology of CCHFV and HAZV

1.2.1 Virion and genome structure

CCHFV and HAZV are enveloped, broadly-spherical viruses with a diameter of 80-100 nm [88, 89, 90]. Electron microscopy (EM) studies show variation in virion shape and dimensions, presenting pleomorphic morphologies that include oval and elongated particles [90]. The surface glycoproteins (G_N and G_C) appear as a continuous layer of protrusions of approximately 10 nm in length from the lipid viral envelope.

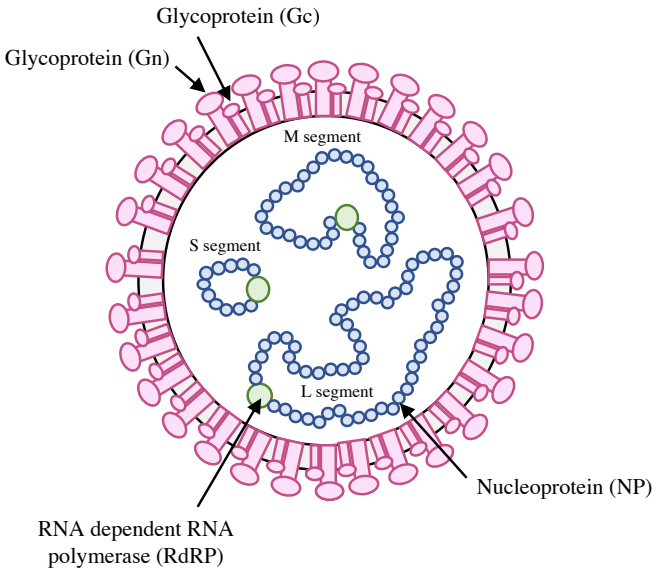


Figure 1.6: Schematic of a nairoviral virion. Nairoviruses are enveloped spherical viruses. The three negative single-stranded segments (S, M and L) are associated with the nucleoprotein (NP) and the RNA dependent RNA polymerase (RdRp) forming ribonucleoprotein (RNP) structures, packaged inside the viral lipid envelope into which the glycoproteins (G_N and G_C) are inserted.

Each infectious virion contains three segments of negative sense RNA named small (S), medium (M) and large (L), which respectively encode the NP, that encapsidates each RNA segment; the glycoprotein precursor (GPC), which is post-translationally cleaved to yield envelope glycoproteins G_N and G_C; and the L protein that contains both RNA dependent RNA polymerase (RdRp) and ovarian tumour (OTU)-like modules with the latter performing immune modulatory roles. The RdRp, together with NP, associates with the genome segments or vRNA forming ribonucleoproteins (RNPs) [91] that act as templates for all viral RNA synthesis activities. It is assumed that each infectious particle must contain at least one copy of each segment in the form of RNP (Figure 1.6). Packing of RNPs inside the virion is thought to be relatively unorganised, suggesting that virion morphology is more likely to be dependent on interactions between the G_N and G_C

heterodimer and the NP component of the RNPs [92].

In common with other members of the *Bunyavirales* order, the three genomic RNA segments of orthonairoviruses contain complementary untranslated regions (UTRs) at the 5' and 3' termini that surround an open reading frame (ORF) (Figure 1.7a) [92]. The complementarity of these terminal regions is believed to contribute to the circular appearance of the bunyavirus genomes as has been evidenced for Uukuniemi phlebovirus (UUKV) and Bunyamwera virus (BUNV) by direct EM observation. However, this has not been yet proven for orthonairoviruses, for which the native structure of viral RNPs (vRNPs) has been poorly studied. Sequence alignment of orthonairoviral 5′ and 3′ UTRs reveals a common arrangement with terminal-proximal regions of high sequence conservation, followed by terminal-distal segment sequences that are more variable. In particular, the nine nucleotides at the 5' and 3' UTR termini are invariant across all segments of all species (Figure 1.7b) as well as being perfectly complementary. Functional analyses of the entire UTRs have revealed the existence of two promoter elements (PEs) separated by a variable spacer region, and detailed mutagenesis of these regions using both mini-genomes and infectious virus has delineated nucleotides that build up the promoters for viral RNA synthesis [93, 94].

1.2.2 Replication cycle

Despite the characterisation of CCHFV replication cycle being challenging due to its high containment requirements, its molecular biology has been better characterised than for HAZV. Due to their structural and genomic analogies, the replication cycle of HAZV is expected to be analogous to that of CCHFV (Figure 1.8).

1.2.2.1 Virus entry and uncoating

CCHFV entry is mediated by the interaction of the glycoproteins with cellular receptors. It is predicted to be predominantly mediated by G_C, which contains a fusion peptide involved in driving membrane fusion [95], and is a target for neutralising monoclonal antibodies [96]. A putative cellular receptor for CCHFV is nucleolin, which has been shown to interact with G_C and be essential for virus entry [97]. Internalisation of CCHFV occurs through clathrin-mediated endocytosis, and is dependent on low pH and cholesterol [98]. HAZV entry has also been demonstrated to be facilitated by cellular cholesterol [99] and K^+ [90].

Following entry, CCHFV particles are transported to early endosomes and

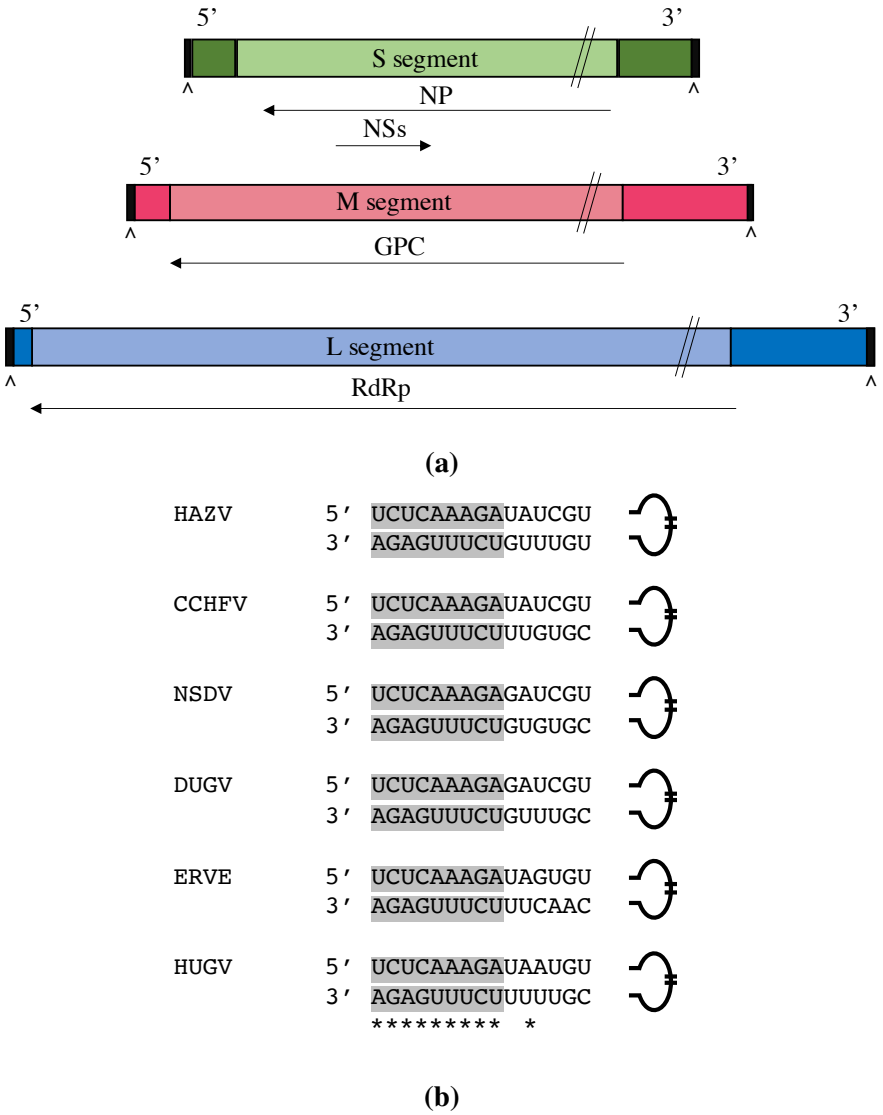


Figure 1.7: Schematic representation of nairovirus genomic segments and consensus terminal sequences. **(a)** Nairoviruses contain three segments of negative sense ssRNA named small (S), medium (M) and large (L), which respectively encode the nucleoprotein (NP), the glycoprotein precursor (GPC) and the RNA dependent RNA polymerase (RdRp). An additional ambisense open reading frame has been described for the CCHFV S segment, encoding the non-structural S protein (NS_S). The three segments contain two untranslated regions (UTRs, represented in darker colours) in the 3' and 5' ends, flanking the respective open reading frames. The 9 terminal nucleotides of the UTRs are highly conserved among the three segments (represented in black, highlighted with circumflexes). **(b)** Representation of the terminal regions of the UTRs from the S segments of different members of the nairovirus genus. The 9 terminal nucleotides (highlighted in grey) are highly conserved, and form canonical Watson-Crick base pairing (highlighted with asterisks), forming an open circular panhandle structure.

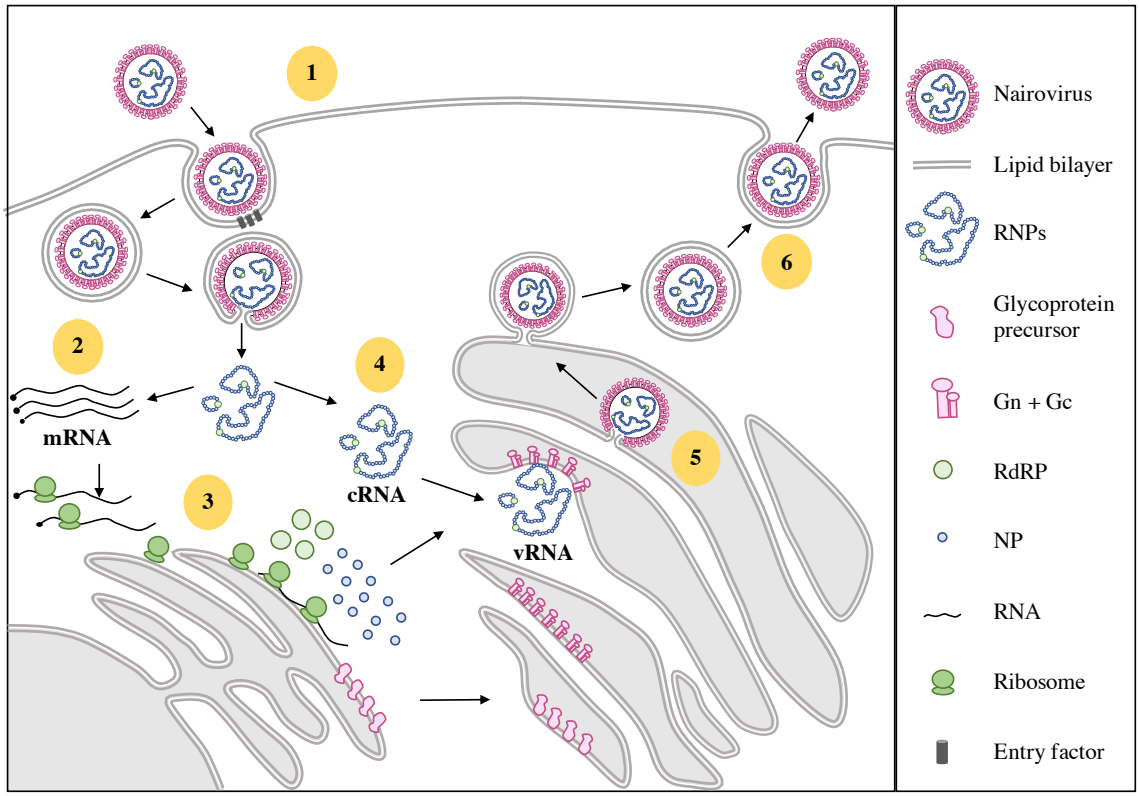


Figure 1.8: Nairovirus life cycle. Schematic representation of the different stages of the nairovirus cycle: attachment and entry (1), transcription (2), translation (3), replication (4), assembly (5) and egress (6).

subsequently to multivesicular bodies (MVB) in a process that is dependent on Rab5 and endosomal sorting complex required for transport (ESCRT) regulators [100, 101]. Blockade of Rab7-dependent trafficking has no effect on CCHFV infection, indicating that the trafficking from early to late endosomes is not important for CCHFV replication cycle, and that the MVBs are likely the main organelle where the CCHFV envelope fuses with the host membranes and releases the RNPs into the cytoplasm. Acidification of early endosomes is thought to induce changes on the conformation of CCHFV G_N and G_C, allowing them to fuse with the cellular membranes [100].

Intact, dynamic microtubules are required for CCHFV internalisation and transport of virion components to sites of replication [102]. Upon infection, CCHFV NP is directed to a perinuclear region that is thought to be the site of CCHFV replication and virus factory formation. Actin filaments have been shown to be involved in trafficking of the NP to these regions [103].

1.2.2.2 RNA transcription and replication

Following fusion with cellular membranes, CCHFV RNPs are released into the cytoplasm and the encapsidated vRNA serves as template for the synthesis of mRNA and the positive sense complementary RNA (cRNA). Although there are no direct studies on CCHFV and HAZV RNA transcription and replication, studies with other nairoviruses, such as DUGV, and other members of the *Bunyavirales* order serve as representative models of this processes.

The first RNA synthesis event is the primer-dependent transcription of a single mRNA from each of the three vRNA segments (Figure 1.9). Primers comprise 7-methylguanylate (m^7G) capped oligoribonucleotides, cleaved from cellular mRNAs [104], generated by an L protein resident endonuclease domain in a process commonly referred to as a ”cap-snatching”. These primers are on average 10 nucleotides in length and present partial base pairing with the 3’ end of the vRNA template. Primary transcription is carried out by the template associated RdRp to yield 5’ capped mRNA [92].

Initiation of RNA synthesis only occurs once per segment, in contrast to non-segmented negative sense RNA viruses where, following transcription of a gene, the polymerase reinitiates transcription at regions that flank each gene, known as gene junctions. This is thought to be due to the fact that the 3’ and 5’ sequences are required for a functional promoter in the bunyavirales genomic segments [105]. A prime-and-realign transcription initiation process has been proposed for bunyaviruses, which

involves annealing the capped primer to an internal region of the 3'UTR containing a tri-nucleotide repeat sequence (3'- AUCAUCAUC-5') [106]. After alignment and extension by one or a few nucleotides, the nascent chain realigns backwards with the 3' proximal triplet repeat on the template RNA. This process can occur several times before processive elongation takes place, leading to a longer complementary primer forming a more stable transcription complex.

Termination of mRNA transcription is thought to occur before the polymerase reaches the 5' terminus of the genomic template, generating truncated mRNAs at their 3' ends relative to the genome template [107]. Highly conserved regions in the 5' UTRs of S and L bunyaviral segments have been characterised as transcription termination signals responsible for this truncation. In the case of the M segments, such a sequence has not been found raising the possibility that mRNA synthesis is terminated via a different mechanism [107].

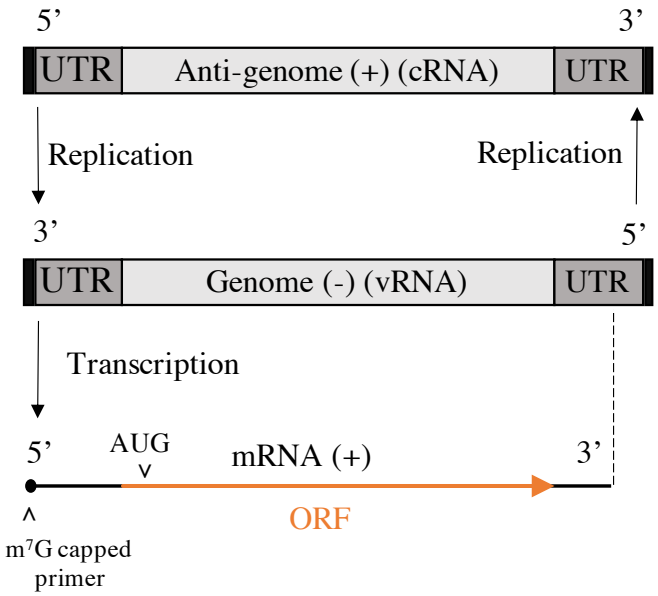


Figure 1.9: Schematic representation of negative sense gene expression strategy. Negative sense genomic RNA (vRNA) acts as a template for two distinct RNA synthesis activities carried out by the viral polymerase: transcription and replication. The viral polymerase utilises a cellular m^7G capped primer for transcription of a positive sense mRNA containing the corresponding open reading frame (ORF) under a start codon (AUG). Replication is primer independent and results in generation of a full-length complementary copy of the viral genome, a positive sense anti-genomic RNA (cRNA), which serves as template for generation of new genomic viral RNA copies (vRNA).

Bunyaviral mRNAs lack a 3' poly-A tail, a common feature of eukaryotic mRNAs that greatly enhances translation efficiency. In the case of BUNV, complex stem-loop secondary structure elements at the 3' mRNA end have been suggested to functionally

replace the poly-A tail in a process where translation is enhanced, independently of poly-A binding protein (PABP) [108]. This stem-loop displays both structural and sequence similarities to the stem–loop present at the 3' ends of histone mRNAs, which are also poly-A-deficient, raising the possibility that BUNV may have hijacked a cellular mechanism for poly-A-independent mRNA translation [92].

After primary transcription of viral mRNAs, the polymerase switches to vRNA replication, generating full-length complementary copies of the viral genome (cRNA), which are further used as a template to obtain genomic RNPs. The signals that initiate this switch are poorly defined, but it is assumed that cytoplasmic NP abundance is critical in this process as, in contrast to viral mRNA, newly synthesised cRNA is packed into RNP complexes. The mechanism for *de novo* initiation is unknown, but it is thought to require the formation of a duplex between the distal 3' and 5' ends. Initiation could occur by direct positioning of the first 3' nucleotide in the polymerase active site, or by a mechanism of internal initiation followed by realignment, similar to the one previously described for the cap-dependent transcription initiation process [109].

Recent models for bunyaviral replication suggest that both the template and the nascent RNA strands remain protected by NP throughout the course of the entire replication process. During elongation, template RNA progressively dissociates from the proximal NP and is translocated through the polymerase entrance tunnel. The RNA-free NP then translocates to the exit tunnel, and binds the protruding RNA, so that no free-RNA is exposed and no NPs need to be removed or added to the template RNP. The nascent RNA copy is thought to be encapsidated by a free-polymerase and free-NP molecules forming new RNPs [109]. For some members of the *Bunyavirales* order, such as phleboviruses, the cRNA can also act as a template for mRNA transcription of ambisense ORFs, and can be packed into virions [110].

Initiation of the replication process depends on inter-terminal complementarity with no apparent sequence specificity, whereas transcription requires the presence of specific nucleotides located at both ends of the RNA template [111, 112, 105]. Mini-genome studies suggest that sequence-independent structural elements formed by inter-terminal base pairing, and also the specific identity of both paired and unpaired nucleotides play a role in promoter functionality for RNA synthesis [92]. Particular sequences in the UTRs of the S, M and L segments specify the overall promoter strength of each segment. The relative genomic promoter strength of the 3 nairoviral segments has been determined as M>S>L using a HAZV mini-genome system [93].

The sub-cellular localisation of the cRNA and vRNA forms during CCHFV infection

has been studied simultaneously with the localisation of the NP. CCHFV cRNA co-localises to a large extent with the NP in the perinuclear area, while vRNA presents a more random distribution in the cytoplasm [113]. The reason for this differential distribution of transcripts remains unclear, but may be related to their different roles within the CCHFV replication cycle.

1.2.2.3 Virus assembly and egress

Bunyaviruses assemble and bud from perinuclear tubular viral factories that are built around the Golgi complex [114, 115]. These factories are thought to physically combine virus replication and assembly and involve the formation of tubular structures that anchor cell organelles to the Golgi compartment [116, 114]. From these stacks, viruses bud into secretory vesicles and are trafficked towards the plasma membrane.

The structural glycoprotein heterodimer G_N-G_C is synthesised in the endoplasmic reticulum (ER) and subsequently post-transcriptionally processed in the membranes of the Golgi complex. Virion formation is thought to occur when vRNPs traffic to the Golgi membrane and interact with the cytoplasmic tails of the G_N-G_C heterodimer, which protrude into the cytoplasm [92]. A direct interaction between RNPs and the G_N-G_C heterodimer has been proposed for orthonairoviruses, as they do not possess any matrix proteins that can mediate this interaction. Nuclear magnetic resonance (NMR) structural studies of the CCHFV G_N cytoplasmic tail revealed a pair of tightly arranged $\beta\beta\alpha$ zinc fingers with the ability to bind vRNA. This observation suggests a possible direct interaction between the G_N cytoplasmic tail and the RNA component of the RNPs, although protein–protein interactions are also possible between G_N and the NP in the RNPs [92, 117].

Transport of RNPs to the budding compartment requires intact actin filaments and microtubules [102, 103]. Destabilisation of microtubules results in a redistribution of both NP and G_N with a reduction of released virus, whereas stabilisation of microtubules leads to accumulation of progeny virions in the cytoplasm, indicating a role for intact and dynamic microtubules in CCHFV budding and egress [98]. Concentration of the major structural proteins of the virus particle within the budding compartment results in virion assembly and budding into secretory vesicles, in which they are trafficked towards the plasma membrane [92].

1.2.3 The L protein

The nairovirus L segment encodes for a single mRNA transcript of approximately 12 kb, which is translated into the L protein. The L protein of nairoviruses (∼450 kDa) is considerably larger than the L proteins of other members of the *Bunyavirales* order (∼250 kDa). While the nairovirus L protein structure has not been solved to date, cryo-EM structures of other members of the *Bunyavirales* order have been recently elucidated, helping to better understand mechanistic insights into bunyavirus RNA replication. These include near-atomic resolution structures of arenavirus (LASV and Machupo virus [118]), phenuivirus (severe fever with thrombocytopenia syndrome virus [119]) and peribunyavirus (LACV [120]) polymerases.

The most conserved region of the nairovirus L protein is the RdRp module, responsible for its main function to transcribe and replicate genomic RNA segments. Sequence analyses and *in silico* structural elucidation of the CCHFV polymerase module have revealed a typical finger-palm-thumb sequence of subdomains (Figure 1.10). The core catalytic domain of the L polymerase module contains six characteristic motifs common to all negative sense ssRNA virus RdRps: Pre-A/F, A, B, C, D, and E [121]. Motif pre-A/F is present in all RdRp and reverse transcriptases. Motifs A, C and D are predicted to be involved in binding of the nucleoside triphosphates (NTPs), while motifs B and E are predicted to take part in template and/or primer positioning [122].

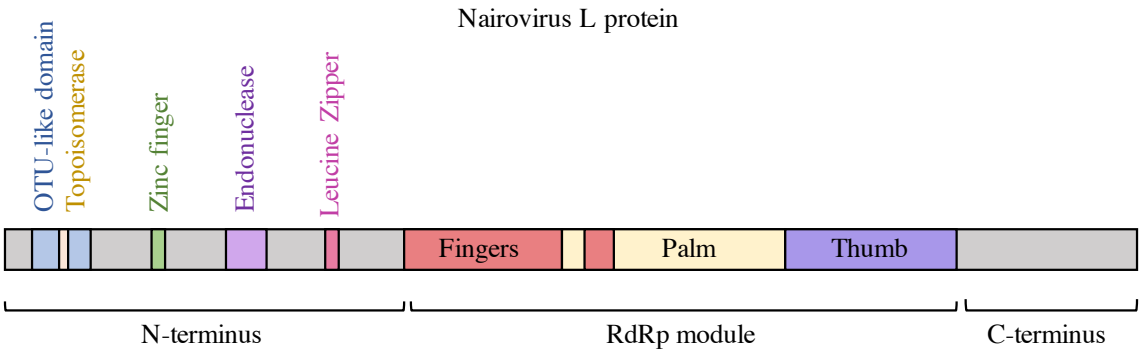


Figure 1.10: Schematic representation of the nairovirus L protein. The N-terminal residues of the nairovirus L proteins contain an ovarian tumour (OTU) cysteine protease, a topoisomerase-like domain, a C2H2 zinc finger, an endonuclease domain (residues 587-895 in CCHFV L protein) and a potential leucine zipper. The RdRp module is formed by the canonical fingers-palm-thumb sequence of subdomains. The function of the C-terminus is unknown and no conserved motifs have been described for this region of the L protein.

The N-terminal residues of the nairovirus L proteins contain domains with non-

classical L functions, such as the OTU cysteine protease, a topoisomerase-like domain potentially involved in RNA structure modification, and a potential leucine zipper and C2H2 zinc finger motifs important for binding of NP [123, 124]. The L protein OTU domain presents both a deubiquitinase and deISGylase activity for the removal of ubiquitin (Ub) and/or IFN-stimulated gene 15 (ISG15) moieties. These activities participate in the suppression of the innate immune responses during viral infection, antagonising the innate antiviral response. The deISGylase protease activity suppresses the antiviral effects of ISG15, mediated through its ability to conjugate to target proteins. The deubiquitinase activity interferes with immune pathways that are regulated by ubiquitination [125]. The crystal structure of the CCHFV OTU domain bound to Ub and ISG15 moieties demonstrates the capability of the OTU binding site to recognise these two different substrates in a similar manner [126]. An additional auto-proteolytic function of the viral OTU protease has been proposed for processing of the L protein. However, a mini-genome study has shown that the OTU protease activity is dispensable for virus RNA replication, suggesting that this domain is not required for the processing of the L protein to generate an efficient RdRp [127].

A potential cap-snatching endonuclease domain has been suggested for CCHFV L protein between the viral OTU and RdRp domains, by structural and sequence analogy with other bunyaviral nucleases. Binding of metal ions and a known endonuclease inhibitor, 2,4-dioxo-4-phenylbutanoic acid (DPBA), support the hypothesis for this endonuclease site [128]. The position of this domain has been functionally proven using a virus-like particle (VLP) system [129]. The endonuclease domain is predicted to be located around amino acid D693, and the alanine-substitution mutant at this position lacks the ability to transcribe mRNA but retains the ability to replicate the mini-genome and assemble nucleocapsids that are packaged into virions. This residue is predicted to form part of the endonuclease active centre and complex the two Mn^{2+} ions that are necessary for endonucleolytic cleavage of host mRNAs [129]. As for other members of the *Bunyavirales* order, nairovirus L proteins contain a significant C-terminal region after the RdRp module with an unknown function. No cross-genera-conserved motifs have been described for this region of the L protein [104].

1.2.4 The glycoprotein precursor (GPC)

The nairovirus M segment encodes for a single gene product of approximately 5kb, which encodes for a polyprotein denominated glycoprotein precursor (GPC). Intracellular processing of the CCHFV GPC has been widely characterised, but the HAZV GPC shows

remarkable differences at the amino acid level, and thus its processing might be different. The CCHFV GPC is co- and post-translationally modified resulting in the mature viral glycoproteins and other non-structural protein products (Figure 1.11). The CCHFV GPC contains an N-terminal signal peptide directing its synthesis to the secretory pathway [130]. Its synthesis starts in the ER, where the signal peptide is removed, the GPC is N-glycosylated and the transmembrane domains span the ER membrane five times. Co-translational cleavage by cellular signalases and intra-membrane cleaving proteases (I-CLiPs) results in the 140-kDa PreG$_N$ (formed by the mucin-like domain (MLD), GP38, and G$_N$), the NS$_M$, and the 85-kDa PreG$_C$ [131, 132].

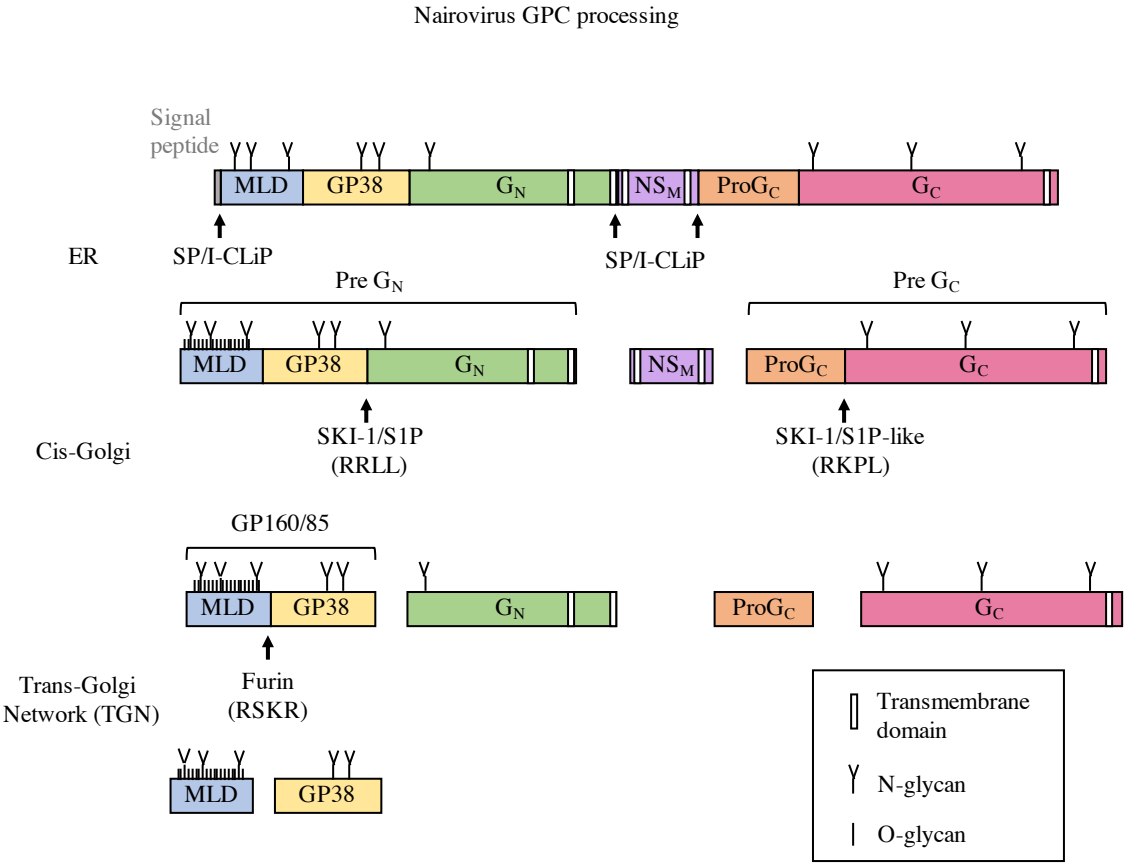


Figure 1.11: Schematic representation of the nairovirus GPC polyprotein processing. The GPC polyprotein is synthesised in the ER where N-glycosylation occurs. Host signal peptidases (SP) and/or intramembrane cleaving proteases (I-CLiPs) cleave GPC resulting in the 140-kDa PreG$_N$ (formed by the mucin-like domain (MLD), GP38, and G$_N$), the NS$_M$, and the 85-kDa PreG$_C$. These proteins traffic to the cis-Golgi where the MLD is O-glycosylated. PreG$_N$ is cleaved at the RRLL motif by subtilisin kexin isozyme-1/site-1 protease (SKI-1/S1P) leading to GP160/85, which is later cleaved at the RSKR motif by furin in the trans-Golgi network (TGN) into MLD and GP38. PreG$_C$ is cleaved at the RKPL motif by a protease with similar specificity to SKI-1/S1P (SKI-1/S1P-like) yielding the structural G$_C$ and the ProG$_C$ non-structural protein product.

$PreG_N$ and $PreG_C$ are then transported to the Golgi complex where $preG_N$ mucin-like domain is O-glycosylated. The interaction of $PreG_N$ with $PreG_C$ is essential for $PreG_C$ to travel from ER to Golgi. The ectodomains of G_N and G_C appear to be sufficient for heterodimer formation and transport to the Golgi [96, 133]. $PreG_N$ is further cleaved at a conserved RRLL motif by a host subtilisin kexin isozyme-1/site-1 protease (SKI-1/S1P) to separate MLD/GP38 of either 160 or 85 kDa from G_N. Cells deficient in SKI-1/S1P are susceptible to infection with CCHFV but infectious particles are not released by these cells, suggesting the cleavage is essential for infectious virus release [132]. The structural G_C glycoprotein and the $ProG_C$ non-structural protein product are cleaved from $PreG_C$ early in the secretory pathway at the conserved RKPL motif, which closely resembles the motif recognised by SKI-1/S1P, but the cellular protease involved in this cleavage remains unidentified [132].

The MLD/GP38 product is further cleaved in the trans-Golgi network (TGN) at a conserved RSKR motif by a furin-like protease in the late secretory pathway, resulting in production of GP38 glycoprotein (38 kDa) and MLD [134]. The function of GP38 remains unknown. Its crystal structure has been recently elucidated, revealing distant homology to the ectodomain of G_N, suggestive of a gene duplication event. The only protective antibody for CCHFV in an adult mouse model reported to date binds GP38 with high affinity, suggesting an important antigenic role for GP38 [135]. The biological function of MLD, GP85, GP160 and the NS_M protein products also remains unknown. Transiently expressed NS_M in cells localises to the Golgi complex, suggesting its possible contribution to retention of the glycoproteins to this cellular compartment [131].

1.2.5 The nucleoprotein (NP) and the non-structural S protein (NS_S)

The S segment of nairoviruses involves overlapping coding regions in opposite directions, and thus is considered to be an ambisense segment. The NP gene is encoded in a negative sense, and the non-structural S protein (NS_S) is encoded in a positive sense, overlapping with NP. Very little is known about the function of NS_S in nairoviruses. Co-localisation studies show that CCHFV NS_S is present in the mitochondria and over-expression of NS_S in mammalian cells induces apoptosis by disrupting the mitochondrial membrane potential, suggesting a possible role in regulation of apoptosis during CCHFV infection [136].

The NP of nairoviruses is significantly larger than the NP of most other *Bunyavirales* order members, apart from members of the *Hantaviridae* family. NP is the most abundant protein in virions, and functions primarily to encapsidate the vRNA and cRNA forming

RNP complexes together with the L protein. Consistent with its structural role, CCHFV NP has been shown to interact with the N- and C-terminal regions of the L protein and to co-localise with tagged L protein in transfected cells [124]. vRNPs are used as a template for all viral RNA synthesis activities. By encapsidating vRNA and cRNA, the NP protects the viral genome from degradation, compacts the RNA and helps to avoid the formation of RNA secondary structures that could potentially form between mRNAs and the corresponding templates [137].

Several cellular partners of CCHFV NP have been described to date including actin filaments, which target NP to perinuclear regions during infection [103], and other cytoskeletal components such as tubulin and vimentin [138]. Interaction of CCHFV NP with the IFN-induced myxovirus resistance protein 1 (MxA) has been also demonstrated by co-immunoprecipitation and co-localisation assays. MxA is thought to play an anti-viral role in CCHFV infection, binding and sequestering NP away from sites of active virus replication. Over-expression of MxA in mammalian cells results in reduction of viral replication [139]. The most abundant interacting partners of both HAZV NP and CCHFV NP are members of the 70 kilodalton heat shock proteins (HSP70) family of adenosine triphosphate (ATP)-dependent cellular chaperones. Reduction of active HSP70 levels in cells by small-molecule inhibitors impairs HAZV replication, demonstrating a beneficial role for HSP70 in the nairovirus replication cycle [138].

1.2.5.1 The structure of CCHFV NP

The crystal structure of CCHFV NP has been determined using the sequence of viral isolates from Iraq [140] (strain Baghdad-12) and China [141, 142] (strains YL04057 and IbAr10200) at 2.1, 2.3 and 3.1 Å resolution, respectively. CCHFV NP presents a structure predominantly built of α-helices with two major domains: a globular core and an extended arm.

The globular domain comprises 23 α-helices with an overall structure similar to the NP of LASV, member of the *Arenaviridae* family. This similarity is supported by the close phylogenetic relationship between the *Arenaviridae* and *Nairoviridae* families, demonstrated by phylogenetic analysis of L and NP proteins of members of the *Bunyavirales* order [140], and also by the fact that arenaviruses and nairoviruses share aspects of their cellular biology that are unique within this order, such as their dependence on cellular SKI-1/S1P protease to process their glycoprotein precursors [143].

The core of the CCHFV NP globular domain is comprised by C-terminal α-helices, which are surrounded by α-helices from the N-terminus (Figures 1.12a, 1.12b). The

core domain presents endonuclease activity for single stranded and double stranded DNA (ss/dsDNA) but not for single stranded or double stranded RNA (ss/dsRNA) [141]. Consistent with the structural similarities between LASV NP and CCHFV NP, the endonuclease activity has also been described for LASV NP, although for dsRNA rather than dsDNA, and it has been proposed this activity antagonises the host cell antiviral response [144]. The ss/dsDNA endonuclease activity of CCHFV NP is somewhat surprising and its role within the replication cycle of CCHFV remains unknown.

The arm domain of CCHFV NP contains two long α-helices that are extended away from the core, with an exposed loop at the apex supported by a small three-helix bundle [140]. The arm domain is highly flexible and presents different positions in the different crystal structures described, with a rotation of up to 120°. The arm is connected to the body domain by a disordered linker region (residues 183-191), for which electron density has been only described in one of all the chains available for CCHFV NP structures [145], indicating a high flexibility in this region.

A caspase 3 recognition site has been described at the apex of the arm domain [145], formed by a DEVD motif at residues 269 to 272 [140]. The cleavage of this recognition site occurs at the late stages of the CCHFV infectious cycle, when the apoptotic pathway is induced, but the functional relevance of this cleavage is currently unknown [146]. Interestingly, the cleavage by caspase 3 *in vitro* does not result in the dissociation of the N- and C-terminal fragments, which remain associated probably by close contact of helical structural elements within the core domain [140].

Although no crystal structure of a nairovirus NP in complex with RNA has been solved to date, the structures of the different oligomeric states formed after crystallisation, as well as the structures obtained from different viral strains, show a conformational shift of the arm domain [147] (Figure 1.13). The flexibility of this domain suggests a possible gating mechanism for RNA binding, similar to the one described for LASV NP [140]. This mechanistic model proposes a controlled binding of the NP to RNA by conformational changes. In the closed RNA-free form of the NP structure, the RNA-binding cleft is not available to accept ssRNA. This could be the conformational state of the RNA-free NP immediately after its synthesis in the host cells, which would prevent it from binding host RNAs. To bind the viral genome, the C-terminal domain must shift away from the RNA-binding cleft to allow RNA to enter and form the RNPs. The factor that triggers this conformational change is still unknown [148].

The electrostatic surface potential of CCHFV NP reveals a continuous positively charged region likely to be the binding site for RNA. Part of the positively charged region

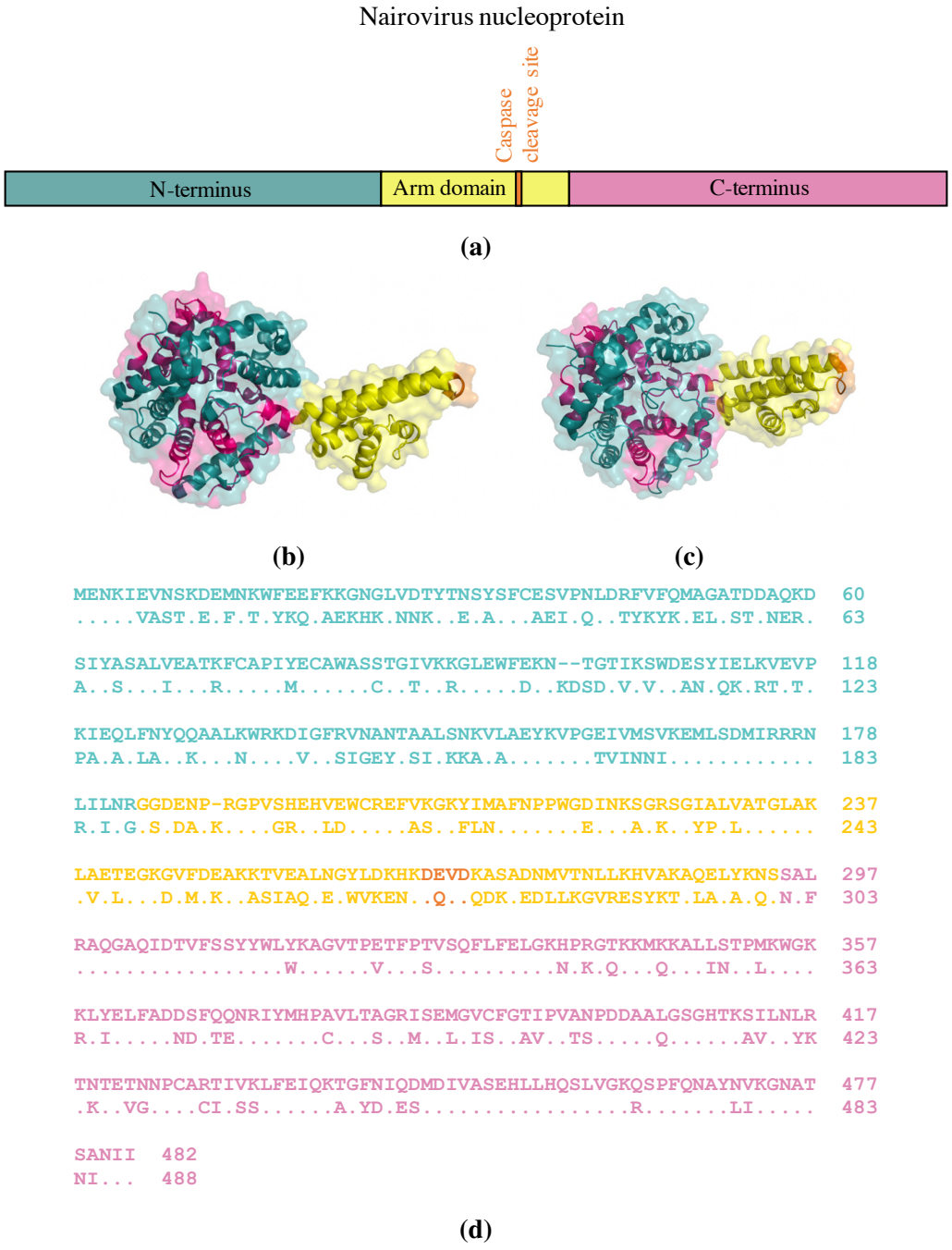


Figure 1.12: The structure of CCHFV and HAZV NP. **(a)** Schematic representation of the nairovirus nucleoprotein sequence. The N-terminus (cyan) and C-terminus (magenta) of the sequence form the globular domain. The central sequence corresponds to the arm domain (yellow) and contains a conserved caspase cleavage site (orange). **(b)** Crystal structure of CCHFV NP (PDB: 4AKL, N-terminus: residues 1-183, central sequence: residues 184-294, C-terminus: residues 295-482). **(c)** Crystal structure of HAZV NP (PDB: 4XZE, N-terminus: residues 1-188, central sequence: residues 189-300, C-terminus: residues 301-488). **(d)** Sequence alignment of CCHFV (query) and HAZV (subject) NPs. The C-terminus core domain of the NPs is more conserved (73% identity) than the solvent exposed N-terminus and arm domains (54% and 50% identity, respectively). Images created using PyMOL.

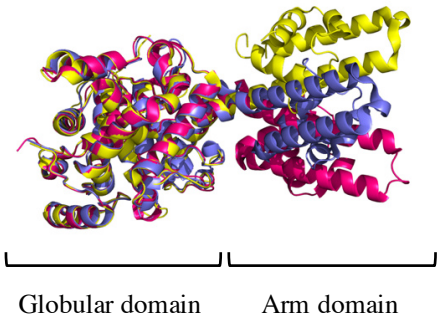
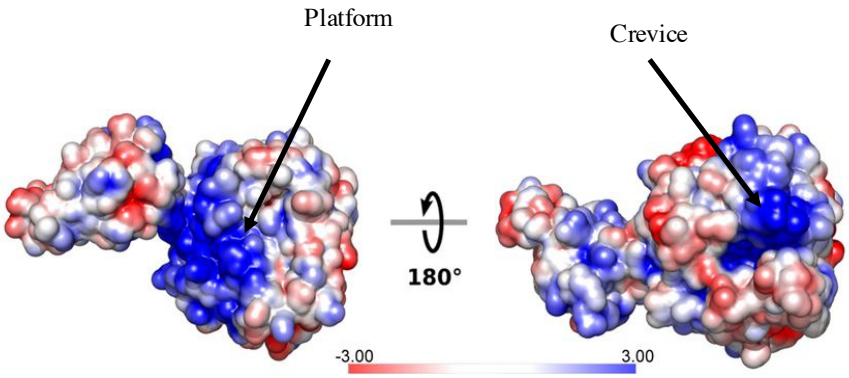


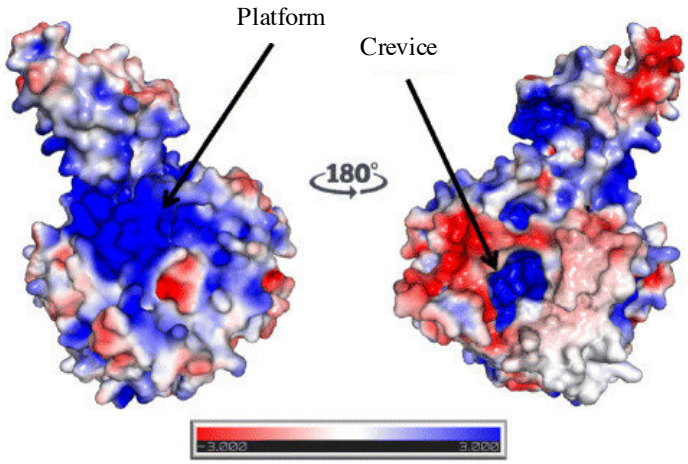
Figure 1.13: Alignment of CCHFV NP crystal structures. Structural alignment of the globular domains of CCHFV NP illustrates the variability of position of the arm domain in respect to the globular body which does not alter conformation. Baghdad-12 (purple, PDB: 4AKL), IbAr10200 (yellow, PDB: 3U3I) and YL04057 (pink, PDB: 4AQF). Images created using PyMOL.

covers the surface of CCHFV NP adjacent to the arm domain and is termed the platform. A second positively charged region, termed the crevice, forms a channel between the base of the globular domain and the beginning of the arm domain [140] (Figure 1.14a). The entire positively charged region comprises residues R45, K72, S149, H197, K222, R225, K237, Q303, K336, R339, K342, K343, K345, R372, and K462 [149]. Residues K132, Q300, and K411 have been shown to be essential for mini-genome replication, and are also likely to play a role in RNA binding due to their charge characteristics and specific location in the NP surface [140]. Two different modes of RNA-binding have been described for CCHFV NP *in vitro*: a higher affinity binding mode ($K_D \approx 50$ nM) for RNA molecules containing the S segment dsRNA panhandle structure and a lower affinity binding mode ($K_D \approx$ 120-150 nM) for ssRNA [150].

Structural analyses of CCHFV NP suggest its oligomerisation plays a role in its RNA binding function. CCHFV NP appears as a monomer in the absence of RNA, but different oligomeric forms of purified recombinant CCHFV NP have been described in samples with high OD_{260}/OD_{280} ratios, suggesting these oligomers are bound to nucleic acid (presumably RNA) of the expression host. The main oligomeric species observed by EM in these samples is the pentameric form, but tetramers, hexamers and heptamers have also been observed [149]. Crystallisation of these oligomeric forms potentially bound to RNA has not been achieved but crystallisation of monomeric CCHFV NP results in higher ordered structures that could just be the result of crystallographic packing artefacts, but may also represent a physiologically relevant organisation of CCHFV NP [145]. The interaction between monomers occurs in a head-tail manner and involves residues 320-354 within the head domain of one molecule and residues 210-219 and 260-272 of the



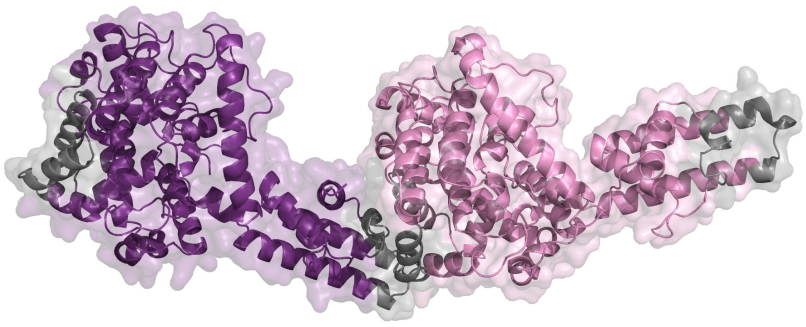
(a)



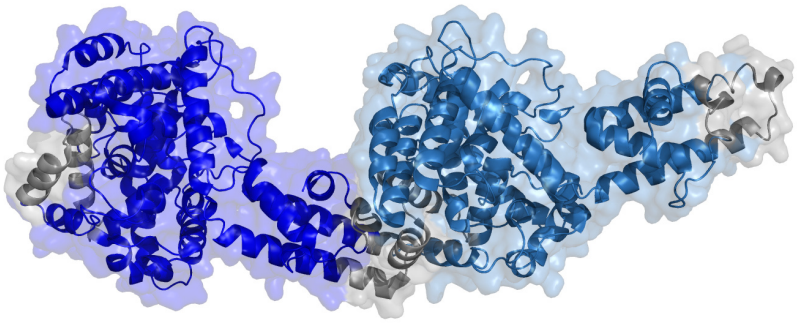
(b)

Figure 1.14: Surface electrostatic potential of CCHFV NP and HAZV NP. Analysis of the electrostatic surface potential of CCHFV NP (solvent-accessible molecular surface, **(a)**) and HAZV NP (straight-up molecular surface, **(b)**) reveals areas of positive charge (blue) and negative charge (red), and two possible RNA binding regions termed the platform and the crevice. Taken from [140] and [151], respectively.

arm of the adjacent molecule [145] (Figure 1.15a).



(a)



(b)

Figure 1.15: Head-tail interactions between CCHFV NP or HAZV NP monomers. The crystal structures of CCHFV NP **(a)** and HAZV NP **(b)** show head-tail interactions between the head domain of one molecule and the arm domain of the adjacent molecule. Head and tail residues involved in the interaction are highlighted in grey. Images created using PyMOL.

There is no high-resolution structure of native RNPs from any of the members of the *Bunyavirales* order. Native CCHFV RNPs have been vaguely observed by EM, showing a flexible and disordered architecture [149]. They present a width that suggests a wound helical structure rather than a ”beads-on-a-string” conformation, but further research needs to be done to determine the structural and oligomeric state of CCHFV NP in biological conditions. Examination of the native form of nairoviral RNPs has been addressed in this project.

1.2.5.2 The structure of HAZV NP

HAZV NP has a similar overall structure to CCHFV NP, with their globular domains superposing with an RMSD = 0.70 Å and a 60% sequence identity. The HAZV NP structure consists of a globular domain formed mostly of α-helices derived from both the N- and C-termini, and an arm domain comprising two long α-helices [151] (Figure 1.12c). Residues 187–196 are missing in the electron density and are thought to form a disordered loop that links the arm domain to the N terminus of the globular domain, in a similar manner that occurs for the equivalent loop in CCHFV NP [151].

A DQVD caspase cleavage site positioned in the apex of the arm domain is also present in HAZV NP, resembling the DEVD motif in CCHFV NP. Cleavage of HAZV NP by caspase 3 has been demonstrated in mammalian cells, where HAZV infection induces apoptosis and HAZV NP caspase cleavage [152]. Interestingly, HAZV infection in tick cells fails to induce apoptosis, and caspase cleavage of HAZV NP is undetectable, suggesting cleavage of the NP is not essential for completion of the virus infectious cycle [152].

Analysis of the electrostatic surface potential of HAZV NP reveals a similar positively charged region to CCHFV NP, also formed by a platform and a crevice, potentially involved in its RNA binding function (Figure 1.14b). The oligomeric forms in the crystal lattice also suggest a head-tail interaction between monomers analogous to the model proposed for CCHFV NP [151] (Figure 1.15b).

1.2.6 Reverse genetics systems

Reverse genetics systems can be defined as tools for the generation and subsequent replication and transcription of full-length virus RNA genomes or truncated viral genome analogues from complementary DNA (cDNA) [153]. Different types of reverse genetics systems have been developed to elucidate mechanisms of transcription, replication, and particle assembly of CCHFV and HAZV, contributing considerably to improve our understanding of their biology.

Full-length infectious clone systems allow for the generation of recombinant viruses, whereas life cycle modelling systems such as mini-genomes and VLPs allow the study of discrete parts of the virus cycle, including transcription, replication, translation, particle budding and entry. Reverse genetics systems can be used to generate reporter viruses, perform mutational analyses for the study of anti-viral drug efficacy or resistance, and elucidate essential catalytic or structural residues. Table 1.1 summarises the currently

available reverse genetics systems for CCHFV and HAZV, further described in this section.

		cDNA	Helper proteins	Modelling capacities	BSL
CCHFV	**Mini-genome [154, 140, 127]**	Truncated genome	NP and L	Genome replication and secondary transcription	BSL-1/ BSL-2
	VLPs [129, 155]	Truncated genome	NP, GPC and L	Particle budding, release, attachment, entry, (primary transcription [155]), secondary transcription and replication	BSL-1/ BSL-2
	Infectious clone [134]	Full-length genome	NP and L	All steps of viral life cycle: genome replication, secondary transcription, morphogenesis, budding, entry, primary transcription	BSL-4
HAZV	**Mini-genome [93]**	Truncated genome	NP and L	Genome replication and secondary transcription	BSL-1/ BSL-2
	Infectious clone [156]	Full-length genome	-	All steps of viral life cycle: genome replication, secondary transcription, morphogenesis, budding, entry, primary transcription	BSL-1/ BSL-2

Table 1.1: Reverse genetics systems available for CCHFV and HAZV. Complementary DNA (cDNA) used, helper proteins provided in trans, modelling capacities and biosafety level (BSL) facility requirements for each system are indicated. NP: nucleoprotein, GPC: glycoprotein precursor, L: L protein, VLP: virus-like particle.

1.2.6.1 CCHFV and HAZV mini-genome systems

CCHFV and HAZV mini-genome systems [154, 140, 127, 93] are based on the use of a truncated form of vRNA containing the viral UTRs of a genomic segment in which the coding region has been replaced with a reporter gene [91]. cDNA of the mini-genome sequences are cloned into expression vectors under the control of the cellular polymerase I (pol I) [154] or the bacteriophage T7 DNA dependent RNA polymerase (T7 pol) promoter [140, 127, 93]. If the system is under control of a T7 promoter, T7 pol can be expressed by co-transfection of a T7 pol-encoding plasmid or by use of a stable T7 pol expression cell line. T7 pol-based expression systems commonly include a hepatitis delta virus ribozyme (HDV RZ) in the 3' end of the RNA to generate a native 3' terminus [153].

The NP and L proteins need to be provided either from expression plasmids or by superinfection with virus. The pol I or T7 pol transcribe the mini-genome RNA, which is encapsidated by the NP and the L protein into virus-like RNPs. These RNPs serve as template for transcription and replication of the mini-genome RNA, resulting in the production of mRNA and complementary RNPs (cRNPs). Translation of mRNA

ultimately results in the generation of a reporter signal, and cRNPs are used for further rounds of replication. The reporter signal provides a quantitative measure of genome transcription and replication. Since mini-genome systems are non-infectious and do not use full viral genomes, they can be handled under low biosafety level containments (BSL-1 or BSL-2, depending on local regulations) [153].

1.2.6.2 CCHFV VLP systems

CCHFV VLP systems have been developed to overcome some of the limitations of mini-genome systems [129, 155] for which analysis is restricted to the formation and activities of the RNPs. In contrast, VLP systems also allow interrogation of factors involved in VLP formation, egress and subsequent entry into new cells. VLPs can mimic a single cycle of infection but cannot establish a multi-cycle infection, making it possible to work with these systems under low biosafety level containments (BSL-1 or BSL-2, depending on local regulations) [153]. VLPs are generated by co-expressing three helper plasmids encoding the NP, GPC, and L proteins together with a mini-genome plasmid, resulting in the incorporation of mini-genome RNPs into VLPs. Incubation of VLPs with target cells results in entry and delivery of mini-genome RNPs. The RNP complexes serve as templates for replication and transcription, which can be supported by transient expression of NP and L helper plasmids, ultimately resulting in reporter activity in the target cells. Reporter signal in the VLP producing cells reflects replication and transcription, whereas reporter signal in the target cells is dependent on efficient budding and release of VLPs by the producer cells and the subsequent entry, transcription and replication of the VLPs in target cells.

Two different systems of CCHFV VLPs have been developed: one of them requires transfection of the target cells with NP and L helper plasmids prior to incubation with the VLPs [129], and the second one is able to generate a robust reporter signal in the target cells without the need to previously express L and NP [155]. The main difference between these two systems is that the first one does not reflect primary transcription, which by definition is performed only by RNP components brought into target cells within virus particles, whereas the second one does.

1.2.6.3 CCHFV and HAZV infectious clone systems

CCHFV and HAZV infectious clone systems have been generated from DNA plasmids containing the complete viral genome sequences under the control of a T7 pol

promoter [134, 156]. Each of the three plasmids is transcribed by the T7 pol producing cRNAs. The CCHFV rescue system utilises helper plasmids designed to produce NP and L proteins, which are provided in trans to encapsidate the cRNAs [134]. In the case of the HAZV rescue system, cRNAs could directly be translated into proteins that encapsidate the cRNAs [156]. The encapsidated cRNPs are used as a replication template for the production of vRNA and viral mRNA. All viral components ultimately assemble into particles resulting in a clonal population of infectious viruses capable of infecting neighbouring cells and performing multiple infection cycles. Consequently, the CCHFV infectious clone system must be used under BSL-4 facilities, whereas the HAZV system can be handled using BSL-2 protocols [134, 156].

1.3 Protein-based scaffolds for high-affinity molecular recognition

Molecules with high affinity and specificity for particular molecular targets are widely used as tools for basic research, diagnostics and therapeutics. Molecular recognition refers to the specific non-covalent interaction between two or more molecules through non-covalent bonding such as hydrogen bonding, metal coordination, hydrophobic forces, van der Waals forces, π-π interactions, halogen bonding, or electrostatic interactions [157]. Different types of molecules, from metal ions and small chemicals to large proteins and higher order protein complexes, whole cells, viruses, or parasites can be involved in molecular recognition interactions.

For many years antibodies have been the binding reagent of choice due to their high affinity binding capacity and their ability to be generated naturally by immunisation of animals with the desired target. However, over the last decades several classes of nucleic acid and small protein-based scaffolds have been developed to artificially generate specific binders that function with micromolar to sub-picomolar dissociation constants, with the aim to overcome some of the limitations of full-size antibodies.

Nucleic acid-based scaffolds, also known as aptamers, are 20-60 nucleotides long ssRNA or ssDNA molecules folded into complex three-dimensional shapes, forming binding pockets and clefts for the specific recognition of very diverse targets. Aptamer selection is based on the process of systematic evolution of ligands by exponential enrichment (SELEX). SELEX refers to a process of selection of ssDNAs or ssRNAs by repeated rounds of binding, partitioning, and amplification of combinatorial oligonucleotide libraries [158]. Some of the main advantages of aptamers include their

low cost of production, small size, high-stability, high variety of targets and their potential chemical modification [159].

Protein-based scaffolds include full-size antibodies and a wide variety of artificial scaffolds that have been developed using antibodies and other proteins present in nature as reference templates or by *de novo* computational design. Protein-based artificial scaffolds can be divided into two different groups depending on the origin of their structure: antibody fragments and non-antibody protein scaffolds. The different types of antibodies and protein-based artificial scaffolds available for molecular recognition are reviewed in this section.

1.3.1 Antibody isotypes

Antibodies are the secreted form of immunoglobulins (Ig), which are essential molecules for the animal adaptive immune response and are expressed only in jawed vertebrates, including fish, amphibians, reptiles, birds, and mammals. Antibodies are Y-shaped heterodimeric glycosylated proteins composed of two heavy (H) and two light (L) chains [160]. In a small number of species, such as sharks and camels, antibodies consisting only of heavy chains have been identified. Both heavy and light chains contain an N-terminal variable (V) and one or more C-terminal constant (C) domains of 12-13 kDa, with variable regions forming the antigen binding sites. Light chains contain one single constant domain, whereas heavy chains contain 3 or 4 of such domains [160].

Ig-antigen interactions typically take place between the paratope, defined as the site on the Ig at which the antigen binds, and the epitope, which is the region of the antigen that is bound. Individual determinants, termed idiotypes, are contained within the V domains [161]. Common determinants specific for the constant portion of the antibody are termed isotypes and allow grouping of Igs into recognised classes. These are responsible for effector functions through interactions with a variety of receptors expressed on immune cells [162]. There are 5 isotypes of soluble Igs or antibodies: IgG, IgM, IgA, IgD and IgE (Figure 1.16).

IgG is the predominant Ig isotype in sera, accounting for 75% of serum Igs, and presents the longest half-life. There are 4 different classes of IgG (IgG1, IgG2, IgG3 and IgG4) based on structural, antigenic and functional differences in the constant region of their heavy chains [163]. IgG1 and IgG3 antibodies are generally induced in response to protein antigens whereas IgG2 and IgG4 are associated with polysaccharide antigens. All of them are expressed in a monomeric form and participate in the secondary immune

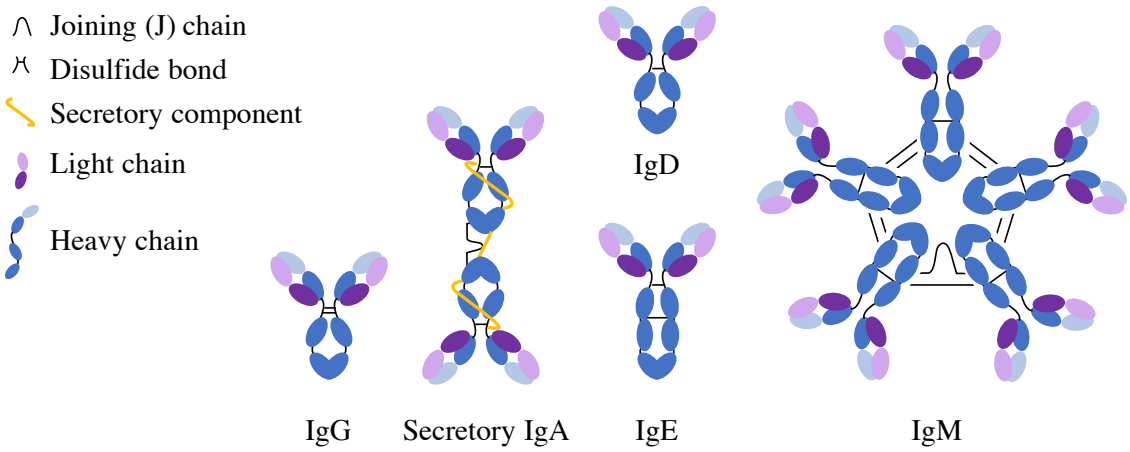


Figure 1.16: Immunoglobulin isotypes. Schematic representation of the five immunoglobulin classes or isotypes in mammals. Mature IgG, IgE and IgA are monomeric, whereas secretory IgA appears as a dimer associated with a joining (J) chain by disulphide bonds on the CH_3 domains and the secretory component, a polypeptide chain bonded to the CH_2 domain. Upon maturation and antigenic stimulation, IgM units link to each other by disulfide bonds in the constant regions forming pentameric or hexameric multimers that are secreted.

response, including neutralisation of toxins and viruses. All subtypes but IgG4 contribute to complement fixation [160].

IgGs can be purified from sera by affinity chromatography using bacterial protein A or protein G. Protein A binds to human IgG1, IgG2 and IgG4 subtypes, whereas protein G binds to the four subtypes. Due to the abundance of IgG in sera and its capacity to modulate the immune response, it is the most commonly used Ig isotype for research, therapeutics and diagnostic applications. All current marketed monoclonal antibodies are of the IgG classes, mainly human IgG1, which binds to both activating and inhibitory Fcγ receptors (FcγRs), and is capable of eliciting pro-inflammatory cytokine production, antibody-dependent cell-mediated cytotoxicity, antibody-dependent cellular phagocytosis, and complement-dependent cytotoxicity [164].

1.3.2 Antibody fragments

Full-size antibodies can only be naturally produced against immunogenic antigens. Their multi-domain nature and the presence of several disulphide bonds in their structure make them relatively unstable and challenging to express and purify in quality. Their unique structural features limit recombinant antibody expression to eukaryotic cell lines in which the proper folding, cysteine oxidation, and post-translational glycosylation can be achieved [165]. Furthermore, the large size of antibodies makes them difficult to

genetically or chemically modify. In clinical applications, the large antibody dimensions interfere with penetration into certain tissues and the high-cost of production makes monoclonal antibody therapies unaccessible for many patients. In the case of diagnostic applications, the constant domain can cause significant nonspecific binding interfering with detection of the analyte of interest.

Multiple functions of antibodies are not always necessary, such as their ability to recruit immune system cells, their bivalency or their high plasma half-life. The fragment crystallizable (Fc) region of antibodies may be advantageous for many therapeutic applications where an immune response needs to be induced, but it can also have unfavourable effects. Such is the case of the antibody-dependent enhancement induced upon infection by some viruses such as Dengue virus and ZIKV [166, 167].

However, the remarkable ability of antibodies to recognise antigens with high affinity and specificity must be preserved in most applications. To overcome full-size antibody disadvantages, most of them related to their structural properties and size, antibody fragments and antibody mimetics have been designed. The first approach used to obtain smaller-size molecules with antibody-binding properties was the cleavage of full-size antibodies with proteases. This was followed by genetic engineering approaches that allowed the production and optimisation of mono or multivalent antibody fragments [168]. Some examples of commonly used antibody fragments are represented in Figure 1.17 and briefly described below.

Fab, Fab', (Fab')$_2$, and Fv. Cleavage of Igs with the protease papain results into two antigen-binding fragments (Fab) fragments, each of which can bind an antigen, and a single Fc fragment. The Fab contains one complete L chain and the V and CH_1 portion of one H chain, and can be further divided into a variable fragment (Fv) composed of the VH and VL domains, and a constant fragment (Fb) composed of the CL and CH_1 domains. Pepsin splits IgG into an Fc fragment and a single dimeric F(ab')$_2$ with two antigen binding regions linked by disulfide bonds. Reduction of F(ab')$_2$ fragments produces 2 monovalent Fab' fragments, which contain a free sulfhydryl group that can be used for conjugation to other molecules [168].

scFv, diabody, minibody and dAb. Genetic engineering methods allow the production of single chain variable fragments (scFv), which are Fv type fragments linked by an engineered flexible linker peptide [169]. Depending on the linker length, these molecules can be associated with a second scFv generating bivalent diabodies, or with more scFv molecules into triabodies or tetrabodies. Multivalent scFvs present an increased binding affinity to their target antigens as a result of having two more

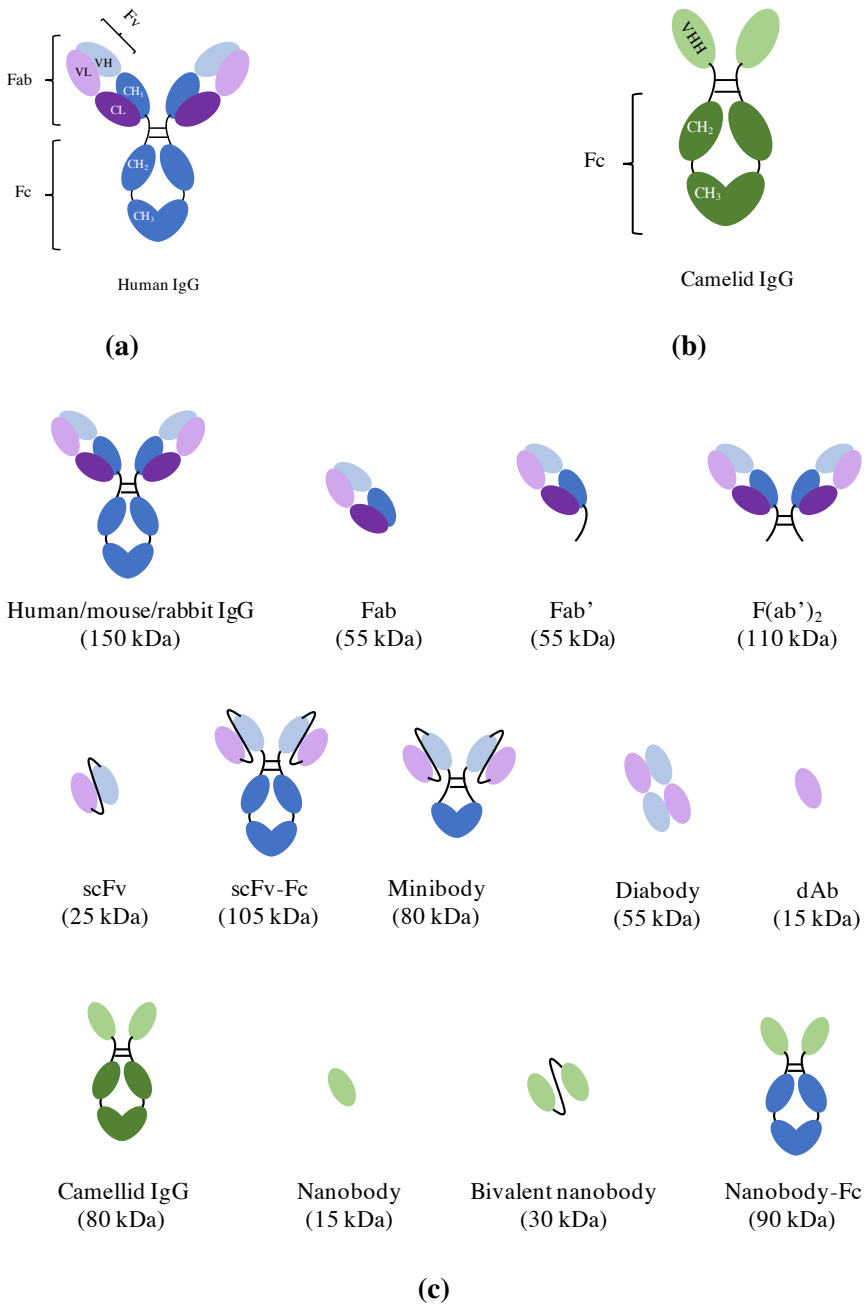


Figure 1.17: Human and camelid IgGs and antibody fragments. **(a)** Schematic representation of a human IgG. Light chain is represented in purple and heavy chain in blue. The different subunits of the IgG are indicated: Fab, Fc, Fv, Variable (VL, light purple) and constant (CL, dark purple) regions of light chains and variable (VH, light blue) and constant (CH_1, CH_2, CH_3 and CH_4, dark blue) regions of heavy chains. **(b)** Schematic representation of a camelid IgG formed only by heavy chains. The different subunits of the IgG are indicated: constant region subunits (CH_2, CH_3) form Fc and are represented in dark green, and variable heavy homodimers (VHH) are represented in light green. **(c)** Schematic representation of some of the most commonly used antibody fragments and their corresponding size (kDa). Colour code is the same as in a) and b).

target antigen binding sites, reducing the off-rate of the antibody fragment. Small scFv fragments with two different variable domains can be generated to produce bispecific bis-scFv fragments capable of binding two different epitopes concurrently [170]. scFvs can be stabilised by an Fc region, forming scFv-Fc, or bound to CH_3 regions forming scFv-CH_3 fusion proteins, known as minibodies, that assemble into bivalent dimers. Domain antibodies (dAb) are fully human unpaired variable domains (either VH or VL), which have been engineered to prevent dimerisation whilst maintaining the specificity and affinity of a canonical antigen binding site [168].

Camelid and shark antibodies and nanobodies. Camelid and shark (*squalidae*) species contain antibodies exclusively composed by H chain homodimers, lacking L chains. The Fab portions of these antibodies, called variable heavy homodimers (VHH) in camelids and variable new antigen receptor (VNAR) in sharks, are the smallest antigen-binding regions naturally found. Nanobodies are VHH-derived recombinant domains able to bind antigens, are very stable and can be easily produced in huge quantity by using common simple protein expression systems such as bacteria. Nanobodies can be engineered to form bivalent nanobodies and nanobody-human Fc chimeras [170].

Recombinant antibody fragments are small and generally non-glycosylated, allowing their production in bacterial recombinant systems, which is faster and cheaper than for full-size antibodies. Because of their small size, antibody fragments have the ability to penetrate tissues, access challenging epitopes, and have potentially reduced immunogenicity, which are important characteristics for therapeutic applications [171]. *In vitro* molecular evolution and selection techniques such as 'phage and ribosome display combined with recombinant expression systems are commonly used for optimisation of antibody properties. Modulation of binding characteristics, such as affinity and specificity maturation of the variable regions, and solubility improvement of the constant domains, has lead to antibody fragments with optimal properties for specific molecular recognition applications [172].

1.3.3 Non-immunoglobulin protein-based scaffolds

With the rise of molecular engineering techniques, the concept of non-antibody protein scaffolds was rapidly developed. New binding protein scaffolds are constantly being created and commercialised using naturally occurring proteins or protein domains as reference for their rational design. These molecules are commonly referred to as antibody mimetics due to their ability to bind to antigens with affinities similar to antibodies, but have no structural relation to them. Their simple structures consist of α-

helices, β-sheets, or random coils that can bind to specific targets. They are derived from synthetic libraries that allow identification of binding reagents without the use of animals. High yields can be produced in recombinant expression systems, and they present smaller size (normally between 6 and 20 kDa) and higher stability than antibodies. Other advantageous characteristics of antibody mimetics include higher solubility, resistance to proteases, or stability toward higher temperatures and extreme pH [173].

Selection of antibody mimetics is normally performed using library screenings. Libraries are designed by employing *in silico* methods and constructed via molecular biology protocols. Site-directed or random mutagenesis strategies are commonly used to mutate the residues involved in ligand binding and generate large libraries [173]. Once the library is designed and constructed, the translated protein scaffolds are screened for isolation of mutants with the desired properties. Some of the most commonly employed display systems include 'phage display, ribosome display, mRNA display, yeast display, and bacterial cell-surface display [174]. The purpose of these display platforms is to establish a physical linkage between phenotype and genotype [174].

The *in vitro* nature of antibody mimetic design, selection, amplification and purification presents several advantages compared to antibodies. *In vitro* design and selection bypasses the experimental use of animals, and permits selection of binders against targets which would normally be challenging for conventional antibodies due to their small size, toxicity or low immunogenicity. The process of selection typically takes only a few weeks, whereas the development of antibodies is a time-consuming process that normally takes several months. Protein display library screening technologies are substantially less expensive, consume less material than *in vivo* development strategies and thus enable low cost production with high yields and reproducible batch-to-batch preparations [165].

The first antibody mimetic obtained from a combinatorial library was described in 1998 [175]. Thioredoxin 1 (TrxA), a modified *E. coli* enzyme, was used as scaffold for the selection of affinity proteins binding cyclin-dependent protein kinase 2 (CDK-2) with nanomolar affinities. Since then, numerous studies have reported the use of thioredoxin-based antibody mimetics and, in the past two decades, the number of novel scaffolds has been constantly growing. Some of the most commonly used protein scaffolds are annotated in table 1.2 and their corresponding structures are presented in Figure 1.18.

Antibody mimetics have been exploited in biomedical research and biotechnology applications, including affinity chromatography techniques, flow cytometry or western blotting [176]. Their easy functionalisation and engineering permits the conjugation to

fluorophores or enzymes which provide single step detection in ELISAs, dot-blots, or immunohistochemical analyses, which normally involve two consequential steps if using antibodies. Their conjugation to fluorophores has also promoted their use in the field of imaging. Non-antibody reagents are small and thus penetrate tissues and can access epitopes in densely packed sub-cellular structures of cells more readily than antibodies, and can place the fluorophore closer to the target of interest, providing an increased spatial resolution in 'super-resolution' approaches [177].

Due to the high affinity that these molecules can present, they can also serve as functional inhibitors or antagonists for biochemical research or in cell biology and have been highly exploited in the area of therapeutics, some of them currently being tested in clinical trials [173]. In some cases, these scaffolds are derived from human proteins, and thus posses very low immunogenic potential. The therapeutic action of these small molecules can be mediated by antagonising receptors by inhibition of their ligand binding sites, binding to ligands to prevent their interaction with cognate receptors, or neutralising toxins or other harmful or infectious agents. They can also be chemically conjugated to drugs for their *in vivo* delivery and they can be easily combined with other scaffolds to generate bispecific or oligomeric reagents [174].

ID	Scaffold	Parental protein	Randomisation	Size	Selection method	Company/References
1	Affibody	Z-domain of staphylococcal protein A (bacterial)	Helices	6 kDa	'Phage display, ribosome display	Affibody AB [178]
2	Affilin	γ-B-crystallin or ubiquitin (human)	Beta strand	8.5 kDa, 20 kDa	'Phage display, ribosome display	SCIL Proteins [179, 180]
3	Affimer	Type I: human protease inhibitor Stefin A. Type II: Consensus plant phytocystatin A	Loop	11 kDa	'Phage display	Avacta [181]
4	Alphabody	computer-based rational design	Helix	10 kDa	'Phage display	Complix [182]
5	Anticalin	Human lipocallin	Loop and beta strand	20 kDa	'Phage display	Pieris AG [183]
6	Atrimer	C-type lectin (tetranectin, human)	Loop	60-70 kDa	'Phage display	Anaphore [184]
7	Avimer	A domains of various membrane receptors (Human/artificial)	Loop	9–18 kDa	'Phage display	Amgen [185]
8	Beta-hairpin mimetics	Diverse (Structural motif)	Loop and beta strand	1.7 kDa	'Phage display	[186]
9	Bicycle peptide	Synthetic peptide	Loop	1.5-2 kDa	'Phage display	Bicycle Therapeutics [187]
10	Cys-knots	Diverse (Structural motif)	Loop	Variable	'Phage display	[188]
11	DARpin	Ankyrin repeats (Human)	Helix, beta-turn	14-18 kDa	Ribosome display	[189]
12	FN3 scaffold (Monobodies/ Adnectins, centyrins, pronectins, Tn3)	Fibronectine type III (human)	Loop and beta strand	10-20 kDa	'Phage display, yeast display	Protelica [190]
13	Fynomer	SH3 domain fyn kinase (human)	Loops	7 kDa	'Phage display, DNA display	Covagen [191]
14	Knottin	Diverse (structural motif)	Loops	3-5 kDa	Yeast display	[192]
15	Kunitz domain	Serine protease inhibitor (human)	Loop	6 kDa	'Phage display	DYAX [193]
16	Nanofitins/Affitins	DNA binding protein Sac7d (archaea Sulfolobus acidocaldarius)	Beta strand	7 kDa	Ribosome display	[194]
17	Repebody	Variable lymphocyte receptors (VLRs, jawless vertebrates)	Loops	Variable number of repeats (20-40 aa/repeat)	'Phage display	[195]
18	Obodies	OB-fold (Archaea Pyrobaculum aerophilum)	Loop and beta-strand	11 kDa	'Phage display	Obodies [196]

Table 1.2: Overview of antibody mimetics scaffolds and some of their basic characteristics.

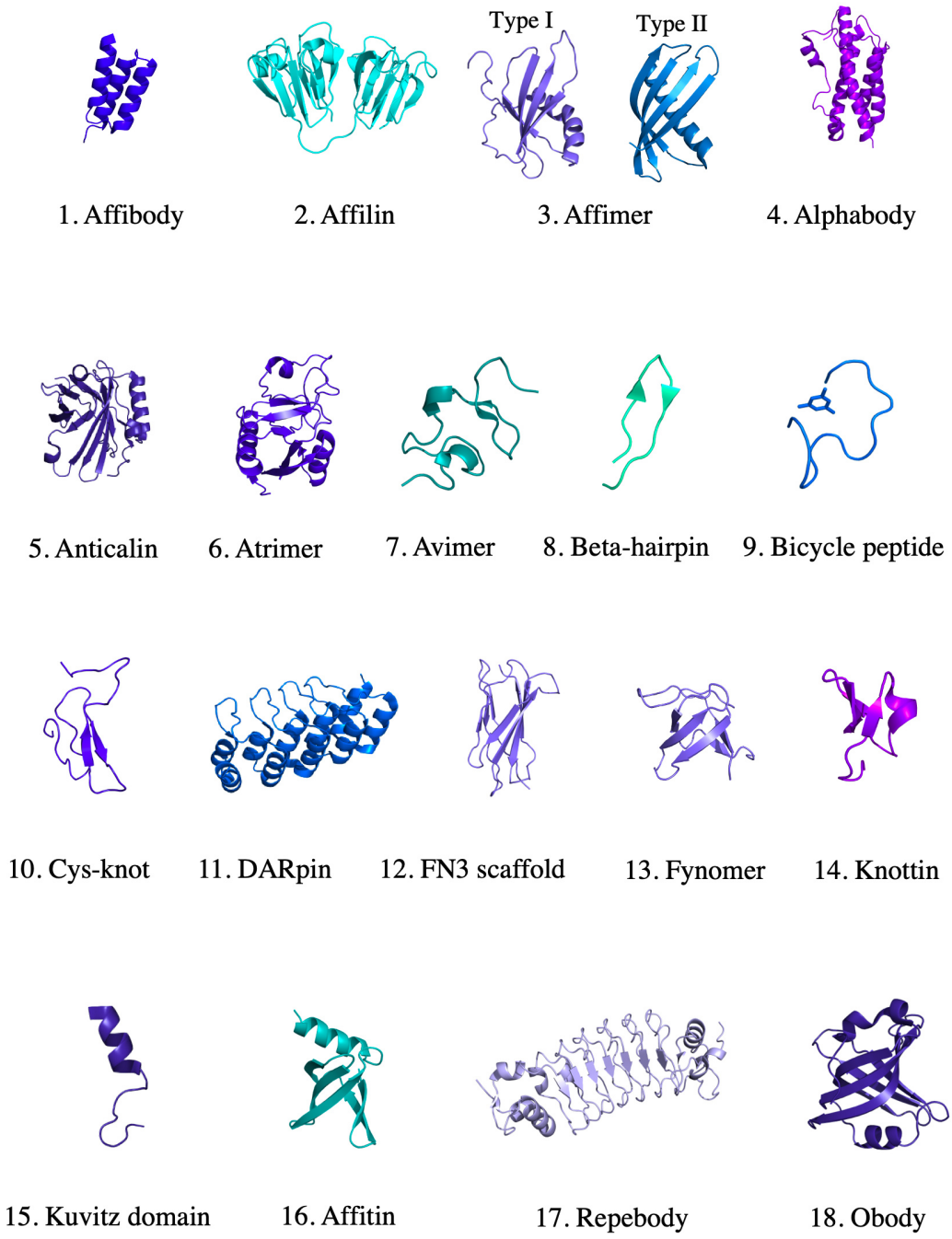


Figure 1.18: Structure of antibody mimetics scaffolds. 1. Affibody (PDB: 3MZW), 2. Affilin (PDB: 2JDG), 3. Affimer (PDB: 1NB5, 4N6T), 4. Alphabody (PDB: 5mj3), 5. Anticalin (PDB: 4GH7), 6. Atrimer (PDB: 3L9J), 7. Avimer (PDB: 1AJJ), 8. Beta-hairpin mimetics (PDB: 2NS4), 9. Bicycle peptide (PDB: 3QN7), 10. Cys-knots (PDB: 1HYK), 11. DARpin (PDB: 4YDY), 12. FN3 scaffold (PDB: 1FNF), 13. Fynomer (PDB: 1M27), 14. Knottin (PDB: 2IT7), 15. Kunitz domain (PDB: 4BQD), 16. Nanofitins/Affitins (PDB: 4CJ2), 17. Repebody (PDB: 4UIP), 18. Obodies (PDB: 4GLV). Images created using PyMOL.

1.3.4 Affimer reagents

One aim of this project has focused on the production and characterisation of Affimers as high-affinity binders for molecular recognition. This section reviews the structure, selection and potential applications of Affimer reagents.

1.3.4.1 Structure and selection of Affimers

Affimer reagents are a class of non-antibody binding proteins that can be classified as type I and type II based on their scaffold. Type I Affimer scaffold was first described in 2010 based on the human stefin A protein [197, 198] and type II was published later in 2014, based on a consensus plant phytocystatin sequence [181]. Both types contain four β-sheets and an α-helix with two variable loops presented between pairs of β-sheets that represent the target recognition sites.

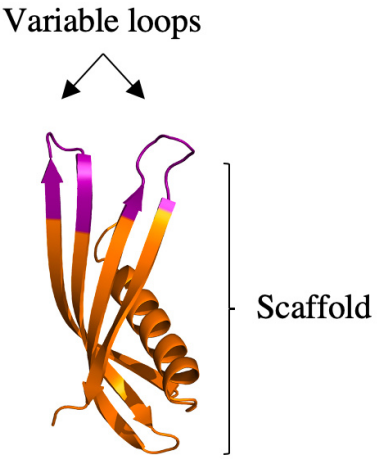


Figure 1.19: Representative crystal structure of an Affimer. Crystal structure of an Affimer against p300 (PDB: 5A0O). The Affimer scaffold, formed by an α-helix and 4 anti-parallel β-sheets is shown in orange and the two variable loops are shown in purple.

Affimers can be selected against diverse targets using large 'phage display libraries based on bacteriophage M13 clones. Each clone of the library expresses a truncated form of the gene 3 protein (g3p) attachment protein fused to an Affimer molecules harbouring a particular pair of variable loops. Affimer libraries directly link each Affimer variant (phenotype) to its cognate gene (genotype) through a discrete 'phage particle [199].

Affimer 'phage display libraries are constructed using sequences encoding nine random amino acids, excluding cysteine, which are introduced at loops 1 and 2 by codon-selected semi-trinucleotide cassette synthesis. Gene splicing by overlap extension is then used to introduce the variable loops into the Affimer scaffold, and the corresponding

products are subcloned into phagemid vectors [200]. The resulting phagemid vectors express a disulfide bond formation protein A (DsbA) secretion signal peptide, the Affimer sequence containing variable loops and the C-terminal half of gene 3 of bacteriophage M13, encoding for g3p attachment protein [181]. Affimer 'phage-display libraries are estimated to contain 1.3 x 10^{10} independent clones. *E. coli* cells are electroporated with the phagemid vectors and infected with a helper 'phage that contains the complete M13 genome encoding all the 'phage proteins needed for capsid production, 'phage assembly, chromosome replication, and budding [199]. Upon infection, the g3p-Affimer fusion protein is incorporated into the 'phage and can be used for Affimer screening by 'phage display selection.

