

Aus dem Institut für Virologie
Geschäftsführender Direktor: Prof. Dr. Stephan Becker
des Fachbereichs Medizin der Philipps-Universität Marburg

**Identification of lead molecules for the development of antivirals
targeting the Ebola virus matrix protein VP40**

Inaugural-Dissertation
zur Erlangung des Doktorgrades
der Naturwissenschaften (Dr. rer. nat.)

dem Fachbereich Medizin der Philipps-Universität Marburg
vorgelegt von

Anke-Dorothee Werner
aus Schwäbisch Hall

Marburg, 2022

Angenommen vom Fachbereich Medizin der Philipps-Universität Marburg am: 20.12.2022

Gedruckt mit Genehmigung des Fachbereichs Medizin

Dekanin: Prof. Dr. Denise Hilfiker-Kleiner

Referentin: Prof. Dr. Stephan Becker

Korreferent: Prof. Dr. Oliver Hantschel

Für mein Vorbild.

Table of contents

List of tables	vii
List of figures	vii
List of Abbreviations.....	x
1 Introduction.....	1
1.1 Taxonomy of filoviruses	1
1.2 Epidemiology.....	1
1.3 Transmission and Clinical Picture.....	2
1.4 Vaccines and Treatments	3
1.5 Morphology and genome organization.....	4
1.6 Viral replication cycle	4
1.7 VP40.....	6
1.7.1 Structure and functions.....	6
1.7.2 VP40 as a drug target	9
1.8 Protein-protein interactions as targets for antiviral intervention	10
1.9 Aims of this work.....	11
2 Methods	13
2.1 Molecular biology methods.....	13
2.1.1 Polymerase Chain reaction.....	13
2.1.2 Preparative and analytical restriction digest.....	14
2.1.3 Preparative and analytical agarose gels electrophoresis	15
2.1.4 Purification of PCR products.....	15
2.1.5 Ligation	15
2.1.6 Site-directed mutagenesis.....	16
2.1.7 Preparation of chemically Z-competent bacteria.....	17
2.1.8 Transformation of chemically competent bacteria.....	17
2.1.9 Preparation of plasmid DNA from bacteria.....	17
2.1.10 Sequencing	18
2.2 Cell culture methods	18
2.2.1 Cultivation of cell lines	18
2.2.2 Transfection of DNA	18
2.2.3 Transfection of peptides.....	18
2.3 Virological methods.....	19
2.3.1 Minigenome assay.....	19
2.3.2 Virus-like particle assays.....	20
2.3.3 Infection of cells with EBOV	21

2.3.4	Harvest of (recombinant) EBOV stocks	21
2.3.5	Extraction of viral RNA from cell culture supernatant	21
2.3.6	Determination of viral titers.....	22
2.3.7	Cloning and rescue of recombinant EBOV	22
2.3.8	Growth curves of infected cells.....	23
2.3.9	Immunofluorescence analysis	23
2.3.10	Transmission electron microscopy and analysis of particle length of recEBOV.....	24
2.3.11	Immunogold labeling of VP40 on whole mounted cells.....	24
2.4	Protein biochemistry	24
2.4.1	Expression in <i>E. coli</i>	24
2.4.2	Protein purification.....	25
2.4.3	Sodium dodecyl sulfate-polyacrylamide gel electrophoresis.....	26
2.4.4	Protein crystallography	27
2.4.5	Disulfide Tethering	35
2.4.6	Molecular docking	38
2.4.7	Thermal shift assay.....	38
2.4.8	Mass spectrometry.....	39
2.4.9	Surface Plasmon Resonance.....	42
2.4.10	Biotin-Switch assay	43
2.4.11	Redox assays.....	44
3	Results	45
3.1	Part I: Identification and characterization of residues involved in protein-protein interactions.....	45
3.1.1	Characterization of selected hot spot amino acids of sVP40 homo-oligomerization ...	45
3.1.2	Previously unresolved C-terminus of Sudan Ebolavirus VP40 exhibits functionally important post-translational redox modifications.....	56
3.2	PartII: Targeting VP40's homo-oligomerization via interface-mimicking peptides.....	71
3.2.1	Introduction.....	71
3.2.2	Results	71
3.2.3	Discussion	78
3.3	Part III: Fragment-based lead discovery.....	81
3.3.1	Site-directed ligand discovery via disulfide tethering	81
3.3.2	Lead discovery via crystal soaking.....	89
4	Summary discussion and outlook.....	118
5	Summary.....	120
7	Zusammenfassung.....	122
8	Materials.....	125

8.1	Cells and viruses	125
8.1.1	Mammalian cells.....	125
8.1.2	Viruses	125
8.1.3	<i>E. coli</i>	125
8.2	Vectors.....	125
8.2.1	Mammalian expression vectors	125
8.2.2	Bacterial expression vectors.....	126
8.3	Oligonucleotides and DNA	127
8.3.1	Primers	127
8.3.2	Nucleic acids and nucleotides	130
8.4	Consumables	130
8.5	Chemicals.....	131
8.6	Buffers, media and gels	132
8.6.1	Buffers and solutions.....	132
8.6.2	Media.....	134
8.6.3	Gels	135
8.7	Antibodies.....	135
8.7.1	Primary antibodies	135
8.7.2	Secondary antibodies	135
8.8	Enzymes.....	135
8.9	Kits	136
8.10	Peptides.....	136
8.11	Equipment	136
8.12	Software	138
8.13	Fragments.....	138
8.13.1	Thiol-containing fragments for DT	138
8.13.2	Fragments for crystal soaking.....	141
9	References.....	a
10	Appendix.....	r
10.1	Supplemental material	r
10.2	Publications, Presentations and Poster.....	ee
10.3	<i>Curriculum vitae</i>	ii
10.4	Ehrenwörtliche Erklärung.....	jj
10.5	Acknowledgements	kk

List of tables

Table 1: DT screening protocol #1. _____	36
Table 2: DT screening protocol #2. _____	37
Table 3: DT optimization protocol. _____	37
Table 4: Data collection and refinement statistics of sVP40 _{Δ43} W95A and L117A. _____	48
Table 5: Data collection and refinement statistics of dimeric sVP40 _{Δ43} WT. _____	57
Table 6: Data collection and refinement statistics of dimeric sVP40 _{Δ43} C314S (AW422) and sVP40 _{Δ43} CCS (AW424). _____	59
Table 7: Sequence of the four core interface-mimicking peptides. _____	71
Table 8: Designed and synthesized peptides. _____	72
Table 9: Immobilisation conditions and final RU. _____	73
Table 10: Overview of sVP40 _{Δ43} WT-peptide crystals. _____	77
Table 11: DT screening hits for sVP40 _{Δ43} N67C. _____	84
Table 12: DT screening hits for sVP40 _{Δ43} N67C-CCA. _____	86
Table 13: Comparison of DT screening hits between sVP40 _{Δ43} N67C and sVP40 _{Δ43} N67C-CCA. _____	87
Table 14: DT optimization results for TK142 and SB100 with sVP40 _{Δ43} N67C-CCA. _____	87
Table 15: DMSO soaking of dimeric sVP40 _{Δ43} WT crystals. _____	92
Table 16: Overview of crystal soaking datasets and refinement strategy. _____	93
Table 17: Data collection and refinement statistics of dimeric sVP40 _{Δ43} WT in complex with SA. _____	97
Table 18: Crystal soaking with SA-derivatives. _____	106
Table 19: Data collection and refinement statistics of dimeric sVP40 _{Δ43} WT in complex with SA-derivatives. _____	108
Table 20: Fragments of the Jena FragXtal Screen with similar building blocks than SA. _____	113
Table 21: Disruptive fragments of the FragXtal screen. _____	114
Table 22: Overview of data obtained by HDX-MS. _____	s
Table 23: Determined affinities [mM] of various peptides to target proteins using SPR. _____	u
Table 24: Selection of the most promising compounds after docking to sVP40. _____	cc

List of figures

Figure 1: Taxonomy of Mononegavirales and Filoviridae. _____	1
Figure 2: Cases and outbreaks of reported EVD caused by Ebola virus species. _____	2
Figure 3: Genome organization and morphology of Ebola virus particles. _____	4
Figure 4: EBOV replication cycle. _____	6
Figure 5: Schematic representation of VP40. _____	7
Figure 6: Work plan of the present thesis following the typical drug design steps. _____	12
Figure 7: EBOV-specific minigenome assay. _____	20
Figure 8: Phase diagram of protein crystallization. _____	28
Figure 9: Vapor diffusion sitting vs. hanging drop. _____	28
Figure 10: Three-dimensional crystal lattice and the building blocks of a protein crystal. _____	30
Figure 11: Schematic of Bragg's law. _____	31
Figure 12: Refinement process for the present work. _____	35
Figure 13: Crystal structures of dimeric and octameric zVP40. _____	46
Figure 14: sVP40 WT and mutants influence reporter gene activity of a MG assay. _____	46
Figure 15: Effect of sVP40 W95A and L117A mutations on VP40-release. _____	47
Figure 16: Assessment of homo-oligomeric state and stability of sVP40 _{Δ43} WT, W95A and L117A via SEC. _____	48
Figure 17: Alignment of sVP40 _{Δ43} W95A and L117A crystal structures with sVP40 _{Δ43} WT. _____	49

Figure 18: HDX-MS for sVP40 _{Δ43} L117A compared to sVP40 _{Δ43} WT.	50
Figure 19: HDX-MS for sVP40 _{Δ43} W95A compared to sVP40 _{Δ43} WT.	51
Figure 20: Thermal stability of sVP40 hot spot mutants.	52
Figure 21: Position and orientation of residue W95 in different VP40 structures.	54
Figure 22: Sequence and structure of the last residues of the VP40-CTD.	57
Figure 23: The mutation of sVP40 _{Δ43} cysteines does not alter homo-oligomerization.	59
Figure 24: sVP40 cysteine mutants influence reporter gene activity in a MG assay and result in a decrease in VLP formation.	61
Figure 25: sVP40 cysteine mutation C320S results in the formation of longer filaments.	62
Figure 26: RecoEBOV carrying VP40 cysteine mutations exhibits slower growth kinetics than the WT.	63
Figure 27: The mutation of cysteines in VP40 of recEBOV leads to the formation of longer filaments in infected cells compared to recEBOV WT-infected cells.	64
Figure 28: Post-translational modification of VP40.	65
Figure 29: VP40 reduction by the thioredoxin reduction system.	66
Figure 30: Comparison of the asymmetric unit of sVP40 _{Δ43} with one protomer of each zVP40 _{Δ43} crystal structure.	68
Figure 31: Structure-based design of inhibitory peptides.	72
Figure 32: Identification of promising peptides using the SPR Single-Level Binding Screen.	73
Figure 33: Co-expression of zVP40 and the core peptides in a MG setting.	75
Figure 34: Effect of transfected peptides on EBOV titers.	75
Figure 35: Dimer/octamer ratio of VP40 co-expressed with different peptides.	76
Figure 36: SEC profile of the sVP40 _{Δ43} -pep1 fusion protein.	77
Figure 37: Principle of lead discovery via DT.	82
Figure 38: Size-exclusion profile of engineered sVP40 _{Δ43} mutants.	83
Figure 39: IP-MS spectra of sVP40 _{Δ43} N67C.	84
Figure 40: IP-MS spectra of sVP40 _{Δ43} N67C-CCA.	86
Figure 41: SEC profile of zVP40 full length WT and sVP40 _{Δ43} WT.	90
Figure 42: Crystals of zVP40 _{Δ43} and sVP40 _{Δ43} dimer.	91
Figure 43: Structure of dimeric sVP40 _{Δ43} WT soaked with SA.	97
Figure 44: Thermal stability of dimeric sVP40 _{Δ43} WT with NaSA.	99
Figure 45: Differences in D ₂ O uptake between dimeric sVP40 _{Δ43} WT in its apo form and in complex with NaSA.	99
Figure 46: NaSA displays only minor effects on sVP40 in functional cell culture assays.	100
Figure 47: Influence of sVP40 pocket mutants on reporter gene activity of a MG assay.	101
Figure 48: Effect of sVP40 pocket mutants on sVP40-release.	102
Figure 49: SEC of sVP40 _{Δ43} pocket mutants.	102
Figure 50: sVP40 R214 sidechain conformations.	105
Figure 51: Docking poses of three exemplary molecules.	105
Figure 52: Structure of dimeric sVP40 _{Δ43} WT with soaked SA-derivatives.	107
Figure 53: LL060 disrupts sVP40's ability to bud VLPs but not its effect on viral RNA synthesis.	110
Figure 54: B-factors analysis of 8B2U and 4LD8 crystal structures.	111
Figure 55: Crystal packing of sVP40 _{Δ43} WT in complex with SA.	113
Figure 56: Model of VP40 inhibition upon binding of a compound in the pocket between the NTD and CTD.	117
Figure 57: Cross-project results for the present work.	119
Figure 58: Comparison of the VP40 sequences.	r
Figure 59: sVP40 WT and mutants influence reporter gene activity of a MG assay.	s

Figure 60: Peptide coverage map of HDX-MS experiments comparing sVP40 _{Δ43} WT with W945A and L117A.	t
Figure 61: HDX-MS of sVP40 _{Δ43} mutants depicting bimodality.	t
Figure 62: Stability of sVP40 _{Δ43} W95A and L117A via SEC.	u
Figure 63: Electron density of dimeric sVP40 _{Δ43} WT (AW122).	u
Figure 64: Size-exclusion profile of the sVP40-linker fusion protein.	v
Figure 65: ESI mass spectra of DT hits for sVP40 _{Δ43} N67C of reaction 1 to 5.	w
Figure 66: ESI mass spectra of DT hits for sVP40 _{Δ43} N67C of reaction 6 to 10.	x
Figure 67: ESI mass spectra of DT hits for sVP40 _{Δ43} N67C of reaction 11 to 19.	y
Figure 68: ESI mass spectra of DT hits for sVP40 _{Δ43} N67C of reaction 22 to 27.	z
Figure 69: ESI mass spectra of DT hits for sVP40 _{Δ43} WT negative control without fragments.	z
Figure 70: ESI mass spectra of DT hits for sVP40 _{Δ43} N67C-CCA of reaction 1 to 4.	aa
Figure 71: ESI mass spectra of DT hits for sVP40 _{Δ43} N67C-CCA of reaction 6 to 8.	bb
Figure 72: Structure of dimeric sVP40 _{Δ43} WT with co-crystallized SA.	cc
Figure 73: Peptide coverage map of HDX-MS experiments comparing sVP40 _{Δ43} WT with and without SA.	cc

List of Abbreviations

6x histidine tag	his-tag
Adsorption, distribution, metabolism and excretion	ADME
Amino acid	aa
Anti	α
Astex' rule of three	Ro3
Asymmetric unit	ASU
Base pair	bp
Biosafety level	BSL
Bombali Ebola virus	BOMV
Bovine serum albumin	BSA
Bundibugyo Ebola virus	BDBV
C-terminal domain	CTD
Calf Intestinal Phosphatase	CIP
Complementary DNA	cDNA
Cytopathic effect	CPE
Dalton	Da
Deletion	Δ
Democratic Republic of Congo	DRC
Desoxyribonucleic acid	DNA
Disulfide Tethering	DT
Dithiothreitol	DTT
Ebola virus disease	EVD
Electrospray ionisation	ESI
Endoplasmatic reticulum	ER
Endosomal sorting complex required for transport	ESCRT
Enhanced high definition MS	HDMS ^E
Enzyme-linked immunosorbent assays	ELISA
Ethylenediaminetetraacetic acid	EDTA
Flavin adenine dinucleotide	FAD
Fetal bovine serum	FBS
Fragment-based lead discovery	FBLD
Fragment-based drug design	FBDD

Glycoprotein	GP
Glutathione, oxidized	GSSG
Glutathione, reduced	GSH
Glutathione reductase	GR
High definition MS	HDMS
Human immunodeficiency viruses	HIV
Immobilized metal affinity chromatography	IMAC
Immunofluorescence analysis	IFA
Infectious virus-like particle	iVLP
Iodoacetamide	IAA
Isopropyl-beta-D-thiogalactopyranoside	IPTG
Isothermal calorimetry	ITC
Late budding domain	L domains
Lipinski's rule of five	Ro5
Lysogeny broth	LB
Marburg virus	MARV
Marburg virus disease	MARD
Marburg virus VP40	mVP40
Melanoma differentiation-associated protein 5	MDA-5
messengerRNA	mRNA
Microscale thermophoresis	MST
Minigenome	MG
Milli Absorbance Unit	mAU
Molecular operating environment	MOE
Molecular Replacement	MR
Monoclonal antibodies	mAbs
Multi-wavelength anomalous diffraction	MAD
Multiplicity of infection	MOI
N-terminal domain	NTD
Neuronal precursor cell-expressed developmentally downregulated 4	Nedd4
New England Biolabs	NEB
Nickel-nitrilotriacetic acid	Ni-NTA
Nicotinamide adenine dinucleotide phosphate	NADPH

Niemann-Pick C1	NPC1
Non-human primate	NHP
Nuclear Magnetic Resonance	NMR
Nucleocapsid complex	NC
Nucleoprotein	NP
Optical density	OD
Pan-Dataset Density Analysis	PanDDA
Phosphate buffered saline	PBS
Polyacrylamide gel electrophoresis	PAGE
Polymerase Chain Reaction	PCR
Polyribonucleotidyl transferase	PRNTase
Post infection	pi
Post transfection	pt
Protein Data Bank	PDB
Quantitative reverse transcription PCR	qRT-PCR
Response units	RU
Reston Ebolavirus	RESTV
Retinoic-acid-inducible gene I	RIG-I
Reverse transcriptase	RT
Ribonucleic acid	RNA
Ribonucleoprotein	RNP
RNA-dependent RNA-polymerase	RdRp
Room temperature	RT
S-nitrosocysteine	Cys-NO
Sample buffer	SB
Salicylic acid	SA
Severe acute respiratory syndrome coronavirus 1	SARS-CoV
Single-stranded RNA	ssRNA
Single-wavelength anomalous diffraction	SAD
Site-directed ligand discovery	SDLD
Size-exclusion chromatography	SEC
Size-exclusion chromatography coupled to multi-angle light scattering	SEC-MALS
Small soluble GP	ssGP

Sodium dodecyl sulfate	SDS
Sodium salicylate	NaSA
Solid-phase peptide synthesis	SPPS
Soluble GP	sGP
Sudan Ebola virus	SUDV
Sudan Ebola virus VP40	sVP40
Surface Plasmon Resonance	SPR
Swiss Light Source	SLS
Tai Forest Ebola virus	TAFV
Thioredoxin	Trx
Thioredoxin reductase	TrxR
Transmission electron microscopy	TEM
Tumor necrosis factor α converting enzyme	TACE
Tumor susceptibility gene 101	tsg101
Viral Protein	VP
Viral RNA	vRNA
Virus-like particle	VLP
Volts	V
Western blot	WB
Zaire Ebola virus	EBOV
Zaire Ebola virus VP40	zVP40

Amino Acid Abbreviations

Alanine	Ala	A	Glycine	Gly	G	Proline	Pro	P
Arginine	Arg	R	Histidine	His	H	Serine	Ser	S
Asparagine	Asn	N	Isoleucine	Ile	I	Threonine	Thr	T
Aspartic acid	Asp	D	Leucine	Leu	L	Tryptophan	Trp	W
Cysteine	Cys	C	Lysine	Lys	K	Tyrosine	Tyr	Y
Glutamine	Gln	Q	Methionine	Met	M	Valine	Val	V
Glutamic acid	Glu	E	Phenylalanine	Phe	F	Any amino acid		X

1 Introduction

1.1 Taxonomy of filoviruses

Viruses within the family of *Filoviridae* exhibit a typical rod-like structure and contain a single-stranded, non-segmented negative-sense RNA genome. The family *Filoviridae* belongs to the order *Mononegavirales* as do the families of *Paramyxoviridae*, *Rhabdoviridae* and *Bornaviridae*. The genera Cuevavirus, Dianlovirus, Marburgvirus and Ebolavirus are part of the filovirus family, Ebolavirus being further divided into the species Zaire ebolavirus (EBOV), Sudan ebolavirus (SUDV), Reston ebolavirus (RESTV), Taï Forest ebolavirus (TAFV), Bundibugyo ebolavirus (BDBV) and Bombali ebolavirus (BOMV) (Figure 1) [316].