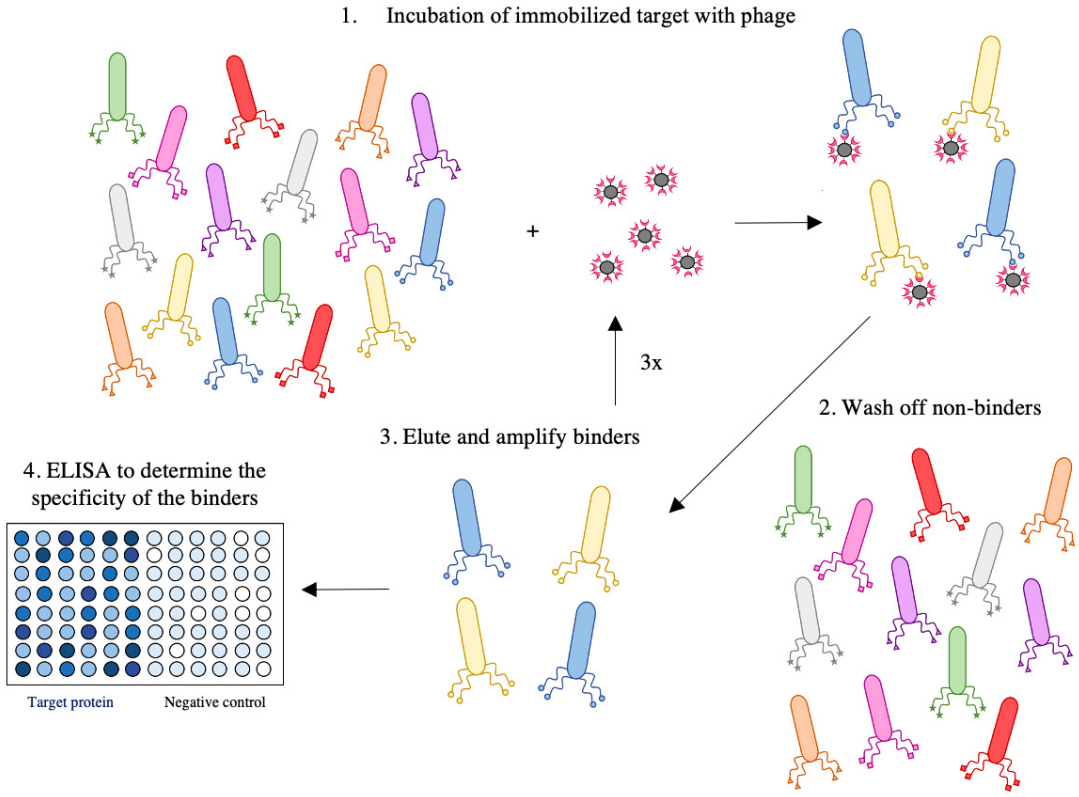


Figure 1.20: Affimers screening using 'phage display libraries. Schematic representation of the Affimers screening process. Target molecule is immobilised in micro-wells or magnetic beads and incubated with Affimer 'phage library (1). Non-bound phages are washed off (2), and bound phages are eluted and amplified in bacterial cells for a consecutive panning round (3). After three rounds of selection, individual phages displaying the selected binders are tested by 'phage ELISA to confirm their ability to bind to the target antigen (4).

'Phage display is a high-throughput method performed by mixing an Affimer 'phage display library with purified target molecules, normally presented as immobilised antigens

on solid surfaces by direct adsorption or by chemical modification of the antigen [199]. Biotin linkers are commonly used to immobilise the antigen on streptavidin or neutravidin coated wells or beads. After binding and washing steps, phages bound to the target molecule are eluted and amplified in *E. coli* cells infected with helper phages. Three panning rounds of selection are commonly performed, and monoclonal phages displaying the selected binders are tested by 'phage ELISA to confirm their ability to bind to the target antigen. These clones are then sequenced, and Affimer coding regions are subcloned into bacterial expression vectors for production of the selected binders [200].

1.3.4.2 Applications of Affimers

Affimer reagents have been raised against more than 350 different targets, ranging from small molecules [201] to recombinant proteins [201, 181] or whole viruses [202], most of them presenting affinities in the nanomolar range [201]. As a result, a broad diversity of applications for Affimers as tools for research, diagnostics, imaging, and potential for therapeutics can be found in the literature. Some examples of Affimer applications are detailed in table 1.3 and reviewed in this section.

The high affinity and specificity of Affimers have promoted their use in basic research and standard procedures of molecular biology. Transfection of plasmids encoding Affimers for transient expression has been successfully used for exploration of intracellular processes and signalling pathways [201], and modulation of function of membrane-bound receptors and ion channels has been explored by incubation of mammalian cells with recombinantly expressed Affimers [201]. Applications in standard molecular biology techniques include ELISA, pull down affinity assays or western blotting [201, 181, 202]. Affimers can be modified on their C-terminal cysteine, which permits their chemical functionalisation. Biotin-malemide moieties are commonly used for this purpose, allowing immobilisation of the binders into beads or surfaces for pull down assays, or conjugation with enzymes such as streptavidin-horseradish peroxidase (HRP) for ELISA or western blot analyses [201, 181].

C-terminal modifications of Affimers have also been successfully exploited for imaging purposes in affinity histochemistry, cell imaging or super-resolution imaging approaches, with examples in the literature of Affimer reagents conjugated to streptavidin-HRP [201], streptavidin-Alexa Fluor [201] or DNA-PAINT [203], respectively. Fluorescently labelled Affimers have also been used for *in vivo* imaging in mice, showing rapid tissue penetration and stability [201]. Potential magnetic resonance imaging (MRI) applications have been explored through trivalent chemical conjugation of Affimers

with the MRI contrast agent gadolinium(III)-1,4,7,10-tetraazacyclododecane-1,4,7,10-tetraacetic acid (Gd-DOTA) [201].

In the field of biosensing, Affimers have been validated in surface plasmon resonance (SPR) [204] and electrochemical biosensors [205, 206, 207] with sub-picomolar protein detection. Affimers can be engineered in their C-terminal cysteine for a controlled and oriented immobilisation on surfaces with high reproducibility, and their small size means that more of these proteins can be packed into a particular surface compared with antibodies [208].

Other applications of Affimers include shape controlled nanoparticle synthesis [209] and detection and discrimination between very similar small molecules such as 2,4,6-trinitrotoluene (TNT) and various dinitrotoluenes (DNT) [201]. The use of Affimers for therapeutic purposes has not been yet fully investigated, but their small size, rapid tissue penetration and high stability and solubility make them attractive candidates for therapeutic applications [181, 210].

Application	**Example**	**Reference**
Modulation of protein function	Anti-vascular endothelial growth factor receptor 2 (VEGFR2) Affimer	[201]
ELISA	Anti-2,4,6-trinitrotoluene (TNT) Affimer	[201]
Pull down affinity assay	Anti-growth factor receptor-bound protein 2 (Grb2) Affimer	[201]
Western blotting	Anti-yeast small ubiquitin-like modifier (SUMO) Affimer	[181]
Cell imaging	Anti-turkey herpesvirus (HVT) Affimer	[201]
Affinity histochemistry	Anti-tenascin C Affimer	[201]
Super-resolution microscopy	Anti-actin Affimer	[203]
In vivo imaging	Anti-tenascin C Affimer	[201]
MRI reagents	Trivalent Gd-1,4,7,10-Tetraazacyclododecane-1,4,7,10-tetraacetic acid (DOTA)	[211]
Biosensors	Anti-myc Affimer	[212]
Detection of small molecules	Anti-TNT Affimer	[201]
Magnetic nanoparticles	Magnetite interacting Affimer	[209]

Table 1.3: Examples of applications of Affimer reagents. Affimer reagents have been used in diverse applications in the fields of basic research, imaging and biosensing among others.

1.4 Project Aims

The work presented in this thesis forms part of the 'advanced diagnostics' work package of the MSCA-ITN HONOURs network on host switching pathogens, infectious outbreaks and zoonosis. The main aim of this project was to contribute to preparedness for potential outbreak scenarios with the development of diagnostic tools and performance of structural and functional studies of emerging zoonotic viruses.

CCHFV is one of the most widespread medically important zoonotic arboviruses and was selected as model example for this study, alongside with the closely-related and non-pathogenic in humans HAZV. Nairovirus NPs are the most abundant proteins in virions, and thus are optimal candidates as diagnostic targets. Furthermore, they perform essential structural roles in virus replication, increasing the relevance of their structural and functional studies. In this context, four main objectives were established for this project:

1. Expression and purification of recombinant CCHFV and HAZV NPs, and production of CCHFV NP-specific binders. Two different approaches were established for this purpose: production of CCHFV NP-specific polyclonal antibodies by a commercial company, and selection of CCHFV NP-specific Affimer reagents by the BioScreening Technology Group (BSTG, University of Leeds).

2. Characterisation and validation of CCHFV NP-specific Affimer reagents using biophysical and biochemical techniques, including pull down affinity precipitation assays, surface plasmon resonance, circular dichroism and X-ray crystallography.

3. Development of diagnostic assays using CCHFV NP-specific antibodies and Affimers. This objective was conceived as a collaboration with the company INGENASA (Eurofins Ingenasa, Madrid, Spain), which forms part of the MSCA-ITN HONOURs network, and was expected to provide the expertise and equipment required for the development of ELISA and lateral flow assays.

4. Structural and functional analyses of nairoviral NPs using a reverse genetics system. We aimed to investigate the incorporation of clonable tags in HAZV NP, for structural and functional analyses of nairoviral NPs in the context of infectious virus, using electron and fluorescence microscopy.

In summary, the focus of this project was on the production and characterisation of CCHFV NP-specific binders for the development of diagnostic tools, and the performance of structural and functional studies of nairoviral NPs using a reverse genetics system.

Chapter 2

Materials and Methods

2.1 Materials

2.1.1 Bacterial strains

Bacterial strains used for cloning, site-directed mutagenesis (SDM), plasmid maintenance and protein expression are detailed in Table 2.1.

Strain	Genotype	Used for
DH5α	F- endA1 glnV44 thi-1 recA1 relA1 gyrA96 deoR nupG Φ80dlacZΔM15 Δ(lacZYA-argF)U169, hsdR17(rK- mK+), λ–	Cloning, site directed mutagenesis and plasmid maintenance
Rosetta 2 (DE3)	F- ompT hsdSB(rB- mB-) gal dcm (DE3) pRARE2 (CamR)	Protein expression
JM109	endA1 glnV44 thi-1 relA1 gyrA96 recA1 mcrB+ Δ(lac-proAB) e14- [F' traD36 proAB+ lacIq lacZΔM15] hsdR17(rK-mK+)	Plasmid maintenance

Table 2.1: Bacterial *E. coli* strains used in this project

2.1.2 Plasmids

Bacterial and mammalian expression plasmids are detailed in table 2.2. The corresponding plasmid maps are illustrated in Appendix B.

	Plasmid amplification		Gene expression		
Plasmid name	**Bacterial strain**	**Antibiotic resistance**	**Mammalian/Bacterial expression**	**Promoter**	**Gene expressed**
Recombinant production					
pET-SUMO-CCHFV NP	DH5α	kan	Bacterial expression	T7	6xHis-SUMO-CCHFV NP
pET-SUMO-HAZV NP	DH5α	kan	Bacterial expression	T7	6xHis-SUMO-HAZV NP
Caspase-3-His	DH5α	amp	Bacterial expression	T7	Caspase-3-6xHis
Affimer expression					
pET-11a-Affimers	DH5α	amp	Bacterial expression	T7	Affimers-cys-8xhis
pET-SUMO-Affimer-NP	DH5α	kan	Bacterial expression	T7	6xHis-SUMO-Affimer-NP
pET-11a-Affimer-actin-RFP	DH5α	amp	Bacterial expression	T7	Affimer-actin-RFP
pMAX-Affimer-actin-RFP	DH5α	kan	Mammalian expression	CMV	Affimer-actin-RFP
pMAX-Affimer-yeast-SUMO-RFP	DH5α	kan	Mammalian expression	CMV	Affimer-yeast-sumo-RFP 6xHis
pMAX-Affimer-NP-RFP	DH5α	kan	Mammalian expression	CMV	Affimer-NP-RFP
HAZV rescue system					
HAZV-Lseg (pMK-RQ-L)	DH5α	kan	Mammalian expression	T7	HAZV L segment (antiviral sense)
HAZV-Mseg (pMK-RQ-M)	JM109	kan	Mammalian expression	T7	HAZV M segment (antiviral sense)
HAZV-Sseg (pMK-RQ-S)	DH5α	kan	Mammalian expression	T7	HAZV S segment (antiviral sense)
HAZV-Sseg-6xHis (pMK-RQ-S-6xHis)	DH5α	kan	Mammalian expression	T7	HAZV S segment (antiviral sense) with NP-6xHis
HAZV-Sseg-eGFP-β11 (pMK-RQ-S-eGFP-β11)	DH5α	kan	Mammalian expression	T7	HAZV S segment (antiviral sense) with NP-eGFP-β11
pCDNA-3.1-eGFP1-10	DH5α	amp	Mammalian expression	CMV	eGFP1-10
CCHFV replicon system					
CCHFV-SsegUTRs-eGFP	DH5α	amp	Mammalian expression	T7	eGFP ORF (neg. sense) flanked by CCHFV S seg-UTRs
CCHFV-NP-support (pC-NP)	DH5α	kan	Mammalian expression	T7	CCHFV NP ORF
CCHFV-NP-6xHis-support (pC-NP-6xHis)	DH5α	kan	Mammalian expression	T7	CCHFV NP ORF-6xHis
CCHFV-L-support (pC-L)	DH5α	spcr	Mammalian expression	T7	CCHFV RdRp ORF

Table 2.2: Plasmids used in this project. kan: kanamycin, amp: ampicillin, spcr: spectinomycin, CMV: human cytomegalovirus.

2.1.3 Mammalian cell lines

Mammalian cell lines, origin, and standard growth media are detailed in table 2.3.

Cell line	Origin	Growth media
SW-13	Human adrenal cortex carcinoma	DMEM supplemented with 10% (v/v) fetal bovine serum (FBS), 100 U/mL penicillin and 100 µg/mL streptomycin
BSR-T7	Baby hamster kidney (BHK) fibroblasts derived cell line stably expressing T7 RNA polymerase	DMEM supplemented with 2.5% (v/v) FBS, 100 U/mL penicillin and 100 µg/mL streptomycin

Table 2.3: Mammalian cell lines used in this project.

2.1.4 Virus strains and recombinant virus

HAZV strain JC280 was used as reference for the production of recombinant WT HAZV. The cRNA sequences of S, M and L segments from HAZV JC280 are available from GenBank with accession numbers M86624.1, DQ813514.1 and DQ076419.1, respectively. Recombinant HAZV (rHAZV) was rescued using the HAZV reverse genetics system previously described [156].

2.1.5 Antibodies

Primary and secondary antibodies used for western blot (WB), lateral flow assays (LFA) or immunofluorescence (IF) are detailed in table 2.4.

2.1.6 DNA and RNA oligonucleotides

Desalted DNA oligonucleotides were provided by Integrated DNA technologies (IDT). Fluorescein (3') labelled oligoribonucleotides were synthesised and HPLC purified by Dharmacon Inc. DNA and RNA oligonucleotide sequences are detailed in tables 2.5 and 2.6, respectively.

2.1.7 Fluorescence *in situ* hybridisation (FISH) probes

Fluorescence *in situ* hybridisation (FISH) DNA probes were designed against the anti-genomic (positive) strand of HAZV S segment RNA (NC_038711.1) excluding the 9

Antibody	Type	Species	Label	Source	Used for (dilution (v/v))
Anti-CCHFV NP	Polyclonal IgG	Rabbit	-	GenScript	WB (1/5K), LFA
Anti-HAZV NP	Polyclonal IgG	Sheep	-	Alta Bioscience	WB (1/3K), IF (1/500)
Anti-GAPDH	Polyclonal IgG	Mouse	-	Cell Signalling	WB (1/5K)
Anti-biotin IgG	Monoclonal IgG	Mouse	-	Operon	LFA
Anti-Rabbit IgG	Polyclonal IgG	Donkey	HRP	Abcam	WB (1/10K)
Anti-His tag	Monoclonal IgG	Mouse	HRP	Abcam	WB (1/1000)
Anti-His tag	Monoclonal IgG	Mouse	Alexa Fluor 647	Thermo Fisher	WB (1/500)
Anti-caspase 3	Monoclonal IgG	Rabbit	-	Cell Signalling	WB (1/1000)
Anti-Golgi 58K	Monoclonal IgG	Mouse	-	Insight Biotech.	IF (1/200)
Anti-Mouse IgG	Polyclonal IgG	Donkey	IRDye 680CW	LI-COR Biosciences	WB (1/15K)
Anti-Goat IgG	Polyclonal IgG	Donkey	IRDye 800CW	LI-COR Biosciences	WB (1/15K)
Anti-Rabbit IgG	Polyclonal IgG	Donkey	IRDye 800CW	LI-COR Biosciences	WB (1/15K)
Anti-Goat IgG	Polyclonal IgG	Donkey	Alexa Fluor 594	Invitrogen	IF (1/500)
Anti-Mouse IgG	Polyclonal IgG	Donkey	Alexa Fluor 488	Invitrogen	IF (1/500)

Table 2.4: Antibodies. Antibody type, species, label, source and dilution of use are indicated. WB: western blot, LFA: lateral flow assay, IF: immunofluorescence.

Oligo Name	Sequence (5'-3')
Affimers	
Affimer-forward	ATGGCTAGCAACTCCCTGGAAATCGAAG
Affimer-reverse-Cys	TTACTAATGCGGCCGCACAAGCGTCACCAACCGGTTTG
pCMV-reverse-NotI	TTACTAATGCGGCCGCACAAGCGTCACCAACCGGTTTG
RFP-reverse	TTTTGAGCTCTCAGTTCAGTTTGTGACCCAGTTTG
HAZV resQ primers	
HAZV-NP-His-forward	CACCACCACTAGGATTGCGGCTAGGTT
HAZV-NP-His-reverse	ATGATGATGGATGATGTTGATGTTGGTGG
HAZV-resQ-NP-HindIII-forward	GGTTAAGCTTGATTGCGGCTAGGTTCA
HAZV-resQ-EcoRI-reverse	CCAGAATTCCCGATGATGTTGATGTTGGTGG
EcoRI-eGFP-β11-HindIII-forward	ATATAGGGAATTCTGGAGGTCGTGACCACATGGTCCTTCATGAGTATGTAAATGCT GCCGGTA
EcoRI-eGFP-β11-HindIII-reverse	TATATAAGCTTAGGTGATACCGGCAGCATTTACATACTCATGAAGGACCATGTGGT CACGACCTCCAGAATTCCCTATAT
NP-C-terminus-forward	GTTTGCTGATAACGA
5'UTR-reverse	TCTCAAAGATATCGTTG
CCHFV replicon	
CCHFV-Sseg-Forward-NotI	AGCTAGCGGCCGCTTTGCTAACACTCAAGAGAAC
CCHFV-Sseg-reverse-NheI	CATGCTAGCTAGCTTACTTGTACAGCTCGTCCAT
eGFP-Forward-NheI	CATGCTAGCTTACTTGTACAGCTCGTCCAT
eGFP-Reverse-NotI	AGCTAGCGGCCGCATGGTGAGCAAGGGCG
CCHFV-NP-sup-forward-His	CAC CAC CAC CAC CAC CAC
CCHFV-NP-sup-reverse-His	GAT GAT GTT GGC GCT GGT G

Table 2.5: DNA oligonucleotides.

Oligo Name	Sequence (5'-3')	Label
27-mer RNA	UCGCUAAAAAUUUUCCUAUUAACAGUU	3'-fluorescein
48-mer RNA	AGUAGUGUACUCCACACUACAAACUUGCUAUUGUUGAAAAUCGCUGUG	3'-fluorescein

Table 2.6: RNA oligonucleotides.

terminal nucleotides of the untranslated regions (UTRs) using Stellaris Probe Designer version 4.2 (LGC Biosearch Technologies). The Stellaris Probe Designer software provided 68 oligonucleotide probe sequences that were 18 nucleotides in length with a minimum spacing of 2 nucleotides. Probes were assessed using NCBI Nucleotide Blast against hamster transcriptome (*Mesocricetus auratus*, taxID: 10036) and against the sequences of the M and L HAZV RNA segments (NC_038710.1 and NC_038709.1, respectively). Probes with 80% or more nucleotides complementarity to non-target RNAs were removed. A final set of 25 Stellaris FISH Probes recognising (+)RNA HAZV S segment labelled with Quasar 670 were purchased from Biosearch Technologies, Inc., Petaluma, CA. Sequences are detailed in table 2.7.

Probe set: positive_sense	
Reporter dye: Quasar 570	
Probe name	**Probe sequence (5'->3')**
positive_sense_1	TGTGTTGGCAGTGAACCT
positive_sense_2	TTATAGTTGAGGACTGCC
positive_sense_3	CATCATCGGGACTGGTGA
positive_sense_4	TCCAAGCTCGGACATTCT
positive_sense_5	GCAAACAGTTCAATGAGC
positive_sense_6	TTTTCTGTCCCTTGGGGT
positive_sense_7	CTAATCCCGTGGCCAAGA
positive_sense_8	TGCACCAGTCAAGATGCT
positive_sense_9	CATCATCGGAGCCACCGT
positive_sense_10	CGAATCATATCGCTGAGC
positive_sense_11	GGAGAAGCCGACATCCTT
positive_sense_12	GGTATGCCAGTAATGCCT
positive_sense_13	TTGGTAGTTGGCATCCCA
positive_sense_14	GTCGAACCATTCCAGACC
positive_sense_15	ATTGTCCGTGCTGGCTAG
positive_sense_16	GAGGAATTTCGGCACAAA
positive_sense_17	GCTGATGTTGTGAGGCGA
positive_sense_18	AGAAAAACAGCAGTAGAG
positive_sense_19	TTGGATCTCAAAGAGGCT
positive_sense_20	TTGAGAGGGGTGTTGATG
positive_sense_21	GGCGTTCGACTGCTTTGC
positive_sense_22	CACCCCTTTTAGGAGGTC
positive_sense_23	AGTTGTGCAATGCTTGCC
positive_sense_24	CCCCTCTAGCTCAACCAG
positive_sense_25	TACTGGTGTACTCCCCAA

Table 2.7: FISH probes against positive strand of HAZV S segment RNA.

2.1.8 Microscopes

For transmission electron microscopy, a FEI Tecnai T12 located in the University of Leeds was used, operating with a LaB_6 electron source, a voltage of 120 keV and a Gatan

US4000/SP detector with a charge coupled device (CCD) photon detector camera.

For fluorescence microscopy, a Zeiss LSM700 inverted microscope, a Zeiss LSM880 upright microscope, and a Zeiss LSM880 Airyscan inverted microscope were used.

2.2 Manipulation of recombinant DNA

2.2.1 Plasmid DNA amplification and isolation

Glycerol stocks or single colonies were used to inoculate overnight cultures of appropriate volume in Luria-Bertani (LB) medium containing antibiotics and DNA was isolated using Plasmid Midiprep or Miniprep kits (Qiagen) according to manufacturer's instructions. These two methods are based on the alkaline-lysis method. Midipreps were performed by first binding the plasmid DNA to a silica-based resin under low-salt and pH conditions. After removal of impurities by a medium-salt wash, plasmid DNA was eluted in a high-salt buffer and was desalted by isopropanol precipitation. In the case of minipreps, adsorption of DNA onto silica was done in the presence of high salt. After washing, the DNA was eluted in low-salt buffer. Purified DNA was quantified by spectrophotometry using a NanoDrop 1000 (Thermo Scientific).

2.2.2 Polymerase chain reaction (PCR)

Polymerase chain reactions (PCRs) were performed in 25 µL total volume and contained 12.5 µL of Q5 High-Fidelity 2X Master Mix (NEB), 1.25 µL of each forward and reverse primers (10 µM) and up to 1 µg of template DNA. Reaction cycles were performed in a thermocycler (Eppendorf) according to the manufacturer's instructions and products were isolated and purified by agarose gel electrophoresis (see section 2.2.4) followed by gel extraction (see section 2.2.5).

2.2.3 DNA restriction digest

Restriction enzyme digests were performed on plasmid DNA for subcloning purposes. Reactions contained 1 µg of DNA, 1X compatible NEBuffer (NEB) and 10 units of each enzyme in a 50 µL total volume. Digestions were carried out at 37°C for 1 h. Digestion products were purified by gel extraction (see section 2.2.5) after agarose gel electrophoresis (see section 2.2.4).

2.2.4 Agarose gel electrophoresis

Size and integrity of plasmid DNA, PCR products and restriction digests was determined using 1% (w/v) agarose gels. Gels were made using 0.5 g of agarose (VWR) in 50 mL 1X Tris-acetate-EDTA (TAE) buffer (40 mM Tris-acetate, 1 mM ethylenediaminetetraacetic acid (EDTA)) and 1:10,000 SYBR Safe DNA Gel Stain (Life Technologies). Samples were mixed with 5x DNA loading dye (Bioline) to a 1X final concentration, and run on gels at 110 V for 1 h in 1X TAE buffer alongside a molecular marker, Hyperladder I or 1kb DNA ladder (Bioline). DNA was visualised using a Safe Imager blue light transilluminator (Invitrogen). Images were taken with an InGenius LHR Gel Imaging System (Syngene) and were visualised using the GeneSnap software (Syngene).

2.2.5 DNA gel extraction

DNA was extracted from agarose gels after electrophoresis using Monarch DNA Gel Extraction Kit (NEB) according to the manufacturer's instructions. DNA fragments were excised from the gel using a scalpel and were dissolved to be further purified using DNA binding columns. DNA was quantified by spectrophotometry using a NanoDrop 1000 (Thermo Scientific).

2.2.6 DNA clean up

DNA clean up was performed using Monarch PCR and DNA Cleanup Kit (NEB) according to the manufacturer's instructions. The kit is based on a bind/wash/elute workflow using DNA binding columns. DNA was quantified by spectrophotometry using a NanoDrop 1000 (Thermo Scientific).

2.2.7 Ligations

Ligations were performed in 20 µL total volume containing 1X Ligase Reaction Buffer (Life Technologies), 4:1 molar ratio of insert:vector DNA and 1 unit of T4 DNA ligase (Life Technologies). Reactions were incubated at room temperature (RT) overnight and then transformed into DH5α *E. coli* cells.

2.2.8 Site-directed mutagenesis (SDM)

Q5 SDM Kit (NEB) was used according to the manufacturer's instructions. Briefly, an exponential amplification (PCR) of the DNA was performed using Q5 Hot start High-fidelity 2X Master Mix, followed by a kinase, ligase and *DpnI* (KLD) Treatment at RT for 5 min. Resulting DNA was transformed into competent bacterial cells.

2.2.9 Sequencing

Purified DNA samples were submitted to GENEWIZ for sequencing by the Sanger sequencing method.

2.3 Bacterial manipulation

2.3.1 Bacterial transformation

DH5α or Rosetta 2 (DE3) *E. coli* cells were transformed with the specified plasmids for plasmid amplification or protein expression, respectively. Cells were thawed on ice and 1 µL of plasmid DNA containing 5-90 ng was added to 50 µL of cells. Transformation mixes were incubated on ice for 30 min, heat shocked at 42°C for 42 s, and cooled on ice for 2 min. Cells were mixed with 450 µL of SOC media and were incubated at 37°C for 1 h with shaking. Transformed bacteria were spread onto antibiotic-containing LB agar plates and incubated at 37°C overnight. Single colonies were picked from plates and bacterial cultures were grown in LB media containing appropriate antibiotic overnight at 37°C.

2.3.2 Generation of glycerol stocks

Overnight bacterial cultures were centrifuged for 30 min at 2,000 xg and 4°C. LB media was removed and pellets were resuspended in 60% (v/v) glycerol in LB. Glycerol stocks were stored at -80°C until needed.

2.4 Analysis of proteins

2.4.1 Protein quantification

Protein concentration was estimated by spectrophotometry using a NanoDrop 1000 (Thermo Scientific) and the theorical extinction coefficient of the corresponding molecules (See Appendix A).

2.4.2 SDS-polyacrylamide gel electrophoresis (SDS-PAGE)

5 mL SDS-PAGE gels were made with 12 or 15% (v/v) acrylamide resolving gel (4 or 5 mL 30% (v/v) acrylamide:bis-acrylamide (29:1), 2.5 mL 1.5 M Tris-HCl pH 8.8, 3.3 or 2.3 mL double-distilled water (ddH_2O), 100 µL 10% (v/v) sodium dodecyl sulfate (SDS), 100 µL 10% (w/v) ammonium persulphate (APS), 10 µL TEMED) and 5% (v/v) acrylamide stacking gel (0.83 mL 30% (v/v) acrylamide, 0.63 mL 1 M Tris-HCl pH 6.8, 3.4 mL ddH_2O, 50 µL 10% (v/v) SDS, 50 µL 10% (w/v) APS, 5 µL TEMED). Protein samples were mixed 1:1 with 4 x lithium dodecyl sulfate (LDS) sample buffer (Invitrogen) containing Coomassie G250 and Phenol Red as tracking dyes, and 50 mM dithiothreitol (DTT) reducing agent. Samples were denatured by heating at 95°C for 5 min and loaded next to 3 µL of color prestained protein standard, broad range (NEB). Electrophoresis was performed in SDS running buffer (25 mM Tris base, 192 mM glycine, 0.1% (v/v) SDS) at 180 V for 1 h. Proteins were visualised using Coomassie staining (section 2.4.3) or western blotting (section 2.4.4).

2.4.3 Coomassie staining

Proteins resolved by SDS-PAGE were fixed in 40% (v/v) ethanol and 10% (v/v) acetic acid for 30 min at RT. Gels were then washed in ddH_2O prior to overnight staining with 15 mL of Coomasie Brilliant blue. A stock solution was prepared by dissolving 1 g of Coomassie brilliant blue (G-250, Thermo Scientific) in 5 mL of ddH_2O and 100 g of ammonium sulphate in 800 mL of ddH_2O, overnight. 20 g of phosphoric acid and the dissolved Coomasie stain were added to the ammonium sulphate mixture and volume was increased to 1 L. A working solution was prepared by mixing 80% (v/v) Coomassie brilliant blue stock solution with 20% (v/v) methanol. Gels were destained in 1% (v/v) acetic acid for 1 h at RT.

2.4.4 Western blotting (WB)

Proteins were transferred from SDS-PAGE gels to polyvinylidene fluoride (PVDF) membranes (Immobilon-FL Transfer membranes, Millipore) using a trans-blot semi-dry cell (Bio-Rad) in Towbin buffer (25 mM Tris, 192 mM glycine, 20% (v/v) methanol) for 1 hr at 15 V. Membranes were blocked overnight at 4°C in 50% (v/v) Odyssey blocking buffer (LiCor) in PBS (10 mM Na_2HPO_4, 2 mM KH_2PO_4, 137 mM NaCl, 2.7 mM KCl). Blocking buffer was then replaced with primary antibodies diluted in blocking buffer (25% (v/v) in PBS) and membranes incubated for 1 h at RT. Membranes were subsequently washed 3x in PBS for 10 min, then incubated with secondary antibodies. After 3x additional washes in PBS, secondary antibodies were directly visualised using a LiCor Odyssey Sa Infrared imaging system (LiCor) and Image Studio 4.0 software.

2.5 Expression and purification of recombinant proteins

2.5.1 CCHFV NP and HAZV NP expression and purification

Rosetta 2 (DE3) *E. coli* cells were transformed with pET-SUMO CCHFV NP or pET-SUMO HAZV NP plasmids and glycerol stocks were generated and stored at -80°C. Small scale starter cultures were grown overnight from stabs of glycerol stocks in 50 mL LB medium supplemented with kanamycin (50 µg/mL) at 37°C. 20 mL starter cultures were used to inoculate 1 L LB medium containing kanamycin (50 µg/mL). Cell cultures were grown at 37°C until optical density (OD) at 600 nm reached 0.6 - 0.8 Au. Expression of 6xHis-SUMO-CCHFV NP or 6xHis-SUMO-HAZV NP was induced by addition of Isopropyl β-D-1-thiogalactopyranoside (IPTG) to a final concentration of 500 µM. Induced cultures were incubated at 18°C for 16 h.

Bacterial cells were harvested by centrifugation at 4,000 xg at 4°C for 45 min using an SLC-6000 rotor (Sorvall). Pellets were resuspended in lysis buffer (500 mM NaCl, 20 mM Tris-HCl pH 7.4, 20 mM imidazole, 0.1% (v/v) Triton X-100, 1 mM $MgCl_2$, 1 mg/mL lysozyme from chicken egg white (Sigma Aldrich), Halt Protease Inhibitor Cocktail (1X), 1 U DNase I/1 L culture, 1 U RNase/1 L culture). 20 mL of lysis buffer were used to lyse the pellet from 1 L of bacterial culture. Pellet was resuspended and left on ice for 30 min. Cell lysis was completed by sonication using a Soniprep 150 sonicator. Sonication was carried out on ice at 10 µm amplitude for 30 cycles, each cycle consisting of 10 s on and 20 s off. Soluble and insoluble fractions were separated by centrifugation at 45,000 xg for 30 min at 4°C.

2.5.1.1 Ni^{2+}-NTA affinity chromatography

5 mL of Super Ni^{2+}-NTA Agarose resin (Generon) were used per 4 L of induced culture and placed in a 50 mL column. Resin was washed with ddH_2O to remove the 20% (v/v) ethanol in which it is stored. Ni^{2+}-NTA resin was equilibrated with binding buffer (500 mM NaCl, 20 mM Tris-HCl pH 7.4, 20 mM imidazole) immediately prior to applying the soluble bacterial lysate. The flow-through was collected and the resin washed with binding buffer with increasing concentrations of imidazole. For HAZV NP, a high-salt wash was performed before high imidazole elution, using high-salt washing buffer (2 M NaCl, 20 mM Tris-HCl pH 7.4, 20 mM imidazole). 6xHis-SUMO-CCHFV NP or 6xHis-SUMO-HAZV NP were eluted from the resin with elution buffer (300 mM or 500 mM imidazole, respectively, 500 mM NaCl, 20 mM Tris-HCl pH 7.4).

The elution fractions were analysed by SDS-PAGE followed by Coomassie staining to check the presence of the protein of interest and were subjected to dialysis in order to remove excess imidazole for SUMO-protease efficiency. 6xHis-SUMO ubiquitin-like specific protease 1 (Ulp1) was produced in-house. The total protein concentration of the elution fractions was estimated by spectrophotometry using a NanoDrop 1000 (Thermo Scientific). 0.5 mg of 6xHis-SUMO Ulp1 protease were added for every 10 mg of eluted 6xHis-SUMO-CCHFV NP or 6xHis-SUMO-HAZV NP. The mixture was decanted into 10k molecular weight cut-off (MWCO) dialysis tubing (Thermo Scientific) and placed in 5 L of dialysis buffer (500 mM NaCl and 20 mM Tris pH 7.4) for 16 h at 4°C on a magnetic stirrer.

After cleavage and dialysis of pooled eluted fractions, a second Ni^{2+}-NTA affinity chromatography was performed. Flow-through containing the protein of interest was collected. 6xHis-SUMO cleavage products and 6xHis-SUMO Ulp1 protease remained bound in the column and were eluted with elution buffer for further analysis.

2.5.1.2 Size exclusion chromatography

Protein sample was concentrated to 5 mL using centrifugal filters (Amicon Ultra 15 10k MWCO regenerated cellulose) and injected onto an equilibrated HiLoad 26/600 Superdex 75 pg (GE Healthcare) using an Akta prime at 0.5 mL/min. Size exclusion chromatography was performed at 4°C and 3 mL fractions were collected after void volume (110 mL). The purity of protein in eluted fractions was analysed by SDS-PAGE followed by Coomassie staining. Pure protein was then concentrated using Amicon Ultra 15 Ultracel 10K Regenerated Cellulose concentrators, mixed with glycerol to a

final concentration of 5% (v/v) and stored at -80°C or lyophilised as described in section 2.5.1.3.

2.5.1.3 Lyophilisation of protein samples

Some of the purified CCHFV NP samples were lyophilised overnight at -80°C under vacuum using an AdVantage 2.0 BenchTop freeze dryer / lyophilizer (SP Scientific).

2.5.2 Caspase 3 expression and purification

2.5.2.1 Ni^{2+}-NTA affinity chromatography

Caspase 3-His was expressed from a plasmid encoding the full-length protein fused to a C-terminal 6xHis tag (pET21b caspase 3, Addgene #90087). Protein expression was induced in the *E. coli* strain Rosetta 2 (DE3) using 200 µM IPTG at 30°C for 4 h. Cells were harvested by centrifugation and proteins were extracted by resuspension in lysis buffer (50 mM Tris-HCl pH 8, 100 mM NaCl, 1% (v/v) Triton X-100, 1 mg/mL chicken egg white lysozyme (Sigma Aldrich), 1 unit (U) DNase and 1 U RNase). Cell lysates were then clarified by centrifugation at 18,000 xg for 30 min at 4°C and supernatant was applied to Ni^{2+}-NTA resin that had been pre-equilibrated in binding buffer (100 mM NaCl, 50 mM Tris-HCl pH 8.0, 20 mM imidazole). Unbound fraction was collected and resin was washed with binding buffer and, subsequently, with washing buffer (500 mM NaCl, 50 mM Tris-HCl pH 8.0, 50 mM imidazole). Elution buffer (100 mM NaCl, 50 mM Tris-HCl pH 8.0, 200 mM imidazole) was used to dissociate 6xHis-caspase 3 from resin. Fractions containing purified caspase 3 were pooled and concentrated to 1 mg/mL. All fractions were analysed by SDS-PAGE and Coomassie staining.

2.5.3 *In vitro* caspase 3 cleavage assay

Different ratios of caspase 3 protease and CCHFV NP or HAZV NP were added to 1.5 mL micro-centrifuge tubes to a total of 10 µg of protein in 20 µL. Caspase 3 digestion buffer (2X) (20 mM piperazine-N,N'-bis(2-hydroxypropanesulfonic acid) (PIPES) pH 7.2, 200 mM NaCl, 20% (w/v) sucrose, 0.2% (w/v) 3-[(3-cholamidopropyl)-dimethylammonio]-1-propanesulfonate (CHAPS), 2 mM EDTA and 20 mM DTT) was added to reach the desired total volume of 40 µL. Reaction mixtures were incubated overnight at RT and NP cleavage was analysed by SDS-PAGE (section 2.4.2) and western blot analysis (section 2.4.4).

2.6 Generation of CCHFV NP-specific polyclonal antibodies

Recombinantly expressed and purified CCHFV NP was used for the generation of polyclonal antibodies. Lyophilised protein was sent to Genscript (USA), for inoculation of two rabbits. The immunisation schedule consisted on a primary immunisation plus three boosts every two weeks. Three months after primary immunisation, the production bleed was carried out. A protein A affinity purification step was carried out by Genscript to further purify the polyclonal IgGs from the serum of the rabbits.

2.7 Affimers production

Affimers screening was performed by the BSTG (University of Leeds) as previously described [200]. Purified CCHFV NP was provided and 3 panning rounds of ′phage display were performed against the biotinylated protein attached to streptavidin coated micro-wells (1st and 3rd selection rounds) or to streptavidin coated magnetic beads (2nd selection round). After propagation of individual ′phages, biotinylated CCHFV NP was immobilised onto neutravidin coated plates that were washed and incubated with medium containing individual ′phages. Results from the corresponding ′phage ELISA and Affimer ORF sequences of the unique binders were supplied by BSTG in a phagemid vector. The second panning round hits from the screening were also used for a screening using HAZV NP as target, with the aim of obtaining CCHFV and HAZV NP cross-reactive Affimer hits.

2.7.1 Affimers subcloning into pET-11a bacterial expression vector

ORFs of the Affimers were amplified from the phagemid vector by PCR (section 2.2.2) using Adhiron-foward and Adhiron-reverse primers (section 2.1.6). Corresponding PCR products and pET-11a expression vector were digested with *Nhe*I and *Not*I restriction enzymes. After purification of digested inserts and vector, a ligation was performed and the resulting products were transformed into DH5α *E. coli* cells. Single colonies were picked from transformation plates and grown in starter cultures containing LB with carbenicillin (100 µg/mL). DNA was isolated using Miniprep kits (Qiagen) and quantified by spectrophotometry using a NanoDrop 1000 (Thermo Scientific). Purified DNA was sent for sequencing (Genewiz, UK) to check the success of the ligations.

2.7.2 Affimers expression and batch Ni^{2+}-NTA affinity chromatography purification

Rosetta 2 (DE3) *E. coli* cells were transformed with pET-11a vectors containing the different ORFs of the Affimers. 1 mL of starter cultures was used to inoculate 50 mL of LB medium containing carbenicilin (100 µg/mL). Cell cultures were grown at 37°C until OD at 600 nm reached 0.6 - 0.8 Au. Expression of 8xHis-Cys-Affimers was induced by addition of IPTG to a final concentration of 100 µM. Induced cultures were incubated at 25°C for 6 h. Cells were harvested by centrifugation at 4,500 xg at 4°C for 20 min using a Thermo Scientific Heraeus Megafuge 16R centrifuge. Supernatant was removed and cell pellet was stored at -80°C until ready for purification. Pellets were resuspended in lysis buffer (50 mM NaH_2PO_4, 300 mM NaCl, 20 mM imidazole, 10% (v/v) glycerol, pH 7.4, 0.1 mg/mL lysozyme from chicken egg white (Sigma Aldrich), Halt Protease Inhibitor Cocktail (1X) (Thermo Scientific), Benzonase Nuclease (Novagen) (10 U/mL), Triton X-100 2% (v/v)). 1 mL of lysis buffer was used to lyse the pellet from 50 mL of bacterial culture. Pellet was resuspended and left on a rotor for 20 min at RT. Non-specific proteins were heat denatured by incubating the samples at 50°C for 20 min. Soluble and insoluble fractions were separated by centrifugation at 4,800 xg for 20 min.

300 µL of Super Ni^{2+}-NTA agarose resin (Generon) were equilibrated with 1 mL of lysis buffer in a micro-centrifuge tube. After centrifugation at 1,000 xg for 1 min, supernatant was removed and soluble fraction was added to the resin. After 1 h of incubation at RT on a rotor, the mixed solution was centrifuged at 1,000 xg for 1 min to sediment the Affimer-bound resin. Wash buffer (50 mM NaH_2PO_4, 500 mM NaCl, 20 mM imidazole, pH 7.4) was used to wash the resin until A280 of eluted samples was $<$0.09 Au. Resin was resuspended in 500 µL of elution buffer (50 mM NaH_2PO_4 pH 7.4, 500 mM NaCl, 300 mM imidazole, 10% (v/v) glycerol). After 5 min of incubation on a rotor at RT, resin was sedimented by centrifugation at 1,000 xg for 1 min. Supernatant containing the purified Affimer was stored in 40% (v/v) glycerol at -80°C or immediately biotinylated as described in section 2.7.3.

2.7.3 Biotinylation of Affimers

Biotinylation of Affimers was performed immediately after their purification. Before starting the biotinylation process, immobilised Tris(2-carboxyethyl)phosphine (TCEP) disulphide reducing gel (Thermo Scientific) was used to reduce any possible Affimer disulphide bonds. 150 µL of immobilised TCEP disulphide reducing gel were washed

3x with PBS containing 1 mM EDTA. 4 µL of PBS containing 50 mM EDTA and 150 µL of 0.5 mg/mL purified Affimer were added to the immobilised TCEP disulphide reducing gel. The mixture was incubated for 1 h at RT on a rotor and then centrifuged at 1,000 xg for 1 min. Supernatant containing reduced Affimers was recovered. For biotin labelling, 6 µL of biotin maleimide in DMSO (5 mg/mL) was immediately added to reduced Affimers and incubated at RT for 2 h. Zepa spin desalting columns (7K MWCO, Thermo Scientific) were used to remove any free biotin according to the manufacturer's instructions. Biotinylation was confirmed by ELISA as described in section 2.7.4.

2.7.4 ELISA to check biotinylation

Serial dilutions of the biotinylated Affimer in PBS were added to Nunc-Immuno MaxiSorp strips (Thermo Scientific) and incubated overnight at 4°C. After 3x washes with PBS-T, 250 µL of 10X blocking buffer were added per well and incubated at 37°C for 3 h. 3x additional washes were done with PBS-T and Streptavidin-HRP in 2X blocking buffer was added. After incubation for 1 h at RT on a vibrating platform shaker, 6x washes with PBS-T were done. 50 µL of 3,3',5,5'- tetramethylbenzidine (TMB) (SeramunBlau fast TMB/substrate solution, Seramun) were added per well and absorbance was measured at 620 nm after 5 min with a plate reader (TECAN).

2.7.5 Large scale expression and purification of Affimer-NP

Affimer-NP ORF was amplified from phagemid vector by PCR and was subcloned into a pET-SUMO bacterial expression vector (pET-SUMO-Affimer-NP) using *BamH*I and *Xho*I restriction sites. Rosetta 2 (DE3) *E. coli* cells were transformed with pET-SUMO-Affimer-NP. 20 mL of starter cultures were used to inoculate 1 L of LB medium containing kanamycin (50 mg/mL). Cell cultures were grown at 37°C until OD at 600 nm reached 0.6-0.8 Au. Expression of 6xHis-SUMO-Affimer-NP was induced by addition of IPTG to a final concentration of 500 mM. Induced cultures were incubated at 18°C for 16 h.

Bacterial cells were harvested by centrifugation at 4,000 xg for 45 min at 4°C using an SLC-6000 rotor (Sorvall). Pellets were resuspended in lysis buffer (500 mM NaCl, 50 mM NaH_2PO_4 pH 7.4, 20 mM imidazole, 0.1% (v/v) Triton X-100, 1 mM $MgCl_2$, 1 mg/mL lysozyme from chicken egg white (Sigma Aldrich), Halt Protease Inhibitor Cocktail (1X), 1 U DNase I/1 L culture and 1 U RNase). Cell lysis was completed by sonication using a Soniprep 150 sonicator, and soluble and insoluble fractions were

separated by centrifugation at 45,000 xg for 30 min at 4°C.

2.7.5.1 Ni^{2+}-NTA affinity chromatography

A first Ni^{2+}-NTA affinity chromatography was performed, using 5 mL of Super Ni^{2+}-NTA Agarose resin (Generon) in a 50 mL column. Resin was equilibrated with binding buffer (500 mM NaCl, 50 mM NaH_2PO_4 pH 7.4, 20 mM imidazole) immediately prior to applying the soluble bacterial lysate. Flow-through was collected and resin was washed with binding buffer with increasing concentrations of imidazole, after which 6xHis-SUMO-Affimer-NP was eluted from resin with elution buffer (500 mM NaCl, 50 mM NaH_2PO_4 pH 7.4, and 300 mM imidazole).

Elution fractions were subjected to dialysis in order to remove excess imidazole for increased SUMO-protease cleavage efficiency. A total of 0.5 mg of 6xHis-SUMO Ulp1 protease produced in-house was added for every 10 mg of eluted 6xHis-SUMO-Affimer-NP. The mixture was decanted into 10 kDa MWCO dialysis tubing (Thermo Scientific) and placed in 5 L of dialysis buffer (500 mM NaCl, 50 mM NaH_2PO_4 pH 7.4) for 16 h at 4°C on a magnetic stirrer.

After cleavage and dialysis of pooled eluted fractions, a second Ni^{2+}-NTA affinity chromatography was performed. Flow-through containing Affimer-NP was collected. 6xHis-SUMO cleavage product and 6xHis-SUMO Ulp1 protease remained bound in the column and were eluted with elution buffer for further analysis.

2.7.5.2 Size exclusion chromatography

Sample containing Affimer-NP was concentrated to 5 mL using centrifugal filters (Amicon Ultra 15 10 kDa MWCO regenerated cellulose) and was injected into an equilibrated HiLoad 26/600 Superdex 75 pg (GE Healthcare) column using an Akta prime at 0.5 mL/min. Size exclusion chromatography was performed at 4°C and 3 mL fractions were collected. The purity of Affimer-NP in eluted fractions was analysed by SDS-PAGE and Coomassie staining. Pure Affimer-NP was then concentrated using Amicon Ultra 15 Ultracel 10 kDa MWCO Regenerated Cellulose concentrators.

2.7.6 Purification of Affimer-NP and CCHFV NP complex

For complex purification, concentrated Affimer-NP and CCHFV NP were mixed and injected onto an equilibrated HiLoad 26/600 Superdex 75 pg (GE Healthcare) column

using an Akta prime at a flow rate of 0.5 mL/min. Size exclusion chromatography was performed at 4°C and 3 mL fractions were collected. The purity of the complex in eluted fractions was analysed by SDS-PAGE followed by Coomassie staining. Pure Affimer-NP and CCHFV NP complex was then concentrated using Amicon Ultra 15 Ultracel 10 kDa MWCO regenerated cellulose concentrators.

2.8 Characterisation of proteins and protein-protein interactions

2.8.1 Pull down affinity precipitation (AP) assays

Dynabeads MyOne Streptavidin T1 (Thermo Fisher) were washed and blocked overnight with 2x blocking buffer (Casein Blocking Buffer 10x, Sigma-Aldrich) in PBS-T. After washing the beads with PBS-T, they were incubated with 2 µg of biotinylated Affimer for 1 h at RT on a rotor. Beads were washed with PBS-T and 10 µg of the target molecule and pull down assays were performed using an automated KingFisher Flex System (Thermo Fisher). The program consisted in 3x washes with washing buffer (PBS-T) and elution of beads in running buffer (40 µL per sample). Samples were boiled for 5 min at 95°C, spun down 5 min at 11,000 xg and left on a magnetic rack for 2 min to remove the magnetic beads. Analysis of samples was done by western blot (section 2.4.4).

2.8.2 Circular Dichroism (CD)

CCHFV NP, HAZV NP, Affimer-NP and CCHFV NP + Affimer-NP complex samples were dialysed overnight in PBS using a Pur-A-Lyzer Mini Dialysis Kit, Mini 6000, capacity 10-250 µL, MWCO 6-8 kDa (Sigma, Cat. No. PURN60100-1KT) at 4°C. Protein samples were diluted to 0.2 mg/mL and were transferred to 1 mm path-length quartz cuvettes. Circular dichroism (CD) analyses and thermal melts were performed in a Chirascan Plus Spectometer (Applied Photophysics) monitoring ellipticity at 190-260 nm at temperatures ranging from 20°C to 90°C in 1°C intervals at 1°C/min. Elipticity values (mdeg) were normalised to mean molar ellipticity per residue (MRE) using the following equation:

$$MRE\ (\frac{deg \times cm^2}{dmol}) = \frac{elpiticity\ (mdeg)}{concentr.\ (mol/litre) \times path - length\ (0.1cm) \times n_{res}} \times \\ \times \frac{1deg}{10^3 mdeg} \times \frac{10^3 cm^3}{1litre} \times \frac{mol}{10dmol} \tag{2.1}$$

where n_{res} is the number of residues of the protein.

2.8.3 Surface Plasmon Resonance (SPR)

Surface plasmon resonance (SPR) analyses were done using a Biacore 3000 (GE Healthcare Life Sciences). Biotinylated Affimer-NP and Affimer-myo were immobilised on flow cells 2 and 3 of a Sensor Chip SA (GE Healthcare Life Sciences) via streptavidin-biotin interaction and flow cell 1 was left empty as a reference surface. For immobilisation of biotinylated Affimers, they were diluted in PBS to a final concentration of 100 nM and were injected into their respective flow cells at a flow rate of 5 µL/min until surface density reached 100 response units (RU). Injection of CCHFV NP and HAZV NP diluted in PBS at different concentrations was done at a flow rate of 20 µL/min for 120 s.

BIAevaluation software was used for double-referencing analysis, subtracting flow cell 1 and buffer only reference injections from each sensogram. Affinity and kinetic constants were calculated using a Langmuir 1:1 binding model.

2.8.4 X-ray crystallography

For crystallisation experiments, Affimer-NP and CCHFV NP complex was concentrated to 10 mg/mL. Crystal trials were performed using a high-throughput sitting drop experimental set up with an NT8 LCP drop setter (Formulatrix). The Joint Centre for Structural Genomics (JCSG) core suite was used for selection of crystallisation conditions mixing 1:1 or 2:1 ratios of protein:mother liquor volumes. Crystal growth was monitored using Rock Maker protein crystal visualisation software, imaging visible light, two-photon excited ultraviolet fluorescence (UV-TPEF) and second-harmonic generation (SHG). Crystals were obtained at RT from a 1:1 (v/v) ratio of protein to reservoir solution containing 0.2 M $NH_4CH_3CO_2$ (Salt), 0.1 M CH_3COONa 4.6 pH (Buffer) and 30% (w/v) PEG 4K (Precipitant). For vitrification, crystals were cryo-protected by transfer to a solution of mother liquor containing 25% (v/v) glycerol.

Diffraction data were collected at Diamond Light Source beamline I04. Structure of the complex was solved at 2.84 Å resolution by molecular replacement using previously published models of the Adhiron scaffold (PDB: 4n6u) and CCHFV NP (PDB: 4akl) with Phaser. Model-building and refinement were performed using Coot and Phenix, and data extraction was done with PDB_extract. The final structure was deposited in the Protein Data Bank (PDB) under accession number 6z0o. Interface residues between Affimer-NP loops and CCHFV NP were quantitated using Chimera and PDBePISA.

2.8.5 Fluorescence anisotropy (FA) assays

For direct binding fluorescence anisotropy assays, 15 µL of RNA binding buffer (100 mM NaCl, 10 mM Tris-HCl 7.5, 0.01% (v/v) Triton X-100) were added to each well of a black 384-well optiplate (Perkin Elmer). 30 µL of CCHFV NP were added to the first well and were subjected to 2 in 3 serial dilutions. 15 µL of 5 nM 3'-fluorescein (Fl) RNA (27-mer or 48-mer RNAs, section 2.1.6) were then added to each well, obtaining a final concentration of 2.5 nM FI-RNA per well. Reactions were incubated at RT for 30 min and fluorescence polarisation was measured using a Spark 10M multimode plate reader (Tecan) with excitation and emission filters selected at 485 nm (20 nm bandwidth) and 535 nm (25 nm bandwidth) wavelengths, respectively. Anisotropy values were calculated from raw emission values for parallel and perpendicular signals using the following equations:

$$Intensity = (2 \times I) + II \tag{2.2}$$

$$Anisotropy = \frac{II - (I \times 1)}{intensity} \tag{2.3}$$

where I = perpendicular signal and II = parallel signal.

Anisotropy values were plotted using GraphPad Prism and fitted to a nonlinear regression curve in order to obtain the K_D affinity values of CCHFV NP for different RNA oligonucleotides. For competition fluorescence anisotropy assays, 15 µL of RNA binding buffer were added to each well of a black 384-well optiplate (Perkin Elmer). 30 µL of Affimer-NP or negative control Affimer-myo were then subjected to 2 in 3 serial dilutions, after which 15 µL of 5 nM 3'-Fl RNA (27-mer or 48-mer) mixed with CCHFV NP were added to each well, obtaining a final concentration of 2.5 nM FI-RNA per well. Reactions were incubated at RT for 30 min and fluorescence polarisation was measured

as previously described. Anisotropy values were plotted using GraphPad Prism and fitted to a nonlinear regression curve to obtain the IC_{50} values.

2.8.6 Effect of Affimer-NP in CCHFV replication

2.8.6.1 Plasmid design

Briefly, the CCHFV S segment replicon was generated using the strain Baghdad-12 (GenBank accession number: AJ538196.1) as reference. Support plasmids (pC-NP and pC-L) expressing CCHFV NP and L proteins were generated by inserting the corresponding cDNA sequences from strain Baghdad-12 (GenBank accession numbers: CAD61342.1 and AY947890.1, respectively) cDNAs downstream of the T7 pol promoter and a cDNA representing the internal ribosome entry site (IRES) of encephalomyocarditis virus (EMCV). The S segment cDNA sequence was incorporated within a plasmid vector flanked by the bacteriophage T7 pol promoter and hepatitis delta virus ribozyme (HDV RZ), and orientated for the expression of viral sense RNA. The S segment major open reading frame (ORF) was initially replaced by a cDNA representing the firefly luciferase ORF, but this sequence was replaced by the enhanced green fluorescence protein (eGFP) ORF to generate pC-SsegUTRs-eGFP expressing eGFP as reporter signal of CCHFV-specific replication. Plasmid backbone containing the Sseg UTRs and eGFP ORF were amplified by PCR using primers CCHFV-Sseg-Forward-NotI, CCHFV-Sseg-reverse-NheI, eGFP-Forward-NheI and eGFP-Reverse-NotI (see section 2.1.6). The eGFP ORF was subcloned into Sseg UTRs backbone using restriction enzyme digestion and ligation to generate pC-SsegUTRs-eGFP.

The Affimer-NP ORF fused to the red fluorescence protein (RFP) ORF was amplified by PCR from pET-11a-Affimer-actin-RFP obtained from the BSTG using Affimer-forward and RFP-reverse primers (Section 2.1.6). PCR products were purified and subcloned into a pMAXcloning mammalian expression vector designed for protein expression under the human cytomegalovirus (CMV) promoter using *Sac*I and *Xba*I restriction sites, resulting in pMAX-Affimer-actin-RFP. Affimer-NP and Affimer-YS ORFs were amplified by PCR using Affimer-forward and Affimer-reverse-Cys primers (section 2.1.6) and subcloned into pMAX-Affimer-actin-RFP to replace Affimer-actin ORF for Affimer-NP or Affimer-YS ORFs using *Nhe*I and *Xho*I restriction sites, resulting in pMAX-Affimer-NP-RFP and pMAX-Affimer-YS-RFP, respectively.

2.8.6.2 Live cell imaging of Affimers and CCHFV replicon

BSR-T7 cells were seeded in 12-well plates 1 day prior to transfection at a density of 10^5 cells per well. Cells were transfected with 450 ng pC-L, 150 ng pC-NP, 150 ng pC-SsegUTRs-eGFP and increasing concentrations of pMAX-Affimer-NP or a negative control pMAX-Affimer-YS-RFP plasmid. TransIT-LT1 Transfection Reagent (Mirus Bio) was used as transfection reagent in a 2.5:1 ratio of TransIT-LT1 (µL) : DNA (µg) according to the manufacturer's instructions. At 6 h post-transfection (hpt) media was replaced with DMEM supplemented with 2.5% (v/v) FBS, 100 U/mL penicillin and 100 µg/mL streptomycin. Phase contrast, and green and red fluorescence were monitored using an IncuCyte Live Cell analysis system. Quantification of green and red unit counts per well was done using the IncuCyte ZOOM software 2018.

2.9 Colorimetric diagnostic assays

2.9.1 Enzyme-linked immunoassays (ELISAs)

96-well high-binding, clear plates (Sigma-Aldrich) were coated overnight at 4°C with 1 µg of Affimer-NP diluted in 100 µL carbonate buffer (50 mM $NaHCO_3$/Na_2CO_3, pH 9.6). Affimer-NP was removed and plates were blocked with 150 µL of ELISA stabilising solution (Surmodics) for 1 h. Plates were then washed 3x with PBS-T and incubated with 100 µL of sample to be tested (buffer or animal/human sera spiked with recombinant NPs at indicated concentrations) for 1 h. After 3x washes with PBS-T, plates were incubated with a primary antibody (CCHFV NP-specific polyclonal rabbit IgG) diluted in PBS-T for 1 h. Plates were then washed 3x with PBS-T and incubated with 100 µL of secondary antibody (goat anti-rabbit polyclonal IgG-HRP) for 1 h. After 3x washes with PBS-T, plates were incubated with 100 µL of TMB Enhanced K-Blue Substrate (Neogen) for 10 min. Oxidation of TMB was stopped with 100 µL of stop solution (0.5 M sulfuric acid). OD at 450 nm was measured with a SpectraMax M5 (Molecular Devices) plate reader using SoftMax Pro 5.2 software.

2.9.2 Lateral flow assays (LFAs)

The LFA test to detect CCHFV NP was based on the use of different coloured carboxyl-modified latex microspheres, which were covalently linked to either CCHFV NP-specific rabbit polyclonal IgGs, Affimer-NP or biotin-BSA. For the test line, CCHFV

NP-specific rabbit polyclonal IgGs or Affimer-NP were conjugated to red latex particles (Ikerlat, 300 nm particle size) and for the control line, BSA-biotin was conjugated to blue latex particles (Ikerlat, 260 nm particle size). For the test line in the nitrocellulose membrane, purified Affimer-NP or CCHFV NP-specific rabbit polyclonal IgGs were dispensed, and for the control line anti-biotin mouse IgG were dispensed.

2.9.2.1 Functionalisation of latex beads

Latex beads were washed in MES 10 mM pH 6, and were centrifuged at 11,000 xg for 15 min. Supernatant was removed and beads were resuspended in the same buffer. Size of the beads was measured by dynamic light scattering (DLS) using a Zetasizer Nano S system (Malvern) and, if necessary, sonified at 15% amplitude using a SFX250 digital sonifier (Branson). Latex beads were then activated using EDC (1-ethyl-3-(3-dimethylaminopropyl) carbodiimide hydrochloride) and NHS (N-hydroxysuccinimide). Beads were centrifuged at 11,000 xg for 15 min, supernatant was removed, and beads were resuspended in MES 10 mM pH 6.0.

The proteins of interest (CCHFV NP-specific rabbit polyclonal IgGs, Affimer-NP or biotin-BSA) were diluted in the same buffer, beads were added drop-wise to the protein solution and were incubated for 2 h at RT. Conjugation was stopped by adding 500 mM imidazole and beads were washed with 0.1% (v/v) Tween-20. After centrifugation, supernatant was removed and the size of the beads was again measured by DLS.

Blue and red functionalised latex beads were mixed at different ratios and were diluted in Tris-HCl 25 mM pH 9.5 buffer. Mixture was dispensed onto a rayon conjugate pad 25 mm (Operon), using a Matrix 1600 dispensing module (Kinematic Automation, Inc.). Conjugate pad was incubated 30 min in a MMUFE 500 universal oven model (Memmert) at 45°C.

2.9.2.2 Membrane dispensation

For the test line, Affimer-NP or CCHFV NP-specific rabbit polyclonal IgGs were diluted at different concentrations in buffer Tris-HCl 20 mM (pH 8.5 or 7.5, respectively), containing 5.0% (w/v) sucrose and 0.095% (w/v) sodium azide as preservative. Anti-biotin control IgG monoclonal antibody, used as control line capture reagent, was diluted at 1 mg/mL in the same buffer (pH 7.5). Both test and control capture reagents were dispensed in two parallel lines on nitrocellulose membrane at 1 µL/cm. Dispensed membranes were dried for 5 min in a MMUFE 500 universal oven model (Memmert)

at 45°C.

2.9.2.3 Assembly of LFAs

A master card was assembled on a 72 mm plastic backing card with adhesive (Kenosha) as follows: HiFlow Plus Nitrocellulose Membrane (HF120) (Millipore), the conjugate pad (Operon, Ltd) and a 25 mm absorbent pad (Ahlstrom, Ltd) were pasted and covered with a 65 mm transparent adhesive protector film (Lohmann, The Bonding Engineers). The master card was then cut to 4.2 mm width strips using a Matrix 2360 machine (Kinematic Automation).

2.9.2.4 LFA sample testing

Different procedures and buffers (detailed in table 2.8) were tested for sample application optimisation. Sequential or simultaneous addition of sample and buffer were tested. 10 min after sample application, results were annotated. For quantitative results, ESEQuant Lateral Flow reader (Qiagen) was used to convert the colorimetric signal in the test line into a 2D peak profile plot. Peak height (mV) was classified as positive, low positive, doubtful or negative.

LFA running buffer	Composition
TDM-36	250 mM Tris-HCl pH 7.2, 1 M NaCl, 1% (v/v) Tween-20, 1% (w/v) BSA, 0.095% (w/v) azide
TDM-42	250 mM Tris-HCl pH 7.5, 150 mM NaCl, 1% (v/v) Tween-20, 1% (w/v) casein, 0.095% (w/v) azide
TDM-45	250 mM Tris-HCl pH 7.5, 150 mM NaCl, 1.5% (v/v) Tween-20, 1% (w/v) casein, 0.095% (w/v) azide

Table 2.8: LFA running buffers tested.

2.10 HAZV reverse genetics system

2.10.1 Plasmid design: C-terminal tagged HAZV S segment

Generation of a plasmid expressing a complementary copy of the HAZV S segment with a 6xHis tag in the C-terminal region of HAZV NP ORF (pMK-RQ-S-6xHis) was achieved by molecular biology techniques starting from the previously described HAZV S segment plasmid (pMK-RQ-S) [156]. A 6xHis tag was inserted between the last codon

of the HAZV NP ORF and its stop codon using HAZV-NP-his-forward and HAZV-NP-his-reverse primers (section 2.1.6) and a Q5 SDM Kit according to the manufacturer's instructions. DNA products were transformed into DH5α *E. coli* cells for amplification, DNA was purified and pMK-RQ-S-6xHis sequence was confirmed by Sanger sequencing (Genewiz).

Generation of a plasmid expressing HAZV NP and eGFP-β11 separated by a 5 aa linker region (pMK-RQ-S-eGFP-β11) was achieved by molecular biology techniques starting from pMK-RQ-S. *EcoR*I and *Hind*III restriction sites were engineered between the last codon of the HAZV NP ORF and its stop codon using HAZV-resQ-NP-HindIII-forward and HAZV-resQ-EcoRI-reverse primers (section 2.1.6) and a Q5 SDM Kit according to the manufacturer's instructions.

Two annealing oligonucleotides encoding a 5 aa linker and the 11th beta strand of the eGFP (eGFP-β11: RDHMVLHEYVNAAGIT) flanked by *EcoR*I and *Hind*III restriction sites, were designed and purchased from Integrated DNA Technologies (EcoRI-eGFP-β11-HindIII-forward and EcoRI-eGFP-β11-HindIII-reverse primers, section 2.1.6). Primers were mixed at equimolar concentrations and were heated at 95°C for 5 min. The mix was let to cool down for 1 h and then transferred to ice.

Annealed primers and pMK-RQ-S plasmid containing *EcoR*I and *Hind*III restriction sites were digested with the corresponding enzymes at 37°C for 1 h. Digested products were resolved on 1% (w/v) agarose gels via electrophoresis and DNA bands corresponding to insert and vector were excised from agarose and purified. Ligations were performed and DNA products were transformed into DH5α *E. coli* cells for amplification. Sequence of purified pMK-RQ-S-eGFP-β11 was confirmed by Sanger sequencing (Genewiz).

2.10.2 Rescue of rHAZV

6-well plates were seeded with 2 x 10^5 BSR-T7 cells/well in 2 mL DMEM supplemented with 2.5% (v/v) FBS, 100 U/mL penicillin and 100 µg/mL streptomycin. After 24 h, cells were transfected with 1.2 µg pMK-RQ-S, pMK-RQ-M, and pMK-RQ-L and 0.6 µg pCAG-T7pol, combined with 2.5 µL Mirus TransIT-LT1 transfection reagent (Mirus Bio) per microgram of DNA, in 200 µL Opti-MEM (Life Technologies). For mutant recovery, the WT plasmid for the S segment (pMK-RQ-S) was replaced with the corresponding mutant plasmid containing a C-terminal tag. A control sample, in which transfection of pMK-RQ-L was omitted, was set up alongside each experiment.

Cell supernatants were collected 4 days post-transfection, and 300 µL supernatant was passaged for 48 h to a 6-well plate of SW13 cells grown in DMEM supplemented with 10% (v/v) FBS, 100 U/mL penicillin and 100 µg/mL streptomycin.

2.10.3 Virus infections

SW13 monolayers were infected with rHAZV at the specified multiplicity of infection (MOI) in serum-free media (SFM: DMEM supplemented with 100 U/mL penicillin and 100 µg/mL streptomycin) at 37°C. After 1 h, the inoculum was removed and cells were washed in PBS. Fresh DMEM supplemented with 2.5% (v/v) FBS, 100 U/mL penicillin and 100 µg/mL streptomycin was then applied for the duration of the infection.

2.10.4 Virus titration

Plaque assays were used to determine HAZV titre. SW13 cells were seeded in 6-well (3 x 10^5 SW13 cells/well) or 12-well (1 x 10^5 SW13 cells/well) plates. After 24 h, HAZV was serially diluted in SFM to generate dilutions varying from 10^{-1} to 10^{-6}. Serial dilutions were used to infect SW13 cells in duplicate. Infections were carried out as previously described with the exception that SFM containing HAZV was immediately replaced with complete media diluted 1:1 with 1.6% (w/v) carboxy-methyl cellulose (CMC, Sigma). CMC prevented HAZV from migrating, ensuring that following release from an infected cell, HAZV particles would only infect neighbouring cells. Cells were then incubated for 6 days at 37°C and 5% (v/v) CO_2, then fixed and stained with crystal violet solution (10% (v/v) formaldehyde, 2% (w/v) crystal violet, 20% (v/v) ethanol) for 30 min. Plaques were visualised using a light box as areas that had not been stained by crystal violet solution. Virus titre was estimated using the following equation:

$$Virus\ titre\ (pfu/mL) = \frac{Average\ number\ of\ plaques}{Dilution\ factor \times Volume\ of\ diluted\ virus\ added} \quad (2.4)$$

where pfu = plaque forming units.

2.10.5 RT-PCR analysis of rHAZV

Viral RNA was first extracted from cell-free supernatant using the QIamp Viral RNA kit (Qiagen) and treated with TURBO DNase (Thermo Fisher) to remove any

contaminating DNA prior to further purification using the RNeasy kit (Qiagen). A cDNA copy was generated using ProtoScript II reverse transcriptase (NEB) according to the manufacturer's instructions, alongside a control in which the reverse transcriptase was omitted. PCR amplification of the C-terminal region of HAZV NP ORF was achieved using NP-C-terminus-forward and 5′ UTR-reverse primers (section 2.1.6) specific to the 5′-terminus of the HAZV S segment and PCR products were resolved on a 1% (w/v) agarose gel and sequenced by Genewiz.

2.10.6 Virus growth curves

To assess viral fitness over time, 1 x 10^5 SW13 cells/well were seeded into 12-well plates one day prior to infection. Virus was used to infect wells at an MOI of 0.01 and supernatants were harvested at 24 h intervals and stored at -80°C until all samples had been collected and stored in a similar manner. Following collection of all time points, samples were analysed for titre of infectious virus as previously described.

2.11 Purification of HAZV

Sucrose cushions were used for concentration and purification of rHAZV from clarified cell culture supernatant. SW13 cells were seeded into T-175 flasks at a confluence of 4 x 10^6 cells. After 18 h, cells were infected with rHAZV at an MOI of 0.001. After 3 days of infection, cell supernatant was harvested and clarified by centrifugation at 4,000 xg for 30 min at 4°C. Supernatant was filtered through a 0.45 µM filter and span down again at 4,000 xg for 30 min at 4°C. A sucrose cushion was prepared with 20% (w/v) sucrose in Tris-NaCl-EDTA (TNE) buffer (10 mM Tris pH 7.4, 100 mM NaCl and 1 mM EDTA) supplemented with Halt Protease Inhibitor Cocktail (1X) (Thermo Scientific), 1 mM $MgCl_2$ and 0.5 mM $CaCl_2$.

30 mL of clarified cell culture supernatant were transferred into a thin-wall, Ultra-Clear, 38.5 mL ultracentrifuge tube (Beckman Coulter). 8 mL of sucrose mixture were carefully layered under the cell supernatant prior to ultracentrifugation at 100,000 xg for 3.5 h at 4°C in an Optima XPN Beckman Coulter ultracentrifuge using an SW 32.1 Ti Swinging-Bucket Rotor (Beckman Coulter). Supernatant was carefully discarded and the resulting pellet was air dried for 5 min to ensure complete removal of supernatant prior to resuspension in 0.1x PBS buffer overnight at 4°C. Resuspended virus was harvested and frozen in aliquots at -80°C or directly used for electron microscopy, RNP purification or

DiD labelling. Samples of supernatant, sucrose cushion and virus pellet were analysed by western blotting as detailed in section 2.4.4.

2.12 Purification of 6xHis-tagged HAZV RNPs

SW13 cells were seeded into T-175 flasks at a confluence of 4 x 10^6 cells. After 18 h, cells were infected with 6xHis-tagged NP rHAZV at an MOI of 0.001. 3 days post-infection, 6xHis-tagged HAZV RNPs were purified either from cell supernatant or from cell lysate as described below.

2.12.1 Harvest of 6xHis-tagged HAZV RNPs from released virus

6xHis-tagged NP rHAZV was purified from cell supernatant as detailed in section 2.11. Purified virus was subjected to 3 freeze-thaw cycles and was mixed with 1.5% (v/v) NP-40 and triton X-100 1.5% (v/v) to break virus membrane and release the RNPs. Sample was mixed 1:5 with binding buffer (200 mM NaCl, 20 mM Tris-HCl pH 7.4, 20 mM imidazole) and used to purify 6xHis-tagged RNPs by Ni^{2+}-NTA affinity chromatography as detailed in section 2.12.3.

2.12.2 Harvest of 6xHis-tagged HAZV RNPs from cell lysate

Cells were washed with ice cold PBS and lysed with radioimmunoprecipitation assay (RIPA) buffer (50 mM Tris-HCl pH 7.5, 150 mM NaCl, 1% (v/v) NP-40 alternative, 0.5% (w/v) sodium deoxycholate, 0.1% (v/v) SDS and Halt Protease Inhibitor Cocktail (1X)). Cell lysate was harvested and clarified by centrifugation at 4,000 xg for 10 min at 4°C. Pellet was discarded and soluble fraction was mixed 1:5 with binding buffer (200 mM NaCl, 20 mM Tris-HCl pH 7.4, 20 mM imidazole) and used for purification of 6xHis-tagged RNPs by Ni^{2+}-NTA affinity chromatography as detailed in section 2.12.3.

2.12.3 Ni^{2+}-NTA affinity chromatography

Super Ni^{2+}-NTA Agarose resin (Generon) was placed in a 20 mL column and washed with ddH_2O to remove the 20% (v/v) ethanol in which it is stored. Ni^{2+}-NTA resin was equilibrated with binding buffer (200 mM NaCl, 20 mM Tris-HCl pH 7.4, 20 mM imidazole) immediately prior to applying the sample containing 6xHis-tagged HAZV

RNPs. Flow-through was collected and resin washed with binding buffer with increasing concentrations of imidazole. 6xHis-HAZV RNPs were eluted from the resin with elution buffer (300 mM imidazole, 200 mM NaCl, 20 mM Tris-HCl pH 7.4). Sample was directly used for EM sample preparation as detailed in section 2.13.1.

2.13 Electron microscopy (EM)

2.13.1 Negative stain of EM grids

Carbon-coated EM grids were rendered hydrophilic immediately prior to sample application by glow discharge in air. EM grids were placed carbon side up on to a parafilm coated glass slide, and subjected to glow discharge for 45 s using a PELCO easyGlow glow discharge system. 3 µL of sample was applied to a glow discharge treated EM grid with continuous carbon film for 1 minute. Excess sample was blotted away using filter paper (Whatman) and the grid was washed with 3 µL of ddH_2O. After blotting away the ddH_2O, grids were equilibrated with 3 µL of negative stain solution (2% (w/v) uranyl acetate) that was subsequently blotted away. A final droplet of 3 µL of stain was applied to the grid and left for 1 min. Excess stain was blotted away and the grid left to air dry before loading into the TEM.

2.13.2 Transmission electron microscopy (TEM)

Negatively stained EM grids were loaded into a single tilt holder and inserted into the TEM. Microscope alignments were performed from the top to the bottom of the TEM column.

2.14 Size and frequency distribution analysis of 6xHis-tagged HAZV RNPs

Length and width of 6xHis-tagged HAZV RNPs particles were analysed using Fiji. Length was analysed using the segmented line tool and width was analysed using the straight line tool. Microsoft Excel was used to analyse the mean and standard deviation (SD) of the 3 different bins of RNP lengths.

2.15 6xHis-tagged-NP CCHFV mini-genome

2.15.1 Plasmid design: 6xHis-tagged CCHFV NP support plasmid

Generation of a support plasmid expressing CCHFV NP tagged with a 6xHis C-terminal tag (pC-NP-6xHis) was achieved by molecular biology techniques starting from the previously described pC-NP support plasmid of the the CCHFV S segment mini-genome system (section 2.8.6). pC-NP support plasmid contained a 6xHis sequence after the stop codon of the CCHFV NP ORF. The sequence between the last codon of the CCHFV NP ORF and the first codon of the 6xHis tag was deleted using CCHFV-NP-sup-forward-His and CCHFV-NP-sup-reverse-His primers (section 2.1.6) and a Q5 SDM Kit (NEB) according to the manufacturer's instructions. DNA products were transformed into DH5α *E. coli* cells for amplification, DNA was purified and pMK-RQ-S-6xHis sequence was confirmed by Sanger sequencing (Genewiz).

2.15.2 Live cell imaging of CCHFV replicon

BSR-T7 cells were seeded in 12-well plates 1 day prior to transfection at a density of 10^5 cells per well. Cells were transfected with 450 ng pC-L, 150 ng pC-NP or 150 ng pC-NP-6xHis, and 150 ng pC-SsegUTRs-eGFP. TransIT-LT1 Transfection Reagent (Mirus Bio) was used as transfection reagent in a 2.5:1 ratio of TransIT-LT1 (µL) : DNA (µg) according to the manufacturer's instructions. At 6 hpt media was replaced with DMEM supplemented with 2.5% (v/v) FBS, 100 U/mL penicillin and 100 µg/mL streptomycin. Phase contrast and green fluorescence were monitored using an IncuCyte Live Cell analysis system. Quantification of green unit counts per well was done using the IncuCyte ZOOM software 2018.

2.16 Production of eGFPcomp HAZV

2.16.1 Live cell analysis of eGFPcomp fluorescent signal

BSR-T7 cells were seeded in 12-well plates (8 x 10^4 cells/well) and after 16 h, cells were transfected with pMK-RQ-S-eGFP-β11 and pcDNA3.1-eGFP(1-10) (Addgene # 70219 [213]) or pcDNA3.1-eGFP(1-10) only, using Mirus TransIT-LT1 transfection reagent (Mirus Bio) according to manufacturer's instructions. After 24 h, media was removed and replaced with fresh media or cells were infected with rHAZV-eGFP-β11

at the specified MOI. Fluorescence signal was monitored using an IncuCyte Live Cell Analysis System (Sartorius) or samples were analysed by immunofluorescence

2.16.2 Vybrant DiD labelling of eGFPcomp HAZV

BSR-T7 cells were seeded in T-175 flasks (4 x 10^6 cells/flask) and after 16 h, cells were transfected with pcDNA3.1-eGFP(1-10) (Addgene No. 70219 [213]) using Mirus TransIT-LT1 transfection reagent (Mirus Bio) according to manufacturer's instructions. After 24 h, media was removed and cells were infected with rHAZV-eGFP-β11 at an MOI of 0.1. After 2 days of infection, cell supernatant was harvested and eGFPcomp rHAZV was purified as described in section 2.11. Resuspended virus was harvested and 5 µL/ml of Invitrogen Vybrant DiD Labelling Solution (Invitrogen) were added. Sample was left on a rocking platform for 2 h at 4°C. A second sucrose cushion purification was performed to remove excess Vybrant DiD label, after which the sample was transferred to a thin-wall Ultra-Clear 4.4 mL ultracentrifuge tube (Beckman Coulter). 1 mL of the sucrose mixture was carefully layered under the cell supernatant prior to ultracentrifugation at 150,000 xg for 2.5 h at 4°C using an SW60 rotor (Beckman Coulter). The resulting pellet was air dried for 5 min to ensure complete removal of supernatant prior to resuspension in 0.1x PBS buffer overnight at 4°C. Resuspended virus was harvested and stored at -80°C. Samples of supernatant, sucrose cushion and virus pellet were analysed by western blotting.

2.17 Confocal microscopy

2.17.1 Fixation and immunostaining of cells

Glass coverslips were washed with 70% (v/v) ethanol, followed by 5x washes in PBS. Coverslips were then placed in 12-well plates and cells were seeded 18 h prior to treatment. At indicated time-points, cells were washed 3x with PBS and were incubated with 4% (v/v) PFA for 10 min at RT. After 3x washes with PBS, cells were permeabilised with 1% (v/v) Triton-X-100 in PBS for 30 min. After 3x subsequent washes with PBS, cells were blocked with 1% (w/v) BSA in PBS for 30 min. Primary antibody diluted in 1% (w/v) BSA in PBS was added to the cells and incubated for 1 h at RT. After 3x washes with PBS, secondary antibody diluted in 1% (w/v) BSA in PBS was added to the cells and incubated for 1 h at RT in the dark. Cells were then washed 5x with PBS, carefully dried using paper towel and mounted onto slides using ProLong Gold Antifade Mountant with DAPI (Thermo Fisher). Mounting media was allowed to cure on a flat surface in the dark

for a couple of hours and slides were then stored at 4°C until examination of the specimen was performed.

2.17.2 Fluorescence *in situ* hybridisation (FISH)

Trypsinised BSR-T7 cells were seeded onto 16 mm round glass coverslips (VWR) in a 12-well plate at 8 x 10^4 cells/well and incubated at 37°C. After 16 h, cells were transfected and/or infected as specified. Media was then removed and cells were fixed and stained following the the manufacturer's instructions, available online at www.biosearchtech.com/stellarisprotocols (Sequential Stellaris FISH and Immunofluorescence using Adherent Cells protocol). Briefly, cells were fixed and stained with primary antibodies targeted to the Golgi complex or HAZV NP followed by incubation with secondary antibodies. Hybridisation of probes was performed and coverslips were mounted using Vectashield Mounting Medium. Coverglass perimeter was sealed with clear nail polish and allowed to dry.

2.17.3 Confocal microscopy

Images were then taken on a Zeiss LSM700 Inverted Microscope, a Zeiss upright LSM880, and a Zeiss LSM880 Airyscan Inverted Microscope. In all multicolour imaging, signals were acquired sequentially.

Chapter 3

Results: Production of nairoviral nucleoproteins and CCHFV NP-specific recognition molecules

3.1 Chapter introduction

Production of recombinant proteins combined with purification techniques allows the isolation of high quantities of proteins for their study, characterisation and multiple applications in basic research, biotechnology and therapeutics. Different expression systems and purification alternatives can be used for this purpose, and the appropriate method must be carefully chosen according to the characteristics of the protein of interest and the final purpose of its production.

Pure protein preparations are commonly used in basic research for analytical characterisation of protein structure and function, and serve as starting material for screening and production of highly-specific recognition molecules. This chapter describes the production and purification of two nairoviral NPs, CCHFV and HAZV NPs, and the generation of two types of recognition molecules specifically targeted to CCHFV NP, namely, the production of polyclonal antibodies and an Affimer screening with subsequent isolation of candidate binders. An initial validation of CCHFV NP-specific recognition molecules was performed by western blot and affinity precipitation (AP) assays, respectively.

CCHFV NP was selected for production due to its indispensable role as structural protein of a highly pathogenic emerging virus and its abundance within CCHFV virions, with the ultimate aim of developing diagnostic tools using the corresponding recognition

molecules to specifically detect this protein in serum samples. The closely related HAZV NP was produced and characterised alongside CCHFV NP, to test possible cross-reactivity of the recognition molecules with this homologous protein. Production of recombinant caspase 3 is also reported in this chapter, alongside its use in an *in vitro* cleavage assay to demonstrate accessibility of the DEVD and DQVD cleavage sites in the NPs, as an initial validation assay.

3.1.1 Expression of recombinant proteins

Genetic engineering and cloning techniques allow the recombinant production of proteins in large quantities, which can be further isolated using simple purification techniques. Commonly used methods for expression of recombinant proteins include bacteria, yeast, insect and mammalian systems [214]. Each method presents several advantages and limitations (detailed in table 3.1), which need to be taken into account when selecting an expression system for a specific protein of interest. Characteristics of the protein such as size, solubility or post-translational modifications, together with other considerations such as amount of protein needed, purpose of protein production, and time to be spent in protein expression must also be taken into account.

A general procedure for recombinant expression starts with the identification of the gene that encodes the protein of interest and its synthesis as cDNA. The genetic sequence is subcloned into an appropriate vector, followed by optimisation of protein expression and purification, and process scale-up, if required. *E. coli* based expression systems are the simplest and most widely used method for recombinant protein production [215]. The short doubling time of bacteria and the inexpensive media used for bacterial growth make this system fast and cheap. Scale-up is very straightforward and genetic manipulation of bacterial strains and expression vectors is easy. Some of the disadvantages of this system are the lack of post-translational modifications and the presence of other cellular proteins that need to be removed from the sample. In some cases, codon optimisation of the sequence of interest is needed due to differences of codon utilisation between bacteria and mammalian cells [214].

However, this method enables the production of high yields of recombinant protein and, when combined with purification techniques, results in high quantities of pure protein suitable for structural studies and characterisation of proteins. For these reasons, and considering previous experience and published examples in the literature [140, 151], *E. coli* expression systems were selected for production of recombinant proteins, later described in this chapter.

Expression system	Advantages	Limitations	Application
Bacterial cells	• Simple conditions • Simple scale-up • Genetic stability • Low cost • Fast growth • High yields	• Protein solubility • Protein-specific optimisation is required • Lack of post-translational modifications	• Functional/structural analyses • Antibody production • Protein interactions
Yeast cells	• Low cost • Simple media required • Scalable • Eukaryotic-type protein processing (folding and post-translational modifications)	• Fermentation required for high yields • Growth conditions may require optimisation • Mannosidation of recombinant protein	• Functional/structural analyses • Antibody production • Protein interactions
Insect cells	• Post-translational modifications (glycosylation)	• Production of recombinant vectors is time-consuming • More demanding culture conditions	• Functional/structural analyses • Expression of protein complexes • Virus production
Mammalian cells	• Post-translational modifications (glycosylation) • Transient/stable expression • Optimised systems with high-yield protein production	• More demanding culture conditions • Grams/litre yields only possible in suspension cultures	• Functional/structural analyses • Antibody production • Expression of complex proteins • Protein interactions • Virus production

Table 3.1: Protein expression systems. Advantages, limitations and applications of each system are detailed.

3.1.1.1 pET expression vectors

A broad variety of bacterial expression vectors is available for protein expression in *E. coli* cells, combining different replication origins, promoters, selection markers, multiple coning sites and fusion proteins or tags [215]. pET vectors are bacterial protein expression vectors commonly used due to their high efficiency for protein production, with expressed recombinant protein presenting up to 50% of the total cell protein content [215].

pET expression systems take advantage of the high activity and specificity of the bacteriophage T7 polymerase (pol) for regulated expression of genes in *E. coli* cells. Protein expression is under control of a T7 pol promoter and a *lac* operator (T7/*lac* promoter), and can be activated by addition of Isopropyl β-D-1-thiogalactopyranoside (IPTG), a non-hydrolyzable lactose analog. A *lac I* gene encoding the *lac* repressor is also included in the expression vector to reduce basal transcription from the T7/*lac* promoter. The T7 pol can be provided with an additional plasmid or can be placed in the genome of the bacterial expression host [215]. Genetically engineered *E. coli* BL21 strains containing the DE3 bacteriophage lambda inserted in their genome are commonly used for this purpose with the T7 pol gene under control of a *lacUV5* promoter inserted into the *int* gene of the bacteriophage. The *lac I* gene encoding for the *lac* repressor is also inserted in the bacterial genome to reduce basal transcription of the T7 pol (figure 3.1).

In the absence of an inducer, the repressor encoded by the *lac I* gene in both the *E. coli* genome and the pET vector acts as a dual inhibitor. This repressor prevents the transcription of the T7 pol by the *E. coli* RNA polymerase, impeding, at the same time, the transcription of the gene of interest in the expression vector by the T7 pol. When an inducer such as IPTG is added to the media, it binds to the repressor and inhibits its activity, inducing gene expression. As a result, T7 pol is produced and drives the expression of the genes of interest in the pET plasmid under its control (figure 3.1). Most pET vectors contain a pBR322 origin, allowing replication and maintenance in *E. coli*. An antibiotic resistance gene is included for selection of *E. coli* containing the plasmid of interest, and a multiple cloning site with different enzyme restriction sites for subcloning of the gene of interest.

In this chapter, different pET vectors are used depending on the protein to be purified and the purity required. pET-SUMO vector was used for production of native nairoviral NPs and untagged Affimers, pET-11a vector was selected for His-tagged Affimers expression and pET-21 was used for His-tagged caspase 3 production.