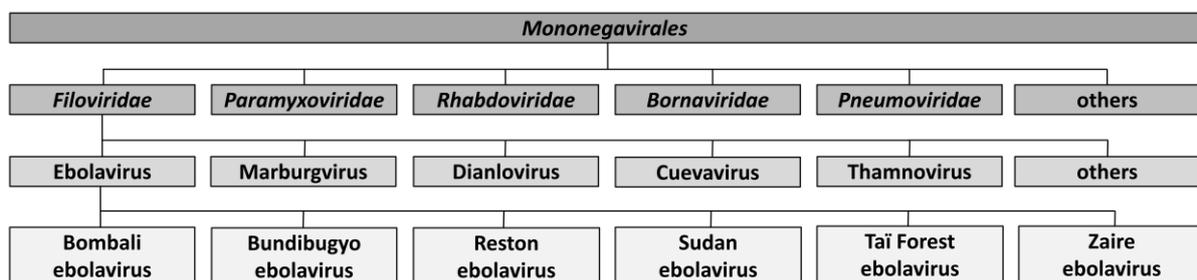


Figure 1: Taxonomy of *Mononegavirales* and *Filoviridae*. The order *Mononegavirales* (top level) contains several viral families (second level from top) besides *Filoviridae*. Ebolavirus (third level from top) is further divided various species (bottom level) [316].

1.2 Epidemiology

Filoviruses were first described in 1967 in Marburg when laboratory workers developed a severe haemorrhagic fever after handling organs and tissues of vervet monkeys (*Cercopithecus aethiops*). This led to the discovery of Marburg virus (MARV) as the causative agent of the Marburg virus disease (MARD) [215]. In the following decades, several outbreaks of MARD were reported. Its close relative, Ebola virus, was first discovered in 1976 during nearly simultaneous outbreaks of EBOV and SUDV in Zaire (now Democratic Republic of Congo, DRC) and Sudan [81,212]. Since then, numerous Ebola virus disease (EVD) outbreaks have been recorded (Figure 2). Most outbreaks occur in Africa, where the presumed natural reservoir hosts of EBOV, bats of the species *Hypsignathus monstrosus*, *Epomops franqueti* and *Myonycteris torquata*, are endemic [193,194,256]. EBOV and SUDV are the most common causes of EVD, while TAFV and BDBV were responsible for only a few human cases [98]. RESTV only occurred outside of Africa and caused asymptomatic human infections [47].

Introduction

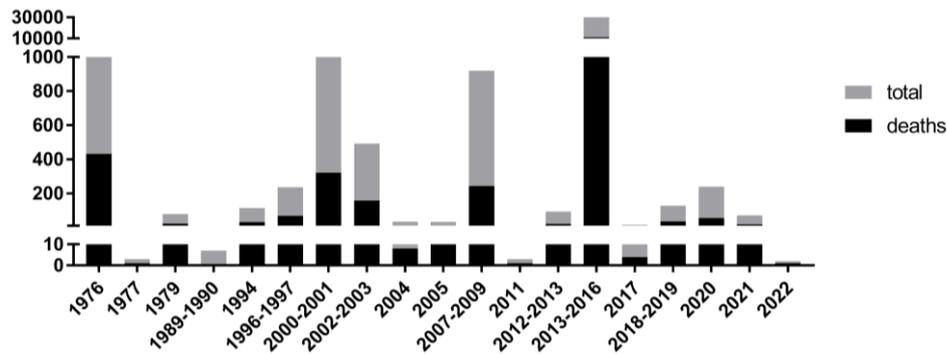


Figure 2: Cases and outbreaks of reported EVD caused by Ebola virus species. No differentiation between symptomatic and asymptomatic cases (the latter for RESTV only) were made. Outbreaks without confirmed human cases are not listed. Modified [51].

1.3 Transmission and Clinical Picture

Transmission Typically, a single spillover event from the natural reservoir into the human population occurs with subsequent human-to-human transmission via infectious body fluids [19,22,262]. Human-to-human transmission is often described to occur during traditional funerals, when EVD patients are being cared for at home and in health-care facilities especially in the beginning when the outbreak has not yet been recognized. Furthermore, sexual transmission has also been described [217]. Airborne transmission has been only reported in animal experiments under special experimental settings [158]. Handling and consumption of bushmeat is considered as the main source of spill-over events [184].

Clinical symptoms After an incubation period of 2 to 21 days, EVD patients first develop a nonspecific febrile illness with subsequent symptoms such as fatigue, anorexia, diarrhoea, arthralgia, headache, abdominal pain, vomiting, myalgia and rashes [59,331]. Later, patients with a fatal outcome develop haemorrhages, shock and multiorgan failure that finally leads to death. The term “Ebola haemorrhagic fever” is considered outdated because haemorrhages occur in only a subset of EVD patients [152,185]. Another feature of acute EVD is the infection of brain [30], eyes [284] and testes [74] and the persistence of the virus in these immune-privileged tissues. Persistence can lead to long-term sequelae such as ocular deficits, headache, depression and insomnia, which is described as the “post Ebola syndrome” [31,63], and in rare cases leads to a relapse of the disease [284]. Infectious viral particles and/or viral RNA were also detected in breast milk [246], semen [74] and cerebrospinal fluid months after the original infection [153].

Filoviruses infect a number of host cells, with macrophages, monocytes and dendritic cells being the primary target cells [36,92,109,279]. By using the lymphatic system, the virus then spreads quickly through the organism [107]. Activated macrophages secrete cytokines such as interleukins IL-1 β , IL-6 and IL-8, which eventually results in endothelial leakage and impaired coagulation [294,314]. Dendritic

Introduction

cells are prompted to induce increased expression of chemokines and fail in their antigen-presenting function, and are therefore unable to activate T cells [36]. This leads to severe “bystander” apoptosis of natural killer and T cells, and impair adaptive immune response. The depletion of lymphocytes very likely contributes to EVD fatal outcomes [13,33,106].

Diagnostic methods include rapid tests [39,145], quantitative reverse transcription polymerase chain reaction (qRT-PCR) [58] and nanopore sequencing, which were also used for field diagnostics during the EBOV outbreak from 2014 to 2016 [140]. Antibody detection via enzyme-linked immunosorbent assays (ELISA) [147,179,180,203,204,244] is primarily used to assess the population-wide seroprevalence [38].

1.4 Vaccines and Treatments

Vaccines The available vaccines target the filoviral glycoprotein (GP) or nucleoprotein (NP). During recent outbreaks, for example during 2013 to 2016 in West Africa, the live-attenuated vaccine rVSV-ZEBOV-GP (ERVEBO®) was tested successfully in a ring-vaccination trial [302,119,133,134]. Further testing of the rVSV-ZEBOV-GP in an outbreak in the DRC (2018 to 2020) resulted in the biologics licence applications, which was granted in 2019 [1]. Another anti-EBOV vaccine was licensed, which consists of a heterologous prime-boost strategy including the Ad26.ZEBOV and MVA-BN-Filo (Zabdeno/Mvabea) by Janssen [16,254] was licensed in 2020 [89]. The similar vaccine platform ChAd3-EBOZ by GlaxoSmithKline, Okairios and NIAID was tested with or without MVA-BN-Filo as an alternative heterologous prime-boost regime [91,299,311].

Treatment To date, there are two licensed therapeutics for EVD, both of them monoclonal antibodies (mAbs) against EBOV: REGN-EB3 (Inmazeb™) by Regeneron Pharmaceuticals is a cocktail of three fully-human monoclonal antibodies (atoltivimab, maftivimab and odesivimab) that targets GP and was approved in 2020 [213,252,289], and Ansuvimab (mAb114, EBanga™), also a recombinant human monoclonal antibody, which was originally isolated from a EVD survivor [66,104]. Both treatments were evaluated only against EBOV during the DRC EVD outbreak 2018 to 2019 in the PALM study [232]. ZMapp, another cocktail consisting of three mAbs [257], was given to EVD patients under emergency use authorization during the 2013 to 2016 EBOV outbreak in the DRC. While being considered promising regarding safety, ZMapp did not meet the expectations for efficacy [70]. In addition, REGN-EB3 and Ansuvimab were shown to be superior to ZMapp during the PALM study [232]. Favipiravir and GS-5734 (Remdesivir) by Gilead are polymerase inhibitors, and the latter was shown to be 100% effective in rhesus macaques when administered in a high dose three days post challenge [321], but was not able to clear persistent infection in EVD survivors [135]. As was the case with ZMapp, treatment with Remdesivir is not being continued after publication of the PALM study [232]. The inhibitory effect on the viral polymerase by Favipiravir was assessed by challenging cynomolgus

Introduction

macaques and their treatment with increasing doses of the drug resulted in a survival rate of 40 to 60% [120,208]. However, its administration to EVD patients indicated no significant effect on viral load or mortality [288]. Other drugs that were intensively tested but not (yet) licensed include the small molecule BCX4430 (Galidesivir) by BioCryst Pharmaceuticals [79], treatment with convalescent plasma [97,310], a small interfering RNA lipid nanoparticle product TKM-130803 [80], interferon- β -1a [175], artesunate-amodiaquine [111] and others [192]. In addition, symptomatic treatment as well as supportive care are administered and are often decisive regarding the outcome [187].

1.5 Morphology and genome organization

The family of *Filoviridae* is named after the thread-like appearance of its members (*filum* in latin). Moreover, 6-shaped or spherical particles are also common. The viral particle is approx. 1 μm long or multiples thereof and 80 nm in diameter (Figure 3) [24,28,108]. Spherical particles have also been observed, which exhibit decreased infectivity [323]. Their negative-sense, single-stranded non-segmented RNA genome of approx. 19 kb is flanked by a leader and trailer region that contain the genomic and antigenomic promoters as well as the encapsidation signals. The open reading frames of the seven structural proteins are also flanked by untranslated regions containing the respective transcriptional start and stop signals. The nucleocapsid is formed by the viral RNA genome (vRNA) together with the nucleoprotein NP, the polymerase L, its co-factor VP35, the transcription factor VP30 and VP24 and represents the template for viral genome replication and transcription [272].

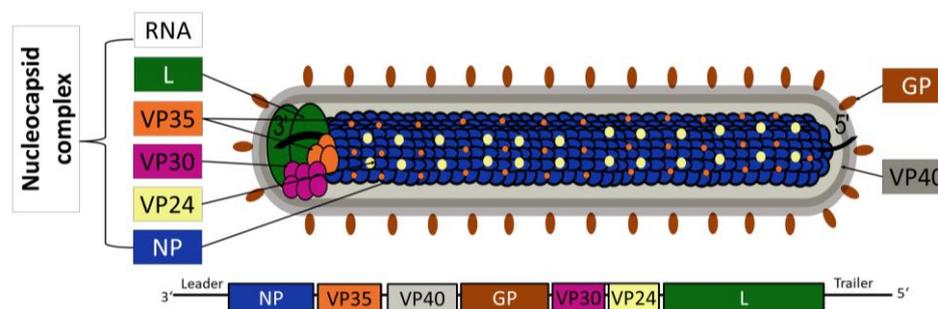


Figure 3: Genome organization and morphology of Ebola virus particles. Trimeric GP is on the outside of the viral particle. VP40 forms the matrix between the lipid membrane and the nucleocapsid complex (top). Schematic representation of the viral genome structure (bottom).

1.6 Viral replication cycle

The filoviral replication cycle can be divided into the three major stages entry, transcription and replication of the viral genome, and assembly and budding.

Entry Initial attachment of viral particles is mediated by the surface glycoprotein GP in its trimeric form [23] consisting of GP1 and GP2 subunits connected via a disulfide bridge [312], similar to other typical class I fusion proteins such as influenza virus hemagglutinin [69]. GP1 contains the receptor binding domain [93,94,181], whereas GP2 mediates fusion with the plasma membrane [191].

Introduction

After binding to cellular attachment factors such as C-type lectins [7] or TIM-1 [40], the virus is internalized via macropinocytosis [237], enters the endosomal pathway and is trafficked through late endosomes [268]. Upon low pH, the proteolytic cleavage of GP1 via the endocytic enzymes cathepsin B and L [53] enables the binding to the intracellular Niemann-Pick C1 (NPC1) receptor. This then results in the fusion of the viral envelope with the endosomal membrane. The ribonucleoprotein (RNP) complex consisting of consists of NP, VP35, VP24, VP30, L and the vRNA is then released into the cytoplasm [49].

Transcription and replication Once the RNP is released, primary transcription can take place. This process encompasses the transcription of viral messenger RNAs (mRNA) through NP, VP35, VP30 and L. Those proteins are incorporated into the particles and therefore available upon release of the RNP into the cytosol [139]. VP35 and L form the viral polymerase complex that enables transcription together with VP30 [231,230] and therefore eventually enables the production of monocistronic, 5'-capped and 3'-polyadenylated mRNA [271]. The latter happens due to the polymerase's stuttering, reading uridines of the transcriptional stop signal only once, but adding adenosines multiple times [324]. Due to a single transcription initiation site in the leader region of the genome, mRNAs of the respective genes are transcribed sequentially in the 3' to 5' order. The transcriptional stop signal at the 5' end of each gene prompts the polymerase to add the poly A tail and with a certain probability the polymerase falls off the template, resulting in high transcription levels for NP and low levels for L [6,35]. Translation of all viral proteins occurs at ribosomes in the cytosols with the exception of GP which is co-translationally translocated into the endoplasmic reticulum (ER) [155,172]. Translation and accumulation of filoviral NP induces the formation of inclusion bodies in the cytosol [173], and, through direct or indirect interaction, NP recruits the other nucleocapsid (NC) proteins VP35, VP30, L and VP24 into the inclusions [237]. Once the level of viral proteins reached a certain level, the replication process of the viral genome can begin [174]. Thereby, a complementary positive sense antigenome is created by the viral polymerase complex by L and VP35, which serves as a template for the production of progeny negative strand viral genomes [230], which in turn are templates for secondary transcription.

Assembly and budding The synthesis of viral RNA and its assembly into NCs takes place in inclusion bodies [141]. The viral proteins interact with the newly synthesized viral genome and form nucleocapsid complexes which are condensed and eventually transported towards the plasma membrane via actin polymerization [280]. The interaction of VP24 with NP is required for assembly and nucleocapsid condensation [15]. Interestingly, for transporting the NCs to the plasma membrane only NP, VP24 and VP35 are needed [297]. In addition, those proteins also determine the structure of the NC, as the expression of only these three proteins resulted in complexes that were morphologically indistinguishable from NCs observed during infection [146]. At the plasma membrane, the NCs interact

Introduction

with VP40, which is solely responsible for the morphology of new virions [28,229,318]. VP40 recruits and hijacks components of the cellular endosomal sorting complex required for transport (ESCRT) such as the neuronal precursor cell-expressed developmentally downregulated 4 (Nedd4) and tumor susceptibility gene 101 (Tsg101) [149,196,287,335] as well as IQGAP1 [202] to support its budding function. However, mutation of the respective interaction sites on VP40 (late domains) still resulted in rescuable recombinant EBOV suggesting alternative interaction sites and/or partners [239]. Simultaneously, GP is also recruited to these VP40-enriched sites. NCs are enveloped by the lipid bilayer of the plasma membrane, which is enriched with VP40 and GP (incorporated into the membrane) and mature virions are eventually released via abscission [83]. This occurs mainly at filopodia but release has also been observed at flat areas of the plasma membrane [245,323]. A schematic depicting of the filoviral replication cycle is shown in Figure 4.

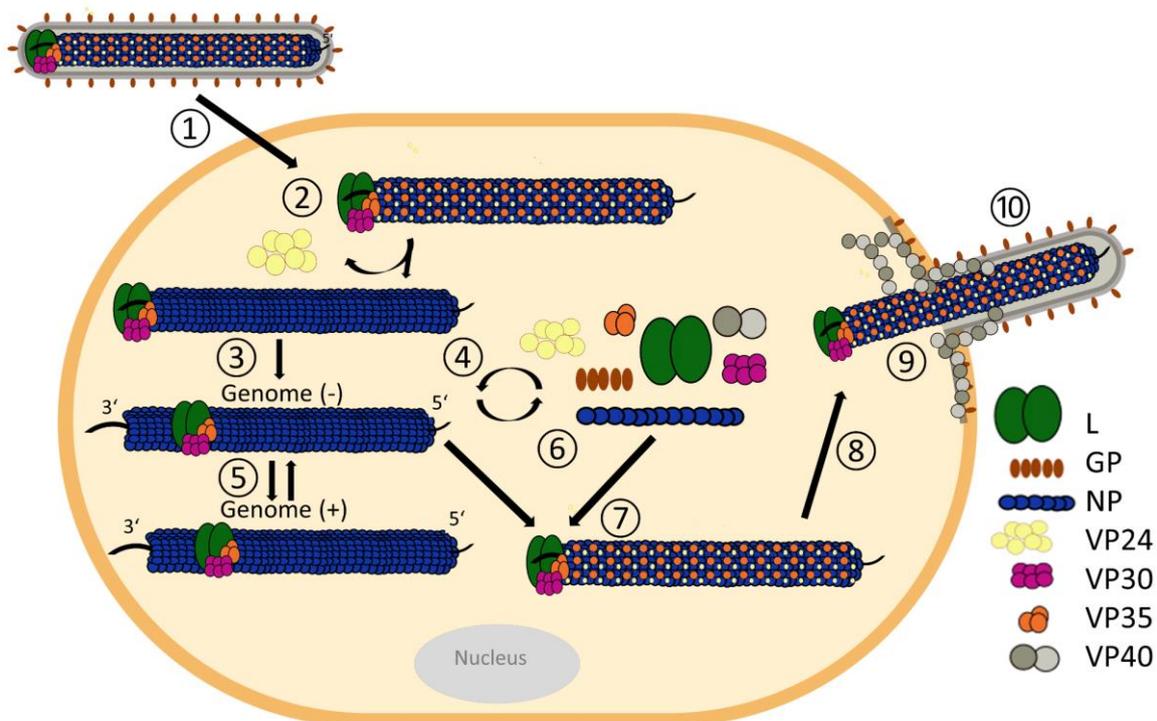


Figure 4: EBOV replication cycle. Fusion of GP with NPC1 and entry via macropinocytosis, 2) decondensation and release of the NC into the cytoplasm, 3) primary transcription, 4) translation of viral proteins, 5) generation of the anti-genome, 6) secondary transcription, 7) formation of the NC, 8) transport to the plasma membrane, 9) assembly and 10) release. Adapted from [137].

1.7 VP40

1.7.1 Structure and functions

VP40 is the filovirus matrix protein (326 aa) which forms different homo-oligomers that have distinct functions in the viral replication cycle [32]. It is also the most abundant protein in the filoviral particles and responsible for budding of new virions [157,305]. VP40's N-terminal domain (NTD) (residues 1 to 194) is connected to its C-terminal domain (CTD) (201 to 326) via a flexible linker (Figure 5 A) [76].

Introduction

Dimeric VP40 adopts a butterfly shape and the two protomers in the crystal are joint via residues 52 to 65 and 108 to 117 in the NTD (Figure 5 B and C) and can homo-oligomerize further into filaments (Figure 5 D and E) or octamers (Figure 5 F and G).

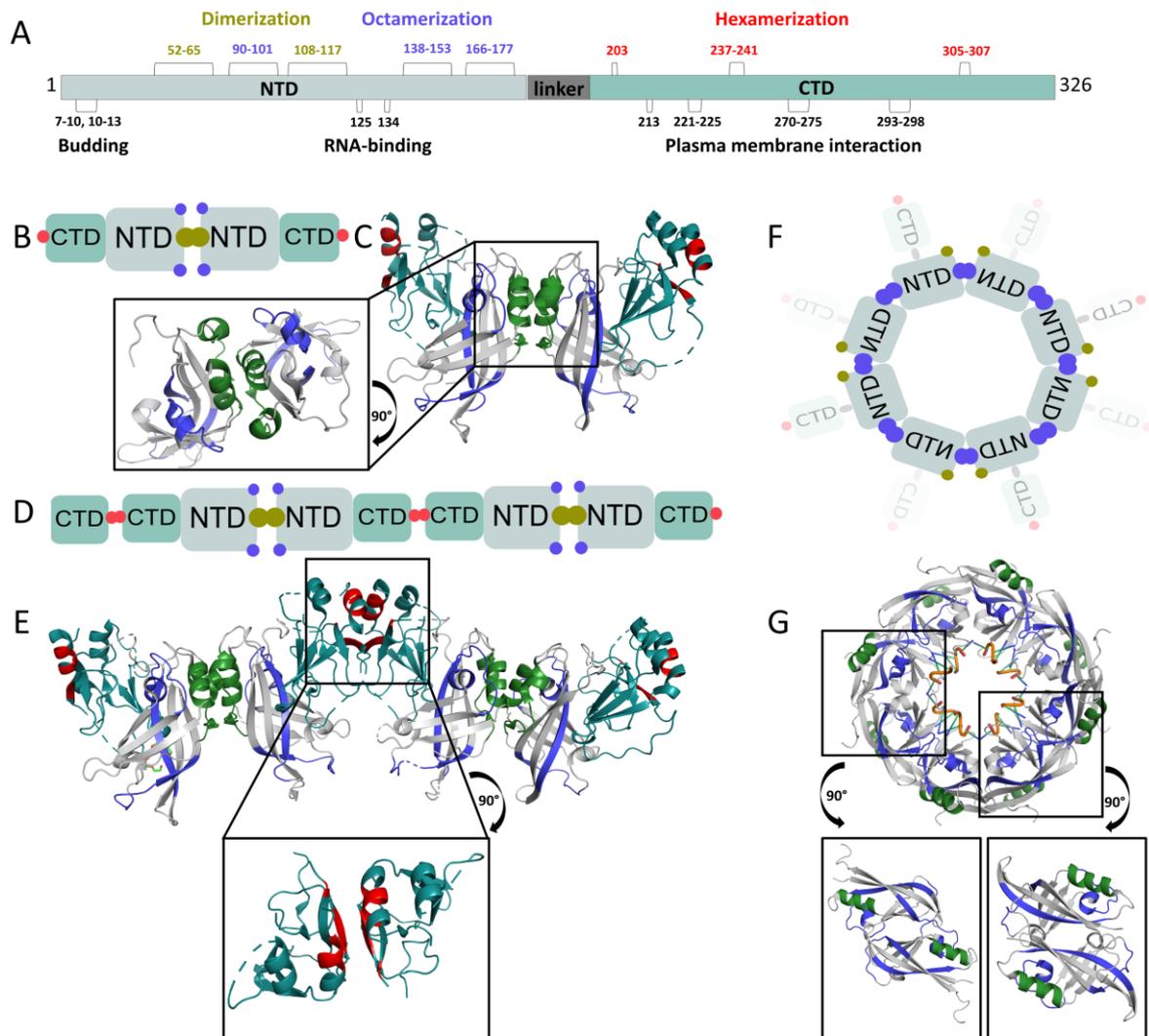


Figure 5: Schematic representation of VP40. A) Residues involved in the homo-oligomerization interfaces are highlighted in green, blue and red (for the dimeric, octameric and filamentous interface, respectively) above the sequence graph whereas further residues of interest are shown below. Different homo-oligomers of VP40 connected via different interfaces coloured in green for dimers (B and C; PDB-code 4LDD), green and red for filaments (D and E, PDB-code 7JZI) and purple for octamer (F and G; PDB-code 1H2C). A close-up of the respective interface(s) is given below the crystal structure of each oligomer.

Dimers are thought to be quickly formed after translation and can be transported to the plasma membrane [32]. This trafficking process was reported to be facilitated by complex formation of residues 303 to 307 of VP40 with Sec24C, a component of the COPII vesicular transport system [332]. It was also described that VP40 interacts with tubulin via residues 223 to 253 [266] and actin for transportation [2]. At the plasma membrane, VP40 dimers are triggered by phosphatidylserine to polymerize into filaments [3] where the dimers are connected through their CTD via residues L203, I237, M241, M305 and I307 (see Figure 5 A). This CTD-to-CTD interface is discontinued and involves

Introduction

only few amino acids suggesting that a high degree of flexibility is needed (Figure 5 D and E). Cryo-EM analysis of virus-like particles (VLPs) and purified virus showed a sequence of VP40 dimers connected only via the CTD-to-CTD interface (Figure 5 D and E). This enables the induction of membrane curvature [318], which is probably exerted by the hydrophobic C-terminal residues L213, I293, L295, and V298 penetrating deeply into the plasma membrane [5]. Other hydrophobic residues described as crucial for the interaction of VP40 with the plasma membrane are lysines at positions 221, 224, 270, 274 and 275 [32] as well as tryptophane at position 191 [160]. In addition to those residues, budding is affected upon mutation of amino acids K127, T129A, and N130 [4] and K212, L213 and R214 [220] (Figure 5 A) suggesting an important role of these residues. While the mechanism of action leading to budding of enveloped particles is not clearly established for some of the mentioned residues, amino acids of VP40's two late domain (L domains) motifs 7-PTAP-10 and 10-PPxY-13 recruit host proteins of the ESCRT complex, which facilitates budding and abscission of vesicles at multivesicular bodies [127]. While the importance of these late domains has been shown multiple times [149,250], they are not essential for the propagation of EBOV [124,239].

The third homo-oligomeric form of VP40 is an octameric ring [115]. A recent study has shown that dimeric VP40 can be triggered into the octamer upon incubation with DNA or RNA [188]. The conformational changes probably consist of the CTDs of dimeric VP40 flipping up and a rearrangement of the NTDs which results in the exposure of residues 90-KQIPIWLPLGVA-101 through which the NTDs reconnect and eventually form an intermediary dimer. Then, four of those dimers arrange to form the octamer via residues 138-LGPGIPDHPLRLLRIG-153 and 166-VQLPQYFTFDLT-177 (Figure 5 F and G). The CTDs are disordered and probably displaced above and below the ring alternately. The octameric structure also includes several triribonucleotides 5'-UGA-3', which are bound to the residues R134 and F125 of each monomer. Octameric VP40 was reported to be involved in the downregulation of viral genome replication and transcription [142]. Crucial residues that contribute mainly to the interface integrity are called hot spot amino acid and include W95 as well as the RNA-binding amino acids R134 and F125 for the octamer. Hot spot amino acids of the dimeric interface (of the dimer), such as T112 and L117, were reported to be indirectly crucial for octamerization as the dimerization-deficient mutants were also unable to form octamers [32]. The interfaces of the various homo-oligomeric forms of VP40 are strongly conserved among Ebola viruses, suggesting that all VP40 proteins are assembled similarly (Figure 58, supplemental material).

Reported post-translational modifications of VP40 include SUMOylation at K326 [21], ubiquitinylation [125] and phosphorylation [103].

1.7.2 VP40 as a drug target

VP40 plays a central role in the viral replication cycle due to its ability to enable budding and to downregulate viral genome replication and transcription. Next to its ability to homo-oligomerize into dimers, octamers and filaments, which are all connected through different interfaces, it interacts with a number of host proteins. Therefore VP40 offers numerous residues, regions or functions that can be targeted for the development of antiviral strategies:

- RNA binding/octamerization
- Transport to the plasma membrane/dimerization
- Interaction with the plasma membrane/egress/filament formation
- Others: Recruitment of and interaction with RNP complexes, post-translational modifications

The following section will give an overview of published strategies using a variety of methods and approaches to target VP40.

RNA-binding/octamerization A virtual screening using a Traditional Chinese Medicine database for binders of octameric VP40 revealed two promising compounds. *In silico* analysis suggested their binding to the RNA-binding residues of VP40, yet *in vitro* data is lacking [165]. In a similar study, sugar alcohols were tested on octameric VP40 using molecular dynamics [236]. Other docking studies investigated compounds against SUDV VP40 (sVP40), especially against its RNA-binding residues [168,206].

Transport/dimerization In a study from 2020, transport and membrane localization of VP40 were assessed in a high-throughput screening using a VP40 oligomerization fluorescent assay. The authors identified sangivamycin as an inhibiting compound and confirmed its potency against other viruses as well [25]. Targeting dimerization specifically was described in a recent publication where molecular dynamics were used as a high-throughput screening to detect molecules binding to the interface. 374 compounds were tested in a luciferase-based assay and the two most promising compounds were shown to impact VLP production across several Ebola virus species as well as MARV [309]. However, no structural information of the binding mode was available in either case.

Interaction with the plasma membrane/egress/filament formation Compounds inhibiting VP40 egress such as quinoxalin-2-mercapto-acetyl-urea analogues were identified using a bimolecular complementation assay as well as budding experiments [201]. Egress was also inhibited using ISG15, a compound that inhibits Nedd4 ligase activity and therefore the VP40-ESCRT interaction [250]. Docking studies targeting K127, T129, and N130 and their role during viral egress revealed a binding compound through molecular dynamics [14].

1.8 Protein-protein interactions as targets for antiviral intervention

Protein-protein interactions (PPIs) play a fundamental role during many signalling cascades, are the basis of the formation of complex protein networks and play a central role during disease development and progression [207,219]. Homo-oligomeric assembly of proteins can result in increased solubility and stability due to the shielding of hydrophobic areas in interfaces and the formation of higher-order oligomers [26], which bring along further properties such as decreased likelihood of aggregation [129]. Self-assembly of proteins is crucial for numerous physiological processes as it can influence active site conformation or even architecture, downstream signalling, and activity levels of the protein of interest [301].

Targeting protein self-association represents a novel strategy to combat infections but also a major challenge for drug design processes due to the typically hydrophobic residues involved in the usually flat interfaces of approx. 1,200 to 4,000 Å² [225]. In addition, these interfaces can be discontinuous and would require either a large or several smaller compounds to target the area of interest [32,48,264,283]. Compounds targeting the self-association interfaces need to exhibit high affinities towards the target structure and be present in potentially very high concentrations to successfully outcompete subunits of the protein from further homo-oligomerization. Homomers with a low affinity of self-association are more suitable for inhibition because interface ligands of highly stable complexes potentially stay undiscovered, especially in cases of molecules with only low affinities [301]. Hot spot residues are defined by an increase of ≥ 2 kcal/mol in free binding energy when mutated to alanines [303]. In addition, they are often surrounded by polar neighbouring residues [195] rendering the interface highly complex and therefore lowering its druggability [10]. Due to these difficulties, targeting allosteric sites is considered an alternative strategy to inhibit homo-oligomerization as these sites often have advantageous properties compared to the self-association interfaces [10,101,131,301]. Numerous benefits result from targeting PPIs of homomeric complexes. For example, interfaces of PPIs were reported to be less conserved than active sites, suggesting a higher selectivity of interface-targeting compounds [44,205]. For conventional compounds such as active-site binders of enzymatic proteins or allosteric inhibitors, compensatory effects were observed, resulting in increased expression of the target protein or adaptation via mutations to avoid the inhibitory effects. In contrast, inhibition of PPIs proved to be a more promising approach for HIV [45] and Kaposi's sarcoma-associated herpesvirus [282]. As mentioned earlier, self-association of proteins and their inhibition can lead to increased stability and decreased degradation which was shown repeatedly [57,291]. Additionally, PPI interfaces are typically large, flat structural areas that potentially offer several sites to target, therefore enabling the inhibition of targets previously considered undruggable, especially for proteins lacking a (known) active site [52,285]. As a final advantage, oligomerization

inhibitors might exert a sub-stoichiometric effect that describes the inhibition of a complex or several copies of the protein of interest by a single compound [102,210].

1.9 Aims of this work

As of today, there is still no licensed drug against EVD. The matrix protein VP40 plays a central role in the viral replication cycle and was investigated as a potential drug target in numerous studies using different approaches and techniques. However, in most cases an inhibition of the interaction between VP40 and host proteins was targeted. In addition, structural information was either lacking and results were not experimentally validated for *in silico* studies. Therefore, the objective of the present study was to develop inhibitors against VP40 using a structure-guided approach. The typical steps of drug discovery include target identification, target validation, lead identification, lead optimization, and preclinical and clinical phases. Structural information is especially crucial for the earlier steps, beginning with lead optimization.

Specifically synthesized peptides or preselected fragments served as lead structures, which were tested for their binding activity to VP40 using various screening methods, such as surface plasmon resonance, crystal soaking experiments, and disulfide tethering (DT). Once a hit was identified, it was optimized for improved affinity towards VP40. This was performed by the generation of a virtual library of derivatives, their molecular docking to VP40, as well as their synthesis by medicinal chemists for further testing (Figure 6).

The present work targets EBOV VP40 (zVP40) and SUDV VP40 (sVP40) depending on the available test systems and was divided in three parts:

Part I: Characterization of residues directly and indirectly involved in VP40 self-assembly

Part II: Inhibition of VP40 using peptides

Part III: Inhibition of VP40 using small molecules

Introduction

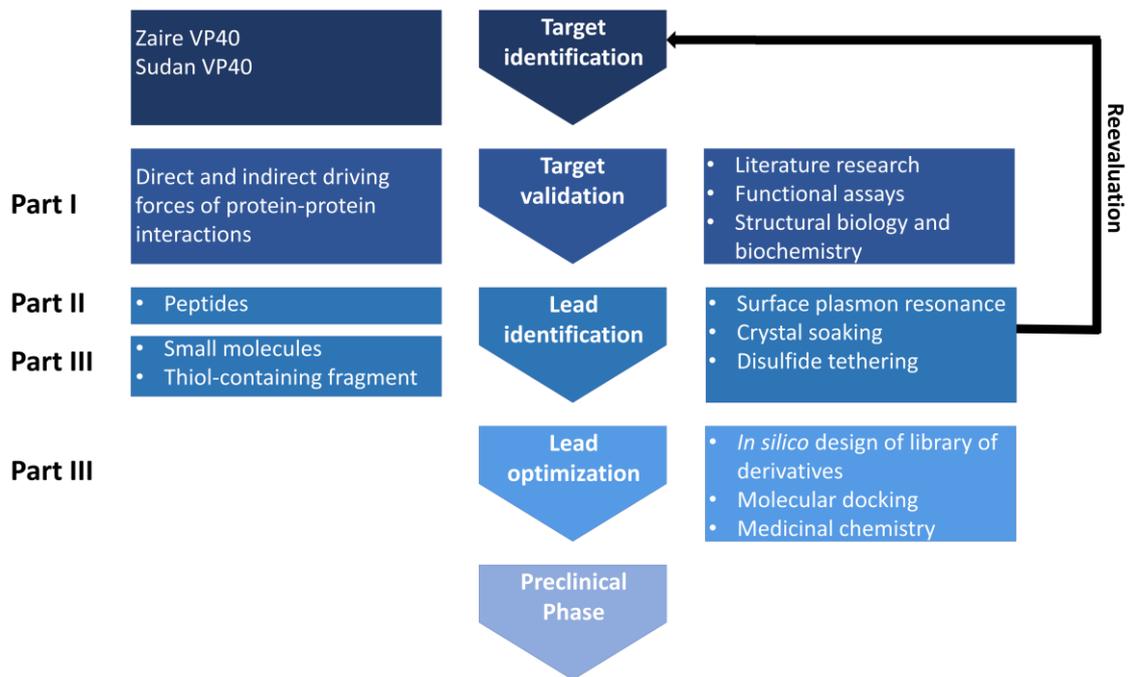


Figure 6: Work plan of the present thesis following the typical drug design steps. Steps include target identification and validation, lead identification and optimization, divided in three major parts.

2 Methods

2.1 Molecular biology methods

2.1.1 Polymerase Chain reaction

A Polymerase chain reaction (PCR) was developed by Kary Mullin in 1987 and is used to enzymatically amplify DNA fragments. In an initial step, the template DNA fragment is denatured at high temperature (usually 94 to 98 °C) so the two strands are separated. Afterwards, the temperature is lowered to allow the hybridisation of two specific primers to the complementary sequences of the single-stranded template which is called annealing. The melting temperature (T_m) of the primers is used to calculate the ideal annealing conditions ($T_m - 5$ °C). For the following amplification phase, the temperature is raised again to typically 72 °C (determined by the used polymerase) and the sequence is extended starting from the primers. Exponential amplification of the DNA sequence stretch of interest flanked by the two primers is achieved by repeating the three steps for usually 35 to 40 cycles. This technique is also used to add molecular tags or chosen restriction enzyme sequences to the final product [234]. PCR was performed typically using the Q5 High-Fidelity DNA Polymerase by New England Biolabs (NEB), Frankfurt. The reaction setup was as follows:

DNA template	100 ng	Step	Temperature	Duration	Cycle
Forward Primer (10 μ M)	1 μ L	Initial denaturation	94 – 98 °C	30 sec	1
Reverse Primer (10 μ M)	1 μ L	Denaturation	94 – 98 °C	30 sec	35 – 40
dNTPs (10 mM each)	1 μ L	Annealing	55-72 °C	2 min	
Q5 reaction buffer	5 μ L	Elongation	72 °C	2 - 6 min	
Q5 polymerase	0.5 μ L	Elongation	72 °C	6 min	1
H ₂ O	ad 50 μ L	storage	8 °C	∞	1

For the amplification of DNA fragments with RNA as a template, a reverse transcriptase PCR (RT-PCR) was employed. For this, the RNA template was first transcribed into complementary DNA (cDNA) using either random hexamer primers or sequence-specific primers. The generated cDNA then served as a template for the following regular PCR as described above. RT-PCRs were performed using the Transcriptor One-Step RT-PCR Kit by Roche. The reaction setup was as follows:

RNA template	3 μ L
Primer (10 μ M)	2 μ L each
5x reaction buffer	10 μ L
Enzyme mix	1 μ L
H ₂ O	32 μ L

Methods

Step	Temperature	Duration	Cycle
Reverse transcription	48 °C	30 min	1
Initial denaturation	94 °C	2 min	1
Denaturation	94 °C	30 sec	10
Annealing	50 °C	30 sec	
Elongation	68 °C	4 min	
Denaturation	94 °C	30 sec	25
Annealing	50 °C	30 sec	
Elongation	68 °C	4 min + 5 sec	
Elongation	68 °C	7 min	1
storage	8 °C	∞	1

Reaction mixtures were either purified via PCR purification (see section 2.1.4) or agarose gel electrophoresis (see section 2.1.3) or directly stored at -20°C until further use.

2.1.2 Preparative and analytical restriction digest

To prepare DNA fragments and vectors for ligation, a digest with restriction enzymes was performed, typically generating sticky single-stranded DNA overhangs. The restriction enzymes used for this study were exclusively type II restriction enzymes (NEB), which have palindromic recognition sites and cut within the restriction sites. By using the same pair of type II enzyme for both DNA fragment and vector, defined overhangs were created and could directly be ligated. To prevent religation of the vector, the DNA was dephosphorylated using calf intestine alkaline phosphatase (CIP).

DNA fragment/vector	x µL
10x restriction buffer	3 µL
Restriction enzyme	1 µL
CIP (only for vector digest)	1 µL
H ₂ O	ad 30 µL

Reaction mixtures were incubated for 1 to 12 h at 37 °C and 300 rpm.

Preparative digested DNA was either purified using the OMEGA DNA purification kit (see section 2.1.4) or via preparative agarose gel electrophoresis (see section 2.1.3). Analytical restriction digest served to control cloned constructs for the presence of the insert. The above-described reaction setup was downscaled sixfold to a total volume of 5 µL and subjected to analytical gel electrophoresis (see section 2.1.3) after incubation for 1 to 2 h at 37 °C and 300 rpm.

Methods

2.1.3 Preparative and analytical agarose gels electrophoresis

Agarose gel electrophoresis was used to either purify or analyze DNA fragments after PCR, analytical or preparative restriction digest. DNA samples were mixed with 6x Orange Loading Dye and applied to a 1% agarose gel in tris acetate EDTA (TAE) buffer along with O'GeneRuler 1kb Plus DNA Ladder. Migration of the negatively charged phosphate backbone in an electric field enabled the separation of DNA fragments according to their size for 1 h at 100 Volts (V) within the agarose matrix. Traditional staining included incubation of the gel in 0.01% ethidium bromide for 30 min and visualization on a UV screen. As an alternative method, staining was performed using an Fisherbrand™ Safeview BLUE Light Transilluminator. Bands were visualized with GelGreen Nucleic Acid Stain after a 1 h incubation period.