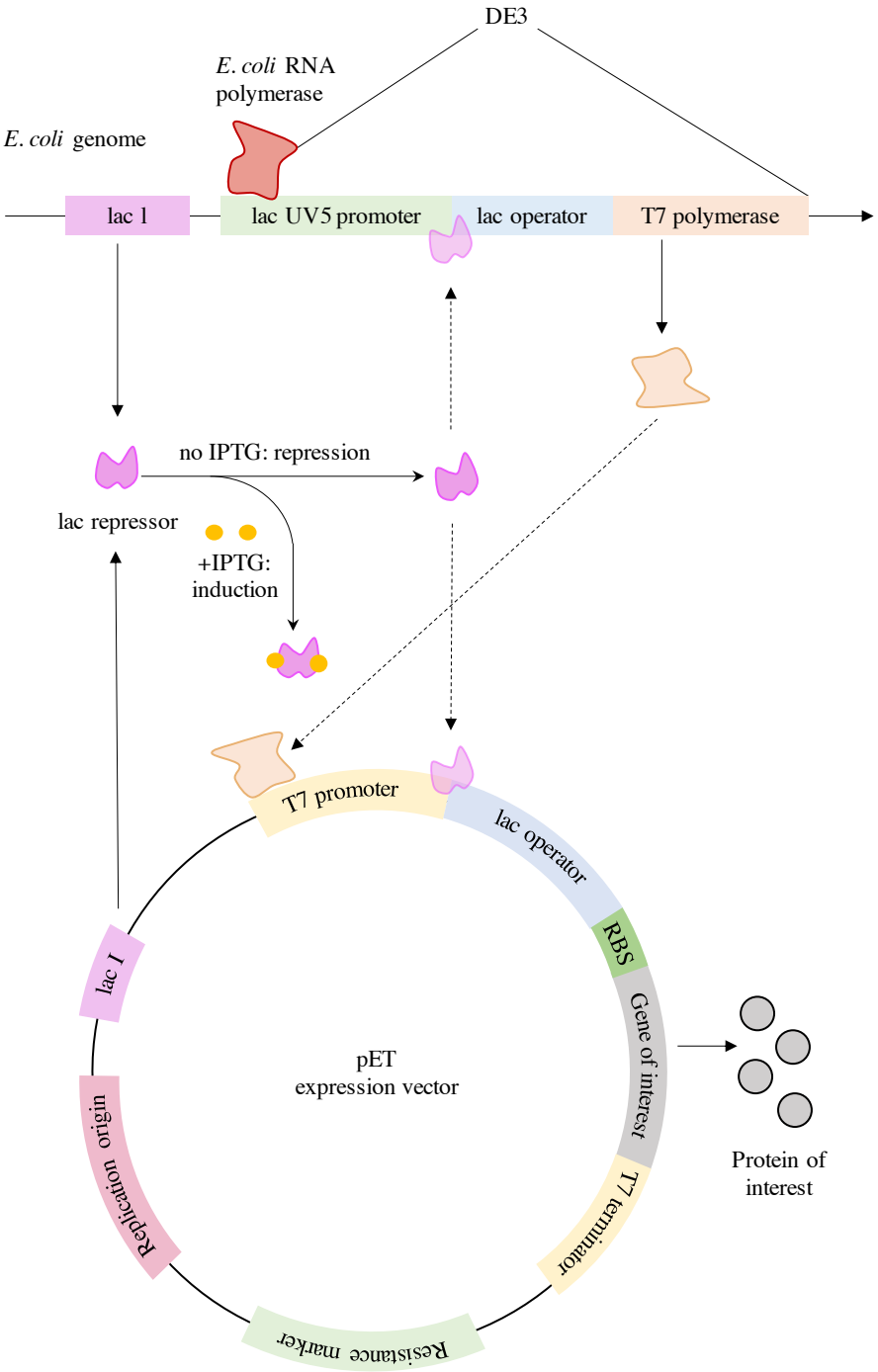


Figure 3.1: Recombinant protein expression system combining pET expression vectors and BL21(DE3) *E. coli* strains. *E. coli* genome is represented at the top, containing a *lac I* gene and a T7 pol gene under control of a *lacUV5* promoter and a *lac* operator. A pET expression vector is also represented with its characteristic elements. If no IPTG is added, *lac* repressor interferes with transcription and expression of T7 pol and the gene of interest are repressed. If IPTG is added, it binds to the repressor inhibiting its activity. As a result, T7 pol is produced and drives the expression of the genes of interest in the pET plasmid under its control. (RBS, ribosome binding site. IPTG, Isopropyl β-D-1-thiogalactopyranoside.)

pET-SUMO vectors contain a kanamycin resistance gene and an N-terminus tagging system. They include a yeast SUMO (small ubiquitin-like modifier) tag (smt3, *Saccharomyces cerevisiae*) as N-terminal fusion partner to the ORF of the gene of interest, improving protein stability and solubility. An additional poly-histidine (6xHis) tag is included in the N-terminal end of the smt3 SUMO tag, allowing purification of the fusion protein by affinity chromatography. Once the fusion protein is produced and purified, cleavage mediated by SUMO ubiquitin-like specific protease 1 (Ulp1) results in separation of the tag from the native protein. Smt3 SUMO fusion protein and SUMO Ulp1 protease contain N-terminal poly-histidine (6xHis) tags that allow their removal from the cleavage reaction using affinity chromatography.

pET-11 and pET-21 vectors contain an ampicillin resistance gene. pET-11 originally does not contain a fusion tag [216], but a modified version of a pET-11a vector, which introduces an 8xHis C-terminal tag using *Nde*I and *Nhe*I restriction sites to insert the gene of interest, was used in this chapter for Affimers expression. pET-21 introduces a C-terminal 6xHis tag in the ORF of the gene of interest, and was commercially obtained from Addgene with the caspase 3 ORF already inserted in the vector [217].

3.1.2 Purification of recombinant proteins

The aim of a protein purification process is not only removal of unwanted material but also concentration of the protein of interest and its transfer to a stable and desired environment. Different procedures can be used and combined for protein purification, including chromatographic techniques, electrophoresis, AP or dialysis. The methodology used for each particular protein must be chosen according to protein behaviour and the final purpose for which the protein is needed.

Chromatographic techniques have become one of the most effective and widely used methods for protein purification [218]. They include a group of separation techniques that involve the physical retardation of molecules due to their interaction with a specific material (stationary phase) with respect to their solvent. The most common experimental configuration consists of a stationary phase packed into a column through which the mobile phase is pumped. The degree to which the molecules interact with the stationary phase determines how fast they will be carried by the mobile phase.

Different physicochemical properties of proteins such as size, charge, hydrophobicity and bio-specific interactions can be used to isolate them from other contaminants. Several types of chromatography can be distinguished depending on the

properties of the protein exploited for its purification and the experimental set up. Some examples include ion-exchange chromatography, affinity chromatography, size exclusion chromatography (SEC), or gas chromatography [219].

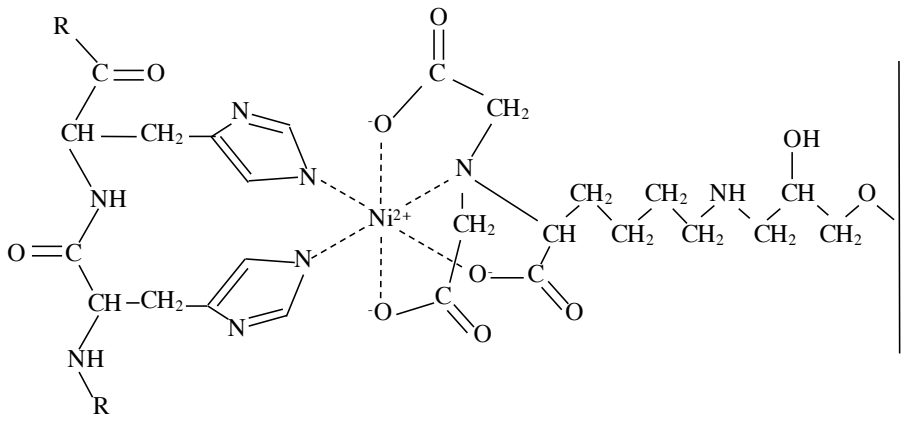
In this chapter, two different chromatographic techniques, affinity and size exclusion chromatography, were used for purification of bacterially expressed recombinant proteins from all other host cellular proteins.

3.1.2.1 Immobilised metal affinity chromatography (IMAC)

Affinity purification of proteins exploits specific interactions between the protein of interest and a specific ligand attached to a matrix, including electrostatic or hydrophobic interactions, van der Waals forces and hydrogen bonds. The binding interaction must be reversible and allow elution of the protein of interest without affecting its biochemical properties. A widely used method for affinity chromatography purification of recombinant proteins is immobilised metal affinity chromatography (IMAC), which involves the use of peptide affinity tags fused to the protein of interest to facilitate its interaction with a metal ligand on the stationary phase [220]. Histidine tails are commonly used as affinity tags due to their strong interaction with immobilised transition metal ion matrices (Co^{2+}, Ni^{2+}, Cu^{2+}, Zn^{2+}).

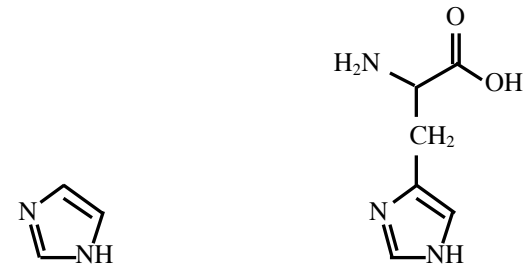
Proteins containing a poly-histidine tail are retained on the metal ion matrix due to the formation of coordination bonds between electron donor groups on the histidine imidazole ring and the immobilised transition metal (figure 3.2a). After washing off non-bound material, the poly-histidine tagged proteins of interest can be eluted by either adjusting the pH or adding free imidazole to the column buffer. A decrease on the pH to 4-5 causes protonation of histidine residues interfering with the interaction between these and the metal ion, whereas imidazole has a chemical structure very similar to histidine (figure 3.2b) and forms coordination bonds with the metal ions, acting as a competitive binder for the resin.

Beaded agarose is the most commonly used material for stationary phases in IMAC. A suitable spacer arm plus a chelator which will coordinate to the metal ion, such as nitrilotriacetic acid (NTA), are incorporated to the matrix (figure 3.2a). Some of the coordination bonds of the metal must remain free after binding to the chelator, so the electron-donor groups on the target protein can access them. The number of histidines interacting with the matrix is directly related with the number of available coordinator sites in the matrix and the number of histidine residues in the protein of interest. Some of the most common clonable tags used for this procedure are 6xHis and 8xHis tails, often



Poly-histidine peptide Nickel-nitriloacetic acid (Ni^{2+}-NTA) resin

(a)



Imidazole Histidine

(b)

Figure 3.2: Chemistry of Ni^{2+}-NTA affinity chromatography. **(a)** Schematic representation of nickel (2+) ion coordination bonds with the poly-histidine peptide in the protein of interest and the nitrilotriacetic acid on the resin. **(b)** Chemical structures of imidazole and histidine. Due to the structural similarities between their aromatic groups, imidazole acts as competitor of histidine for Ni^{2+}-NTA resin binding.

resulting in purified samples with over 90% purity [218].

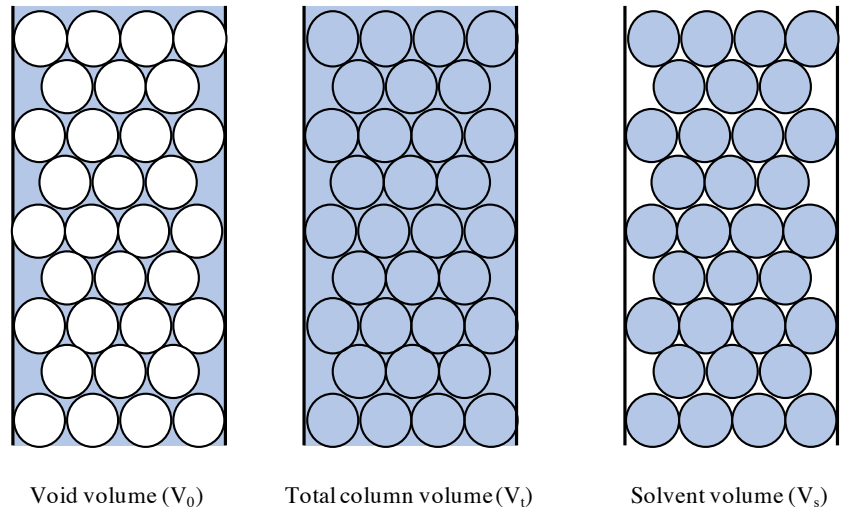
3.1.2.2 Size exclusion chromatography

Size exclusion chromatography (SEC) is based on the separation of molecules according to their size and shape due to their interaction with a porous matrix. The ability of proteins to enter all, some, or none of the channels in the porous beads determines their elution volume. Smaller molecules can access smaller pores in the beads and are retarded compared to larger molecules which, for steric reasons, are excluded from the pores and pass quickly between the beads. Differential exclusion or inclusion within porous particles also depends on the shape of the molecules. Linear molecules, such as DNA or polysaccharides, tend to have larger hydrodynamic volumes than globular molecules, such as most proteins [218].

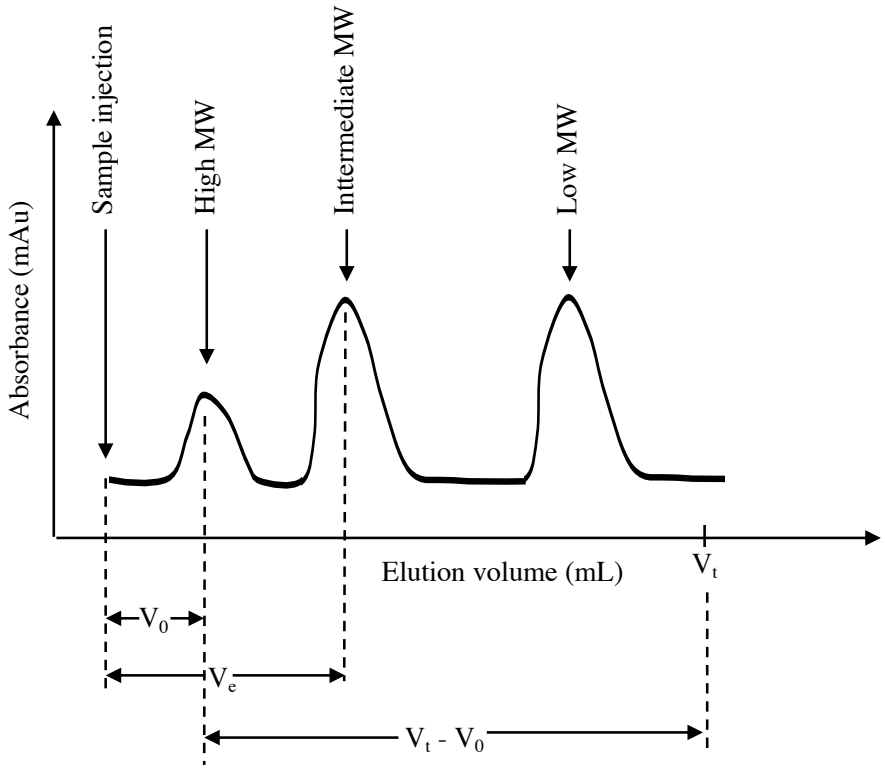
Selectivity of SEC depends on the inherent porosity of the material. Commonly used matrices include natural polymers, such as agarose or dextran, or synthetic ones, such as polyacrylamide. Gels can be formed by cross-linking these polymers, resulting in three-dimensional networks [218]. Different pore sizes can be obtained depending on the degree of cross-linking used for the manufacture of the resins. The first commercial SEC stationary phase, Sephadex, is composed of dextran cross-linked with epichlorohydrin.

Specific terminology is employed to describe SEC experimental concepts and understand SEC chromatograms, including terms like void volume (V_0), solvent volume (V_s), total volume (V_t) and elution volume (V_e) (figure 3.3) [221]. Void volume (V_0) is defined as the space taken by solvent surrounding the gel beads. Solvent volume (V_s) refers to space occupied by solvent inside the medium particles. Total volume (V_t) is the sum of V_0 and V_s. The elution volume (V_e) of a certain molecule is the volume of mobile phase entering the column between the sample injection and the elution of the molecule. In the case of proteins, elution from the column can be monitored by measuring absorbance at 280 nm with an ultraviolet (UV) light lamp. Results of SEC are normally represented in chromatograms showing the variation in protein concentration (absorbance at 280 nm) versus elution volume (mL) (figure 3.3b).

Some of the most common applications of SEC include buffer exchange, where proteins are excluded and low molecular weight components such as salts are retained in the column; protein fractionation, for separation of molecules with different sizes and shapes; and qualitative characterisation of proteins, such as determination of molecular size, oligomeric state or study of protein interactions.



(a)



(b)

Figure 3.3: Terminology of size exclusion chromatography and representation of a chromatogram. **(a)** Visual representation of SEC terminology. Void volume (V_0) is defined as the space taken by solvent surrounding the gel beads. Solvent volume (V_s) refers to space occupied by solvent inside the medium particles. Total volume (V_t) is the sum of V_0 and V_s. **(b)** Schematic representation of a SEC chromatogram. High molecular weight (MW) molecules elute before intermediate MW molecules, and these before low MW molecules. V_0, V_s and V_t are represented.

3.2 Expression and purification of nairoviral nucleoproteins

3.2.1 CCHFV NP expression and purification

CCHFV NP was recombinantly expressed in Rosetta 2 (DE3) *E. coli* cells. Cells were transformed with a pET-SUMO plasmid containing the CCHFV NP ORF (see plasmid map in Appendix B) and protein expression was induced with IPTG. Cell lysates were separated into soluble and insoluble fractions, and samples were analysed by SDS-PAGE and Coomassie staining (figure 3.4a). 6xHis-SUMO-CCHFV NP was mainly present in the soluble fraction, which was further used for purification of CCHFV NP.

3.2.1.1 Ni^{2+}-NTA affinity chromatography

6xHis-SUMO-CCHFV NP was purified from soluble fraction of cell lysate by Ni^{2+}-NTA affinity chromatography. Washes with increasing imidazole concentration buffers were done to remove non-specifically bound proteins. Elution of 6xHis-SUMO-CCHFV NP was achieved with 500 mM imidazole buffer.

Samples were analysed by SDS-PAGE and Coomassie staining (figure 3.4a). A protein band corresponding to full length 6xHis-SUMO-CCHFV NP (67 kDa) was observed in the elution fractions, alongside a contaminant protein band of $\sim$20 kDa corresponding to an unknown protein product (highlighted with an asterisk in figure 3.4a). Elution fractions containing 6xHis-SUMO-CCHFV NP were pooled and 6xHis-SUMO tag was proteolytically removed by addition of 6xHis-SUMO Ulp1 protease. SDS-PAGE analysis and Coomassie staining of the cleaved sample revealed a 27 kDa protein band corresponding to 6xHis-SUMO Ulp1 protease and two other protein bands of 54 kDa and 13 kDa corresponding to CCHFV NP and 6xHis-SUMO fusion tag cleavage products, respectively (figure 3.4b). The aforementioned contaminant product of 20 kDa was not present in the sample after cleavage with 6xHis-SUMO Ulp1 protease.

A second Ni^{2+}-NTA affinity chromatography purification step was performed to remove the 6xHis-SUMO fusion tag and the 6xHis-SUMO Ulp1 protease. Briefly, sample was applied to Ni^{2+}-NTA resin and flow-through containing CCHFV NP was collected. The remaining sample in the column was eluted with 500 mM imidazole buffer and all fractions were analysed by SDS-PAGE and Coomassie staining (figure 3.4b). After the second purification step, there was a clear reduction in the amount of 6xHis-SUMO tag and 6xHis-SUMO Ulp1 protease present in the flow-through sample.

3.2.1.2 Size exclusion chromatography

As not all the 6xHis-SUMO tag and protease were removed, SEC was performed to further purify CCHFV NP using an S75 sephadex column. Successful separation of the proteins present in the sample was achieved due to their distinct molecular weights. Three different peaks were observed in the size exclusion chromatogram, corresponding to 6xHis-SUMO tag (13 kDa), 6xHis-SUMO Ulp1 protease (27 kDa) and CCHFV NP (54 kDa) (figure 3.5a). Fractions corresponding to CCHFV NP peak were analysed by SDS-PAGE gel and Coomassie staining (figure 3.5b), and showed efficient removal of non-desired products. The ratio of absorbance values at 260 nm versus 280 nm was measured and determined as A260/A280$<$1 in purified CCHFV NP fractions. Fractions containing the protein of interest (corresponding to lanes 3-13 in figure 3.5b) were pooled and concentrated to 1 mg/mL.

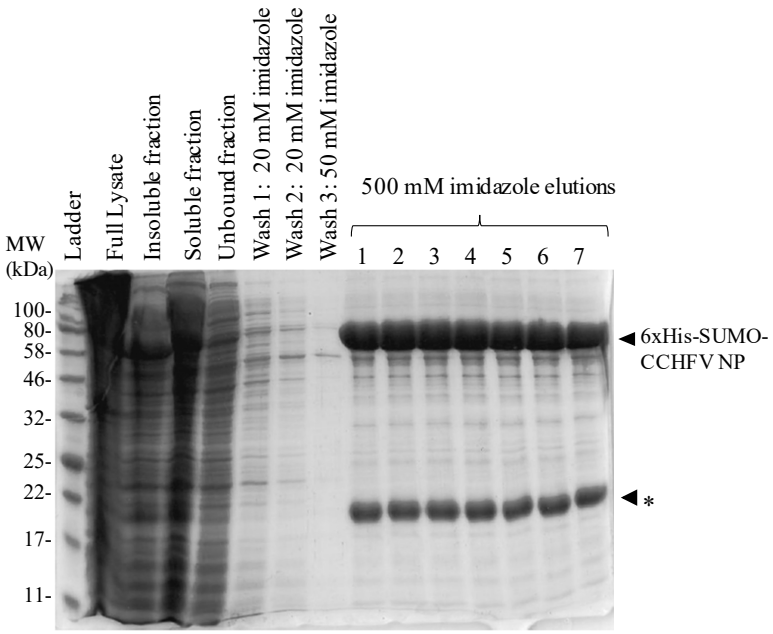
3.2.2 HAZV NP expression and purification

Recombinant HAZV NP was expressed using a similar procedure to the above in Rosetta 2 (DE3) *E. coli* cells. Cells were transformed with a pET-SUMO plasmid containing the HAZV NP ORF (see plasmid map in Appendix B) and protein expression was induced with IPTG. Cell lysates were separated into soluble and insoluble fractions, and samples were analysed by SDS-PAGE and Coomassie staining (figure 3.6a). A protein band corresponding to 6xHis-SUMO-HAZV NP (67 kDa) was present in soluble and insoluble fractions.

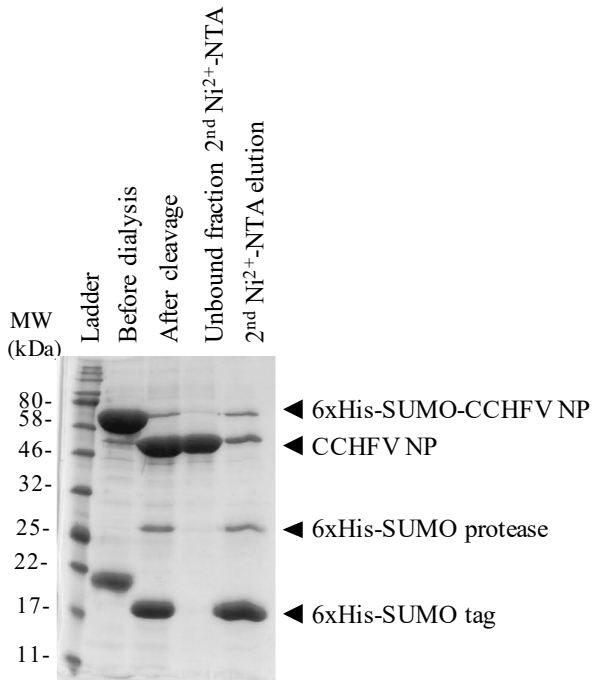
3.2.2.1 Ni^{2+}-NTA affinity chromatography

6xHis-SUMO-HAZV NP was purified from soluble fraction of cell lysate using Ni^{2+}-NTA affinity chromatography. Washes with low imidazole concentration buffers were done to remove non-specifically bound proteins, followed by a high-salt wash to remove any potentially bound nucleic acids. 6xHis-SUMO-HAZV NP was eluted with 300 mM imidazole buffer and samples were analysed by SDS-PAGE and Coomassie staining (figure 3.6a). A protein band corresponding to full length 6xHis-SUMO-HAZV NP (67 kDa) was observed in the elution fractions, alongside a contaminant double protein band of $\sim$20 kDa (highlighted with an asterisk in figure 3.6a).

Fractions containing 6xHis-SUMO-HAZV NP were pooled and 6xHis-SUMO tag was proteolytically removed by addition of 6xHis-SUMO Ulp1 protease. SDS-PAGE analysis and Coomassie staining of cleaved sample revealed a 27 kDa protein band

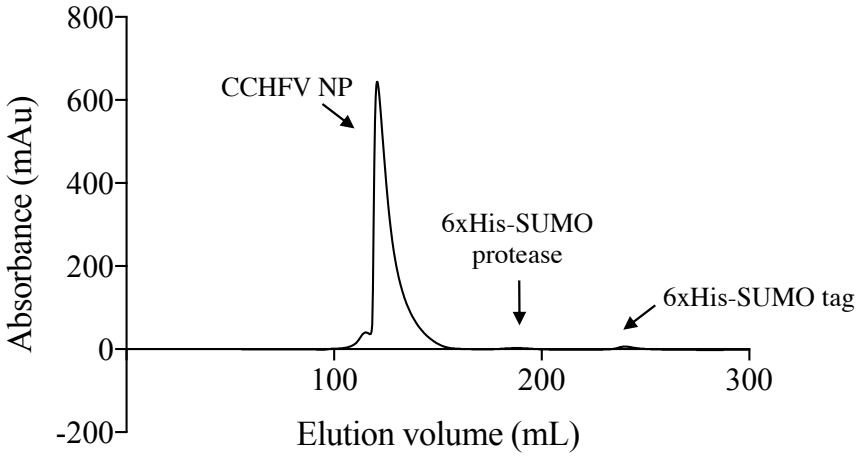


(a)

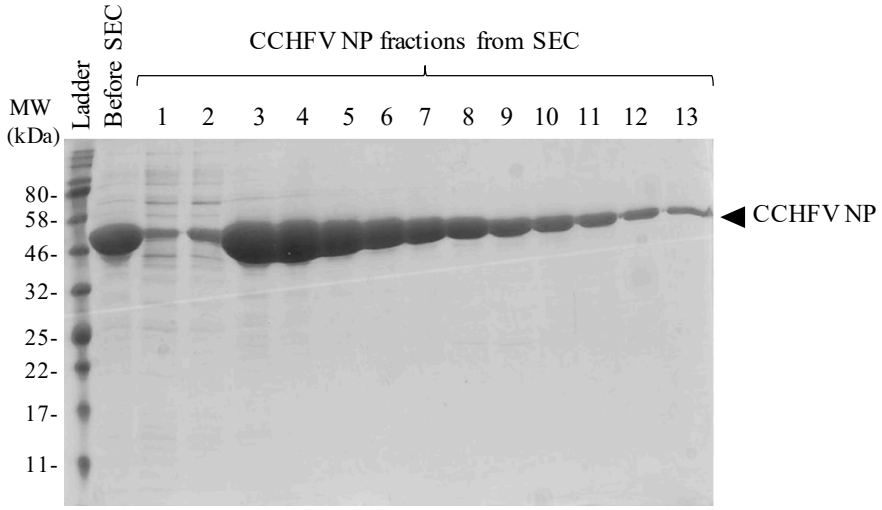


(b)

Figure 3.4: CCHFV NP purification by Ni^{2+}-NTA affinity chromatography. **(a)** SDS-PAGE analysis and Coomassie staining of fractions from different steps of the expression and purification processes of 6xHis-SUMO-CCHFV NP by Ni^{2+}-NTA affinity chromatography. **(b)** SDS-PAGE analysis and Coomassie staining of purified 6xHis-SUMO-CCHFV NP after dialysis and cleavage with 6xHis-SUMO Ulp1 protease, and a second purification step by Ni^{2+}-NTA affinity chromatography. Arrows indicate migration distances of specified proteins. * indicates an unknown protein product.



(a)



(b)

Figure 3.5: CCHFV NP purification by size exclusion chromatography. **(a)** Chromatogram of the SEC purification of CCHFV NP sample after the second Ni^{2+}-NTA affinity chromatography. **(b)** SDS-PAGE analysis and Coomassie staining of SEC fractions containing CCHFV NP. Arrows indicate elution volume and migration distances of specified proteins.

corresponding to 6xHis-SUMO Ulp1 protease and two other protein bands of 54 kDa and 13 kDa corresponding to HAZV NP and 6xHis-SUMO fusion tag cleavage products, respectively. 6xHis-SUMO-HAZV NP cleavage was incomplete, with uncleaved protein remaining present in sample after cleavage (figure 3.6b). The contaminant product of 20 kDa was not present after cleavage with 6xHis-SUMO Ulp1 protease.

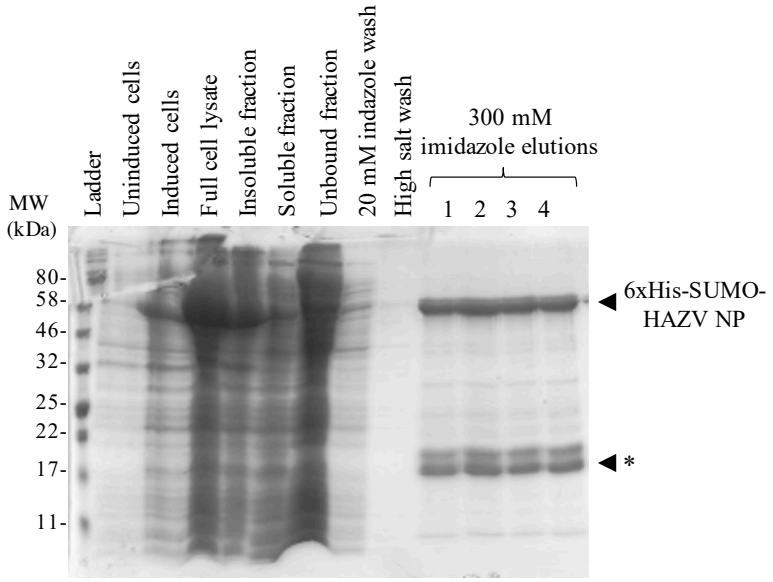
A second Ni^{2+}-NTA purification step was performed to remove uncleaved protein, 6xHis-SUMO tag and 6xHis-SUMO Ulp1 protease. Briefly, sample was applied to Ni^{2+}-NTA resin and flow-through containing HAZV NP was collected. The remaining sample in the column was eluted with 300 mM imidazole and all fractions were analysed by SDS-PAGE and Coomassie staining (figure 3.6b). After the second Ni^{2+}-NTA affinity chromatography purification step, most 6xHis-SUMO Ulp1 protease and 6xHis-SUMO tag were successfully removed, but a significant amount of uncleaved 6xHis-SUMO-HAZV NP remained present in the sample.

3.2.2.2 Size exclusion chromatography

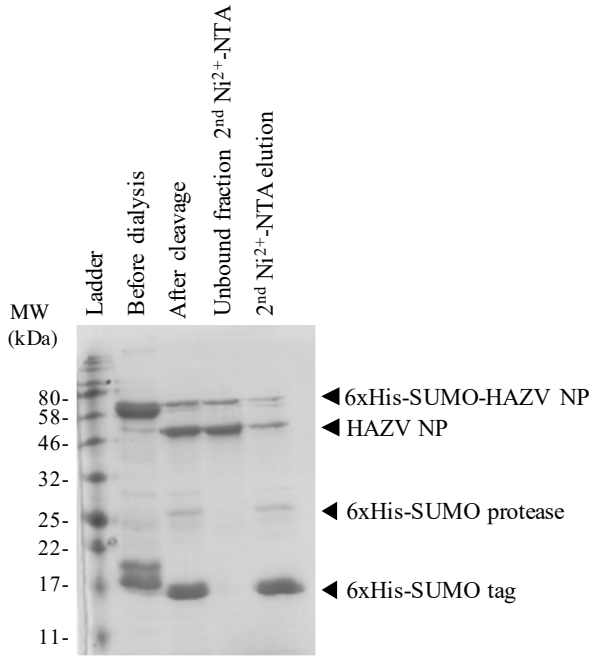
To further purify HAZV NP, SEC was performed using an S75 sephadex column. Successful separation of the proteins present in the sample was achieved due to their different molecular sizes. The size exclusion chromatogram presented four peaks corresponding to 6xHis-SUMO tag (13 kDa), 6xHis-SUMO Ulp1 protease (27 kDa), HAZV NP (54 kDa) and 6xHis-SUMO-HAZV NP (67 kDa) (figure 3.5a). There was an additional and considerably large peak with a migration volume corresponding to the void volume of the column (110 mL). Fractions corresponding to HAZV NP and 6xHis-SUMO-HAZV NP were analysed by SDS-PAGE and Coomassie staining (figure 3.5b). The ratio of absorbance values at 260 nm versus 280 nm was measured and determined as A260/A280$<$1 in purified HAZV NP fractions. Fractions containing the protein of interest (corresponding to lanes 11-18 in figure 3.7b) were pooled and concentrated to 1 mg/mL.

3.2.3 Caspase 3 cleavage validation

In vitro cleavage assays can be performed using recombinant caspases, which are active under appropriate buffer conditions. Caspase 3 is known to cleave CCHFV and HAZV NPs during the course of viral infection and hence, a caspase 3-mediated cleavage assay was performed to initially validate the folding and accessibility of the caspase 3 cleavage site in the purified recombinant nairoviral NPs. Recombinant caspase

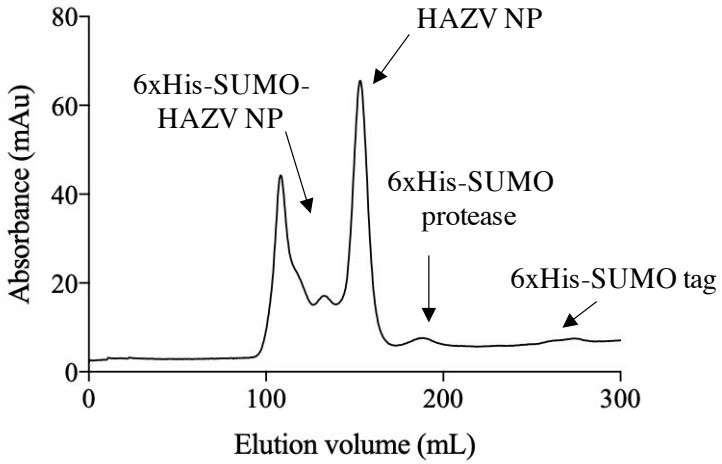


(a)

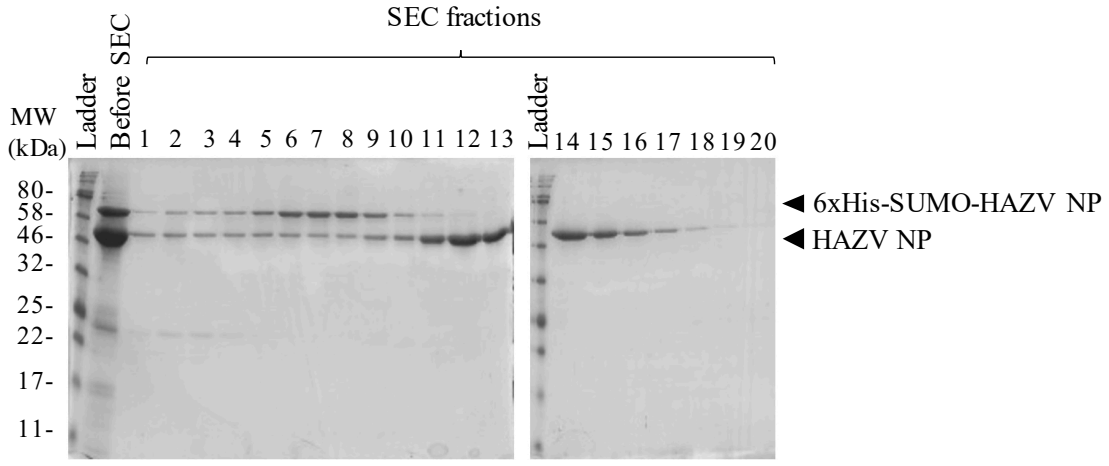


(b)

Figure 3.6: HAZV NP purification by Ni^{2+}-NTA affinity chromatography. **(a)** SDS-PAGE analysis and Coomassie staining of fractions from different steps of the expression and purification processes of 6xHis-SUMO-HAZV NP by Ni^{2+}-NTA affinity chromatography. **(b)** SDS-PAGE analysis and Coomassie staining of purified 6xHis-SUMO-HAZV NP after dialysis and cleavage with 6xHis-SUMO Ulp1 protease, and a second purification step by Ni^{2+}-NTA affinity chromatography. Arrows indicate migration distances of specified proteins. * indicates an unknown protein product.



(a)



(b)

Figure 3.7: HAZV NP purification by size exclusion chromatography. **(a)** Chromatogram of SEC purification of HAZV NP sample after the second Ni^{2+}-NTA affinity chromatography. **(b)** SDS-PAGE analysis and Coomassie staining of SEC fractions containing HAZV NP. Arrows indicate elution volume and migration distances of specified proteins.

3 was expressed in Rosetta 2 (DE3) *E. coli* cells and purified by Ni^{2+}-NTA affinity chromatography as previously described [222]. Briefly, Rosetta 2 (DE3) *E. coli* cells were transformed with a plasmid encoding the full length caspase 3 protein fused to a C-terminal 6xHis tag (pET21b-caspase-3, Addgene, see plasmid map in Appendix B). Protein expression was induced with IPTG and cell lysates were separated into soluble and insoluble fractions for purification of soluble caspase 3.

3.2.3.1 Caspase 3 purification by Ni^{2+}-NTA affinity chromatography

The soluble fraction containing full length caspase 3 (35 kDa) was applied to Ni^{2+}-NTA sepharose resin. Non-specifically bound proteins were removed with washes of increasing imidazole concentration buffers and elution of caspase 3 was achieved with 200 mM imidazole buffer. Fractions containing purified caspase 3 were pooled and concentrated to 1 mg/mL. All fractions were analysed by SDS-PAGE and Coomassie staining (figure 3.8). Induced caspase 3 lysates showed a predominant 35 kDa protein band corresponding to full length caspase 3, whereas elution fractions presented two protein bands corresponding to the large (p17) and small (p12) subunits yielded by auto-processed cleavage of caspase 3 [223].

3.2.3.2 Caspase 3 cleavage assay

Recombinant CCHFV and HAZV NPs were incubated overnight with purified caspase 3 in either equimolar or 2:1 ratios of caspase 3:NP. Products of digestion were separated by SDS-PAGE and identified by western blotting using CCHFV NP-specific polyclonal IgGs (described in section 3.3) or HAZV NP-specific antisera, respectively. Cleaved caspase 3 antisera was used to confirm the presence of active protease.

Recombinant caspase 3 mediated cleavage of both nairoviral NPs. SDS-PAGE and western blot analyses of the corresponding samples revealed the presence of two prominent NP cleavage products having approximate molecular masses of $\sim$20 kDa and 30 kDa (figure 3.9). These corresponded precisely to the apparent masses of the CCHFV and HAZV NPs products, previously identified in infected mammalian cells [222, 152].

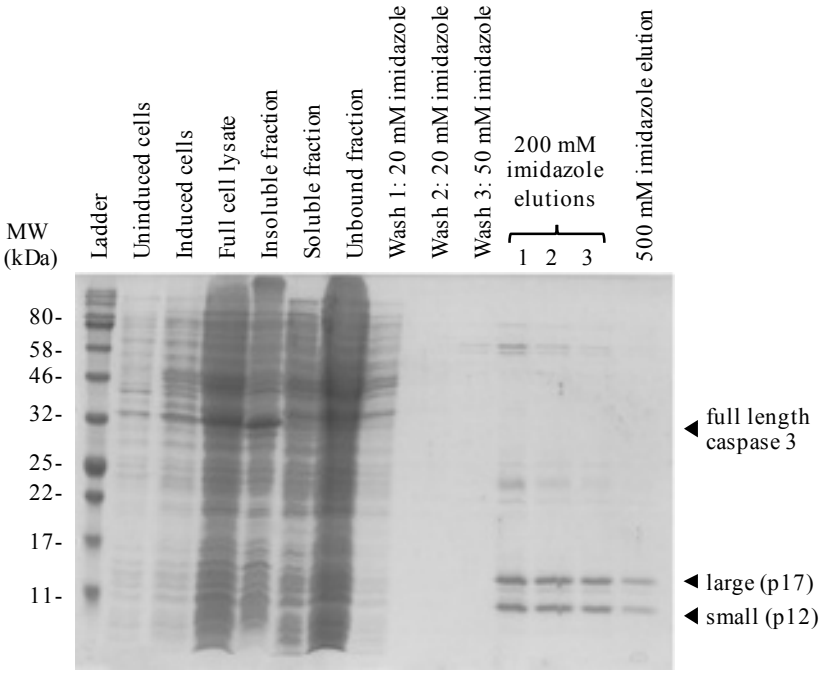


Figure 3.8: Caspase 3 purification by Ni^{2+}-NTA affinity chromatography. SDS-PAGE analysis and Coomassie staining of fractions from different steps of the expression and purification processes of caspase 3 by Ni^{2+}-NTA affinity chromatography. Caspase 3 is initially expressed as full length, and subsequently auto-processed yielding large (p17) and small (p12) subunits. Arrows indicate elution volume and migration distances of specified proteins.

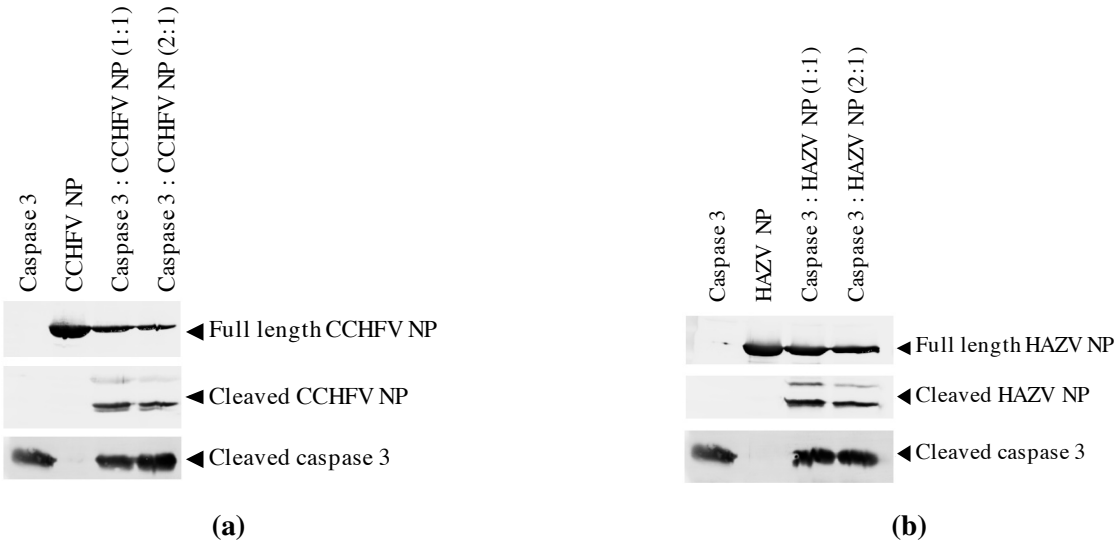


Figure 3.9: Caspase 3 mediated cleavage of recombinant CCHFV and HAZV NPs. Western blot analysis of the caspase 3 cleavage assay of recombinant CCHFV **(a)** or HAZV NP **(b)**. Anti-CCHFV NP polyclonal antibodies, anti-HAZV NP antisera and anti-cleaved caspase 3 antibodies were used as primary antibodies.

3.3 CCHFV NP-specific antibody generation and characterisation

Recombinantly expressed and purified CCHFV NP was used for the generation of polyclonal antibodies. Briefly, lyophilised recombinant CCHFV NP was sent to Genscript, a company specialised in custom polyclonal antibody services. Two rabbits were inoculated with the protein, and the immunisation schedule consisted of a primary immunisation plus three boosts every two weeks. Three months after primary immunisation of rabbits, a production bleed was carried out. Polyclonal IgGs were purified by the company from the serum of the rabbits using protein A affinity purification. A total volume of 86 mL of antibody preparation at a concentration of 6.69 mg/mL were obtained from the company.

CCHFV NP-specific rabbit polyclonal IgGs were first characterised by western blotting to test their specificity and sensitivity. Briefly, different amounts of recombinant CCHFV NP were analysed by SDS-PAGE and Coomassie staining or western blot. A sample of myoglobin (Mb) was included as negative control. Western blotting was performed using CCHFV NP-specific rabbit polyclonal IgGs as primary antibody, successfully detecting from 10 to 100 ng of CCHFV NP and showing lack of reactivity to myoglobin (figure 3.10a).

To study the specificity of the rabbit polyclonal IgGs, SDS-PAGE and western blot analysis was performed using 1µg of recombinant CCHFV NP and 1µg of its homologue HAZV NP. CCHFV NP-specific rabbit polyclonal IgGs and HAZV NP-specific goat antisera were separately tested as primary antibodies. CCHFV NP-specific polyclonal IgGs showed cross-reactivity for CCHFV and HAZV NPs (figure 3.10b), presenting a stronger signal for CCHFV NP than for HAZV NP. Cross-reactivity was also observed when HAZV NP-specific sheep antisera was used as primary antibody (figure 3.10c), with a stronger signal obtained for HAZV NP.

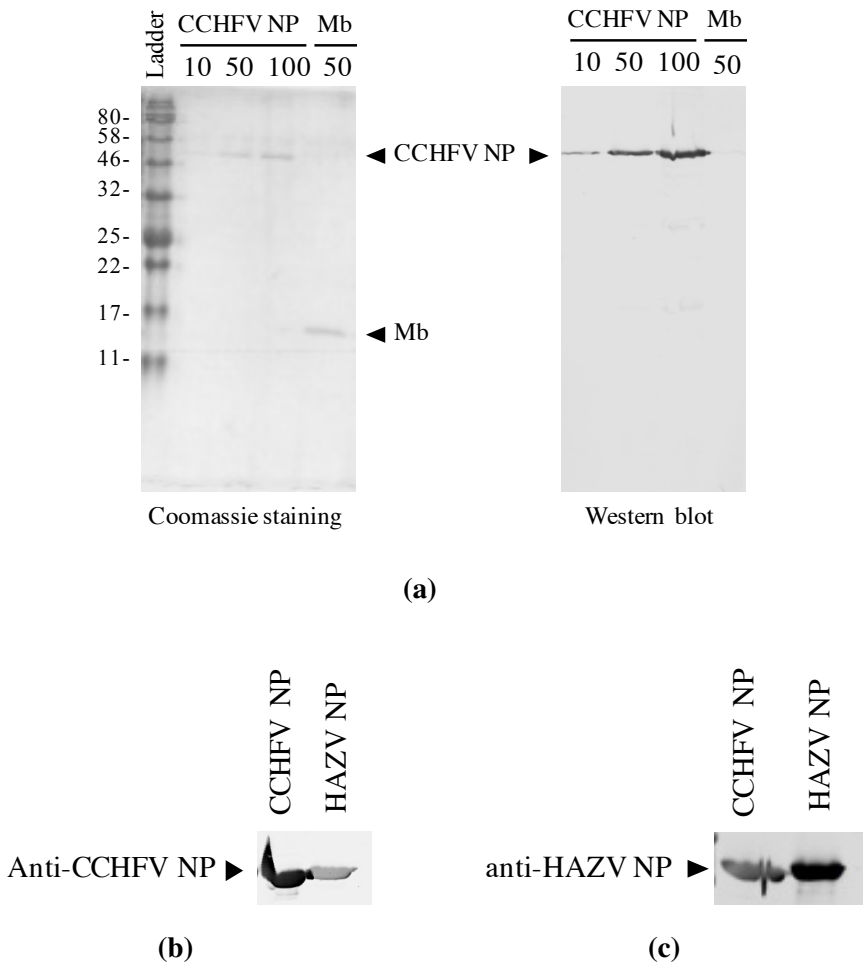


Figure 3.10: Use of CCHFV NP-specific polyclonal antibodies for western blot analysis. **(a)** SDS-PAGE analysis followed by Coomassie staining (left) or western blot analysis (right) of different amounts of recombinant CCHFV NP and myoglobin (Mb). CCHFV NP-specific rabbit polyclonal IgGs were used as primary antibody for western blot. Numbers above lanes indicate amount of protein in ng. **(b-c)** SDS-PAGE and western blot analysis of recombinant CCHFV and HAZV NPs using CCHFV NP-specific polyclonal antibodies **(b)** or HAZV NP-specific antisera **(c)** as primary antibodies, respectively.

3.4 CCHFV NP-specific Affimers production and selection

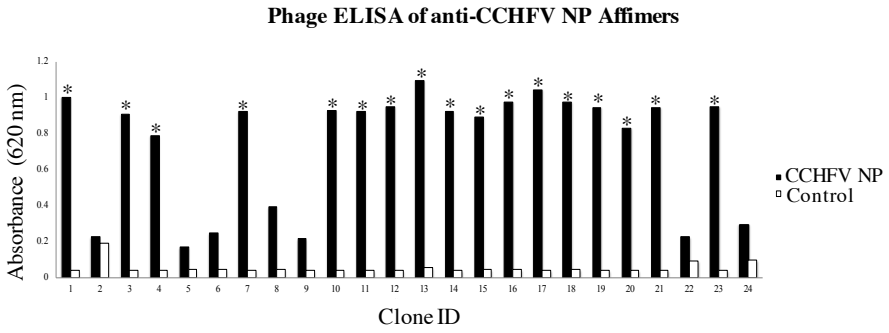
Purified CCHFV NP was provided to the BioScreening Technology Group (BSTG, University of Leeds) for an Affimer library screening of CCHFV NP-specific binders, as previously described [200]. After 3 panning rounds of selection, a 'phage ELISA was performed to test the specificity of 24 isolated binders compared to a control empty well (figure 3.11a).

Eight different sequences corresponding to unique binders were obtained (Affimers 1, 2, 5, 6, 8, 9, 22 and 24). Clone number 2 was discarded due to unspecific binding to control wells with similar levels to CCHFV NP coated wells, resulting in a final set of seven unique binders of interest. One of the hits was present 17 times (Affimer-1, highlighted with asterisks in figure 3.11a) suggesting a distinct reactivity for CCHFV NP.

Alignment of the seven isolated sequences showed hydrophobic residues in positions 3, 4 and 6 of loop 1 for all unique binders (figure 3.11b). 80% of the sequences presented hydrophobic residues in positions 1, 2, 5, and 6 of the first loop, and positive residues in positions 3 and 4. These findings suggest a tendency of loop 1 to present hydrophobic residues, likely binding to a hydrophobic region of CCHFV NP. For Loop 2, a predominant physicochemical tendency of the amino acid residues was not identified.

3.4.1 Subcloning of CCHFV NP-specific Affimers into a bacterial expression vector

The ORFs of the seven isolated Affimers were subcloned into bacterial expression vectors for recombinant expression. Briefly, PCR-amplified Affimer ORFs were subcloned into pET-11a vectors by restriction enzyme digestion and ligation (figure 3.12), and corresponding plasmids were analysed by Sanger sequencing (figure 3.11c). A unique cysteine was incorporated in the C-terminal region of the Affimer ORFs by means of the reverse PCR primer, and a C-terminal 8xHis tag encoded by the pET-11a expression vector was kept in frame with the Affimer ORFs for downstream affinity chromatography purification of the corresponding protein products (see plasmid map in Appendix B).

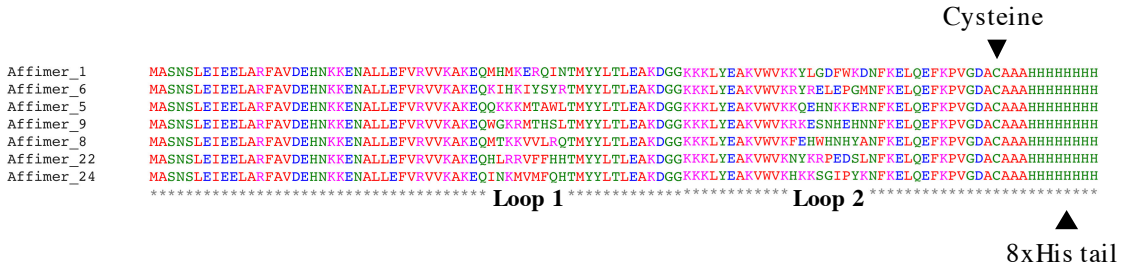


(a)

	Loop 1	Loop 2
100% consensus	..hh.h...	....
90% consensus	..hh.h...	....
80% consensus	hh++hh...	p.ct .....

Abbreviation	Physicochemical property	Amino acids included
c	charged	D, E, H, K, R
h	hydrophobic	A, C, F, G, H, I, K, L, M, R, T, V, W, Y
p	polar	C, D, E, H, K, N, Q, R, S, T
+	positive	H, K, R
s	small	A, C, D, G, N, P, S, T, V
t	turn-like	A, C, D, E, G, H, K, N, Q, R, S, T

(b)



(c)

Figure 3.11: 'Phage ELISA, alignment and consensus residues of CCHFV NP-specific Affimer sequences. (a) CCHFV NP 'phage ELISA results after 3 consecutive panning rounds of selection. Histogram representation of absorbance at 620 nm for 24 monoclonal 'phages. Results from neutravidin wells coated with biotinylated CCHFV NP are represented in black and control blank wells are represented in white. Asterisks (*) indicate hits containing Affimer-1-'phages, present 17 times. **(b)** Consensus alignment of the amino acid sequence of loops 1 and 2 from seven unique binders obtained from 'phage-display library screening. Table indicates physicochemical properties analysed, abbreviations used and amino acids included in each of the groups analysed. **(c)** Alignment of the amino acid sequences of the seven isolated binders subcloned into pET-11a expression vectors. Arrows indicate carboxyl-terminal cysteines and 8xHis tails. The multiple sequence alignment was generated using Clustal omega and visualised using Jalview.

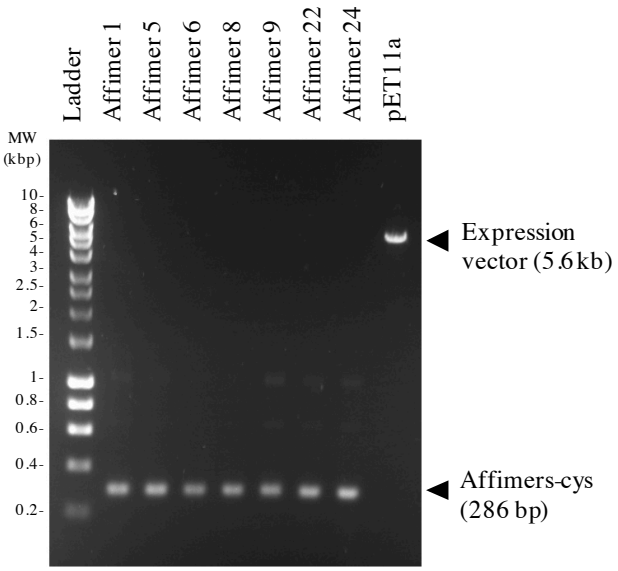


Figure 3.12: Digested PCR products and expression vector for Affimers subcloning. UV transillumination imaging of agarose gel electrophoresis analysis of the 7 Affimer ORFs and the pET-11a vector after *Nhe*I and *Not*I restriction enzymes digestion. DNA was stained using SYBR safe DNA stain. Arrows indicate migration distances, name and size of the corresponding DNA products.

3.4.2 Expression and purification of CCHFV NP-specific Affimers

Rosetta 2 (DE3) *E. coli* cells were transformed with pET-11a vectors containing the seven different Affimer ORFs. A myoglobin-specific Affimer (Affimer-myo) was used as negative control. Protein expression was induced with IPTG and Affimers were expressed for 6 h at 25°C (figure 3.13). Protein expression was also tested overnight at 18°C, as previously performed for the expression of CCHFV and HAZV NPs, but best yields of Affimer expression were obtained at 25°C for 6 h.

Bacterial cell lysates were separated into soluble and insoluble fractions. The soluble fractions were incubated at 50°C to denature non-specific proteins prior to batch purification of Affimers using Ni^{2+}-NTA sepharose affinity chromatography. Non-specifically bound proteins were washed off with low imidazole buffer, before elution of Affimer proteins was achieved with 300 mM imidazole elution buffer. Samples were analysed by SDS-PAGE and Coomassie staining (figure 3.14). Elution fractions presented a predominant protein band of 13 kDa, corresponding to Affimers-Cys-8xHis. An additional protein band of 25 kDa was observed in the Coomassie staining analysis of some of the Affimer elution fractions (e.g. Affimer 1, Affimer 5). Fractions containing the proteins of interest were pooled and stored or immediately biotinylated.

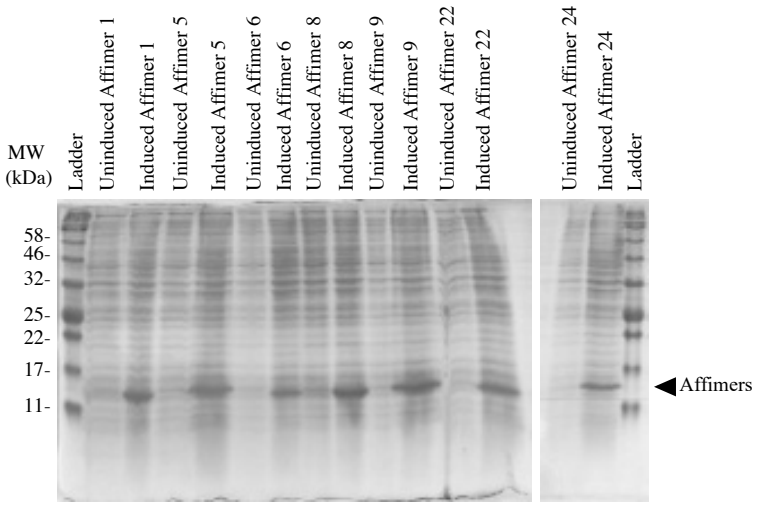
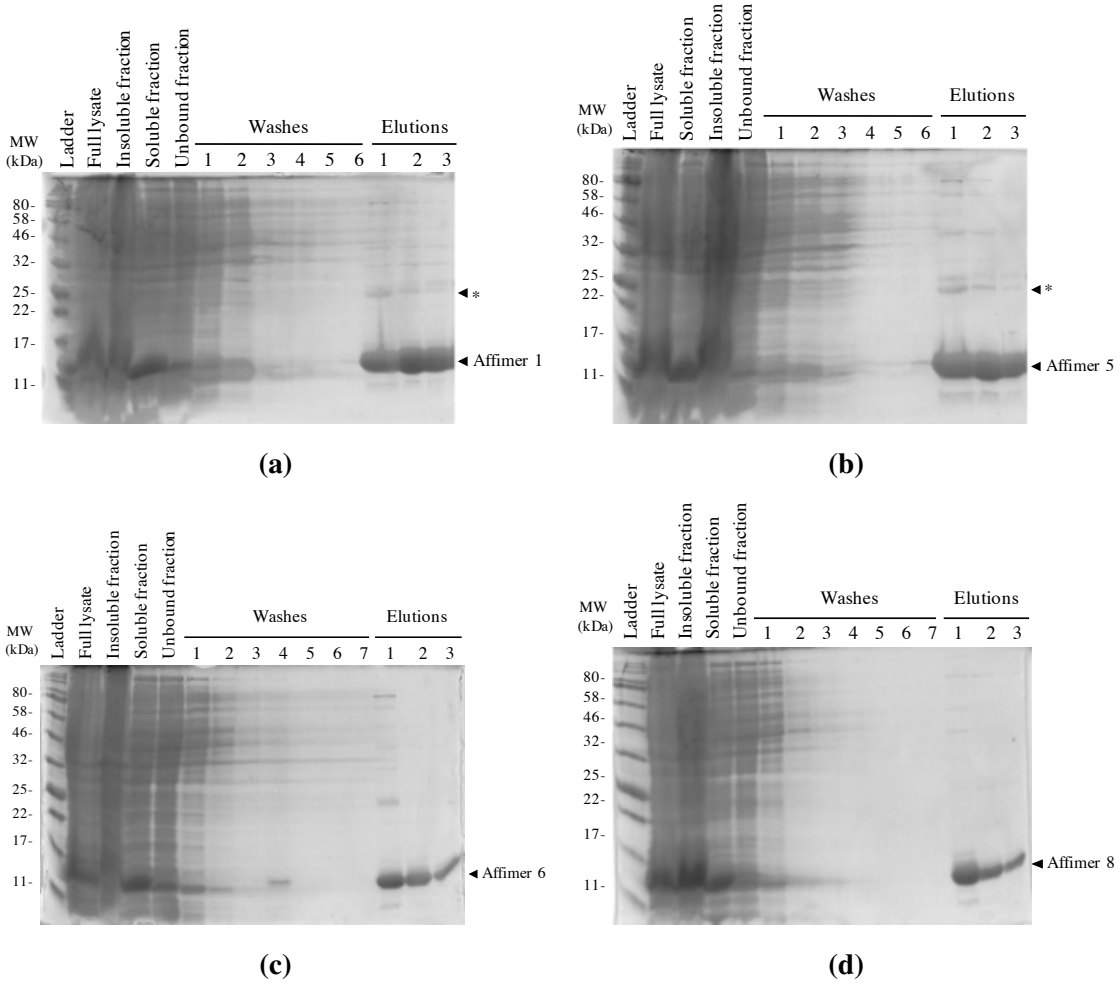


Figure 3.13: Bacterial expression of Affimers. SDS-PAGE and Coomassie staining analysis of uninduced and IPTG-induced Rosetta 2 (DE3) *E. coli* cells transformed with pET-11a vectors for the expression of CCHFV NP-specific Affimers.



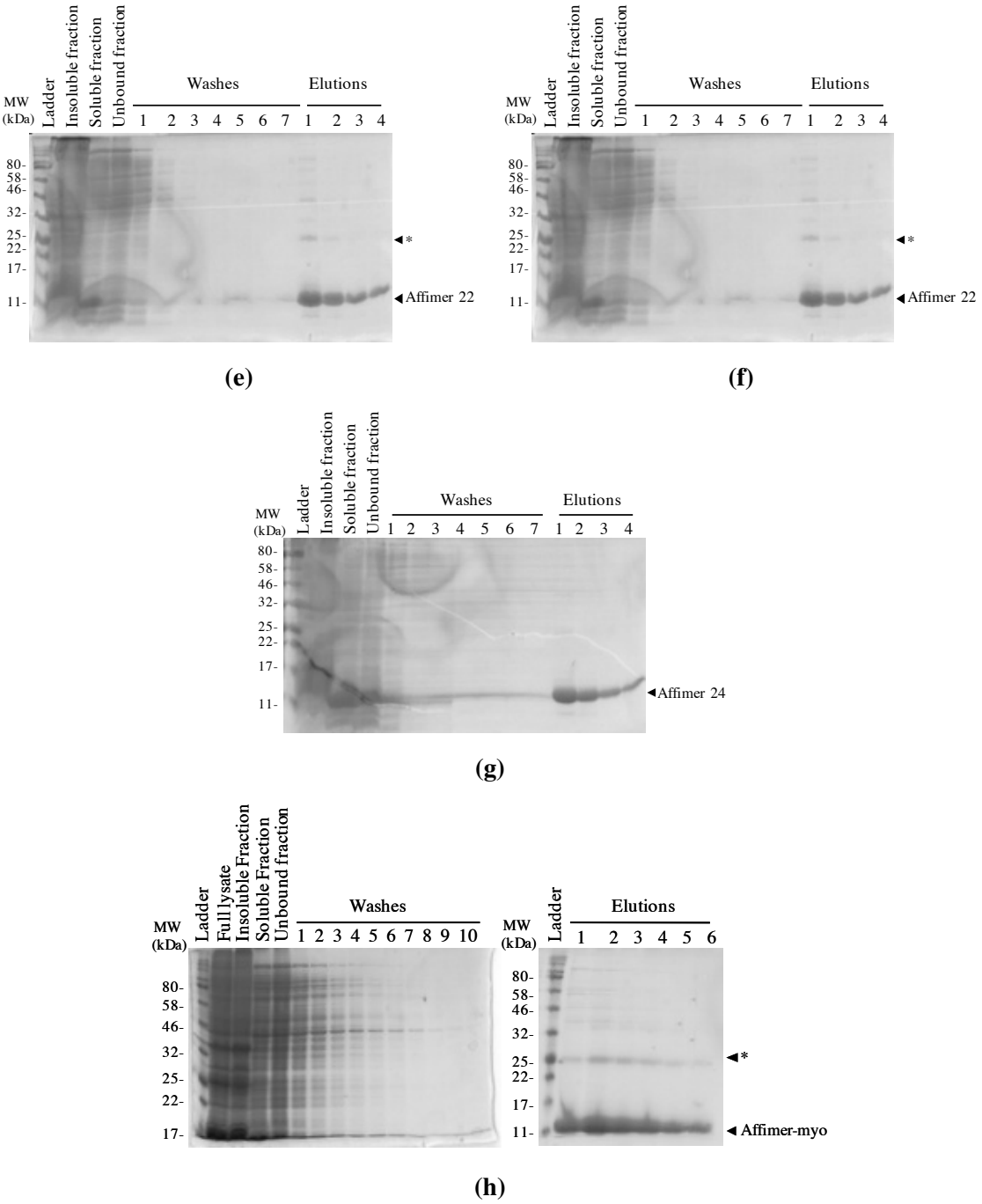


Figure 3.14: Batch Ni^{2+}-NTA sepharose affinity chromatography purification of Affimers. **(a-h)** SDS-PAGE analysis and Coomassie staining of fractions corresponding to the different steps from the Ni^{2+}-NTA batch affinity chromatography purification of Affimers 1, 5, 6, 8, 9, 22, 24 and the negative control Affimer-myo. * indicates an unknown protein product.

3.4.3 Biotinylation of CCHFV NP-specific Affimers

Affimers were biotinylated immediately after purification to avoid possible formation of disulphide bonds by thiol groups of cysteine residues. A maleimide moiety was employed to covalently link a biotin molecule to the free C-terminal cysteine residue (figure 3.15a). Tris(2-carboxyethyl)phosphine (TCEP) disulphide-reducing gel was used to ensure reduction of the sulfhydryl groups from the cysteine residues and Affimers were linked to biotin-maleimide.

An ELISA-like test was performed using streptavidin conjugated to horseradish peroxidase (HRP) to check biotinylation of Affimers. Briefly, micro-wells were incubated with serial dilutions of biotinylated Affimers, alongside a non-biotinylated Affimer used as negative control. Micro-wells were subsequently incubated with streptavidin conjugated to horseradish peroxidase (HRP) and 3,3',5,5'-tetramethylbenzidine (TMB) substrate, which changes colour upon oxidation by the conjugated HRP (figure 3.15b). All biotinylated samples presented a concentration-dependent increase in absorbance at 620 nm, corresponding to the appearance of the coloured oxidised substrate and confirming the presence of the biotin moiety (figure 3.16). Coloured substrate and absorbance at 620 nm were not observed in non-biotinylated samples, as expected due to absence of biotin.

3.4.4 Pull down affinity precipitation (AP) assays

The ability of the Affimers to bind CCHFV NP was first assessed by pull down affinity precipitation (AP) assays. Briefly, streptavidin-conjugated beads were coated with the 8 biotinylated Affimers and incubated with CCHFV NP. After washing off unbound CCHFV NP, samples were analysed by SDS-PAGE and western blotting using CCHFV NP-specific rabbit polyclonal IgGs and anti-His tag antibodies. Only Affimer-1 showed clear binding to CCHFV NP. The rest of the beads tested showed lack of binding to CCHFV NP, including the negative control Affimer (Affimer-myo), and the beads-only control (figure 3.17). Affimer-1 (renamed to Affimer-NP) was consequently chosen as the CCHFV NP binder of interest.

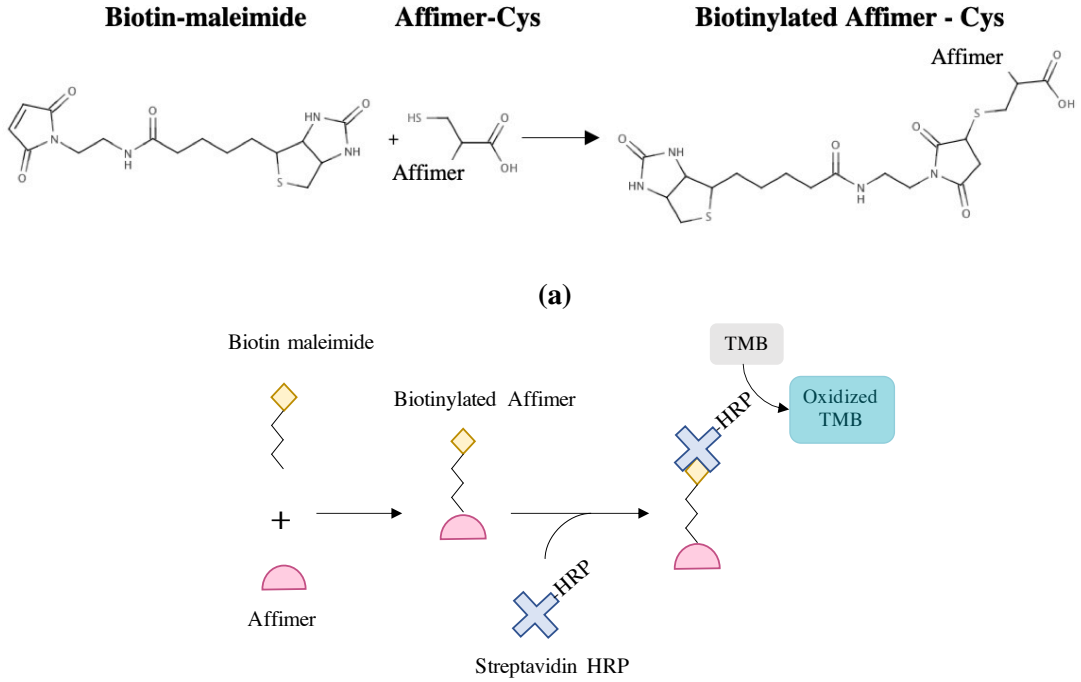


Figure 3.15: Biotinylation of Affimers. **(a)** Schematic representation of the biotinylation process of Affimers on their C-terminal cysteine residue. **(b)** Schematic representation of the biotinylation process of Affimers and confirmation of biotinylation by ELISA-like test using streptavidin conjugated to horseradish peroxidase (HRP) and 3,3',5,5'- tetramethylbenzidine (TMB) substrate.

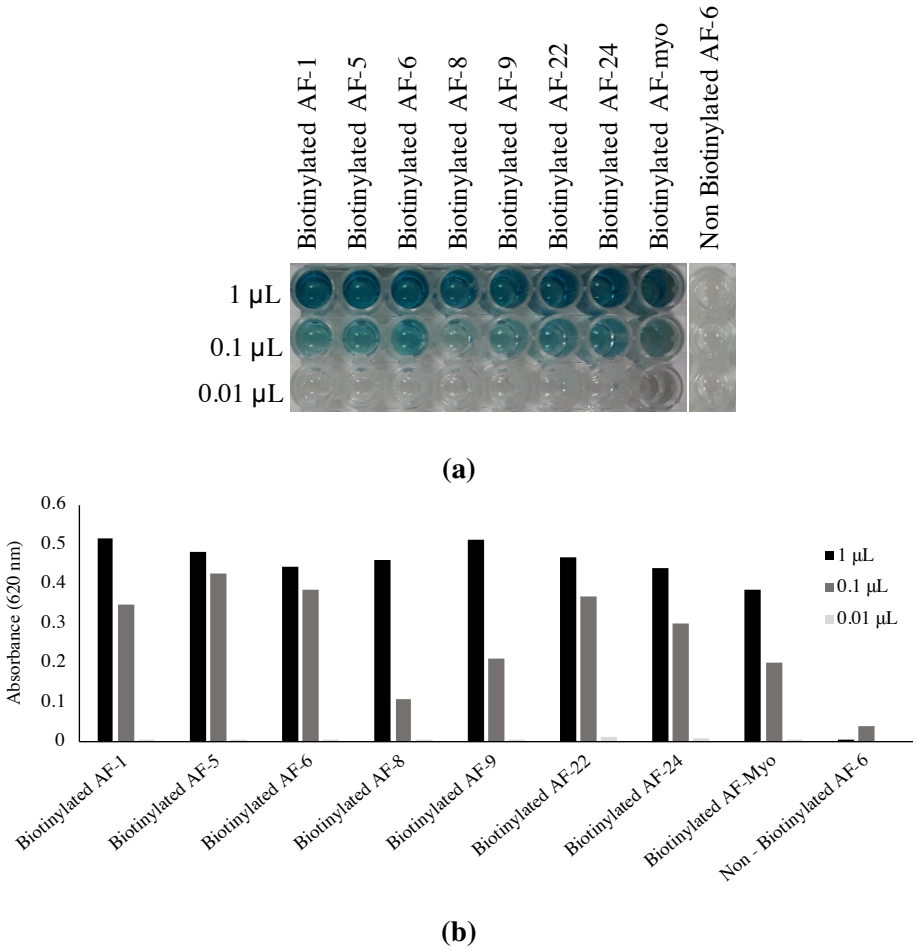


Figure 3.16: ELISA to check biotinylation of Affimers. **(a)** Colorimetric result of ELISA-like test to check biotinylation of Affimers. Different amounts (1, 0.1 or 0.01 µL) of biotinylated Affimers (0.5 mg/mL) were incubated with streptavidin-HRP and TMB. **(b)** Quantification of colorimetric output of results presented in *(a)*. Histogram representation of absorbance at 620 nm for different amounts of biotinylated Affimers and a non-biotinylated Affimer negative control.

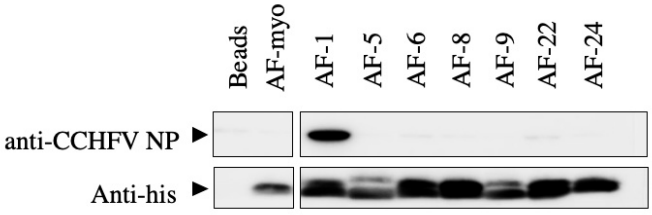


Figure 3.17: CCHFV NP affinity precipitation assay. Western blot analysis of samples corresponding to an affinity precipitation experiment performed using biotinylated Affimers captured on streptavidin coated beads and CCHFV NP as target molecule. A negative control Affimer (Affimer-myo) and a beads-only control were included.

3.5 Large-scale production of Affimer-NP

3.5.1 pET-11a large-scale production of Affimer-NP

To further characterise the interaction between CCHFV NP and Affimer-NP, larger quantities of the latter were required. Large-scale purification of Affimer-NP-Cys-8xHis was performed using the pET-11a expression vector. The C-terminal 8xHis tag was used for a first purification process by Ni^{2+}-NTA affinity chromatography, followed by a second SEC purification step.

3.5.1.1 Ni^{2+}-NTA affinity chromatography

Briefly, Affimer-NP-Cys-8xHis was purified from soluble fraction of cell lysate by Ni^{2+}-NTA affinity chromatography. Non-specifically bound proteins were washed off with low imidazole concentration buffer and elution of Affimer-NP-Cys-8xHis was achieved with 300 mM imidazole buffer. Samples were analysed by SDS-PAGE and Coomassie staining (figure 3.18). A protein band corresponding to the size of Affimer-NP-Cys-8xHis (13 kDa) was mainly present in elution fractions. An additional band at 25 kDa (highlighted with an asterisk in figure 3.18), previously observed in small scale Affimer purifications (section 3.4.2), was also present in these fractions.

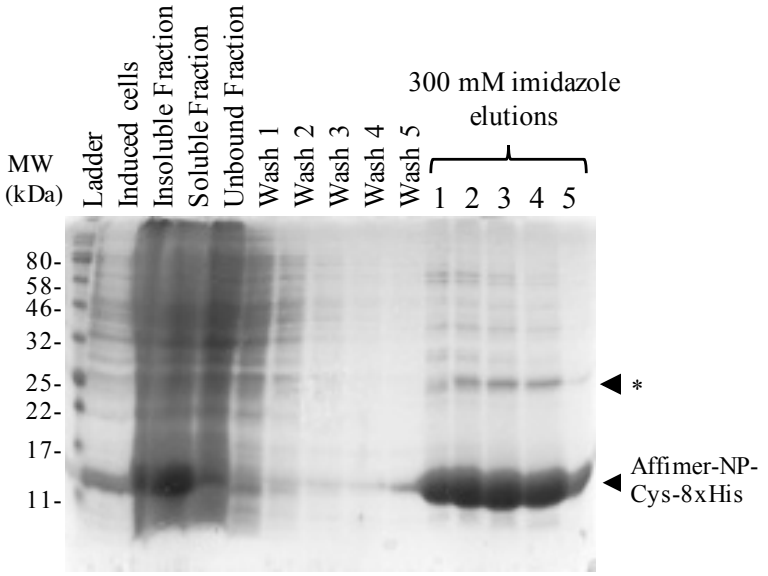


Figure 3.18: Affimer-NP-Cys-8xHis Ni^{2+}-NTA purification. SDS-PAGE and Coomassie staining analysis of fractions from Ni^{2+}-NTA affinity chromatography purification of Affimer-NP-Cys-8xHis. * indicates an unknown protein product.

3.5.1.2 Size exclusion chromatography

Fractions containing Affimer-NP-Cys-8xHis were pooled and a SEC purification step was performed using an S75 sephadex column. Two peaks were observed in the size exclusion chromatogram (figure 3.19a) and fractions corresponding to each peak were analysed by SDS-PAGE and Coomassie staining (figure 3.19b). Fractions corresponding to the first protein peak presented two protein bands at 13 kDa and 25 kDa, whereas those corresponding to the second peak showed a single 13 kDa protein band. The larger migration volume and the single 13 kDa protein band suggested the second peak exclusively contained the monomeric form of Affimer-NP-Cys-8xHis. Fractions corresponding to the second protein peak were collected and concentrated to 1 mg/mL.

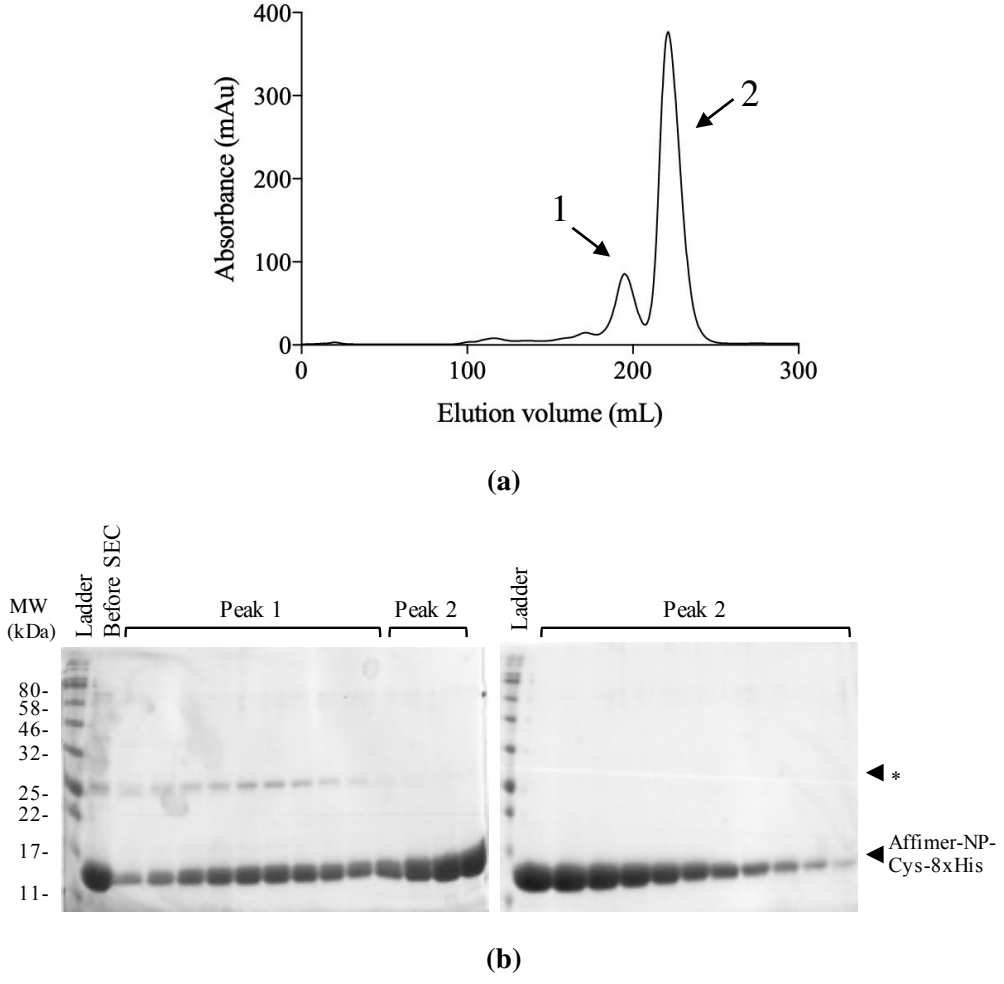


Figure 3.19: Affimer-NP-Cys-8xHis purification by size exclusion chromatography. (a) Chromatogram of the SEC purification of Affimer-NP-Cys-8xHis after Ni^{2+}-NTA affinity chromatography. **(b)** SDS-PAGE analysis and Coomassie staining of SEC fractions containing Affimer-NP-Cys-8xHis. Arrows indicate elution volume and migration distances of specified proteins. * indicates an unknown protein product.

3.5.2 pET-SUMO large-scale production of Affimer-NP

The size of the 25 kDa contaminant protein band and its SEC migration profile suggested this protein product could correspond to an Affimer dimer formed by disulphide bonds between C-terminal cysteine residues. To avoid the possible formation of disulphide bonds, Affimer-NP was subcloned into a pET-SUMO expression vector without the C-terminal cysteine. Briefly, PCR-amplified Affimer-NP ORF was subcloned into pET-SUMO vector by restriction enzyme digest and ligation (figure 3.20, see plasmid map in Appendix B). Rosetta 2 (DE3) *E.coli* cells were transformed with Affimer-NP-pET-SUMO plasmid for expression and purification of Affimer-NP.

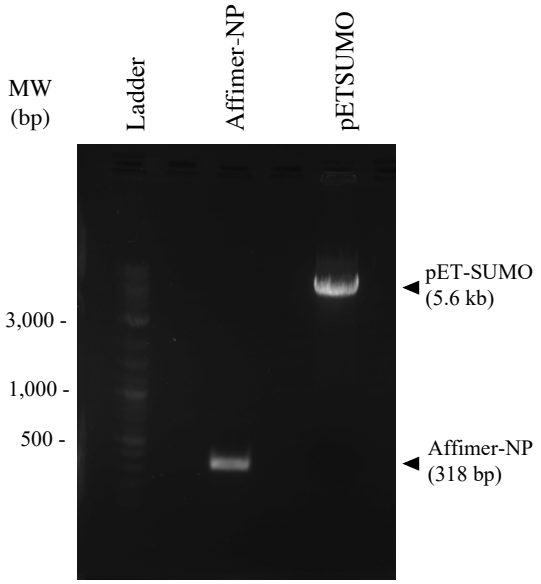


Figure 3.20: Digested Affimer-NP PCR products and pET-SUMO expression vector. Agarose gel electrophoresis analysis of digested Affimer-NP PCR products and pET-SUMO expression vector. Arrows indicate migration distances, names and sizes of specified DNA products.

3.5.2.1 Ni^{2+}-NTA affinity chromatography

Protein expression was induced with IPTG and 6xHis-SUMO-Affimer-NP was purified from soluble fraction of lysate by Ni^{2+}-NTA affinity chromatography. Non-specifically bound proteins were washed off with low imidazole concentration buffer and elution of 6xHis-SUMO-Affimer-NP was achieved with 300 mM imidazole buffer. Samples were analysed by SDS-PAGE and Coomassie staining (figure 3.21a) and a protein band corresponding to full length 6xHis-SUMO-Affimer-NP (24 kDa) was observed in elution fractions.

Fractions containing 6xHis-SUMO-Affimer-NP were pooled and 6xHis-SUMO tag was proteolytically removed by addition of 6xHis-SUMO Ulp1 protease. Cleavage resulted in the generation of two protein products: Affimer-NP (10 kDa) and 6xHis-SUMO tag (13 kDa). A second Ni^{2+}-NTA affinity chromatography purification step was performed to remove 6xHis-SUMO fusion tag and 6xHis-SUMO Ulp1 protease. Briefly, sample was applied to Ni^{2+}-NTA resin and flow-through containing Affimer-NP was collected. The remaining sample in the column was eluted with 500 mM imidazole buffer and all fractions were analysed by SDS-PAGE and Coomassie staining (figure 3.21b). After the second affinity chromatography purification step, there was a clear reduction in the amount of 6xHis-SUMO tag present in the Affimer-NP sample.

3.5.2.2 Size exclusion chromatography

A SEC purification step was performed to further purify Affimer-NP using an S75 sephadex column. Two peaks were observed in the corresponding size exclusion chromatogram, possibly corresponding to monomeric and dimeric forms of Affimer-NP (figure 3.22a). The first peak, potentially corresponding to a dimeric form of the protein, was significantly smaller than in the case of Affimer-NP produced from the pET-11a expression vector (section 3.5.1). Fractions corresponding to the second peak were analysed by SDS-PAGE gel and Coomassie staining (figure 3.22b). A single band with a migration distance corresponding to the size of Affimer-NP (10 kDa) was observed, confirming the efficient removal of non-desired products. Fractions containing the protein of interest were collected and concentrated to 1 mg/mL.

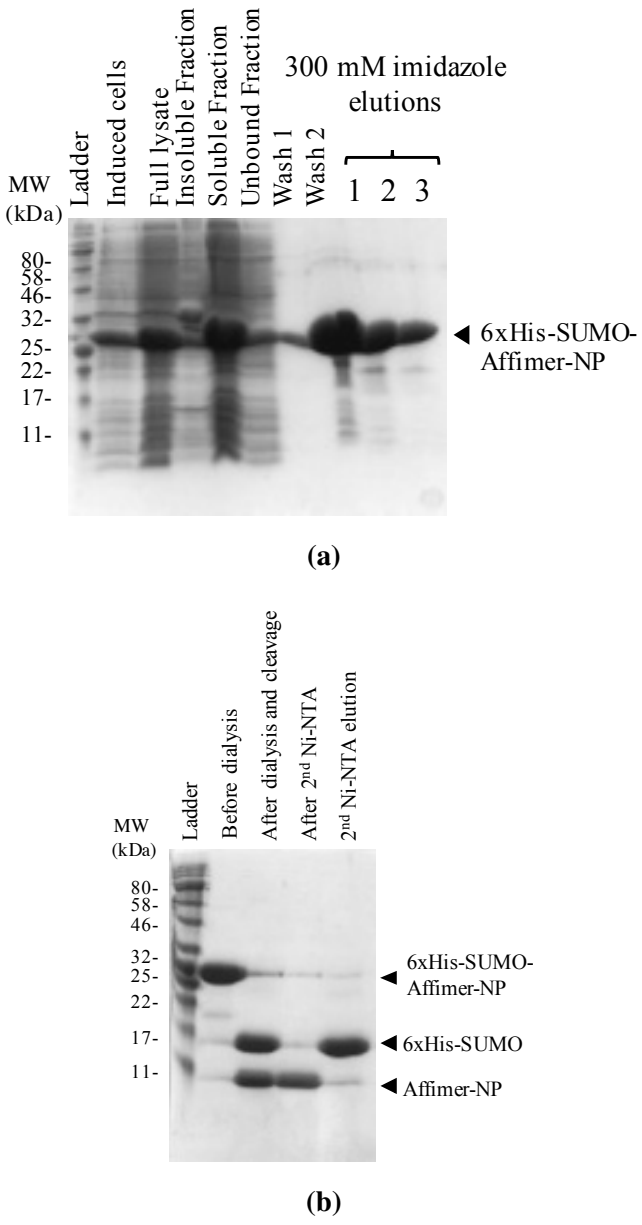


Figure 3.21: Affimer-NP purification by Ni^{2+}-NTA affinity chromatography. (a) SDS-PAGE analysis and Coomassie staining of fractions from different steps of the expression and purification processes of 6xHis-SUMO-Affimer-NP by Ni^{2+}-NTA affinity chromatography. **(b)** SDS-PAGE analysis of purified 6xHis-SUMO-Affimer-NP after dialysis and cleavage with 6xHis-SUMO Ulp1 protease and a second purification step by Ni^{2+}-NTA affinity chromatography. Arrows indicate migration distances of specified proteins.