For preparative agarose gel electrophoresis, the DNA bands of interest were cut out with a scalpel and purified using the E.Z.N.A.® Gel Extraction Kit by Omega Bio-Tek Inc., Norcross (USA), according to manufacturer's instructions. Briefly, the cut-out agarose was dissolved in Binding Buffer (XP2) at 65 °C for 10 min, loaded onto a DNA binding column, washed with DNA washing buffer and eventually eluted in 30 to 50 µL dH₂O.

2.1.4 Purification of PCR products

After completion of the PCR or RT-PCR, the reaction mixture was purified using the OMEGA Cycle Purification Kit according to manufacturer's instructions and either stored at 4 °C or frozen at -20 °C for long-term storage. This served to remove residual salts, enzymes, dNTPs of buffers from the DNA. Briefly, the PCR reaction was mixed with 5 x volumes of CP buffer, applied to a DNA binding column and washed twice with DNA washing buffer. To elute the DNA, 30 to 50 µL dH₂O were added, followed by a final centrifugation step. The underlying principle of such purifications is the binding of DNA to the anion-exchange column, followed by elution at low salt concentrations with dH₂O.

2.1.5 Ligation

Ligation is the process of inserting one or several cut DNA fragments into a vector that is cut with the same restriction enzymes using a T4 DNA ligase (NEB). This enzyme catalyzes the formation of phosphodiester bonds between nucleotides which is usually most effective when joining cohesive ends. The enzyme is ATP and Mg²⁺ dependent, components which are supplied with the 10x reaction buffer. To avoid religation of the vector, a surplus of insert is used for the reaction (usually 1:5 to 1:20 insert to vector). The setup was as follows:

Insert	x µL
Vector	y µL
T4 DNA ligase	1 µL
10x T4 DNA ligase reaction buffer	1.5 µL
H ₂ O	ad 15 µL

Methods

The reaction was then incubated at 16 °C or room temperature (RT) for 1 to 16 h and stored at 4 °C until transformation (see section 2.1.8) or at -20 °C for permanent storage.

2.1.6 Site-directed mutagenesis

To introduce one or several mutations such as small deletions, insertions or amino acid switches into a plasmid, the Quickchange Multi Site-directed mutagenesis kit and Quickchange Lightning Multi Site-directed mutagenesis kit by Agilent were used. The following formula was used to determine the ideal primer length and composition which reflects the melting temperature T_m :

$$T_m = 81.5 + 0.41 (\% \text{ GC}) - (675/N) - \% \text{ mismatch}$$

with N depicting the primer length in bases and values for % GC and % mismatch being whole numbers. The T_m and the length of the primer used should not exceed 78 °C and 45 bases, respectively.

The setup was as follows:

QuikChange	
ds-DNA template	x μL
Quiksolution	0.5 – 1.5 μL
Primer (1 μM)	1 μL
dNTP-Mix	1 μL
10x quikchange multi reaction buffer	2.5 μL
Quikchange (Lightning) multi enzyme blend	1 μL
H ₂ O	ad 25 μL

Step	QuikChange			QuikChange Lightning		
	Temperature	Duration	Cycle	Temperature	Duration	Cycle
Initial denaturation	95 °C	1 min	1	95 °C	2 min	1
Denaturation	95 °C	1 min	30	95 °C	20 sec	30
Annealing	55 °C	1 min		55 °C	10 sec	
Elongation	65 °C	12 min		65 °C	2,25 min	
Elongation	65 °C	12 min	1	65 °C	5 min	1
storage	8 °C	∞	1	8 °C	2 min	1

After the cycling process, 1 to 2 μL *DpnI* is added to the reaction mixture. This endonuclease served to digest the parental and non-methylated DNA template. After an incubation of 1 to 5 h, usually 3 μL were transformed into chemically competent *XL10 Gold E. coli* cells, supplied with the kit according to manufacturer's instructions (see section 2.1.8.2).

Methods

2.1.7 Preparation of chemically Z-competent bacteria

Z-competent bacteria were prepared by using the Z-Competent *E. coli* Transformation Kit™ and Buffer Set by Zymo Research, Orange (USA), according to manufacturer's instructions. Briefly, a preculture of XL1 Blue cells were grown in SOB medium overnight at 37 °C and 300 rpm and transferred to the main culture which was incubated overnight at RT without shaking until an optical density (OD) measured at a wavelength of 600 nm (OD₆₀₀) of 0.4 to 0.5 was reached. The cell suspension was cooled down on ice, spun down at 4,000 rpm and the bacterial pellet washed twice with washing buffer. After the final centrifugation step, cells were resuspended in competent buffer, aliquoted and frozen at -80 °C until transformation.

2.1.8 Transformation of chemically competent bacteria

2.1.8.1 Transformation of XL1 Blue cells

Transformation of cloned DNA fragments was performed by adding 5 to 10 µL of the ligation mixture to 50 µL of Z-competent thawed bacteria. Cells were incubated on ice for at least 20 min and plated out on lysogeny broth (LB) agar plates containing the appropriate antibiotic and colonies were grown for 16 h at 37 °C.

2.1.8.2 Transformation of XL10 Gold cells

After performing site-directed mutagenesis, the transformation of mutated DNA fragments into *E. coli* was performed according to manufacturer's instructions. Briefly, cells were treated with β-mercaptoethanol and incubated for 30 min on ice after addition of the DNA. To induce uptake of the plasmid, cells were heated to 42 °C for 30 sec. After mixing the cells with 100 to 500 µL NZY⁺ medium, they were incubated for at least 1 h at 37 °C and 300 rpm before being plated on LB agar plates with the appropriate antibiotic and colonies were grown for 16 h at 37 °C.

Afterwards, 2 to 5 mL of LB medium containing the appropriate antibiotic were inoculated with a single colony from the agar plate and again incubated for 16 h at 37 °C and 200 rpm. These so-called precultures were either used for the inoculation of a main culture for heterologous expression of recombinant proteins or for plasmid preparations (see section 2.4.1 and 2.1.9, respectively).

2.1.9 Preparation of plasmid DNA from bacteria

To extract and purify extrachromosomal plasmids from *E. coli*, the commercially available E.Z.N.A.® Plasmid DNA Mini Kit or E.Z.N.A.® Fast Filter Plasmid DNA Maxi Kit by Omega Bio-Tek Inc., Norcross (USA), were used. In short, 2 to 5 mL or 100 mL (for small- or large-scale purification, respectively) of a bacterial culture were centrifuged and the pellet resuspended in lysis buffer containing RNase A. Afterwards, solution II containing SDS and NaOH was added to destroy bacterial cell walls and denature proteins. By neutralizing the lysate with buffer III containing potassium acetate and glacial acetic acid, plasmid DNA renatured whereas genomic DNA precipitated. The mixture was then centrifuged for 10 min at RT and 13,200 rpm. Supernatants were applied to DNA binding columns with a silica matrix,

Methods

washed with HBC and washing buffer, containing isopropanol and ethanol respectively. Purified plasmids were eventually eluted using 30 to 1,500 μL dH_2O . Purified plasmid DNA was then used to transform bacterial strains for heterologous expression, subjected to sequencing (see section 2.1.10) or stored at -20°C .

2.1.10 Sequencing

To check for successful cloning of plasmid DNA or correct EBOV genome sequence, samples of 12 μL DNA or cDNA were mixed with 3 μL of a selected primer (10 mM stock concentration) and sent to Microsynth Seqlab, Schweiz. Analysis of DNA sequences was done using BioEdit and the NCBI Sequence Alignment tool BLAST (www.ncbi.nlm.nih.gov/BLAST/).

2.2 Cell culture methods

2.2.1 Cultivation of cell lines

For the present work, HuH7 (human hepatocyte-derived carcinoma cell line), HEK293 F (Human Embryonic Kidney) and VeroE6 cells (kidney epithelial cells of African green monkeys) were used. All cell lines were cultivated in T75 flasks in Dulbecco's Modified Eagle's Medium (DMEM) supplemented with 10% fetal bovine serum (FBS), 2 mM L-glutamine (Q), 100 U/ml penicillin and 100 $\mu\text{g}/\text{ml}$ streptomycin (P/S) and grown at 37°C and 5% CO_2 and passaged every 3 to 4 days. For splitting cells, they were first washed twice in phosphate buffered saline (PBS_{def}) and incubated with 2 ml Trypsin/EDTA until they start to detach visibly from the flask's bottom. Cells were then resuspended in 8 mL DMEM + 10% FBS + Q + P/S, counted and seeded into new flasks or cell culture plates.

2.2.2 Transfection of DNA

The transfection of plasmid DNA was performed with the transfection agent TransIT[®]-LT1 (Mirus), which enables complex formation of the negatively charged DNA backbone with cationic lipids. The complexes can then attach to and eventually pass cell membranes. The ideal cell confluency was between 50 and 70%. For the present work, HuH7 or HEK293 cells were transfected by the following protocol: 100 μL Opti-MEM[®] were mixed with 3 μL TransIT[®]/ μg of DNA to be transfected and incubated for 5 min at RT. Afterwards 100 μL Opti-MEM[®] were added to the plasmid DNA. Both reaction volumes were mixed, incubated for 15 min at RT and then added to the cells. An optional cell culture media change was performed 4 h post transfection (pt) to DMEM + 3% FBS + Q + P/S. For VLP production for mass spectrometry analyses (see section 2.4.8.1), the following alterations were made: cell culture medium of T175 HEK293 cells was exchanged to 8 ml Opti-MEM[®] and the transfection mixes contained 750 μL Opti-MEM[®] instead of 100 μL .

2.2.3 Transfection of peptides

To efficiently transfect peptides into mammalian cell lines, the commercially available agent Chariot[™] was used. This peptide carrier consists of a tryptophan-rich domain responsible for the interaction with

Methods

the cellular membrane, a lysine-rich domain which serves to increase solubility and efficient transport as well as a spacer between the two domains. The Chariot™ transfection agent is mixed with peptides, forming an initially stable complex which dissolves upon entry into the cell, releasing the peptides [228].

Peptides were reconstituted in DMSO in varying concentrations and 0.35 μL diluted in 96.5 μL PBS_{def}. To reach maximum concentrations of DMSO of 0.1% (v/v), peptides were diluted 1:1000 in the final volume of 350 μL (see below). As stated by the manufacturer, Chariot™ is supposed to be diluted 1:10 in dH₂O for the transfection of peptides and 4 μL mixed with 96 μL dH₂O. This solution was then mixed with the peptide-PBS_{def} solution and incubated for 1 h at RT. HuH7 cells were seeded in 12 well plates with a confluency of 50 to 60%. Shortly before transfection, cell culture medium was aspirated, cells washed once with PBS_{def} and 150 μL DMEM without FBS added to each well. The final volume therefore sums up to 350 μL . The Chariot™-peptide mix was then added to each well and then cells incubated for approx. 1 h at 37 °C and 5% CO₂. Cells were infected with EBOV under Biosafety level (BSL) 4 conditions, described in section 2.3.3. Samples of the supernatant were taken at 0 and 2 days post infection (dpi) and viral titers analyzed via TCID50 (see section 2.3.6).

2.3 Virological methods

2.3.1 Minigenome assay

A minigenome (MG) system was devised to study viral genome replication and transcription under BSL-2 conditions [138]. The EBOV genome is replaced by a reporter gene coding for a Renilla Luciferase, flanked by the EBOV genome's 3'-leader and 5'-trailer, the regions necessary for the initiation of these processes and cloned into a mammalian expression vector (pANDY 3E5E). Transfection of the plasmid coding for the minigenome as well as plasmids containing genes coding for the filoviral proteins L, VP30, VP35 and NP leads to their recombinant expression in mammalian cells, enabling transcription and replication of the minigenome. Additionally, a plasmid encoding the T7 polymerase was co-transfected to generate negative-sense RNA of the MG that can then be recognized by the viral replication machinery, which delivers positive-sense RNA that then serves as a template for the transcription and cellular translation of the Renilla luciferase (Figure 7) and also further replication. The enzymatic activity of the Renilla luciferase was measured and the signal intensity indicates the level of viral genome replication and/or transcription. As a transfection control, a plasmid encoding the firefly luciferase is co-transfected which can also be quantitatively measured but was translated independently of the viral proteins.

Methods

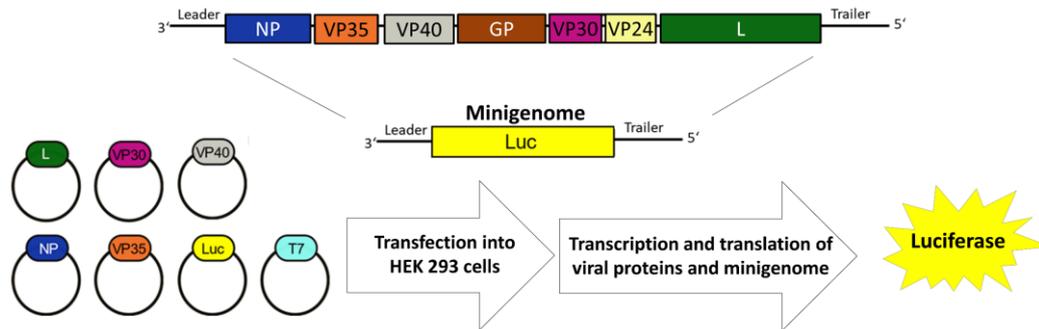


Figure 7: EBOV-specific minigenome assay. The sequences coding for the seven structural proteins are exchanged for the Renilla luciferase gene, flanked by the leader and trailer regions and called the minigenome (MG). Transfection of the MG plasmid as well as the necessary components for viral genome replication and transcription (L, VP35, VP30, NP) as well as a plasmids coding for a T7 polymerase and VP40 into HEK 293 cells results in the transcription and translation of the viral protein and the T7 polymerase. The latter generates the negative-sense RNA of the MG, which serves as a template for replication as well as transcription and eventually translation. Reporter gene activity indirectly indicates levels of viral genome replication and transcription.

To measure the inhibitory effect of VP40 on viral genome replication and transcription, cells at a confluency of approximately 50% were cultured in 12 (or 6) well plates and transfected with 500 ng pCAGGS L, 62.5 ng pCAGGS VP35, 50 ng pCAGGS VP30, 62.5 ng pCAGGS NP, 125 ng pCAGGS T7, 125 ng pANDY 3E5E, 50 ng pGL4 firefly luciferase and 400 ng sVP40 WT or mutants (or double the amounts for 6 well plates). Differences in the absolute amount of transfected plasmid DNA were compensated by the addition of empty pCAGGS vector. 48 h pt, cells were washed twice with ice-cold PBS_{def}, the cell pellet resuspended in Lysis Juice (Promega) and frozen at -20 °C. After thawing, the cell debris was spun down for 10 min at 4 °C and 13,200 rpm. After diluting the supernatant 1:100 in Lysis Juice (PJK), 10 µL were used for the measurement of reporter gene activity of both Renilla and firefly luciferase using the Beetle-Juice and Renilla-Juice BIG KITS (PJK) and measured with a Centro LB 960 luminometer (Berthold Technologies). For this, 50 µL of substrate was added to the 10 µL sample. Firefly luciferase signals were used to normalize for transfection efficiency. The sample without VP40 was considered to be the positive control and set to 100%. In addition to the luminometer measurements, 30 µL were mixed with 10 µL 4x sample buffer (SB) and used for Western blot (WB) analysis (see section 2.4.3.2) to control expression levels of the viral proteins.

2.3.2 Virus-like particle assays

8x10⁵ HEK293 cells were transfected with 500 ng per well of sVP40 WT or mutants along with 500 ng GP per well to increase yields [245]. Media was changed 4 h pt to 3 mL DMEM containing 3% FBS. 24 or 72 h pt cell culture medium was collected, cell debris removed via centrifugation for 10 min at 2,500 rpm and the supernatants subjected to ultra-centrifugation for 2 h at 4 °C and 40,000 rpm with a 20% sucrose cushion using an SW41 or SW60 rotor (Beckman Coulter). Pellets were then resuspended in 50 µL SB and analyzed via WB analysis (see section 2.4.3.2). Cells were washed once with PBS_{def} and resuspended in 100 µL 1x Lysis Juice (Promega) and frozen at -20 °C. After thawing, the

Methods

samples were centrifuged for 10 min at 4 °C and 13,200 rpm and the supernatants were used for WB analysis together with the resuspended VLPs.

2.3.3 Infection of cells with EBOV

All infection experiments using (recombinant) EBOV were performed in the BSL4 laboratory of the Institute for Virology, Philipps-University Marburg according to national and international regulations.

2.3.3.1 Infection of VeroE6 cells with EBOV for virus propagation

For infection studies with (recombinant) EBOV, VeroE6 or HuH7 cells are typically used as they are highly susceptible and display a clear cytopathic effect (CPE). Cells were seeded in T75 flasks or 24, 12 or 6 well plates to approx. 50% confluency and transferred to the BSL4 laboratory. After aspirating the media, (recombinant) EBOV was added at multiplicity of infection (MOI) between 0.01 to 5, which describes the ratio of the virus titer as assessed (for the present work) via tissue culture infectious dose 50 (TCID₅₀; see in section 2.3.6) divided by number of cells. The required virus volume was mixed with DMEM without FBS (500 µL for well plates and 5 mL for T75 flasks) and added to the cells. The inoculum was removed after an incubation period of 1 h at 37 °C and 5% CO₂, the cells washed once with PBS_{def} and DMEM with 3% FBS was added to the cells, which were kept at 37 °C and 5% CO₂ for 5 to 10 days until harvesting of cells and/or supernatant.

2.3.3.2 Infection of HuH7 cells with EBOV for immunofluorescence analysis

For immunofluorescence analysis, HuH7 cells were seeded on cover slips to 50% confluency in 6 well plates and transferred to the BSL4 laboratory. After aspirating the media, (recombinant) EBOV was added at MOI 1 in 500 µL DMEM without FBS to the cells and incubated for 1 h. The inoculum was then removed, the cells washed once with PBS_{def} and DMEM with 3% FBS added and cells kept at 37 °C and 5% CO₂. 1 day post infection (dpi) the supernatant was removed, the cells washed once with PBS_{def} and covered with 4% PFA in DMEM for 24 h. The next day, the fixed coverslips were transferred to a 24 well plate and covered again in 4% PFA in DMEM and the samples were exported from the BSL4 laboratory. After another incubation period for 24 h at 4 °C samples were ready for subsequent analyses as described in section 2.3.9.

2.3.4 Harvest of (recombinant) EBOV stocks

VeroE6 cells were infected with (recombinant) EBOV as described in section 2.3.3. 5 to 7 dpi, the supernatant was harvested and centrifuged for 10 min at 2,500 rpm and 4 °C to remove cell debris. The supernatant was then aliquoted and flash-frozen in liquid nitrogen until further use. Viral titers were assessed using TCID₅₀ analysis as described in section 2.3.6.

2.3.5 Extraction of viral RNA from cell culture supernatant

Viral RNA was usually extracted for subsequent sequencing using the QIAamp Viral RNA Mini Kit (QIAGEN) according to manufacturer's instructions. In short, 140 µL of cell culture supernatant was

Methods

incubated with a mixture of AVL buffer, Carrier-RNA and Bromphenol Blue (provided with the kit). After a 10 min incubation time, 560 μL of 100% ethanol was added and again incubated. Samples were then transferred to new tubes which were surface decontaminated for removal out of the BSL4 laboratory. Viral RNA was then applied to a RNA-binding column, washed, eventually eluted and finally transcribed into cDNA and used for further analysis (see section 2.1.10).

2.3.6 Determination of viral titers

Titers of the (recombinant) EBOV were assessed by TCID₅₀ assay. For this, 180 μL of a VeroE6 cell suspension in DMEM with 3% FBS were pipetted in 96 well plates to reach a confluency of approx. 50% and plates were transferred to the BSL4 laboratory. 20 μL of the virus-containing suspension (with or without prior dilution) were added to 4 to 6 wells. A serial dilution of 1:10 from one row to the next was performed and finally 20 μL were removed from the last row and discarded thus resulting in usually 8 dilution steps containing each 180 μL with 4 to 6 replicative wells. As a mock control, one setup was performed with only DMEM without FBS without virus. Cells were then incubated at 37 °C and 5% CO₂. Six dpi cells were analyzed for CPE and viral titers (TCID₅₀/mL) calculated using the Spearman & Kärber algorithm (Hierholzer & Killington 1996). If needed, the procedure was repeated with 1:5 steps dilution using 160 μL cell culture medium and transferring 40 μL from one row to the next.

2.3.7 Cloning and rescue of recombinant EBOV

Reverse genetics systems are powerful tools to study the effect of amino acid changes on the virus itself and allow its characterization. The underlying principle was the transfection of a DNA plasmid that encodes the full-length viral genome in the presence of a T7 polymerase and the plasmids that are necessary for viral genome replication and transcription, namely L, NP, VP30 and VP35.

2.3.7.1 Cloning of recombinant EBOV

Cloning of plasmids containing the full-length EBOV genome was performed using three cloning cassettes that were eventually fused together and ligated into the final plasmid pAMP [118,240]. Fragment 1 contains the sequences of leader sequences, NP, VP35, VP40 and GP, fragment 2 GP, VP30, VP24 and the first the beginning of L and fragment 3 contains the rest of the gene coding for L, the trailer sequence and ribozyme site. As a first step, selected mutations in the *vp40* gene in fragment 1 were introduced via multisite-directed mutagenesis (see section 2.1.6) with subsequent transformation, cultivation of bacterial cultures, DNA extraction and sequencing. Bacterial cultures containing the correct VP40 sequence were used for downstream cloning. Each fragment and the final vector pAMP were digested using the appropriate restriction enzymes and ligated in a single reaction. An analytical restriction digest with subsequent analytical agarose gel electrophoresis revealed positive clones which were partially sequenced.

Methods

2.3.7.2 Rescue of recombinant EBOV

Rescue of recombinant viruses was performed in HuH7 and VeroE6 cells as described in previous studies [141,177]. Briefly, HuH7 cells (passage 0, P₀) in a 6 well format and at a confluency of approx. 50% were transferred to the BSL4 laboratory and transfected in duplicates with 1 µg full-length plasmid for recEBOV WT or recEBOV VP40 mutants (pAMP), 125 ng pCAGGS NP, 1 µg pCAGGS L, 125 ng pCAGGS VP35, 100 ng pCAGGS VP30 and 250 ng pCAGGS T7 with 7.8 µL TransIT[®] per well (see also section 2.2.2). Wells not transfected with pCAGGS T7 or pCAGGS L served as negative controls. 24 h pt the media was changed to DMEM containing 3% FBS. 7 days pt, supernatants from transfected cells were centrifuged for 10 min at 4 °C and 2500 rpm and used as inoculum for VeroE6 cells (passage 1, P₁) in a 6 well format at 50% confluency. This procedure was repeated for two further infection passages in T25 and T75 flasks containing 50% confluent VeroE6 cells (P₂ and P₃, respectively) at 6 dpi. Successful rescue events were confirmed via WB analysis of cell lysates and supernatants of P₀ 6 days pt and clearly visible cytopathic effects for P₂ and P₃. All recombinant viruses were sequenced to completion. P₃ was used as a virus stock culture (see section 2.3.4) and the titers of the recombinant mutants were assessed by TCID₅₀ assay (see section 2.3.6).

2.3.8 Growth curves of infected cells

To compare the growth kinetics of recombinant mutants with recEBOV WT, 3x10⁵ VeroE6 cells were infected at MOI 0.01, cells washed with PBS_{def} after 1 h of incubation at 37 °C and 3.5 mL DMEM containing 3% FBS added. Samples of 500 µL supernatant were taken 0, 3, 4, 5 and 6 dpi and fresh medium of the same amount added. Viral titers of each sample were assessed via TCID₅₀ (see section 2.3.6).

2.3.9 Immunofluorescence analysis

For immunofluorescence analysis (IFA), 1x10⁵ HuH7 cells (50% confluency) were seeded onto cover slips and transfected with 500 ng plasmid coding for sVP40 WT or mutants in pCAGGS. 24 h pt cells were fixed with 4% PFA for 20 min. Fixed cells were treated with 100 mM glycine in PBS_{def} and permeabilized using 0.1% triton X-100 in PBS_{def} before incubation in blocking buffer (2% bovine serum albumin (BSA), 5% glycerol, 0.2% Tween-20, 0.05% Sodium azide in PBS_{def}) for 10 min.

Protein-specific staining was performed using polyclonal rabbit anti- (α) sVP40 or monoclonal mouse α-zVP40 (Sino Biological) antibodies, diluted 1:1000, and goat α-rabbit Alexa 488 at a dilution of 1:300 or goat α-mouse Alexa 488 as a secondary antibodies, together with DAPI for nucleus staining. Cellular distribution of VP40 filaments and clusters was assessed using a Leica DMI6000 CS and different VP40 phenotypes were quantified. Confocal images were acquired using Leica TCS SP5 confocal microscope. Images were processed using ImageJ 1.52p and the length of VP40 filaments was measured. In total, 55-76 filaments per condition of 2 independent experiments were measured. Rout outlier test was performed (Q=0.5%) and Kruskal-Wallis test was performed followed by Dunn's multiple comparison

Methods

test. For IFA of recEBOV, infection was performed as described in section 2.3.3.2. These experiments were performed with the kind help of Dr. Michael Klüver, Becker lab, Institute for Virology, Philipps-University Marburg.

2.3.10 Transmission electron microscopy and analysis of particle length of recEBOV

For Transmission electron microscopy (TEM), HuH7 cells were infected at MOI 1-2 and incubated in DMEM containing 3% FBS. At 2 dpi, the supernatants were first clarified with low-speed centrifugation and then subjected to ultracentrifugation for 2 h at 4 °C and 25,000 rpm using an SW32 rotor (Beckman Coulter). Released virus particles were fixed twice for 24 h each with 4% PFA and removed from the BSL4 laboratory after surface decontamination. Fixed viral particles were then negatively stained with 2% phosphotungstic acid. Electron microscopy was carried out at 80 kV on a JEOL JEM-1400 transmission electron microscope (JEOL, Tokyo, Japan) equipped with a TemCam-F416 camera (TVIPS, Gauting, Germany). The length of released virus particles was measured on micrographs with the help of analytical tools in ImageJ [278]. For each virus more than 350 particles were measured in at least three independent experiments. Lengths of viral particles were analysed using Tukey's multiple comparisons test (average length) and Dunn's multiple comparisons test (percentage of viral particles $\geq 1.3 \mu\text{M}$). Experiments described here and in section 2.3.11 were performed with the kind help of Dr. Martin Schauflinger, Becker lab, Institute for Virology, Philipps-University Marburg.

2.3.11 Immunogold labeling of VP40 on whole mounted cells

HuH7 cells were grown on formvar coated TEM nickel grids, and transfected with 500 ng plasmid coding for sVP40 constructs the next day. 1 dpt, cells were fixed with 4% PFA. For immunogold labeling, cells were treated with 0.1% Triton-X 100, blocked with 1% BSA-c. VP40 was detected using antibodies (rabbit α -VP40 Sudan) followed by secondary antibodies conjugated to colloidal gold (α -rabbit 10 nm). Then, cells were briefly treated with 2% phosphotungstic acid and imaged with a JEM-1400 at 80kV.

2.4 Protein biochemistry

2.4.1 Expression in *E. coli*

For the overexpression of a protein of interest in bacteria, plasmids were transformed into an *E. coli* strain suitable for heterologous expression. A single colony was used for the inoculation of a preculture, usually 5 mL LB medium containing the appropriate antibiotic (as all following cultures). This preculture was incubated for several hours at 37 °C and 200 rpm and then used for the inoculation of the overnight culture (100 mL LB medium, typically) which was shaken overnight at 37 °C and 140 rpm. For the final culture, 1 L LB medium in a baffled Erlenmeyer flask was inoculated with approx. 35 mL of the overnight culture. Cells were grown at 37 °C and 140 rpm until an OD_{600} of > 0.6 was reached. The culture was cooled down at RT for usually 3 h and the production of recombinant proteins was induced by adding isopropyl-beta-D-thiogalactopyranoside (IPTG) to a final concentration of 1 mM. After induction, the culture was shaken for 16 h at RT and 140 rpm before being transferred to

Methods

1 L centrifugation bottles and centrifuged at 2,500 rpm and 4 °C for 30 min. The bacterial pellet was washed once with PBS_{def} and frozen at -20 °C until further use. This protocol was adapted from Bornholdt *et al.* 2013.

2.4.2 Protein purification

2.4.2.1 Cell lysis

Bacterial pellets were thawed on ice and resuspended in lysis buffer (25 mM Tris, 300 mM NaCl, pH 8, and 10 mM imidazole) and lysed mechanically using a microfluidizer (Microfluidics), kindly provided by the Bange lab, Philipps-University Marburg. Equilibration of the microfluidizer was achieved with ice-cold lysis buffer. Afterwards, the bacterial suspension was introduced into the system. After 20 pumps, the lysed suspension was collected and the tubes were washed with ice-cold lysis buffer. The lysate was then subjected to centrifugation for 30 min at 4°C and 20,000 rpm. The cleared supernatant was mixed with 1 mL nickel-nitrilotriacetic acid (Ni-NTA; 1 ml/L expression culture) slurry and incubated overnight at 4 °C using an overhead shaker. To determine the ratio of soluble vs. insoluble protein, the pellet was resuspended in the same volume of PBS_{def} as the volume of the cleared supernatant. Samples of both the pellet and supernatant were taken for downstream analysis.

2.4.2.2 Purification via immobilized metal affinity chromatography

The purification of proteins of interest via immobilized metal affinity chromatography (IMAC) exploits the binding of the protein fused to a 6x histidine tag (his-tag) to the nickel ions of the resin, Ni-NTA coupled to Sepharose®CL-6B beads in 20% ethanol. The protein-slurry mixture was applied to a chromatography column using the column batch purification method. After collection of the flow-through, the matrix was washed with 4x 10 mL washing buffer (25 mM Tris, 300 mM NaCl, pH 8, and 20 mM imidazole) and the protein of interest was eventually eluted with 2x 10 mL elution buffer (25 mM Tris, 300 mM NaCl, pH 8, and 250 mM imidazole). The underlying principle was the competitive binding of imidazole to the Ni-NTA beads due to its structural similarity with the histidines. High imidazole concentrations strip his-tagged proteins off the matrix, whereas low concentrations (10 to 20 mM) can avoid unspecific binding. A 30 µL sample of each fraction was taken and mixed with 4x SB, incubated for 5 min at 95 °C and afterwards a sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) with a subsequent Coomassie staining was performed.

2.4.2.3 Protein concentration

Elution fractions of both IMAC and SEC (see sections 2.4.2.2 and 2.4.2.4) were judged on both the purity and intensity of SDS bands of interest and chosen fractions were pooled and concentrated using the Amicon Ultra concentrators with a molecular weight cut-off at 10,000 Da. Concentrators were first equilibrated with protein buffer (25 mM Tris, 300 mM NaCl, pH 8) by centrifugation for 10 min at 4,500 rpm and 4 °C. The protein solution was then applied to the concentrator and the centrifugation steps were repeated until the desired volume or concentration was reached, usually 2 mL of final

Methods

volume for downstream size exclusion chromatography (SEC) purification of further analysis. The protein concentration was determined using a NanoDrop photometer (Peglab).

2.4.2.4 Purification via size-exclusion chromatography

Using SEC, proteins can be further separated from impurities or other homo-oligomeric forms. This method is based on the different hydrodynamic diameters of each analyte. This means that proteins in the native form exhibit different elution volumes which roughly depends on their size and shape [143]. For the SEC procedures of the present study, a HiLoad 16/60 Superdex 200 column (Cytiva) was used. This column was prepacked with Superdex resin beads which consisted of a dextran matrix bound to cross-linked agarose. The column was equilibrated with degassed and filtered buffer (25 mM Tris, 300 mM NaCl, pH 8) before the concentrated protein solution was applied into the chromatography system using a syringe. The proteins passed through the column with a flow rate of 1 ml/min where proteins with smaller hydrodynamic radii were retarded longer by the column matrix due to the ability to diffuse into pores of the column resin. The presence of proteins in the buffer was measured at a wavelength of 280 nm and was depicted in the UNICORN software as milli absorbance units (mAU). The eluted protein was collected in 1 mL fractions. Fractions of interest were analyzed via SDS-PAGE and Coomassie staining to check for the correct bands and purity. Selected fractions were pooled and concentrated using Amicon Ultra concentrators until the desired volume or concentration was reached. Proteins could be stored at 4 °C for up to five weeks or mixed with 15% glycerol, aliquoted and frozen at -80 °C for long-term storage until further use. For the present study, SEC was used for preparative protein purification as well as analytical investigation of the homo-oligomerization state.

2.4.3 Sodium dodecyl sulfate-polyacrylamide gel electrophoresis

To separate proteins according to their molecular weight, a SDS-PAGE was performed. The formation of negatively charged complexes of proteins with SDS enabled the migration in an electrical field and the addition of β -mercaptoethanol reduces disulfide bonds. A protein ladder indicated the approximate size of the denatured proteins. For SDS-PAGE, proteins in solution were mixed with 4x SB (final concentration 1x), whereas whole cells lysates were resuspended in 1x SB. The protein-SB mixes were additionally incubated for 5 min at 95 °C and usually 5 to 10 μ L were applied to the pockets of the gel after cooling down. After the samples passed the 4% stacking gel, the actual separation took place in the 12% separation gel. The gel was run in a SDS electrophoresis chamber SDS-PAGE running buffer for 45 to 50 min at 200 V. Afterwards, the proteins were detected via Coomassie staining (see section 2.4.3.1) or via WB analysis (see section 2.4.3.2).

2.4.3.1 Coomassie staining

To roughly determine both the purity and the amount of a protein (mixture) of interest, Coomassie staining after SDS-PAGE was performed. For this, the gel was placed in H₂O until it cooled down and residual SB traces were destained. The gel was then incubated in Coomassie staining solution for

Methods

approx. 15 min (or 5 min when heated up), followed by the Coomassie destaining solution until protein bands were clearly visible.

2.4.3.2 Western blot

To specifically detect proteins of interest with protein-specific antibodies, WB analysis was performed where proteins are blotted from the gel onto a membrane using the so-called semi-dry blot system. The setup of the blotting sandwich between anode (bottom) and cathode (top) was as follows (from bottom to top): 3 Whatman papers, SDS-PAGE, nitrocellulose membrane, 3 Whatman papers. Each component was soaked in WB transfer buffer before stacking of the sandwich. Excessive buffer was removed and one of the following protocols applied: 1) Standard: 25 V; 1 A; 30 min or 2) Turbo: 25 V; 2.5 A; 7 min. Afterwards, the blotting success was checked by staining with Ponceau S solution. For this, the gel was placed in Ponceau S staining solution for 1 min and washed with H₂O until bands were clearly visible. To completely destain the membrane, WB washing buffer (PBS_{def} with 0.1% Tween 20) was applied. Blocking of unoccupied sites on the membrane by incubating in 10% skim milk for 1 h at RT or overnight at 4 °C reduced unspecific antibody binding. Primary antibodies were diluted 1:500 to 1:4,000 in 1% skim milk and incubated with the membrane again for either 1 h at RT or overnight at 4 °C which enabled binding to the epitope of the proteins of interest. After washing the membrane for 15 min in PBS_{def} with 0.1% Tween 20, the primary antibodies were added at a 1:5,000 dilution in 1% skim milk, coupled to a fluorophore and targeting the species of the primary antibody. The final washing steps included 15 min in PBS_{def} with 0.1% Tween 20 followed by 5 min in PBS_{def}. For imaging and quantification of protein bands the Odyssey® Infrared Imaging Scanner with the ImageStudio™ Acquisition software was used.

2.4.4 Protein crystallography

Crystallography is one of the most widely used methods for structural determination of biological macromolecules, next to nuclear magnetic resonance (NMR) and (cryo) electron microscopy (EM). Usually, all datasets of protein crystals are deposited into the openly accessible Protein Data Bank (PDB) prior to publication. *De novo* protein structure determination as well as identification of crystallographic binders such as small molecules are two of the main goals of protein crystallography.

2.4.4.1 Principle

Proteins can be forced into the crystalline state by slowly decreasing the solubility of the protein and simultaneously raising the concentration of precipitants in a controlled fashion. Figure 8 depicts a crystallization phase diagram.

Methods

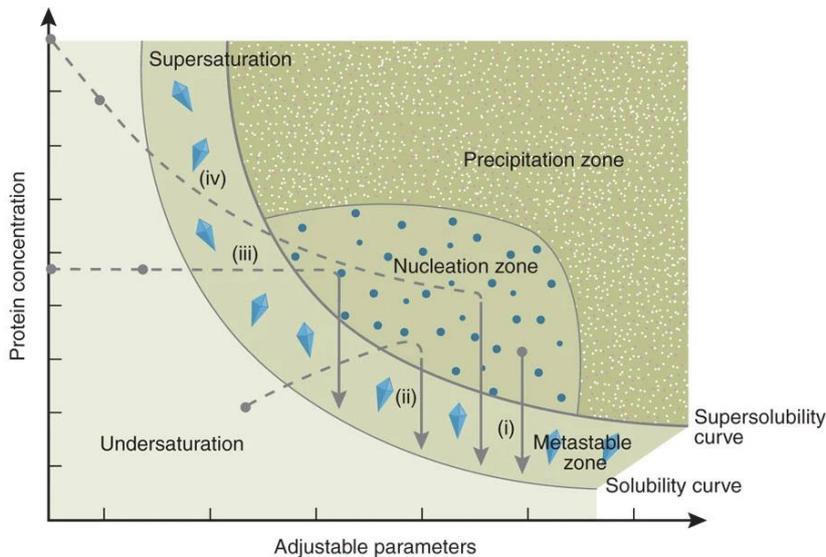


Figure 8: Phase diagram of protein crystallization. Adjustable parameters are precipitant and its concentration, pH, temperature, additives and their concentration. The crystallization curves of the four major routes are depicted: (i) microbatch, (ii) vapor diffusion, (iii) dialysis and (iv) free-interface diffusion (liquid/liquid diffusion). Each method begins at the starting conditions indicated by the large black dots and uses a different route to reach the nucleation and metastable zone and eventually the solubility curve [55].

If successful, crystallization begins when a protein is in the nucleation zone (oversaturation of the protein), forming nuclei, thereby lowering the protein concentration in the solution, resulting in the meta-stable zone. These nuclei then serve as a starting point for protein crystals. Both the nucleation event and the growth of the crystal can take from days to several weeks [32].

2.4.4.2 Crystallization

Vapor diffusion is the most commonly used technique for producing protein crystals. It is usually performed in well plates and each well is filled with the so-called mother or reservoir solution containing the chosen crystallization conditions such as selected buffers, salts and precipitants. The protein solution as a hanging or sitting drop slowly equilibrates against the reservoir solution, thereby increasing the concentration of the precipitating agents and decreasing the solubility of the protein in the drop over time, thus favoring crystallization of the protein (Figure 9) [223].

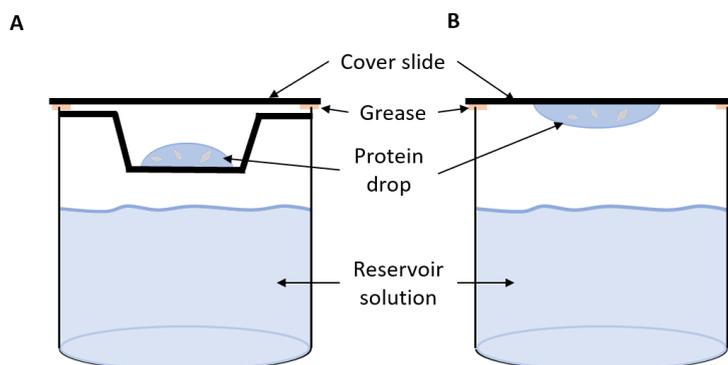


Figure 9: Vapor diffusion sitting vs. hanging drop. The drop containing the protein-reservoir mix is either placed on a bridge (sitting drop, A) or on a cover slide (hanging drop, B). The well is sealed enabling a closed microenvironment for slow equilibration. Modified [223].