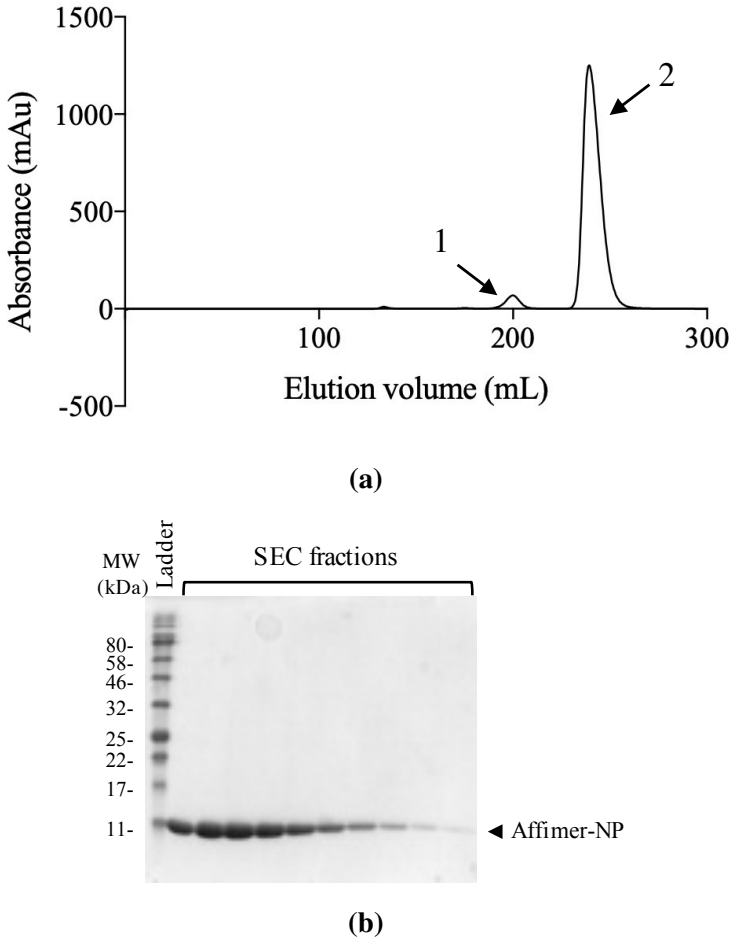


Figure 3.22: Affimer-NP purification by size exclusion chromatography. **(a)** Chromatogram of the SEC purification of Affimer-NP sample after Ni^{2+}-NTA affinity chromatography. **(b)** SDS-PAGE analysis and Coomassie staining of SEC fractions containing Affimer-NP. Arrows indicate elution volume and migration distances of specified proteins.

3.6 Chapter summary and discussion

This chapter describes the optimisation of expression and purification of CCHFV and HAZV NPs, and the production of two types of CCHFV NP-specific recognition molecules: rabbit polyclonal IgGs and Affimer reagents.

Purification of CCHFV and HAZV NPs has been previously described in the literature to obtain large quantities of pure protein for crystallisation experiments, among others. Expression of CCHFV NP has been described and optimised using a pET-SUMO construct [151], whereas purification of HAZV NP has been previously achieved using a pGEX vector for N-terminal glutathione S-transferase (GST) tagging and glutathione sepharose affinity chromatography purification [224]. In this chapter, the expression of CCHFV and HAZV NPs using a pET-SUMO vector was described. Due to physicochemical similarities between both proteins, including analogous electrostatic surface potential maps and similar theoretical isoelectric points (CCHFV NP pI: 8.46, HAZV NP pI: 8.73, see Appendix A), the purification protocol used for CCHFV NP was adapted to HAZV NP and a similar procedure was used for purification of both nairoviral NPs.

6xHis-SUMO-tagged NPs were purified using a first Ni^{2+}-NTA affinity chromatography step. A high-salt wash was included in HAZV NP protein preparations to remove any bacterial nucleic acids potentially bound to HAZV NP. This step was not necessary for CCHFV NP protein preparations as, based on previous experience, DNase and RNase addition to CCHFV NP-induced cell lysates is thought to be sufficient for efficient removal of bacterial DNA and RNA present in the sample. Absence of nucleic acids was corroborated by measuring the ratio of absorbance at 260 nm versus 280 nm (A260/A280) of elution samples, which resulted in <1 values in all of the cases. After cleavage of 6xHis-SUMO tag with 6xHis-SUMO Ulp1 protease, a second Ni^{2+}-NTA affinity chromatography step was performed to remove the tag and protease, followed by SEC. CCHFV and HAZV NPs showed migration volumes corresponding to $\sim$50 kDa in the size exclusion chromatograms, suggesting a monomeric state of the recombinant proteins, according to previously described results [151, 224].

An unknown protein product of $\sim$20 kDa was observed in SDS-PAGE and Coomassie staining analyses of elution fractions from the first Ni^{2+}-NTA affinity chromatography for both CCHFV and HAZV NPs preparations. After 6xHis-SUMO Ulp1 protease cleavage this product was not observed, suggesting it could possibly be constituted by 6xHis-SUMO-tag and a small fragment of the NPs. Cleavage of the $\sim$20 kDa contaminant product mediated by the protease would potentially lead to a 6xHis-

SUMO tag fragment of ∼13 kDa and a smaller fragment of ∼7 kDa, not be visible in the SDS-PAGE analysis and Coomassie staining due to its small size. In any case, this product was not observed in any downstream analyses and did not seem to interfere with the final protein preparations.

Protein preparations for CCHFV NP were in general cleaner than for HAZV NP, the latter presenting less solubility, partial 6xHis-SUMO Ulp1 protease mediated cleavage, a remarkable peak at the void volume of the size exclusion chromatography column and less purity in the final protein preparation. These observations suggest a particular tendency of HAZV NP to aggregate, interfering with solubility and protease mediated cleavage, and forming aggregates of high molecular weight. However, large amounts (mg quantities) of both nairoviral NPs could be successfully purified with an acceptable purity.

An initial validation of the purified NPs was performed with a caspase 3 cleavage assay. CCHFV and HAZV NPs present a caspase 3 recognition site (DEVD and DQVD, respectively) positioned in the apex of the arm domain. Cleavage of these recognition sites occurs at the late onset of the CCHFV and HAZV infectious cycle, when the apoptotic pathway is induced, yielding two protein products of ∼20 and 30 kDa [146, 152]. Caspase 3 mediated cleavage of CCHFV NP has been previously demonstrated *in vitro* [140]. To explore the caspase 3 cleavage in the analogous HAZV NP site, caspase 3 was expressed from a pET21 vector and purified using a previously described protocol for Ni^{2+}-NTA affinity chromatography [222]. Caspases are synthesised in cells as catalytically inactive zymogens, and must undergo an activation process. The activation of caspase 3 occurs through cleavage, to separate the large and small subunits. Caspase 3 was initially expressed as full length, and subsequently auto-processed yielding large (p17) and small (p12) subunits.

Caspase 3 mediated cleavage of both nairoviral NPs, yielding two prominent NP cleavage products with approximate molecular weights of ∼20 and 30 kDa, which corresponded precisely to the apparent mass of the CCHFV and HAZV NPs products identified in apoptotic mammalian cells [146, 152]. These results indicate that the cleavage site in the apex of the arm domains is exposed and accessible for caspase cleavage, suggesting the adequate folding of the recombinant NPs, and corroborating that the ∼20 and 30 kDa fragments observed in CCHFV and HAZV infected cells can be generated by caspase 3-mediated cleavage.

Recombinant CCHFV NP was lyophilised and used for generation of CCHFV NP-specific polyclonal IgGs. Freeze-dried protein samples present greater storage stability than proteins in liquid forms, and thus, lyophilisation of CCHFV NP ensured

the appropriate conditions of the sample upon delivery to the commercial company. Furthermore, lyophilised protein samples can be easily resuspended in appropriate buffers facilitating the antigen preparation process for immunisation. The corresponding CCHFV NP-specific polyclonal antibodies were first validated as primary antibodies in western blotting, demonstrating the ability to specifically detect down to 10 ng of CCHFV NP. Cross-reactivity for HAZV NP was observed, and conversely, HAZV NP-specific antisera also demonstrated reactivity for CCHFV NP. CCHFV and HAZV form part of the same serogroup and share common antigenic determinants. Notably, CCHFV and HAZV NPs share 60% amino acid sequence identity and possess remarkable structural similarities, supporting the cross-reactivity observed for both CCHFV- and HAZV-specific polyclonal antibodies.

The second approach explored for the generation of CCHFV NP-specific binding molecules was an Affimer screening. Seven Affimer sequences were determined as unique CCHFV NP binders. One of the sequences, corresponding to Affimer-1, was present 17 times. 'Phages displaying Affimer-1 showed a remarkably high absorbance signal in the 'phage ELISA results, suggesting that this Affimer variant exhibits a distinct reactivity for CCHFV NP. 'Phages displaying other Affimer variants presented distinctive binding to CCHFV NP-coated and blank wells, but the 'phage ELISA signal was significantly lower than for 'phages displaying Affimer-1, suggesting these binders possess less ability to recognise CCHFV NP.

The corresponding Affimer proteins were expressed and purified by Ni^{2+}-NTA affinity chromatography, alongside a myoglobin-specific Affimer used as negative control. Purity of Affimer preparations was evaluated by SDS-PAGE and Coomassie staining and elution fractions presented a predominant band of 13 kDa, as expected for Affimer-Cys-8xHis products. In some cases, a second band of $\sim$25 kDa was also observed, suggesting the presence of Affimer dimers in the elution sample, despite the use of reduction agents in SDS-PAGE analyses.

To assess the binding of the Affimers to CCHFV NP, AP assays were performed. Only Affimer-1 demonstrated clear binding to CCHFV NP, whereas the rest of the Affimer molecules and the negative controls presented lack of binding. These results are in concordance with the previously described 'phage ELISA results and corroborate the strong binding of Affimer-1 to CCHFV NP. Binding of the other isolated Affimer molecules to CCHFV NP could not be detected by AP assays, suggesting a weak or unspecific interaction. Taking these results, Affimer-1 (renamed to Affimer-NP) was selected as CCHFV NP-specific Affimer of interest.

To further study the binding characteristics of Affimer-NP to CCHFV NP, high quantities of Affimer-NP were required. Affimer-NP was produced in large-scale using the pET-11a expression vector, and was purified by Ni^{2+}-NTA affinity chromatography and SEC. Ni^{2+}-NTA affinity chromatography elution fractions presented a predominant band of 13 kDa, corresponding to Affimer-NP-Cys-8xHis, but the aforementioned 25 kDa protein product was also present. The corresponding size exclusion chromatogram presented two peaks, the first of which contained both the 13 and 25 kDa protein products, whereas the second one presented only the 13 kDa protein product, as confirmed by SDS-PAGE and Coomassie staining analysis of the corresponding samples. The size of the 25 kDa product, together with the lower migration volume of this product in SEC, supported the hypothesis that this product corresponds to a dimer of Affimers potentially formed by a disulphide bond on their C-terminal cysteines, despite the presence of reduction agents in the SEC buffer and SDS-PAGE analyses.

To overcome the presence of a potential dimeric Affimer product, which could interfere in downstream analyses for which high protein sample homogeneity is required, Affimer-NP was subcloned in a pET-SUMO vector excluding the C-terminal cysteine. Two consecutive Ni^{2+}-NTA affinity chromatography purifications were performed, followed by SEC. In this case, the 25 kDa product was absent in the Ni^{2+}-NTA affinity chromatography elution fractions, suggesting that removal of the C-terminal cysteine residue decreased the amount of dimeric Affimer produced. A small peak was observed in the SEC, potentially corresponding to a dimeric form of the Affimer, which could be explained by intermolecular β-sheet interactions between monomers, as has been previously described for other Affimer reagents [225], but this has not been experimentally confirmed for Affimer-NP. The large migration volume of the second peak of the SEC and the presence of a single protein band at 10 kDa in the corresponding fractions suggested successful purification of monomeric Affimer-NP with high purity.

In summary, the results presented in this chapter meet objective 1 of this project (see project aims in section 1.4). Here, the successful purification of recombinant CCHFV and HAZV NPs with a preliminary validation of these molecules by a caspase 3 mediated cleavage assay was described. CCHFV NP-specific polyclonal antibodies and Affimers were generated using recombinant protein, and initially validated by western blot and AP assays, respectively. Finally, large-scale production of Affimer-NP was optimised for high purity and homogeneity. Further characterisation of the interaction between Affimer-NP and CCHFV NP is addressed in Chapter 4, and application of both CCHFV NP-specific polyclonal antibodies and Affimer-NP in diagnostic assays is covered in Chapter 5.

Chapter 4

Results: Characterisation of the interaction between CCHFV NP and Affimer-NP

Characterisation of protein-protein interactions and their consequent functional effects requires the combination of interdisciplinary research methods within the fields of physics, chemistry and biology. Physicochemical techniques are commonly performed using purified proteins in simple buffer conditions, allowing the study of physical properties such as the secondary and tertiary structures of the interacting proteins or the affinity and kinetics of the binding interaction. The functional effects derived from these interactions can be addressed using biochemical analyses in simple *in vitro* experimental set-ups, or more complex and biologically relevant conditions using *in cell* or *in vivo* models.

This chapter describes different approaches used for characterisation of the interaction between Affimer-NP and CCHFV NP, including pull down affinity precipitation (AP) assays, surface plasmon resonance (SPR), circular dichroism (CD), size exclusion chromatography (SEC) and X-ray crystallography (summarised in table 4.1). Functional studies were also addressed to determine the effect of Affimer-NP in CCHFV NP RNA binding function and CCHFV-specific gene expression, using *in vitro* fluorescence anisotropy (FA) assays and a mini-genome system in mammalian cells, respectively.

	Detection of protein binding	**Purification of complex in large quantities**	**Affinity and kinetics**	**Structural analysis**
AP	**+**	-	-	-
SPR	+	-	**+**	-
CD	**+**	-	-	**+**
Size exclusion chromatography	**+**	**+**	-	-
X-ray crystallography	**+**	-	-	**+**

Table 4.1: Methods used for the physicochemical characterisation of the interaction between Affimer-NP and CCHFV NP. Pull down affinity precipitation (AP) assays, surface plasmon resonance (SPR), circular dichroism (CD), size exclusion chromatography (SEC) and X-ray crystallography were used for detection of protein binding, study of affinity and kinetics, purification of complex in large quantities and structural analyses.

4.1 Specificity: pull down affinity precipitation (AP) assays

Introduction to AP: Pull down affinity precipitation (AP) assays are small-scale affinity purification procedures used for the co-precipitation of interacting proteins. AP assays can be performed using different experimental set ups according to the type of proteins and matrix to be used, the latter usually consisting of agarose or magnetic beads.

If the interaction of interest involves antibodies, these can be immobilised using a matrix functionalised with specific antibody-binding proteins, such as protein A, which contains IgG binding domains [226]. For non-antibody interactions, recombinant proteins can be immobilised as capture molecules on the resin. Expression of proteins with clonable tags (e.g. poly-histidine or GST tags) allows immobilisation of the protein onto the appropriate matrix (e.g. metal or glutathione (GSH) functionalised agarose beads, respectively) [227]. Specific post-translational modifications of recombinant proteins can also be introduced for their immobilisation onto certain matrices (e.g. immobilisation of biotinylated proteins on streptavidin-coated beads).

Once the protein of interest is immobilised onto the matrix and excess unbound protein is washed off, functionalised matrix is incubated with a solution containing the candidate interactor protein (e. g. purified protein samples or cell lysates). After incubation and washing, matrix containing the capture protein and its interacting partners is collected (figure 4.1). Different techniques can be used to analyse the protein complex

bound to the matrix (e.g. SDS-PAGE analysis is commonly followed by western blotting or mass spectrometry [228] to characterise the interacting partners).

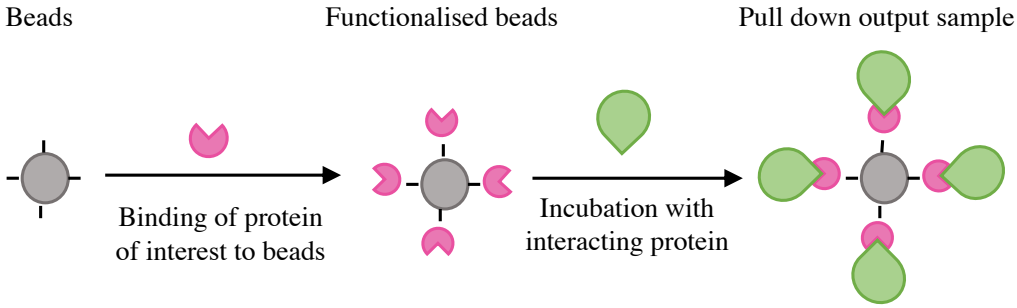


Figure 4.1: Pull down affinity precipitation assay procedure. Protein of interest (represented in pink) is immobilised onto beads and incubated with a solution containing interacting partners (represented in green). Pull down output sample contains the complex formed by the interacting proteins bound onto the matrix.

Results: The specificity of Affimer-NP for CCHFV NP was tested using pull down AP assays. CCHFV NP and its homologue HAZV NP were used as target molecules, alongside a negative control Affimer-myo. Briefly, streptavidin coated beads were used for the immobilisation of biotinylated Affimer-NP or biotinylated Affimer-myo, and incubated with CCHFV or HAZV NPs. Unbound NPs were washed off, and samples analysed by SDS-PAGE and Coomassie staining (figure 4.2a) or western blotting (figure 4.2b). CCHFV NP-specific rabbit polyclonal IgGs (previously described in section 3.3) were used for detection of CCHFV NP and HAZV NP, and an anti-His tag antibody was used for the detection of Affimers. Affimer-NP showed binding to CCHFV NP but not to HAZV NP, and the negative control Affimer-myo showed no binding to either NP (figure 4.2). These results demonstrate that Affimer-NP selectively recognises CCHFV NP, showing no cross-reactivity with HAZV NP, despite their close structural similarity.

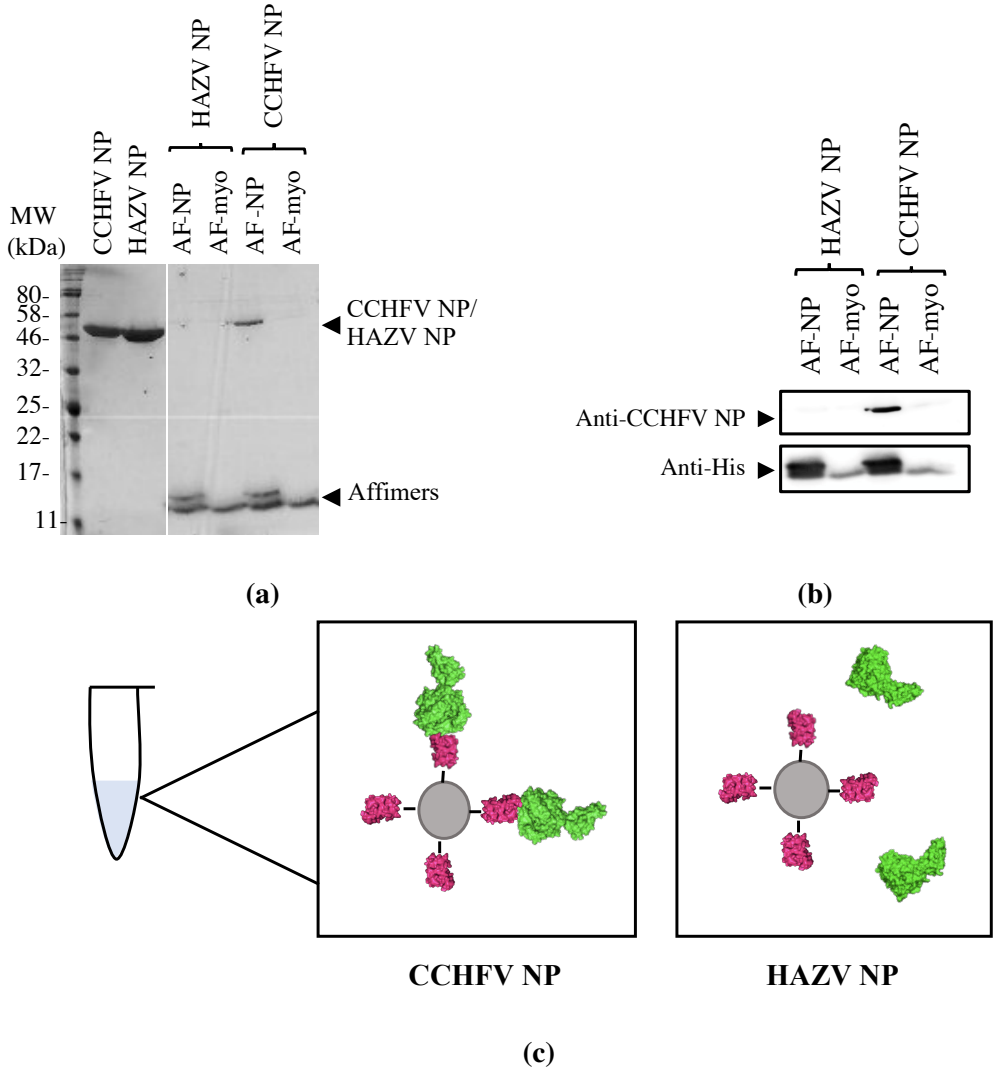


Figure 4.2: CCHFV and HAZV NP pull down affinity precipitation assay. (a) SDS-PAGE and Coomassie staining analysis of input CCHFV and HAZV NP samples (lanes 1, 2) and pull down AP samples (lanes 3-6). Affimer-NP and Affimer-myo were immobilised onto streptavidin beads, subsequently incubated with CCHFV NP or HAZV NP. **(b)** Western blot analysis of pull down samples using CCHFV NP-specific and anti-His tag antibodies as primary detection molecules for nairoviral NPs and Affimers, respectively. Affimer-NP commonly appeared as a doublet in Coomassie and WB analyses for unknown reasons. **(c)** Schematic representation of the pull down AP assays using CCHFV or HAZV NP as potential binding partners of Affimer-NP. Affimer-NP binds specifically to CCHFV NP, and shows no cross-reactivity with HAZV NP.

4.2 Affinity and binding kinetics: SPR

Introduction to SPR: SPR is a physical phenomenon that consists of the resonant oscillation of electrons at the interface between negative (e.g. metal) and positive (e.g. air, vacuum or water) permittivity materials under stimulation by polarised light. SPR changes can be used to monitor the binding, specificity, affinity and kinetics of molecular interactions taking place on the surface of the negative permittivity material [229].

Stimulation of a metal with polarised light results in a non-radiative electromagnetic surface wave, named surface resonance polaritron, which is very sensitive to changes in the refractive index of the metal surface such as the adsorption of molecules to the conducting surface [229]. SPR-based sensors use a metal surface positioned next to a glass prism to generate parallel polarised (p-polarised) light to the plane of incidence with a given wavelength and angle (figure 4.3a) [230]. Most of the incident light is totally reflected and not absorbed by the surface. At a certain incident angle (resonance angle), the surface plasmon is resonant with the vibration of the irradiated light, resulting in the absorption of light at that angle (evanescent wave) and the consequent absence of reflected light with that angle (resonance or SPR angle) (figure 4.3b). Immobilisation of a molecule on an SPR-sensor metal surface results in a specific SPR angle. A shift on the SPR angle can be detected when changes such as molecular binding and dissociation events take place on or near the metal film [229]. These changes can be monitored over time and converted to a sensogram that describes relative response units (RU), proportional to the molecular mass on the surface, versus time (figure 4.4).

SPR-based sensors normally consist of a glass slide coated with a thin layer of silver or gold. The surface is covered by a linker layer, commonly a matrix of modified dextran, that allows coupling of a wide variety of ligands. Covalent immobilisation of ligands to the surface is possible using amine, thiol or aldehyde groups of the ligand chemically bound to the sensor surface. High affinity binding is an alternative for ligand coupling that can be performed in physiological buffer conditions. For permanent capture of biotinylated molecules, for example, sensors containing streptavidin covalently attached to dextran matrix can be used (SA chip, figure 4.5). The affinity of streptavidin for biotin is extremely high, with an equilibrium dissociation constant of about 10^{-15} M, and dissociation of biotinylated ligands from the surface is generally negligible. If reversible capture of biotinylated ligand is required, sensors containing oligonucleotides covalently attached to dextran matrix can be used in combination with streptavidin-carrying complementary oligonucleotides [231]. Histidine-tagged ligands can also be high-affinity captured on chips functionalised with Ni^{2+}-NTA covalently attached to

dextran matrix.

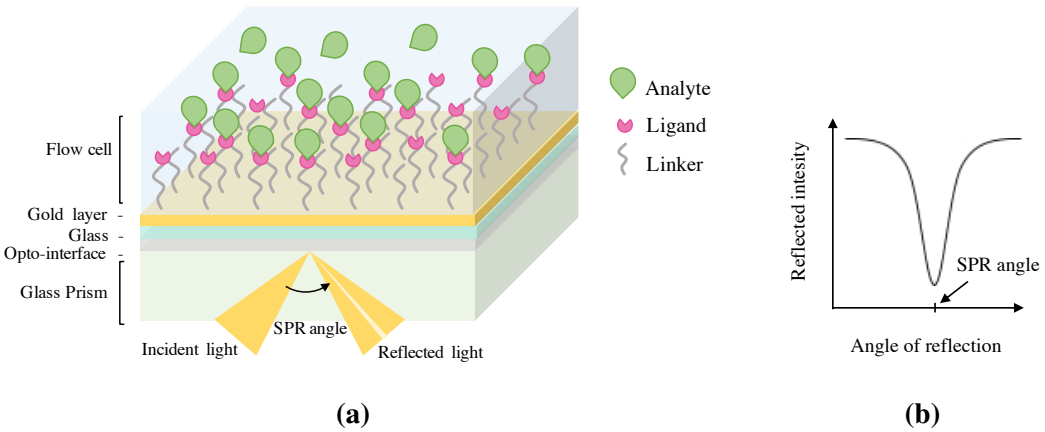


Figure 4.3: Study of affinity and binding kinetics by SPR. (a) Schematic representation of an SPR-based sensor and its components. Gold surface is stimulated with parallel polarised (p-polarised) light at different angles. At a certain incident angle, the surface plasmon is resonant with the vibration of the irradiated light, resulting in the absorption of light and the consequent absence of reflected light with that angle. **(b)** SPR scanning angle response. SPR causes an intensity dip in the reflected light at the sensor surface. A shift in the curve represents molecular binding.

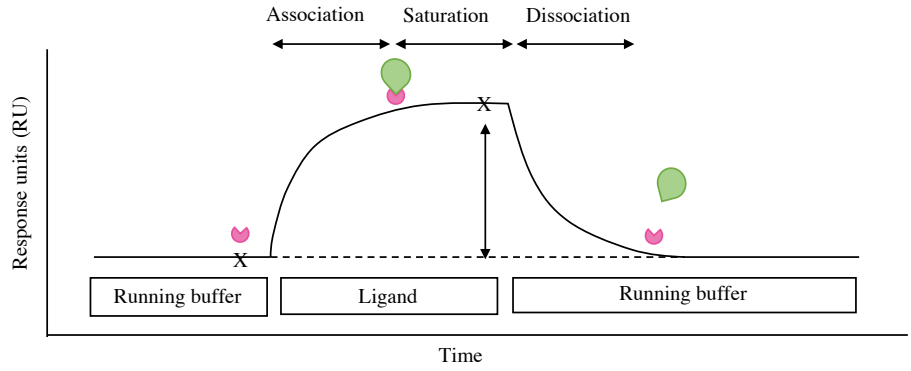


Figure 4.4: Typical SPR sensogram and steps of an analytical cycle. Injection of ligand onto the SPR cell results in three binding phases: association, saturation and dissociation.

Once the ligand is attached to the surface of the chip, the candidate interactor protein (analyte) is injected onto the cell containing the immobilised ligand molecule. The candidate protein is also injected onto a blank cell and a cell containing a non-interacting ligand, which are used for referencing analysis. If the candidate interactor protein binds to the immobilised ligand on the surface, the resonance angle shifts accompanying the increase in density of the proteins bound to the surface [229]. Therefore, it is possible to examine the binding and dissociation of proteins by analysing the shift of the resonance angle over time.

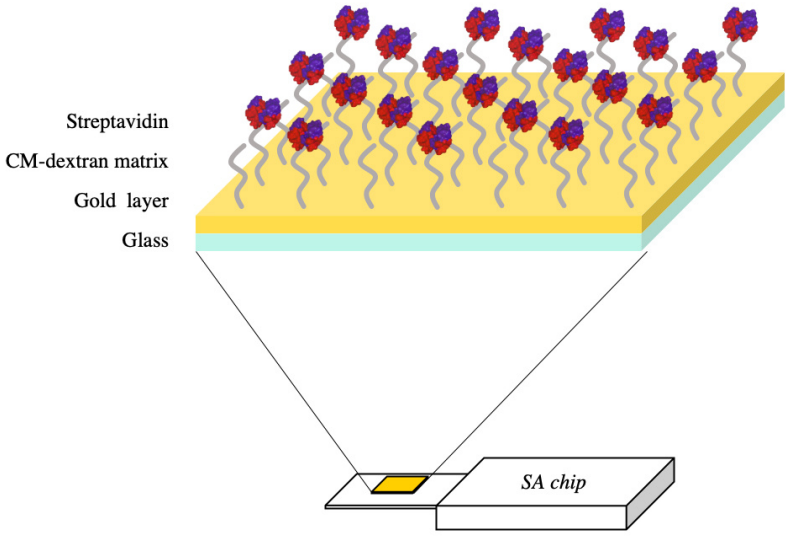


Figure 4.5: Schematic of a streptavidin SPR sensor chip. Streptavidin is covalently attached to a carboxymethyl-dextran (CM-dextran) matrix on the sensor surface for capture of biotinylated compounds.

Kinetics and affinity analyses of the binding interaction between ligand and analyte can be done by fitting experimental data to interaction models. The most commonly used model is the Langmuir model for 1:1 binding interactions between analyte A and ligand B (Equation 4.1) [232].

$$\mathrm{A} + \mathrm{B} \underset{\mathrm{k_{off}}}{\overset{\mathrm{k_{on}}}{\rightleftharpoons}} \mathrm{AB} \tag{4.1}$$

where: k_{on} is the association rate constant ($M^{-1}s^{-1}$) and k_{off} is the dissociation rate constant (s^{-1}).

The concentration of the complex AB varies along time according to the differential equation 4.2 where R and R_{max} correspond to the observed response and the theoretical maximal response that would be observed if an infinite concentration of A was injected, respectively.

$$dR/dT = k_{on}[A](R_{max} - R) - k_{off}R \tag{4.2}$$

If a plateau is reached for analyte injection, equation 4.2 can be simplified to equation 4.3 where R_{eq} corresponds to response reached at equilibrium. This equation can be further rearranged to equation 4.4, in which K_D is the thermodynamic dissociation constant.

$$k_{\text{on}}[A](R_{\text{max}} - R_{\text{eq}}) - k_{\text{off}}R_{\text{eq}} = 0 \tag{4.3}$$

$$R_{\text{eq}} = ([A]R_{\text{max}})/(K_{\text{D}} + [A]) \tag{4.4}$$

For simple 1:1 binding models, the affinity constant can also be obtained directly from kinetics analyses, as the ratio of the rate constants (Equation 4.5).

$$K_{\text{D}} = k_{\text{off}}/k_{\text{on}} \tag{4.5}$$

Results: To determine the binding kinetics and affinity of Affimer-NP for CCHFV NP, surface plasmon resonance (SPR) based experiments were performed, alongside Affimer-myo and HAZV NP used as negative controls. Different cells of a streptavidin sensor chip were coated with biotinylated Affimer-NP and Affimer-myo, and different concentrations (10 nM, 100 nM and 1 µM) of CCHFV or HAZV NP were injected into each cell (figure 4.6).

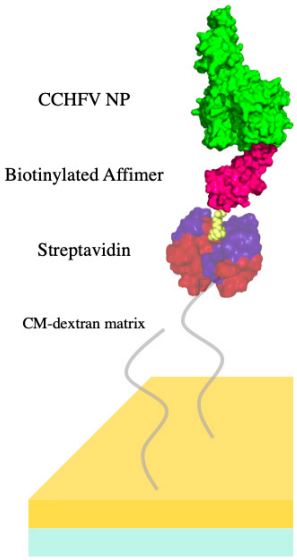


Figure 4.6: Schematic representation of the SPR-based analysis of CCHFV NP/Affimer-NP affinity and binding kinetics. A streptavidin chip was used to immobilise biotinylated Affimers as ligands. Nairoviral NPs were injected as analyte molecules.

CCHFV NP sensograms showed binding to Affimer-NP but not to Affimer-myo and conversely, HAZV NP sensograms showed no binding to either Affimer-NP nor Affimer-myo (figure 4.7). Double-referencing analysis of the SPR data was done, fitting CCHFV NP sensograms to a Langmuir 1:1 binding model (figure 4.8). The most accurate fitting corresponded to the 10 nM dilution of CCHFV NP, and lowest Chi^2 values were obtained

relative to the response units of the sensogram (Chi^2 = 0.104). The fitted curve reveals a K_D in the low nanomolar range (K_D = 5.69 nM), demonstrating a high affinity interaction between Affimer-NP and CCHFV NP with a fast rate association (k_{on} = 1.17 x 10^6 M^{-1} s^{-1}) and a slow rate dissociation (k_{off} = 6.67 x 10^{-3} s^{-1}) (figures 4.8a and 4.8e).

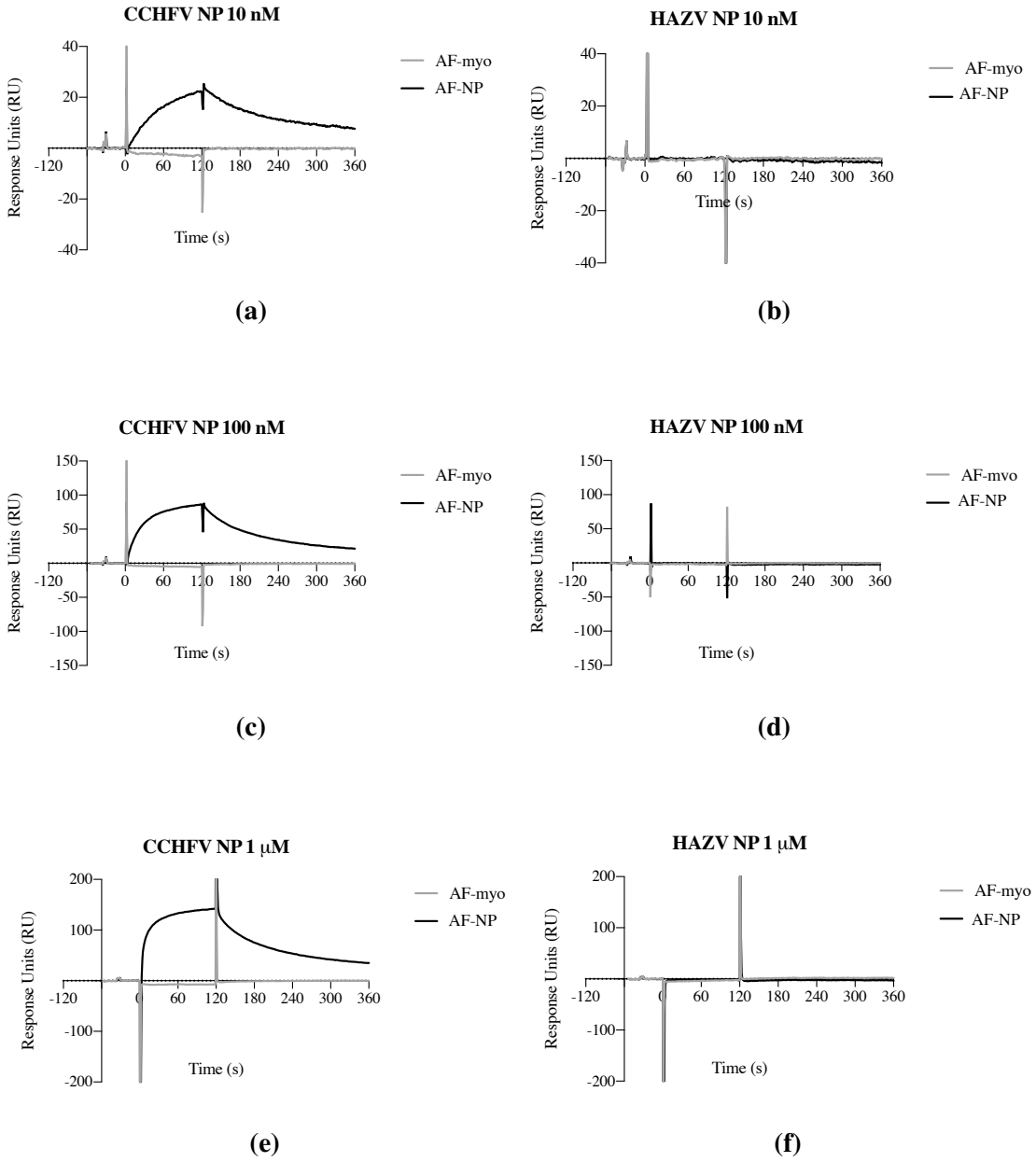
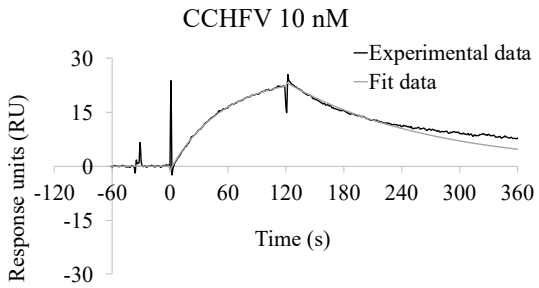
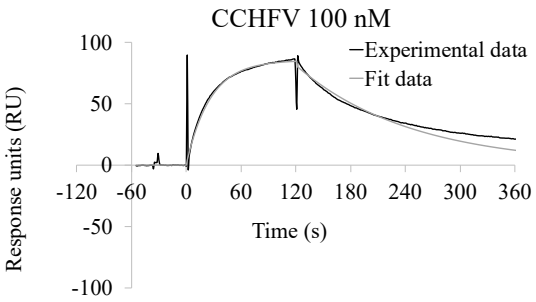


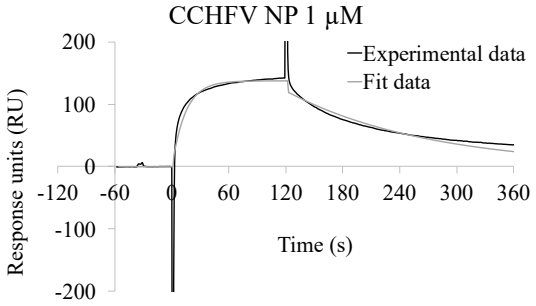
Figure 4.7: SPR analysis of the binding kinetics between Affimers and nairoviral nucleoproteins. **(a, c, e)** SPR sensograms corresponding to ten-fold dilutions of CCHFV NP injected into Affimer-NP or Affimer-myo coated sensor chips. **(b, d, f)** SPR sensograms corresponding to ten-fold dilutions of HAZV NP injected into Affimer-NP or Affimer-myo coated sensor chips.



(a)



(b)



(c)

[CCHFV NP]	K_D	Chi^2
10 nM	5.69 x 10^{-9} M	0.104
100 nM	5 x 10^{-8} M	2.29
1 µM	1.18 x 10^{-7} M	19.4

(d)

[CCHFV NP]	10 nM
k_{on}	1.17 x 10^{6} M^{-1} s^{-1}
k_{off}	6.67 x 10^{-3} s
K_D	5.69 x 10^{-9} M
Chi^2	0.104

(e)

Figure 4.8: SPR data fitting. **(a-c)** Fitting of the SPR sensogram corresponding to the binding between CCHFV NP 10 nM **(a)**, 100 nM **(b)** or 1 µM **(c)** and Affimer-NP to a Langmuir 1:1 binding model. **(d)** K_D and Chi^2 values of the SPR data corresponding to binding between ten-fold dilutions of CCHFV NP and Affimer-NP, fitted to a Langmuir 1:1 binding model. **(e)** Association rate (k_{on}), dissociation rate (k_{off}) and affinity (K_D) constants and Chi^2 value obtained from the fit curve in (a).

4.3 Secondary structure: CD

Introduction to CD: CD spectroscopy measures the differential absorption of right and left handed circularly polarised light by optically active or chiral samples [233]. Circularly polarised light rays travel through optically active medium with different velocities due to the different indices of refraction for right and left handed circularly polarised light, named optical rotation or circular birefringence [234]. Variation of optical rotation as a function of wavelength is called optical rotary dispersion (ORD). The extinction coefficients of a sample for right and left handed polarised light can also be different at certain wavelengths, resulting absorption of the two polarised lights to different extents, phenomenon known as CD [233].

After passing through an optically active sample, the counter-rotating electric (E) vector components of circularly right and left handed polarised light will be different due to the different molar extinction coefficients for these two elements. Transmitted light will still contain two circular electric components, but their magnitude will be unequal, resulting in elliptical polarised light (figure 4.9) [233]. The tangent of the ratio of the minor to major elliptical axis is defined as ellipticity, and it is the unit for measurement of CD. There will also be a directly related rotation of the major axis of the ellipse due to differences in refractive indices (ORD) (represented as α in figure 4.9).

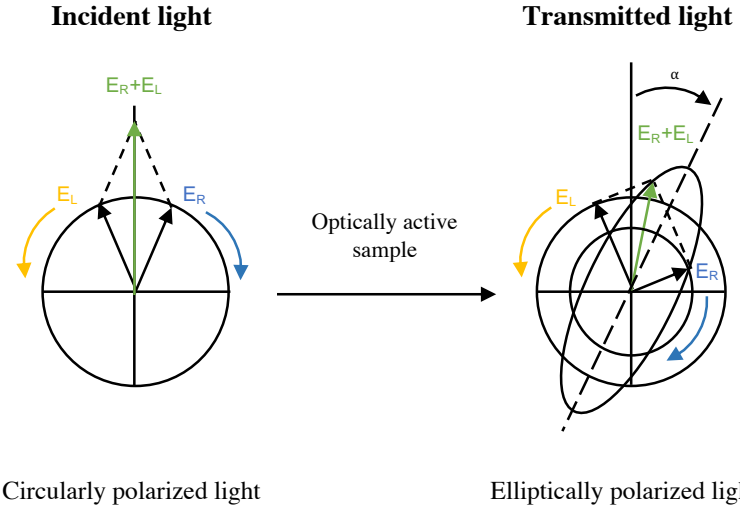


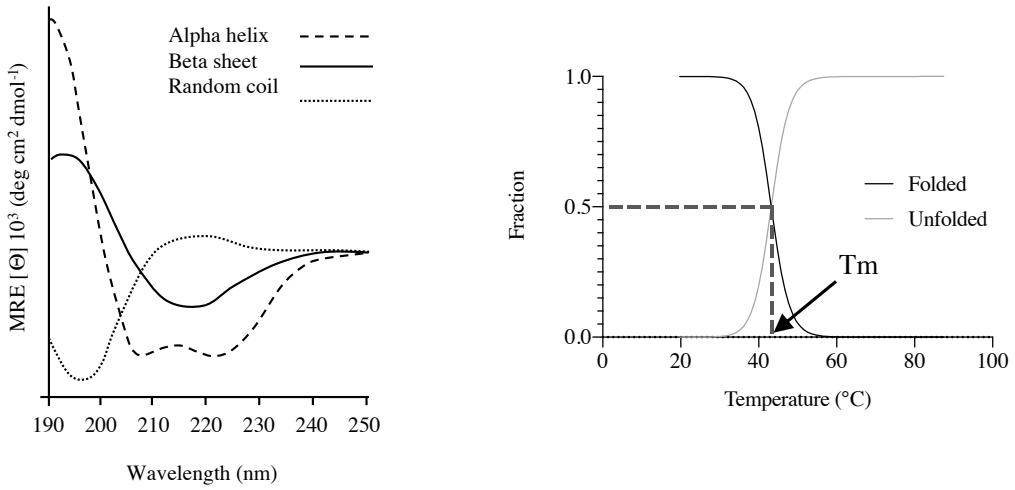
Figure 4.9: Principles of circular dichroism. Circularly polarised light passing through an optically active sample results in different absorptions of the right and left polarisation. E_R and E_L are the magnitudes of the electric field vectors of the right-circularly and left-circularly polarised light, respectively. Elliptical polarised light (right) is composed of unequal contributions of right (E_R) and left (E_L) circular polarised light.

CD can be used for determination of protein secondary structure and monitoring of structural changes. Protein structure can be subdivided into four different levels of organisation [235]. The first level, known as primary structure, consists on the specific sequence of amino acids of each protein. The next level, designed as secondary structure, refers to regular arrangements of the backbone of the polypeptide chain into α-helices, β-strands or random coils. Secondary structures are folded into a compact molecule for the entire polypeptide chain, referred to as tertiary structure. The last level of organisation, namely quaternary structure, describes the arrangement of multiple protein subunits in a regular manner driven by non-covalent forces [235].

Characteristic CD patterns in the far UV (180-250 nm) can be associated to the different secondary structure elements in a protein. α-helical structures present a strong double minimum at 222 and 208-210 nm, and a maximum at 191-193 nm [236]. The CD pattern of beta strands consists on a single minimum between 210 and 225 nm, and a positive maximum between 190 and 200 nm. Unordered proteins and random coils typically show a CD spectrum with a strong minimum near 200 nm and some weak positive or negative bands between 220 and 230 nm (figure 4.10a) [235].

For quantitative secondary structure prediction, ellipticity is normalised to mean residue ellipticity (MRE), which is independent of protein length, protein concentration, and path-length. A CD dataset of MRE versus wavelength can be used to quantify secondary structure content of a protein sample using different approaches and mathematical models that are combined in secondary structure prediction softwares [235]. Additional information about the thermal stability of a protein can be obtained analysing the CD spectra of proteins at different temperatures [237]. The folded and unfolded protein fractions of a sample can be predicted at each temperature. A melting temperature (Tm), defined as the temperature at which half of the sample is unfolded, can be determined as an indication of protein thermal stability (figure 4.10b). Additional information can be obtained from protein CD analyses, such as protein structural changes upon different treatments or binding of ligands [237].

Results: CD analyses of CCHFV and HAZV NPs, Affimer-NP and a 1:1 molar mix of CCHFV NP and Affimer-NP were used for analysis of secondary structure and thermal stability. CD spectra were measured in the range of 190-260 nm at temperatures ranging between 20°C and 90°C, and secondary structure content was predicted using the online prediction tool BeStSel [238]. CCHFV and HAZV NP spectra showed two minima at 208 and 222 nm (figures 4.11a and 4.12a), indicating predominance of α-helices in concordance with both secondary structure prediction analyses (figures 4.15a and 4.15b) and previously published crystal structures [224] (PDB: 4akl and 4xze, respectively,



(a) **(b)**

Figure 4.10: Study of protein secondary structure and thermostability by circular dichroism. **(a)** CD study of protein secondary structure. Characteristic CD patterns in the far UV (180-250 nm) can be associated to the different secondary structure elements in a protein: α-helix, β-sheet or random coil. **(b)** CD study of protein thermostability. Information about the thermal stability of a protein can be obtained analysing CD of proteins at different temperatures. The folded and unfolded protein fractions of a sample can be predicted at each temperature to determine the melting temperature (Tm), defined as the temperature at which half of the sample is unfolded.

figures 4.11d and 4.15b). The Tm obtained for CCHFV and HAZV NP was 31.74°C and 35.63°C, respectively (figures 4.11b, 4.11c, 4.12b and 4.12c), also in agreement with previously published data [151].

Affimer-NP spectra presented a single minimum at 218 nm, indicating a predominance of β-sheets (figure 4.13a), in agreement with the secondary structure prediction software (figure 4.15c), and consistent with previously published structures of the Affimer scaffold (PDB: 4n6u, figure 4.13d). The melting temperature (Tm) obtained for Affimer-NP was 74.37°C, indicating high thermostability of this molecule (figures 4.13b and 4.13c), also previously observed [181].

The CD spectra of the CCHFV NP and Affimer-NP mixture showed two minima at 208 and 222 nm, indicating a predominance of α-helices (figures 4.14a and 4.14d) as expected due to a greater global contribution of CCHFV NP α-helices relative to Affimer-NP β-sheets. This predominance was consistent with the data obtained with the secondary structure prediction software (figure 4.15d). The Tm obtained for the mixture (43.35°C, figures 4.14b and 4.14c) was higher than the Tm of CCHFV NP alone, indicating the presence of Affimer-NP caused a delay in the melting of α-helical structures, as confirmed with the secondary structure prediction software (figure 4.15e), predominantly corresponding to CCHFV NP. Taken together, these CD analyses confirm the appropriate folding of the recombinant proteins and suggests that Affimer-NP induces thermal stabilisation of CCHFV NP.

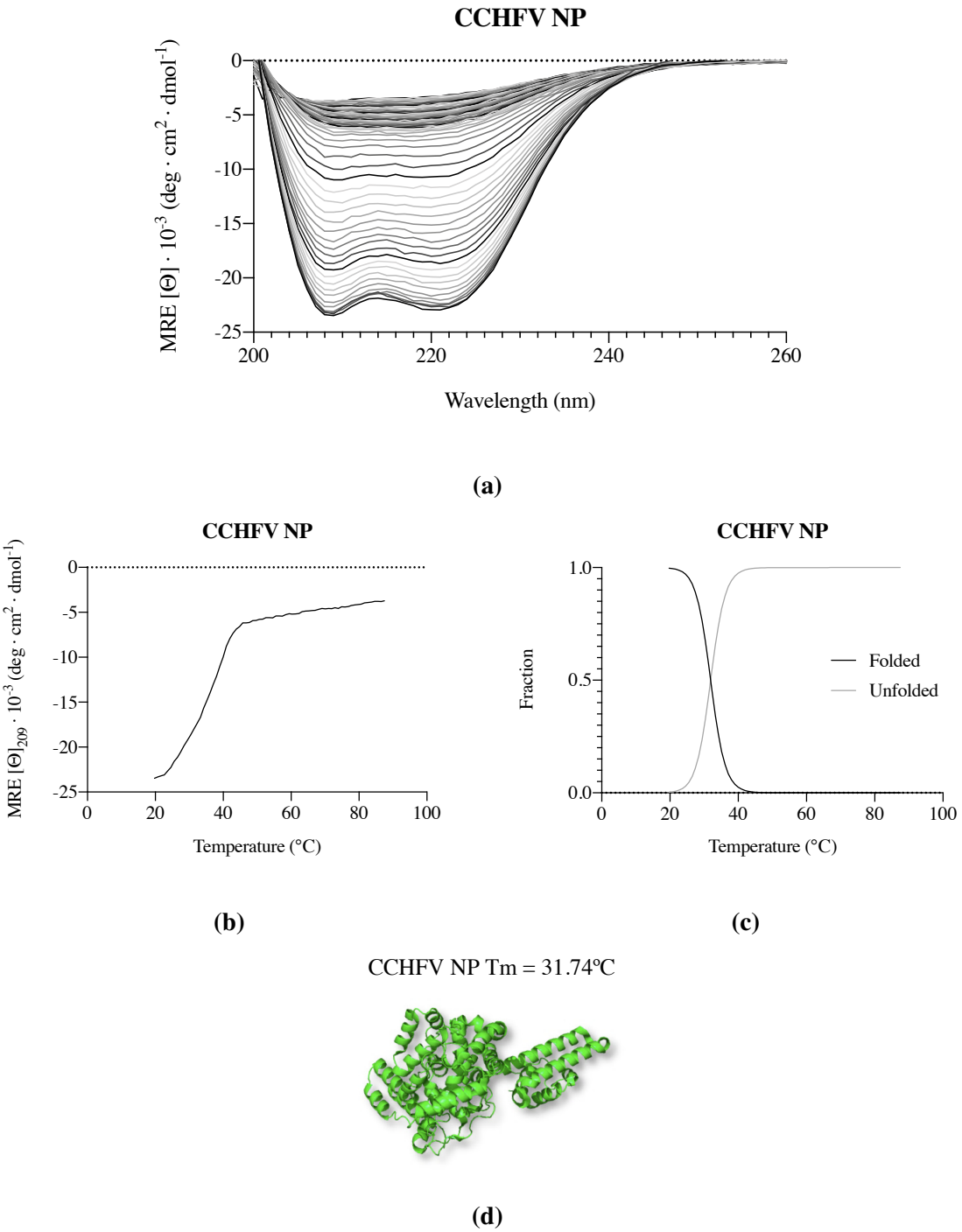


Figure 4.11: Circular dichroism analysis of recombinant CCHFV NP. (a) Circular dichroism curves representing the ellipticity across the far UV spectra of CCHFV NP at different temperatures (20°C to 90°C). **(b)** Representation of CCHFV NP ellipticity values at 209 nm versus temperature. **(c)** Representation of folded and unfolded species of CCHFV NP versus temperature. **(d)** Crystal structure of CCHFV NP (PDB: 4AKL).

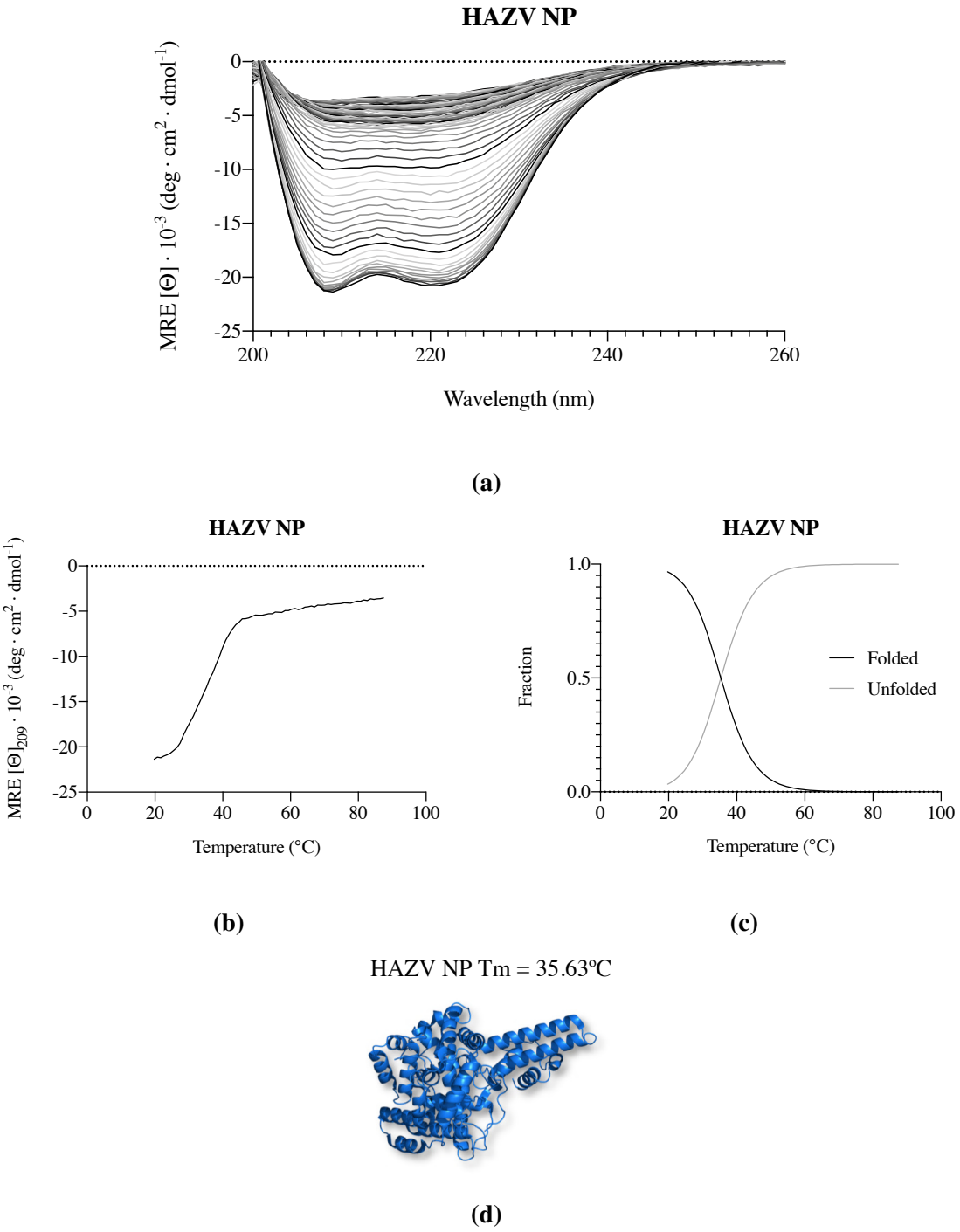


Figure 4.12: Circular dichroism analysis of recombinant HAZV NP. **(a)** Circular dichroism curves representing the ellipticity across the far UV spectra of HAZV NP at different temperatures (20°C to 90°C). **(b)** Representation of HAZV NP ellipticity values at 209 nm versus temperature. **(c)** Representation of folded and unfolded species of HAZV NP versus temperature. **(d)** Crystal structure of HAZV NP (PDB: 4XZE).

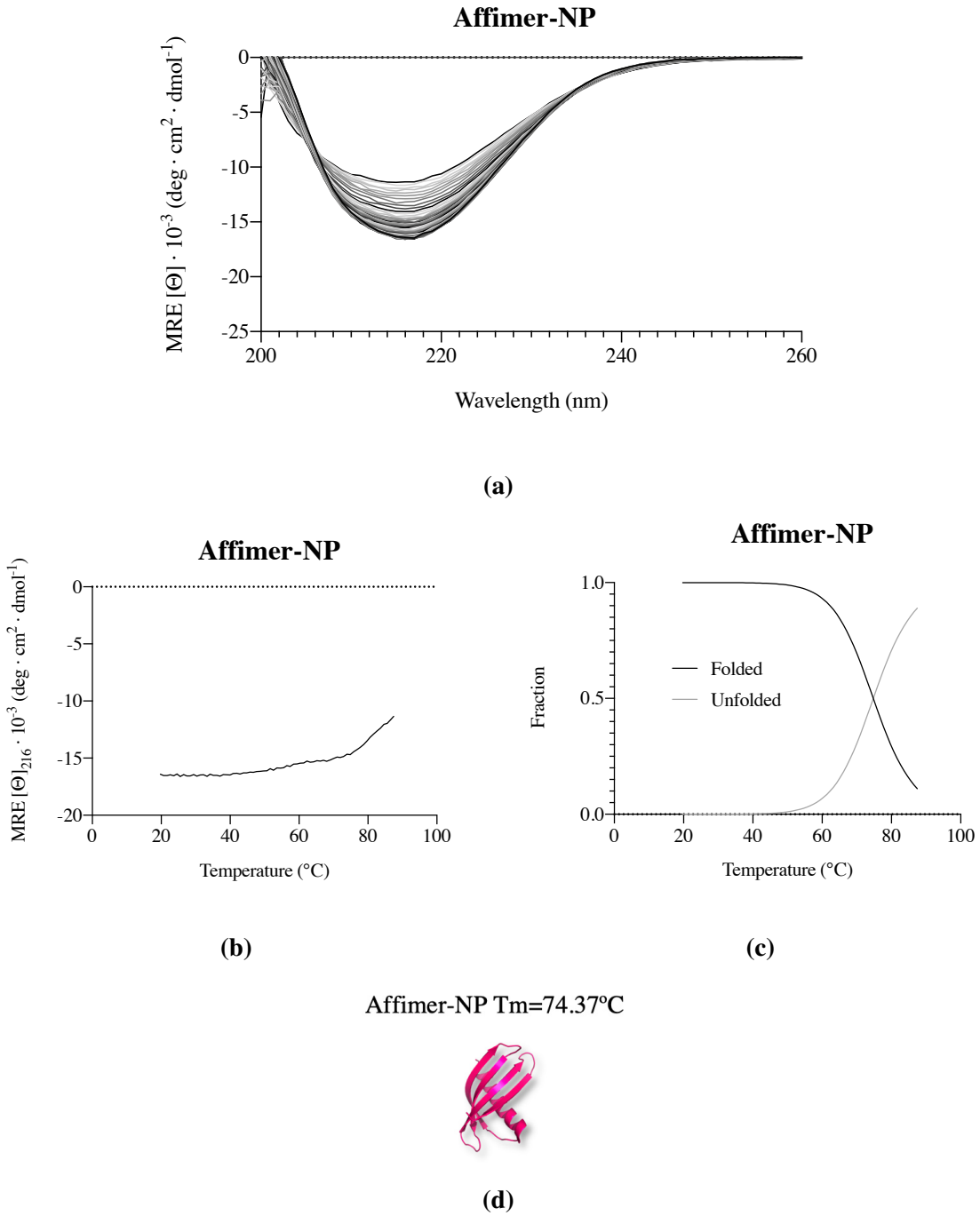


Figure 4.13: Circular dichroism analysis of Affimer-NP. **(a)** Circular dichroism curves representing the ellipticity across the far UV spectra of Affimer-NP at different temperatures (20°C to 90°C). **(b)** Representation of Affimer-NP ellipticity values at 216 nm versus temperature. **(c)** Representation of folded and unfolded species of Affimer-NP versus temperature. **(d)** Crystal structure of Affimer-NP (PDB: 4N6U).

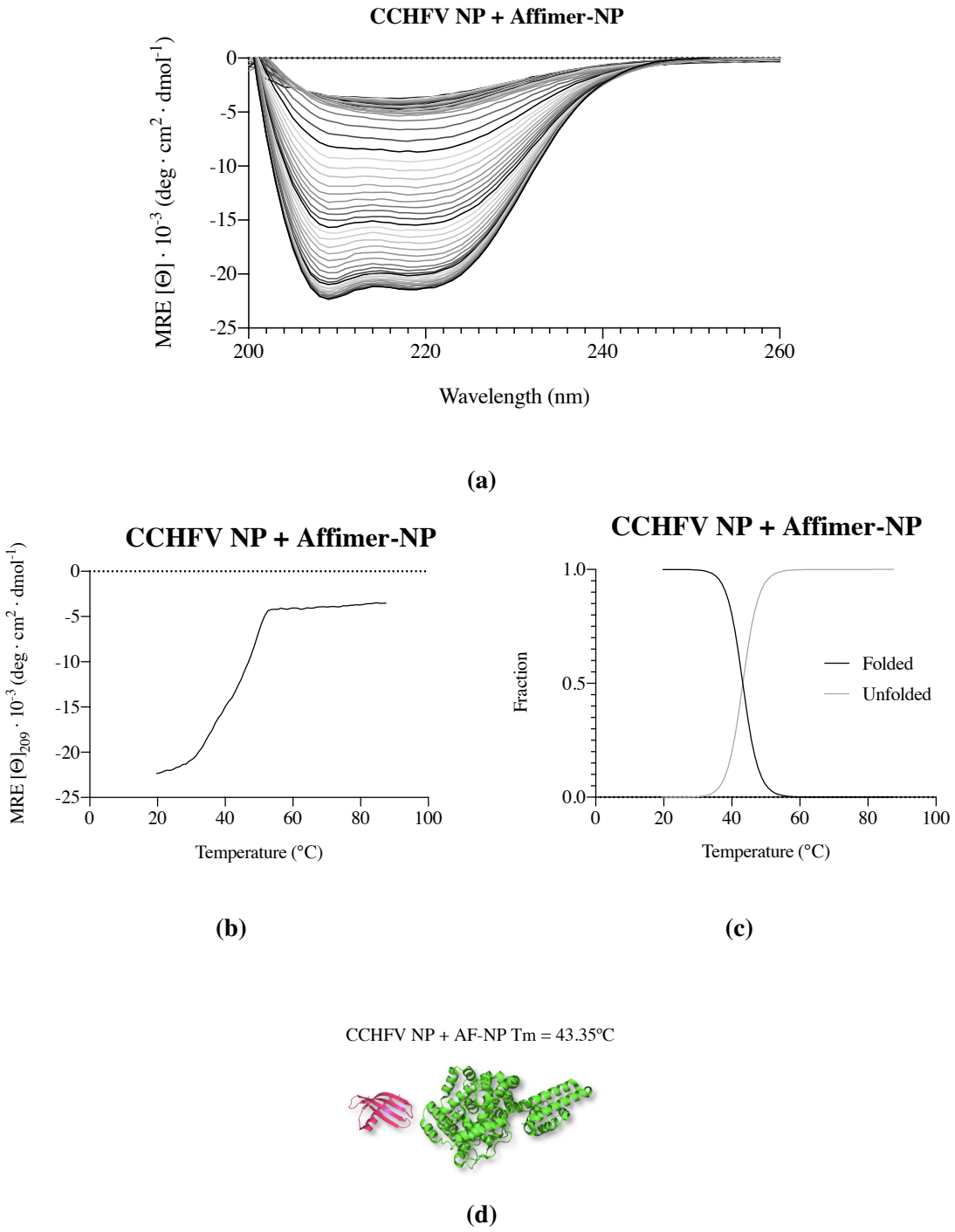


Figure 4.14: Circular dichroism analysis of CCHFV NP and Affimer-NP complex. **(a)** Circular dichroism curves representing the ellipticity across the far UV spectra of CCHFV NP and Affimer-NP complex at different temperatures (20°C to 90°C). **(b)** Representation of CCHFV NP and Affimer-NP complex ellipticity values at 209 nm versus temperature. **(c)** Representation of folded and unfolded species of CCHFV NP and Affimer-NP complex versus temperature. **(d)** Crystal structure of CCHFV NP and Affimer scaffold (PDB: 4AKL and 4N6U, respectively).

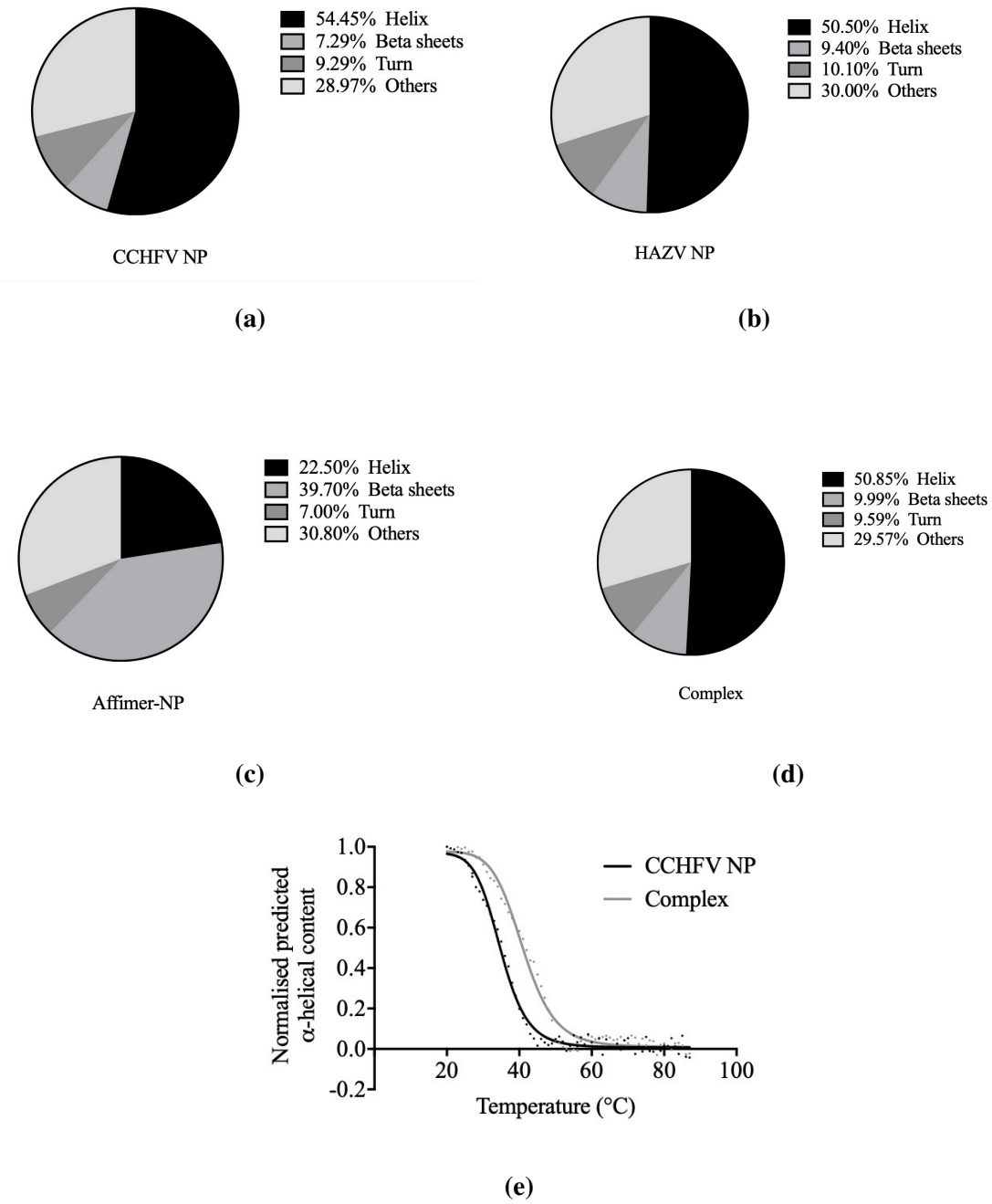


Figure 4.15: Prediction of secondary structure elements. **(a-c)** Percentage of α-helices, β-sheets, turns and other secondary structure elements of CCHFV NP **(a)**, HAZV NP **(b)**, Affimer-NP **(c)** and Affimer-NP/CCHFV NP complex **(d)** at 20°C. **(e)** Normalised predicted α-helical content of CCHFV NP and CCHFV NP + Affimer-NP complex at different temperatures (20°C to 90°C).

4.4 Complex purification: SEC

Introduction to SEC of protein complexes: As previously described in section 3.1.2.2, gel filtration or SEC is a technique commonly used to separate molecules based on their molecular weight. Protein interactions can be studied by SEC mixing a solution of a purified protein with a solution containing a purified candidate interacting protein and performing gel filtration on the mixture [239]. If the two proteins form a complex, an absorbance peak can be detected at the elution volume corresponding to the sum of molecular weight of the two proteins (figure 4.16).

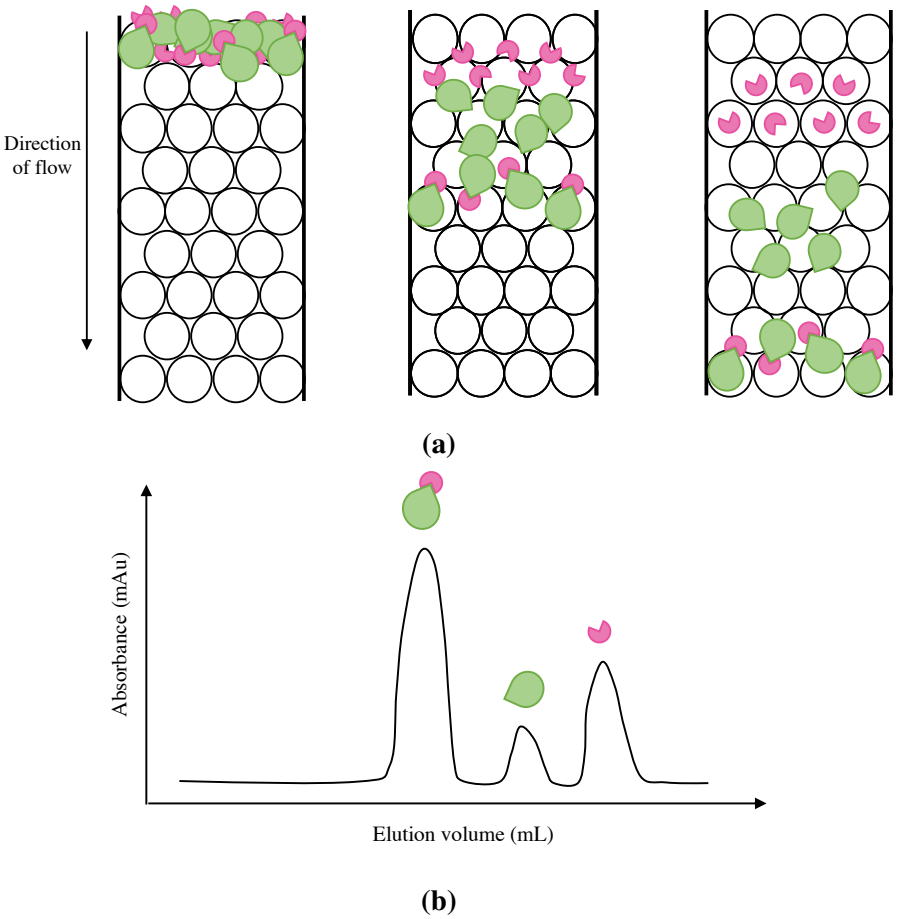


Figure 4.16: Complex purification by size exclusion chromatography. **(a)** Schematic representation of SEC purification of a 1:1 protein complex and migration of the different components. **(b)** Schematic representation of a size exclusion chromatogram corresponding to the purification of a 1:1 protein complex.

Some of the critical matters to be taken into account for studying protein-protein interactions by SEC are the resolution of the column and the buffer to be used. Resolution directly depends on several factors, including column dimensions, pore size of the particles, medium packing density, sample volume, flow rate, and viscosity of the sample

and buffer. The running buffer to be used must be able to maintain the interactions of interest with an appropriate pH, detergent concentration and ion strength [240]. When studying high-affinity interactions, quantitative analyses of the interaction can also be performed by SEC. An idea of the stoichiometry of the interaction can be obtained when mixing different molar ratios of the proteins and performing SEC. A consideration that must be taken into account is that equilibrium may be disturbed during running of sample due to changes in concentration and diffusion.

Fractions containing the molecular complex can be collected to obtain a purified sample and further analysis can be done using additional techniques (e.g. composition analyses by SDS-PAGE, western blot or mass spectrometry, or detailed structural analyses using NMR, EM or X-ray crystallography).

Results: To further characterise the high-affinity binding interaction between Affimer-NP and CCHFV NP, SEC was used for the stoichiometric study of the protein complex and the production of large quantities of purified complex for crystallographic experiments. Affimer-NP expressed from the pET11a vector, with its correspondent C-terminal cysteine and 8xHis-tail, presented several peaks in the SEC analysis when mixed with CCHFV NP, potentially formed by different combinations of monomers or dimers of Affimer-NP bound to CCHFV NP (results not shown). Due to the possibility of the terminal cysteines forming disulphide bonds that increased the heterogeneity of the sample, complex was purified using Affimer-NP produced with the pET-SUMO expression vector, without the additional cysteine nor the 8xHis-tail (as detailed in section 3.5.2).

Recombinantly expressed Affimer-NP and CCHFV NP were mixed and the complex was purified by SEC (figure 4.17). An excess of Affimer-NP (10 kDa) was mixed with CCHFV NP, expecting a 1:1 interaction, due to its easier separation from the 1:1 complex (64 kDa) as a result of a larger difference in molecular weight compared to monomeric CCHFV NP (54 kDa). The corresponding size exclusion chromatogram presented a predominant peak corresponding to the protein complex (peak 1, figure 4.17a) and a small peak corresponding to the excess of Affimer-NP (peak 2). Fractions corresponding to the complex were analysed by SDS-PAGE and Coomassie staining (figure 4.17b), pooled and concentrated to 10 mg/ml. These results corroborate the strong binding interaction between CCHFV NP and Affimer-NP and suggest a 1:1 interaction between these molecules.

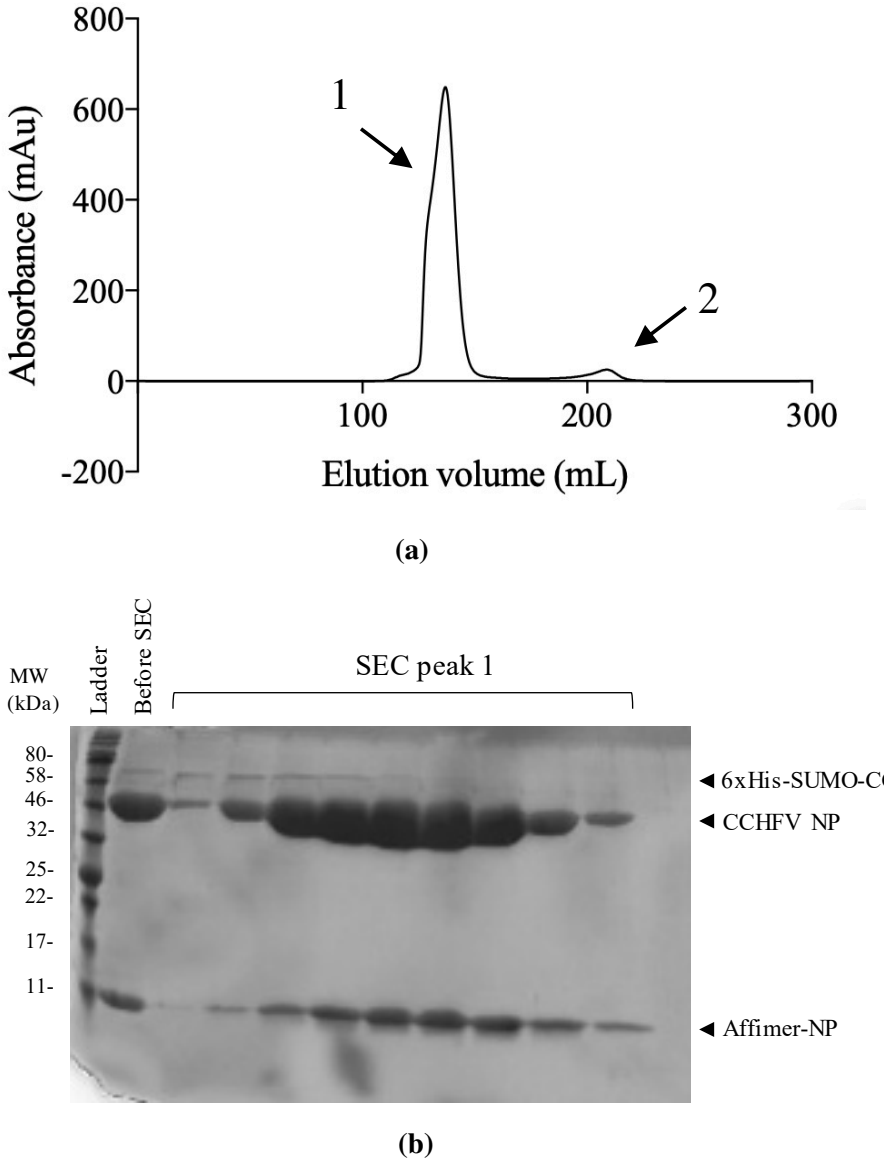


Figure 4.17: Affimer-NP and CCHFV NP complex purification. **(a)** Chromatogram of the size exclusion chromatography of CCHFV NP and Affimer-NP. **(b)** SDS-PAGE analysis and Coomassie staining of the size exclusion chromatography fractions containing CCHFV NP and Affimer-NP (peak 1 of the size exclusion chromatogram in (a)). The left shoulder observed in peak 1 may be caused by the presence of uncleaved 6xHis-SUMO-CCHFV NP (67 kDa), detected in the SDS-PAGE and Coomassie staining analysis of the corresponding fractions. Arrows indicate the elution volume and migration distances of specified proteins.

4.5 Interference of Affimer-NP in RNA binding of CCHFV NP

Introduction to FA: FA is a measure of the extent of polarised light emitted when a fluorophore is excited with a polarised light. This phenomenon is commonly used to study macromolecular interactions, such as RNA binding of proteins [241]. Fluorescently labelled RNAs in solution present a rotational diffusion that causes depolarisation of emitted fluorescence when the fluorophore molecule is excited with a polarised light, presenting low anisotropy values [242]. The extent of depolarisation is dependent on the size and shape of the rotating molecule, along with the viscosity of the solution in which the reaction occurs. In contrast, when a protein binds to the fluorescently labelled RNA, the rotation of the fluorophore decreases, increasing the likelihood that the fluorophore will be in the same plane at the time of light emission as it was at the time of excitation, emitting highly polarised light with increased anisotropy values (figure 4.18). Anisotropy can be quantified from raw emission values parallel to and perpendicular to the excitation polarisation [242].

FA is commonly used to analyse the equilibrium and real-time kinetics of the direct interaction between protein and RNA, by quantifying the anisotropy values at increasing concentrations of protein (figure 4.19a) [243]. The binding capacity of the protein can be quantified by its dissociation constant (K_D). K_D refers to the concentration of protein at which 50% of RNA bound to protein is achieved at equilibrium.

Competition binding studies can also be performed by FA adding a potential competition candidate that interferes with the binding of the RNA to the protein [242]. Increasing concentrations of the interfering molecule will decrease the number of RNA molecules bound to the protein, resulting in a decrease of anisotropy (figure 4.19b). A quantitative measurement of the competition interference can be obtained by analysing the half maximal inhibitory concentration (IC_{50}) of the inhibitor. IC_{50} refers to the concentration of inhibitor required to obtain a 50% reduction in binding.

Results: FA assays were performed to assess whether Affimer-NP interferes with the RNA binding function of CCHFV NP. RNA binding FA assays were performed to assess direct binding of CCHFV NP to three different synthetic RNA oligonucleotides (8 mer, 27 mer and 48 mer RNAs, see table 2.6 in section 2.1.6) labelled with fluorescein (figure 4.20). CCHFV NP showed incomplete binding to 8 mer oligonucleotides, and saturation was not achieved at high CCHFV NP concentration. In the case of the 27 mer and 48 mer RNA oligonucleotides, saturation was achieved at high CCHFV NP concentration and the

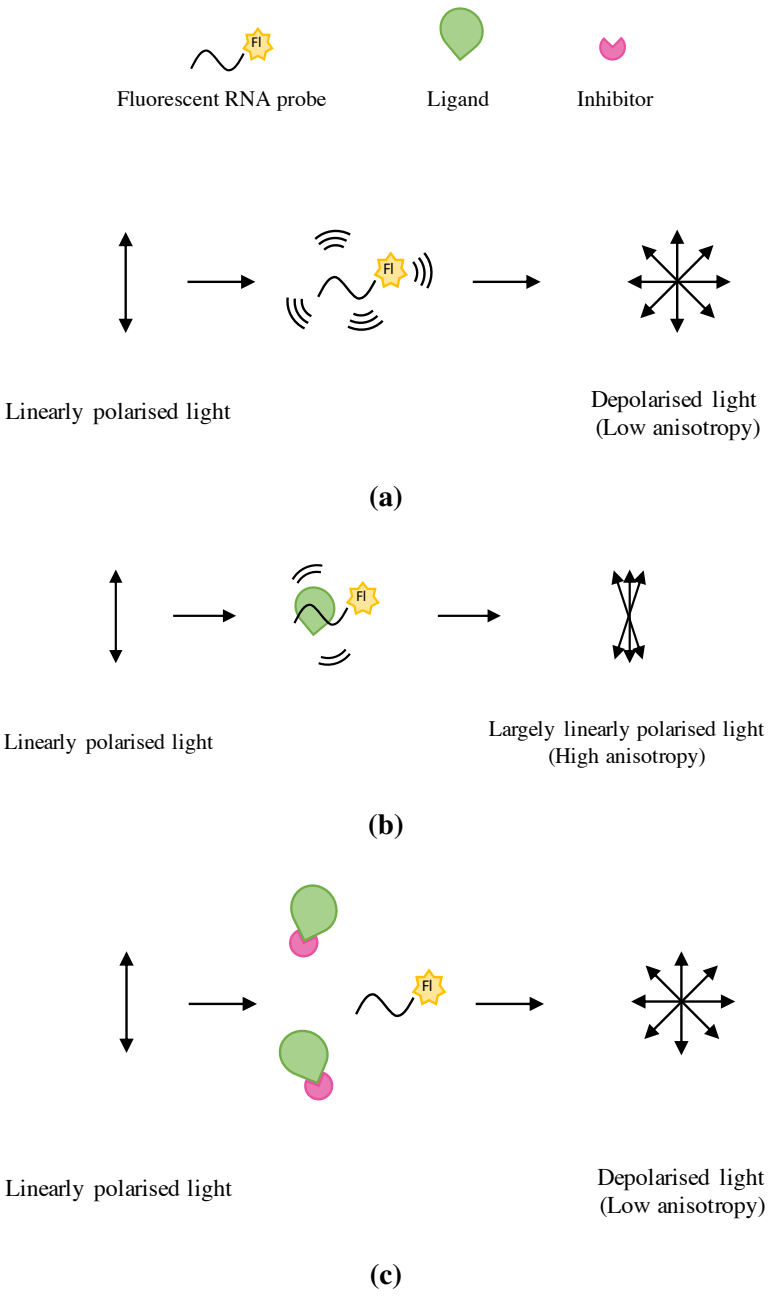


Figure 4.18: Principles of fluorescence anisotropy for the study of RNA-protein interactions. **(a)** Fluorescently labelled RNAs in solution present a rotational diffusion that causes depolarisation of emitted fluorescence when the fluorophore molecule is excited with a polarised light, presenting low anisotropy values. **(b)** When a protein binds to the fluorescently labelled RNA, the rotation of the fluorophore decreases, increasing the likelihood that the fluorophore will be in the same plane at the time of light emission as it was at the time of excitation, emitting highly polarised light with increased anisotropy values. **(c)** Competitive inhibitors of the RNA-protein interaction induce the release of free RNA molecules, decreasing the polarisation of emitted light with reduced anisotropy values.