The initial screening process for crystallization conditions for recombinant VP40 was performed at the Sapphire lab, SCRIPPS Institute in La Jolla, California, USA. The plasmids from the original publication by Bornholdt *et al.* [32] were used and included expression plasmids (pET46 EK/LIC) containing genes coding for both full-length and truncated (deletion of the first 43 a

Methods

N-terminal amino acids, $\Delta 43$) versions of z/sVP40. The expression was performed as reported (see also section 2.4.1). However, purification and crystallization were optimized and simplified. Recombinant protein (z/sVP40 $\Delta 43$) was purified as described in sections 2.4.2.2 and 2.4.2.4 and frozen at $-80\text{ }^{\circ}\text{C}$ in 15% glycerol. After thawing the protein, it was equilibrated in the corresponding protein buffer using Zeba Spin 7K MWCO desalting columns until the desired concentration was reached. For crystal screenings, the protein of interest was sent to the MarXtal crystallization facility (Philipps-University Marburg). Both protein and screening buffers of commercially available crystallization screens were pipetted into MRC 2 well plates (Swissci) using a Honeybee 963 crystallization robot (Digilab). The final volume of the so-called mother solution in the well was 80 μL of which 0.3 μL were mixed with 0.3 μL protein as a sitting drop experiment. Plates were sealed and incubated at $18\text{ }^{\circ}\text{C}$ in a Rock Imager (Formulatrix) crystallization imager. The formation and growth of crystals was documented and checked daily. For manual crystallization set-ups, the protein was then mixed 1:1 with crystallization buffer in a total volume of 2 μL on a cover slip and crystallized using the vapor diffusion hanging drop method with 500 μL crystallization buffer as the reservoir solution. Crystals of sVP40 $\Delta 43$ WT usually grew overnight at $18\text{ }^{\circ}\text{C}$.

2.4.4.3 Co-crystallization

To crystallize protein-peptide or protein-fragment complexes, co-crystallizations were set up. Peptides were dissolved in DMSO (final concentration of 100 mM), diluted in protein buffer to a final concentration of 1 mM and mixed 1:1 with VP40. Fragments were dissolved in protein buffer to a final concentration of 500 mM but were usually hardly soluble. The suspension was centrifuged and the supernatant mixed with sVP40 $\Delta 43$. Co-crystallization mixtures were incubated on ice and crystallization drops were set as described in the previous section.

2.4.4.4 Harvest and soaking of protein crystals

Soaking of crystals served to identify low-binding ligands or fragments without the need to optimize the co-crystallization conditions for each protein-ligand combination. As fragments were dissolved in DMSO, preliminary tests involved soaking sVP40 $\Delta 43$ WT crystals with varying DMSO concentrations (1 to 20 mM) diluted in VP40 crystallization buffer with varying soaking durations (overnight to 24 h).

Fragments from the Jena Biosciences Frag Xtal Screen were kindly supplied by the Klebe lab (Institute of Pharmaceutical Chemistry, Philipps-University Marburg). For the composition of the library see section 8.13.2. 1 M stock solutions of fragments in DMSO were diluted 1:10 in crystallization buffer to a final concentration of 100 mM. Crystals of dimeric sVP40 $\Delta 43$ WT were soaked in 1 μL of fragment solution in crystallization buffer and then transferred to the same solution with 20% ethylene glycol as a cryoprotectant (crystallization cryo buffer). If the crystal endured visible damage during soaking such as cracks or rounding of its edges or if the diffraction data did not meet the desired criteria, soaking

Methods

times were adjusted to either 1 h or 5 sec. In this case of short incubation times crystals were directly soaked with fragments in crystallization cryo buffer. Untreated crystals were harvested by either soaking them briefly in crystallization cryo buffer or flash-freezing them directly in liquid nitrogen. Crystals were then shipped to the beamline or stored in liquid nitrogen.

2.4.4.5 X-Ray diffraction and data collection

This section covers both the key concepts of crystal geometry as well as the underlying principles of X-ray diffraction of protein crystals.

Protein crystals are highly ordered assemblies of molecules with certain geometric symmetries. Each protein crystal can be assigned to one of 65 space groups, as described in the International Tables for Crystallography, Volume A. For the determination of the space group, the crystals are regarded as a complex of three-dimensional lattices that can be described using the vectors a , b and c in directions x , y and z with angles α , β and γ which define the unit lattice (Figure 10 A). When combined with a molecular motif, the unit cells are formed, which are infinitely assembled units in each direction in the protein crystal, thereby repeating the same number of molecules with the same geometry. In most cases, the unit cells consist of several protein molecules. The smallest building block of a protein crystal is the asymmetric unit (ASU). This unit cannot be divided into smaller fragments but is repeated by various symmetry operations such as rotation, reflection or translation. In the easiest case, the ASU consists of a single protein molecule in its monomeric form. Several ASUs then make up the unit cell as the ASU contains all the information needed to generate the unit cell when all the required symmetry operations are applied (Figure 10 B). Depending on this symmetry and the combination of translation and rotation, the crystal belongs a certain space group, such as space group 5, also called C2. This space group is the most common one for the present work and represents a monoclinic symmetry with $a \neq b \neq c$ and $\alpha = \gamma = 90^\circ$ and $\beta \neq 90^\circ$ [265].

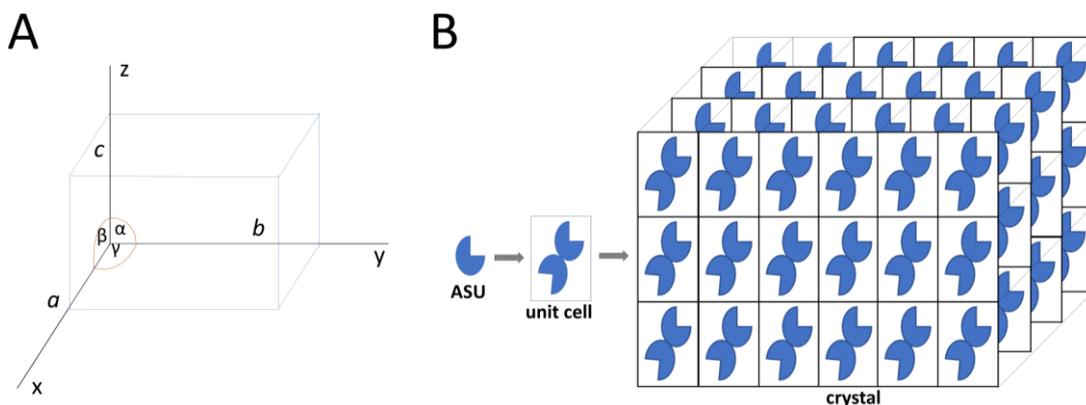


Figure 10: Three-dimensional crystal lattice and the building blocks of a protein crystal. A) The vectors a , b and c describe the crystal lattice in 3D. B) The ASU is the smallest building block of a protein crystal and usually several ASUs are found in the unit cells which is repeated in all directions to form the crystal. Adapted from [265].

Methods

When protein crystals are subjected to X-rays, the beams are diffracted and generate a diffraction pattern on the detector screen. This diffraction can be explained by the Bragg equation. Protein crystals are a set of parallelly spaced lattice planes (Miller planes (hkl)) with lattice indices hkl and each plane mirrors the incident beam which results in constructive interference of the beams, when they remain in phase, or destructive interference. Bragg's law describes the condition under which the cumulative effect of constructive interference is maximised using the following formula in Figure 11 [34,259]:

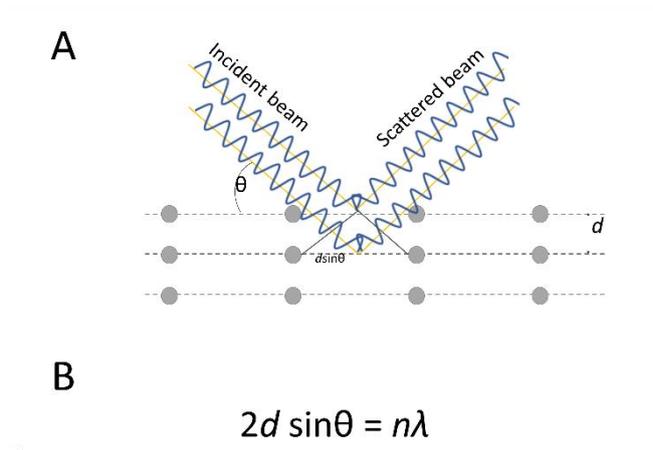


Figure 11: Schematic of Bragg's law. A) X-rays hit scattering elements on the Bragg planes. B) If the equation is true, Bragg's law is fulfilled and constructive interference occurs, resulting in diffraction patterns. d = spacing between crystallographic planes; θ = angle between plane and X-ray; n = integer; λ = X-ray wavelength. Adapted from [259].

X-ray beams with the wavelength λ meet scatterers at the imaginary *Bragg planes* – which are separated by the distance d – at an angle of incidence θ . When the total path length difference $2d \sin \theta$ is an integer number of λ , Bragg's law is fulfilled and constructive interference will occur. All datasets for this work were collected at the Swiss Light Source (SLS) of the Paul-Scherrer Institute, Villigen, Switzerland, at the PXI beamline. Crystal handling was performed by an automatic sample changer. Protein crystals were placed on the goniometer in a constant stream of liquid nitrogen, its position carefully adjusted and subjected to very bright monochromatic X-ray beams which were scattered by the electrons in the protein and projected onto a detector. All produced reflexions on the detector make up the diffraction pattern. The distance between the detector and the crystal was changed according to the intensity of reflexions and the resolution (the higher the resolution the shorter the distance). To achieve a full three-dimensional picture, the crystal was rotated on the goniometer so diffraction patterns of different angles can be detected and radiation damage is limited. To assess ideal exposure times, distance between crystal and detector and covered angles, two test exposures at typically 0° and 90° were performed and results interpreted visually based on observed spots and also with the help of the SLS *DA+* software [327] which estimated resolution and space group among other criteria. Parameters for data collection was adjusted according to the two test exposures. For the present work, a resolution range of usually 1.4 to 3.5 Å (as determined by detector distance), exposure time of 0.01 sec, oscillation angle 0.1 or 0.2, oscillation range 180 to 360° and 70% beam intensity at a wavelength of approx. 1 Å were employed, resulting in 1800 to 3600 frames per dataset.

Methods

2.4.4.6 Processing of diffraction data

Directly after data collection, XDS [163] was launched, a program suite to process crystallographic data. In most cases, datasets had to be reprocessed manually with adjusted parameters so the quality of the resolution shells meets the selected criteria (described below). This was performed by modifying the XDS input file (XDS.INP). Typical corrections were made for software requirements, path to the raw data of each dataset, detector information, resolution (subdivided into resolution shells) and possibly also space group. In addition, Friedel's law was changed to "true" and jobs performed by XDS were switched to "DEFPIX INTEGRATE CORRECT" instead of "XYCORR INIT COLSPOT DEFPIX INTEGRATE CORRECT". The latter resulted in rerunning the jobs with less stringent criteria and was only performed in a failure of XDS to fully process the data as indicated by "Insufficient percentage (<20%) of indexed reflections, automatic data processing stopped". The CORRECT.LP file was used for data interpretation and typically, the resolution had to be adjusted based the following criteria:

- Data completeness as high as possible, with an overall completeness of at least 90%
- R-factor as low as possible
- I/σ as high as possible, but at least > 0.6 in the outer shell
- $CC(1/2) > 0.5$ in the outer shell

Usually, a compromise between resolution and the above mentioned parameters could be reached, with strict cut-off only defined for $CC(1/2)$ [164]. Processing of collected X-ray diffraction data consisted of several steps. First, spatial correction values were calculated for each pixel (XYCORR), the background was determined and differentiated towards the pixels (INIT), spots with high intensities were identified (COLSPOT) and used for the estimation of possible space groups (indexing, IDXREF), removal of areas outside of the defined resolution (DEFPIX) and the actual intensities of each spot were calculated (INTEGRATE) and corrected for decay, absorption or variations in the sensitivity of the detector surface (CORRECT). The last step also produced the "CORRECT.LP" file containing processing statistics for each resolution shell and the final output "XDS_ASCII.HKL" file [163]. The latter was then used for data reduction using AIMLESS within the CCP4i suite [90] during which the space group is assessed/confirmed for a second time and a .mtz file is generated which was used for molecular replacement (see section 2.4.4.7). In addition, a so-called free-R set was created using 5% of the reflections which was used to compare the model with the raw data to avoid the introduction of modelling bias [255].

2.4.4.7 Phasing

In protein crystallography, one differentiates between the real space R and the reciprocal space R^* . The crystal itself belongs to the real space, as does its lattice and the electron density. The diffraction pattern on the other hand is part of the reciprocal space. R can be used to calculate R^* but not the

Methods

other way around, as the critical phase information of the X-ray waves is missed during data acquisition and this is referred to as the “phase problem” of crystallography [67]. The phase problem can be solved via various methods such as anomalous dispersion, where the introduction of anomalous scatterers such as heavy metals are introduced into the crystal, called single-wavelength anomalous diffraction (SAD) or multi-wavelength anomalous diffraction (MAD) experiments, or molecular replacement (MR). Only MR was used for phasing for the present work and it describes the usage of structure factors of homologous or similar proteins as a phasing model. In the case of isomorphous phasing models, the phases of the template protein are directly introduced in the Fourier sum which is used to calculate the real space R – the electron density of the unit cell – from the reciprocal space R*:

$$\rho(x, y, z) = \frac{1}{V} \sum_h \sum_k \sum_l |F_{hkl}^{target}| e^{-2\pi i(hx+ky+lz-\alpha'_{hkl}^{model})}$$

In this equation, $\rho(x, y, z)$ describes the function of the electron density at position xyz, V is the volume of the unit cell, $|F^{target}(hkl)|$ signifies the structure factor amplitudes of the new protein and $\alpha'^{model}(hkl)$ stands for the phases of the template protein. This phasing step serves to generate the electron density map into which the amino acids chain can be placed [259]. For the present work *Phaser* of the *Phenix* suite was used for molecular replacement [221].

2.4.4.8 Structure Refinement

As the model is only roughly aligned to the electron density, an iterative process of reciprocal structure refinement with manual model building follows phasing: positions of amino acid back bone and side chains are corrected, occupancies are adjusted, additional residues are built in as well as water/solvent molecules or possibly ligands using *coot* [84]. The resulting .pdb file is then subjected to several refinement rounds to eventually improve the density fit between the modelled structure (interpretation of the original data) and the electron density map (original data) (Figure 12 A). There are several statistical parameters that indicate how well the refinement process is progressing, the most important ones being the two R-factors, R_{work} and R_{free} . R_{work} describes how well the electron density is explained by the model but this might include the introduction of overfitting bias which is indicated by R_{free} . As mentioned above, by default 5% of the original experimental data is not included in the refinement process but is used for comparing the agreement between calculated and observed data. The equation is as follows:

$$R = \frac{\sum ||F_{obs}| - |F_{calc}|}{\sum |F_{obs}|}$$

F_{obs} indicates the structure factor amplitudes of the experimental data, whereas F_{calc} describes the structure factor amplitudes calculated from the model. R_{free} is always higher than R_{work} (approx. 2 to 5%) but during refinement they should both be lowered more or less equally. Overfitting is indicated

Methods

by a strong difference between the two values (typically >5%). Other criteria include B-factors, occupancies, angle and distance restraints, bond lengths, binding and torsion angles, the existence of Ramachandran or rotamer outliers and clash scores [197].

Refinement restraints for the present work usually included XYZ coordinates, Individual B-factors, TLS parameters, Occupancies and Optimize X-ray/stereochemistry weight. TLS parameters for each structure were determined using the Phenix TLS tool.

Data processing, phasing and refinement for one crystallographic data set (AW122) were performed using autoPROC [313] and BUSTER [37] with the assistance of Dr. Michael Norris, Institute for Biochemistry, University of Toronto.

For screening purposes, three automated pipelines were used. While the tool *dimple* (Difference map pipeline) is usually used for the identification of positive electron density, for the present work, it was used to refine a large number of datasets as a preparation for further investigation using PanDDA (see below). It is executed in the ccp4i suite and requires a template model for molecular replacement (.pdb and .mtz) file as well as the .mtz file of the new structure. It runs several different steps of refinement (<https://github.com/ccp4/dimple>) [328] and the dimple output files were then used for the analysis of potential “blobs” of positive electron density using the PanDDA suite (Pan-Dataset Density Analysis), a tool for the identification of ligand binding and so-called events of numerous datasets simultaneously [253] (Figure 12 B).

For the execution of PanDDA, a separate folder had to be created for each dataset containing the refined .pdb and .mtz files from dimple. The folders were all named in the same style (“AW**.*dimple.pdb*” und “AW**.*dimple.mtz*”). The analysis was started using an Ubuntu system with the commands:

```
pandda.analyse data_dirs=~ /Dokumente/PANDDA/* pdb_style=*.dimple.pdb  
mtz_style=*.dimple.mtz out_dir= ~/Dokumente/PANDDA/PanDDA_pep cpus=40
```

PanDDA’s output included the identification of datasets with events which mean regions that differed (structurally) from the average of all input structures. These event datasets were first checked visually in the PanDDA output maps and then a typical data analysis was performed using *Phaser* for molecular replacement, *Phenix Refine* and *coot*. Several refinement cycles were performed until the positive electron density was either identified as belonging to residues missing from the structure, water/solvent molecules or ligands. Once ligands were identified, PanDDA was rerun without the ligand-containing maps containing the additional command:

```

pandda.analyse data_dirs=~ /Dokumente/PANDDA/* existing_datasets=ignore
pdb_style=*.dimple.pdb mtz_style=*.dimple.mtz out_dir=~/Dokumente/PANDDA/ cpus=40

```

Additionally, the automated refinement pipeline developed by Schiebel *et al.* was performed [276]. As this tool was written only for certain (older) versions of XDS, ccp4i and Phenix, this analysis was kindly performed by Prof. Dr. Andreas Heine (Institute of Pharmaceutical Chemistry, Philipps-University Marburg) with the respective software requirements. As input files, all .mtz files of interest as well as one template model for molecular replacement were needed. The pipeline output consisted of a new folder for each dataset with data for each performed step. Steps included 1) MR, 2) Simulated annealing, 3) standard refinement, 4) TLS refinement, 5) water inclusion, 6) deletion of selected waters, 7) anisotropic B except water, 8) anisotropic B, 9) hydrogens included, 10) water inclusion and 11) deletion of selected waters. Log-files of each step of each dataset were evaluated and the step with the best R-factor chosen for further visual inspection. Again, as for the promising datasets identified by PanDDA, the typical procedure of manual molecular replacement and refinement cycles were performed (Figure 12 B).

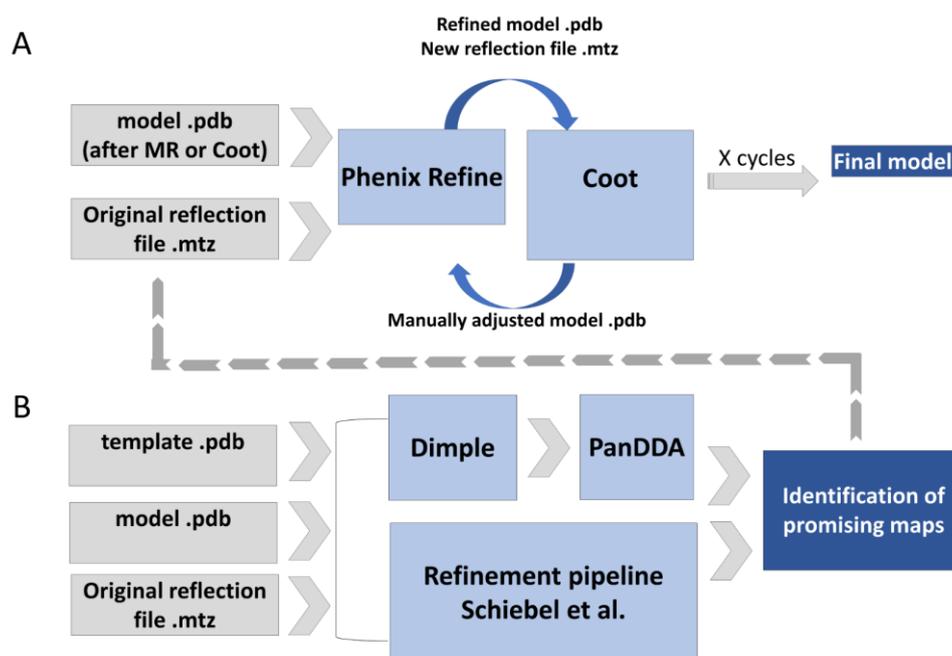


Figure 12: Refinement process for the present work. A) The typical refinement process started with the original reflection and model file. After molecular replacement, several iterative cycles of refinement and manual building were performed which eventually resulted in the final model. B) Screening for crystallographic binders was performed using two approaches for the identification of promising maps. These maps were then manually evaluated by following the steps outlined in A.

2.4.5 Disulfide Tethering

As described in the introductory part, disulfide tethering (DT) is a method used for site-directed ligand discovery and exploits the formation of disulfide bonds between the protein of interest and a thiol-

Methods

containing ligand, thus greatly enhancing the affinity and enabling the analysis via intact protein mass spectrometry. The addition of a reducing agent facilitates the binding of only favorable ligands that form more interactions towards the target protein than just the disulfide bridge.

As sVP40 has two cysteines at the very end of the C-terminal domain (C314 and C320), an additional cysteine was introduced in the vicinity of the dimeric interface and DT experiments were performed on sVP40 WT as well as two mutants carrying the N67C mutation, one including the two native cysteines (sVP40 N67C) and one carrying additional mutations of the cysteines to alanines (sVP40 N67C-C314A-C320A (CCA)). The thiol-containing fragments were kindly provided by the Klebe lab, Institute of Pharmaceutical Chemistry, Philipps-University Marburg (see section 8.13.1 for the complete list of used fragments). Dr. Andreas Nguyen kindly assisted with advice [241].

2.4.5.1 DT Screening protocol

The reaction setup for the initial screening was as follows: 30 μM sVP40 $_{\Delta 43}$ N67C (CCA) were added to 66.7/133/200 μM ligand and 500 μM DTT, ad 30 μL VP40 buffer. Reactions without ligands were used as negative controls for each reaction batch. The mix was then incubated for 1 to 2 h on ice, followed a 2 h incubation step at RT. The reaction was then desalted using Zebaspin desalting columns to remove DTT and unbound ligands. The flow-through was collected and sent to the *MarMass* facility, Marburg and analyzed via intact protein mass spectrometry (see section 2.4.8.2).

As the fragments had different molecular masses, it was possible to combine up to three ligands per reaction. The following table shows the fragment combinations (with molecular weights in Dalton (Da)) for the initial screenings (negative controls not included) (Table 1). SB100 and SB101 were available as filtrates and recess. By accident, reaction cocktail #22 contained two fragments with identical masses.

Table 1: DT screening protocol #1. 30 μM sVP40 $_{\Delta 43}$ N67C was incubated with up to three thiol-containing fragments and 500 μM DTT. Reactions without ligands were used as negative controls. The mix was then incubated for 1 to 2 h on ice, followed by 2 h at RT. The reaction was desalted using Zebaspin desalting columns to remove DTT and unbound ligands. The flow-through was collected and sent to the *MarMass* facility, Marburg, and analyzed via intact protein mass spectrometry.

#	Fragment 1	Mass [Da]	Fragment 2	Mass [Da]	Fragment 3	Mass [Da]
1	SB104	145.22	SB115	231.32	TK103	265.38
2	SB109	207.3	TK094	250.32	SB124	307.15
3	SB127	161.27	SB114	211.28	TK169	251.35
4	TK135	171.22	SB228	237.32	TK163	269.37
5	SB106	173.16	TK148	237.32	TK098	273.40
6	SB095	181.26	SB103	225.31	TK109	262.38
7	SB123	187.31	SB233	237.34	TK106	276.4
8	TK149	211.28	TK170	251.35	SB227	285.41
9	TK157	211.28	TK096	251.37	TK176	283.39
10	SB101	213.28	TK095	252.31	TK108	294.20
11	SB101	213.28	TK100	264.35	TK162	243.39
12	SB120	229.39	SB229	280.17	TK099	313.47
13	Sb128	250.37	TK105	308.23		
14	Sb107	234.32	TK102	265.40		
15	TK112	236.29	TK101	266.34		

Methods

16	SB100	195.29	SB113	237.37
17	SB100	195.29	TK111	238.29
18	TK147	197.26	TK097	239.34
19	TK164	199.27	TK158	239.34
20	SB119	215.36	TK104	253.37
21	TK136	234.32	TK175	254.35
22	TK110	225.31	TK140	225.31
23	SB096	226.26	TK142	255.34
24	SB097	227.28	TK171	255.34
25	TK141	229.37	TK107	259.37
26	TK165	229.37	SB099	260.15
27	TK174	239.34	TK173	262.38
28	SB102	240.28		
29	SB226	245.35		
30	SB232	248.35		
31	TK150	248.35		

The protocol outlined in Table 2 was used for sVP40_{Δ43} N67C-CCA:

Table 2: DT screening protocol #2. 30 μM sVP40_{Δ43} N67C-CCA was incubated with up to three thiol-containing fragments and 500 μM DTT. Reactions without ligands were used as negative controls. The mix was then incubated for 1 to 2 h on ice, followed by 2 h at RT. The reaction was desalted using Zebaspin desalting columns to remove DTT and unbound ligands. The flow-through was collected and sent to the MarMass facility, Marburg, and analyzed via Intact protein mass spectrometry.

#	Fragment 1	Mass [Da]	Fragment 2	Mass [Da]	Fragment 3	Mass [Da]
1	SB115	231,32	TK103	265,38	SB124	307,15
2	SB101	213,28	TK169	251,35	SB227	285,41
3	SB101	213,28	TK148	237,32	SB229	280,17
4	TK158	239,34	TK163	269,37	TK099	313,47
5	SB100	195,29	TK112	236,29	TK098	273,40
6	TK110	225,31	TK096	251,37	TK106	276,40
7	TK147	197,26	SB096	226,26	TK109	262,38
8	TK141	229,37	TK174	239,34	TK095	252,31
9	SB113	237,37	TK142	255,34	SB231	207.30

As a first step, the analysis of the mass spectra involved determining the mass of unmodified protein as observed for the negative controls (± 2 Da): 32,430 Da for sVP40_{Δ43} WT, 32,417 Da for sVP40_{Δ43} N67C and 32,356 Da for sVP40_{Δ43} N67C-CCA. The binding of a ligand to the target protein could be detected as a mass shift from the unmodified mass to the addition of the fragment's mass (± 2 Da).

2.4.5.2 DT Optimization protocol

The setup for optimization is shown in Table 3. The protocol was performed as described above.

Table 3: DT optimization protocol. 30 μM sVP40_{Δ43} N67C-CCA was incubated various concentrations of fragments and DTT. The mix was then incubated for 1 to 2 h on ice, followed by 2 h at RT. The reaction was desalted using Zebaspin desalting columns to remove DTT and unbound ligands. The flow-through was collected and sent to the MarMass facility, Marburg, and analyzed via intact protein mass spectrometry.

#	Fragment [μM]	DTT [μM]
1	100	250
2	200	250

Methods

3	500	250
4	1000	250
5	100	500
6	200	500
7	500	500
8	1000	500
9	100	750
10	200	750
11	500	750
12	1000	750

2.4.6 Molecular docking

Molecular docking describes an *in silico* method used for drug design which aims to predict the structural relationship between a protein of interest and a potential ligand. In addition to identifying or generating the so-called docking pose – where and how a molecule binds to the receptor – it ranks the interactions between the protein and a ligand based on a scoring system which depends on the software suite of interest. Scores are based on criteria such as the introduction of clashes, molecular weight of the docked molecule, accuracy of binding site and number of hydrogen bonds or salt bridges, to name only a few. As docking studies are usually employed when a large number of molecules are being investigated, the scoring function and its reliability are of central importance. However, due to the disadvantages of most program suites, determining the most promising docking poses for potential drug leads based on the highest scores remains a challenge. A major downside is the lacking protein-flexibility [224,267].

For the present work, a virtual library was created using the online tool PINGUI [60], accessible via <https://scubidoo.pharmazie.uni-marburg.de/pingui/view/pingui.php>. The resulting molecules were listed using their SMILES codes and redundant compounds were deleted with the kind help of Kai Schmidt of the Kolb lab, Institute of Pharmaceutical Chemistry, Philipps-University Marburg. The preparation of ligands (molecules of the generated virtual library) and receptors (VP40) as well as the actual docking analysis was very kindly performed by Maria Giovanna Papadopoulos with support from Prof. Dr. Peter Kolb, Kolb lab, Institute of Pharmaceutical Chemistry, Philipps-University Marburg. The OpenEye Hybrid Molecular Docking suite was used which uses crystallographic information of a ligand bound to a protein for its algorithms and tries to superimpose docking poses of ligand derivatives with the original ligand. In addition, it allows several different conformations of the same receptor protein to be used [222].

2.4.7 Thermal shift assay

The thermal shift assay (TSA) is usually used to study protein stability or protein-ligand interactions and exploits the stabilizing effects of the ligand on the protein observed as a shift in the protein's

Methods

melting temperature. This shift is detected by SYPRO Orange which exhibits high fluorescence when excited in the presence of non-polar components. Upon denaturation, the protein unfolds and exposes its central non-polar core, which therefore leads to an increase in fluorescence signal which is measured as a melting curve from 25 to 99 °C with a continuous increase of 0.05 °C/sec using a StepOne™ Real-Time PCR system. Expression and purification of the recombinant protein was described in sections 2.4.1 and 2.4.2. The setup was as follows: 20 μM VP40 were added to 10% (v/v) DMSO (optional), 10% (v/v) or 10 to 100 μM fragment/compound and 50X SYPRO Orange, ad 20 μL VP40 buffer. Each reaction was measured in triplicates and 20 μL were used per well of a 96 well plate. Negative controls included samples without protein or without compound.

2.4.8 Mass spectrometry

The proteomics analysis was performed by the *MarMass* facility, Marburg. The protocols for each of the used procedures were supplied by Dr. Uwe Linne (sections 2.4.8.1 and 2.4.8.2) or Dr. Wieland Steinchen (section 2.4.8.3), indicated by quotation marks.

2.4.8.1 Qualitative protein identification mass spectrometry (DQ-MS)

The investigation of post-translational modifications of both zVP40 and sVP40 released as VLPs was already described in section 2.3.2. Briefly, HEK293 cells at 50% confluency were seeded in a T175 flask and transfected with 10 μg plasmid coding for s/zVP40 WT or mutants along with 10 μg plasmid coding for zGP, to increase VLP-yields. 3 dpt, the medium was collected and centrifuged for 10 min at 4 °C and 2,500 rpm to remove cell debris. The supernatant was afterwards subjected to ultracentrifugation for 2 h at 4 °C and 30,000 rpm using an SW32 rotor (Beckman Coulter) to spin down VLPs. The pellet was resuspended in 25 mM Tris, 300 mM NaCl, pH 8 and 0.1% SDS and sent to the *MarMass* facility.

“Proteins were digested by the addition of Sequencing Grade Modified Trypsin (Serva) and incubated at 37 °C overnight. Peptides were desalted and concentrated using Chromabond C18WP spin columns (Macherey-Nagel, Part No. 730522). Finally, peptides were dissolved in 25 μL water with 5% acetonitrile and 0.1% formic acid. The mass spectrometric analysis of the samples was performed using a timsTOF Pro mass spectrometer (Bruker Daltonic). A nanoElute HPLC system (Bruker Daltonics), equipped with an Aurora column (25cm x 75μm) C18 RP column filled with 1.7 μm beads (IonOpticks) was connected online to the mass spectrometer. A portion of approx. 200 ng of peptides corresponding to 2 μL was injected directly on the separation column. Sample Loading was performed at a constant pressure of 800 bar. Separation of the tryptic peptides was achieved at 50 °C column temperature with the following gradient of water/0.1% formic acid (solvent A) and acetonitrile/0.1% formic acid (solvent B) at a flow rate of 400 nl/min: linear increase from 2% B to 17% B within 18 min, followed by a linear gradient to 25% B within 9 min and linear increase to 37% solvent B in additional 3 min. Finally, B was increased to 95% within 10 min and hold for additional 10 min. The built-in “DDA

Methods

PASEF-standard_1.1sec_cycletime" method developed by Bruker Daltonics was used for mass spectrometric measurement. Data analysis was performed using Proteome Discoverer 2.4 (Thermo Scientific) with Sequest search engine together with the Byonic node (Proteinmetrics). Uniprot databases were used. Alternatively, data were analyzed using PEAKS (Bioinformatics Inc.)" (Dr. Uwe Linne, Philipps-University Marburg)

2.4.8.2 Intact protein mass spectrometry (IP-MS)

Intact protein mass spectrometry was primarily used for the screening method disulfide tethering (see section 2.4.5). Briefly, sVP40_{Δ43} WT, N67C or N67C-CCA were expressed and purified as explained in section 2.4.1 and 2.4.2 and were incubated with DTT and thiol-containing fragments. Afterwards, the mixture was purified using Zebaspin Desalting columns to remove unbound fragments and sent to the *MarMass* facility. "Depending on their concentration and the expected protein masses, 1-10 μL of the buffered protein solutions were desalted online using a Waters ACQUITY H-Class HPLC-system equipped with a MassPrep column (Waters). Desalted proteins were eluted into the electrospray ionisation (ESI) source of a Synapt G2Si mass spectrometer (Waters) by the following gradient of buffer A (water/0.05% formic acid) and buffer B (acetonitrile/0.045% formic acid) at a column temperature of 60 °C and a flow rate of 0.1 ml/min: isocratic elution with 5% A for 2 min, followed by a linear gradient to 95% B within 8 min and holding 95% B for additional 4 min. Positive ions within the mass range of 500-5000 m/z were detected. Glu-Fibrinopeptide B was measured every 45 sec for automatic mass drift correction. Averaged spectra were deconvoluted after baseline subtraction and eventually smoothing using MassLynx instrument software with MaxEnt1 extension." (Dr. Uwe Linne, Philipps-University Marburg)

2.4.8.3 Hydrogen-deuterium exchange mass spectrometry

Hydrogen-deuterium exchange is typically coupled to mass spectrometry (HDX-MS) and enables indirect structural determination, conformational dynamics and folding. The naturally occurring proton exchange of backbone amides can be exploited by incubating the protein of interest in deuterated solutions and measuring the exchange rate of hydrogens to deuterium. This exchange depends on both accessibility of the amide in question as well as temperature and pH. Amides strongly involved in the in PPIs or deeply buried will therefore show slow to hardly any exchange [46,216,238]. For the present work, HDX-MS experiments were performed with the kind help of Dr. Wieland Steinchen, LOEWE Center for Synthetic Microbiology & Department of Chemistry, Philipps-University Marburg.

For the following experiments, sVP40_{Δ43} WT and mutants were expressed and purified as explained in section 2.4.1 and 2.4.2 and sent to the *MarMass* facility. "sVP40_{Δ43} WT was mixed prior HDX-MS with sodium salicylate to reach final concentrations of 50 μM and 10 mM, respectively. For studies on the

Methods

impact of site-directed variants on the conformation of sVP40, 25 μ M concentrated protein solutions of sVP40 $_{\Delta 43}$ WT/L117A/W95A were employed. HDX-MS experiments were conducted essentially as described previously aided by a two-arm robotic autosampler (LEAP Technologies) [251]. In brief, 7.5 μ l of sVP40 $_{\Delta 43}$ WT/L117A/W95A solution were mixed with 67.5 μ l of D₂O-containing buffer (25 mM Tris-Cl pH 8, 300 mM NaCl) to initiate the exchange reaction and incubated for 10, 30, 100, 1,000 or 10,000 s at 25 °C. Afterwards, 55 μ l of sample were withdrawn from the reaction and mixed with an equal volume of quench buffer (400 mM KH₂PO₄/H₃PO₄, 2 M guanidine-HCl, pH 2.2) kept cold at 1 °C, and 95 μ l of the resulting mixture was injected into an ACQUITY UPLC M-Class System with HDX Technology (Waters) [315]. Undeuterated samples were prepared similarly by 10-fold dilution in H₂O-containing buffer. sVP40 $_{\Delta 43}$ WT/L117A/W95A was digested online employing a column (2 mm x 2 cm) packed with immobilized porcine pepsin at 12 °C under a constant flow (100 μ l/min) of water + 0.1% (v/v) formic acid and the resulting peptic peptides trapped on an ACQUITY UPLC BEH C18 VanGuard Pre-column, 130 Å, 1.7 μ m, 2.1 mm x 5 mm (Waters) that was kept at 0.5 °C. After 3 min, the trap column was placed in line with an ACQUITY UPLC BEH C18 1.7 μ m 1 x 100 mm column (Waters), and the peptic peptides eluted at 0.5 °C with a gradient of water + 0.1% (v/v) formic acid (A) and acetonitrile + 0.1% (v/v) formic acid (B) at 60 μ l/min flow rate as follows: 0 to 7 min/95-65% A, 7-8 min/65-15% A, 8-10 min/15% A. The eluting peptides were ionized by ESI (capillary temperature 250 °C, spray voltage 3.0 kV) and mass spectra acquired over a range of 50 to 2,000 m/z on a G2-Si HDMS mass spectrometer with ion mobility separation (Waters) in enhanced high definition MS (HDMS^E) or high definition MS (HDMS) mode for undeuterated and deuterated samples, respectively [110]. [Glu1]-Fibrinopeptide B standard (Waters) was employed for lock mass correction. During each run, the pepsin column was washed three times with 80 μ l of 4% (v/v) acetonitrile and 0.5 M guanidine hydrochloride and blanks were performed between each sample. Three technical replicates (independent H/D exchange reactions) were measured per incubation time.

Peptides were identified with the software ProteinLynx Global SERVER (PLGS, Waters) from the non-deuterated samples acquired with HDMS^E by employing low energy, elevated energy and intensity thresholds of 300, 100 and 1,000 counts, respectively. The identified ions were matched to peptides with a database containing the amino acid sequences of the respective proteins (sVP40 $_{\Delta 43}$ and variants thereof), porcine pepsin and their reversed sequences with the following search parameters: peptide tolerance = automatic; fragment tolerance = automatic; min fragment ion matches per peptide = 1; min fragment ion matches per protein = 7; min peptide matches per protein = 3; maximum hits to return = 20; maximum protein mass = 250,000; primary digest reagent = non-specific; missed cleavages = 0; false discovery rate = 100. Deuterium incorporation into peptides was quantified with DynamX 3.0 software (Waters). Only peptides that were identified in all undeuterated samples and with a minimum intensity of 10,000 counts, a maximum length of 30 amino acids, a minimum number of two products,

Methods

a maximum mass error of 25 ppm and retention time tolerance of 0.5 min were considered for analysis. All spectra were manually inspected and, if necessary, peptides omitted (e.g., in case of low signal-to-noise ratio or presence of overlapping peptides).” (Dr. Wieland Steinchen, Philipps-University Marburg)

An overview of parameters and characteristics of the HDX-MS experiments are given in Table 22, Supplemental Material.

2.4.9 Surface Plasmon Resonance

2.4.9.1 Principle

Surface Plasmon Resonance (SPR) is a highly sensitive method to detect label-free interactions between proteins, peptides, small molecules, nucleic acids, whole cells or viruses in real-time. During the measurement, plasmons – electron charge density waves – are generated when polarized light is directed on a gold film at the interface of two media with a different refractive index. As a result, the intensity of the reflected light is reduced depending on the incidence angle (resonance angle). This reduction is proportional to the mass on the sensor surface. For this, proteins are typically immobilized on a gold sensor surface and called ligands. The analyte in solution is injected into the capillary system and passed over the gold surface and results in SPR signal upon binding, visible as response units (RU) in a sensorgram. Due to the constant flow, binding, saturation of interaction and dissociation can be measured in real time and allows the calculation of kinetic binding constants (k_a and k_d) as well as equilibrium binding constants (KD) [243,293]. For the present work, a Biacore 3000 system (Cytiva) was used. Data acquisition and analysis were performed with the kind help of Dr. Julia Hahn and Melissa Dillenberger from the Becker/Przyborski lab (Institute of Nutritional Science, Justus-Liebig University, Giessen, Germany). A typical SPR experiment includes pH scouting to determine ideal buffer and protein conditions, immobilisation of the protein of interest on the chip surface and analysis of analytes binding to the ligand. In addition, the appropriate washing steps have to be determined to avoid contamination between different analytes.