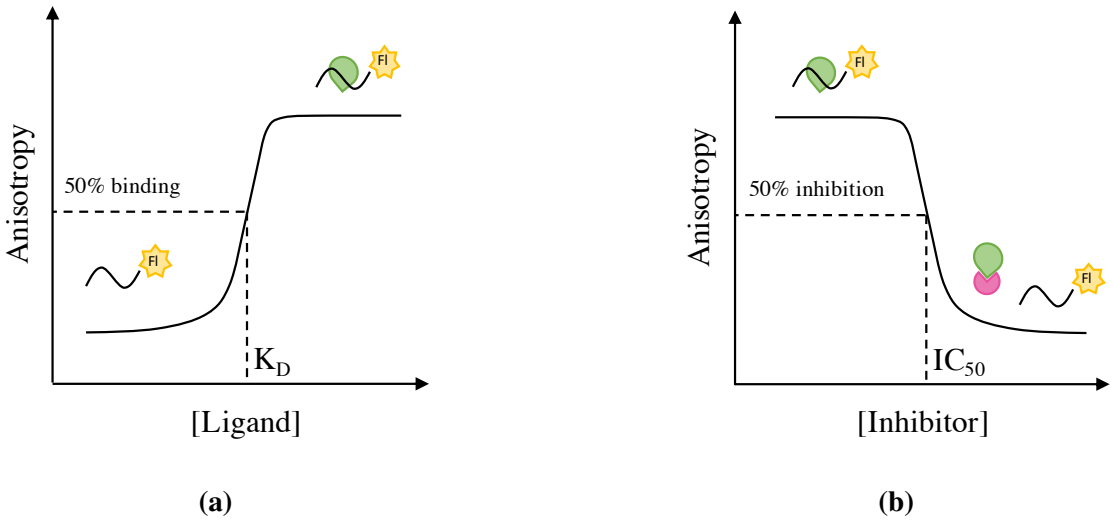
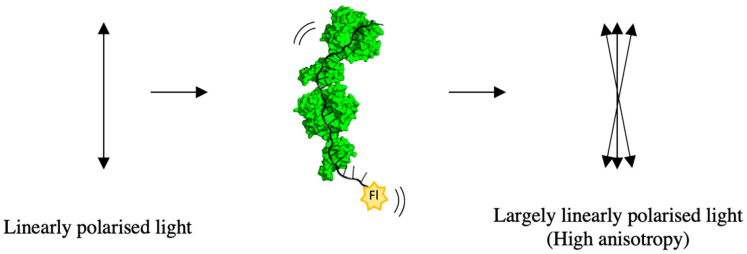


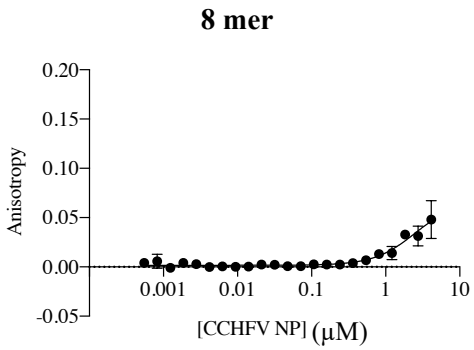
Figure 4.19: Schematic representation of changes in fluorescence anisotropy values versus concentration of ligand or inhibitor. **(a)** Increased concentration of protein results in more protein bound to the fluorescently labelled RNA, consequently emitting highly polarised light with increased anisotropy values. K_D refers to the concentration of protein at which 50% of RNA bound to protein is achieved. **(b)** Increased concentration of competitive inhibitors of the RNA-protein interaction induces the release of free RNA molecules, consequently decreasing the polarisation of emitted light with reduced anisotropy values. IC_{50} refers to the concentration of inhibitor required to obtain a 50% reduction in binding.

K_D of the interaction was determined fitting the binding curves to a non-linear regression model: $K_{D\ 27mer}$ = 0.190 μM ± 0.0062 and $K_{D\ 48mer}$ = 0.230 μM ± 0.008 (figure 4.20e).

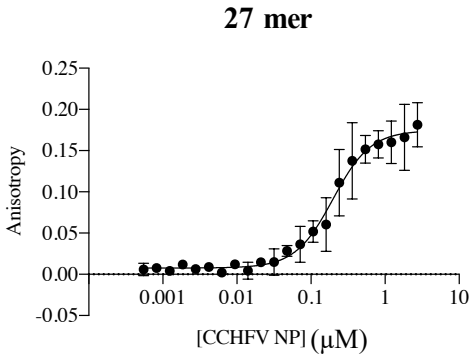
Once the K_D of the direct binding interactions was determined, a competitive assay was performed using Affimer-NP alongside the negative control Affimer-myo. Two different concentrations of CCHFV NP were tested, the highest of the K_D values (230 nM) and twice this concentration (460 nM), to ensure that the interference of Affimer-NP could be properly measured. Affimer-NP showed interference in the RNA binding of CCHFV NP (230 nM) with IC_{50} values of $IC_{50\ 27mer}$ = 0.197 μM ± 0.0084 and $IC_{50\ 48mer}$ = 0.167 μM ± 0.0046, respectively (figure 4.21, table 4.2). For higher concentration of CCHFV NP (460 nM), IC_{50} values of $IC_{50\ 27mer}$ = 0.790 μM ± 0.0029 and $IC_{50\ 48mer}$ = 0.570 μM ± 0.0034 were obtained (figure 4.22, Table 4.2). The negative control Affimer-myo, showed lack of interference in the RNA binding function of CCHFV NP. These observations support a specific interference of Affimer-NP in the RNA binding function of CCHFV NP.



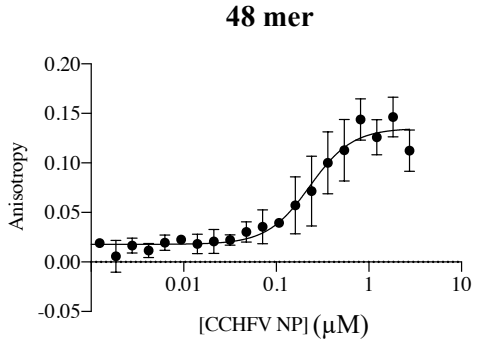
(a)



(b)



(c)



(d)

RNA	K_D
27 mer	0.190 μM ± 0.0062
48 mer	0.230 μM ± 0.008

(e)

Figure 4.20: Fluorescence anisotropy analysis of the RNA binding function of CCHFV NP. **(a)** Binding of CCHFV NP to fluorescently labelled RNA causes an increase in polarised light emission and anisotropy. **(b-d)** Binding of CCHFV NP to 8 mer, 27 mer (a) or 48 mer (b) synthetic RNA molecules (2.5 nM). **(e)** K_D values obtained for the binding of CCHFV NP to 27 mer and 48 mer RNA molecules. Data are presented as mean ± SD (n = 3 replicates) and are fitted to a non-linear regression curve.

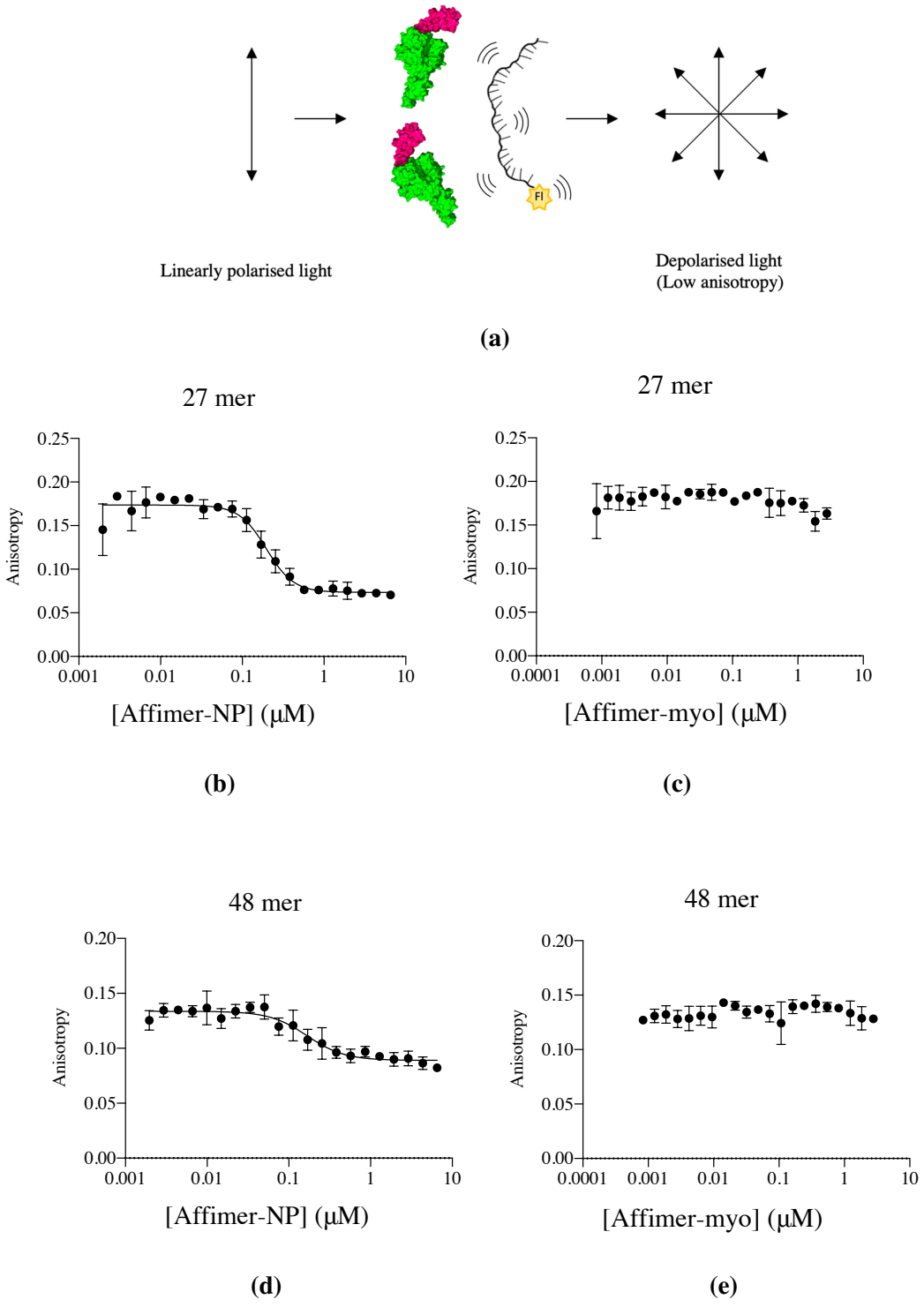


Figure 4.21: Interference of Affimer-NP in CCHFV NP RNA binding function. (a) Binding of Affimer-NP to CCHFV NP induces the release of fluorescently labelled RNA molecules increasing the emission of depolarised light and subsequently reducing anisotropy values. **(b-e)** Fluorescence anisotropy analysis of the interference of Affimer-NP in CCHFV NP RNA binding using 27mer (b) or 48mer (d) RNA molecules (2.5 nM) and a concentration of 230 nM CCHFV NP. Negative control experiments using Affimer-myo are represented in (c) and (e). Data are presented as mean ± SD (n = 3 replicates) and are fitted to a non-linear regression curve.

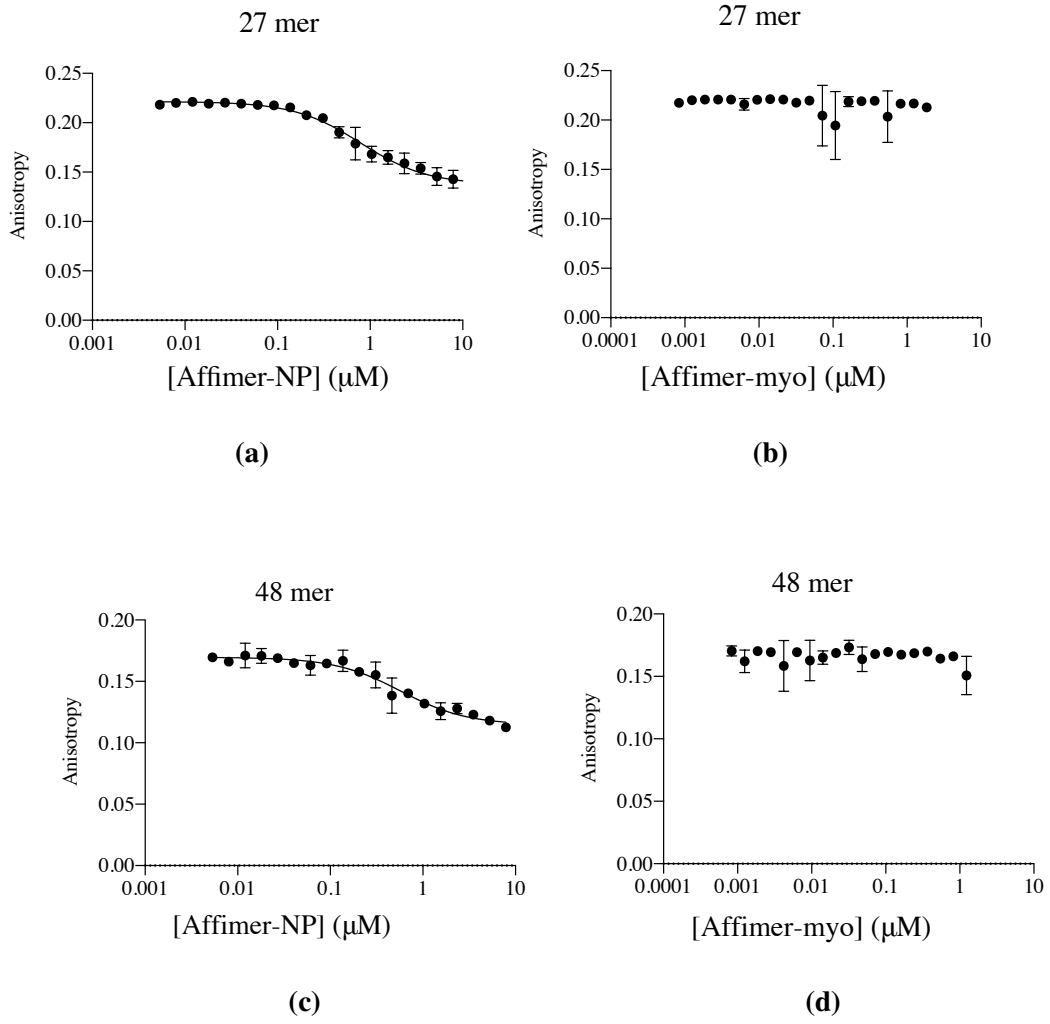


Figure 4.22: Interference of Affimer-NP in CCHFV NP RNA binding function. (a-d) Fluorescence anisotropy analysis of the interference of Affimer-NP in CCHFV NP RNA binding using 27mer (a) or 48mer (c) RNA molecules (2.5 nM) and a concentration of 460 nM CCHFV NP. Negative control experiments using Affimer-myo are represented in (b) and (d). Data are presented as mean ± SD (n = 3 replicates) and are fitted to a non-linear regression curve.

[CCHFV NP]	230 nM	460 nM
$IC_{50\ 27mer}$	0.197 μM ± 0.0084	0.790 μM ± 0.0029
$IC_{50\ 48\ mer}$	0.167 μM ± 0.0046	0.570 μM ± 0.0034

Table 4.2: IC_{50} values for the interference of Affimer-NP in CCHFV NP RNA binding function. IC_{50} values obtained for the interference of Affimer-NP in CCHFV NP binding to 27 mer or 48 mer RNA molecules (2.5 nM). Two different concentrations of CCHFV NP (230 and 460 nM) were tested.

4.6 Interference of Affimer-NP in CCHFV replication

The competition of Affimer-NP in the RNA binding activity of CCHFV NP, which is critical for formation of CCHFV RNPs, raised the possibility that Affimer-NP would also interfere with CCHFV gene expression. To test this hypothesis, a CCHFV mini-genome system that uses eGFP as reporter of gene expression was established. This system was combined with transient mammalian expression of Affimer-NP to determine its effects in CCHFV-specific genome replication.

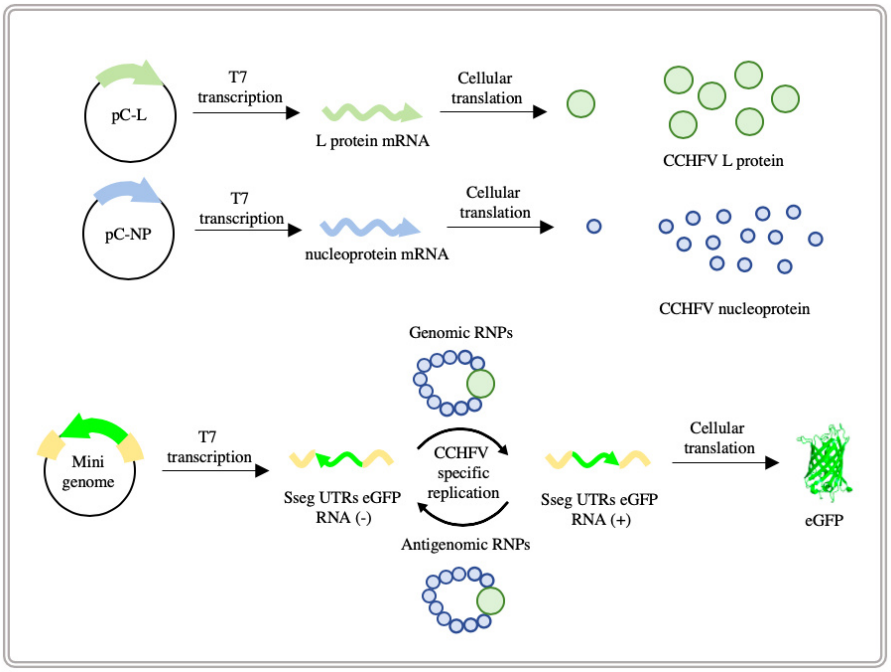


Figure 4.23: CCHFV mini-genome system. The CCHFV mini-genome system is formed by three plasmids: a NP support plasmid (pC-NP), an L protein support plasmid (pC-L) and a mini-genome plasmid with eGFP as reporter of CCHFV-specific viral replication (pC-SsegUTRs-eGFP). The NP and L protein support plasmids produce the corresponding proteins, which assemble into virus-like RNPs with the mini-genome RNA containing a negative sense copy of the eGFP ORF. After virus-like replication, positive sense mini-genome RNA is produced and translated into eGFP. eGFP fluorescent signal can be measured as a reporter for CCHFV-specific replication.

Plasmid construction: Briefly, support plasmids (pC-NP and pC-L) expressing CCHFV NP and L proteins were generated by inserting the corresponding cDNA sequences from strain Baghdad-12 downstream of the T7 pol promoter and a cDNA representing the internal ribosome entry site (IRES) of encephalomyocarditis virus (EMCV).

The CCHFV S segment replicon was generated using the strain Baghdad-12 as

reference. The S segment cDNA sequence was incorporated within a plasmid vector flanked by the T7 pol promoter and HDV RZ, and orientated for the expression of viral sense RNA. The S segment major ORF was replaced by the enhanced green fluorescence protein (eGFP) ORF to generate pC-SsegUTRs-eGFP expressing eGFP as reporter signal of CCHFV-specific replication. Plasmid backbone containing the Sseg UTRs and eGFP ORF were amplified by PCR using primers containing *Nhe*I and *Not*I restriction sites (figure 4.24a). The eGFP ORF was subcloned into the Sseg UTRs plasmid backbone using restriction enzyme digestion and ligation to generate pC-SsegUTRs-eGFP (see plasmid maps in Appendix B).

The coding sequence of Affimer-NP fused to RFP was subcloned into a pMAXcloning vector for its expression in mammalian cells (pMAX-AF-NP-RFP). A negative control plasmid was generated substituting the Affimer-NP sequence for a yeast SUMO specific Affimer (pMAX-AF-YS-RFP). Briefly, AF-YS-RFP sequence was amplified by PCR and subcloned into pMAX expression vector using restriction enzyme digestion and ligation to generate pMAX-AF-YS-RFP. The ORF of Affimer-YS-RFP was replaced by PCR-amplified Affimer-NP-RFP ORF using restriction enzyme digestion and ligation to generate pMAX-AF-NP-RFP (figure 4.24b, see plasmid map in Appendix B).

Results: The CCHFV mini-genome system was initially validated by transfection of the three plasmids (pC-NP, pC-L and pC-SsegUTRs-eGFP) into BSR-T7 cells, alongside a control transfection (pC-NP and pC-SsegUTRs-eGFP) omitting transfection of the essential pC-L plasmid. eGFP signal was monitored over time using an Incucyte live cell imaging system. An increase in the fluorescent signal was observed in cells transfected with all mini-genome plasmids, compared to non-transfected and negative control cells lacking the pC-L plasmid (figure 4.25).

Different concentrations of the Affimer-NP-RFP plasmid (50, 100, 250, 500 and 750 ng) were co-transfected along with plasmids encoding the CCHFV mini-genome components into BSR-T7 cells. *In vivo* green and red fluorescence signals were monitored using an Incucyte system, corresponding to the eGFP reporter of the mini-genome system and the Affimer-RFP expression, respectively. Co-transfection of the replicon with 750 ng of the control Affimer-YS-RFP plasmid resulted in no statistically significant change of the eGFP reporter signal (figures 4.26a and 4.26b). Conversely, co-transfection of increasing amounts of Affimer-NP-RFP with CCHFV replicon was correlated with a statistically significant decrease in the eGFP mini-genome reporter signal (figure 4.26). These results illustrate a specific and dose dependent interference of Affimer-NP in CCHFV mini-genome replication.

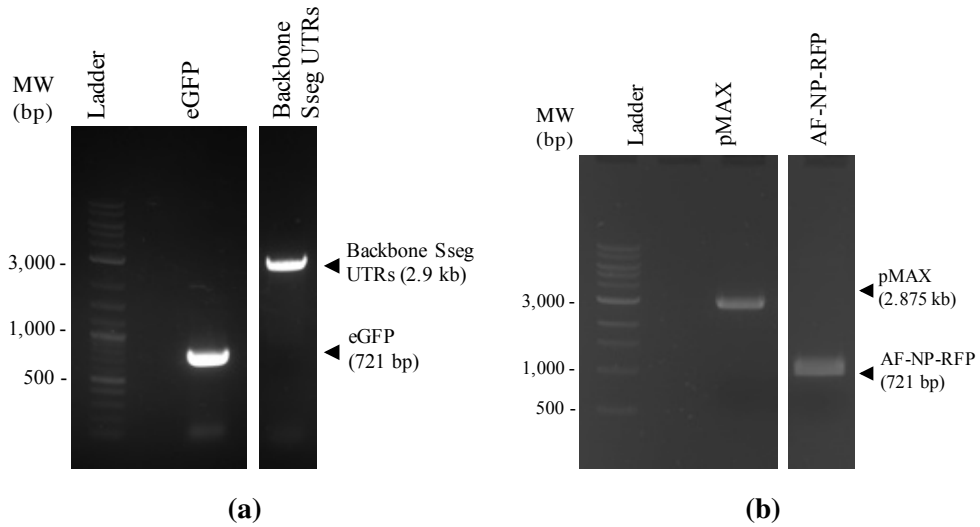


Figure 4.24: Digested DNA products for subcloning of CCHFV S segment eGFP replicon and Affimer-NP mammalian expression vector. **(a)** UV transillumination imaging of agarose gel electrophoresis analysis of the eGFP ORF and S segment UTRs backbone vector after *Nhe*I and *Not*I restriction enzymes digestion. **(b)** UV transillumination imaging of agarose gel electrophoresis analysis of the Affimer-NP ORF and pMAXcloning backbone after *Sac*I and *Xba*I restriction enzymes digestion. DNA was stained using SYBR safe DNA stain. Arrows indicate migration distances, name and size of the corresponding DNA products.

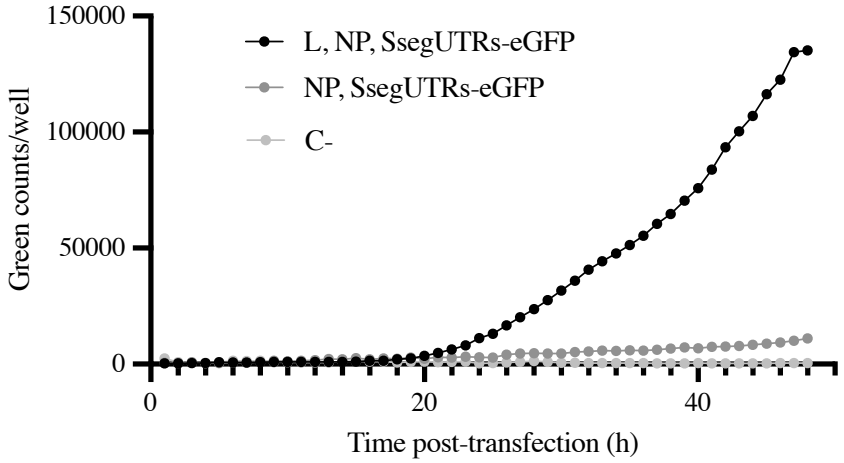


Figure 4.25: eGFP reporter signal associated to CCHFV-specific gene expression. **(a)** BSR-T7 cells were transfected with the CCHFV mini-genome plasmids (pC-NP, pC-L and pC-SsegUTRs-eGFP), alongside a control transfection (pC-NP and pC-SsegUTRs-eGFP) omitting transfection of the essential pC-L plasmid and non-transfected cells (C-). eGFP signal (green counts/well) was monitored over time using an Incucyte live cell imaging system.

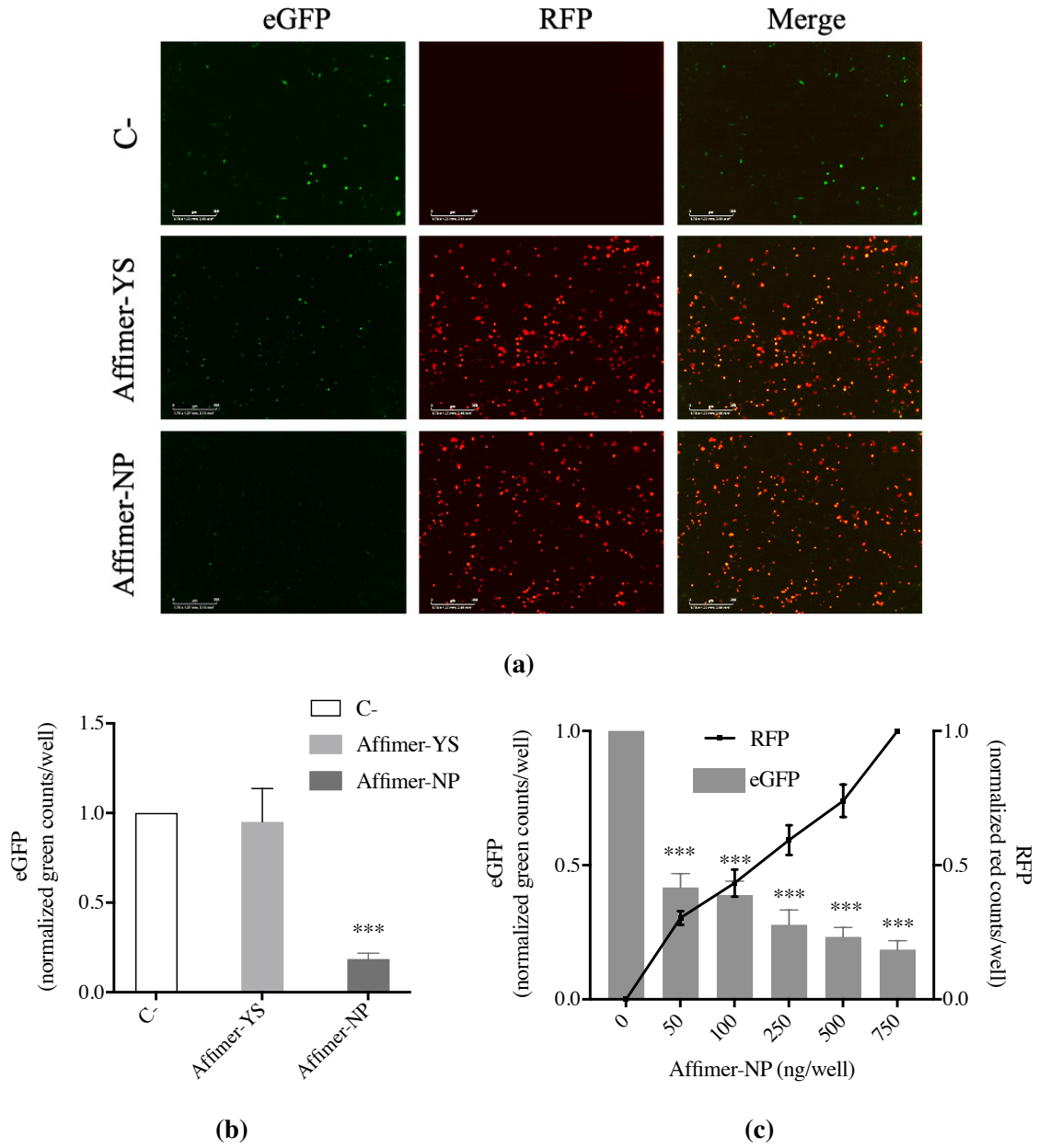


Figure 4.26: Inhibition of CCHFV-specific gene expression by Affimer-NP (a) IncuCyte fluorescence pictures taken 72 h post-transfection of CCHFV replicon plus 750 ng of Affimer-NP, Affimer-YS or no Affimer (C-). Scalebar: 300 µm. **(b)** CCHFV replicon eGFP reporter signal (normalised green counts per well) 72 h post-co-transfection with 750 ng of Affimer-NP, Affimer-YS or no Affimer (C-). **(c)** eGFP and RFP signal (normalised green or red counts per well) corresponding to CCHFV replicon reporter and Affimer-NP-RFP, respectively, 72 h post-co-transfection of CCHFV replicon plasmids with different amounts of Affimer-NP (50, 100, 250, 500 and 750 ng). Data are presented as mean ± SD (n = 3 replicates). Statistical analysis of data in (b) and (c) was performed using a paired t-test (ns = non-significant, *** = $p<0.001$).

4.7 Tertiary structure: X-ray crystallography

Introduction to X-ray crystallography: Tertiary structure of proteins can be determined using different techniques that include NMR, cryo-EM or X-ray crystallography, the latter being considered the most accurate technique, with cryo-EM approaching similar resolution limits.

X-ray protein crystallography is based on the analysis of the diffraction pattern of X-rays when they pass through a protein crystal [244]. Most of the photons that pass through the crystal are totally totally transmitted, without interacting with the protein molecules, but occasionally, a photon from the X-ray source interacts with the electron density around the atoms of the proteins contained within the crystal and is scattered [245]. Emerging rays interact with each other either in a constructive or a destructive way, resulting in waves with higher or lower amplitudes, respectively. Rays passing through the crystal hit a detector placed in front of the crystal where the reflection pattern is recorded, including the intensities of all reflected waves (figure 4.27). The dataset of reflections and their intensities is known as diffraction pattern, and it is the raw data used for the determination of the tertiary structure of the molecules within the crystal [244].

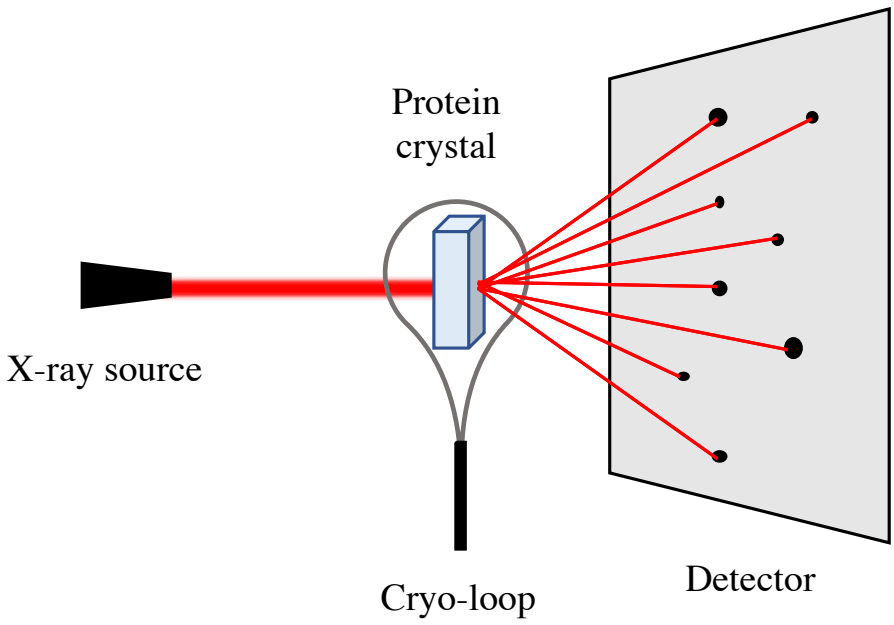


Figure 4.27: X-ray protein crystallography experimental set up. X-ray protein crystallography is based on the analysis of the diffraction pattern of X-rays when they pass through a protein crystal, commonly positioned in front of the X-ray source using a cryo-loop. Emerging rays hit a detector placed in front of the crystal where the reflection pattern is recorded, including the intensities of all reflected waves.

4.7.0.1 Crystal growth

The first, and probably most challenging, step of an X-ray crystallography experiment is the growth of protein crystals. Large quantities of homogeneous, pure, soluble and concentrated protein are mixed with different buffers, precipitants and salts with the aim of obtaining well ordered, high quality and well-diffracting protein crystals [246]. High-throughput techniques can be used for this purpose, including automated systems such as multi-pipetting robots for setting up crystal screens, making buffers and monitoring crystal growth. Crystal growth is a necessary step due to the fact that the intensity of diffracted X-rays from a single protein molecule is too weak to be detected even by the most sophisticated equipment [245]. Instead, protein crystals contain multiple copies of the molecules, arranged in a periodic, well-ordered lattice, which amplify the signal of the diffracted X-rays enabling their detection.

Crystal growth occurs when an undersaturated solution of protein and precipitant is concentrated to supersaturation without reaching precipitation, remaining in a phase known as the nucleation zone [246]. When the phase of the solution moves into the supersaturated nucleation zone, spontaneous nucleation or ordered aggregation can occur. Once nucleation crystals are formed, concentration of protein in solution decreases and the phase of the solution changes to a metastable zone, promoting the growth in size of the crystals (figure 4.28).

The most common procedure for crystal growth is known as vapour diffusion [246]. A droplet of a solution containing protein and precipitant is placed in a closed system next to a reservoir of the same solution containing a higher concentration of precipitant (mother liquor solution). Differences in precipitant concentration promote water evaporation from the droplet, leading to an equilibrium on the concentration of precipitants between the two solutions (droplet and reservoir). As water evaporates, the concentration of protein and precipitant in the droplet increases, and the phase of the solution moves into supersaturation, sometimes leading to spontaneous nucleation [245]. Concentration of protein in the droplet subsequently decreases, moving the phase of the droplet solution into the metastable zone and promoting the growth in size of the crystal. Droplets can be suspended underneath coverslips over a mother liquor reservoir (hanging drop set up) or can be placed on platforms over a mother liquor reservoir (sitting drop set up) (figures 4.29a and 4.29b).

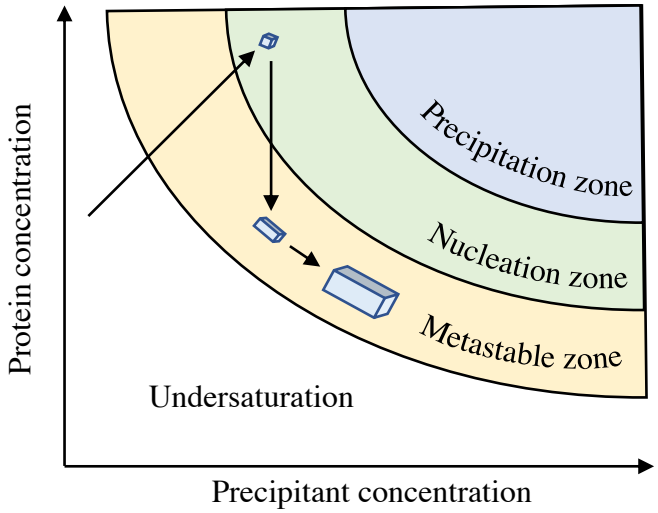


Figure 4.28: Crystallisation phase diagram. Schematic representation of a two-dimensional phase diagram, illustrating the change of protein concentration versus precipitant concentration. The concentration space is divided by the solubility curve into two areas corresponding to undersaturated and supersaturated state of a protein solution. The supersaturated area comprises of the metastable, nucleation and precipitation zones. Crystal growth occurs when an undersaturated solution of protein and precipitant is concentrated to supersaturation without reaching precipitation. Spontaneous nucleation occurs in this phase, decreasing concentration of protein in solution and changing the phase of the solution to a metastable zone, promoting the growth in size of the crystals.

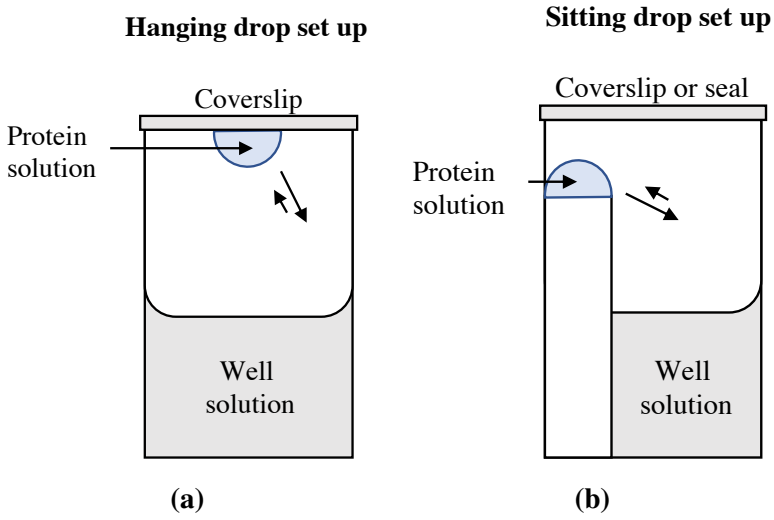


Figure 4.29: Crystal growth by vapour diffusion. (a-b) Schematic representation of crystal growth by vapour diffusion using hanging drop (a) or sitting drop (b) experimental set ups.

4.7.0.2 Crystal arrangement

Protein crystals are ordered and symmetric (figure 4.30a). The asymmetric unit (AU) is defined as the smallest portion of a crystal structure to which symmetry operations (rotations, translations and screw axes (combinations of rotation and translation)) can be applied to recreate the entire unit cell [245].

One or more AUs form the unit cell, which is defined as the smallest crystal repeating unit. Unit cells are described by vectors a, b and c along axes x, y and z, respectively, and angles α, β and γ (figure 4.30b). The space group of the crystal describes how the AUs are arranged in the unit cell. There is a finite number of space groups as there is a finite number of ways in which molecules can be arranged in space.

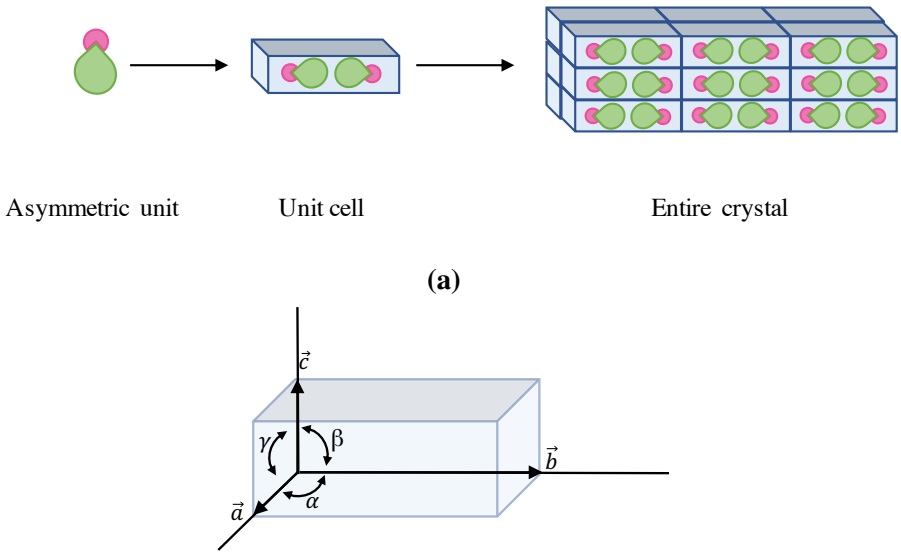


Figure 4.30: Crystal arrangement and the unit cell. (a) The asymmetric unit is defined as the smallest portion of a crystal structure to which symmetry operations can be applied. One or more asymmetric units form the unit cell, which is defined as the smallest crystal repeating unit. **(b)** Unit cells are described by vectors a, b and c along axes x, y and z, respectively, and angles α, β and γ.

4.7.0.3 Bragg's law and reflection intensities

Bragg's law defines whether a particular set of reflected waves will be constructively combined to generate a reflection. Each unit cell within the crystal lattice can be intersected by a set of parallel planes (Bragg's planes) at a fixed point along their axis [247]. Taking d as the distance between two Bragg's planes, two incoming X-ray waves of wavelength λ and angle of incidence θ will only remain in phase, and thus combine

constructively, after reflecting from the crystal if the additional distance the second wave must travel (*2d sinθ*) is a multiple (*n*) of the incoming wavelength (λ) (figure 4.31):

$$n\lambda = 2dsin\theta \tag{4.6}$$

where:
n is an integer
λ is the wavelength of X-rays
d is the spacing between the planes in the atomic lattice
θ is the angle between the incident ray and the scattering planes

Each diffracted wave in every unit cell throughout the whole crystal constructively combines in a wave that generates a reflection on the detector. This is due to the fact that crystals are symmetrically arranged and Bragg's planes intersect each unit cell in the same way, therefore each unit cell contributes equally to the emerging wave [245].

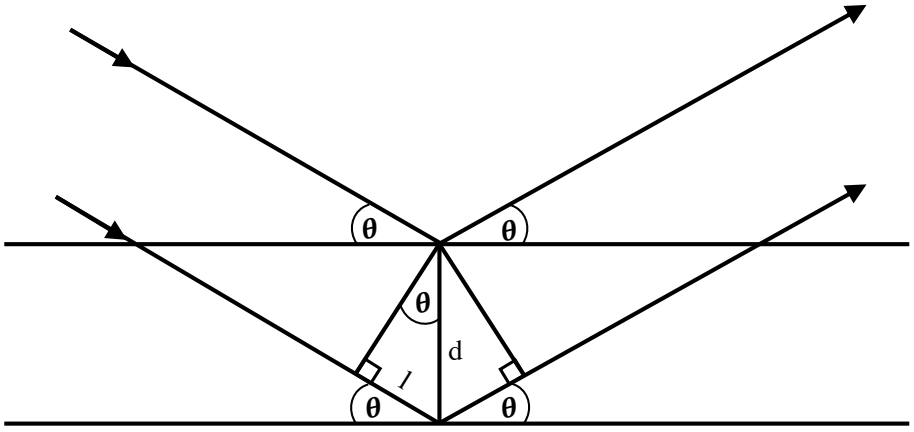


Figure 4.31: Bragg's law reflection. Two incoming X-ray waves of wavelength λ and angle of incidence θ will only remain in phase, and thus combine constructively, after reflecting from the crystal if the additional distance the second wave must travel (*2d sinθ = 2l*) is a multiple (*n*) of the incoming wavelength (λ).

The way in which Bragg's planes intersect each unit cell is described by Miller indices (hkl), which are assigned to each plane that contributes to the diffraction pattern. All the contributing waves across the whole crystal that result in a single reflection can be added together in the structure factor equation, which results in a structure factor F_{hkl} being assigned to each individual reflection, which describes its amplitude and phase [245]. The structure factor amplitude $|F_{hkl}|$ is proportional to the square root of

the reflection intensity. During data collection, each diffraction pattern only provides information about the intensities and their positions from a single direction of the crystal, therefore the crystal must be rotated to collect a complete dataset. A different set of resultant waves will be detected at each angle as the crystal is rotated, as a different set of planes will satisfy Bragg's law at each angle [245].

4.7.0.4 The phase problem: molecular replacement

The spread of individual intensities of all symmetry-equivalent reflections, contributing to the same unique reflection, is averaged in a process known as 'merging', and results in a set of unique reflection intensities, each accompanied by an estimate of error [244]. But there is a vital piece of information necessary to generate the corresponding electron density map that is not provided by the diffraction pattern, namely the phase of the waves that generate each reflection. The 'phase problem' can be solved by different methods depending on the available information on related protein structures [248]. In the majority of cases phases are derived either by using the atomic coordinates of a structurally similar protein using a technique known as molecular replacement (MR) or by finding the positions of heavy atoms that are intrinsic to the protein or that have been added (direct methods, isomorphous replacement or anomalous scattering).

MR enables the solution of the crystallographic phase problem by providing initial estimates of the phases of the new structure from atomic coordinates of a previously known homologous structure [248]. This method relies on the examination of the different possible orientations and positions of the known molecule in the unknown crystal AU to find those where the predicted diffraction best matches the observed diffraction.

For this purpose, MR is based on the utilisation of Patterson maps [248]. These are derived from the Fourier transform (defined as a way of describing a wave as a sum of all its component waves) of the structure factor amplitudes, setting all phases to zero, and describe the distance vectors between atoms. There are two types of vectors, intramolecular vectors, representing the distance between two atoms within a molecule, and intermolecular vectors, representing the distance between two atoms of adjacent molecules. Intramolecular vectors are dependent on orientation, whereas intermolecular vectors depend on both the location and orientation of the molecule. Intermolecular vectors are thus commonly used for determination of the translation function once the orientation is known [245].

The way in which the vectors are presented in a Patterson map derives from the protein structure, and during MR the orientation and location of the unknown molecule

and model are compared through correlation of the intra- and intermolecular vectors of their corresponding Patterson maps [248]. Each molecule needs six parameters to define its orientation and position: 3 rotation angles, which define a rotation matrix that moves the model molecule into the proper orientation, and 3 translation parameters, which need to be specified to place the model structure in the correct location in the AU. If there are N molecules in the AU, then a total of 6N parameters are needed to define the solution.

4.7.0.5 Model-building and refinement process

Once the phase and intensity of the waves that generate each reflection are known, an electron density map can be obtained from the experimental structure factors by a Fourier transform. Comparing the electron density maps of the known and unknown structures, features in the calculated electron density map that are missing in the known molecule can be built using the sequence of the molecules in the crystal [245]. The process of model-building can be initiated by automated methods, but it generally involves laborious manual adjustments, including modifications on the position of residues, torsion angles or bond lengths, especially tedious for low resolution data sets. The accuracy and completeness of model building is a function of the quality of the starting electron density map [244].

The initial macromolecular model then undergoes a refinement procedure in which the parameters of the model are adjusted to best fit the experimental data and stereochemical expectations such as Ramachandran constrains (explained below in this section). As the model improves, the electron density maps improve as well and a better model may be fitted [244]. Model building and refinement are therefore linked to each other, and should be regarded as one unified process. The refinement process improves the calculated phases and minimises differences between the calculated and observed structure factor amplitudes by improving the atomic positions in the model. The corresponding electron density map is then used together with the experimental diffraction pattern to improve the model that fits it, in an iterative cycle [245]. This process can introduce bias into the model, so different statistical calculations and quality indicators are used to assess the models generated [244]. Some of the parameters that determine the quality of the diffraction data set and the corresponding atomic model are briefly summarised below:

Quality of the diffraction data

Resolution limits. Resolution is measured in Å and can be defined as the minimum spacing of crystal lattice planes that provide measurable diffraction of X-rays. Resolution defines the level of detail, or the minimum distance between structural features that can be distinguished in the electron-density maps.

Number of unique reflections. Multiple observations of the same and symmetry related reflections are merged into one unique reflection during the 'scaling and merging' process. The unique reflections number is defined by the resolution of the diffraction data, the size of the unit cell and the space group.

Redundancy. The total number of reflections divided by the number of unique measurements defines the redundancy of the data. This metric lists how many times on average each unique reflection was measured, providing an estimate of the accuracy one should expect from these measurements.

Completeness. Defines how many of the total possible number of unique reflections were indeed measured. Because of the properties of Fourier transforms, each value of the electron-density map is correctly calculated only with the contribution of all reflections, thus lack of completeness will negatively influence the quality and interpretability of the maps computed.

Mean((I)/sd(I)). 'I' refers to measured intensity values for reflections, while 'sd(I)' is the estimated standard deviation (SD) in the measurement of intensity values. Thus, (I)/sd(I), or signal to noise, refers to the average degree to which measured reflection intensities stand out over background. Structures with high signal to noise values (e.g. 15–20) indicate that the data are strong and imply that the quality of the data is high.

Half-set correlation CC(1/2). CC1/2 is the Pearson correlation coefficient (a statistic measurement of the linear correlation between two variables) obtained by comparing two sets of intensity estimates obtained by partition of the measurements used to estimate the intensities of each reflection into two non-overlapping sets of equal size. CC1/2 is a primary indicator for selection of high resolution cut-off for data processing and is related to the effective signal to noise of the data.

Quality of the structure model

R-work. The R-work value is a measure of the agreement between the structure factor amplitudes of the crystallographic model and the structure factor amplitudes of the experimental data. In practice, acceptable R-work values are below the 10% of the resolution of the data (e.g. for a dataset that diffracted to 2.8 Å, the R-value should be $<$ 0.28).

R-free. Before refinement begins, about 5% of the experimental observations are removed from the data set, and refinement is then performed using the remaining 95%. The R-free value is a measure of how well the refined model can predict the original structure factor amplitudes from the 5% data that was omitted from the refinement process. This is considered to be less biased than the R-work and is commonly 5-7% higher than the R-work.

RMS(angles). This parameter determines how well the final crystallographic model conforms to expected values of bond angles. It is a statistical measurement (root mean squared deviation) of the deviation of the bond angles in the model compared to an accepted set of values based on the geometry of small organic molecules. The variation of the refined model should not be more than 2° from the accepted values.

RMS(bonds). This parameter is analogous to RMS(angles) but refers to bond lengths, and determines how well the final crystallographic model conforms to expected values of bond lengths. It is a statistical measurement (root mean squared deviation) of the deviation of the bond lengths in the model compared to an accepted set of values based on the geometry of small organic molecules. Typical values are smaller than 0.02 Å.

Ramachandran outliers and favoured residues. Peptide bonds are planar, and other backbone conformational angles (φ: N – Cα and ψ: Cα– C , where C is the carbonyl carbon) are also restricted due to steric repulsion between adjacent side chains. The Ramachandran diagram represents the allowed pairs of angles (φ and ψ) on either side of the Cα residue with no steric repulsion and the regions with modest and high repulsion.

Rotamer outliers. Protein side chains take on preferred conformations which fall into distinct local energy minima known as rotamers. Rotamer outliers define the residues with a conformation that lies outside the outlier contours of a reference dataset.

Cβ outliers. This parameter determines distortion of the Cβ position from the ideal one. Cβ deviation is sensitive to incompatibilities between side chain and backbone caused by misfit conformations or inappropriate refinement restraints.

Molprobity clashscore and overall score. The **Molprobity clashscore** is derived

from the number of pairs of atoms in the model that are unusually close to each other. This value is expressed as the number of serious clashes per 1000 atoms. The **overall score** is a single number that represents the central MolProbity protein quality statistics [249]. It is a log-weighted combination of the clashscore, percentage of not-favoured Ramachandran residues and percentage of bad side-chain rotamers, giving one number that reflects the crystallographic resolution at which those values would be expected. A structure with a numerically lower MolProbity score than its actual crystallographic resolution is, quality-wise, better than the average structure at that resolution.

Results: To further characterise the interaction between Affimer-NP and CCHFV NP and determine the molecular basis for the interference of Affimer-NP with CCHFV NP RNA binding and gene expression functions, the high-resolution crystal structure of the CCHFV NP/Affimer-NP complex was solved (figure 4.32). Purified complex was concentrated and crystal trials were performed using a high-throughput sitting-drop experimental set up. Optimisation of crystal growth was attempted through micro-seeding and hanging drop experiments, but best diffraction data was achieved with crystals belonging to the Joint Centre for Structural Genomics (JCSG) core suite screen (1:1 (v/v) ratio of purified protein solution to reservoir solution containing 0.2 M $NH_4CH_3CO_2$ (salt), 0.1 M CH_3COONa 4.6 pH (buffer) and 30% (w/v) PEG 4K (precipitant), figure 4.33). For vitrification, crystals were cryo-protected by transfer to a solution of mother liquor containing 25% (v/v) glycerol.

Diffraction data were collected at Diamond Light Source beamline I04 and the structure of the complex was solved at 2.84 Å resolution by MR using previously published models of the Affimer scaffold (PDB: 4n6u) and CCHFV NP (PDB: 4akl) with Phaser (data collection and refinement statistics are detailed in table 4.3). Two complexes of Affimer-NP and CCHFV NP were elucidated per AU (figure 4.34). The structure shows the variable loops of Affimer-NP bound to the globular domain of CCHFV NP, mainly interacting with the residues Q329, F330, F332, E333, K336, R339, K343, A347, T351, N399, D402 and L405 (figure 4.35). The interaction between CCHFV NP and Affimer-NP involved 5 hydrogen bonds and 4 salt bridges, which were quantitated using Chimera and PDBePISA (figure 4.36).

As previously published structures show, the CCHFV NP structure contains two major domains: a globular core and an extended flexible stalk. Although no crystal structure of CCHFV NP in complex with RNA has been solved to date, analysis of the electrostatic surface potential of the NP has revealed a continuous positively charged region, likely to be a binding site for the viral RNA. Some of the residues described to be involved in the RNA binding of CCHFV NP include R45, K72, S149, H197, K222, R225,

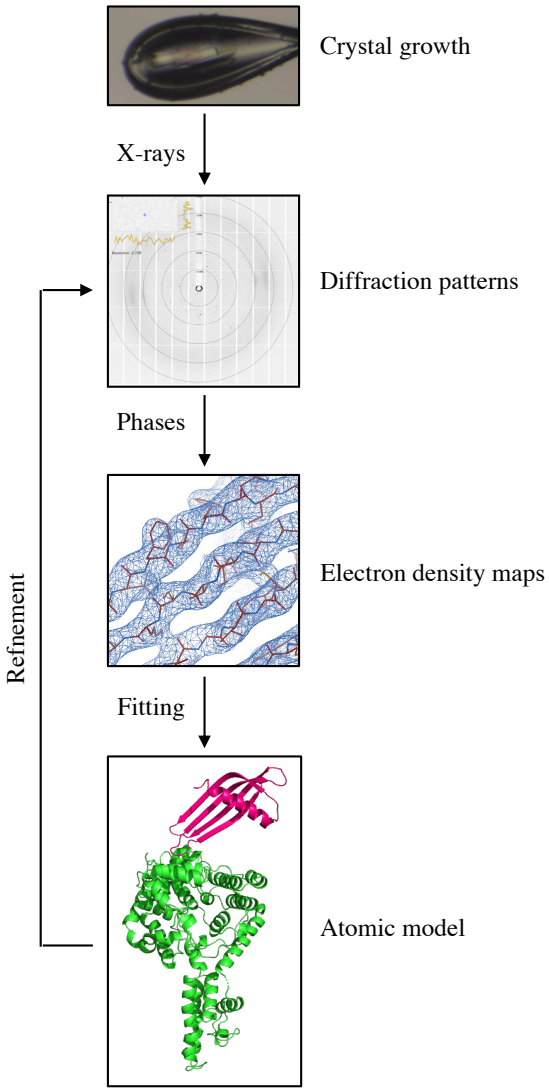
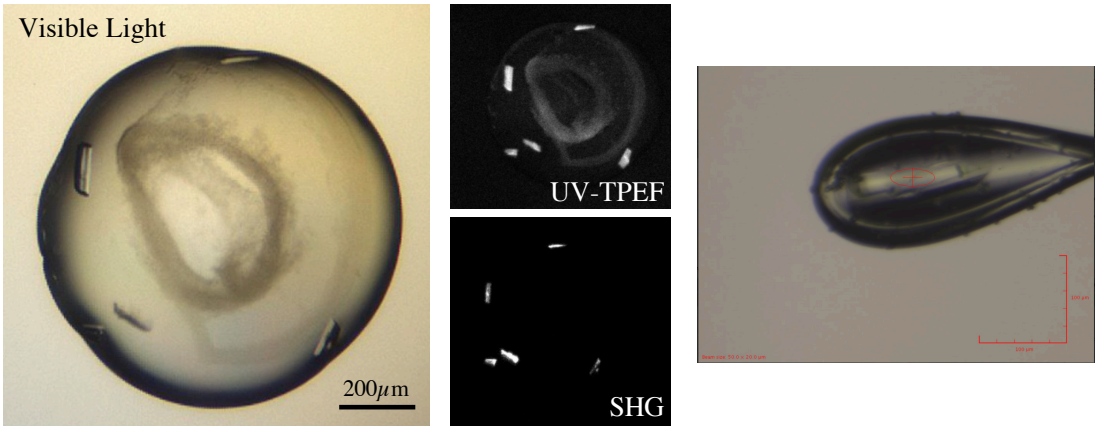
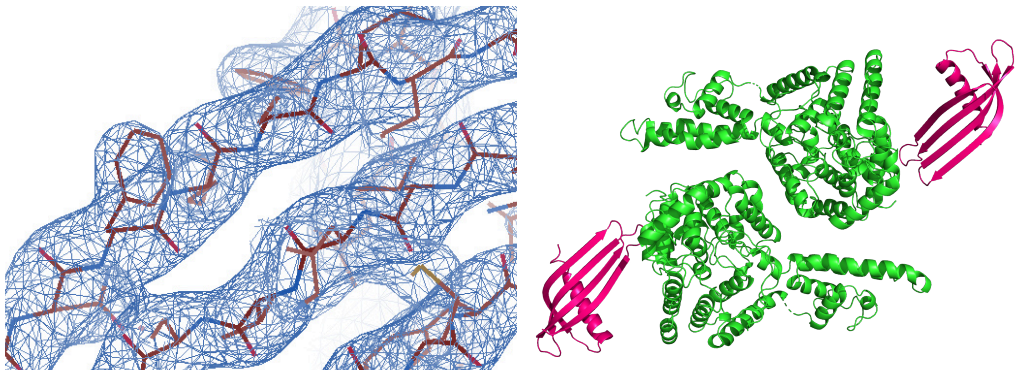


Figure 4.32: Stages of structure determination by X-ray crystallography. Protein expression and purification is followed by crystal growth. When quality crystals are obtained, diffraction data is collected to obtain the reflection intensities. After solution of the phase problem, an electron density map is calculated and an atomic model is built. After several rounds of refinement, a valid structural model can be obtained.



(a) **(b)**

Figure 4.33: Protein crystals containing CCHFV NP/Affimer-NP complex. **(a)** Visible light, two-photon excited ultraviolet fluorescence (UV-TPEF) and second-harmonic generation (SHG) pictures of the protein crystals taken by Rock Imager. **(b)** Cryo-loop containing one of the CCHFV NP/Affimer-NP complex protein crystals at the Diamond Light Source synchrotron.



(a) **(b)**

Figure 4.34: Electron density map and structural model of the asymmetric unit of protein crystals containing CCHFV NP/Affimer-NP complex. **(a)** Atomic model and electron density map of the complex with a contour level of 1σ. Model visualised using Coot. **(b)** Representation of an asymmetric unit containing two CCHFV NP/Affimer-NP complexes. Image created using PyMOL.

Data collection			
Unit cell parameters	a = 90.33 b = 73.98 c = 95.76 (Å) α = 90.00° β= 100.92° γ= 90.00°		
Space group	P 21		
	Overall	**Inner**	**Outer**
Low resolution limit	94.03	94.03	2.71
High resolution limit	2.60	9.00	2.6
Number unique reflections	26028	965	126
Mean((I)/sd(I))	10.3	23.3	1.1
Half-set correlation CC(1/2)	0.998	0.998	0.219
Completeness %	67.1	98.9	2.6
Overall, resolution estimate	2.84Å		
Refinement			
R-work	0.2331		
R-free	0.2665		
RMS (angles)	0.96		
RMS (bonds)	0.005		
Ramachandran outliers	0.6%		
Ramachandran favoured	93.1%		
Rotamer outliers	5.2%		
C-beta outliers	0		
Molprobity Clashscore	4.02		
Molprobity Overall score	2.17		
PDB ID	6Z0O		

Table 4.3: CCHFV NP/Affimer-NP complex X-ray crystallography data collection and refinement statistics.

K237, Q303, K336, R339, K342, K343, K345, R372, and K46236 (highlighted in blue in figure 4.35). The variable loops of Affimer-NP show interaction contacts with some of these residues (K336, R339 and K343), suggesting a potential direct interference in the RNA binding function of the CCHFV NP.

Structural analyses of CCHFV NP also suggest that its oligomerisation plays a role in its RNA binding function. CCHFV NP appears as a monomer in the absence of RNA, but different oligomeric forms of the protein have been described when bound to RNA of the expression host [149]. The interaction between monomers involves residues 320 to 354 at the base of the head domain of one molecule and residues 210 to 219 and 260 to 272 of the arm of the adjacent molecule [145] (residues highlighted in pink in figure 4.35). Some of the residues of the globular domain that interact with the variable loops of Affimer-NP lie between NP residues 320 to 354, suggesting a potential interference of Affimer-NP in the oligomerisation of CCHFV NP.

The specificity of Affimer-NP for CCHFV NP, showing no cross-reactivity with HAZV NP, may also be explained by the crystal structure of the complex. Of the 12 main residues of the CCHFV NP involved in the interaction with the Affimer-NP variable loops (highlighted in orange in figure 4.37), 3 are not conserved in the HAZV NP (N399, L405, K336, highlighted in blue in figure 4.37). Affimers present a high selectivity, even for different analogues of small molecules [201], suggesting that these 3 amino acid changes could be responsible for the selectivity of Affimer-NP for CCHFV NP and not for HAZV NP.

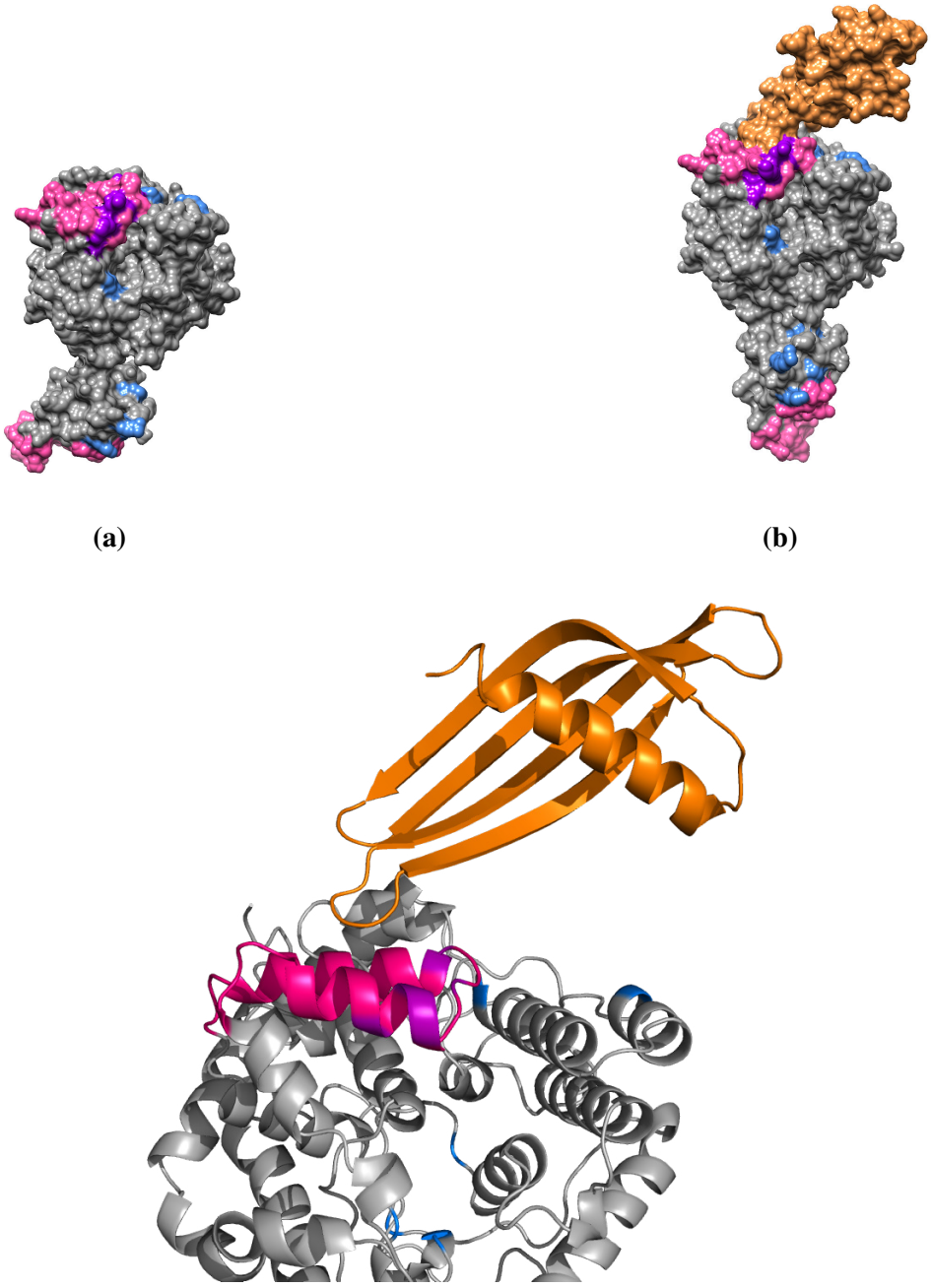


Figure 4.35: Tertiary structure analysis of the Affimer-NP and CCHFV NP complex. (a) Crystal structure of the CCHFV NP (PDB = 4AKL). Residues predicted to be involved in either RNA binding, monomer-monomer interactions, or both, are highlighted in blue, pink and purple respectively. **(b)** Crystal structure of Affimer-NP bound to CCHFV NP. Variable loops of Affimer-NP bind the globular domain of CCHFV NP. Colour code of CCHFV NP is the same as in (a) and Affimer-NP is represented in orange. **(c)** Closer look of the interaction between Affimer-NP and the CCHFV NP globular domain. Colour code is the same as in (b). Images created using PyMOL.

Hydrogen bonds		
##	Affimer-NP	CCHFV NP
1	TYR 103 [HH]	GLN 329 [OE1]
2	ASP 106 [H]	GLU 333 [OE1]
3	PHE 107 [H]	GLU 333 [OE2]
4	TRP 108 [H]	GLU 333 [OE2]
5	ASP 110 [OD2]	LYS 336 [HZ2]

(a)

Salt bridges		
##	Affimer-NP	CCHFV NP
1	ASP 110 [OD1]	LYS 336 [NZ]
2	ASP 110 [OD1]	ARG 339 [NH2]
3	ASP 110 [OD2]	LYS 336 [NZ]
4	ASP 110 [OD2]	ARG 339 [NH2]

(b)

Figure 4.36: Hydrogen bond and salt bridge interactions between Affimer-NP and CCHFV NP. (a-b) Interface residues between Affimer-NP loops and CCHFV NP were quantitated using Chimera and PDBePISA. Hydrogen bond **(a)** and salt bridge **(b)** interactions between Affimer-NP and CCHFV NP are shown.

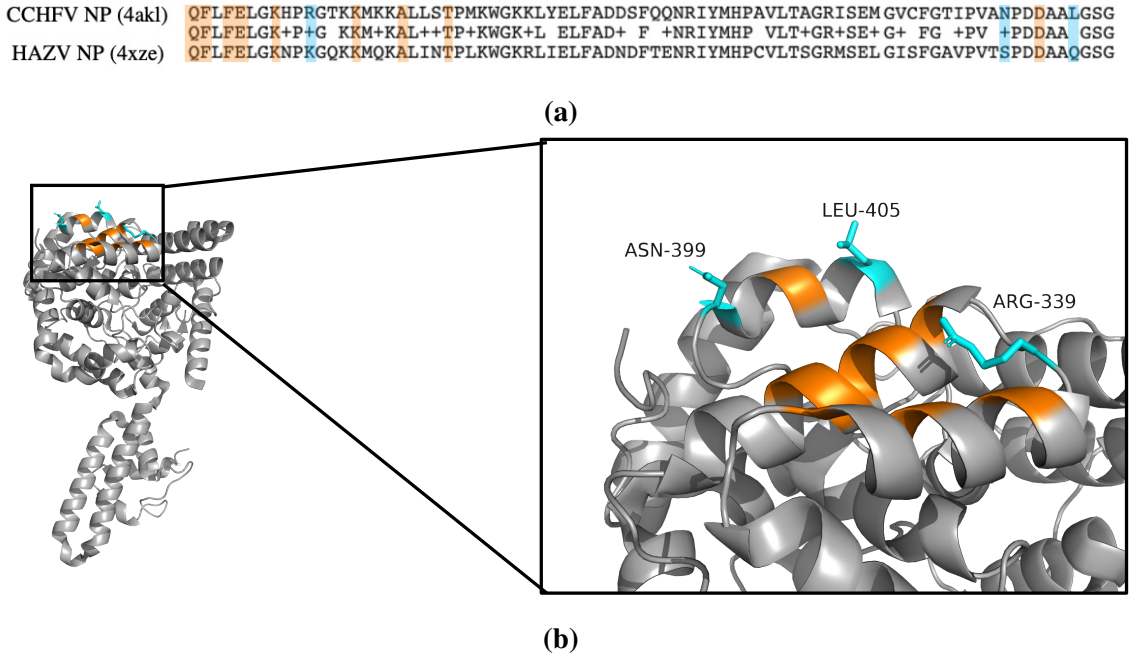


Figure 4.37: Alignment between CCHFV NP residues involved in Affimer-NP interaction and HAZV NP. (a) Alignment of a partial sequence of CCHFV NP (residues 329-408, PDB:4akl) and HAZV NP (PDB:4xze). Residues involved in the Affimer-NP variable loop contacts are highlighted in orange if they are conserved in the two proteins or in cyan if they are different. **(b)** Close look of the residues from the globular domain of CCHFV NP involved in the interaction with Affimer-NP loops. Colour code is the same as in (d). Residues not conserved between CCHFV NP and HAZV (cyan) are labelled. Images created using PyMOL.

4.8 Chapter summary and discussion

In this chapter, the characterisation of the interaction between CCHFV NP and Affimer-NP using different biochemical and physicochemical techniques was described. Affimers for more than 100 targets have been previously described in the literature with high specificity and low nanomolar affinities [181]. Likewise, Affimer-NP presents a high specificity for CCHFV NP, showing no cross-reactivity with its close homologue HAZV NP, and exhibiting high affinity with a K_D in the nanomolar range, as observed using AP and SPR analyses. Appropriate folding of Affimer-NP and CCHFV NP molecules was confirmed by CD studies, which revealed Affimer-NP imparted a stabilising interaction on CCHFV NP when in complex. An increase in the melting temperature of a complex in comparison with the unmixed components is a frequently observed phenomenon in CD analyses [237].

CCHFV NP is predicted to bind RNA in a sequence-independent manner, but the number of nucleotides each NP can accommodate is not yet known. FA analysis of the RNA binding of recombinant CCHFV NP revealed high affinity binding to 27 and 48 mer RNA oligonucleotides, whereas low affinity was observed for 8 mer oligonucleotides. This observation suggests that monomeric CCHFV NP preferably binds RNA molecules of more than 8 oligoribonucleotides in length. A specific interference of Affimer-NP in the RNA binding function of CCHFV NP and CCHFV replication was demonstrated using FA binding assays and a newly developed CCHFV mini-genome system, respectively, suggesting a potential antiviral effect of this small protein. The antiviral drug ribavirin is the only currently available therapeutic agent for the treatment of CCHFV infections, but its mechanism of action remains to be elucidated and case–control studies have not been conducted to date [35]. The use of Affimers for therapeutic purposes has not been yet fully investigated, but their small size, rapid tissue penetration and high stability and solubility make them an attractive candidate for therapeutic applications [181, 210].

A better understanding of the CCHFV NP/Affimer-NP interaction was achieved by the elucidation of the crystal structure of the complex. The variable loops of Affimer-NP bind to the globular domain of CCHFV NP. These interaction sites are around the surface that has been previously predicted to participate in both NP-RNA interactions and the oligomerisation of CCHFV NP, suggesting a possible interference of Affimer-NP in these two fundamental NP functions. The specificity of Affimer-NP for CCHFV NP and the lack cross-reactivity with HAZV NP can also be explained by the crystal structure of the complex. Affimers have shown high selectivity in previous studies [250, 251], being able to differentiate even between analogues of small molecules [201]. From the predominant

12 residues involved in the interaction contacts between Affimer-NP and CCHFV NP, only 3 are not conserved in HAZV NP. These 3 amino acid changes seem to be enough to prevent the recognition of HAZV NP by Affimer-NP.

In summary, the results presented in this chapter meet objective 2 of this project (see project aims in section 1.4). It can be concluded that the interaction between CCHFV NP and Affimer-NP is characterised by a high affinity and specificity, and the functional effects of this interaction include the interference of Affimer-NP in the RNA binding function of CCHFV NP and CCHFV-specific replication.

Chapter 5

Results: Development of CCHFV NP diagnostic assays

5.1 Chapter introduction

Timely detection of viral infections is crucial, not only for the early diagnosis and adequate treatment of individual patients, but also for the rapid response to outbreaks before they can develop and proliferate. In the particular case of CCHF disease, fatal outcome is associated with higher viral loads [44, 45, 46] and late diagnosis of patients decreases treatment efficacy and increases the risk of fatal outcome and nosocomial spread [68], making quick and accurate detection of antigen important for disease management. Although presence of viral RNA can be detected in human samples using molecular methods such as RT-PCR, rapid detection of CCHFV antigen in remote areas or in low-resource settings is an urgent and unmet need. For these reasons, the development of a commercially available, bio-safe, rapid point-of-care test is considered as urgent need for research [35].

ELISAs and lateral flow assays (LFAs) are two types of colorimetric immunoassays used to detect the presence of antigens in liquid samples. While performance of PCR-based diagnostic assays requires specialised lab equipment and training, ELISA and LFAs are simple to carry-out and only require basic equipment, being more suitable for low-resource settings. The high affinity, specificity and binding kinetics of Affimer-NP for CCHFV NP suggested it could represent a successful candidate as recognition molecule for the development of *in vitro* diagnostic tests. The experiments described in this chapter explored the potential application of Affimer-NP as recognition molecule in ELISA and LFAs, in combination with CCHFV NP-specific polyclonal antibodies, for the detection

of CCHFV NP in spiked human and animal sera.

5.2 Development of a sandwich ELISA for the detection of CCHFV NP

Introduction to sandwich ELISAs. Sandwich ELISA is a commonly used format of ELISA testing for the detection of antigen in heterogeneous samples. This system relies on the binding of two different recognition molecules specific for the target antigen, which recognise different epitopes. Sandwich ELISAs are commonly performed in 96-well plates, in which each well surface is coated with a purified capture molecule specific for the antigen of interest. After blocking the plate surface to avoid non-specific adsorption of other proteins, the antigen in the test sample is allowed to react with the immobilised capture molecule. An antigen-specific primary antibody is added to the wells and binds to the captured antigen. Enzyme-linked secondary antibodies are then bound to the primary antibody. A specific substrate of the linked enzyme is added and enzymatically converted to a coloured substrate. The presence of the coloured substrate can be quantified by measuring absorbance of each individual well. These assays can be used for quantification of antigen, as the colorimetric output signal is proportional to the amount of antigen present in the test sample, and are characterised by a high specificity, due to the involvement of two recognition molecules in the detection of the target antigen.

Results. Affimer-NP was tested as capture molecule for development of a sandwich ELISA in combination with CCHFV NP-specific polyclonal IgGs, used as recognition molecule. A secondary anti-rabbit IgG antibody conjugated to horseradish peroxidase (HRP) was used to provide a colorimetric output (figure 5.1). Primary antibody concentration was optimised to obtain the best signal-to-background ratios. Briefly, 96-well plates were coated with 1 µg of Affimer-NP and incubated with different concentrations of CCHFV NP or HAZV NP, the latter used as negative control, diluted in diluent buffer. Two different dilutions (1/5000 and 1/10000) of CCHFV NP-specific polyclonal IgGs in diluent buffer were tested (figure 5.2). The 1/5000 dilution demonstrated a better signal-to-background ratio, and thus was selected for following experiments.

To further characterise the ability of the Affimer-NP-based ELISA test to detect CCHFV NP in sera samples, animal and human sera were spiked with CCHFV NP. Other bunyaviral NPs, namely HAZV, Rift Valley fever virus (RVFV) and Schmallenberg virus (SBV), were also tested as negative controls. Briefly, 96-well plates were coated with 1 µg

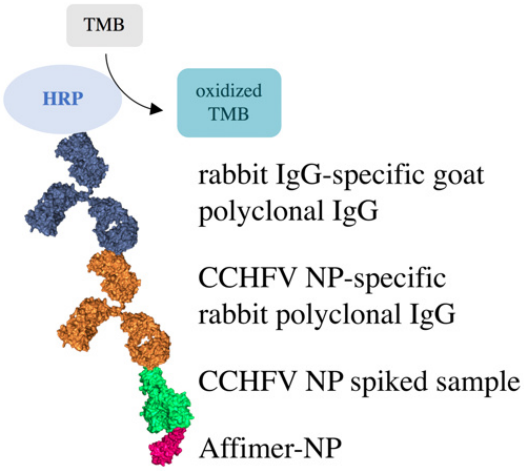


Figure 5.1: Schematic representation of Affimer-NP-based sandwich ELISA for the recognition of CCHFV NP. Affimer-NP was used as capture molecule for CCHFV NP in spiked sera. CCHFV NP-specific IgGs were used as detection molecule. Anti-rabbit IgG goat polyclonal IgGs conjugated to HRP were used to develop the colorimetric output using TMB as substrate.

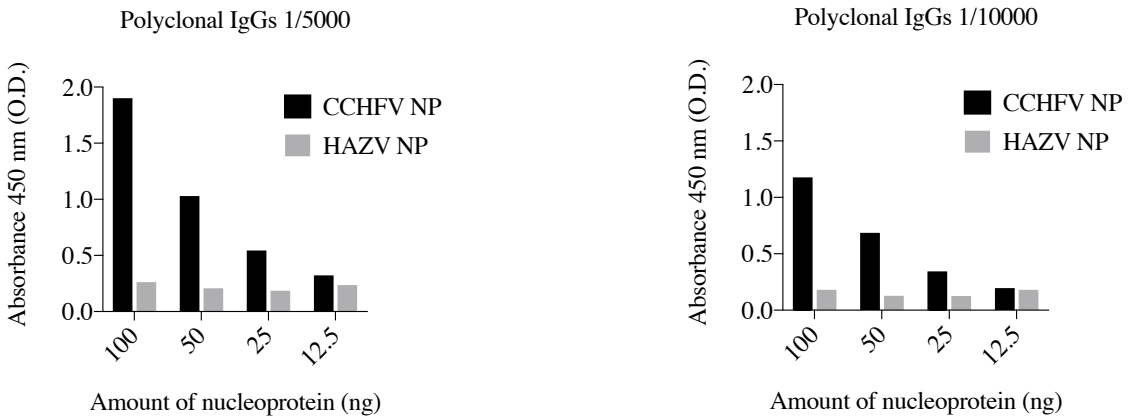


Figure 5.2: Optimisation of CCHFV NP-specific polyclonal IgGs dilution used for the recognition of CCHFV NP in sandwich ELISA. (a-b) Micro-well strips were coated with Affimer-NP and incubated with different concentrations of CCHFV NP or HAZV NP. Two different dilutions, 1/5000 (a) and 1/10000 (b), of CCHFV NP-specific polyclonal IgGs in diluent buffer were tested.

of Affimer-NP and then incubated with serial dilutions of CCHFV, HAZV, RVFV or SBV NPs (figure 5.3a). CCHFV NP-specific IgGs were used as primary recognition molecule (1/5000 in diluent buffer), and anti-rabbit IgG antibodies conjugated to HRP were used to provide a colorimetric output with 3,3',5,5'-tetramethylbenzidine (TMB) substrate (figure 5.3b).

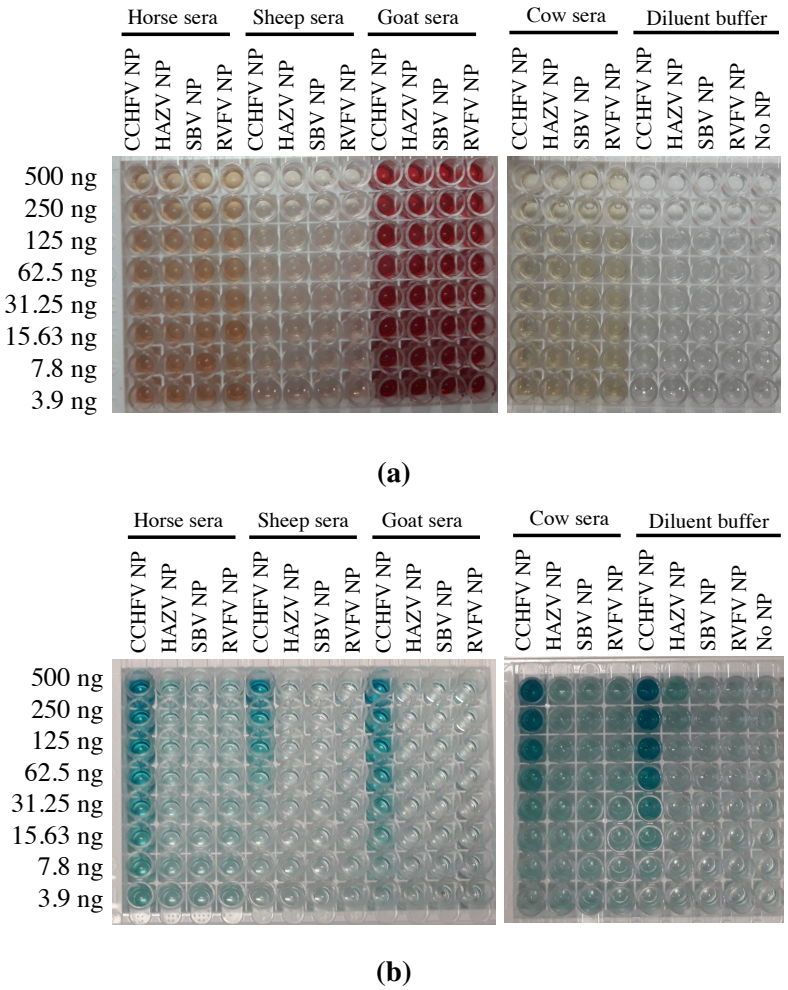


Figure 5.3: Representative pictures of different steps from the Affimer-NP-based sandwich ELISA for the recognition of CCHFV NP. (a) Picture of micro-well strips coated with Affimer-NP and incubated with animal sera spiked with different concentrations of bunyaviral NPs (ng of recombinant protein/well are indicated on the left side). **(b)** Picture of micro-well strips in the final step of the sandwich ELISA after development of a colorimetric output by addition of TMB substrate.

The Affimer-NP-based sandwich ELISA showed the ability to detect CCHFV NP in spiked sera from horse, goat, sheep, cow and human sources. The limit of detection (LOD) varied for the different sera samples, ranging from 31.25 ng/well (sheep sera) to less than 3.9 ng/well (horse sera). No cross-reactivity was shown when testing the negative control NPs of HAZV, RVFV and SBV diluted in sera. Some cross-reactivity

was observed for high concentrations of HAZV NP in diluent buffer (500 ng/100 µL) (figure 5.4).

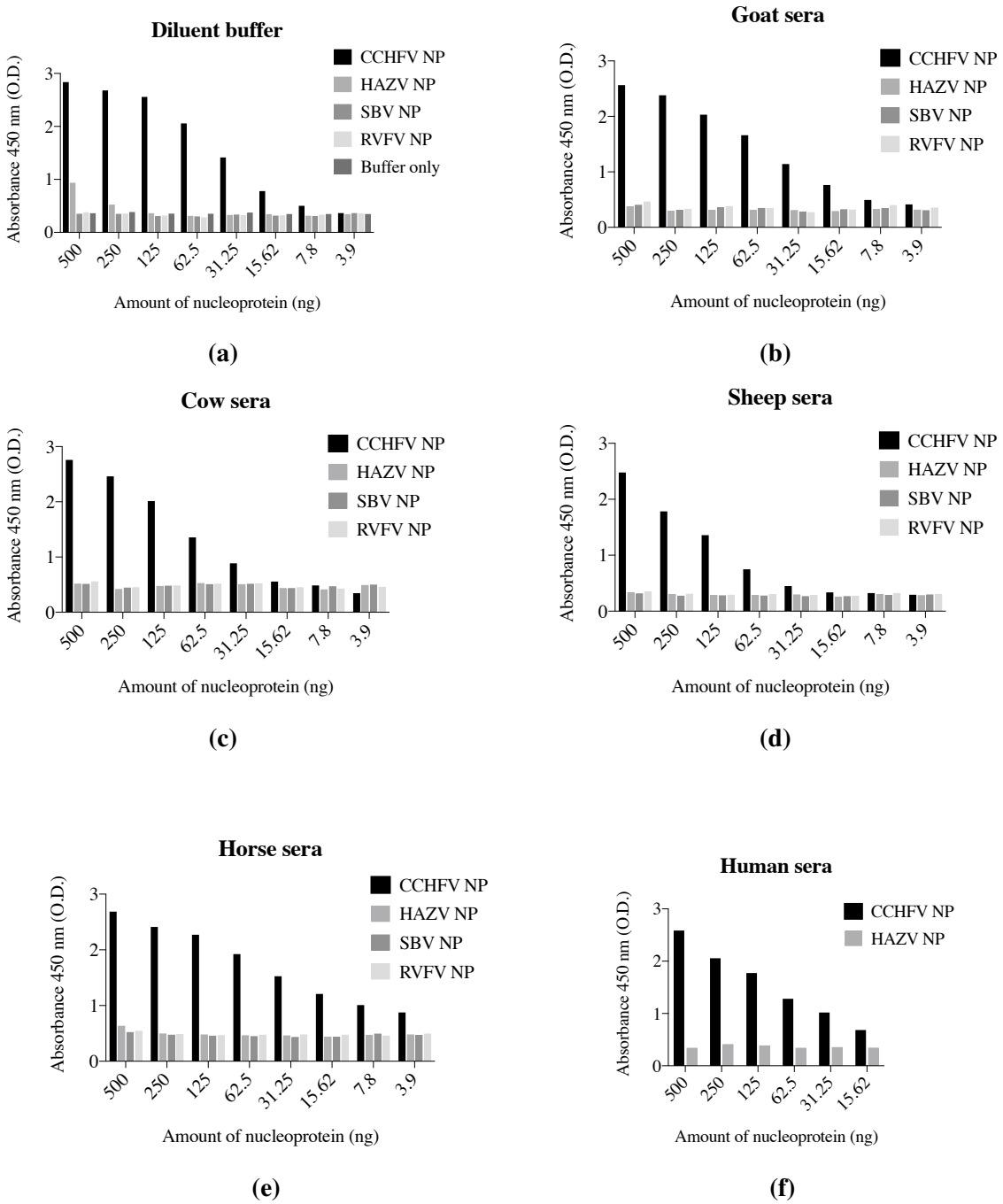


Figure 5.4: Affimer-NP-based sandwich ELISA results. (a-d) ELISA results testing diluent buffer (a) or sera from animals (b-e) or human (f) spiked with different concentrations of CCHFV, HAZV, SBV or RVFV NPs (ng of recombinant protein/well are indicated). The optical density (OD) of each well was quantified by measurement of absorbance at 450 nm. Results correspond to a single experiment representative of three independent replicates.

5.3 Development of LFAs for the detection of CCHFV NP

Introduction to LFAs. LFAs are the most commonly used and user-friendly format of point-of-care diagnostic tests. LFAs are based on the chromatographic separation of molecular complexes conjugated to coloured beads to analyse the presence of a specific molecule within a sample. The colorimetric output of LFAs can be detected within minutes of sample application and is easy to interpret by non-specialised users. LFA performance is based on the deposition of a liquid sample containing the analyte of interest on the device, which migrates by capillary action through various zones of polymeric strips on which recognition molecules specific for the analyte are deposited.

A typical lateral flow test strip for direct detection of a specific antigen consists of overlapping membranes that are mounted on a backing card for better stability and handling. The sample is applied at one end of the strip, on the conjugate release pad, which contains molecules specific to the target analyte conjugated to coloured particles, most commonly colloidal gold or latex microspheres. The sample, together with the conjugated recognition molecule bound to the target analyte, migrates along the strip into the detection zone. This is a porous membrane, usually composed of nitrocellulose, with specific biological components immobilised in lines. Recognition of the sample analyte results in the association of the analyte to the test line and consequent accumulation of coloured particles in the test line. The colorimetric output on the control line indicates the proper liquid flow through the strip. Read-out can be assessed by eye or using a dedicated reader. The liquid flow across the device is enhanced by the capillary force of the strip material and an absorbent pad placed at the end of the strip, which absorbs excess reagents and prevents back-flow of the liquid.

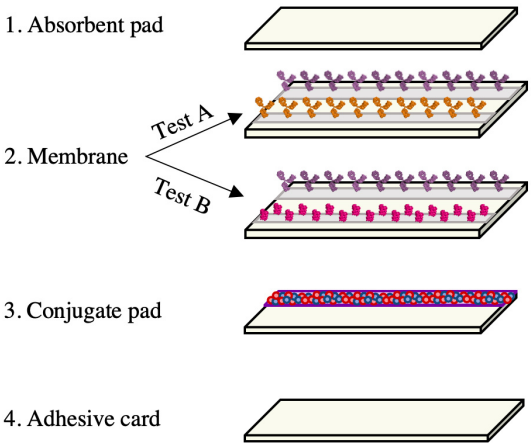
Some of the main advantages of LFAs compared to ELISA tests include rapidity and one step analysis, low operational cost, simple instrumentation, user-friendly format, long term stability under different environmental conditions and portability. On the downside, LFAs commonly present a lower sensitivity due to the lack of signal amplification steps.

Results. The suitability of using Affimer-NP for the development of an LFA combined with CCHFV NP-specific polyclonal IgGs was tested. Two different possibilities were tested: the use of Affimer-NP as capture molecule in the conjugate pad and its use as detection molecule in the test line (figure 5.5).

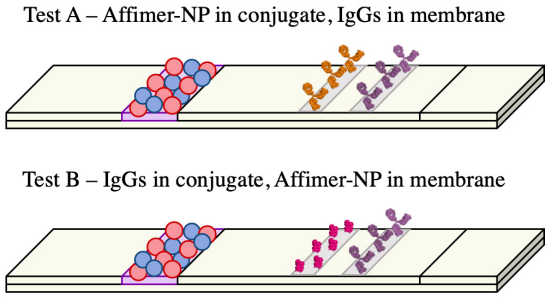
Briefly, red latex beads were coated with CCHFV NP-specific IgGs or Affimer-NP and were mixed with control blue latex beads, coated with biotin-BSA. For the test line, either Affimer-NP or CCHFV NP-specific IgGs were deposited in the LFA membrane,

Conjugated latex beads	Membrane	Pair used for
Affimer-NP coated red latex beads	Anti-CCHFV NP rabbit polyclonal IgG	Test line - test A
CCHFV NP-specific rabbit polyclonal IgG coated red latex beads	Affimer-NP	Test line - test B
BSA-biotin coated blue latex beads	Anti-biotin mouse IgG	Control line - tests A and B

(a)



(b)



(c)

Figure 5.5: Design of CCHFV NP-specific LFA. (a) Two different possibilities were tested: the use of Affimer-NP as capture molecule in the conjugate pad, and its use as detection molecule in the detection zone. Red latex beads were conjugated with Affimer-NP (test A) or CCHFV NP-specific polyclonal IgGs (test B). Blue latex beads were conjugated with BSA biotin. The test line in the nitrocellulose membrane contained CCHFV NP-specific polyclonal IgGs (test A) or Affimer-NP (test B). Control line contained biotin-specific mouse IgGs. **(b)** Schematic representation of the different components used for the fabrication of the CCHFV NP-specific LFAs: absorbent pad, nitrocellulose membranes, conjugate pads and adhesive cards. **(c)** Schematic representation of fully assembled LFA tests A and B.

and anti-biotin IgGs were deposited in the control line. The test containing Affimer-NP in the conjugate pad and CCHFV NP-specific IgGs in the test line will be referred to as test A (Affimer-NP in conjugate pad), and the test containing CCHFV NP-specific IgGs in the conjugate pad and Affimer-NP in the test line will be referred to as test B (Affimer-NP in test line, figure 5.5). In both cases, when a CCHFV NP spiked sample is applied, the NP binds to the recognition molecule on the red beads and subsequently binds to the recognition molecule in the test line resulting in a red coloured line. When a negative sample that does not contain CCHFV NP is applied, red beads migrate until the end of the strip, resulting in an uncoloured test line. In all cases, proper liquid flow through the strip results in a blue coloured control line (figure 5.6).

Once the conjugate and the test and control lines were dispensed, membrane, conjugate pad and absorbent pad were pasted onto a backing card and were covered with a transparent adhesive protector film. Assembled master cards were then cut to 4.2 mm width strips (figure 5.7).

5.3.0.1 Optimisation of test line composition and LFA sample buffer

Different concentrations of Affimer-NP or CCHFV NP-specific polyclonal IgGs were used for the preparation of the LFA membrane. LFA tests were assembled and test line signal was analysed by eye (data not shown). The highest signal for both test lines (Affimer-NP and CCHFV NP-specific IgGs) was obtained when using a concentration of 1 mg of protein/mL. This concentration was used for further production of the LFAs.

Three different running buffers (namely TDM-36, TDM-42 and TDM-45) were tested for sample application. Composition of the buffers is detailed in table 5.1. Briefly, TDM-36 contained high-salt concentration (1 M NaCl) and BSA as blocking agent, whereas TDM-42 and TDM-45 contained casein for the same purpose. The only difference between TDM-42 and TDM-45 was that the latter contained an extra 0.5% (v/v) of detergent (Tween-20). Buffers were tested and signal was analysed by eye (data not shown). TDM-42 and TDM-45 buffers showed best performance as diluent buffer for test A (Affimer-NP in conjugate pad) and test B (Affimer-NP in test line), respectively. These buffers were consequently selected as optimal and used for following experiments.

5.3.0.2 Optimisation of LFA conjugate composition

Red latex beads were coated with CCHFV NP-specific IgGs or Affimer-NP and were mixed with control blue latex beads coated with biotin-BSA. The size of latex beads was

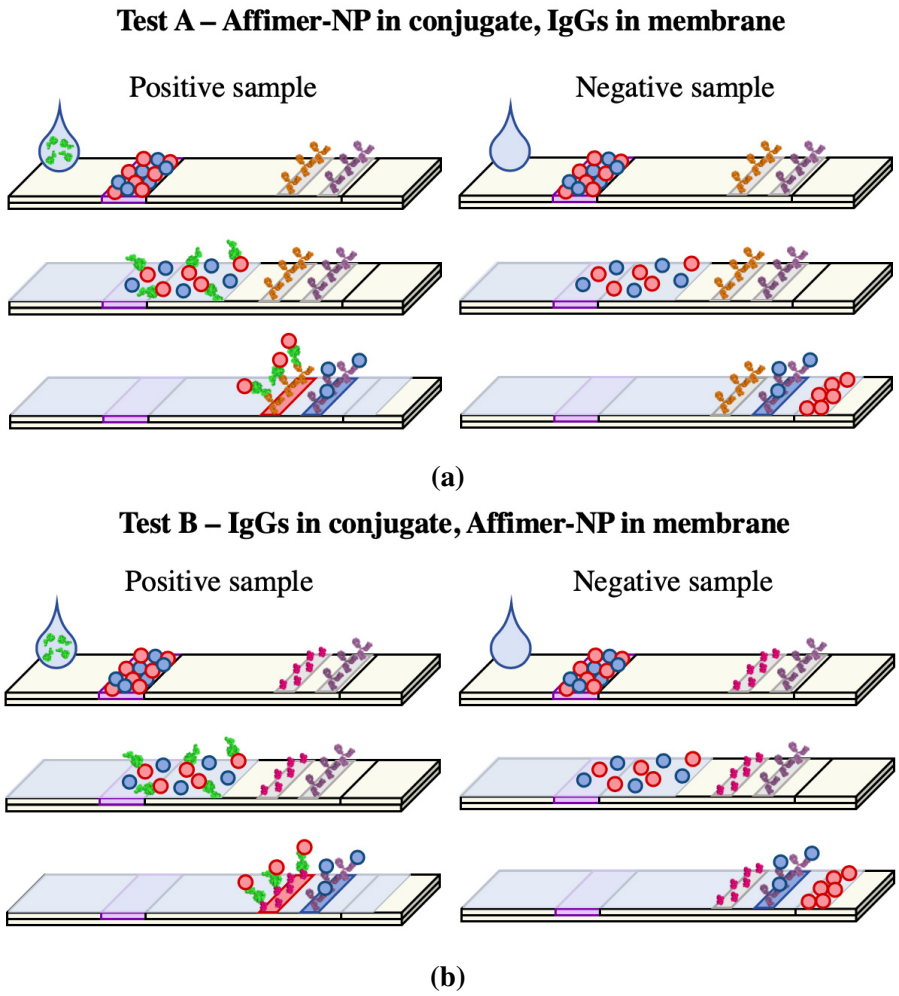


Figure 5.6: Mode of operation of CCHFV NP-specific LFA tests A and B. (a-b) The test containing Affimer-NP in the conjugate and CCHFV NP-specific IgGs in the test line will be referred to as Test A (Affimer-NP in conjugate pad, (a)), and the test containing CCHFV NP-specific IgGs in the conjugate and Affimer-NP in the test line as Test B (Affimer-NP in test line, (b)). In both cases, when a CCHFV NP spiked sample is applied, migration of beads results in a red coloured test line, whereas negative samples do not result in a coloured test line. In all cases, proper liquid flow through the strip results in a blue coloured control line.

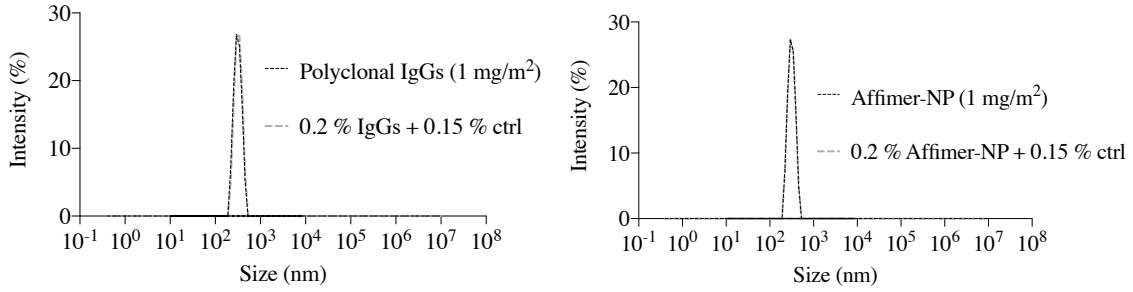
LFA running buffer	Composition
TDM-36	250 mM Tris-HCl pH 7.2, 1 M NaCl, 1% (v/v) Tween-20, 1% (w/v) BSA, 0.095% (w/v) azide
TDM-42	250 mM Tris-HCl pH 7.5, 150 mM NaCl, 1% (v/v) Tween-20, 1% (w/v) casein, 0.095% (w/v) azide
TDM-45	250 mM Tris-HCl pH 7.5, 150 mM NaCl, 1.5% (v/v) Tween-20, 1% (w/v) casein, 0.095% (w/v) azide

Table 5.1: LFA running buffers tested. Chemical composition of the different running buffers tested for the CCHFV NP-specific LFAs.

(a) **(b)**

Figure 5.7: Manufacturing process of CCHFV NP-specific LFA tests. (a) Picture showing the different components used for the fabrication of the CCHFV NP-specific LFAs: 1. absorbent pad, 2. nitrocellulose membranes, 3. conjugate pads and 4. adhesive cards. **(b)** Pictures of a fully assembled LFA cardboard (top left), strips cutting process (right) and ready-to-use LFA strips (bottom left).

monitored by dynamic light scattering (DLS) to avoid the use of aggregated particles. DLS determines the size distribution profile of small particles in suspension by measuring the different intensities of scattered light after passing through the sample. The Brownian motion of particles in suspension is dependent on particle size, and thus, scattered light intensity fluctuations caused by this motion can be measured and related to particle size. DLS analyses were performed before and after functionalisation of the latex beads, and after mixing red test beads with blue control beads. Particle size in the 300 nm range was considered as adequate. If the DLS analysis showed aggregated particles, sonication of the sample was done until particles were dispersed to an appropriate size (figure 5.8).



(a) **(b)**

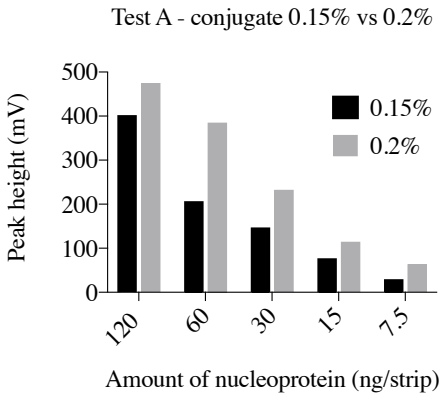
Figure 5.8: Dynamic light scattering analysis of functionalised latex beads. (a-b) Size distribution of beads functionalised with CCHFV NP-specific IgGs (a) or Affimer-NP (b) and mixes of these functionalised latex beads (0.2% (v/v)) and control biotin-BSA beads (ctrl, 0.15% (v/v)). Results correspond to a single representative experiment.

Different concentrations of Affimer-NP and CCHFV NP-specific polyclonal IgGs, and different mixed amounts of red and blue beads were tested using serial dilutions of CCHFV NP in diluent buffer. For quantitative results, ESEQuant Lateral Flow reader (Qiagen) was used. The LFA reader scans a light dot across the LFA strip and measures the intensity of the reflected light using an optical detection system that converts the light into an electrical current. The intensity of the reflected light can be represented in 2D profile plots as voltage current (mV) versus distance (mm). Light absorbance by coloured beads in the test and control lines decreases the intensity of the reflected light, resulting in reduced voltage peaks as the coloured lines in the LFA strip are scanned. The absolute value of the intensity peak caused by the presence of captured coloured beads in the test line is directly proportional to the amount of analyte present in the sample. Thus, peak height (mV) can be used to quantify the presence of analyte and was used to classify test results as positive, low positive, doubtful or negative by the Lateral Flow Studio Software.

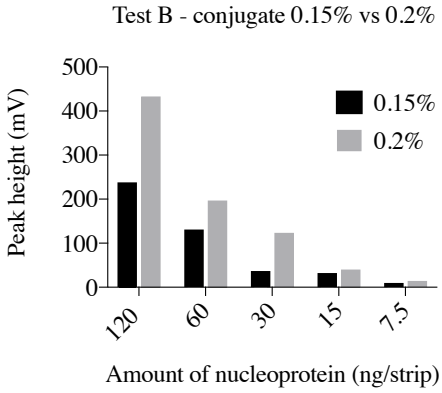
Figure 5.9a represents an optimisation experiment of the amount of Affimer-NP functionalised red beads mixed with biotin-BSA functionalised red beads. Two different percentages of Affimer-NP red beads were tested (0.15 and 0.2% (v/v)) mixed with 0.2% (v/v) of control biotin-BSA beads. The LOD obtained with the mix containing 0.2% (v/v) Affimer-NP red beads (LOD: 7.5 ng CCHFV NP/strip) was lower than with the mix containing 0.15% (v/v) of Affimer-NP red beads (LOD: 15 ng/strip) (figure 5.9a). A similar optimisation procedure was performed for test B: two different concentrations of CCHFV NP-specific IgGs functionalised beads were tested (0.15 and 0.2% (v/v)). The LOD obtained with the mix containing 0.2% (v/v) of CCHFV NP-specific IgG functionalised beads (LOD: 30 ng/strip) was lower than the one obtained for the mix containing 0.15% (v/v) (60 ng/strip) (figure 5.9b). Consequently, a concentration of 0.2% (v/v) of red latex beads was selected as optimal for both tests. Overall, test A (Affimer-NP in conjugate pad) presented more sensitivity than test B (Affimer-NP in test line) for CCHFV NP diluted in diluent buffer (figure 5.9).

5.3.0.3 Optimisation of LFA sample application

Due to its better sensitivity for CCHFV NP in diluent buffer, test A was selected for optimisation of sample application. Three different methods of sample application were tested using sera spiked with CCHFV NP: (1) sequential application of 20 µL of CCHFV NP spiked sera and 100 µL of buffer, (2) application of 120 µL premixed CCHFV NP spiked sera with buffer, and (3) sequential application of 10 µL of CCHFV NP spiked sera and 110 µL of buffer (figure 5.10). Test results were classified by the Lateral Flow



(a)



(b)

Test A 0.15%			Test A 0.2%		
120 ng/strip	402.1	POSITIVE	120 ng/strip	475	POSITIVE
60 ng/strip	206.8	POSITIVE	60 ng/strip	385.2	POSITIVE
30 ng/strip	147.3	POSITIVE	30 ng/strip	232.5	POSITIVE
15 ng/strip	77.3	LOW POSITIVE	15 ng/strip	114.8	POSITIVE
7.5 ng/strip	29.8	NEGATIVE	7.5 ng/strip	63.8	LOW POSITIVE

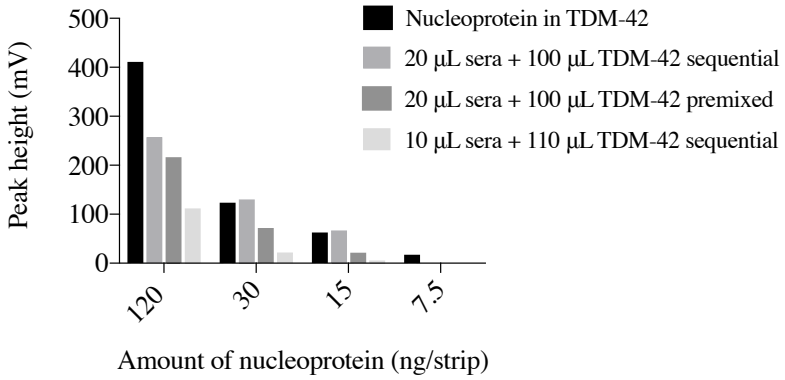
(c)

Test B 0.15%			Test B 0.2%		
120 ng/strip	238.1	POSITIVE	120 ng/strip	433	POSITIVE
60 ng/strip	131.2	POSITIVE	60 ng/strip	197	POSITIVE
30 ng/strip	36.7	DOUBTFUL	30 ng/strip	123.7	POSITIVE
15 ng/strip	32.1	NEGATIVE	15 ng/strip	40	DOUBTFUL
7.5 ng/strip	9.6	NEGATIVE	7.5 ng/strip	14.3	NEGATIVE

(d)

Figure 5.9: Optimisation of LFA conjugate composition. **(a-b)** Optimisation of conjugate composition for test A (a) and B (b). Two different concentrations of test conjugate (0.15 or 0.2% (v/v)) were tested. **(c-d)** Quantification of results in (a) and (b), and corresponding test results classified by the Lateral Flow Studio Software as positive, low positive, doubtful or negative.

Studio Software as positive, low positive, doubtful or negative according to the strength of the signal in the test line. The best sensitivity was obtained with sequential application of 20 µL of CCHFV NP spiked sera and 100 µL of buffer. This was consequently selected as optimal sample application method and was used for further characterisation of the LFA tests.



(a)

	120 µl Nucleoprotein in TDM-42		20 µl sera + 100µl TDM-42 sequential		20 µl sera + 100 µl TDM-42 premixed		10 µl sera + 110 µl TDM-42 sequential	
120 ng/strip	411.1	POSITIVE	257.4	POSITIVE	216.6	POSITIVE	111.8	POSITIVE
60 ng/strip	225.3	POSITIVE	52.5	DOUBTFUL	144.9	POSITIVE	77.5	LOW POSITIVE
30 ng/strip	123.7	POSITIVE	130.1	POSITIVE	71.8	LOW POSITIVE	21.7	NEGATIVE
15 ng/strip	62.8	LOW POSITIVE	66.9	LOW POSITIVE	21.3	NEGATIVE	5.6	NEGATIVE
7.5 ng/strip	17.4	NEGATIVE	0	NEGATIVE	0	NEGATIVE	0	NEGATIVE

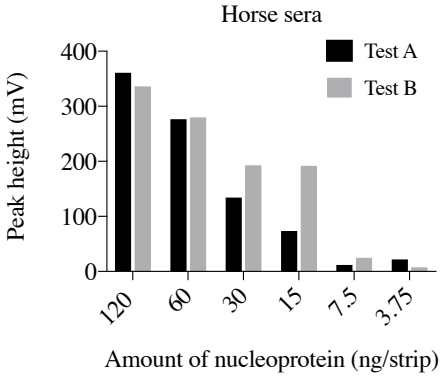
(b)

Figure 5.10: Optimisation of LFA sample application. **(a)** Three different ways of sample application were tested using sera spiked with CCHFV NP: (1) sequential application of 20 µL of CCHFV NP spiked sera and 100 µL of buffer, (2) application of 120 µL premixed CCHFV NP spiked sera with buffer, (3) sequential application of 10 µL of CCHFV NP spiked sera and 110 µL of buffer. **(b)** Quantification of results in (a) and correspondent test results classified by the Lateral Flow Studio Software as positive, low positive, doubtful or negative.

5.3.1 Specificity and sensitivity of the LFAs

Sensitivity of the tests was firstly analysed using CCHFV NP spiked horse sera (figure 5.11). Briefly, CCHFV NP was serially diluted in horse sera and 20 µL of sample were applied to the LFA strips followed by application of 100 µL of running buffer. Buffer TDM-42 was used for Test A and buffer TDM-45 was used for Test B. Test results were classified by the Lateral Flow Studio Software as positive, low positive, doubtful

or negative according to the strength of the signal in the test line. The LOD for tests A and B was 15 ng of CCHFV NP/strip. In the case of test A the signal obtained for this concentration of CCHFV NP was lower (low positive: 73.5 mV) than for test B (positive: 191.7 mV) (figure 5.11). These results suggest that test B performs better than test A for the detection of CCHFV NP in spiked horse sera.



(a)

	Test A			Test B	
120 ng/strip	360.7	POSITIVE	120 ng/strip	336.3	POSITIVE
60 ng/strip	276.5	POSITIVE	60 ng/strip	279.9	POSITIVE
30 ng/strip	134.2	POSITIVE	30 ng/strip	192.9	POSITIVE
15 ng/strip	73.5	LOW POSITIVE	15 ng/strip	191.7	POSITIVE
7.5 ng/strip	11.8	NEGATIVE	7.5 ng/strip	24.8	NEGATIVE
3.75 ng/strip	21.7	NEGATIVE	3.75 ng/strip	7.5	NEGATIVE

(b)

Figure 5.11: Limit of detection of LFA tests A and B. (a) Sera spiked with different amounts of CCHFV NP were used to determine the limit of detection of LFA tests A and B. **(b)** Quantification of results in (a) and correspondent test results classified by the Lateral Flow Studio Software as positive, low positive, doubtful or negative.

To determine the specificity of the tests, spiked sera samples were prepared using different concentrations of recombinant NPs from different bunyaviruses. NPs of HAZV, RVFV and SBV were tested as negative controls, diluted in animal or human sera. Samples were applied to the strips (120 ng of NP/strip) followed by the application of 100 µL of running buffer. The strength of the control line was evaluated by eye (figure 5.12a) or quantified using an ESE Quant Lateral Flow reader as previously described (figure 5.12b).

Test A was capable of detecting CCHFV NP in horse, sheep, and goat sera; but failed to detect CCHFV NP in the cow serum tested. Test B was capable of detecting CCHFV NP in all the different spiked sera tested, showing higher peak height values (mV) than test A. No cross-reactivity was shown when testing these sera spiked with

		Test A		Test B	
Horse	CCHFV NP	279.6	POSITIVE	386.7	POSITIVE
	HAZV NP	21.9	NEGATIVE	18.8	NEGATIVE
	SBV NP	29.8	NEGATIVE	16.1	NEGATIVE
Sheep	CCHFV NP	69.3	LOW POSITIVE	96.3	LOW POSITIVE
	HAZV NP	8.8	NEGATIVE	14.8	NEGATIVE
	SBV NP	30.6	NEGATIVE	15.5	NEGATIVE
Goat	CCHFV NP	56.4	LOW POSITIVE	216.9	POSITIVE
	HAZV NP	18.1	NEGATIVE	13.6	NEGATIVE
	SBV NP	2.9	NEGATIVE	2.5	NEGATIVE
Cow	CCHFV NP	16.8	NEGATIVE	95.2	LOW POSITIVE
	HAZV NP	30	NEGATIVE	4.4	NEGATIVE
	SBV NP	16.1	NEGATIVE	13	NEGATIVE

(a)

(b)

Figure 5.12: Performance of LFA tests A and B with spiked horse sera samples. **(a)** LFA positive line intensity measured after applying different animal sera spiked with CCHFV, HAZV or SBV NP (120 ng NP/strip). Table represents quantification of results and correspondent test results classified by the Lateral Flow Studio Software as positive, low positive, doubtful or negative. **(b)** Picture of LFAs (test B) 10 min after application of horse sera samples spiked with CCHFV, HAZV, SBV or RVFV NPs.

negative control NPs from other bunyaviruses. Test B was selected as best performing test for spiked sera samples, and was further used to test 10 different field sera of each animal (horse, sheep, goat cow) and 10 human sera. None of the field sera resulted in false positives, demonstrating a 100% specificity of the LFAs for CCHFV NP under these conditions. An additional experiment was performed testing human sera spiked with the different NPs (120 ng of NP/test). Test B showed the capability of detecting CCHFV NP in spiked human sera, and no false positives were obtained when human sera spiked with other bunyaviral NPs (HAZV, SBV or RVFV NPs) were tested (data not shown).

5.3.2 Comparison of performance between the ELISA and the LFAs

The LOD of both Affimer-NP-based diagnostic assays was compared using horse sera spiked with different concentrations of CCHFV NP (figure 5.13). The Affimer-NP-based sandwich ELISA showed the ability to detect down to 1.5 ng of recombinant CCHFV NP per well (0.015 µg/mL). The lowest amount of CCHFV NP that could be detected as a positive result with the LFA was 15 ng/strip (0.75 µg/mL).

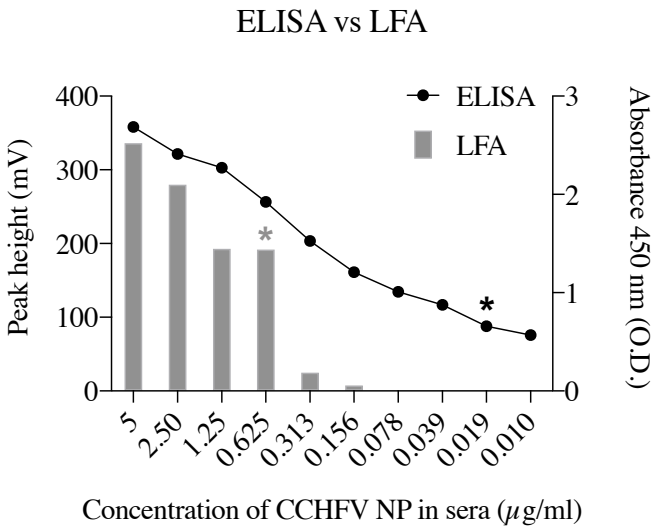


Figure 5.13: Comparison of the ELISA and LFA using spiked horse sera. Comparison of performance between ELISA and LFAs using horse sera spiked with CCHFV NP. Primary axes correspond to LFAs and secondary axes to ELISA. Limits of detection are highlighted with an asterisk.

5.4 Chapter summary and discussion

CCHFV is an important public health issue and an urgent need for accelerated research in its diagnosis has been declared [252], with special emphasis on the necessity to develop rapid diagnostics tests. The high affinity, specificity and stability of Affimer-NP suggested that it was a promising candidate as recognition molecule for the detection of CCHFV NP in diagnostic assays. In this chapter, the performance of Affimer-NP as a recognition molecule in ELISA and LFAs were tested.

This small protein showed an efficient performance as a capture molecule in a sandwich ELISA format, detecting down to 1.5 ng of recombinant CCHFV NP per 100 µL of human and animal sera. A previous study has shown that an ELISA test with a LOD of 2 ng of recombinant CCHFV NP per 100 µL was suitable for the detection of native CCHFV in acute sera of CCHFV-infected patients [253]. According to these observations, the sandwich ELISA shown here using Affimer-NP as capture molecule is theoretically capable of detecting native CCHFV NP in acute patient samples, but this has not been experimentally demonstrated due to limitations on accessibility to patient samples.

Two different prototypes were designed to test Affimer-NP as recognition molecule in LFAs: test A (Affimer-NP in test line) and test B (Affimer-NP in conjugate pad). Test A demonstrated the best sensitivity for CCHFV NP diluted in buffer TDM-42, whereas test B showed the best performance using buffer TDM-45, containing 0.5% (v/v) additional detergent concentration. Overall, the performance of test A was better than test B for samples of CCHFV NP diluted in buffer, and conversely, the sensitivity of test B was better than test A for spiked sera samples. A possible explanation for this observation is that the higher concentration of detergent used for test B may be favourable in the context of sera samples, helping to increase the solubility and migration of this type of matrices, which are characterised for presenting high viscosity and background components that may interfere with test sensitivity.

The Affimer-NP-based LFAs presented in this study require minimal specimen processing, have time to results of minutes and are suitable for field testing and low-infrastructure settings. The high thermal stability of Affimer-NP (previously analysed in section 4.3) allows its immobilisation onto the LFA membrane, which involves a baking step at 45°C, and gives an advantageous stability to the tests, which might remain unused for long periods or may be exposed to different temperatures. In our hands, these tests present 100% specificity for CCHFV NP and their LOD is 15 ng of recombinant CCHFV NP per strip (LOD of test B).

There are no previous studies describing the LOD needed for this kind of test in order to successfully detect CCHFV NP in acute patient samples. No reference antigen preparation standards are available for hemorrhagic fever viral NPs [254] but different point-of-care diagnostic tests have shown the ability to detect them in virus isolates and patient samples. Such is the case for LASV [255], EBOV [256] or RVFV [257]. In order to determine if the sensitivity of this assay is sufficient to detect CCHFV NP in real patient samples, these tests should be experimentally tested and optimised using patient samples under high-level biosafety containment.

Nonetheless, our results serve as proof of concept for the use of Affimers in LFAs and present the first point-of-care diagnostic test for the detection of CCHFV NP described to date. These findings present a possible starting point for the future development of diagnostic assays for the detection of CCHFV NP, contributing to the preparedness for potential future outbreak scenarios.

In summary, the results presented in this chapter meet objective 3 of this project (see project aims in section 1.4). Here, the successful application of Affimer-NP in the development of ELISA and LFA colorimetric tests for the detection of CCHFV NP in spiked sera was described.

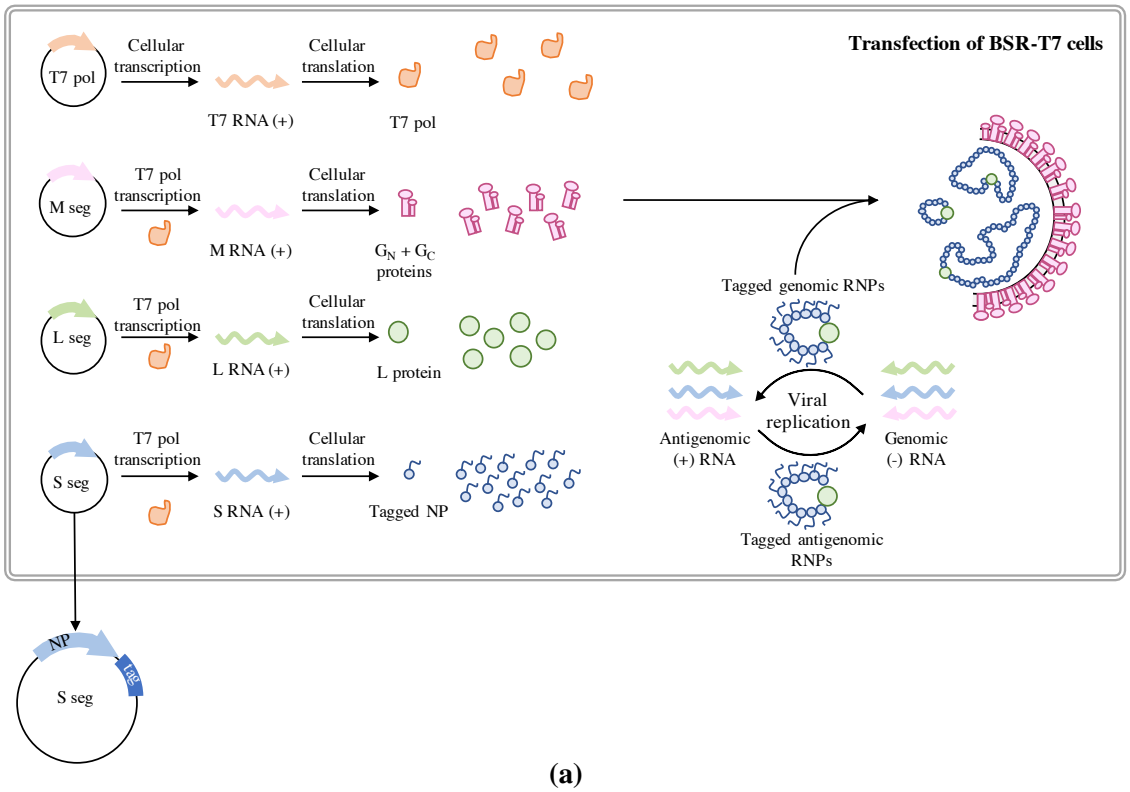
Chapter 6

Results: Structural and localisation studies of nairoviral components using a reverse genetics system

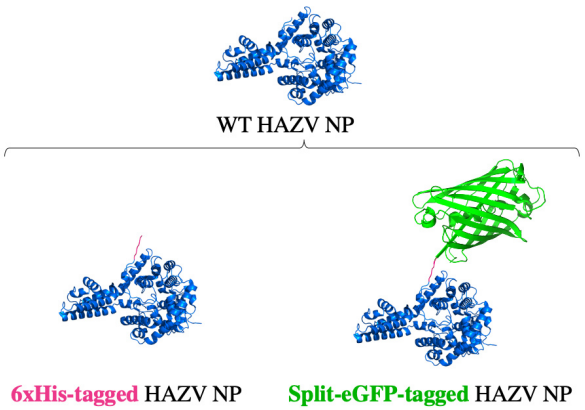
Full length infectious clone systems allow the generation of infectious recombinant viruses and present a valuable tool for the generation of reporter viruses, the performance of mutational analyses, and the elucidation of essential catalytic or structural residues, among other applications. Our laboratory has recently reported the establishment of a reverse genetics system for the efficient generation of infectious HAZV from recombinant sources [156]. The HAZV rescue system is formed by three plasmids in which expression of S, M, and L antigenomic RNAs is driven by T7 pol in hamster-origin BSR-T7 cells. Transfection of these plasmids into cells allows transcription and translation of S, M, and L antigenomes followed by assembly of HAZV RNPs, and subsequent generation of infectious recombinant HAZV (rHAZV). An additional plasmid to express exogenous T7 pol is also included, which increases rescue efficiency.

The HAZV rescue system allows the modification of the different plasmids to obtain genetically modified recombinant viruses. The experiments reported in this chapter describe the insertion of two different clonable tags in the S segment plasmid of the HAZV rescue system, engineered downstream of the ORF of the HAZV NP. Transfection of these plasmids alongside the rest of the components of the rescue system resulted in the recovery of genetically engineered recombinant viruses bearing the clonable tags in their RNA genome and the corresponding protein tag in their NPs (figure 6.1a). The two clonable tags explored in this chapter are a 6xHis tag for the purification and structural studies of HAZV RNPs using transmission electron microscopy (TEM), and a split-

enhanced green fluorescence protein (split-eGFP) tagging system to track HAZV NP dynamics inside infected cells using confocal microscopy (figure 6.1b).



(a)



(b)

Figure 6.1: Using a HAZV reverse genetics system to tag HAZV NP. (a) The HAZV reverse genetics system can be used for the insertion of clonable tags downstream of the ORF of HAZV NP, resulting in genetically engineered recombinant viruses containing the corresponding nucleotide sequence encoded in their RNA genome and the corresponding amino acid sequence incorporated in their NPs. **(b)** The crystal structure of wild type (WT) HAZV NP (PDB: 4akl) and schematic representation of the two different clonable tags incorporated in the C-terminal end of the protein, namely, a 6xHis tag and a split-eGFP-tag. Models created using Chimera and PyMol.

6.1 Development of 6xHis-NP rHAZV

The crystal structure of HAZV NP and NPs of several CCHFV strains have been previously elucidated and present in common a globular domain and a flexible arm domain. Different oligomeric forms of CCHFV NP have also been described, including tetramers, pentamers, hexamers and heptamers observed by EM [149], and higher ordered crystallised structures that could just be the result of crystal packing artefacts, but may also represent a physiologically relevant organisation of CCHFV NP [145]. The interaction between monomers occurs in a head-tail manner and involves residues 320-354 within the head domain of one molecule and residues 210-219 and 260-272 of the arm of the adjacent molecule [145].

However, the structure of native nairoviral RNPs has not been elucidated to date. It is unknown if nairoviral NPs associate as monomers with viral RNA in a 'beads-on-a-string' structure or as a more complex helical structure based on any of the observed oligomeric forms. With the aim of investigating the structure of native nairoviral RNPs, a 6xHis tag was engineered downstream of the HAZV NP ORF in the S segment plasmid of the HAZV rescue system for the purification of native RNPs by affinity chromatography. Examination of the HAZV NP crystal structure [151] suggested that a C-terminal 6xHis tag would be exposed and accessible for affinity chromatography purification (figure 6.1b).

6.1.1 Rescue and characterisation of 6xHis-NP rHAZV

Briefly, a 6xHis tag sequence was incorporated between the last codon of the HAZV NP ORF and its adjacent stop codon within the previously described HAZV S segment expression plasmid (pMK-RQ-S) [156], using site directed mutagenesis (SDM) to generate pMK-RQ-S-6xHis (see plasmid maps in Appendix B). Co-transfection of modified plasmid pMK-RQ-S-6xHis, and wild type (WT) M and L segment expressing-plasmids was performed alongside a transfection comprising all three WT plasmids for comparison. In both cases, a T7 pol expression plasmid was also included. At 96 h post-transfection (hpt), supernatants were harvested and used to re-infect fresh monolayers of SW13 cells for 48 h. Western blot analysis of corresponding lysates revealed abundant HAZV NP expression confirming successful rescue of both WT and 6xHis-NP rHAZV viruses (figure 6.2a). The presence of the 6xHis tag could be detected in 6xHis-NP rHAZV infected lysates by western blot using an anti-His tag antibody. As expected, no HAZV NP was detected in control rescues in which the L segment expressing plasmid

was omitted, and no His tag was detected in any of the WT rHAZV infected lysates.

To confirm the presence of the 6xHis tag in the RNA of engineered viruses, RT-PCR analysis was performed on RNA extracted from WT and 6xHis-NP rHAZV stocks to amplify a cDNA fragment corresponding to the C-terminal portion of the HAZV NP ORF. The cDNAs were analysed on a 1% (w/v) agarose gel (figure 6.2b) and the sequence of the tag was further confirmed by subsequent nucleotide sequencing of the RT-PCR-amplified fragment.

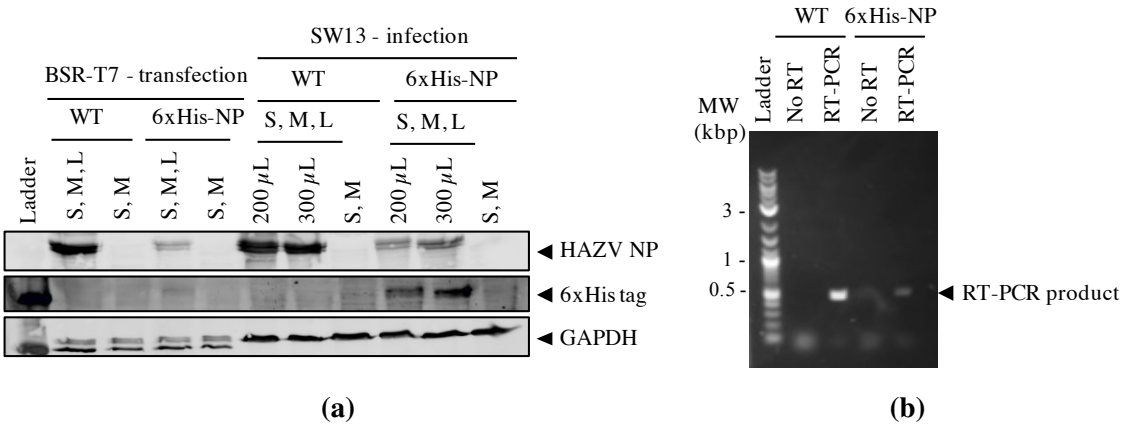


Figure 6.2: Rescue of WT and 6xHis-NP rHAZV infectious viruses. (a) BSR-T7 cells were transfected with the T7 pol support plasmid and the three rescue plasmids (SML) for recovery of wild-type (WT) or 6xHis-NP rHAZV. At 96 hpt, 200 or 300 µL of supernatants were used to infect SW13 monolayers for 48h. Lysates from BSR-T7 and SW13 cells were analysed for expression of HAZV NP by SDS-PAGE and western blot analysis using HAZV NP antisera. Lysates were also collected from non-transfected cells (MOCK) and recovery experiments in which the L segment expression plasmid was omitted (S, M). Presence of the His tag was analysed using an anti-His tag antibody, and glyceraldehyde-3-phosphate dehydrogenase (GAPDH) abundance was included as a loading control. **(b)** UV transillumination imaging of agarose gel electrophoresis analysis of the RT-PCR products from WT or 6xHis NP rHAZV infected cell supernatants. RNA was extracted and treated with DNase, and RT-PCR was performed. For each sample, a negative control was performed omitting the reverse transcriptase (No RT). DNA was stained using SYBR safe DNA stain. Arrow indicates migration distances of the corresponding protein or DNA products.

Next, the growth properties of WT and 6xHis-NP rHAZV were compared to determine the impact of the addition of the 6xHis tag on overall virus fitness. Stocks of both viruses were titrated by plaque assay and used to infect fresh SW13 cells at an MOI of 0.01. Cell lysates and supernatants were collected at 24 h intervals across a 5 day period. Western blot analysis of corresponding lysates revealed abundant HAZV NP expression in both WT and 6xHis-NP rHAZV-infected cells from day 2 post-infection (figure 6.3a). The presence of the 6xHis tag could be detected in 6xHis-NP rHAZV

infected lysates by western blot using an anti-His tag antibody. As expected, no HAZV NP was detected in control non-infected cell lysates, and no His tag was detected in any of the WT rHAZV infected lysates.

The titres of released viruses were further assessed by plaque assays. This analysis revealed that titres for WT and 6xHis-NP rHAZV were similar across the entire 5-day time period, with both viruses exhibiting peak titres at 3-4 days post-infection (figure 6.3b). These findings suggest that WT and 6xHis-NP rHAZV possess broadly similar growth parameters, demonstrating that the addition of the 6xHis tag does not interfere with virus multiplication and 6xHis-NP rHAZV represents a useful tool to study nairovirus biology.

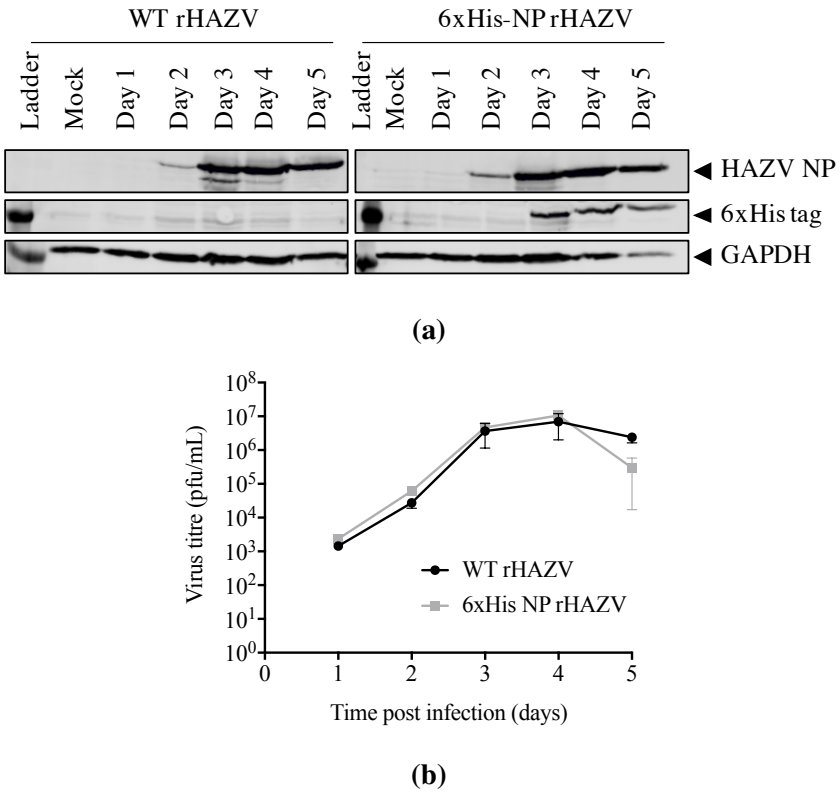


Figure 6.3: Growth kinetics of WT versus 6xHis-NP rHAZV. (a) SDS-PAGE and western blot analysis of cell lysates harvested at 24 h intervals for 5 days following infection with WT or 6xHis-NP rHAZV. GAPDH abundance was included as a loading control. **(b)** Titres of infectious WT and 6xHis-NP rHAZV harvested at 24 h intervals for 5 days following infection. pfu, plaque forming units.

6.1.2 Study of HAZV RNPs structure by TEM

Introduction to TEM. TEM utilises a beam of electrons transmitted through a specimen to form an image. An electron gun is used to generate the electron beam, which

is focused by the condenser lens in the column of the microscope and is restricted by the condenser aperture, which excludes high-angle electrons [258]. During this procedure, electrons only within a small range of energy can pass through, leading to a well-defined energy electron beam. To enhance the mean free path (average distance travelled between successive impacts) of electrons, the column in the microscope should maintain an extremely high vacuum [259].

Transmitted electrons are applied to the specimen placed in the back focal plane of the objective lens, which is responsible for forming the image. Specimens are prepared in TEM grids and are inserted into the microscope using a sample holder [259]. A mechanical arm is used to control the position of the specimen. Electrons can be scattered by the sample in two ways: elastic and inelastic scattering [258]. The path of elastically scattered electrons is altered with no energy loss, whereas inelastic scattering causes energy loss of the electron and results in X-ray emission, formation of free radicals and rearrangement of chemical bonds, consequently damaging the sample, increasing noise and limiting the resolution of the collected image [258]. The thickness of a TEM specimen usually should be within 100 nm for electrons to pass through and lower the chances of multiple scattering events.

The transmitted electrons are refocused by the objective lens and then magnified by an electromagnetic lens system consisting of the intermediate and projector lenses, and are projected on a detector or phosphorescent screen to convert the electron image information to a visible form (figure 6.4). Additional objective apertures can be used to enhance the contrast by blocking out high-angle diffracted electrons [259]. The phosphorescent plate glows when being hit by electrons during the TEM imaging process. Darker areas of the image represent those areas of the sample where fewer electrons are transmitted through, while lighter areas of the image represent those areas of the sample where more electrons are transmitted [259].

The biological specimen to be visualised is placed on a circular grid typically 3 mm in diameter, consisting of an ordered metal mesh with regularly spaced transparent square areas for data collection. A continuous layer of carbon is commonly deposited on top of the grid providing a support for biomolecules. Prior to sample application TEM grids are made hydrophilic by glow discharge treatment [260]. The specimen sample is then adsorbed to the grid and excess sample is blotted before the sample is stained with a heavy metal salt solution [261]. Excess of staining solution is then blotted and the grid is dried before loading into the microscope.

Negative staining is the most common procedure to stain biomolecules at room

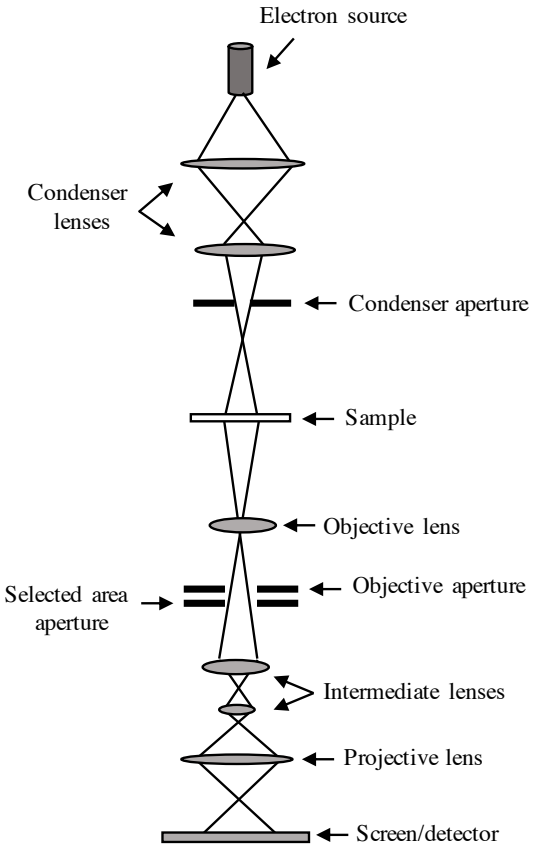


Figure 6.4: Schematic representation of a transmission electron microscope. The electron beam path is directed by electromagnetic lenses through apertures down the central axis of the microscope to produce a magnified image view on either the fluorescent screen or the detector, all held under high vacuum.

temperature [261]. Embedding the macromolecules in a layer of a heavy metal solution dehydrates the specimen preventing its collapse under the high vacuum of the microscope. The heavy metal stain causes a loss of electrons from reaching the detector, drastically increasing the amplitude contrast, which does not occur strongly from atoms with low atomic mass that make up the majority of biological complexes (O, N, C, H) [261]. Negative stain is a quick and simple procedure to assess the purity, overall morphology and homogeneity of isolated biological specimens. If the specimen adsorbs to the grid in random orientations, the data acquired can be used to produce 3D reconstructions. For high-resolution structure elucidation, cryogenic TEM (cryo-TEM) can be used, vitrifying the specimen and preserving its native hydrated state. Under vitrified conditions, chemical bonds can be broken while the atomic nuclei remain approximately in place, and much larger electron doses can be applied before destroying the specimen [262].

TEM overcomes many of the limitations of crystallography and NMR spectroscopy. Large macromolecular complexes are ideal for structural studies by TEM, although molecules smaller than 50 kDa are currently difficult to work with. In comparison with X-ray crystallography and NMR, much less specimen is required: 3 µL at µg scale are used per EM grid. Other advantages of TEM above X-ray crystallography or NMR include the requirement of samples with less purity and the possibility to perform structural studies of different conformational stages and highly flexible macromolecular structures.

The major hurdles for TEM single particle reconstruction are the high levels of noise obtained during data acquisition due to the use of limited electron doses to minimise radiation damage and the presence of conformational heterogeneity in the sample because of the flexible nature of the molecules and their different orientations in the 3D space [263]. As a result, most of the single particle reconstructions suffer from uneven distribution of resolution within the 3D EM maps, with the molecular core region normally having higher resolution than the periphery. In contrast, crystallisation of molecules or their subunits locks the molecules in the highly ordered crystal lattice and thus enforces them to stay in a rigid and ordered conformational state, allowing the reconstruction of electron density maps with a more evenly distributed resolution limit. EM and crystallography are thus complementary and both can contribute to reveal in-depth structural characteristics of biomolecules.

Results. An initial small-scale affinity purification experiment was performed to check that the 6xHis tag in the NP of 6xHis-NP rHAZV was accessible to bind the Ni^{2+}-NTA resin. Briefly, SW13 cells were infected with either WT or 6xHis-NP rHAZV. After 3 days, the cell lysate was harvested and used for purification of 6xHis-tagged RNPs by batch Ni^{2+}-NTA affinity chromatography. Samples corresponding to each of the

fractions from the different purification steps were analysed by SDS-PAGE and western blot using HAZV NP antisera as primary antibody (figure 6.5). WT rHAZV infected lysate fractions showed the presence of HAZV NP in the unbound material, whereas presence of HAZV NP could be detected in the high imidazole elution fractions of 6xHis-NP rHAZV infected cell lysate. These results suggest that the 6xHis tag in HAZV NP is exposed and accessible for its specific binding to Ni^{2+}-NTA resin.

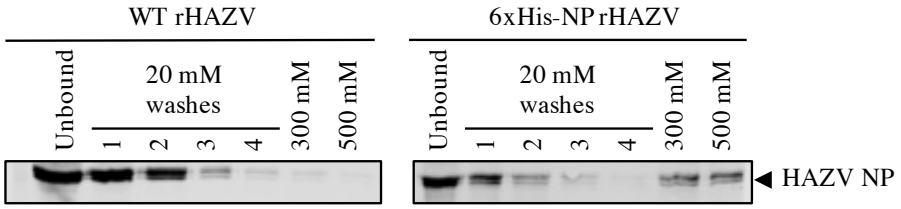


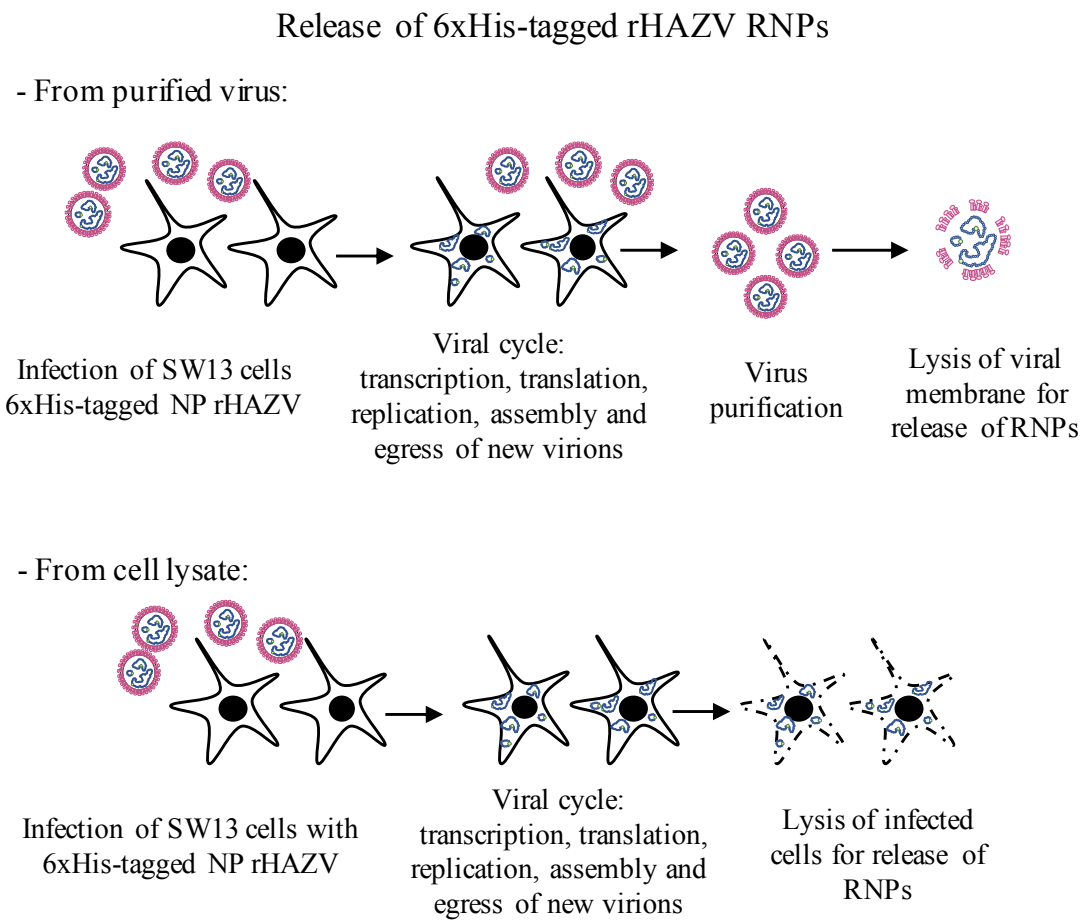
Figure 6.5: Batch Ni^{2+}-NTA affinity chromatography of WT and 6xHis-NP rHAZV infected cell lysates. SDS-PAGE and western blot analysis of samples corresponding to the different steps of WT (left) and 6xHis-NP (right) rHAZV batch affinity chromatography purification. SW13 cells were infected with WT or 6xHis-NP rHAZV and cell lysates were harvested 3 days post-infection.

With the aim of obtaining a purified sample of native HAZV RNPs, and to analyse any potential structural differences between virion- and cell-derived HAZV RNPs, two different procedures for large-scale purification of 6xHis-tagged RNPs were tested; harvesting of RNPs using affinity chromatography either from 6xHis-NP rHAZV purified virus samples, or from 6xHis-NP rHAZV infected cell lysates (figure 6.6).

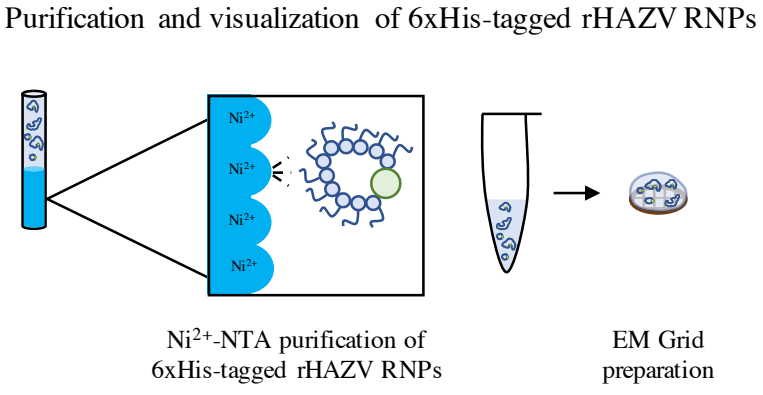
6.1.2.1 Analysis of HAZV RNPs from purified virus

6xHis-tagged RNPs purification was first attempted from purified virions. Briefly, large-scale SW13 cell cultures were infected with 6xHis-NP rHAZV, and 3 days post-infection virus was harvested and purified by sucrose cushion centrifugation. Fractions corresponding to the different purification steps were analysed by SDS-PAGE and western blot using HAZV NP antisera as primary antibody (figure 6.7a). The presence of HAZV NP was detected in the pellet sample after the sucrose cushion purification process, suggesting the successful purification of 6xHis-NP rHAZV. Negative stain EM analysis of the purified virus sample revealed the presence of pleomorphic virions, further corroborating the successful purification of the virus (figure 6.7b, 6.7c).

Viral membrane was lysed by freeze-thaw and detergent treatment to release 6xHis-tagged RNPs from virions. 6xHis-tagged RNPs were purified by Ni^{2+}-NTA affinity chromatography. Increasing imidazole concentration buffers were used to wash off unbound material and elution of 6xHis-tagged RNPs was achieved using high imidazole



(a)



(b)

Figure 6.6: Purification of 6xHis-tagged HAZV RNPs from purified virus and cell lysates. **(a)** Schematic representation of the two different protocols used for the release of 6xHis-tagged HAZV RNPs: from purified virus or cell lysate. **(b)** In both cases, released RNPs were purified by Ni^{2+}-NTA affinity chromatography and elution fractions were used for negative stain EM grid preparation.

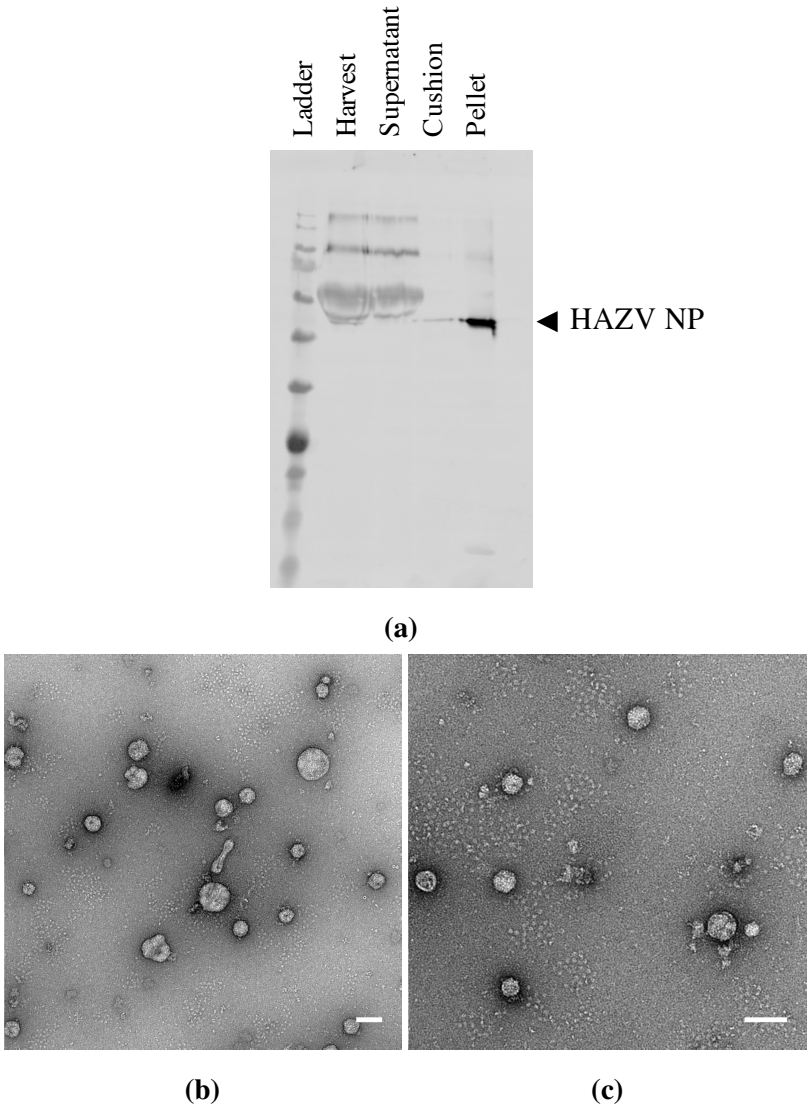


Figure 6.7: 6xHis-NP rHAZV virus purification. **(a)** SDS-PAGE and western blot analysis of samples corresponding to the different steps of 6xHis-NP rHAZV purification. SW13 cells were infected with 6xHis-NP rHAZV and supernatants were harvested 3 days post-infection. Virus was purified from supernatants by sucrose cushion purification. **(b-c)** TEM micrographs of purified 6xHis-NP rHAZV virions. Scalebar: 100 nm.

concentration buffer. Samples from all purification steps were analysed by SDS-PAGE and western blot analysis using HAZV NP antisera as primary antibody (figure 6.8a). HAZV NP was detected in elution fractions and these were used for negative stain EM sample preparation and analysis by TEM. TEM micrographs showed the presence of highly disordered coiled filaments within a considerably large amount of background material (figures 6.8b, 6.8c, 6.8d, 6.8e).

6.1.2.2 Analysis of HAZV RNPs purified from cell lysate

For 6xHis-tagged RNPs purification from infected cell lysate, large-scale SW13 cell cultures were infected with 6xHis-NP rHAZV, and 3 days post-infection cells were lysed with radioimmunoprecipitation assay buffer (RIPA) buffer. 6xHis-tagged RNPs were purified from cell lysate by Ni^{2+}-NTA affinity chromatography. Unbound material was washed off with increasing imidazole concentration buffers and elution of RNPs was achieved with high imidazole concentration buffer. Samples from all purification steps were analysed by SDS-PAGE and western blot analysis using HAZV NP antisera as primary antibody (figure 6.9a). The presence of HAZV NP was detected in elution fractions and these were used for negative stain EM sample preparation and analysis by TEM. TEM micrographs showed the presence of RNP filaments of different sizes, and non-specific background material was remarkably less abundant than in the case of 6xHis-tagged RNPs purified from virions (figure 6.9b).

To further characterise these filaments, their length was measured using the software Fiji for image analysis. Three different classes of RNPs could be distinguished according to their different lengths. The relative lengths of these three classes were directly proportional to the length in kilobasepairs (kbp) of the three different RNA segments of the virus. These findings suggest that each of the three filament classes that could be distinguished corresponds to RNPs packaging each of the three HAZV RNAs, namely the S, M and L segments (figure 6.10). The average RNA packaging extent of these filaments was calculated by the ratio of the length (nm)/kbp of RNA of each group of filaments. For the three different classes the ratio nm/kbp was determined as 100-105 nm/kbp, suggesting that the extent to which RNA is packaged is equal for all three RNA segments, and three times more compacted than linear ssRNA (300 nm/kbp) [264]. The abundance of particles observed for each of the respective classes was also analysed: 51.47% of the particles corresponded to the S segment RNP class, whereas particles of the M and L segments only represented the 27.94% and 22.06% of the RNPs observed, respectively (figures 6.10b, 6.10c). Since particles were manually picked, this difference

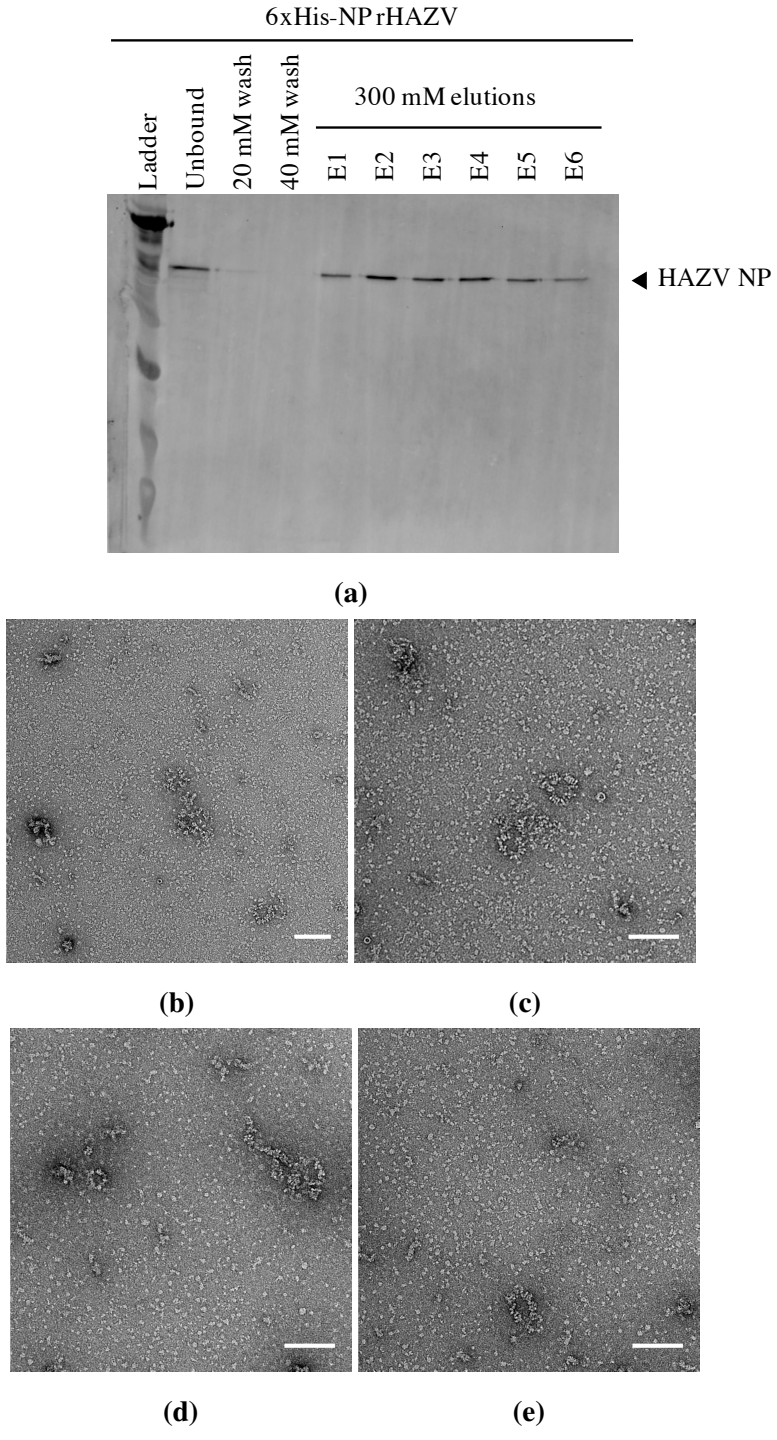


Figure 6.8: Purification of 6xHis-NP rHAZV RNPs from purified virus. **(a)** SDS-PAGE and western blot analysis of samples corresponding to the different steps of 6xHis-NP rHAZV RNPs purification from lysed virions. **(b-e)** TEM micrographs of purified 6xHis-NP rHAZV RNPs from lysed virions. Scalebar: 100 nm.

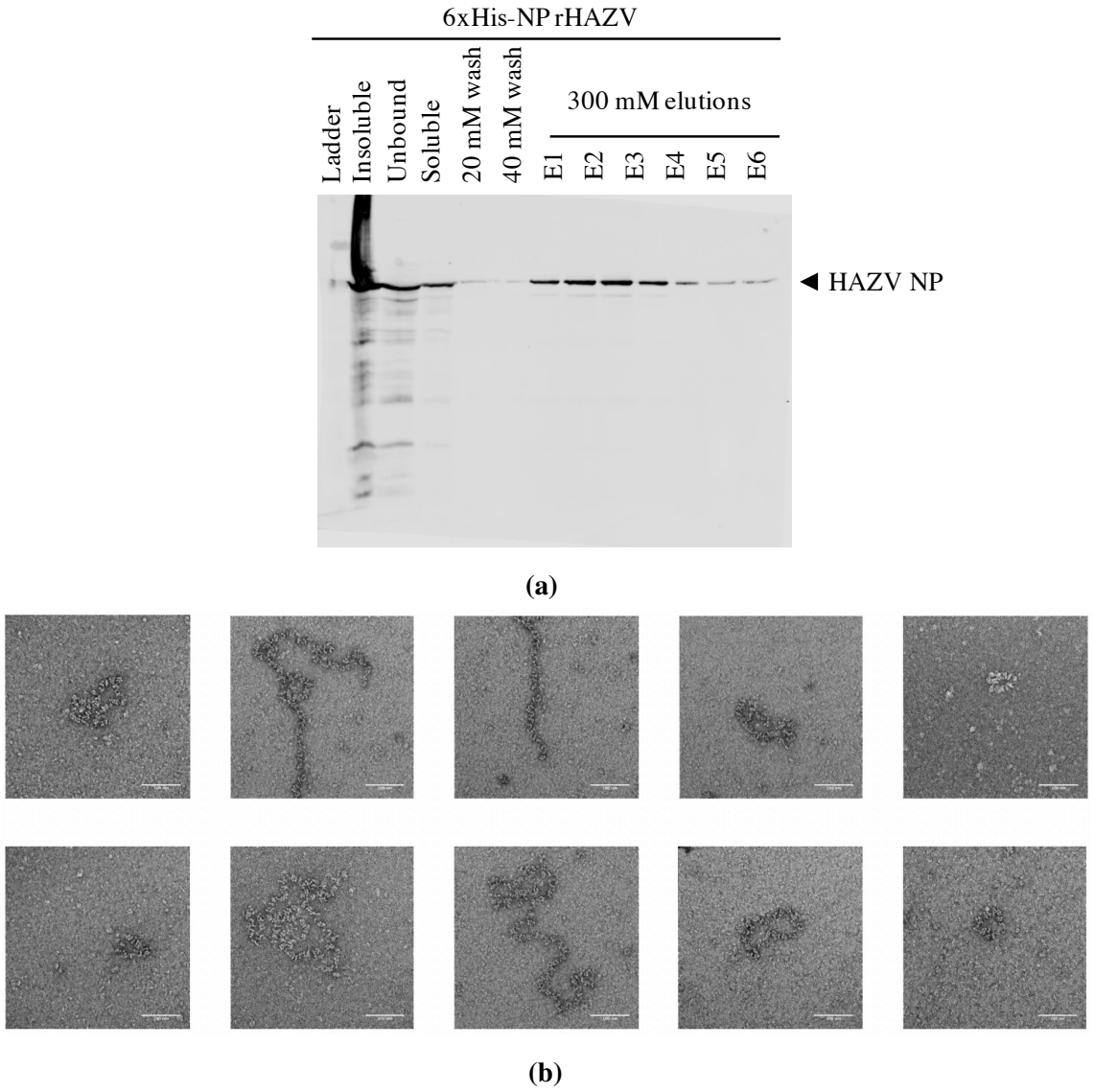


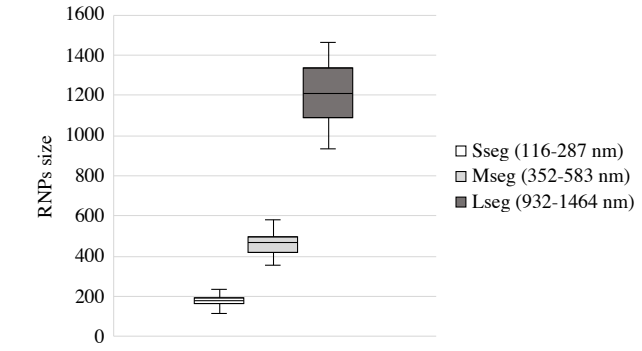
Figure 6.9: Purification of 6xHis-NP rHAZV RNPs from infected cell lysate. **(a)** SDS-PAGE and western blot analysis of samples corresponding to the different steps of 6xHis-NP rHAZV RNP purification from infected cell lysate. **(b)** TEM micrographs of purified 6xHis-NP rHAZV RNPs from infected cell lysate. Scalebar: 100 nm.

may be due to experimental bias and thus cannot be definitively attributed to a biologically relevant event.

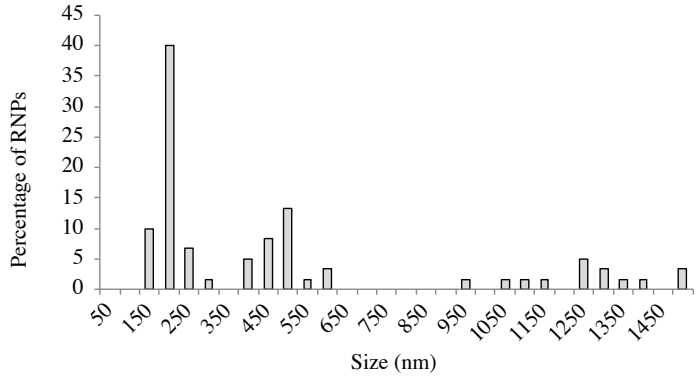
The width of the RNP particles was measured and the three different classes observed presented a similar width of 16 nm (figure 6.11a). This width is analogous to the one described for recombinant CCHFV NP pentamers. Due to similarities in the structure of CCHFV and HAZV NPs, a HAZV NP pentamer is expected to present a similar width (figure 6.11), suggesting that native HAZV RNPs are potentially assembled in helical filaments based on pentameric turns (figure 6.11d).

To examine whether a similar 6xHis tag could be incorporated in replication-competent CCHFV RNPs, the activity of the CCHFV mini-genome system (previously described in section 2.8.6) in the presence of a 6xHis-tagged NP support plasmid was tested. The support plasmid expressing CCHFV NP tagged with a 6xHis C-terminal tag (pC-NP-6xHis) was generated by SDM starting from the previously described pC-NP support plasmid (see plasmid maps in Appendix B). BSR-T7 cells were transfected with the CCHFV mini-genome system plasmids (pC-NP, pC-L and pC-SsegUTRs-eGFP), or replacing the pC-NP plasmid with the 6xHis-tagged CCHFV NP support plasmid (pC-NP-6xHis). A control transfection (pC-NP and pC-SsegUTRs-eGFP) omitting the essential pC-L plasmid and non-transfected cells (C-) were also included. eGFP signal (green counts/well) was monitored over time using an IncuCyte live cell imaging system.

CCHFV-specific replicon activity was similar using WT CCHFV NP support and 6xHis-NP support plasmids. These results suggest that 6xHis-tagged NPs can support CCHFV replication, and thus could potentially be used to purify and study native CCHFV RNPs.



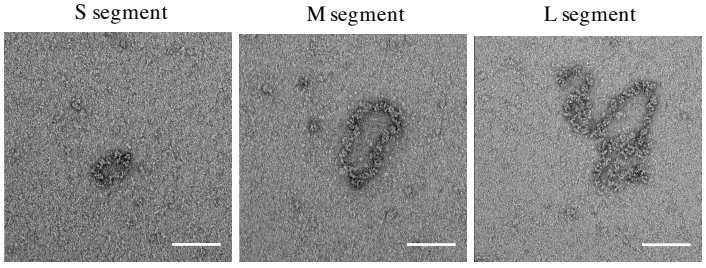
(a)



(b)

	S segment	M segment	L segment
Size ± SD (nm)	176.34 ± 32.16	459.23 ± 61.6	1219.48 ± 158.16
kbp	1.677	4.575	11.98
nm/kbp	105.154817	100.378717	101.793209
% of RNPs	51.47	27.94	22.06

(c)



(d) **(e)** **(f)**

Figure 6.10: Length and frequency distribution analyses of RNP particles. **(a)** Three different classes of RNPs can be distinguished according to their lengths. Data obtained from analysis of 68 particles. **(b)** RNPs size frequency distribution. Data obtained from analysis of 68 particles. **(c)** Table representing the length in nm and kbp, the ratio of these (nm/kbp) and percentage inside the whole population (%) of each RNP class. **(d-f)** Representative TEM micrographs of each of the RNP classes, corresponding to S, M and L segments. Scalebar: 100 nm.

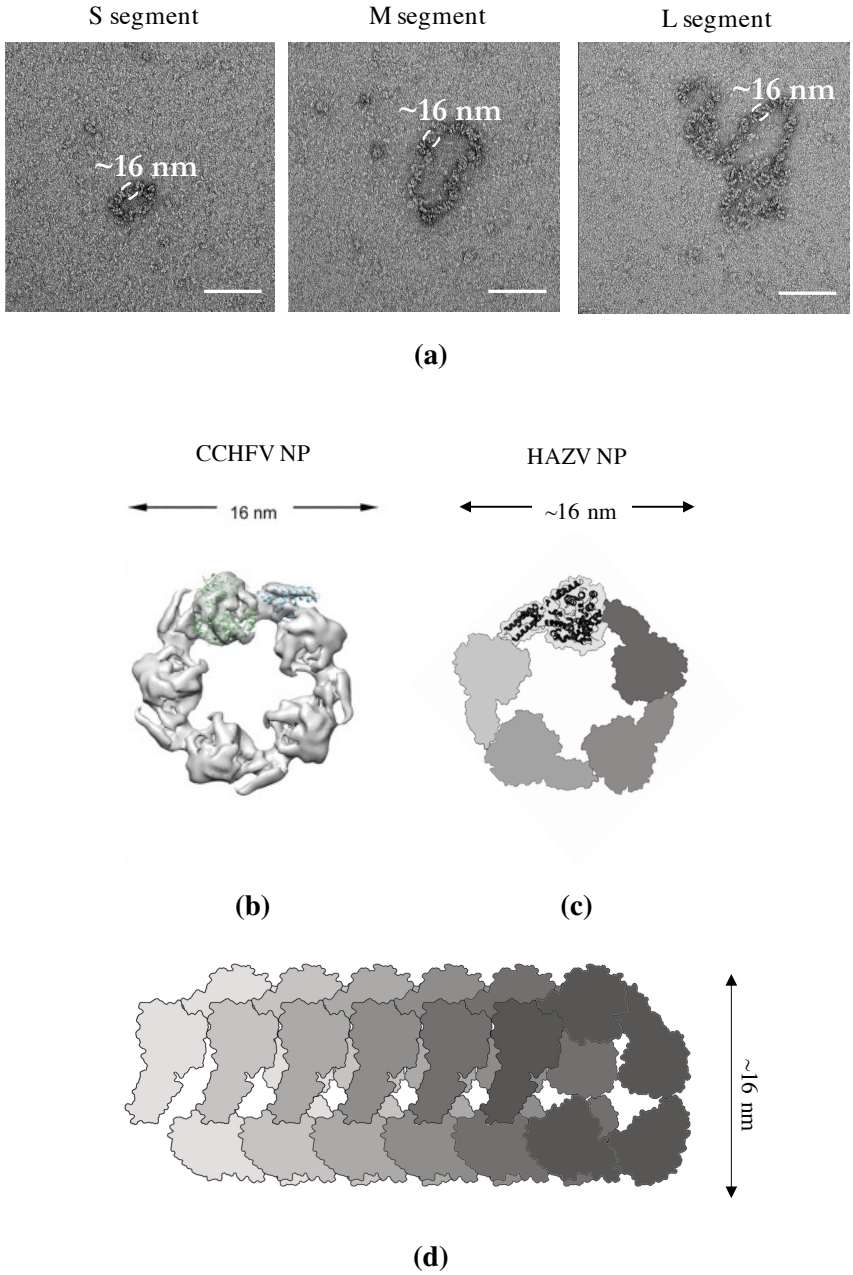


Figure 6.11: Schematic of a potential RNP packaging model. **(a)** The three RNP classes, corresponding to S, M and L segments, present a similar width of 16 nm. **(b)** Structural model of recombinant CCHFV NP pentamer, modified from [149]. **(c)** Potential model of HAZV NP pentamer using CCHFV NP as reference model. Due to similarities in CCHFV and HAZV NPs structure, HAZV NP pentamer is expected to present a width of 16 nm. **(d)** Potential arrangement of HAZV NP in native HAZV RNPs. The width observed for HAZV RNPs suggests a possible helical arrangement based on pentamers. Scalebar: 100 nm.

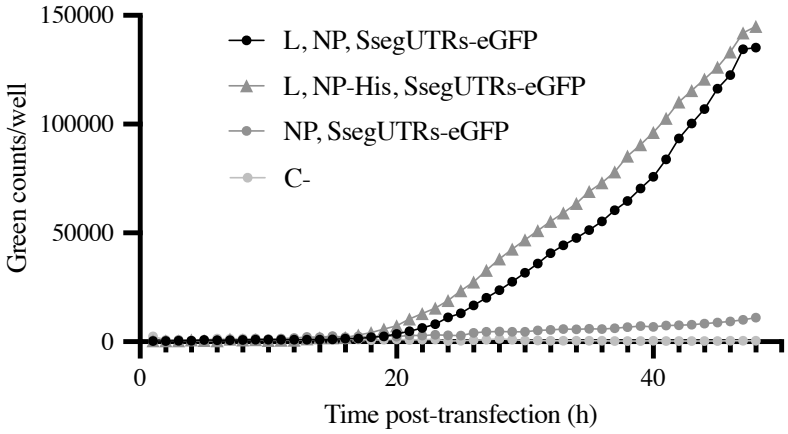


Figure 6.12: Analysis of CCHFV mini-genome replication using a 6xHis-tagged CCHFV NP support plasmid. BSR-T7 cells were transfected with the CCHFV mini-genome plasmids (pC-NP, pC-L and pC-SsegUTRs-eGFP), or the system replacing the pC-NP plasmid with a 6xHis-tagged CCHFV NP support plasmid (pC-NP-6xHis). Control transfection (pC-NP and pC-SsegUTRs-eGFP) omitting transfection of the essential pC-L plasmid and non-transfected cells (C-) were also tested. eGFP signal (green counts/well) was monitored over time using an IncuCyte live cell imaging system.

6.2 Development of eGFP-β11-tagged NP rHAZV

Once it was confirmed that infectious HAZV could be recovered with a small tag incorporated in the NP, the feasibility of generating a rHAZV for which NP molecules within the RNP were appended to a fluorescent tag was explored, which would allow live cell tracking of RNPs throughout the entire nairovirus replication cycle. The possibility of tagging HAZV NP with components of the split-eGFP system was tested, for which full length eGFP (eGFPcomp) assembles by interaction of one small and one large constituent fragment. The small component is a 16-amino acid C-terminal eGFP peptide (eGFP-β11) corresponding to the 11^{th} beta strand of the eGFP beta barrel, and the large complementary eGFP fragment (eGFP-β1-10) corresponds to the 10 remaining beta strands of the eGFP molecule [265, 266, 213]. Examination of the HAZV NP crystal structure [151] suggested that a C-terminal eGFP-β11 tag would be exposed and accessible for trans-complementation with the eGFP-β1-10 fragment, possibly resulting in the assembly of rHAZV for which its NP was appended with an intact complemented eGFP molecule (eGFPcomp) capable of emitting fluorescence (figure 6.1b).

6.2.1 Rescue and characterisation of eGFP-β11-tagged NP rHAZV

First, the ability of eGFP-β1-10 and eGFP-β11 fragments to assemble and fluoresce in cells when appended to HAZV NP was explored. To do this, the eGFP-β11 fragment sequence and a flexible linker were inserted downstream of the C-terminus of the HAZV NP ORF within the previously described HAZV S segment expression plasmid (pMK-RQ-S) [156], using SDM, restriction enzyme digestion and ligation to generate pMK-RQ-S-eGFP-β11 (figure 6.13a, see plasmid maps in Appendix B). Next, the pMK-RQ-S-eGFP-β11 plasmid was co-transfected into BSR-T7 cells with pcDNA3.1-eGFP(1-10), which expresses the large complementary eGFP fragment (eGFP-β1-10, figure 6.13b). Emission of NP-eGFPcomp fluorescence was monitored over time using an IncuCyte live cell imaging system, and a pcDNA3.1-eGFP(1-10) only control was performed alongside to measure any basal fluorescence of the eGFP-β1-10 fragment on its own. Only co-transfection of the pMK-RQ-S-eGFP-β11 and the pcDNA3.1-eGFP(1-10) resulted in an increase of the number of fluorescent cells over time (figures 6.13c, 6.13d), not observed in the pcDNA3.1-eGFP(1-10) only control. The fluorescent signal showed a diffuse cytoplasmic pattern, expected due to the unspecific localisation of the transiently expressed trans-complemented HAZV NP-eGFPcomp. These results demonstrate that the eGFP-β11 tag in HAZV NP-eGFP-β11 is accessible and can functionally-associate with eGFP-β1-10 supplied by transient expression.

Based on the finding that HAZV NP-eGFP-β11 can functionally-associate with eGFP-β1-10 in cells, the next step was to establish whether a rHAZV expressing NP fused to the eGFP-β11 fragment could be rescued as infectious virus. To examine this possibility, the previously described system for the rescue of infectious rHAZV from cDNA was used [156]. Transfection of modified plasmid pMK-RQ-S-eGFP-β11 along with wild type (WT) M and L segment expressing-plasmids, and the T7 pol expression plasmid was performed alongside a transfection comprising all three WT plasmids for comparison. At 96 hpt, supernatants were harvested and used to re-infect fresh monolayers of SW13 cells for 48 h. Western blot analysis of corresponding lysates revealed abundant HAZV NP expression, confirming successful rescue of both rHAZV and rHAZV-eGFP-β11 viruses (figure 6.14a). As expected, no HAZV NP was detected in control rescues in which the L segment expressing plasmid was omitted.

To confirm the presence of the eGFP-β11 tag, RT-PCR analysis was performed on RNAs extracted from WT rHAZV and rHAZV-eGFP-β11 stocks to amplify a cDNA fragment corresponding to the C-terminal portion of the HAZV NP ORF. The cDNAs were analysed on a 1% (w/v) agarose gel, which revealed the fragment amplified from

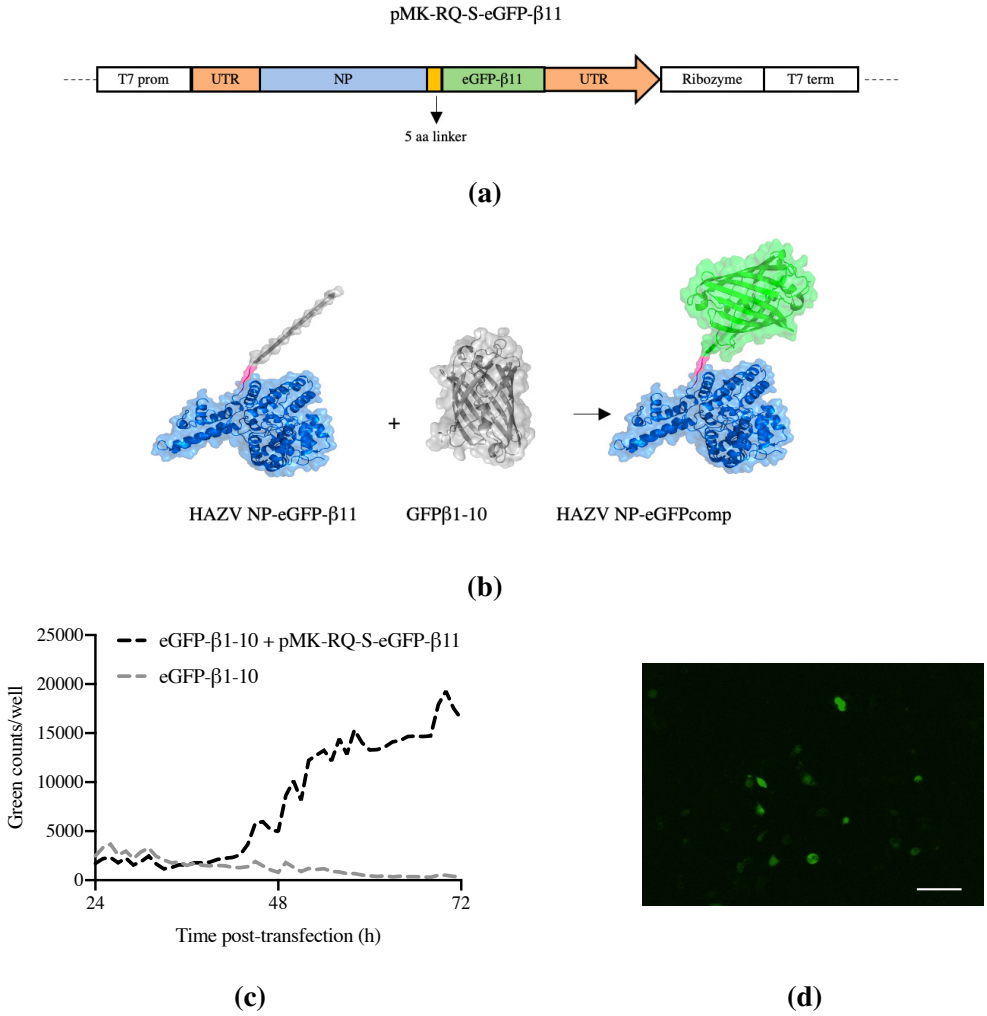


Figure 6.13: Trans-complementation of HAZV NP-eGFP-β11 with eGFP-β1-10. **(a)** Schematic representation of the pMK-RQ-eGFP-β11 plasmid, designed to express a modified HAZV S segment containing the ORF of the HAZV NP tagged with a C-terminal eGFP-β11 fragment under a T7 pol promoter (T7 prom). A 5 aa flexible linker and the eGFP-β11 fragment sequence were inserted between the HAZV NP ORF and the adjacent viral UTR. The T7 pol terminator (T7 term) and hepatitis delta virus ribozyme (Ribozyme) in the 3' end ensure transcription termination and generation of a native 3' terminus, respectively. **(b)** Schematic representation of the split-eGFP system for HAZV NP fluorescent tagging. The eGFP-β11 tag (grey) and 5 aa linker (pink) are extended from the C-terminal end of the HAZV NP forming HAZV NP-eGFP-β11. Trans-complementation of HAZV NP-eGFP-β11 with eGFP-β1-10, which does not fluoresce on its own, results in the assembly of HAZV NP-eGFPcomp, formed by a complete eGFP fluorescent moiety attached to HAZV NP. **(c)** Transient expression of eGFP-β1-10 and pMK-RQ-eGFP-β11 results in an increase of green fluorescent signal (green object counts/well) over time. BSR-T7 cells were co-transfected with eGFP-β1-10 and pMK-RQ-eGFP-β11 expression plasmids and fluorescent signal was monitored over time. Transfection of eGFP-β1-10 only did not result in fluorescent signal. **(d)** Live cell image of BSR-T7 cells transfected with eGFP-β1-10 and pMK-RQ-eGFP-β11 taken at 60 hpt. Scale bar: 200 µm.

rHAZV-eGFP-β11 was increased in size relative to WT rHAZV, indicating maintenance of the tag (figure 6.14b). The sequence of the tag was further confirmed by subsequent nucleotide sequencing of the RT-PCR-amplified fragment.

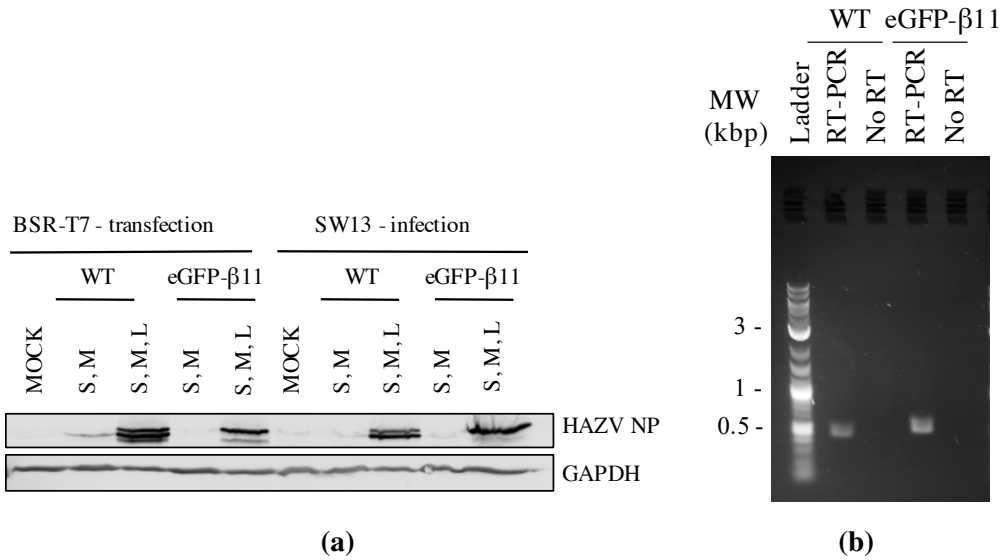


Figure 6.14: Rescue of WT rHAZV and rHAZV-eGFP-β11 infectious viruses. (a) BSR-T7 cells were transfected with the T7 pol support plasmid and the three rescue plasmids (SML) for recovery of wild-type (WT) or rHAZV-eGFP-β11 virus (eGFP-β11). At 96 hpt supernatants were used to infect SW13 monolayers for 48 h. Lysates from BSR-T7 and SW13 cells were analysed for expression of HAZV NP by western blot analysis using HAZV NP antisera. Lysates were also collected from non-transfected cells (MOCK) and a recovery experiment in which the L segment expression plasmid was omitted (S, M). GAPDH abundance was included as a loading control. **(b)** RT-PCR analysis of recovered virus. Virus stocks were used for RNA extraction, reverse transcription (RT) and PCR amplification of the C-terminal region of the HAZV NP ORF and the adjacent UTR. Corresponding DNA products were analysed on a 1% (w/v) agarose gel showing a slightly bigger size for pMK-RQ-S-eGFP-β11 than for pMK-RQ-S (WT) due to the presence of the C-terminal tag, further confirmed by Sanger sequencing.

Next, whether eGFP trans-complementation could also occur in the context of rHAZV-eGFP-β11 infection was explored. BSR-T7 cells were transfected with the eGFP-β1-10 fragment and subsequently infected with different MOIs (0.1, 0.01, 0.001) of rHAZV-eGFP-β11. An increase in the number of fluorescent cells was observed over time, with a higher number of fluorescent cells observed when higher MOIs were employed (figure 6.15a). This indicated assembly of the trans-complemented HAZV NP-eGFPcomp as well as suggesting spread of infectious virus through the culture. The highest fluorescence signal was obtained at 72 hours post-infection (hpi), demonstrating rHAZV-eGFP-β11 presented similar growth kinetics to WT rHAZV [156].

In contrast to the previous experiment where the two trans-complementing plasmids were co-transfected to form NP-eGFPcomp (figure 6.13), in the context of rHAZV-eGFP-β11 infection, the fluorescent signal was not unspecific but instead showed a characteristic localisation pattern (figure 6.15b); at early time points of infection the eGFPcomp signal was observed in a discrete perinuclear region, and as the infection advanced the pattern acquired a more diffuse cytoplasmic distribution concluding with the apoptosis of the infected cells, suggesting viral processes posed a dominant influence on the localisation of the NP-eGFPcomp fusion (figure 6.15c).

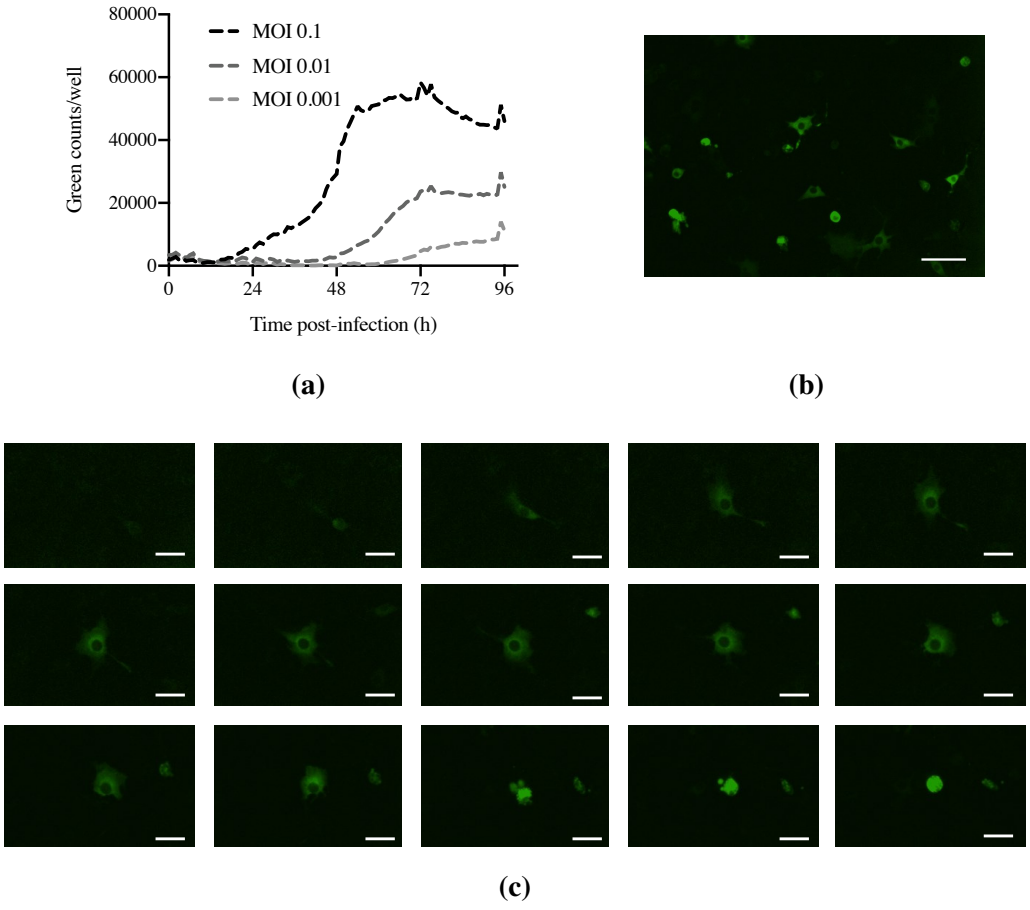


Figure 6.15: Trans-complementation of HAZV NP-eGFP-β11 with eGFP-β1-10 results in HAZV NP fluorescent tagging in the context of virus infection. **(a)** Transient expression of eGFP-β1-10 and infection with rHAZV-eGFP-β11 results in an increase of green fluorescent signal (green object counts/well) over time. BSR-T7 cells were transfected with eGFP-β1-10 and, at 24 hpt, infected with rHAZV-eGFP-β11 at the specified MOIs. Fluorescent signal was monitored over time. **(b)** Live cell image of BSR-T7 cells transfected with eGFP-β1-10 and infected with rHAZV-eGFP-β11 taken at 60 hpi. Scale bar: 200 µm. **(c)** Live cell images of BSR-T7 cells transfected with eGFP-β1-10 and infected with rHAZV-eGFP-β11. All images represent the same area in the well, and were taken at consecutive 1h intervals. Scale bar: 50 µm.

To further confirm that the intracellular distribution of the eGFPcomp signal was associated with HAZV NP, immunofluorescence analyses were performed using HAZV NP antisera and confocal microscopy. BSR-T7 cells were transfected with the eGFP-β1-10 expressing plasmid and subsequently infected with rHAZV-eGFP-β11 at an MOI of 1 for 24 h. Cells were fixed and immunostained using HAZV NP antisera. All cells showing eGFP signal also presented staining for HAZV NP (figure 6.16) with a similar cytoplasmic distribution pattern, demonstrating a similar localisation for both eGFPcomp and HAZV NP and showing that rHAZV-eGFP-β11 infection was a requisite for the presence of eGFPcomp signal, as expected. The presence of cells positive for HAZV NP staining but lacking eGFPcomp signal, suggested the transfection and transient expression of the eGFP-β1-10 fragment was not efficient in all cells (white arrow in figure 6.16). Taken together, these observations confirm the successful trans-complementation of HAZV NP-eGFP-β11 with the eGFP-β1-10 fragment, resulting in eGFPcomp signal that associates with HAZV NP in the context of viral infection.

6.2.2 Study of the intracellular localisation of HAZV components by confocal microscopy

Introduction to confocal microscopy. Confocal microscopy is an optical imaging technique that relies on the use of a spatial pinhole to block out-of-focus light in image formation [267]. The principle of confocal microscopy arose to overcome some of the limitation of traditional wide-field fluorescence microscopy. In conventional fluorescence microscopy, areas of the specimen above and below the focal plane contribute to the resulting fluorescence emission, which is detected by the microscope's photodetector or camera, and includes a large unfocused background signal [267]. Confocal microscopy uses point illumination and a pinhole in an optically conjugate plane in front of the detector to eliminate out-of-focus signa [268]. A pair of oscillating mirrors raster scan a point of laser light across the specimen via the objective. Fluorescence emitted by the specimen passes back through the mirror systems to a beam splitter which rejects any reflected excitation wavelengths, and then through the pinhole (figure 6.17) [268]. The beam is scanned across the sample in the horizontal plane, the detector records the brightness of fluorescence at each raster point and maps this into a 2D image [267]. Slower scans provide a better signal-to-noise ratio, resulting in better contrast.

As only light produced by fluorescence very close to the focal plane can be detected, the optical resolution of the image, particularly in the sample depth direction, is much better than that of wide-field microscopes. However, as much of the light from sample

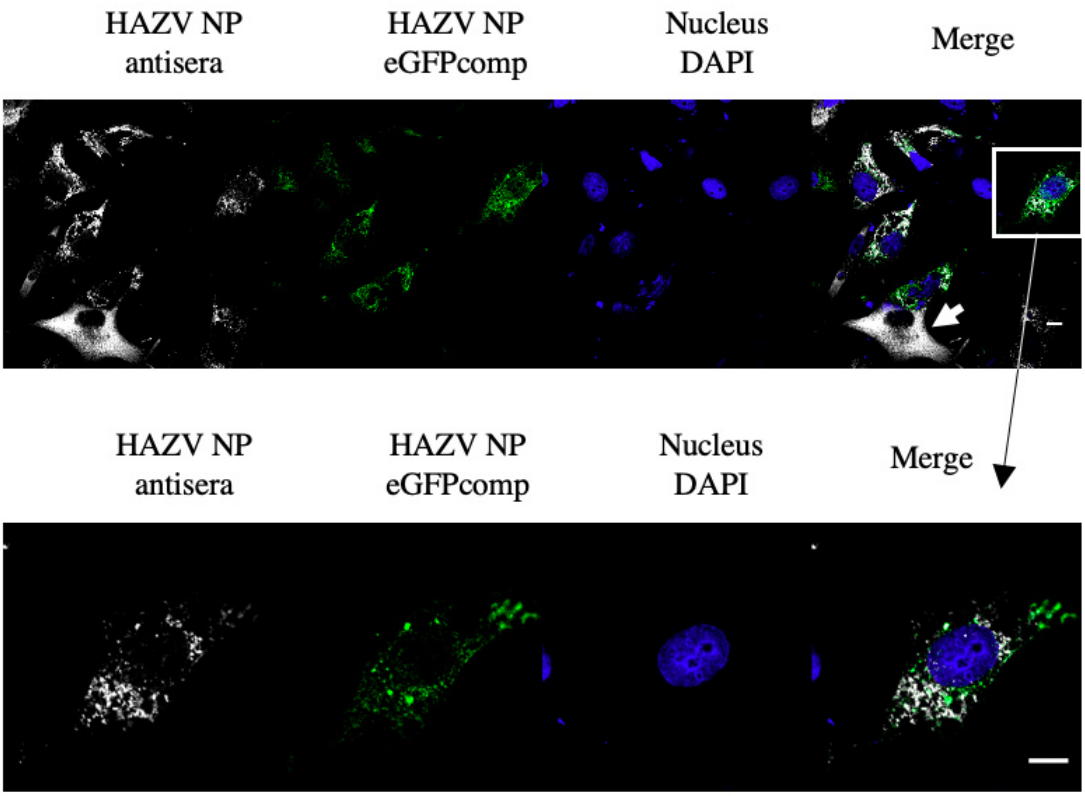


Figure 6.16: eGFPcomp co-localises with HAZV NP. Confocal microscopy images of BSR-T7 cells transfected with eGFP-β1-10 and infected with HAZV NP-eGFP-β11 and closer look of the area highlighted with a square. Cells were fixed at 24 hpi and stained with HAZV NP antisera (white) and DAPI (blue). All cells showing eGFP signal also presented staining for HAZV NP. White arrow indicates the presence of a cell positive for HAZV NP staining but lacking eGFPcomp signal, suggesting that the transfection and transient expression of the eGFP-β1-10 fragment was not efficient in all cells. Scale bar: 10 μm.

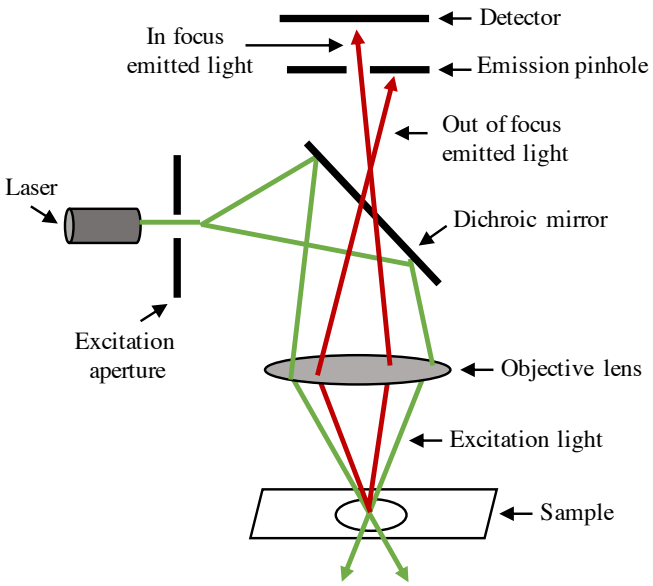


Figure 6.17: Schematic representation of a confocal microscope. Light emitted by the laser system passes through an excitation pinhole aperture, is reflected by a dichromatic mirror and is scanned across the specimen in a defined focal plane. Secondary fluorescence emitted from points on the specimen pass back through the dichromatic mirror and are focused as a confocal point at the detector pinhole aperture, which acts as a physical barrier to out of focus emitted fluorescence.

fluorescence is blocked at the pinhole, this increased resolution is at the cost of decreased signal intensity [267]. Sensitive detectors such as photomultiplier tubes (PMT) or avalanche photodiodes transform the light signal into an electrical one. As only one point in the sample is illuminated at a time, scanning over the specimen is required to obtain 2D or 3D images.

6.2.2.1 Study of HAZV NP localisation

The results of the previous section showed that rHAZV-eGFP-β11 was able to complete its multiplication cycle in eGFP-β1-10 expressing cells, and the split-eGFP fragments were able to assemble while appended to HAZV NP. These results suggested attachment of eGFPcomp to HAZV NP did not ablate any critical NP functions. Next, the ability of the eGFP-β1-10 fragment to be incorporated into infectious rHAZV-eGFP-β11 particles was tested, which would allow the formation of fluorescently-labelled eGFPcomp-RNPs that could be passed into fresh cells.

To do this, large-scale virus amplification was performed by inducing transient expression of the eGFP-β1-10 fragment in BSR-T7 cells that were infected with rHAZV-eGFP-β11. At 72 hpi, supernatants were collected and used for the concentration and

purification of rHAZV-eGFPcomp by sucrose cushion centrifugation, with the viral envelope subsequently labelled with Vybrant DiD, a lipophilic fluorescent dye with peak emission at 665 nm, distinct from eGFPcomp (509 nm). A second sucrose cushion was performed to remove any excess dye (figure 6.18a) and further increase virus purity. Western blot analyses of fractions taken from the sucrose cushions, supernatant and pellets after ultracentrifugation revealed the presence of HAZV NP in the pellet, suggesting purification of the dually-labelled rHAZV-eGFPcomp virus was successful (figure 6.18b).

Purified, dually-labelled rHAZV-eGFPcomp was used to infect a fresh set of BSR-T7 cells, in which eGFP-β1-10 was not supplied, thus any eGFPcomp fluorescence must have originated within the infecting virus particles. Cells were subsequently fixed at 8 or 24 hpi and stained with HAZV NP antisera. At 8 hpi, signal corresponding to HAZV NP-eGFPcomp was clearly detected in a perinuclear region, which also co-localised with signals corresponding to HAZV NP antisera and Vybrant DiD (figure 6.18c). This confirmed the ability of the rHAZV-eGFP-β11 to incorporate eGFP-β1-10 into virions and furthermore demonstrated these particles were capable of carrying out multiple stages of the virus replication cycle, including assembly, egress, internalisation, entry and delivery of fluorescent RNPs into new cells. The assembly of rHAZV-eGFPcomp now allowed the location of the infecting RNPs to be determined throughout the entire internalisation and entry process.

Above, the results showed that at 8 hpi the signals for eGFPcomp closely localised with signals corresponding to HAZV NP antisera and Vybrant DiD within perinuclear regions. However, visualising infected cells at the later time point of 24 hpi revealed that while the HAZV NP-eGFPcomp and Vybrant DiD signals remained in a discrete perinuclear region, the distribution of HAZV NP detected by immunostaining was different. At 24 hpi, immunostaining revealed the presence of additional NP, the majority of which was not detected by eGFPcomp-associated fluorescence. This observation is consistent with the ability of the input rHAZV-eGFPcomp to undergo gene expression, producing newly synthesised HAZV NP. However, while input NP-eGFPcomp can be detected by fluorescent emission, any newly produced HAZV NP cannot, due to the lack of supply of eGFP-β1-10 for trans-complementation in these cells. Thus, fluorescence microscopy of rHAZV-eGFPcomp infected cells, in combination with immunostaining, allowed the differentiation between input and newly synthesised HAZV NP.

Interestingly, the localisation of this newly synthesised NP lacking eGFPcomp excluded the perinuclear sites where eGFP and Vybrant DiD signals were present (figure 6.19), indicating an inverse spatial relationship between the input signals (membrane: Vybrant DiD label, HAZV NP-eGFP) and the newly synthesised HAZV NP (antisera

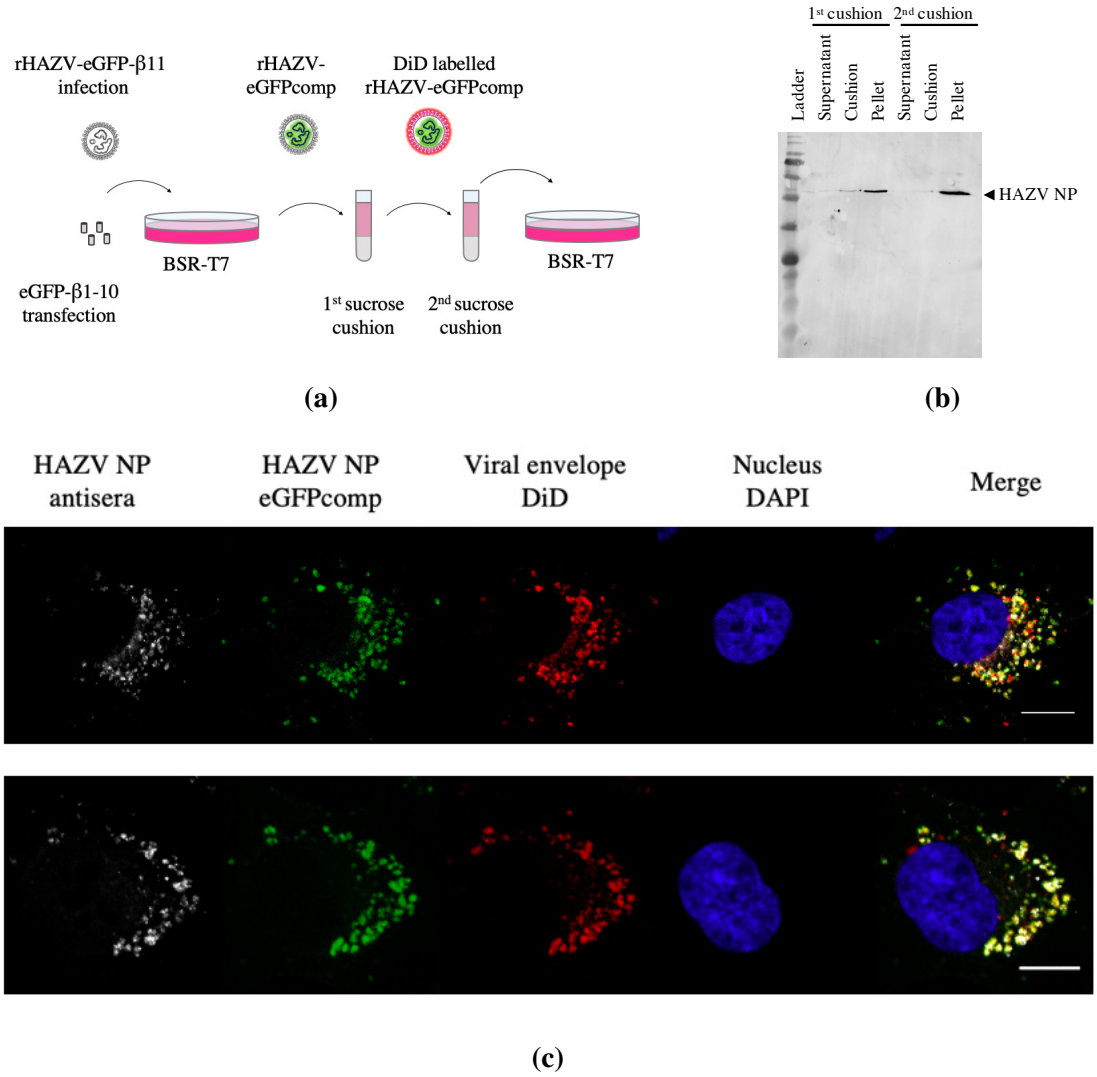


Figure 6.18: Dual labelling of HAZV and sub-cellular localisation of viral components at 8 hpi. **(a)** Schematic representation of the workflow followed for dual labelling of rHAZV-eGFP-β11. BSR-T7 cells were transfected with eGFP-β1-10 and infected with rHAZV NP-eGFP-β11. rHAZV NP-eGFPcomp was purified from supernatant by sucrose cushion. Viral membrane was subsequently labelled with Vybrant DiD dye. After a second sucrose cushion, purified Vybrant DiD labelled rHAZV-eGFPcomp was used for infection of fresh BSR-T7 monolayers and confocal microscopy analyses. **(b)** SDS-PAGE and western blot analysis of samples belonging to sucrose cushions, supernatant and pellets after ultracentrifugation. HAZV NP was detected using HAZV NP antisera. **(c)** Confocal microscopy images of BSR-T7 cells infected with Vybrant DiD labelled rHAZV NP-eGFPcomp. Cells were fixed at 8 hpi and stained with HAZV NP antisera (white) and DAPI (blue). Fluorescence signal corresponding to HAZV NP-eGFPcomp co-localised with signals corresponding to HAZV NP antisera and Vybrant DiD in a perinuclear region. Scale bar: 10 µm.

labelling). Taken together, these results reveal the localisation of input viral components is perinuclear, where they remain for the duration of the infection. These sites are distinct from newly synthesised HAZV NP, which migrates away from the perinuclear region to display a more diffuse cytoplasmic distribution at later times of the infection cycle. The maintenance of a discrete DiD signal at late timepoints post-infection suggests that some of the virus particles may have not successfully fused with the endosomes and dead-end infection events are potentially being observed simultaneously to productive infections.

6.2.2.2 Study of (+) sense HAZV RNA localisation

Next, the intracellular location of viral RNA synthesis activities was examined, to determine whether viral RNA was associated with the perinuclear input HAZV NP or with the diffusely distributed cytoplasmic newly synthesised HAZV NP. To do this, a fluorescence *in situ* hybridisation (FISH) probe set was designed, comprising multiple non-overlapping fluorescently-labelled short oligonucleotides with near complete coverage of the HAZV S segment positive sense RNA target. As input HAZV RNAs are negative sense, this FISH probe set would only detect newly synthesised HAZV-specific RNAs either as mRNA transcripts or anti-genomic replication products.

BSR-T7 cells were infected with WT rHAZV at an MOI of 1, fixed and immunostained with HAZV NP antisera, and then subjected to FISH analysis. At 8 hpi, immunostaining for HAZV NP showed input HAZV RNPs were localised to the perinuclear regions of the cell (figures 6.20a), as expected from our findings presented above (figure 6.18). FISH analysis at this time point showed that (+) sense HAZV RNA was also localised in close proximity to NP. These findings suggest that input RNPs that localise to perinuclear regions are able to generate (+) sense RNA, thus identifying this region of the cell as the initial site of HAZV RNA synthesis.

At 24 hpi, the HAZV NP antisera staining revealed a diffuse cytoplasmic distribution, as expected from the results described above (figure 6.19). Interestingly, FISH analysis revealed this distribution was also reflected in that of (+) sense RNA, revealing the sites of HAZV RNA synthesis were not confined to perinuclear regions, but instead had expanded to occupy a greater area within the cell (figures 6.20b. Taken together, these findings suggest newly made NP assembles into RNPs that are active for RNA synthesis, which later migrate away from the initial perinuclear sites of RNA synthesis, resulting in the expansion of the intracellular space occupied by the virus factory.

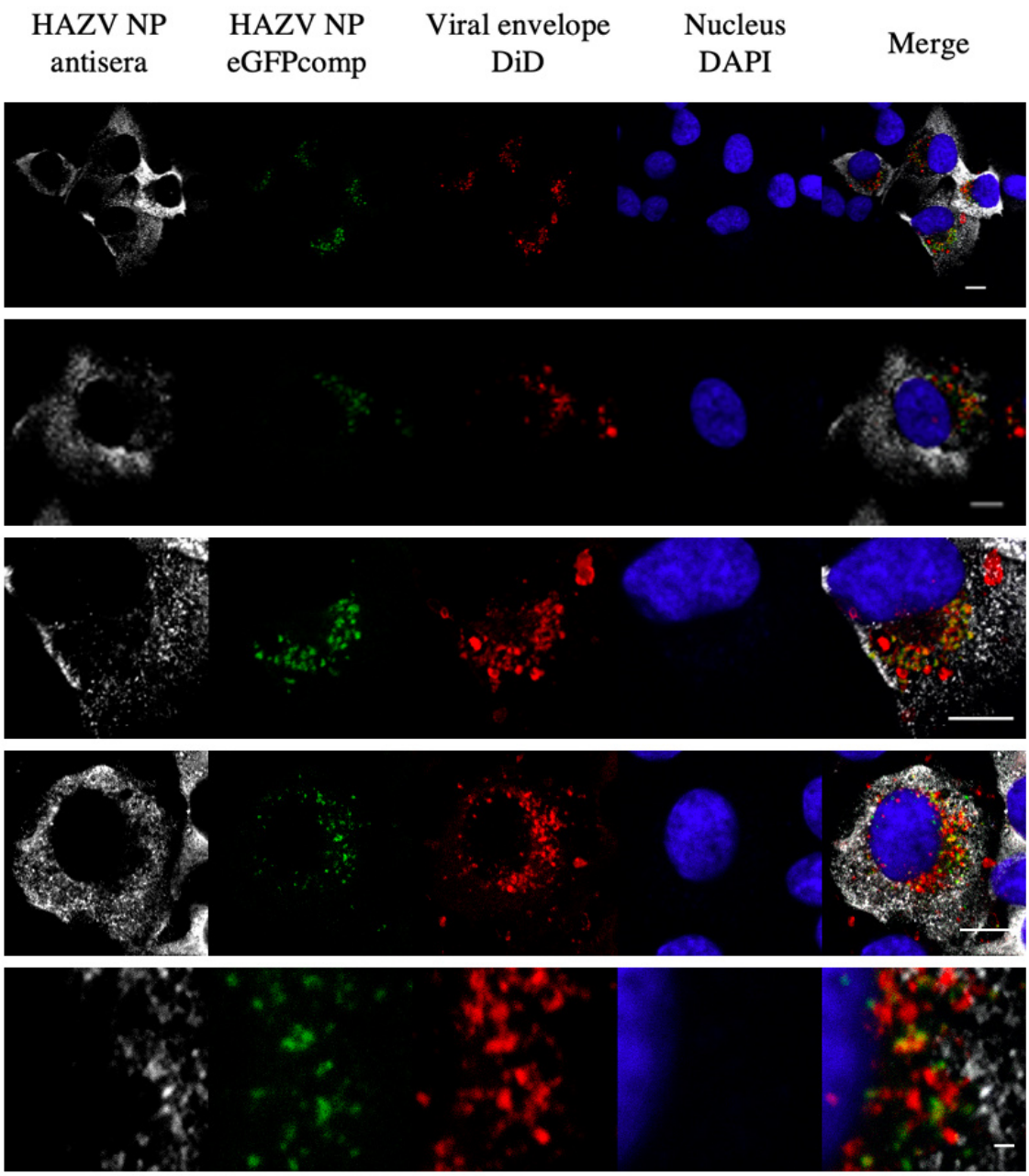


Figure 6.19: Sub-cellular localisation of viral components at 24 hpi. Confocal microscopy images of BSR-T7 cells infected with Vybrant DiD labelled rHAZV NP-eGFPcomp. Cells were fixed at 24 hpi and stained with HAZV NP antisera (white) and DAPI (blue). Newly synthesised HAZV NP labelled with antisera presented a cytoplasmic distribution and excluded the perinuclear sites where eGFP and Vybrant DiD signals were present. Scale bar: 10 µm.

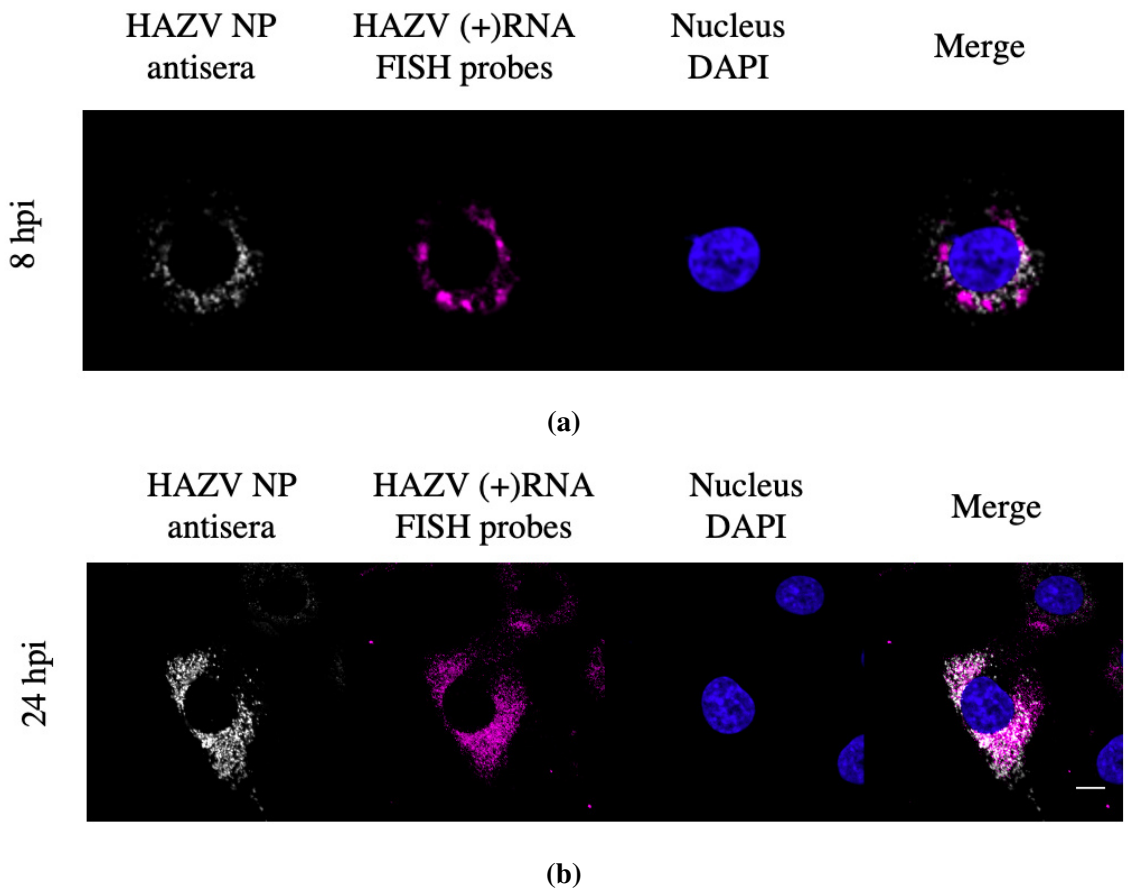


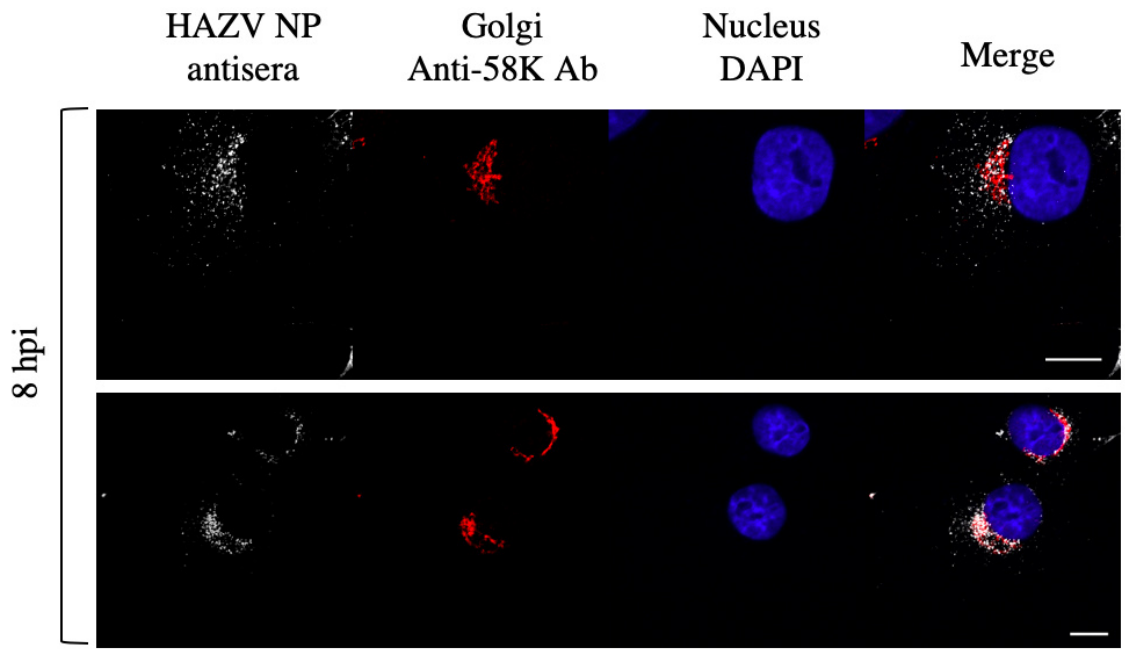
Figure 6.20: Sub-cellular localisation of HAZV NP and (+) sense HAZV RNA at 8 and 24 hpi. **(a)** Confocal microscopy images of BSR-T7 cells infected with WT HAZV. Cells were fixed at 8 hpi and stained with HAZV NP antisera (white), (+) sense HAZV RNA FISH probes (magenta) and DAPI (blue). The (+) sense HAZV RNA signal was localised in close proximity to HAZV NP in a perinuclear region. **(b)** Confocal microscopy images of BSR-T7 cells infected with WT HAZV. Cells were fixed at 24 hpi and stained with HAZV NP antisera (white), (+) sense HAZV RNA FISH probes (magenta) and DAPI (blue). The (+) sense HAZV RNA and HAZV NP sigals presented a cytoplasmic distribution that excluded a perinuclear region. Scale bar: 10 µm.

6.2.2.3 The site for input viral components trafficking is associated to the Golgi complex

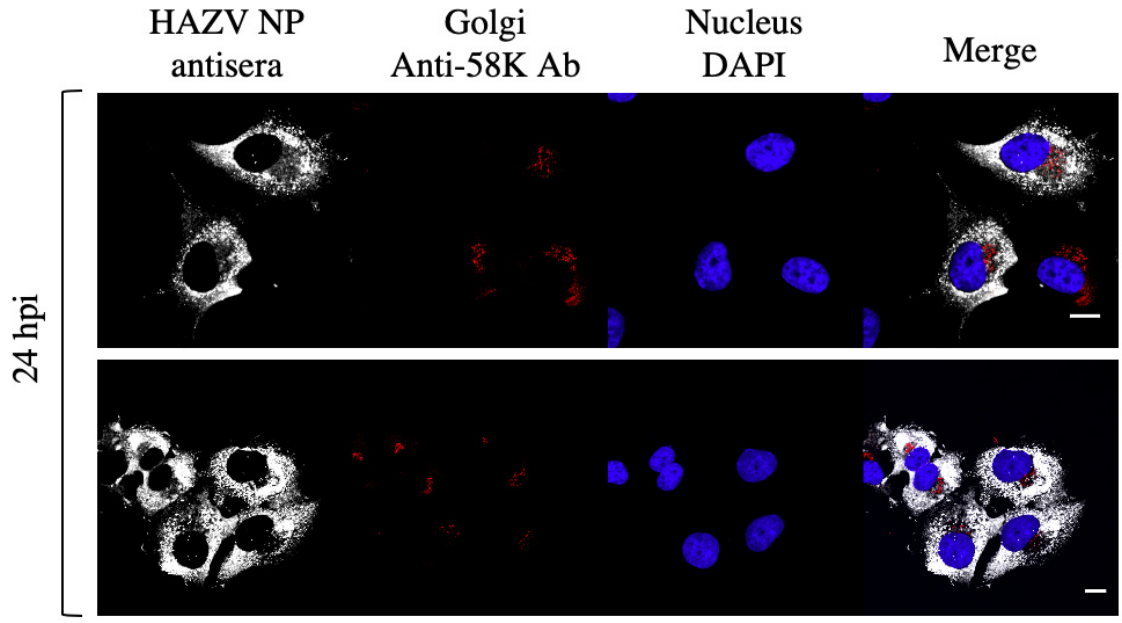
The next aim was to more precisely identify the intracellular compartment within the perinuclear regions where HAZV RNPs were initially localised, and where initial RNA synthesis occurred. As previous studies determined that other members of the *Bunyavirales* order establish their replication factories in and around the Golgi complex [114, 269], we first tested whether Golgi and HAZV NP co-localised using immunofluorescence and confocal imaging analyses.

BSR-T7 cells were infected with WT HAZV at an MOI of 1, fixed and immunostained with HAZV NP antisera and an anti-58K Golgi protein antibody that binds to the Golgi-associated formiminotransferase cyclodeaminase (FTCD) enzyme. At 8 hpi, the HAZV NP staining, previously shown to colocalise with HAZV (+) sense RNA and input viral components at this stage of the virus multiplication cycle, showed close co-localisation with the Golgi marker (figures 6.21a), suggesting that at early stages of HAZV infection, RNPs are trafficked to or near the Golgi complex, where (+) sense RNA synthesis takes place.

At the later time-point (24 hpi), while NP staining had expanded to occupy much of the cytoplasm, the Golgi staining was still restricted to perinuclear regions, as at 8 hpi. Thus, expansion of the HAZV NP immunostaining (figures 6.21b) was not accompanied with expansion of the Golgi. Consequently, we conclude that the site for input viral components trafficking is associated with the Golgi complex, from which newly synthesised HAZV NP and (+) sense HAZV RNA are exported at later stages of the virus multiplication cycle.



(a)



(b)

Figure 6.21: Sub-cellular localisation of HAZV NP and 58K Golgi marker at 8 and 24 hpi. **(a)** Confocal microscopy images of BSR-T7 cells infected with WT HAZV. Cells were fixed at 8 hpi and stained with HAZV NP antisera (white), Golgi marker (red) and DAPI (blue). The Golgi complex signal was localised in close proximity to HAZV NP in a perinuclear region. **(b)** Confocal microscopy images of BSR-T7 cells infected with WT HAZV. Cells were fixed at 24 hpi and stained with HAZV NP antisera (white), Golgi marker (red) and DAPI (blue). HAZV NP staining presented a cytoplasmic distribution that excluded a perinuclear region positively stained for the Golgi complex. Scale bar: 10 µm.

6.3 Chapter summary and discussion

In this chapter, the use of two different clonable tags for the production of genetically engineered rHAZV were described, and the first evidence for the recovery of a genetically engineered infectious bunyavirus containing small clonable tags in the most abundant structural protein of the virions, the NP, was presented. The recovery of infectious viruses with an artificially engineered tag in the NP would likely not be possible for other bunyaviruses due to the structural properties of their NPs, which contain two N- and C- terminal arms critical for NP-NP interactions [270]. Nairoviral NPs possess a distinct structure, with their C-terminal residues not being involved in the head-tail interaction between monomers nor the predicted RNA binding site [151, 145, 224, 141], plausibly explaining the maintenance of the structural activities of HAZV NP and the viability of the recovered infectious virus presented here, despite the presence of small artificially engineered C-terminal tags.

Both engineered recombinant viruses presented growth kinetics similar to WT virus, and the incorporation of the tags in vRNA was checked by RT-PCR and sequencing. The presence of the 6xHis tag could also be detected in the NP by western blot analysis using an anti-His tag antibody.

Purification of native viral RNPs was achieved using 6xHis-NP rHAZV and affinity chromatography, followed by structural characterisation using TEM. Two different purification procedures were tested: from purified virions and from infected cell lysate samples. Purification of RNPs from virions involved an ultracentrifugation step for virus purification, followed by freeze-thaw and detergent treatment to release the RNPs from the viral particles. These treatments may have interfered with the assembly of the RNPs, which showed an unorganised structure with high background of non-specific material when observed under TEM.

RNPs purified from cell lysates showed a clearer structure under TEM. RNP filaments of different sizes could be distinguished, most of them presenting a closed circularised arrangement, as expected by the complementarity of the terminal regions of the RNA segments and by analogy with other bunyaviruses. Linear filaments were rare, representing less than one percent of the imaged RNPs. The length and width of the filaments was analysed and 3 different populations could be distinguished according to their size, with lengths (in nm) proportional to the length (in kbp) of each of the RNA segments of the virus (S, M and L). The width of these filaments was similar to the width of previously described recombinant CCHFV NP pentamers, suggesting native nairoviral RNPs present a helical structure based on pentameric turns. One of the main functions

of the NPs of ssRNA viruses in the assembly of RNPs is the compaction of the RNA. The compaction of these RNPs was similar for the three different segments, presenting a packaging ratio (100 nm/kbp) three times more condensed than linear ssRNA (300 nm/kbp) [264].

A high resolution structure of the arrangement of these filaments could not be obtained due to the heterogeneity and high flexibility they presented. In order to obtain a high resolution structural model, sample preparation must be optimised to achieve a purified sample with higher concentration of these filaments suitable for cryo-TEM analysis. The ability of 6xHis-tagged NP to support CCHFV-specific replication was confirmed using a mini-genome system, suggesting a similar procedure could potentially be used to purify and study CCHFV native RNPs.

Studies on the intracellular trafficking of nairoviral RNPs and other viral components are currently lacking by the absence of tools enabling their visualisation during infection in single cells. We describe a split-eGFP-based recombinant virus that enables direct observation of the sub-cellular localisation of input HAZV NP and its discrimination from newly synthesised HAZV NP when combined with immunofluorescence staining. Using confocal imaging of this system in combination with FISH techniques, our results reveal new insights into the localisation of nairovirus components during different stages of the infectious cycle.

Transient expression of eGFP-β1-10 and subsequent infection with infectious rHAZV-eGFP-β11 allowed the real-time localisation of newly synthesised NP in the context of a viral infection in live cells. At early times of infection the fluorescent signal associated to HAZV NP presented a perinuclear localisation, further extended through the cytoplasm as the infection progressed, concluding with apoptosis of the infected cell as has previously been described [152]. We demonstrate that the eGFP signal obtained is associated to HAZV NP, validating the split-GFP system for fluorescent tagging of the NP in infectious rHAZV.

Labelling of the resulting rHAZV-eGFPcomp, containing the fluorescent eGFPcomp moiety attached to viral RNPs, with a marker for the viral membrane, alongside immunofluorescence and FISH techniques, permitted discrimination between input viral components (HAZV NP labelled with eGFPcomp, viral membrane labelled with Vybrant DiD) and newly synthesised viral components (HAZV NP labelled with HAZV NP antisera and newly synthesised (+) sense HAZV RNA labelled with FISH probes).

Early during infection, input HAZV NP showed a perinuclear localisation, accompanied by the viral membrane. (+) sense HAZV RNA is present in this region

at early timepoints, suggesting this sub-cellular compartment is the site for the first RNA synthesis activities of the virus. This perinuclear localisation was closely associated to the Golgi complex, potentially being the site where HAZV establishes its viral factories upon infection, in concordance with previous studies for other bunyaviruses [114, 269]. Of note, the co-localisation of the HAZV NP and the Golgi complex was not completely perfect, suggesting that this perinuclear region could also be associated to recycling endosomes or other juxtanuclear cellular compartments.

At later stages, the localisation of input HAZV NP and viral membrane remained perinuclear, suggesting input virus remains intact for the entire replication cycle. However, newly synthesised HAZV NP and (+) sense HAZV RNA presented a diffuse cytoplasmic localisation. In concordance with these findings, cytoplasmic co-localisation of CCHFV (+) sense RNA with the NP has previously been described [113].

Two possible scenarios arise here, HAZV NP and (+) sense HAZV RNA could be synthesised in the viral factories in the Golgi complex early during infection and sequentially be transported to the cytoplasmic area, or they could be directly synthesised in the cytoplasmic area during late stages of the viral cycle. The co-localisation of NP and (+) sense HAZV RNA during the entire cycle suggests that HAZV NP could be associated to (+) sense HAZV RNA in actively transcribing RNPs. This could potentially be explained by an initial phase where HAZV NP is involved in primary transcription in the Golgi site, where the viral factories are established. At later stages, the cytoplasmic HAZV NP could be involved in late transcription activities synthesising mRNA and consequently producing additional cytoplasmic HAZV NP.

The distinction between input and newly synthesised HAZV NP is a unique feature provided by this system, and would not be possible using standard immunofluorescence techniques. This system has further potential as a tool for direct visualisation of infectious virions during multiple stages of the viral entry process, including virus binding, uptake into endosomes, fusion, RNP release, RNP trafficking and establishment of replication factories using live cell imaging. Additional techniques utilising fluorescent molecules to track proteins are also plausible with this system, such as correlative light electron microscopy (CLEM), which could be used for morphological studies of HAZV viral factories. From a broader point of view, our study opens up the possibility of inclusion of many other small clonable tags for further studies of the nairoviral cycle, including other fluorescent tags (e.g. SNAP, CLIP), affinity tags for purification and co-immunoprecipitation studies or metal-binding tags for EM structural analyses.

In summary, the results presented in this chapter meet objective 4 of this project

(see project aims in section 1.4). The study of nairoviral RNPs and viral factories was described using two different clonable tags in the NP of rHAZV.

Chapter 7

Concluding Remarks

The main goal of this project was to contribute to the preparedness for potential outbreak scenarios with the development of diagnostic tools and the performance of structural and functional studies of emerging zoonotic viruses, as part of the MSCA-ITN HONOURs network. Four different objectives, previously described in section 1.4, were established for this purpose and have been addressed and fulfilled during this project, as briefly discussed below.

Objective 1: Expression and purification of recombinant CCHFV and HAZV NPs, and production of CCHFV NP-specific binders.

Nairovirus NPs are the most abundant proteins in virions, and thus represent optimal candidates as diagnostic targets. They perform essential structural roles in virus replication, increasing the relevance of their structural and functional studies. Large quantities of highly pure CCHFV and HAZV NPs were successfully achieved using bacterial expression systems and chromatographic techniques. CCHFV NP was used for the production of highly-specific binders with the aim of using these reagents for the development of diagnostic tools. Two different approaches were pursued for the production of high-affinity binders: production of CCHFV NP-specific polyclonal antibodies by a commercial company, and selection of CCHFV NP-specific Affimer reagents by the BioScreening Technology Group (BSTG, University of Leeds).

CCHFV NP-specific polyclonal antibodies were produced and initially validated using western blot analyses. Polyclonal antibodies recognise different epitopes in their target molecule, and thus tend to show cross-reactivity with analogous proteins. Likely, cross-reactivity of CCHFV NP-specific polyclonal antibodies was observed against HAZV NP, which shares 60% sequence identity with CCHFV NP.

Affimers are non-antibody binding molecules that can be screened using large ′phage libraries that allow the rapid selection of specific binders by ′phage display. Here, the selection and production of an Affimer reagent, Affimer-NP, that presents a distinguished binding capacity for CCHFV NP were described.

Objective 2: Characterisation and validation of CCHFV NP-specific Affimer reagents using biophysical and biochemical techniques.

Different physicochemical techniques were used for the characterisation of the interaction between CCHFV NP and Affimer-NP, including pull down AP assays, SPR, CD and X-ray crystallography. Unlike CCHFV NP-specific polyclonal antibodies, Affimer-NP specifically recognised CCHFV NP in AP assays and SPR analyses, showing no cross-reactivity to HAZV NP. Affimers are synthesised as monoclonal binders and thus recognise a single epitope, which enhances their characteristic specificity.

The affinity and kinetics of the CCHFV NP/Affimer-NP interaction were further studied using SPR. The binding interaction between CCHFV NP and Affimer-NP was characterised by a high affinity and fast-on/slow-off kinetic rates. These characteristics are ideal for a capture molecule in diagnostic tests. A high affinity ensures tight binding to the analyte of interest and, alongside a high specificity, helps to avoid unspecific binding to other molecules present in the sample that could lead to false positive results. Additionally, a quick binding interaction (fast-on rate) durable in time (slow-off rate) allows the quick detection of the analyte of interest bound to the corresponding capture molecule within a wide (in terms of molecular interaction time-scales) time window.

The secondary structure of the nairoviral NPs, Affimer-NP, and the CCHFV NP/Affimer-NP complex was analysed using CD analyses. These analyses confirmed the proper folding of the molecules, and suggested a stabilising interaction between CCHFV NP and Affimer-NP. The interference of Affimer-NP in the RNA binding function of CCHFV NP was determined using FA experiments, and its inhibition of CCHFV-specific gene expression was assessed in mammalian cells using a mini-genome system. The CCHFV mini-genome established for this purpose uses an eGFP molecule as reporter signal of CCHFV-specific replication, allowing the easy, quick and real-time characterisation of the system by live cell fluorescence imaging.

Solution of the crystallographic structure of the CCHFV NP/Affimer-NP complex at 2.84 Å resolution revealed the structural basis for the interference of Affimer-NP in the RNA binding function of CCHFV NP and CCHFV-specific replication, and explained the specificity of this small protein for CCHFV NP and lack of cross-reactivity with HAZV NP. The binding interface of Affimer-NP and CCHFV NP involves residues of the NP

previously described to be part of the RNA binding interaction and NP-NP oligomeric interface. The interference of Affimer-NP in its RNA binding and oligomerisation functions further corroborates the importance of this region of CCHFV NP in what are two of its essential functions.

Currently, few treatment options exist for CCHF, and most are palliative. Therefore, there is an urgent need for novel therapeutics against CCHFV. The interference that Affimer-NP shows in the RNA binding function of CCHFV NP and CCHFV-specific replication suggests its potential translation to antiviral therapeutics. The small size of Affimers, rapid tissue penetration and high stability and solubility make them attractive candidates as therapeutic molecules. Intracellular delivery of Affimer-NP would need to be achieved for its efficacy as inhibitor of viral replication. *In vivo* intracellular delivery of Affimers has not been described to date but possible approaches that could be explored for *in vivo* targeting of small proteins into cells include delivery of protein DNA through injection or molecular engineering of proteins with cell-penetrating peptides.

Objective 3: Development of diagnostic assays using CCHFV NP-specific antibodies and Affimers.

The *in vitro* application of Affimer-NP, combined with CCHFV NP-specific antibodies, for the development of ELISA and LFAs was validated, presenting the first published prototype point-of-care test able to detect recombinant CCHFV NP in spiked human and animal sera. This objective was carried out as a collaboration with the company INGENASA (Eurofins Ingenasa, Madrid, Spain), which provided the expertise and equipment required for the development of ELISA and LFAs.

The *in vitro* nature of Affimer design, selection, amplification and purification presents several advantages compared to antibody production. *In vitro* selection bypasses the experimental use of animals, in a process that typically takes only a few weeks, whereas the development of antibodies is a time-consuming process that normally takes several months. ′Phage display library screening is substantially less expensive, consumes less material than *in vivo* development strategies and thus enables low cost production with high yields and reproducible batch-to-batch preparations [165]. The high thermal stability of Affimer-NP also presents a particular advantage in the context of LFA development, which involves a baking step at 45°C for the immobilisation of Affimer-NP into the LFA membrane. Additionally, this feature of Affimer-NP suggests that the diagnostic tests will be robust in the field. This is a remarkable advantage of Affimer reagents especially relevant in low-resource settings where refrigerated storage of diagnostic tests may not be possible.

Rapid diagnosis of CCHFV is a critical aspect of disease management, but PCR equipment and expert personnel are commonly limited in low-income countries and remote areas. The LFA tests presented here require minimal specimen processing, have time to results of minutes and are suitable for field testing and low-infrastructure settings. Although these diagnostic tests have not been validated with real patient samples, the results described validate the use of Affimers for the development of ELISA and LFA diagnostics and present a possible starting point for the future development of diagnostic assays for the detection of CCHFV NP.

Objective 4: Structural and functional analyses of nairoviral NPs using a reverse genetics system.

A novel approach was designed for the incorporation of clonable tags in the NP of infectious rHAZV using a previously described HAZV rescue system. The introduction of a 6xHis tag in the C-terminal region of HAZV NP allowed the affinity purification of 6xHis tagged RNPs for their structural analysis using TEM. The corresponding micrographs represent the first evidence of assembled native nairoviral RNPs from which filaments corresponding to each viral RNA segment can be distinguished.

The width of the RNP filaments suggests a higher order helical structure likely based on pentameric turns. Reconstitution of RNPs with recombinant nucleoprotein in the absence of the viral polymerase results in the association of the protein to the RNA in the form of 'beads-on-a-string' [271]. The findings presented here suggest that the helical structure of the HAZV RNP complexes is likely assembled with help from the viral polymerase, and thus, the purification of RNP complexes from replication-competent virus is key to determine the native structure of these filaments. Additionally, the higher order structure observed for the RNPs supports previous presumptions that the oligomerisation of nairoviral NPs plays a role in their RNA binding function [149].

Finally, a split-eGFP-based recombinant virus that enables direct observation of the sub-cellular localisation of HAZV NP was designed. Using confocal imaging of this system in combination with immunofluorescence and FISH techniques, new insights into the localisation of nairovirus components during different stages of the infectious cycle were revealed. Two different phases of the nairoviral cycle could be distinguished using this system. During the early phase of the infectious cycle all viral components (input HAZV NP, viral membrane, newly synthesised HAZV NP and (+) sense HAZV RNA) are localised in the Golgi complex. At later stages of the cycle, input virus components (input HAZV NP, viral membrane) remain perinuclear co-localising with the Golgi complex, but newly synthesised components (newly synthesised HAZV NP and (+) sense HAZV RNA)

redistribute to a diffuse cytoplasmic localisation.

These results present the first evidence of the Golgi complex as the site for input virus localisation and potential cellular compartment for the establishment of nairoviral factories. The perinuclear localisation of nairoviral NPs at early stages of the infection and their co-localisation with (+) sense viral RNA has been previously described for CCHFV [113]. Here, we present the first study on the intracellular localisation of HAZV NP and (+) sense HAZV RNA during different stages of the viral cycle. Using the split-eGFP-tagged NP rHAZV system, the intracellular localisation of input and newly synthesised HAZV NP could be distinguished, which would have not been possible using standard immunofluorescence techniques. This system has further potential as a tool for direct visualisation of infectious virions during multiple stages of the viral cycle, and thus represents a valuable tool for the study of nairovirus biology.

In conclusion, the results presented in this project represent a step forward on the development of therapeutics and diagnostics for the highly pathogenic CCHFV, and the understanding of the native assembly, function and cellular localisation of nairoviral NPs. The tools and workflow used for these discoveries could be translated to other emerging viruses, for which therapeutic and diagnostic tools may be needed, and help to fill the basic biology knowledge gaps concomitant to any novel pathogen, contributing to the preparedness for potential future outbreak scenarios.

Every real story is a never ending story.
Michael Ende, The Neverending Story

Bibliography

[1] A. Lustig and A. J. Levine, "One hundred years of virology.," *Journal of virology*, vol. 66, no. 8, p. 4629, 1992.

[2] L. P. Villarreal *et al.*, *Viruses and the evolution of life.* ASM press, 2005.

[3] M. Woolhouse, F. Scott, Z. Hudson, R. Howey, and M. Chase-Topping, "Human viruses: discovery and emergence," *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 367, no. 1604, pp. 2864–2871, 2012.

[4] K. J. Olival, P. R. Hosseini, C. Zambrana-Torrelio, N. Ross, T. L. Bogich, and P. Daszak, "Host and viral traits predict zoonotic spillover from mammals," *Nature*, vol. 546, no. 7660, p. 646, 2017.

[5] J. M. Hassell, M. Begon, M. J. Ward, and E. M. Fèvre, "Urbanization and disease emergence: dynamics at the wildlife–livestock–human interface," *Trends in ecology & evolution*, vol. 32, no. 1, pp. 55–67, 2017.

[6] I. M. Fund, "World economic outlook, april 2020: The great lockdown."

[7] L. van der Hoek, E. Verschoor, M. Beer, D. Höper, K. Wernike, M. Van Ranst, J. Matthijnssens, P. Maes, P. Sastre, P. Rueda, *et al.*, "Host switching pathogens, infectious outbreaks and zoonosis: A marie skłodowska-curie innovative training network (honours)," *Virus research*, vol. 257, pp. 120–124, 2018.

[8] E. H. Loh, C. Zambrana-Torrelio, K. J. Olival, T. L. Bogich, C. K. Johnson, J. A. Mazet, W. Karesh, and P. Daszak, "Targeting transmission pathways for emerging zoonotic disease surveillance and control," *Vector-Borne and Zoonotic Diseases*, vol. 15, no. 7, pp. 432–437, 2015.

[9] A. E. Paniz-Mondolfi, A. J. Rodriguez-Morales, G. Blohm, M. Marquez, and W. E. Villamil-Gomez, "Chikdenmazika syndrome: the challenge of diagnosing arboviral infections in the midst of concurrent epidemics," 2016.

[10] L. Braack, A. P. G. de Almeida, A. J. Cornel, R. Swanepoel, and C. De Jager, "Mosquito-borne arboviruses of african origin: review of key viruses and vectors," *Parasites & vectors*, vol. 11, no. 1, p. 29, 2018.

[11] M. Venter, "Assessing the zoonotic potential of arboviruses of african origin," *Current opinion in virology*, vol. 28, pp. 74–84, 2018.

[12] N. Cleton, M. Koopmans, J. Reimerink, G.-J. Godeke, and C. Reusken, "Come fly with me: review of clinically important arboviruses for global travelers," *Journal of Clinical Virology*, vol. 55, no. 3, pp. 191–203, 2012.

[13] N. Grashchenkov, "Investigation of etiology, pathogenesis, and clinical symptomatology of crimean hemorrhagic fever," *Reports on the 1944 Scientific Investigation of the Institute of Neurology*, pp. 100–107, 1945.

[14] H. Hoogstraal, "The epidemiology of tick-borne crimean-congo hemorrhagic fever in asia, europe, and africa," *Journal of medical entomology*, vol. 15, no. 4, pp. 307–417, 1979.

[15] D. A. Bente, N. L. Forrester, D. M. Watts, A. J. McAuley, C. A. Whitehouse, and M. Bray, "Crimean-congo hemorrhagic fever: history, epidemiology, pathogenesis, clinical syndrome and genetic diversity," *Antiviral research*, vol. 100, no. 1, pp. 159–189, 2013.

[16] C. A. Whitehouse, "Crimean–congo hemorrhagic fever," *Antiviral research*, vol. 64, no. 3, pp. 145–160, 2004.

[17] A. Butenko, M. Chumakov, V. Rubin, and D. Stolbov, "Isolation and investigation of astrakhan strain ('drozdov') of crimean hemorrhagic fever virus and data on serodiagnosis of this infection," tech. rep., NAVAL MEDICAL RESEARCH UNIT NO 3 FPO NEW YORK 09527, 1968.

[18] J. Casals, "Antigenic similarity between the virus causing crimean hemorrhagic fever and congo virus," *Proceedings of the Society for Experimental Biology and Medicine*, vol. 131, no. 1, pp. 233–236, 1969.

[19] M. Chumakov, S. Smirnova, E. Tkachenko, *et al.*, "Relationship between strains of crimean haemorrhagic fever and congo viruses.," *Acta virologica*, vol. 14, no. 1, pp. 82–5, 1970.

[20] J. C. Morrill, "Crimean-congo hemorrhagic fever: A global perspective," *Vector-Borne and Zoonotic Diseases*, vol. 8, no. 1, pp. 123–124, 2008.

[21] S. B. Appannanavar and B. Mishra, "An update on crimean congo hemorrhagic fever," *Journal of global infectious diseases*, vol. 3, no. 3, p. 285, 2011.

[22] A. Negredo, F. de la Calle-Prieto, E. Palencia-Herrejón, M. Mora-Rillo, J. Astray-Mochales, M. P. Sánchez-Seco, E. Bermejo Lopez, J. Menárguez, A. Fernández-Cruz, B. Sánchez-Artola, *et al.*, "Autochthonous crimean–congo hemorrhagic fever in spain," *New England Journal of Medicine*, vol. 377, no. 2, pp. 154–161, 2017.

[23] F. Begum, C. Wisseman Jr, and J. Casals, "Tick-borne viruses of west pakistan: Ii. hazara virus, a new agent isolated from ixodes redikorzevi ticks from the kaghan valley, w. pakistan," *American journal of epidemiology*, vol. 92, no. 3, pp. 192–194, 1970.

[24] F. Begum, C. Wisseman Jr, and J. Casals, "Tick-borne viruses of west pakistan: Iv. viruses similar to, or identical with, crimean hemorrhagic fever (congo-semunya), wad medani and pak argas 461 isolated from ticks of the changa manga forest, lahore district, and of hunza, gilgit agency, w. pakistan," *American journal of epidemiology*, vol. 92, no. 3, pp. 197–202, 1970.

[25] F. Begum, C. Wisseman Jr, and R. Traub, "Tick-borne viruses of west pakistan: I. isolation and general characteristics," *American journal of epidemiology*, vol. 92, no. 3, pp. 180–191, 1970.

[26] S. Shimada, K. Aoki, T. Nabeshima, Y. Fuxun, Y. Kurosaki, K. Shiogama, T. Onouchi, M. Sakaguchi, T. Fuchigami, H. Ono, *et al.*, "Tofla virus: A newly identified nairovirus of the crimean-congo hemorrhagic fever group isolated from ticks in japan," *Scientific reports*, vol. 6, p. 20213, 2016.

[27] A. Abudurexiti, S. Adkins, D. Alioto, S. V. Alkhovsky, T. Avšič-Županc, M. J. Ballinger, D. A. Bente, M. Beer, É. Bergeron, C. D. Blair, *et al.*, "Taxonomy of the order bunyavirales: update 2019," *Archives of virology*, vol. 164, no. 7, pp. 1949–1965, 2019.

[28] A. Albornoz, A. B. Hoffmann, P.-Y. Lozach, and N. D. Tischler, "Early bunyavirus-host cell interactions," *Viruses*, vol. 8, no. 5, p. 143, 2016.

[29] M. Baron and B. Holzer, "Nairobi sheep disease virus/ganjam virus," *Rev Sci Tech*, vol. 34, no. 2, pp. 411–7, 2015.

[30] M. B. Crabtree, R. Sang, and B. R. Miller, "Kupe virus, a new virus in the family bunyaviridae, genus nairovirus, kenya," *Emerging infectious diseases*, vol. 15, no. 2, p. 147, 2009.

[31] D. Watts, T. Ksiazek, K. Linthicum, H. Hoogstraal, and T. Monath, "The arboviruses: epidemiology and ecology," *Boca Raton*, vol. 177, p. 222, 1988.

[32] A. Gargili, A. Estrada-Pena, J. R. Spengler, A. Lukashev, P. A. Nuttall, and D. A. Bente, "The role of ticks in the maintenance and transmission of crimean-congo hemorrhagic fever virus: A review of published field and laboratory studies," *Antiviral research*, vol. 144, pp. 93–119, 2017.

[33] J. R. Spengler, A. Estrada-Peña, A. R. Garrison, C. Schmaljohn, C. F. Spiropoulou, É. Bergeron, and D. A. Bente, "A chronological review of experimental infection studies of the role of wild animals and livestock in the maintenance and transmission of crimean-congo hemorrhagic fever virus," *Antiviral research*, vol. 135, pp. 31–47, 2016.

[34] A. Papa, K. Tsergouli, K. Tsioka, and A. Mirazimi, "Crimean-congo hemorrhagic fever: tick-host-virus interactions," *Frontiers in Cellular and Infection Microbiology*, vol. 7, p. 213, 2017.

[35] J. R. Spengler, D. A. Bente, M. Bray, F. Burt, R. Hewson, G. Korukluoglu, A. Mirazimi, F. Weber, and A. Papa, "Second international conference on crimean-congo hemorrhagic fever," *Antiviral research*, vol. 150, pp. 137–147, 2018.

[36] J. P. Messina, D. M. Pigott, N. Golding, K. A. Duda, J. S. Brownstein, D. J. Weiss, H. Gibson, T. P. Robinson, M. Gilbert, G. William Wint, *et al.*, "The global distribution of crimean-congo hemorrhagic fever," *Transactions of the Royal Society of Tropical Medicine and Hygiene*, vol. 109, no. 8, pp. 503–513, 2015.

[37] Y. Ince, C. Yasa, M. Metin, M. Sonmez, E. Meram, B. Benkli, and O. Ergonul, "Crimean-congo hemorrhagic fever infections reported by promed," *International Journal of Infectious Diseases*, vol. 26, pp. 44–46, 2014.

[38] J. R. Spengler, É. Bergeron, and C. F. Spiropoulou, "Crimean-congo hemorrhagic fever and expansion from endemic regions," *Current opinion in virology*, vol. 34, pp. 70–78, 2019.

[39] B. Wahid, S. Altaf, N. Naeem, N. Ilyas, and M. Idrees, "Scoping review of crimean-congo hemorrhagic fever (cchf) literature and implications of future research," *Journal of the College of Physicians and Surgeons Pakistan*, vol. 29, no. 6, pp. 563–573, 2019.

[40] D. W. Hawman and H. Feldmann, "Recent advances in understanding crimean–congo hemorrhagic fever virus," *F1000Research*, vol. 7, 2018.

[41] S. Shayan, M. Bokaean, M. R. Shahrivar, and S. Chinikar, "Crimean-congo hemorrhagic fever," *Laboratory medicine*, vol. 46, no. 3, pp. 180–189, 2015.

[42] T. Schwarz, H. Nsanze, and A. Ameen, "Clinical features of crimean-congo haemorrhagic fever in the united arab emirates," *Infection*, vol. 25, no. 6, pp. 364–367, 1997.

[43] Y. Yu-Chen, K. Ling-Xiong, L. Ling, Z. Yu-Qin, L. Feng, C. Bao-Jian, and G. Shou-Yi, "Characteristics of crimean-congo hemorrhagic fever virus (xinjiang strain) in china," *The American journal of tropical medicine and hygiene*, vol. 34, no. 6, pp. 1179–1182, 1985.

[44] D. Duh, A. Saksida, M. Petrovec, S. Ahmeti, I. Dedushaj, M. Panning, C. Drosten, and T. Avšič-Županc, "Viral load as predictor of crimean-congo hemorrhagic fever outcome," *Emerging infectious diseases*, vol. 13, no. 11, p. 1769, 2007.

[45] A. Papa, C. Drosten, S. Bino, E. Papadimitriou, M. Panning, E. Velo, M. Kota, A. Harxhi, and A. Antoniadis, "Viral load and crimean-congo hemorrhagic fever," *Emerging infectious diseases*, vol. 13, no. 5, p. 805, 2007.

[46] M. A. Çevik, A. Erbay, H. Bodur, S. S. Eren, E. Akinci, K. Şener, P. Öngürü, and A. Kubar, "Viral load as a predictor of outcome in crimean-congo hemorrhagic fever," *Clinical Infectious Diseases*, vol. 45, no. 7, pp. e96–e100, 2007.

[47] E. Akıncı, H. Bodur, and H. Leblebicioglu, "Pathogenesis of crimean-congo hemorrhagic fever," *Vector-Borne and Zoonotic Diseases*, vol. 13, no. 7, pp. 429–437, 2013.

[48] A. Connolly-Andersen, "Pathogenesis of an emerging pathogen-crimean-congo hemorrhagic fever," *Thesis*, 2010.

[49] O. Ergonul, S. Tuncbilek, N. Baykam, A. Celikbas, and B. Dokuzoguz, "Evaluation of serum levels of interleukin (il)–6, il-10, and tumor necrosis factor–α in patients with crimean-congo hemorrhagic fever," *The Journal of infectious diseases*, vol. 193, no. 7, pp. 941–944, 2006.

[50] D. A. Bente, J. B. Alimonti, W.-J. Shieh, G. Camus, U. Ströher, S. Zaki, and S. M. Jones, "Pathogenesis and immune response of crimean-congo hemorrhagic fever

virus in a stat-1 knockout mouse model," *Journal of virology*, vol. 84, no. 21, pp. 11089–11100, 2010.

[51] N. Tasdelen Fisgin, T. Fisgin, E. Tanyel, L. Doganci, N. Tulek, N. Guler, and F. Duru, "Crimean-congo hemorrhagic fever: five patients with hemophagocytic syndrome," *American journal of hematology*, vol. 83, no. 1, pp. 73–76, 2008.

[52] M. Habjan, I. Andersson, J. Klingström, M. Schümann, A. Martin, P. Zimmermann, V. Wagner, A. Pichlmair, U. Schneider, E. Mühlberger, *et al.*, "Processing of genome 5 termini as a strategy of negative-strand rna viruses to avoid rig-i-dependent interferon induction," *PloS one*, vol. 3, no. 4, p. e2032, 2008.

[53] I. Andersson, H. Karlberg, M. Mousavi-Jazi, L. Martínez-Sobrido, F. Weber, and A. Mirazimi, "Crimean-congo hemorrhagic fever virus delays activation of the innate immune response," *Journal of medical virology*, vol. 80, no. 8, pp. 1397–1404, 2008.

[54] F. Weber and A. Mirazimi, "Interferon and cytokine responses to crimean congo hemorrhagic fever virus; an emerging and neglected viral zonoosis," *Cytokine & growth factor reviews*, vol. 19, no. 5-6, pp. 395–404, 2008.

[55] M. A. Darwish, H. Hoogstraal, T. J. Roberts, R. Ghazi, and T. Amer, "A sero-epidemiological survey for bunyaviridae and certain other arboviruses in pakistan," *Transactions of the Royal Society of Tropical Medicine and Hygiene*, vol. 77, no. 4, pp. 446–450, 1983.

[56] S. Smirnova, "A comparative study of the crimean hemorrhagic fever-congo group of viruses," *Archives of virology*, vol. 62, no. 2, pp. 137–143, 1979.

[57] A. Papa, "Diagnostic approaches for crimean-congo hemorrhagic fever virus," *Expert review of molecular diagnostics*, vol. 19, no. 6, pp. 531–536, 2019.

[58] B. Atkinson, J. Chamberlain, C. H. Logue, N. Cook, C. Bruce, S. D. Dowall, and R. Hewson, "Development of a real-time rt-pcr assay for the detection of crimean-congo hemorrhagic fever virus," *Vector-Borne and Zoonotic Diseases*, vol. 12, no. 9, pp. 786–793, 2012.

[59] A. Saksida, D. Duh, B. Wraber, I. Dedushaj, S. Ahmeti, and T. Avšič-Županc, "Interacting roles of immune mechanisms and viral load in the pathogenesis of crimean-congo hemorrhagic fever," *Clin. Vaccine Immunol.*, vol. 17, no. 7, pp. 1086–1093, 2010.

[60] D. T. Mourya, P. D. Yadav, A. M. Shete, Y. K. Gurav, C. G. Raut, R. S. Jadi, S. D. Pawar, S. T. Nichol, and A. C. Mishra, "Detection, isolation and confirmation of crimean-congo hemorrhagic fever virus in human, ticks and animals in ahmadabad, india, 2010–2011," *PLoS neglected tropical diseases*, vol. 6, no. 5, p. e1653, 2012.

[61] M. Saijo, Q. Tang, B. Shimayi, L. Han, Y. Zhang, M. Asiguma, D. Tianshu, A. Maeda, I. Kurane, and S. Morikawa, "Recombinant nucleoprotein-based serological diagnosis of crimean–congo hemorrhagic fever virus infections," *Journal of medical virology*, vol. 75, no. 2, pp. 295–299, 2005.

[62] P. Emmerich, A. Mika, R. von Possel, A. Rackow, Y. Liu, H. Schmitz, S. Günther, K. Sherifi, B. Halili, X. Jakupi, *et al.*, "Sensitive and specific detection of crimean-congo hemorrhagic fever virus (cchfv)—specific igm and igg antibodies in human sera using recombinant cchfv nucleoprotein as antigen in μ-capture and igg immune complex (ic) elisa tests," *PLoS neglected tropical diseases*, vol. 12, no. 3, p. e0006366, 2018.

[63] N. Shrivastava, A. Shrivastava, S. Ninawe, S. Sharma, S. I. Alam, S. Sharma, P. K. Dash, *et al.*, "Development of multispecies recombinant nucleoprotein based indirect elisa for high throughput screening of crimean-congo hemorrhagic fever virus specific antibodies," *Frontiers in microbiology*, vol. 10, p. 1822, 2019.

[64] I. Schuster, M. Mertens, B. Köllner, T. Korytář, M. Keller, B. Hammerschmidt, T. Müller, N. Tordo, P. Marianneau, C. Mroz, *et al.*, "A competitive elisa for species-independent detection of crimean-congo hemorrhagic fever virus specific antibodies," *Antiviral research*, vol. 134, pp. 161–166, 2016.

[65] M. A. Sas, L. Comtet, F. Donnet, M. Mertens, Z. Vatansever, N. Tordo, P. Pourquier, and M. H. Groschup, "A novel double-antigen sandwich elisa for the species-independent detection of crimean-congo hemorrhagic fever virus-specific antibodies," *Antiviral research*, vol. 151, pp. 24–26, 2018.

[66] F. Burt, R. Swanepoel, and L. Braack, "Enzyme-linked immunosorbent assays for the detection of antibody to crimean-congo haemorrhagic fever virus in the sera of livestock and wild vertebrates," *Epidemiology & Infection*, vol. 111, no. 3, pp. 547–558, 1993.

[67] A. Shepherd, R. Swanepoel, and P. Leman, "Antibody response in crimean-congo hemorrhagic fever," *Reviews of infectious diseases*, vol. 11, no. Supplement_4, pp. S801–S806, 1989.

[68] C. Escadafal, S. Ölschläger, T. Avšič-Županc, A. Papa, J. Vanhomwegen, R. Wölfel, A. Mirazimi, A. Teichmann, O. Donoso-Mantke, and M. Niedrig, "First international external quality assessment of molecular detection of crimean-congo hemorrhagic fever virus," *PLoS neglected tropical diseases*, vol. 6, no. 6, p. e1706, 2012.

[69] O. Ergonul, "Treatment of crimean-congo hemorrhagic fever," *Antiviral research*, vol. 78, no. 1, pp. 125–131, 2008.

[70] Ö. Ergönül, Ş. Keske, M. G. Çeldir, İ. A. Kara, N. Pshenichnaya, G. Abuova, L. Blumberg, and M. Gönen, "Systematic review and meta-analysis of postexposure prophylaxis for crimean-congo hemorrhagic fever virus among healthcare workers," *Emerging infectious diseases*, vol. 24, no. 9, p. 1642, 2018.

[71] L. Oestereich, T. Rieger, M. Neumann, C. Bernreuther, M. Lehmann, S. Krasemann, S. Wurr, P. Emmerich, X. de Lamballerie, S. Ölschläger, *et al.*, "Evaluation of antiviral efficacy of ribavirin, arbidol, and t-705 (favipiravir) in a mouse model for crimean-congo hemorrhagic fever," *PLoS neglected tropical diseases*, vol. 8, no. 5, p. e2804, 2014.

[72] A. Papa, E. Papadimitriou, and I. Christova, "The bulgarian vaccine crimean-congo haemorrhagic fever virus strain," *Scandinavian journal of infectious diseases*, vol. 43, no. 3, pp. 225–229, 2011.

[73] S. D. Dowall, M. W. Carroll, and R. Hewson, "Development of vaccines against crimean-congo haemorrhagic fever virus," *Vaccine*, vol. 35, no. 44, pp. 6015–6023, 2017.

[74] N. Canakoglu, E. Berber, S. Tonbak, M. Ertek, I. Sozdutmaz, M. Aktas, A. Kalkan, and A. Ozdarendeli, "Immunization of knock-out α/β interferon receptor mice against high lethal dose of crimean-congo hemorrhagic fever virus with a cell culture based vaccine," *PLoS neglected tropical diseases*, vol. 9, no. 3, p. e0003579, 2015.

[75] K. R. Buttigieg, S. D. Dowall, S. Findlay-Wilson, A. Miloszewska, E. Rayner, R. Hewson, and M. W. Carroll, "A novel vaccine against crimean-congo haemorrhagic fever protects 100% of animals against lethal challenge in a mouse model," *PloS one*, vol. 9, no. 3, p. e91516, 2014.

[76] T. Aligholipour Farzani, K. Földes, A. Hanifehnezhad, B. Yener Ilce, S. Bilge Dagalp, N. Amirzadeh Khiabani, K. Ergünay, F. Alkan, T. Karaoglu,

H. Bodur, *et al.*, "Bovine herpesvirus type 4 (bohv-4) vector delivering nucleocapsid protein of crimean-congo hemorrhagic fever virus induces comparable protective immunity against lethal challenge in ifnα/β/γr-/- mice models," *Viruses*, vol. 11, no. 3, p. 237, 2019.

[77] M. M. Sahib, "Rapid development of optimized recombinant adenoviral vaccines for biosafety level 4 viruses," *MSc Thesis, University of Manitoba, Canada*, 2010.

[78] M. Zivcec, D. Safronetz, D. P. Scott, S. Robertson, and H. Feldmann, "Nucleocapsid protein-based vaccine provides protection in mice against lethal crimean-congo hemorrhagic fever virus challenge," *PLoS neglected tropical diseases*, vol. 12, no. 7, p. e0006628, 2018.

[79] S. E. Rodriguez, R. W. Cross, K. A. Fenton, D. A. Bente, C. E. Mire, and T. W. Geisbert, "Vesicular stomatitis virus-based vaccine protects mice against crimean-congo hemorrhagic fever," *Scientific reports*, vol. 9, no. 1, p. 7755, 2019.

[80] T. Aligholipour Farzani, A. Hanifehnezhad, K. Földes, K. Ergünay, E. Yilmaz, H. Hashim Mohamed Ali, and A. Ozkul, "Co-delivery effect of cd24 on the immunogenicity and lethal challenge protection of a dna vector expressing nucleocapsid protein of crimean congo hemorrhagic fever virus," *Viruses*, vol. 11, no. 1, p. 75, 2019.

[81] A. R. Garrison, C. J. Shoemaker, J. W. Golden, C. J. Fitzpatrick, J. J. Suschak, M. J. Richards, C. V. Badger, C. M. Six, J. D. Martin, D. Hannaman, *et al.*, "A dna vaccine for crimean-congo hemorrhagic fever protects against disease and death in two lethal mouse models," *PLoS neglected tropical diseases*, vol. 11, no. 9, p. e0005908, 2017.

[82] K. Spik, A. Shurtleff, A. K. McElroy, M. C. Guttieri, J. W. Hooper, and C. Schmaljohn, "Immunogenicity of combination dna vaccines for rift valley fever virus, tick-borne encephalitis virus, hantaan virus, and crimean congo hemorrhagic fever virus," *Vaccine*, vol. 24, no. 21, pp. 4657–4666, 2006.

[83] T. Aligholipour Farzani, K. Földes, K. Ergünay, H. Gurdal, A. Bastug, and A. Ozkul, "Immunological analysis of a cchfv mrna vaccine candidate in mouse models," *Vaccines*, vol. 7, no. 3, p. 115, 2019.

[84] S. M. Ghiasi, A. Salmanian, S. Chinikar, and S. Zakeri, "Mice orally immunized with a transgenic plant expressing the glycoprotein of crimean-congo hemorrhagic fever virus," *Clin. Vaccine Immunol.*, vol. 18, no. 12, pp. 2031–2037, 2011.

[85] J. Kortekaas, R. P. Vloet, A. J. McAuley, X. Shen, B. J. Bosch, L. De Vries, R. J. Moormann, and D. A. Bente, "Crimean-congo hemorrhagic fever virus subunit vaccines induce high levels of neutralizing antibodies but no protection in stat1 knockout mice," *Vector-Borne and Zoonotic Diseases*, vol. 15, no. 12, pp. 759–764, 2015.

[86] S. Dowall, K. Buttigieg, S. Findlay-Wilson, E. Rayner, G. Pearson, A. Miloszewska, V. Graham, M. Carroll, and R. Hewson, "A crimean-congo hemorrhagic fever (cchf) viral vaccine expressing nucleoprotein is immunogenic but fails to confer protection against lethal disease," *Human vaccines & immunotherapeutics*, vol. 12, no. 2, pp. 519–527, 2016.

[87] J. Hinkula, S. Devignot, S. Åkerström, H. Karlberg, E. Wattrang, S. Bereczky, M. Mousavi-Jazi, C. Risinger, G. Lindegren, C. Vernersson, *et al.*, "Immunization with dna plasmids coding for crimean-congo hemorrhagic fever virus capsid and envelope proteins and/or virus-like particles induces protection and survival in challenged mice," *Journal of virology*, vol. 91, no. 10, pp. e02076–16, 2017.

[88] M. L. Martin, H. Lindsey-Regnery, D. R. Sasso, J. McCormick, and E. Palmer, "Distinction between bunyaviridae genera by surface structure and comparison with hantaan virus using negative stain electron microscopy," *Archives of virology*, vol. 86, no. 1-2, pp. 17–28, 1985.

[89] J. P. Clerx, J. Casals, and D. H. Bishop, "Structural characteristics of nairoviruses (genus nairovirus, bunyaviridae)," *Journal of General Virology*, vol. 55, no. 1, pp. 165–178, 1981.

[90] E. K. Punch, S. Hover, H. T. Blest, J. Fuller, R. Hewson, J. Fontana, J. Mankouri, and J. N. Barr, "Potassium is a trigger for conformational change in the fusion spike of an enveloped rna virus," *Journal of Biological Chemistry*, vol. 293, no. 26, pp. 9937–9944, 2018.

[91] M. Zivcec, F. E. Scholte, C. F. Spiropoulou, J. R. Spengler, and É. Bergeron, "Molecular insights into crimean-congo hemorrhagic fever virus," *Viruses*, vol. 8, no. 4, p. 106, 2016.

[92] C. T. Walter and J. N. Barr, "Recent advances in the molecular and cellular biology of bunyaviruses," *Journal of General Virology*, vol. 92, no. 11, pp. 2467–2484, 2011.

[93] Y. Matsumoto, K. Ohta, D. Kolakofsky, and M. Nishio, "A minigenome study of hazara nairovirus genomic promoters," *Journal of virology*, vol. 93, no. 6, pp. e02118–18, 2019.

[94] D. F. Mega, J. Fuller, B. Álvarez-Rodríguez, J. Mankouri, R. Hewson, and J. N. Barr, "Mutagenic analysis of hazara nairovirus non-translated regions during single and multi-step growth identifies both attenuating and functionally-critical sequences for virus replication," *Journal of Virology*, 2020.

[95] C. E. Garry and R. F. Garry, "Proteomics computational analyses suggest that the carboxyl terminal glycoproteins of bunyaviruses are class ii viral fusion protein (beta-penetrenes)," *Theoretical Biology and Medical Modelling*, vol. 1, no. 1, p. 10, 2004.

[96] A. Bertolotti-Ciarlet, J. Smith, K. Strecker, J. Paragas, L. A. Altamura, J. M. McFalls, N. Frias-Stäheli, A. García-Sastre, C. S. Schmaljohn, and R. W. Doms, "Cellular localization and antigenic characterization of crimean-congo hemorrhagic fever virus glycoproteins," *Journal of virology*, vol. 79, no. 10, pp. 6152–6161, 2005.

[97] X. Xiao, Y. Feng, Z. Zhu, and D. S. Dimitrov, "Identification of a putative crimean-congo hemorrhagic fever virus entry factor," *Biochemical and biophysical research communications*, vol. 411, no. 2, pp. 253–258, 2011.

[98] M. Simon, C. Johansson, and A. Mirazimi, "Crimean-congo hemorrhagic fever virus entry and replication is clathrin-, ph-and cholesterol-dependent," *Journal of General Virology*, vol. 90, no. 1, pp. 210–215, 2009.

[99] F. W. Charlton, S. Hover, J. Fuller, R. Hewson, J. Fontana, J. N. Barr, and J. Mankouri, "Cellular cholesterol abundance regulates potassium accumulation within endosomes and is an important determinant in bunyavirus entry," *Journal of Biological Chemistry*, vol. 294, no. 18, pp. 7335–7347, 2019.

[100] A. R. Garrison, S. R. Radoshitzky, K. P. Kota, G. Pegoraro, G. Ruthel, J. H. Kuhn, L. A. Altamura, S. A. Kwilas, S. Bavari, V. Haucke, *et al.*, "Crimean–congo hemorrhagic fever virus utilizes a clathrin-and early endosome-dependent entry pathway," *Virology*, vol. 444, no. 1-2, pp. 45–54, 2013.

[101] O. Shtanko, R. A. Nikitina, C. Z. Altuntas, A. A. Chepurnov, and R. A. Davey, "Crimean-congo hemorrhagic fever virus entry into host cells occurs through the

multivesicular body and requires escrt regulators," *PLoS pathogens*, vol. 10, no. 9, 2014.

[102] M. Simon, C. Johansson, Å. Lundkvist, and A. Mirazimi, "Microtubule-dependent and microtubule-independent steps in crimean-congo hemorrhagic fever virus replication cycle," *Virology*, vol. 385, no. 2, pp. 313–322, 2009.

[103] I. Andersson, M. Simon, Å. Lundkvist, M. Nilsson, A. Holmström, F. Elgh, and A. Mirazimi, "Role of actin filaments in targeting of crimean congo hemorrhagic fever virus nucleocapsid protein to perinuclear regions of mammalian cells," *Journal of medical virology*, vol. 72, no. 1, pp. 83–93, 2004.

[104] J. Reguera, F. Weber, and S. Cusack, "Bunyaviridae rna polymerases (l-protein) have an n-terminal, influenza-like endonuclease domain, essential for viral cap-dependent transcription," *PLoS pathogens*, vol. 6, no. 9, 2010.

[105] A. Kohl, E. F. Dunn, A. C. Lowen, and R. M. Elliott, "Complementarity, sequence and structural elements within the 3' and 5' non-coding regions of the bunyamwera orthobunyavirus s segment determine promoter strength," *Journal of General Virology*, vol. 85, no. 11, pp. 3269–3278, 2004.

[106] D. Garcin, M. Lezzi, M. Dobbs, R. M. Elliott, C. Schmaljohn, C. Y. Kang, and D. Kolakofsky, "The 5'ends of hantaan virus (bunyaviridae) rnas suggest a prime-and-realign mechanism for the initiation of rna synthesis.," *Journal of virology*, vol. 69, no. 9, pp. 5754–5762, 1995.

[107] J. N. Barr, J. W. Rodgers, and G. W. Wertz, "Identification of the bunyamwera bunyavirus transcription termination signal," *Journal of general virology*, vol. 87, no. 1, pp. 189–198, 2006.

[108] G. Blakqori, I. van Knippenberg, and R. M. Elliott, "Bunyamwera orthobunyavirus s-segment untranslated regions mediate poly (a) tail-independent translation," *Journal of virology*, vol. 83, no. 8, pp. 3637–3646, 2009.

[109] P. Gerlach, H. Malet, S. Cusack, and J. Reguera, "Structural insights into bunyavirus replication and its regulation by the vrna promoter," *Cell*, vol. 161, no. 6, pp. 1267–1279, 2015.

[110] T. Ikegami, S. Won, C. Peters, and S. Makino, "Rift valley fever virus nss mrna is transcribed from an incoming anti-viral-sense s rna segment," *Journal of virology*, vol. 79, no. 18, pp. 12106–12111, 2005.

[111] J. N. Barr and G. W. Wertz, "Role of the conserved nucleotide mismatch within 3'- and 5'-terminal regions of bunyamwera virus in signaling transcription," *Journal of virology*, vol. 79, no. 6, pp. 3586–3594, 2005.

[112] J. N. Barr, J. W. Rodgers, and G. W. Wertz, "The bunyamwera virus mrna transcription signal resides within both the 3' and the 5' terminal regions and allows ambisense transcription from a model rna segment," *Journal of virology*, vol. 79, no. 19, pp. 12602–12607, 2005.

[113] C. Andersson, S. Henriksson, K.-E. Magnusson, M. Nilsson, and A. Mirazimi, "In situ rolling circle amplification detection of crimean congo hemorrhagic fever virus (cchfv) complementary and viral rna," *Virology*, vol. 426, no. 2, pp. 87–92, 2012.

[114] R. R. Novoa, G. Calderita, R. Arranz, J. Fontana, H. Granzow, and C. Risco, "Virus factories: associations of cell organelles for viral replication and morphogenesis," *Biology of the Cell*, vol. 97, no. 2, pp. 147–172, 2005.

[115] E. Kuismanen, K. Hedman, J. Saraste, and R. Pettersson, "Uukuniemi virus maturation: accumulation of virus particles and viral antigens in the golgi complex.," *Molecular and cellular biology*, vol. 2, no. 11, pp. 1444–1458, 1982.

[116] J. Fontana, N. López-Montero, R. M. Elliott, J. J. Fernández, and C. Risco, "The unique architecture of bunyamwera virus factories around the golgi complex," *Cellular microbiology*, vol. 10, no. 10, pp. 2012–2028, 2008.

[117] D. F. Estrada and R. N. De Guzman, "Structural characterization of the crimean-congo hemorrhagic fever virus gn tail provides insight into virus assembly," *Journal of Biological Chemistry*, vol. 286, no. 24, pp. 21678–21686, 2011.

[118] R. Peng, X. Xu, J. Jing, M. Wang, Q. Peng, S. Liu, Y. Wu, X. Bao, P. Wang, J. Qi, *et al.*, "Structural insight into arenavirus replication machinery," *Nature*, vol. 579, no. 7800, pp. 615–619, 2020.

[119] D. Vogel, S. R. Thorkelsson, E. R. Quemin, K. Meier, T. Kouba, N. Gogrefe, C. Busch, S. Reindl, S. Günther, S. Cusack, *et al.*, "Structural and functional characterization of the severe fever with thrombocytopenia syndrome virus l protein," *Nucleic acids research*, vol. 48, no. 10, pp. 5749–5765, 2020.

[120] B. Arragain, G. Effantin, P. Gerlach, J. Reguera, G. Schoehn, S. Cusack, and H. Malet, "Pre-initiation and elongation structures of full-length la crosse virus polymerase reveal functionally important conformational changes," *bioRxiv*, 2020.

[121] M. U. Mirza, M. Vanmeert, M. Froeyen, A. Ali, S. Rafique, and M. Idrees, "In silico structural elucidation of rna-dependent rna polymerase towards the identification of potential crimean-congo hemorrhagic fever virus inhibitors," *Scientific reports*, vol. 9, no. 1, pp. 1–18, 2019.

[122] L. Lasecka and M. D. Baron, "The molecular biology of nairoviruses, an emerging group of tick-borne arboviruses," *Archives of virology*, vol. 159, no. 6, pp. 1249–1265, 2014.

[123] J. E. Honig, J. C. Osborne, and S. T. Nichol, "Crimean–congo hemorrhagic fever virus genome l rna segment and encoded protein," *Virology*, vol. 321, no. 1, pp. 29–35, 2004.

[124] J. M. L. Macleod, H. Marmor, A. García-Sastre, and N. Frias-Staheli, "Mapping of the interaction domains of the crimean–congo hemorrhagic fever virus nucleocapsid protein," *The Journal of general virology*, vol. 96, no. Pt 3, p. 524, 2015.

[125] F. E. Scholte, M. Zivcec, J. V. Dzimianski, M. K. Deaton, J. R. Spengler, S. R. Welch, S. T. Nichol, S. D. Pegan, C. F. Spiropoulou, and É. Bergeron, "Crimean-congo hemorrhagic fever virus suppresses innate immune responses via a ubiquitin and isg15 specific protease," *Cell reports*, vol. 20, no. 10, pp. 2396–2407, 2017.

[126] T. W. James, N. Frias-Staheli, J.-P. Bacik, J. M. L. Macleod, M. Khajehpour, A. García-Sastre, and B. L. Mark, "Structural basis for the removal of ubiquitin and interferon-stimulated gene 15 by a viral ovarian tumor domain-containing protease," *Proceedings of the National Academy of Sciences*, vol. 108, no. 6, pp. 2222–2227, 2011.

[127] É. Bergeron, C. G. Albariño, M. L. Khristova, and S. T. Nichol, "Crimean-congo hemorrhagic fever virus-encoded ovarian tumor protease activity is dispensable for virus rna polymerase function," *Journal of virology*, vol. 84, no. 1, pp. 216–226, 2010.

[128] T. Holm, J.-D. Kopicki, C. Busch, S. Olschewski, M. Rosenthal, C. Uetrecht, S. Günther, and S. Reindl, "Biochemical and structural studies reveal differences and commonalities among cap-snatching endonucleases from segmented negative-strand rna viruses," *Journal of Biological Chemistry*, vol. 293, no. 51, pp. 19686–19698, 2018.

[129] S. Devignot, E. Bergeron, S. Nichol, A. Mirazimi, and F. Weber, "A virus-like particle system identifies the endonuclease domain of crimean-congo hemorrhagic fever virus," *Journal of virology*, vol. 89, no. 11, pp. 5957–5967, 2015.

[130] A. J. Sanchez, M. J. Vincent, and S. T. Nichol, "Characterization of the glycoproteins of crimean-congo hemorrhagic fever virus," *Journal of virology*, vol. 76, no. 14, pp. 7263–7275, 2002.

[131] L. A. Altamura, A. Bertolotti-Ciarlet, J. Teigler, J. Paragas, C. S. Schmaljohn, and R. W. Doms, "Identification of a novel c-terminal cleavage of crimean-congo hemorrhagic fever virus pregn that leads to generation of an nsm protein," *Journal of virology*, vol. 81, no. 12, pp. 6632–6642, 2007.

[132] É. Bergeron, M. J. Vincent, and S. T. Nichol, "Crimean-congo hemorrhagic fever virus glycoprotein processing by the endoprotease ski-1/s1p is critical for virus infectivity," *Journal of virology*, vol. 81, no. 23, pp. 13271–13276, 2007.

[133] S. Haferkamp, L. Fernando, T. F. Schwarz, H. Feldmann, and R. Flick, "Intracellular localization of crimean-congo hemorrhagic fever (cchf) virus glycoproteins," *Virology journal*, vol. 2, no. 1, p. 42, 2005.

[134] E. Bergeron, M. Zivcec, A. K. Chakrabarti, S. T. Nichol, C. G. Albariño, and C. F. Spiropoulou, "Recovery of recombinant crimean congo hemorrhagic fever virus reveals a function for non-structural glycoproteins cleavage by furin," *PLoS pathogens*, vol. 11, no. 5, 2015.

[135] A. K. Mishra, C. L. Moyer, D. M. Abelson, D. J. Deer, K. El Omari, R. Duman, L. Lobel, J. J. Lutwama, J. M. Dye, A. Wagner, *et al.*, "Structure and characterization of crimean-congo hemorrhagic fever virus gp38," *Journal of Virology*, 2020.

[136] B. Barnwal, H. Karlberg, A. Mirazimi, and Y.-J. Tan, "The non-structural protein of crimean-congo hemorrhagic fever virus disrupts the mitochondrial membrane potential and induces apoptosis," *Journal of Biological Chemistry*, vol. 291, no. 2, pp. 582–592, 2016.

[137] J. Reguera, S. Cusack, and D. Kolakofsky, "Segmented negative strand rna virus nucleoprotein structure," *Current opinion in virology*, vol. 5, pp. 7–15, 2014.

[138] R. Surtees, S. D. Dowall, A. Shaw, S. Armstrong, R. Hewson, M. W. Carroll, J. Mankouri, T. A. Edwards, J. A. Hiscox, and J. N. Barr, "Heat shock protein

70 family members interact with crimean-congo hemorrhagic fever virus and hazara virus nucleocapsid proteins and perform a functional role in the nairovirus replication cycle," *Journal of virology*, vol. 90, no. 20, pp. 9305–9316, 2016.

[139] I. Andersson, L. Bladh, M. Mousavi-Jazi, K.-E. Magnusson, Å. Lundkvist, O. Haller, and A. Mirazimi, "Human mxa protein inhibits the replication of crimean-congo hemorrhagic fever virus," *Journal of virology*, vol. 78, no. 8, pp. 4323–4329, 2004.

[140] S. D. Carter, R. Surtees, C. T. Walter, A. Ariza, É. Bergeron, S. T. Nichol, J. A. Hiscox, T. A. Edwards, and J. N. Barr, "Structure, function, and evolution of the crimean-congo hemorrhagic fever virus nucleocapsid protein," *Journal of virology*, vol. 86, no. 20, pp. 10914–10923, 2012.

[141] Y. Guo, W. Wang, W. Ji, M. Deng, Y. Sun, H. Zhou, C. Yang, F. Deng, H. Wang, Z. Hu, *et al.*, "Crimean-congo hemorrhagic fever virus nucleoprotein reveals endonuclease activity in bunyaviruses," *Proceedings of the National Academy of Sciences*, vol. 109, no. 13, pp. 5046–5051, 2012.

[142] W. Wang, X. Liu, X. Wang, H. Dong, C. Ma, J. Wang, B. Liu, Y. Mao, Y. Wang, T. Li, *et al.*, "Structural and functional diversity of nairovirus-encoded nucleoproteins," *Journal of virology*, vol. 89, no. 23, pp. 11740–11749, 2015.

[143] O. Lenz, J. ter Meulen, H.-D. Klenk, N. G. Seidah, and W. Garten, "The lassa virus glycoprotein precursor gp-c is proteolytically processed by subtilase ski-1/s1p," *Proceedings of the National Academy of Sciences*, vol. 98, no. 22, pp. 12701–12705, 2001.

[144] K. M. Hastie, L. B. King, M. A. Zandonatti, and E. O. Saphire, "Structural basis for the dsrna specificity of the lassa virus np exonuclease," *PloS one*, vol. 7, no. 8, 2012.

[145] Y. Wang, S. Dutta, H. Karlberg, S. Devignot, F. Weber, Q. Hao, Y. J. Tan, A. Mirazimi, and M. Kotaka, "Structure of crimean-congo hemorrhagic fever virus nucleoprotein: superhelical homo-oligomers and the role of caspase-3 cleavage," *Journal of virology*, vol. 86, no. 22, pp. 12294–12303, 2012.

[146] H. Karlberg, Y.-J. Tan, and A. Mirazimi, "Induction of caspase activation and cleavage of the viral nucleocapsid protein in different cell types during crimean-congo hemorrhagic fever virus infection," *Journal of Biological Chemistry*, vol. 286, no. 5, pp. 3227–3234, 2011.

[147] H. Zhou, Y. Sun, Y. Guo, and Z. Lou, "Structural perspective on the formation of ribonucleoprotein complex in negative-sense single-stranded rna viruses," *Trends in microbiology*, vol. 21, no. 9, pp. 475–484, 2013.

[148] K. M. Hastie, T. Liu, S. Li, L. B. King, N. Ngo, M. A. Zandonatti, V. L. Woods, J. C. de la Torre, and E. O. Saphire, "Crystal structure of the lassa virus nucleoprotein–rna complex reveals a gating mechanism for rna binding," *Proceedings of the National Academy of Sciences*, vol. 108, no. 48, pp. 19365–19370, 2011.

[149] X. Wang, B. Li, Y. Guo, S. Shen, L. Zhao, P. Zhang, Y. Sun, S.-F. Sui, F. Deng, and Z. Lou, "Molecular basis for the formation of ribonucleoprotein complex of crimean-congo hemorrhagic fever virus," *Journal of structural biology*, vol. 196, no. 3, pp. 455–465, 2016.

[150] S. Jeeva, S. Pador, B. Voss, S. S. Ganaie, and M. A. Mir, "Crimean-congo hemorrhagic fever virus nucleocapsid protein has dual rna binding modes," *PloS one*, vol. 12, no. 9, 2017.

[151] R. Surtees, A. Ariza, E. K. Punch, C. H. Trinh, S. D. Dowall, R. Hewson, J. A. Hiscox, J. N. Barr, and T. A. Edwards, "The crystal structure of the hazara virus nucleocapsid protein," *BMC structural biology*, vol. 15, no. 1, p. 24, 2015.

[152] J. Fuller, R. Surtees, A. Shaw, B. Álvarez-Rodríguez, G. Slack, L. Bell-Sakyi, J. Mankouri, T. Edwards, R. Hewson, and J. Barr, "Hazara nairovirus elicits differential induction of apoptosis and nucleocapsid protein cleavage in mammalian and tick cells," *The Journal of general virology*, vol. 100, no. 3, pp. 392–402, 2019.

[153] T. Hoenen, A. Groseth, F. de Kok-Mercado, J. H. Kuhn, and V. Wahl-Jensen, "Minigenomes, transcription and replication competent virus-like particles and beyond: reverse genetics systems for filoviruses and other negative stranded hemorrhagic fever viruses," *Antiviral research*, vol. 91, no. 2, pp. 195–208, 2011.

[154] J. Zhao, H. Xia, Y. Zhang, S. Yin, Z. Zhang, S. Tang, Z. Kou, J. Yu, Z. Fan, and T. Li, "Mini-genome rescue of crimean-congo hemorrhagic fever virus and research into the evolutionary patterns of its untranslated regions," *Virus research*, vol. 177, no. 1, pp. 22–34, 2013.

[155] M. Zivcec, M. G. Metcalfe, C. G. Albarino, L. W. Guerrero, S. D. Pegan, C. F. Spiropoulou, and E. Bergeron, "Assessment of inhibitors of pathogenic crimean-

congo hemorrhagic fever virus strains using virus-like particles," *PLoS neglected tropical diseases*, vol. 9, no. 12, 2015.

[156] J. Fuller, R. Surtees, G. Slack, J. Mankouri, R. Hewson, and J. Barr, "Rescue of infectious recombinant hazara nairovirus from cdna reveals the nucleocapsid protein dqvd caspase cleavage motif performs an essential role other than cleavage," *Journal of virology*, vol. 93, no. 15, pp. e00616–19, 2019.

[157] D. Z. Wang, U. Wille, and E. Juaristi, *Encyclopedia of Physical Organic Chemistry:...* Wiley Online Library, 2017.

[158] C. Tuerk and L. Gold, "Systematic evolution of ligands by exponential enrichment: Rna ligands to bacteriophage t4 dna polymerase," *science*, vol. 249, no. 4968, pp. 505–510, 1990.

[159] S. Kedzierski, M. Khoshnejad, and G. T. Caltagirone, "Synthetic antibodies: the emerging field of aptamers," *Bioprocess. J*, vol. 11, pp. 46–49, 2012.

[160] H. W. Schroeder Jr and L. Cavacini, "Structure and function of immunoglobulins," *Journal of Allergy and Clinical Immunology*, vol. 125, no. 2, pp. S41–S52, 2010.

[161] I. Sela-Culang, V. Kunik, and Y. Ofran, "The structural basis of antibody-antigen recognition," *Frontiers in immunology*, vol. 4, p. 302, 2013.

[162] Y. Sun, Z. Liu, L. Ren, Z. Wei, P. Wang, N. Li, and Y. Zhao, "Immunoglobulin genes and diversity: what we have learned from domestic animals," *Journal of animal science and biotechnology*, vol. 3, no. 1, p. 18, 2012.

[163] G. Vidarsson, G. Dekkers, and T. Rispens, "Igg subclasses and allotypes: from structure to effector functions," *Frontiers in immunology*, vol. 5, p. 520, 2014.

[164] W. R. Strohl and L. M. Strohl, *Therapeutic antibody engineering: current and future advances driving the strongest growth area in the pharmaceutical industry*. Elsevier, 2012.

[165] X. Yu, Y.-P. Yang, E. Dikici, S. K. Deo, and S. Daunert, "Beyond antibodies as binding partners: the role of antibody mimetics in bioanalysis," *Annual Review of Analytical Chemistry*, vol. 10, pp. 293–320, 2017.

[166] W. Dejnirattisai, P. Supasa, W. Wongwiwat, A. Rouvinski, G. Barba-Spaeth, T. Duangchinda, A. Sakuntabhai, V.-M. Cao-Lormeau, P. Malasit, F. A. Rey, *et al.*, "Dengue virus sero-cross-reactivity drives antibody-dependent enhancement of infection with zika virus," *Nature immunology*, vol. 17, no. 9, p. 1102, 2016.

[167] L. M. Paul, E. R. Carlin, M. M. Jenkins, A. L. Tan, C. M. Barcellona, C. O. Nicholson, S. F. Michael, and S. Isern, "Dengue virus antibodies enhance zika virus infection," *Clinical & translational immunology*, vol. 5, no. 12, p. e117, 2016.

[168] M. Johnson, "Antibody structure and fragments," *Mater. Methods*, vol. 3, no. 160), 2013.

[169] Z. A. Ahmad, S. K. Yeap, A. M. Ali, W. Y. Ho, N. B. M. Alitheen, and M. Hamid, "scfv antibody: principles and clinical application," *Clinical and developmental immunology*, vol. 2012, 2012.

[170] U. Brinkmann and R. E. Kontermann, "The making of bispecific antibodies," in *MAbs*, vol. 9, pp. 182–212, Taylor & Francis, 2017.

[171] A. Bates and C. A. Power, "David vs. goliath: the structure, function, and clinical prospects of antibody fragments," *Antibodies*, vol. 8, no. 2, p. 28, 2019.

[172] F. Ducancel and B. H. Muller, "Molecular engineering of antibodies for therapeutic and diagnostic purposes," in *MAbs*, vol. 4, pp. 445–457, Taylor & Francis, 2012.

[173] R. Vazquez-Lombardi, T. G. Phan, C. Zimmermann, D. Lowe, L. Jermutus, and D. Christ, "Challenges and opportunities for non-antibody scaffold drugs," *Drug discovery today*, vol. 20, no. 10, pp. 1271–1283, 2015.

[174] R. Simeon and Z. Chen, "In vitro-engineered non-antibody protein therapeutics," *Protein & cell*, vol. 9, no. 1, pp. 3–14, 2018.

[175] B. A. Cohen, P. Colas, and R. Brent, "An artificial cell-cycle inhibitor isolated from a combinatorial library," *Proceedings of the National Academy of Sciences*, vol. 95, no. 24, pp. 14272–14277, 1998.

[176] A. Skerra, "Alternative non-antibody scaffolds for molecular recognition," *Current opinion in biotechnology*, vol. 18, no. 4, pp. 295–304, 2007.

[177] R. Bedford, C. Tiede, R. Hughes, A. Curd, M. McPherson, M. Peckham, and D. C. Tomlinson, "Alternative reagents to antibodies in imaging applications," *Biophysical reviews*, vol. 9, no. 4, pp. 299–308, 2017.

[178] J. Löfblom, J. Feldwisch, V. Tolmachev, J. Carlsson, S. Ståhl, and F. Y. Frejd, "Affibody molecules: engineered proteins for therapeutic, diagnostic and biotechnological applications," *FEBS letters*, vol. 584, no. 12, pp. 2670–2680, 2010.

[179] H. Ebersbach, E. Fiedler, T. Scheuermann, M. Fiedler, M. T. Stubbs, C. Reimann, G. Proetzel, R. Rudolph, and U. Fiedler, "Affilin–novel binding molecules based on human γ-b-crystallin, an all β-sheet protein," *Journal of molecular biology*, vol. 372, no. 1, pp. 172–185, 2007.

[180] A. Hoffmann, M. Kovermann, H. Lilie, M. Fiedler, J. Balbach, R. Rudolph, and S. Pfeifer, "New binding mode to tnf-alpha revealed by ubiquitin-based artificial binding protein," *PLoS One*, vol. 7, no. 2, 2012.

[181] C. Tiede, A. A. Tang, S. E. Deacon, U. Mandal, J. E. Nettleship, R. L. Owen, S. E. George, D. J. Harrison, R. J. Owens, D. C. Tomlinson, *et al.*, "Adhiron: a stable and versatile peptide display scaffold for molecular recognition applications," *Protein Engineering, Design and Selection*, vol. 27, no. 5, pp. 145–155, 2014.

[182] J. Desmet, K. Verstraete, Y. Bloch, E. Lorent, Y. Wen, B. Devreese, K. Vandenbroucke, S. Loverix, T. Hettmann, S. Deroo, *et al.*, "Structural basis of il-23 antagonism by an alphabody protein scaffold," *Nature communications*, vol. 5, no. 1, pp. 1–12, 2014.

[183] M. Gebauer, A. Schiefner, G. Matschiner, and A. Skerra, "Combinatorial design of an anticalin directed against the extra-domain b for the specific targeting of oncofetal fibronectin," *Journal of molecular biology*, vol. 425, no. 4, pp. 780–802, 2013.

[184] P. Byla, M. H. Andersen, T. L. Holtet, H. Jacobsen, M. Munch, H. H. Gad, H. C. Thøgersen, and R. Hartmann, "Selection of a novel and highly specific tumor necrosis factor α (tnfα) antagonist insight from the crystal structure of the antagonist-tnfα complex," *Journal of Biological Chemistry*, vol. 285, no. 16, pp. 12096–12100, 2010.

[185] J. Silverman, Q. Lu, A. Bakker, W. To, A. Duguay, B. M. Alba, R. Smith, A. Rivas, P. Li, H. Le, *et al.*, "Multivalent avimer proteins evolved by exon shuffling of a family of human receptor domains," *Nature biotechnology*, vol. 23, no. 12, pp. 1556–1561, 2005.

[186] J. A. Robinson, "β-hairpin peptidomimetics: design, structures and biological activities," *Accounts of chemical research*, vol. 41, no. 10, pp. 1278–1288, 2008.

[187] E. Iaccarino, L. Calvanese, G. Untiveros, L. Falcigno, G. D'Auria, D. Latino, L. Strizzi, M. Ruvo, A. Sandomenico, *et al.*, "Structure-based design of small bicyclic peptide inhibitors of cripto-1 activity.," *The Biochemical Journal*, 2020.

[188] B. G. Lui, N. Salomon, J. Wüstehube-Lausch, M. Daneschdar, H.-U. Schmoldt, Ö. Türeci, and U. Sahin, "Targeting the tumor vasculature with engineered cystine-knot miniproteins," *Nature Communications*, vol. 11, no. 1, pp. 1–11, 2020.

[189] M. A. Seeger, R. Zbinden, A. Flütsch, P. G. Gutte, S. Engeler, H. Roschitzki-Voser, and M. G. Grütter, "Design, construction, and characterization of a second-generation darpin library with reduced hydrophobicity," *Protein science*, vol. 22, no. 9, pp. 1239–1257, 2013.

[190] T. Mitchell, G. Chao, D. Sitkoff, F. Lo, H. Monshizadegan, D. Meyers, S. Low, K. Russo, R. DiBella, F. Denhez, *et al.*, "Pharmacologic profile of the adnectin bms-962476, a small protein biologic alternative to pcsk9 antibodies for low-density lipoprotein lowering," *Journal of Pharmacology and Experimental Therapeutics*, vol. 350, no. 2, pp. 412–424, 2014.

[191] M. Silacci, N. Baenziger-Tobler, W. Lembke, W. Zha, S. Batey, J. Bertschinger, and D. Grabulovski, "Linker length matters, fynomer-fc fusion with an optimized linker displaying picomolar il-17a inhibition potency," *Journal of Biological Chemistry*, vol. 289, no. 20, pp. 14392–14398, 2014.

[192] S. J. Moore and J. R. Cochran, "Engineering knottins as novel binding agents," in *Methods in enzymology*, vol. 503, pp. 223–251, Elsevier, 2012.

[193] M. S. Dennis, A. Herzka, and R. A. Lazarus, "Potent and selective kunitz domain inhibitors of plasma kallikrein designed by phage display," *Journal of Biological Chemistry*, vol. 270, no. 43, pp. 25411–25417, 1995.

[194] M. Krehenbrink, M. Chami, I. Guilvout, P. M. Alzari, F. Pécorari, and A. P. Pugsley, "Artificial binding proteins (affitins) as probes for conformational changes in secretin puld," *Journal of molecular biology*, vol. 383, no. 5, pp. 1058–1068, 2008.

[195] S.-C. Lee, K. Park, J. Han, J.-j. Lee, H. J. Kim, S. Hong, W. Heu, Y. J. Kim, J.-S. Ha, S.-G. Lee, *et al.*, "Design of a binding scaffold based on variable lymphocyte receptors of jawless vertebrates by module engineering," *Proceedings of the National Academy of Sciences*, vol. 109, no. 9, pp. 3299–3304, 2012.

[196] J. D. Steemson, M. Baake, J. Rakonjac, V. L. Arcus, and M. T. Liddament, "Tracking molecular recognition at the atomic level with a new protein scaffold based on the ob-fold," *PLoS One*, vol. 9, no. 1, 2014.

[197] T. Hoffmann, L. K. J. Stadler, M. Busby, Q. Song, A. T. Buxton, S. D. Wagner, J. J. Davis, and P. Ko Ferrigno, "Structure-function studies of an engineered scaffold protein derived from stefin a. i: Development of the sqm variant," *Protein Engineering, Design & Selection*, vol. 23, no. 5, pp. 403–413, 2010.

[198] L. K. J. Stadler, T. Hoffmann, D. C. Tomlinson, Q. Song, T. Lee, M. Busby, Y. Nyathi, E. Gendra, C. Tiede, K. Flanagan, *et al.*, "Structure- function studies of an engineered scaffold protein derived from stefin a. ii: Development and applications of the sqt variant," *Protein Engineering, Design & Selection*, vol. 24, no. 9, pp. 751–763, 2011.

[199] L. Ledsgaard, M. Kilstrup, A. Karatt-Vellatt, J. McCafferty, and A. H. Laustsen, "Basics of antibody phage display technology," *Toxins*, vol. 10, no. 6, p. 236, 2018.

[200] A. Ah-San Tang, C. Tiede, D. J. Hughes, M. J. McPherson, and D. C. Tomlinson, "Isolation of isoform-specific binding proteins (affimers) by phage display using negative selection," *Sci. Signal.*, vol. 10, no. 505, p. eaan0868, 2017.

[201] C. Tiede, R. Bedford, S. J. Heseltine, G. Smith, I. Wijetunga, R. Ross, D. AlQallaf, A. P. Roberts, A. Balls, A. Curd, *et al.*, "Affimer proteins are versatile and renewable affinity reagents," *Elife*, vol. 6, p. e24903, 2017.

[202] E. L. Hesketh, C. Tiede, H. Adamson, T. L. Adams, M. J. Byrne, Y. Meshcheriakova, I. Kruse, M. J. McPherson, G. P. Lomonossoff, D. C. Tomlinson, *et al.*, "Affimer reagents as tools in diagnosing plant virus diseases," *Scientific reports*, vol. 9, no. 1, pp. 1–9, 2019.

[203] T. Schlichthaerle, A. S. Eklund, F. Schueder, M. T. Strauss, C. Tiede, A. Curd, J. Ries, M. Peckham, D. C. Tomlinson, and R. Jungmann, "Site-specific labeling of affimers for dna-paint microscopy," *Angewandte Chemie International Edition*, vol. 57, no. 34, pp. 11060–11063, 2018.

[204] J. J. Davis, J. Tkac, R. Humphreys, A. T. Buxton, T. A. Lee, and P. Ko Ferrigno, "Peptide aptamers in label-free protein detection: 2. chemical optimization and detection of distinct protein isoforms," *Analytical chemistry*, vol. 81, no. 9, pp. 3314–3320, 2009.

[205] P. Estrela, D. Paul, Q. Song, L. K. Stadler, L. Wang, E. Huq, J. J. Davis, P. K. Ferrigno, and P. Migliorato, "Label-free sub-picomolar protein detection with field-effect transistors," *Analytical chemistry*, vol. 82, no. 9, pp. 3531–3536, 2010.

[206] A. Johnson, Q. Song, P. Ko Ferrigno, P. R. Bueno, and J. J. Davis, "Sensitive affimer and antibody based impedimetric label-free assays for c-reactive protein," *Analytical chemistry*, vol. 84, no. 15, pp. 6553–6560, 2012.

[207] P. Zhurauski, S. K. Arya, P. Jolly, C. Tiede, D. C. Tomlinson, P. K. Ferrigno, and P. Estrela, "Sensitive and selective affimer-functionalised interdigitated electrode-based capacitive biosensor for her4 protein tumour biomarker detection," *Biosensors and Bioelectronics*, vol. 108, pp. 1–8, 2018.

[208] P. K. Ferrigno, "Non-antibody protein-based biosensors," *Essays in biochemistry*, vol. 60, no. 1, pp. 19–25, 2016.

[209] A. E. Rawlings, J. P. Bramble, A. A. Tang, L. A. Somner, A. E. Monnington, D. J. Cooke, M. J. McPherson, D. C. Tomlinson, and S. S. Staniland, "Phage display selected magnetite interacting adhirons for shape controlled nanoparticle synthesis," *Chemical science*, vol. 6, no. 10, pp. 5586–5594, 2015.

[210] P. J. Carter, "Introduction to current and future protein therapeutics: a protein engineering perspective," *Experimental cell research*, vol. 317, no. 9, pp. 1261–1269, 2011.

[211] M. J. Fisher, D. J. Williamson, G. M. Burslem, J. P. Plante, I. W. Manfield, C. Tiede, J. R. Ault, P. G. Stockley, S. Plein, A. Maqbool, *et al.*, "Trivalent gd-dota reagents for modification of proteins," *RSC advances*, vol. 5, no. 116, pp. 96194–96200, 2015.

[212] M. Raina, R. Sharma, S. E. Deacon, C. Tiede, D. Tomlinson, A. Davies, M. McPherson, and C. Wälti, "Antibody mimetic receptor proteins for label-free biosensors," *Analyst*, vol. 140, no. 3, pp. 803–810, 2015.

[213] D. Kamiyama, S. Sekine, B. Barsi-Rhyne, J. Hu, B. Chen, L. A. Gilbert, H. Ishikawa, M. D. Leonetti, W. F. Marshall, J. S. Weissman, *et al.*, "Versatile protein tagging in cells with split fluorescent protein," *Nature communications*, vol. 7, no. 1, pp. 1–9, 2016.

[214] M. Rai and H. Padh, "Expression systems for production of heterologous proteins," *CURRENT SCIENCE-BANGALORE-*, vol. 80, no. 9, pp. 1121–1128, 2001.

[215] G. L. Rosano and E. A. Ceccarelli, "Recombinant protein expression in escherichia coli: advances and challenges," *Frontiers in microbiology*, vol. 5, p. 172, 2014.

[216] I. MANUAL, "pet system vectors and hosts,"

[217] C. Pop, Y.-R. Chen, B. Smith, K. Bose, B. Bobay, A. Tripathy, S. Franzen, and A. C. Clark, "Removal of the pro-domain does not affect the conformation of the procaspase-3 dimer," *Biochemistry*, vol. 40, no. 47, pp. 14224–14235, 2001.

[218] M. Hedhammar, A. E. Karlström, and S. Hober, "Chromatographic methods for protein purification," *Stockholm: Royal Institute of Technology*, pp. 1–31, 2006.

[219] O. Coskun, "Separation techniques: chromatography," *Northern clinics of Istanbul*, vol. 3, no. 2, p. 156, 2016.

[220] J. A. Bornhorst and J. J. Falke, "[16] purification of proteins using polyhistidine affinity tags," in *Methods in enzymology*, vol. 326, pp. 245–254, Elsevier, 2000.

[221] A. Biosciences, "Gel filtration: Principles and methods," *Amersham Biosciences, Piscataway, NJ*, 2002.

[222] J.-B. Denault and G. S. Salvesen, "Expression, purification, and characterization of caspases," *Current protocols in protein science*, vol. 30, no. 1, pp. 21–13, 2002.

[223] T. Fernandes-Alnemri, G. Litwack, and E. S. Alnemri, "Mch2, a new member of the apoptotic ced-3/ice cysteine protease gene family," *Cancer research*, vol. 55, no. 13, pp. 2737–2742, 1995.

[224] S. Carter, J. Barr, and T. Edwards, "Expression, purification and crystallization of the crimean–congo haemorrhagic fever virus nucleocapsid protein," *Acta crystallographica section F: structural biology and crystallization communications*, vol. 68, no. 5, pp. 569–573, 2012.

[225] M. A. Michel, K. N. Swatek, M. K. Hospenthal, and D. Komander, "Ubiquitin linkage-specific affimers reveal insights into k6-linked ubiquitin signaling," *Molecular cell*, vol. 68, no. 1, pp. 233–246, 2017.

[226] T. Moks, L. ABRAHMSÉN, B. NILSSON, U. HELLMAN, J. SJÖQUIST, and M. UHLÉN, "Staphylococcal protein a consists of five igg-binding domains," *European journal of biochemistry*, vol. 156, no. 3, pp. 637–643, 1986.

[227] M. E. Kimple, A. L. Brill, and R. L. Pasker, "Overview of affinity tags for protein purification," *Current protocols in protein science*, vol. 73, no. 1, pp. 9–9, 2013.

[228] K. Meyer and M. Selbach, "Quantitative affinity purification mass spectrometry: a versatile technology to study protein–protein interactions," *Frontiers in genetics*, vol. 6, p. 237, 2015.

[229] Y. Tang, X. Zeng, and J. Liang, "Surface plasmon resonance: an introduction to a surface spectroscopy technique," *Journal of chemical education*, vol. 87, no. 7, pp. 742–746, 2010.

[230] D. Wang, J. F. C. Loo, J. Chen, Y. Yam, S.-C. Chen, H. He, S. K. Kong, and H. P. Ho, "Recent advances in surface plasmon resonance imaging sensors," *Sensors*, vol. 19, no. 6, p. 1266, 2019.

[231] A. Dupont-Filliard, A. Roget, T. Livache, and M. Billon, "Reversible oligonucleotide immobilisation based on biotinylated polypyrrole film," *Analytica chimica acta*, vol. 449, no. 1-2, pp. 45–50, 2001.

[232] L. Masson, A. Mazza, and G. De Crescenzo, "Determination of affinity and kinetic rate constants using surface plasmon resonance," in *Bacterial Toxins: Methods and Protocols*, pp. 189–201, Springer, 2000.

[233] N. Berova, K. Nakanishi, and R. W. Woody, *Circular dichroism: principles and applications*. John Wiley & Sons, 2000.

[234] P. Fischer and F. Hache, "Nonlinear optical spectroscopy of chiral molecules," *Chirality: the pharmacological, biological, and chemical consequences of molecular asymmetry*, vol. 17, no. 8, pp. 421–437, 2005.

[235] S. Y. Venyaminov and J. T. Yang, "Determination of protein secondary structure," in *Circular dichroism and the conformational analysis of biomolecules*, pp. 69–107, Springer, 1996.

[236] W. C. Johnson Jr, "Protein secondary structure and circular dichroism: a practical guide," *Proteins: Structure, Function, and Bioinformatics*, vol. 7, no. 3, pp. 205–214, 1990.

[237] N. J. Greenfield, "Using circular dichroism collected as a function of temperature to determine the thermodynamics of protein unfolding and binding interactions," *Nature protocols*, vol. 1, no. 6, p. 2527, 2006.

[238] A. Micsonai, F. Wien, É. Bulyáki, J. Kun, É. Moussong, Y.-H. Lee, Y. Goto, M. Réfrégiers, and J. Kardos, "Bestsel: a web server for accurate protein secondary structure prediction and fold recognition from the circular dichroism spectra," *Nucleic acids research*, vol. 46, no. W1, pp. W315–W322, 2018.

[239] K. Miura, "An overview of current methods to confirm protein-protein interactions," *Protein and peptide letters*, vol. 25, no. 8, pp. 728–733, 2018.

[240] Y. Bai, "Detecting protein-protein interactions by gel filtration chromatography," in *Protein-Protein Interactions*, pp. 223–232, Springer, 2015.

[241] J. M. Pagano, C. C. Clingman, and S. P. Ryder, "Quantitative approaches to monitor protein–nucleic acid interactions using fluorescent probes," *Rna*, vol. 17, no. 1, pp. 14–20, 2011.

[242] W. A. Lea and A. Simeonov, "Fluorescence polarization assays in small molecule screening," *Expert opinion on drug discovery*, vol. 6, no. 1, pp. 17–32, 2011.

[243] D. D. Licatalosi, X. Ye, and E. Jankowsky, "Approaches for measuring the dynamics of rna–protein interactions," *Wiley Interdisciplinary Reviews: RNA*, vol. 11, no. 1, p. e1565, 2020.

[244] A. Wlodawer, W. Minor, Z. Dauter, and M. Jaskolski, "Protein crystallography for non-crystallographers, or how to get the best (but not more) from published macromolecular structures," *The FEBS journal*, vol. 275, no. 1, pp. 1–21, 2008.

[245] G. Rhodes, *Crystallography made crystal clear: a guide for users of macromolecular models*. Elsevier, 2010.

[246] M. A. Dessau and Y. Modis, "Protein crystallization for x-ray crystallography," *JoVE (Journal of Visualized Experiments)*, no. 47, p. e2285, 2011.

[247] M. M. Woolfson and M. M. Woolfson, *An introduction to X-ray crystallography*. Cambridge University Press, 1997.

[248] G. L. Taylor, "Introduction to phasing," *Acta Crystallographica Section D: Biological Crystallography*, vol. 66, no. 4, pp. 325–338, 2010.

[249] V. B. Chen, W. B. Arendall, J. J. Headd, D. A. Keedy, R. M. Immormino, G. J. Kapral, L. W. Murray, J. S. Richardson, and D. C. Richardson, "Molprobity: all-atom structure validation for macromolecular crystallography," *Acta Crystallographica Section D: Biological Crystallography*, vol. 66, no. 1, pp. 12–21, 2010.

[250] J. I. Robinson, E. W. Baxter, R. L. Owen, M. Thomsen, D. C. Tomlinson, M. P. Waterhouse, S. J. Win, J. E. Nettleship, C. Tiede, R. J. Foster, *et al.*, "Affimer proteins inhibit immune complex binding to fcγriiia with high specificity through competitive and allosteric modes of action," *Proceedings of the National Academy of Sciences*, vol. 115, no. 1, pp. E72–E81, 2018.

[251] D. J. Hughes, C. Tiede, N. Penswick, A. Ah-San Tang, C. H. Trinh, U. Mandal, K. Z. Zajac, T. Gaule, G. Howell, T. A. Edwards, *et al.*, "Generation of specific inhibitors of sumo-1–and sumo-2/3–mediated protein-protein interactions using affimer (adhiron) technology," *Sci. Signal.*, vol. 10, no. 505, p. eaaj2005, 2017.

[252] M. S. Mehand, F. Al-Shorbaji, P. Millett, and B. Murgue, "The who r&d blueprint: 2018 review of emerging infectious diseases requiring urgent research and development efforts," *Antiviral research*, vol. 159, pp. 63–67, 2018.

[253] M. Saijo, Q. Tang, B. Shimayi, L. Han, Y. Zhang, M. Asiguma, D. Tianshu, A. Maeda, I. Kurane, and S. Morikawa, "Antigen-capture enzyme-linked immunosorbent assay for the diagnosis of crimean-congo hemorrhagic fever using a novel monoclonal antibody," *Journal of medical virology*, vol. 77, no. 1, pp. 83–88, 2005.

[254] G. Mattiuzzo, E. M. Bentley, and M. Page, "The role of reference materials in the research and development of diagnostic tools and treatments for haemorrhagic fever viruses," *Viruses*, vol. 11, no. 9, p. 781, 2019.

[255] M. L. Boisen, J. N. Hartnett, J. G. Shaffer, A. Goba, M. Momoh, J. D. Sandi, M. Fullah, D. K. Nelson, D. J. Bush, M. M. Rowland, *et al.*, "Field validation of recombinant antigen immunoassays for diagnosis of lassa fever," *Scientific reports*, vol. 8, no. 1, p. 5939, 2018.

[256] R. Yoshida, S. Muramatsu, H. Akita, Y. Saito, M. Kuwahara, D. Kato, K. Changula, H. Miyamoto, M. Kajihara, R. Manzoor, *et al.*, "Development of an immunochromatography assay (quicknavi-ebola) to detect multiple species of ebolaviruses," *The Journal of infectious diseases*, vol. 214, no. suppl_3, pp. S185–S191, 2016.

[257] C. Cêtre-Sossah, A. Pédarrieu, M. Juremalm, P. J. Van Vuren, A. Brun, A. B. O. E. Mamy, J.-M. Héraud, C. Filippone, J.-P. Ravalohery, H. Chaabihi, *et al.*, "Development and validation of a pen side test for rift valley fever," *PLoS neglected tropical diseases*, vol. 13, no. 9, 2019.

[258] S. Amelinckx, D. Van Dyck, J. Van Landuyt, and G. Van Tendeloo, *Electron microscopy: principles and fundamentals*. John Wiley & Sons, 2008.

[259] L. E. Franken, K. Grünewald, E. J. Boekema, and M. C. Stuart, "A technical introduction to transmission electron microscopy for soft-matter: Imaging,

possibilities, choices, and technical developments," *Small*, vol. 16, no. 14, p. 1906198, 2020.

[260] U. Aebi and T. D. Pollard, "A glow discharge unit to render electron microscope grids and other surfaces hydrophilic," *Journal of electron microscopy technique*, vol. 7, no. 1, pp. 29–33, 1987.

[261] D. F. Barreto-Vieira and O. M. Barth, *Negative and positive staining in transmission electron microscopy for virus diagnosis*. InTech, 2015.

[262] R. F. Thompson, M. Walker, C. A. Siebert, S. P. Muench, and N. A. Ranson, "An introduction to sample preparation and imaging by cryo-electron microscopy for structural biology," *Methods*, vol. 100, pp. 3–15, 2016.

[263] K. Murata and M. Wolf, "Cryo-electron microscopy for structural analysis of dynamic biological macromolecules," *Biochimica et Biophysica Acta (BBA)-General Subjects*, vol. 1862, no. 2, pp. 324–334, 2018.

[264] W. Saenger, *Principles of nucleic acid structure*. Springer Science & Business Media, 2013.

[265] S. Cabantous, T. C. Terwilliger, and G. S. Waldo, "Protein tagging and detection with engineered self-assembling fragments of green fluorescent protein," *Nature biotechnology*, vol. 23, no. 1, pp. 102–107, 2005.

[266] K. P. Kent, W. Childs, and S. G. Boxer, "Deconstructing green fluorescent protein," *Journal of the American Chemical Society*, vol. 130, no. 30, pp. 9664–9665, 2008.

[267] D. Semwogerere and E. R. Weeks, "Confocal microscopy," *Encyclopedia of biomaterials and biomedical engineering*, vol. 23, pp. 1–10, 2005.

[268] D. M. Shotton, "Confocal scanning optical microscopy and its applications for biological specimens," *Journal of Cell Science*, vol. 94, no. 2, pp. 175–206, 1989.

[269] P. Rwambo, M. Shaw, F. Rurangirwa, and J. DeMartini, "Ultrastructural studies on the replication and morphogenesis of nairobi sheep disease virus, anairovirus," *Archives of virology*, vol. 141, no. 8, pp. 1479–1492, 1996.

[270] Y. Sun, J. Li, G. F. Gao, P. Tien, and W. Liu, "Bunyavirales ribonucleoproteins: the viral replication and transcription machinery," *Critical reviews in microbiology*, vol. 44, no. 5, pp. 522–540, 2018.

[271] F. Niu, N. Shaw, Y. E. Wang, L. Jiao, W. Ding, X. Li, P. Zhu, H. Upur, S. Ouyang, G. Cheng, *et al.*, "Structure of the leanyer orthobunyavirus nucleoprotein–rna complex reveals unique architecture for rna encapsidation," *Proceedings of the National Academy of Sciences*, vol. 110, no. 22, pp. 9054–9059, 2013.

Appendix A

Protein sequences and physicochemical parameters

Amino acid sequence, molecular weight and theoretical isoelectric points (pI) of recombinantly produced proteins are detailed in this section. * indicates a stop codon.

CCHFV NP

MENKIEVNSKDEMNKWFEEFKKGNGLVDTYTNSYSFCESVPNLDRFVFQM
AGATDDAQKDSIYASALVEATKFCAPIYECAWASSTGIVKKGLEWFEKNTGTIK
SWDESYIELKVEVPKIEQLFNYQQAALKWRKDIGFRVNANTAALSNKVLAEYK
VPGEIVMSVKEMLSDMIRRRNLILNRGGDENPRGPVSHEHVEWCREFVKGKYI
MAFNPPWGDINKSGRSGIALVATGLAKLAETEGKGVFDEAKKTVEALNGYLDK
HKDEVDKASADNMVTNLLKHVAKAQELYKNSSALRAQGAQIDTVFSSYYWLY
KAGVTPETFPTVSQFLFELGKHPRGTKKMKKALLSTPMKWGKKLYELFADDSF
QQNRIYMHPAVLTAGRISEMGVCFGTIPVANPDDAALGSGHTKSILNLRTNTET
NNPCARTIVKLFEIQKTGFNIQDMDIVASEHLLHQSLVGKQSPFQNAYNVKGNA
TSANII*

Number of amino acids: 482

Molecular weight: 53958.53 Da

Theoretical pI: 8.46

Ext. coefficient: 73715 M^{-1} cm^{-1}

6xHis-SUMO-CCHFV NP

HHHHHHSSGLVPRGSHMSDSEVNQEAKPEVKPEVKPETHINLKVSDGSSEI
FFKIKKTTPLRRLMEAFAKRQGKEMDSLRFLYDGIRIQADQTPEDLDMEDNDIIE
AHREQIGGSMENKIEVNSKDEMNKWFEEFKKGNGLVDTYTNSYSFCESVPNLD
RFVFQMAGATDDAQKDSIYASALVEATKFCAPIYECAWASSTGIVKKGLEWFE
KNTGTIKSWDESYIELKVEVPKIEQLFNYQQAALKWRKDIGFRVNANTAALSNK
VLAEYKVPGEIVMSVKEMLSDMIRRRNLILNRGGDENPRGPVSHEHVEWCREF
VKGKYIMAFNPPWGDINKSGRSGIALVATGLAKLAETEGKGVFDEAKKTVEAL
NGYLDKHKDEVDKASADNMVTNLLKHVAKAQELYKNSSALRAQGAQIDTVFS
SYYWLYKAGVTPETFPTVSQFLFELGKHPRGTKKMKKALLSTPMKWGKKLYE
LFADDSFQQNRIYMHPAVLTAGRISEMGVCFGTIPVANPDDAALGSGHTKSILN
LRTNTETNNPCARTIVKLFEIQKTGFNIQDMDIVASEHLLHQSLVGKQSPFQNAY
NVKGNATSANII*

Number of amino acids: 597

Molecular weight: 67090.20

Theoretical pI: 7.09

Ext. coefficient: 75205 M^{-1} cm^{-1}

HAZV NP

MENKIVASTKEEFNTWYKQFAEKHKLNNKYTESASFCAEIPQLDTYKYKM
ELASTDNERDAIYSSALIEATRFCAPIMECAWASCTGTVKRGLEWFDKNKDSD
TVKVWDANYQKLRTETPPAEALLAYQKAALNWRKDVGFSIGEYTSILKKAVA
AEYKVPGTVINNIKEMLSDMIRRRNRIINGGSDDAPKRGPVGREHLDWCREFAS
GKFLNAFNPPWGEINKAGKSGYPLLATGLAKLVELEGKDVMDKAKASIAQLEG
WVKENKDQVDQDKAEDLLKGVRESYKTALALAKQSNAFRAQGAQIDTVFSSY
YWLWKAGVTPVTFPSVSQFLFELGKNPKGQKKMQKALINTPLKWGKRLIELFA
DNDFTENRIYMHPCVLTSGRMSELGISFGAVPVTSPDDAAQGSGHTKAVLNYK
TKTEVGNPCACIISSLFEIQKAGYDIESMDIVASEHLLHQSLVGKRSPFQNAYLIK
GNATNINII*

Number of amino acids: 485

Molecular weight: 54216.88

Theoretical pI: 8.73

Ext. coefficient: 86330 M^{-1} cm^{-1}

6xHis-SUMO-HAZV NP

HHHHHHSSGLVPRGSHMSDSEVNQEAKPEVKPEVKPETHINLKVSDGSSEI
FFKIKKTTPLRRLMEAFAKRQGKEMDSLRFLYDGIRIQADQTPEDLDMEDNDIIE
AHREQIGGSMENKIVASTKEEFNTWYKQFAEKHKLNNKYTESASFCAEIPQLDT
YKYKMELASTDNERDAIYSSALIEATRFCAPIMECAWASCTGTVKRGLEWFDK
NKDSDTVKVWDANYQKLRTETPPAEALLAYQKAALNWRKDVGFSIGEYTSILK
KAVAAEYKVPGTVINNIKEMLSDMIRRRNRIINGGSDDAPKRGPVGREHLDWC
REFASGKFLNAFNPPWGEINKAGKSGYPLLATGLAKLVELEGKDVMDKAKASI
AQLEGWVKENKDQVDQDKAEDLLKGVRESYKTALALAKQSNAFRAQGAQIDT
VFSSYYWLWKAGVTPVTFPSVSQFLFELGKNPKGQKKMQKALINTPLKWGKRL
IELFADNDFTENRIYMHPCVLTSGRMSELGISFGAVPVTSPDDAAQGSGHTKAV
LNYKTKTEVGNPCACIISSLFEIQKAGYDIESMDIVASEHLLHQSLVGKRSPFQN
AYLIKGNATNINII*

Number of amino acids: 600

Molecular weight: 67348.54

Theoretical pI: 7.98

Ext. coefficient: 87820 M^{-1} cm^{-1}

6xHis-Caspase-3

MENTENSVDSKSIKNLEPKIIHGSESMDSGISLDNSYKMDYPEMGLCIIINNK
NFHKSTGMTSRSGTDVDAANLRETFRNLKYEVRNKNDLTREEIVELMRDVSKE
DHSKRSSFVCVLLSHGEEGIIFGTNGPVDLKKITNFFRGDRCRSLTGKPKLFIIQA
CRGTELDCGIETDSGVDDDMACHKIPVEADFLYAYSTAPGYYSWRNSKDGSWF
IQSLCAMLKQYADKLEFMHILTRVNRKVATEFESFSFDATFHAKKQIPCIVSMLT
KELYFYHLEHHHHHH*

Number of amino acids: 285

Molecular weight: 32673.00

Theoretical pI: 6.27

Ext. coefficient: 26400 M^{-1} cm^{-1}

Affimer-NP-Cys-8xHis

MASNSLEIEELARFAVDEHNKKENALLEFVRVVKAKEQMHMKERQINTM
YYLTLEAKDGGKKKLYEAKVWVKKYLGDFWKDNFKELQEFKPVGDACAAAH
HHHHHHH*

Number of amino acids: 107

Molecular weight: 12671.50

Theoretical pI: 7.91

Ext. coefficient: 16960 M^{-1} cm^{-1}

Affimer-NP

NSLEIEELARFAVDEHNKKENALLEFVRVVKAKEQMHMKERQINTMYYLT
LEAKDGGKKKLYEAKVWVKKYLGDFWKDNFKELQXFKPV*

Number of amino acids: 89

Molecular weight: 10707.64

Theoretical pI: 9.08

Ext. coefficient: 16960 M^{-1} cm^{-1}

6xHis-SUMO-Affimer-NP

HHHHHHSSGLVPRGSHMSDSEVNQEAKPEVKPEVKPETHINLKVSDGSSEI
FFKIKKTTPLRRLMEAFAKRQGKEMDSLRFLYDGIRIQADQTPEDLDMEDNDIIE
AHREQIGGSASNSLEIEELARFAVDEHNKKENALLEFVRVVKAKEQMHMKERQ
INTMYYLTLEAKDGGKKKLYEAKVWVKKYLGDFWKDNFKELQXFKPV*

Number of amino acids: 206

Molecular weight: 23997.47

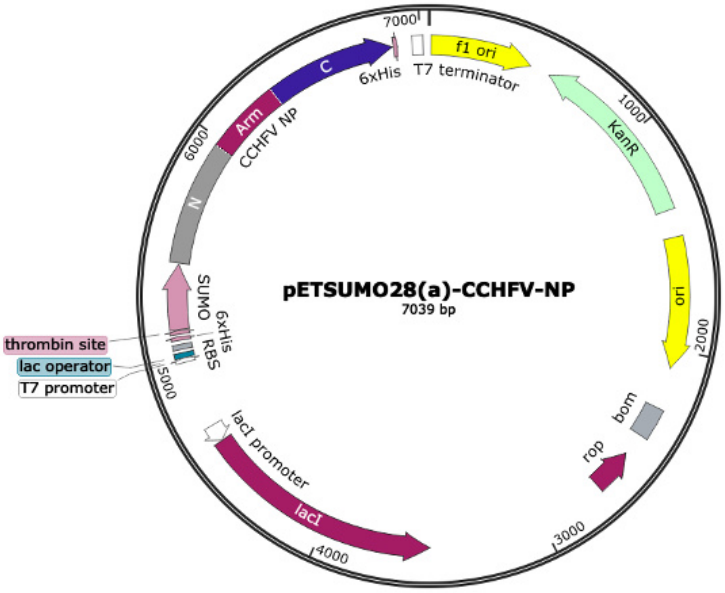
Theoretical pI: 6.65

Ext. coefficient: 18450 M^{-1} cm^{-1}

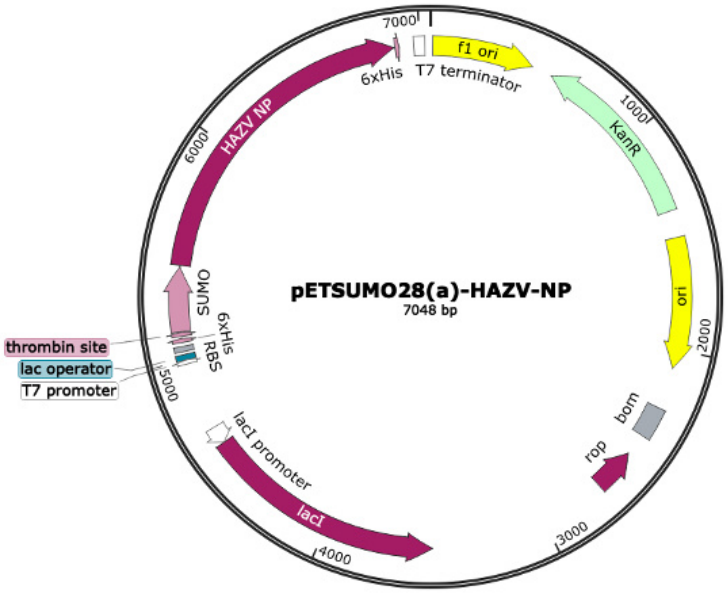
Appendix B

Plasmid maps

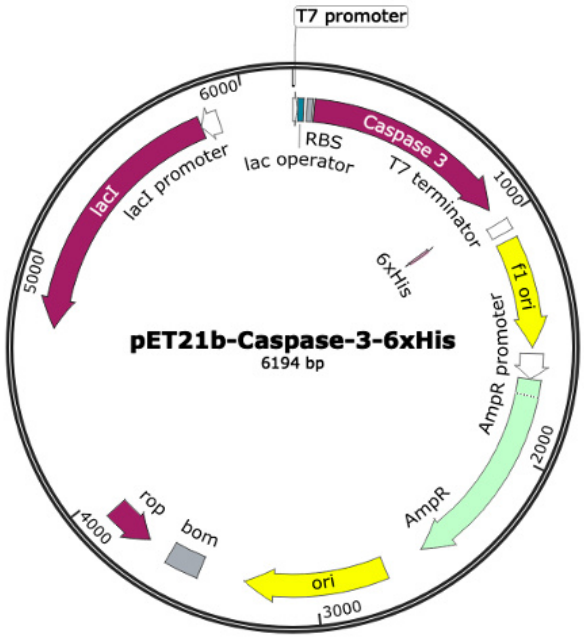
Plasmid maps of expression vectors for recombinant production, Affimer expression, HAZV rescue system and CCHFV replicon system (see plasmids table in section 2.1.2) are detailed below in figure B.1.



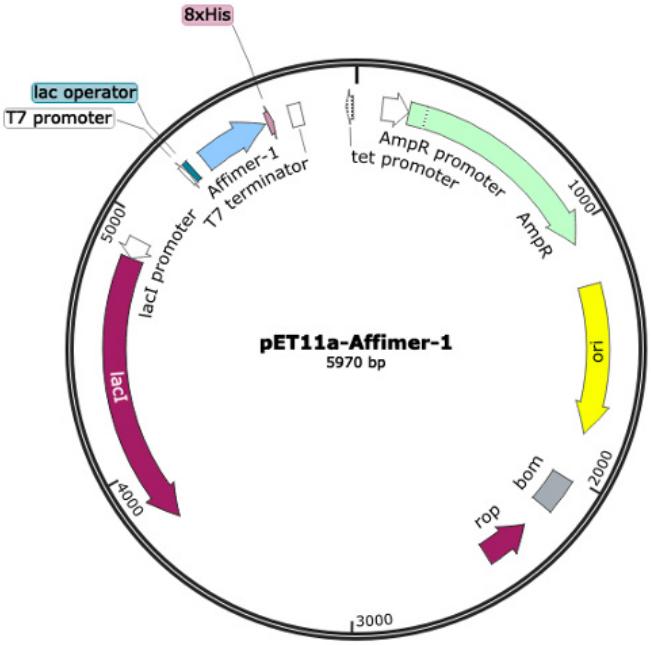
(a)


pETSUMO28(a)-HAZV-NP
7048 bp
7000
f1 ori
6xHis
T7 terminator
1000
KanR
6000
HAZV NP
ori
2000
SUMO
6xHis
RBS
thrombin site
lac operator
T7 promoter
5000
bom
rop
lacI promoter
lacI
4000
3000

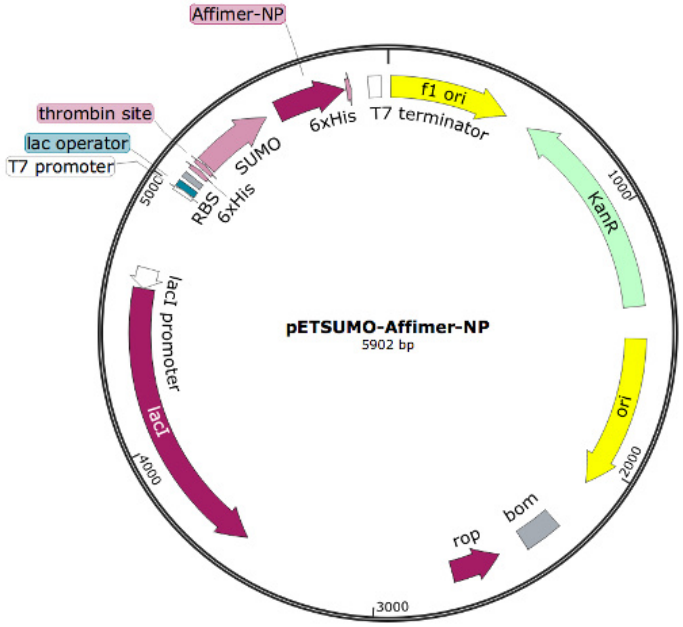
(b)


T7 promoter
6000
RBS
Caspase 3
lac operator
T7 terminator
1000
lacI
lacI promoter
6xHis
f1 ori
5000
pET21b-Caspase-3-6xHis
6194 bp
AmpR promoter
2000
AmpR
rop
4000
bom
ori
3000

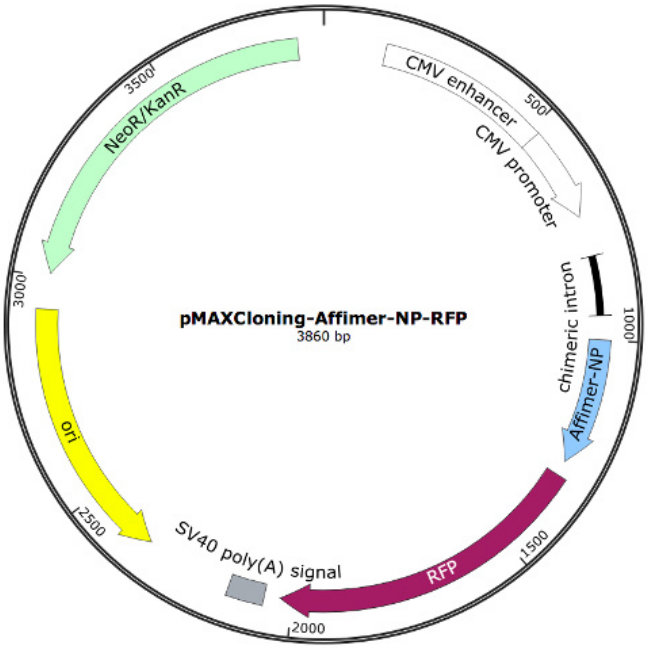
(c)


8xHis
lac operator
T7 promoter
Affimer-1
T7 terminator
AmpR promoter
tet promoter
AmpR
1000
lacI promoter
5000
lacI
pET11a-Affimer-1
5970 bp
ori
bom
rop
2000
4000
3000

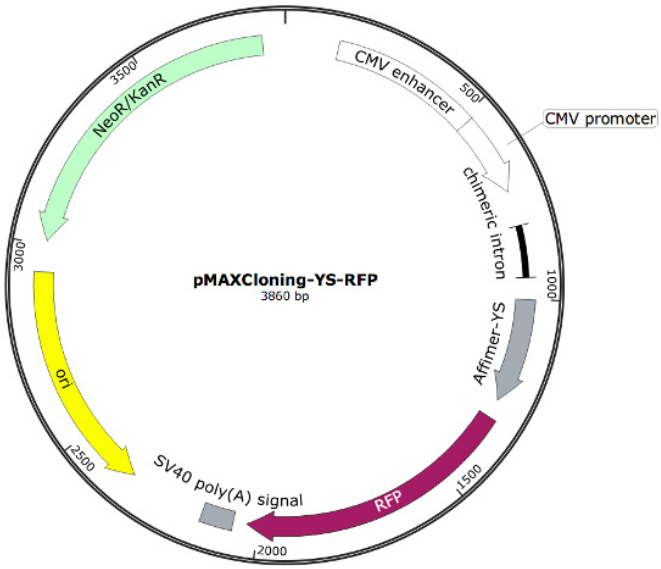
(d)


Affimer-NP
thrombin site
lac operator
T7 promoter
SUMO
6xHis
f1 ori
T7 terminator
RBS
6xHis
KanR
1000
5000
lacI promoter
lacI
pETSUMO-Affimer-NP
5902 bp
ori
bom
rop
2000
4000
3000

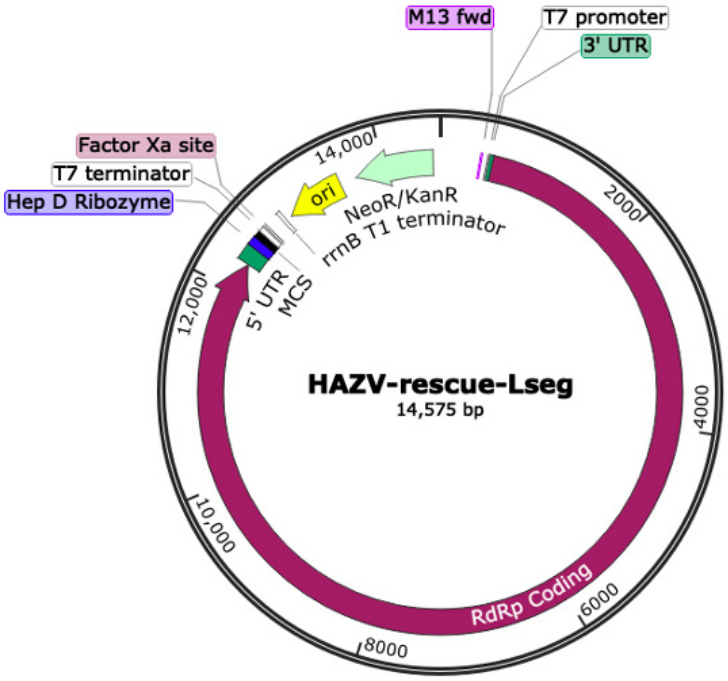
(e)


pMAXCloning-Affimer-NP-RFP
3860 bp
CMV enhancer
CMV promoter
chimeric intron
Affimer-NP
RFP
SV40 poly(A) signal
ori
NeoR/KanR
500
1000
1500
2000
2500
3000
3500

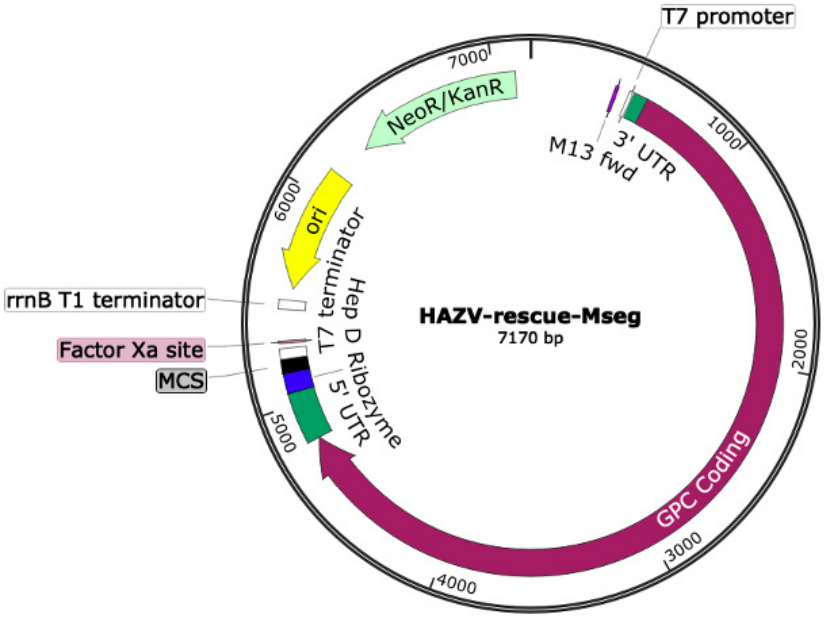
(f)


pMAXCloning-YS-RFP
3860 bp
CMV enhancer
CMV promoter
chimeric intron
Affimer-YS
RFP
SV40 poly(A) signal
ori
NeoR/KanR
500
1000
1500
2000
2500
3000
3500

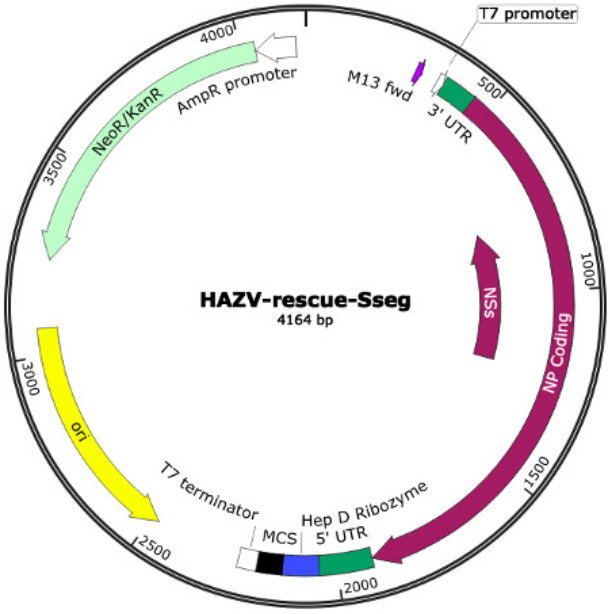
(g)


M13 fwd
T7 promoter
3' UTR
Factor Xa site
T7 terminator
Hep D Ribozyme
14,000
ori
NeoR/KanR
rrnB T1 terminator
2000
12,000
5' UTR
MCS
HAZV-rescue-Lseg
14,575 bp
4000
10,000
RdRp Coding
6000
8000

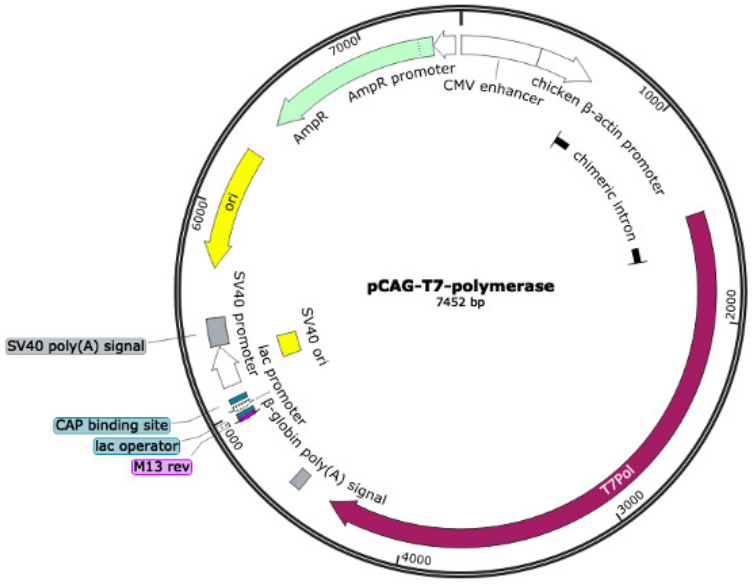
(h)


T7 promoter
7000
NeoR/KanR
3' UTR
M13 fwd
1000
6000
ori
T7 terminator
Hep D Ribozyme
rrnB T1 terminator
HAZV-rescue-Mseg
7170 bp
Factor Xa site
MCS
5' UTR
5000
2000
GPC Coding
3000
4000

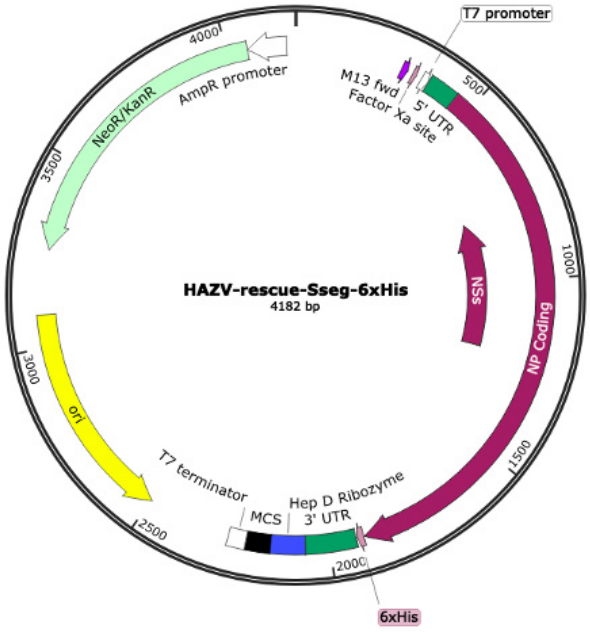
(i)


T7 promoter
4000
AmpR promoter
M13 fwd
500
3' UTR
3500
NeoR/KanR
HAZV-rescue-Sseg
4164 bp
SSN
1000
NP Coding
3000
ori
T7 terminator
1500
Hep D Ribozyme
MCS
5' UTR
2500
2000

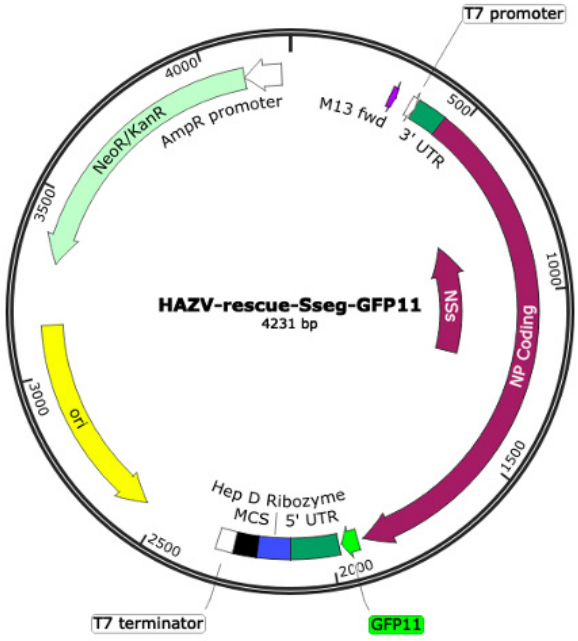
(j)


7000
AmpR promoter
AmpR
CMV enhancer
chicken β-actin promoter
1000
chimeric intron
ori
6000
pCAG-T7-polymerase
7452 bp
SV40 promoter
SV40 ori
2000
SV40 poly(A) signal
lac promoter
CAP binding site
β-globin poly(A) signal
lac operator
M13 rev
T7Pol
3000
4000

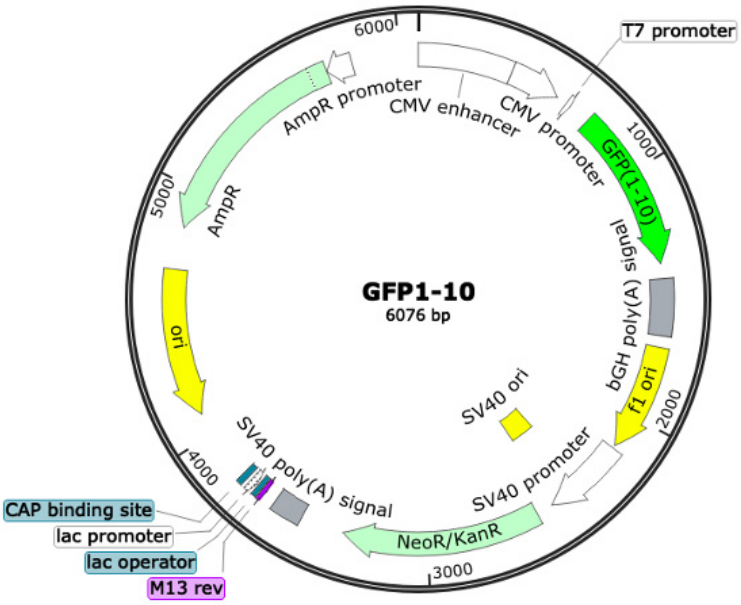
(k)


T7 promoter
4000
AmpR promoter
NeoR/KanR
3500
M13 fwd
Factor Xa site
5' UTR
500
1000
NP Coding
1500
SSN
HAZV-rescue-Sseg-6xHis
4182 bp
3000
ori
T7 terminator
Hep D Ribozyme
MCS
3' UTR
2500
2000
6xHis

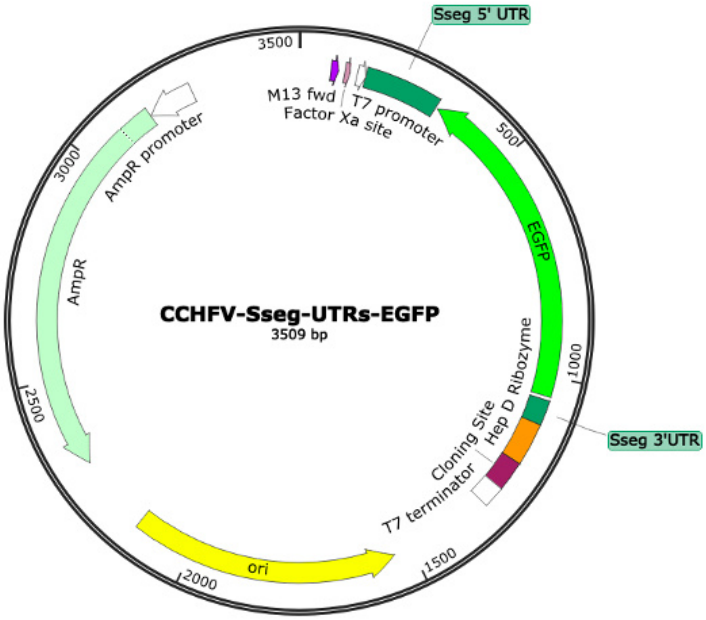
(l)


T7 promoter
4000
AmpR promoter
NeoR/KanR
3500
M13 fwd
3' UTR
500
1000
NP Coding
1500
SSN
HAZV-rescue-Sseg-GFP11
4231 bp
3000
ori
Hep D Ribozyme
MCS
5' UTR
2500
2000
T7 terminator
GFP11

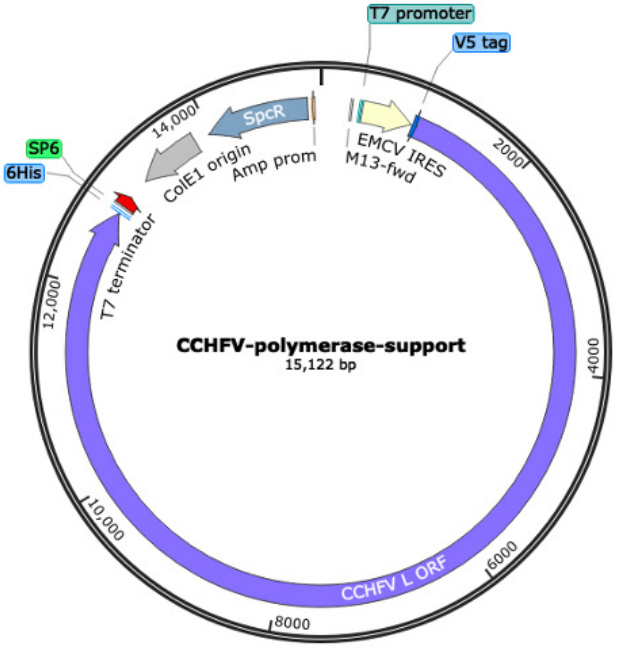
(m)


GFP1-10
6076 bp
6000
1000
2000
3000
4000
5000
T7 promoter
CMV enhancer
CMV promoter
GFP(1-10)
bGH poly(A) signal
f1 ori
SV40 ori
SV40 promoter
NeoR/KanR
SV40 poly(A) signal
AmpR promoter
AmpR
ori
CAP binding site
lac promoter
lac operator
M13 rev

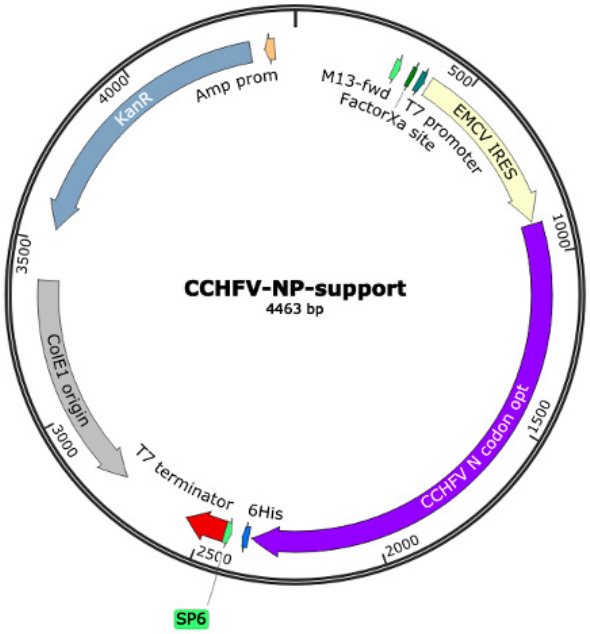
(n)


CCHFV-Sseg-UTRs-EGFP
3509 bp
3500
500
1000
1500
2000
2500
3000
Sseg 5' UTR
M13 fwd
Factor Xa site
T7 promoter
EGFP
Hep D Ribozyme
Cloning Site
T7 terminator
Sseg 3'UTR
ori
AmpR
AmpR promoter

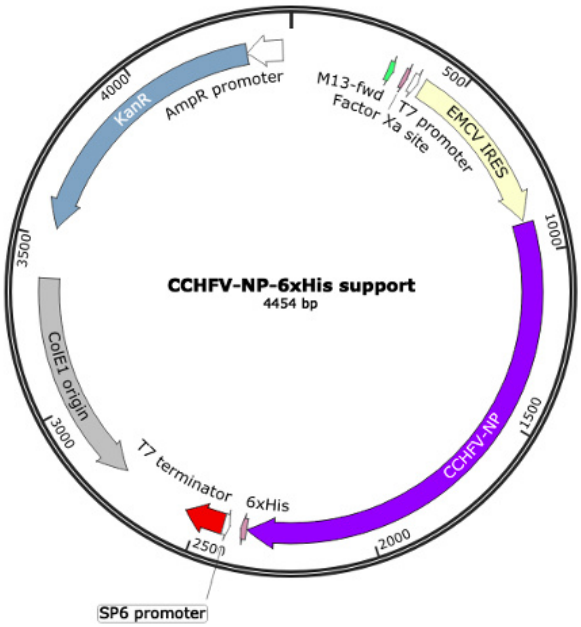
(o)


T7 promoter
V5 tag
SpcR
EMCV IRES
M13-fwd
ColE1 origin
Amp prom
SP6
6His
T7 terminator
CCHFV-polymerase-support
15,122 bp
CCHFV L ORF
14,000
2000
4000
6000
8000
10,000
12,000

(p)


KanR
Amp prom
M13-fwd
FactorXa site
T7 promoter
EMCV IRES
CCHFV-NP-support
4463 bp
ColE1 origin
T7 terminator
6His
CCHFV N codon opt
SP6
500
1000
1500
2000
2500
3000
3500
4000

(q)



(r)

Figure B.1: Plasmid maps. **(a-c)** Expression vectors for recombinant production. **(d-g)** Affimer expression vectors. **(h-n)** HAZV rescue system expression vectors. **(o-r)** CCHFV replicon system expression vectors.