2.4.9.2 pH Scouting

As an initial experiment, pH scouting is performed to determine the ideal protein concentrations and buffer composition for the follow-up immobilization. For this, the recombinant protein of the desired concentration is passed over the sensor chip and being concentrated without being coupled. A CM5-chip was selected for the analysis of both z/sVP40 Δ 43 WT dimer and octamer, with concentrations varying from 50 to 500 μ g/mL. Buffers included acetate pH 4, acetate pH 4.5, sodium acetate pH 5.5, sodium phosphate pH 6, maleate pH 6, sodium phosphate pH 6.5, maleate pH 6.5, sodium phosphate pH 7 and HEPES pH 7, all in a concentration of 10 mM. The running buffer consisted of 10 mM PBS pH 7.4, 0,005% Surfactant P20, 5% DMSO and protein in the respective buffers was run through the system

Methods

with a contact time of 180 sec with 5 $\mu\text{L/s}$. Between each condition, the chip surface was regenerated using 50 mM NaOH.

2.4.9.3 Immobilisation

After having defined ideal conditions during pH scouting, the protein was immobilized on the sensor chip by applying the amine coupling. For this, three processes followed: 1) activation of the dextran surface of the CM5 sensor chip using N-hydroxysuccinimide and N-ethyl-N'-(dimethylaminopropyl) carbodiimide to create N-Hydroxy-succinimide esters, enabling the formation of covalent bonds with other molecules, 2) coupling of the injected ligand onto the surface to enable interaction and 3) deactivation of the remaining active groups on the sensor surface by injecting a low molecular reactive group [161]. The success rate of the immobilisation is indicated by RU and is usually between 1,000 and 10,000 RU.

2.4.9.4 Binding Analyses

Once the protein of interest was immobilised on the sensor surface, a preliminary analysis called single level binding screen was performed to identify promising binders. Synthetic peptides were dissolved in DMSO and diluted in SPR running buffer to a final concentration of 500 μM . After correcting for solvent effects, the RU values for each peptide were normalised to the negative control and their respective molecular weights. Based on these results, selected peptides were tested using a concentration range optimized to determine their affinities towards VP40 and binding kinetics.

2.4.10 Biotin-Switch assay

The Biotin-Switch assay was performed as previously described with minor modifications [99,154]. Recombinant dimeric sVP40 $_{\Delta 43}$ (WT and cysteine mutants) at a concentration of approx. 0.1 mg/mL were reduced with 5 mM DTT for 30 min on ice. The protein was then desalted using ZebaTM Spin Desalting columns (Thermo Scientific) and incubated with 100 μM S-nitrosocysteine (Cys-NO) for 1 h at RT covered in tin foil. Blocking of free thiols was done by adding 200 mM iodoacetamide (IAA) and incubating the mixture for 45 min at 50 °C covered in tin foil, followed by precipitation and washing in acetone. For labelling with biotin, samples were mixed with 200 mM sodium ascorbate and 2 mM Iodoacetyl-PEG2-Biotin and incubated for 1.25 h at RT covered in tin foil. In between steps, samples were precipitated by adding 100% acetone and incubating for 30 min at -20 °C followed by washing with 70 °C acetone via centrifugation for 5 min at 9,200 rpm and 4 °C. Subsequently, WB analysis (see section 2.4.3.2) was performed with 10% sucrose as a blocking reagent and mouse α -biotin antibody (Santa Cruz Biotechnology), diluted 1:1000 in 1% sucrose. As a secondary antibody goat α -mouse IRDye[®] 800 was used. Afterwards, the same membrane was stained using a rabbit α -sVP40 (Sino Biological), diluted 1:1000 in 1% sucrose and goat α -rabbit IRDye[®] 680 as a secondary antibody.

Methods

2.4.11 Redox assays

Absorbance assays to assess VP40's oxidation state *in vitro* and to identify potential redox systems that interact with VP40 were performed with the kind help of Dr. Christina Brandstädter and Melissa Dillenberger, Becker/Przyborski lab, Justus-Liebig University Gießen. All proteins (except for VP40) and reagents were kindly provided by the Becker/Przyborski lab. Extraction of human Thioredoxin reductase (TrxR) as well as expression and purification of human Thioredoxin (Trx) and human glutathione reductase (GR) were performed as described previously [100,117,247].

2.4.11.1 Ellman's assay

The oxidation state of VP40 was assessed using the 5,5'-dithiobis(2-nitrobenzoic acid) (DTNB) assay (Ellman's assay) [260]. 500 μ M DTNB were mixed with buffer (100 mM potassium phosphate, 2 mM EDTA, pH 7.4) and its absorbance at 412 nm was measured as baseline. Addition of 20 μ M dimeric s/zVP40 $_{\Delta 43}$ WT/CCS resulted in an increased absorbance and the number of free thiol groups per molecule VP40 were calculated using the formula: free thiols = $(\Delta A * \text{mM} * \text{cm} * V_0) / (14.15 * v_i * d [\text{cm}])$ (ΔA = change of absorbance at 412 nm; V_0 = total assay volume; 14.15 = extinction coefficient (ϵ) [$\text{mM}^{-1} * \text{cm}^{-1}$]; v_i = volume of VP40 solution; d = layer thickness of the cuvette (in this case 0.556)).

2.4.11.2 Thioredoxin reduction assay

For the thioredoxin reduction assay [338], varying concentrations of z/sVP40 $_{\Delta 43}$ WT/CCS were mixed with assay buffer (100 mM potassium phosphate, 2 mM EDTA, pH 7.4), 100 μ M nicotinamide adenine dinucleotide phosphate (NADPH; extinction coefficient 6.22 $\text{mM}^{-1} * \text{cm}^{-1}$) and 0,25 U/ml TrxR. After monitoring the baseline, 10 μ M Trx were added to the reaction mix and the change in absorbance was monitored at 340 nm. Absorbances were all measured at 25 °C using a Tecan infinite M200 multiplate reader with a total assay volume of 200 μ L. As a statistical analysis, unpaired t-test was performed.

2.4.11.3 Glutathione reduction assay

For the GR assay [50], human GR was mixed with buffer (20.5 mM KH_2PO_4 , 26.5 mM K_2HPO_4 , 1 mM EDTA, 200 mM potassium chloride, pH 6.9) and 100 μ M NADPH. sVP40 $_{\Delta 43}$ WT was added after 5 min in varying concentrations and the absorbance was monitored at 340 nm. To assess a potential reduction of VP40 by reduced glutathione (GSH), 10 μ M oxidized glutathione (GSSG) were added to the reaction mix and the baseline monitored at 340 nm. Absorbances were all measured at 25 °C using a Tecan infinite M200 multiplate reader with a total assay volume of 200 μ L.

3 Results

The following chapters are divided into three parts, each covering different approaches of identifying lead structures of VP40. Part I covers the identification of crucial interface residues as well as amino acids involved in allosteric regulation of VP40's functions. Part II describes the investigation of peptides that mimic the amino acid sequence of the different interfaces and their impact on the various VP40 functions. Fragment-based lead discovery is the subject of Part III, where different approaches are described in the search and development of a VP40-specific inhibitor. While the different parts are connected with regards to methods but also to content, the scientific questions are fundamentally different for each part. Therefore, each chapter has a separate introduction, results and discussion section.

3.1 Part I: Identification and characterization of residues involved in protein-protein interactions

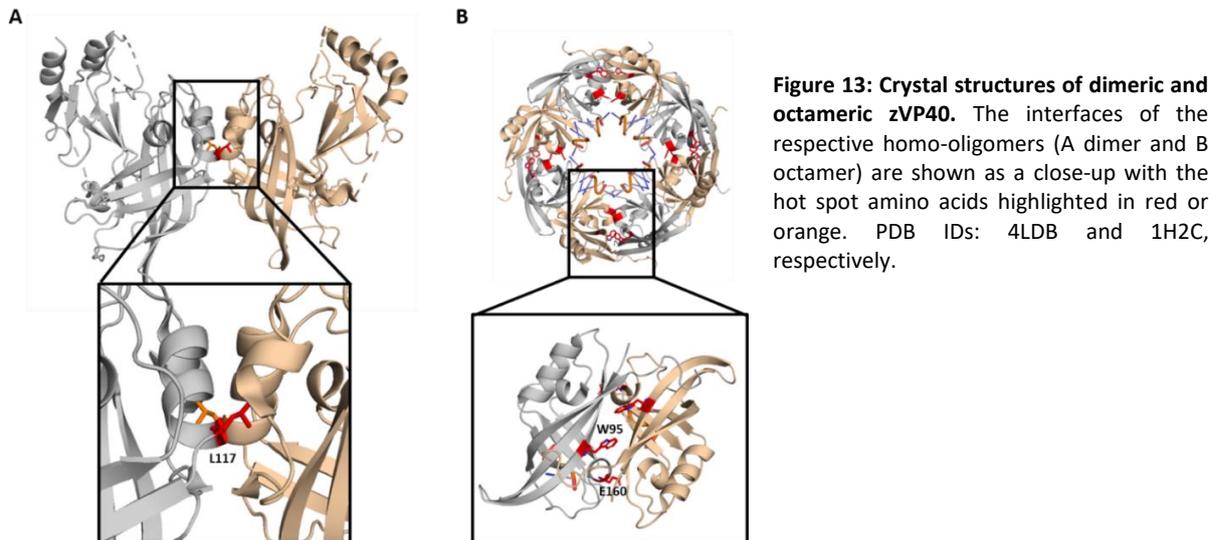
Part I covers both the characterization of previously described interface residues (section 3.1.1) as well as novel amino acids of interest within VP40 (section 3.1.1 and 3.1.2). As mentioned in section 1.9, this chapter aims to characterize sVP40 and serves as the "target validation" step in the drug design process.

3.1.1 Characterization of selected hot spot amino acids of sVP40 homo-oligomerization

3.1.1.1 Introduction

VP40 adopts several crucial homo-oligomeric forms that are formed upon conformational rearrangement of monomers and reconnect through different interfaces. Hot spot amino acids of the various interfaces were the subject of several studies, but described solely for zVP40 [32,136]. L117 at the dimeric interface was reported to be crucial for the stability and integrity of the dimer (Figure 13 A), and inhibited zVP40 transport to the plasma membrane and VLP budding [32]. Mutation of zVP40 W95 was reported to be important for octamer formation and stability (Figure 13 B) [115]. To investigate whether L117 and W95 serve the same purpose in sVP40 as in zVP40, W95 and L117 of sVP40 were mutated in both mammalian and bacterial expression vectors, and investigated using different methods (see also Figure 58).

Results



3.1.1.2 Results

Functional analyses of hot spot mutants of sVP40 in functional cell culture assays

Since the aim of this work was to inhibit the homo-oligomerization of VP40, the function of the above-mentioned residues were characterized for sVP40. First, an EBOV Zaire-based MG assay was used to analyze the function of sVP40, as a MG reporter assay system was not available for SUDV. The principle and protocol are outlined in section 2.3.1. As expected, sVP40 WT inhibited reporter gene expression in a dose-dependent manner, although double amounts of sVP40 (400 ng) compared to zVP40 (200 ng) had to be transfected in order to reduce the levels of reporter gene activity to similar levels (Figure 14 A). Upon transfection of 400 ng of plasmid DNA coding for sVP40, the ectopic expression of sVP40 W95A lowered reporter gene activity even stronger than observed for sVP40 WT (Figure 14 B). In contrast, sVP40 L117A was not able to inhibit viral genome replication and transcription.

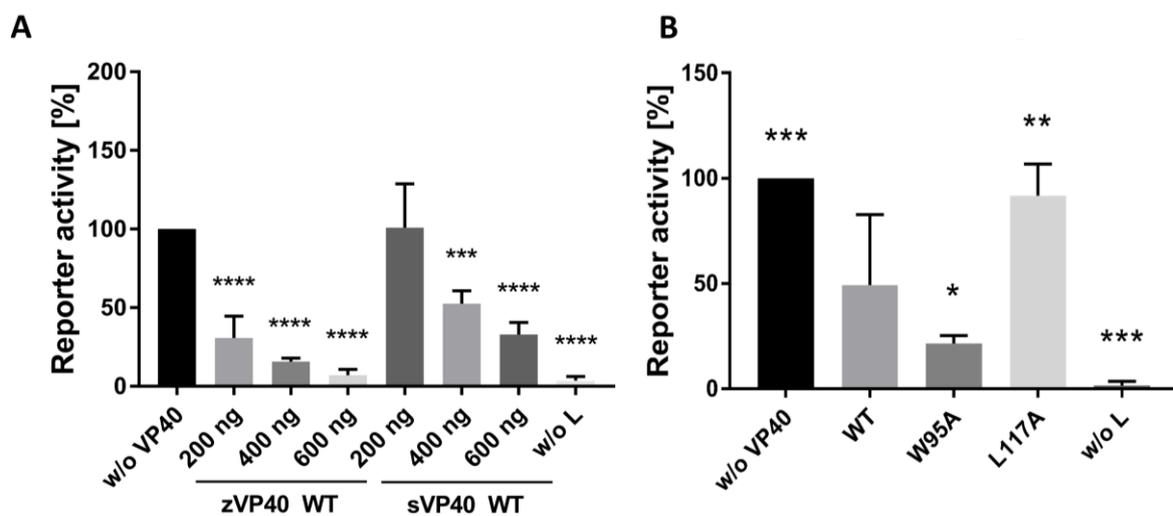


Figure 14: sVP40 WT and mutants influence reporter gene activity of a MG assay. A) Comparison of zVP40 and sVP40 WT in a Zaire-based MG assay and B) VP40 W95A and L117A mutations of sVP40 show statistically significant changes in reporter gene activity compared to sVP40 WT. For both experiments, HEK293 cells were transfected with pCAGGS s/zVP40 WT or mutants along with the MG assay components. Cells were lysed 48 h pt, reporter gene activity was measured and normalized

Results

to the sample without VP40 (set to 100%). For the negative control, no polymerase L was added to the transfection mix. Bars indicate the mean \pm SD of at least four independent experiments, asterisks indicate statistical significance (One-way ANOVA) compared to the positive control without VP40 for graph A or to sVP40 WT for graph B as follows: * $P < 0.05$, ** $P < 0.005$, *** $P < 0.0005$ and **** $P < 0.0001$.

The same experiment performed using 200 ng plasmid DNA coding for VP40 is shown in the supplemental material (Figure 59).

To test whether these mutants were still able to be released from the cell as VLPs, sVP40 WT and mutants were co-transfected with GP (Zaire) to improve the release of VP40-induced VLPs [245]. WB analysis showed a strongly increased budding efficiency for sVP40 W95A whereas sVP40 L117A was hardly able to induce the release of VLPs into the supernatant (Figure 15).

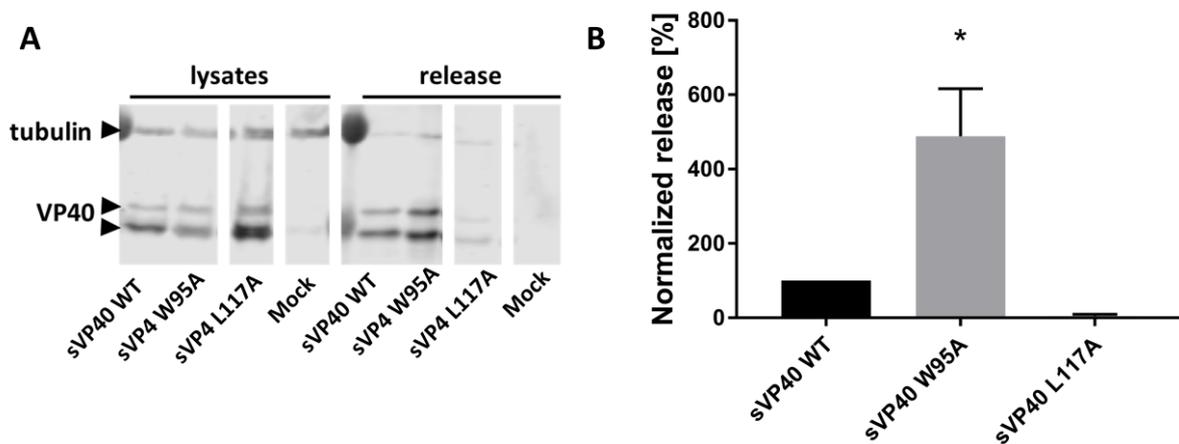


Figure 15: Effect of sVP40 W95A and L117A mutations on VP40-release. A) Cells were transfected with pCAGGS zGP and sVP40 WT or mutants. Supernatants were collected and cells lysed 72 h pt WB analyses of cellular and released VP40 were performed using rabbit α -VP40 and mouse α -tubulin primary and donkey α -mouse and goat α -rabbit IRDye[®] 680 as secondary antibodies. B) Quantification of VLPs. Released VP40 was normalized to the amount of VP40 in lysates and sVP40 WT set to 100%. Bars indicate the mean \pm SD of three independent experiments, asterisks indicate statistical significance (One-way ANOVA) compared to sVP40 WT as follows: * $P < 0.05$.

Hot spot mutants are important for the formation of dimers and monomers in solution

To investigate an effect on homo-oligomerization upon mutation or L117 or W95, SEC experiments were performed using recombinant protein expressed in *E. coli*. For this, the first 43 amino acids of sVP40 were deleted for improved (future) crystallization. This truncated construct (sVP40 _{Δ 43}) was used for the experiments of subsequent chapters as well. The mutations replacing W95 to A and L117 to A were introduced into the sVP40 gene in the bacterial expression plasmid pET46 (kindly provided by the Sapphire lab, La Jolla). The mutant sVP40 _{Δ 43} molecules were expressed in Rosetta2 cells and purified via IMAC, followed by SEC analysis (2.4.2.4). sVP40 _{Δ 43} WT formed mainly dimers with only little amounts of octamers (Figure 16 A). sVP40 _{Δ 43} W95A still formed primarily dimers (Figure 16 B). which was expected because residue W95 is not involved in the dimer interface. sVP40 _{Δ 43} L117A formed mainly monomers (Figure 16 C) with only minor amounts of higher-order oligomers, which is in line with the

Results

literature for zVP40 L117R, where higher amounts of both octamer and monomer, but hardly any dimers were reported [32].

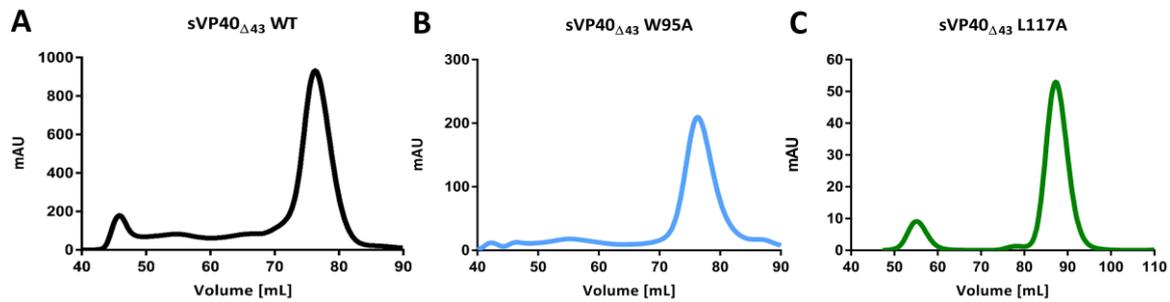


Figure 16: Assessment of homo-oligomeric state and stability of sVP40_{Δ43} WT, W95A and L117A via SEC. Elution profiles of sVP40_{Δ43} WT (A, black), sVP40_{Δ43} W95A (B, blue) and sVP40_{Δ43} L117A (C, green) using a HiLoad 16/60 column.

Hot spot mutants exhibit the same crystal packing as the WT

SEC fractions corresponding to sVP40_{Δ43} W95A and L117A were collected and used for the setup of crystallization screens (performed by the MarXtal facility) to investigate possible structural changes upon mutation. Several datasets of crystals grown in different conditions were collected for both recombinant mutants of sVP40_{Δ43} and the crystals with the highest diffraction were selected for further processing. sVP40_{Δ43} W95A (AW44) was crystallized using the MBC I screen in 0.5 M potassium chloride, 50 mM MOPS, 12% PEG4000 (w/v), 20% glycerol (w/v) (well F10), with a concentration 3.6 mg protein/ml. For sVP40_{Δ43} L117A (AW47), crystals were formed in 90 mM HEPES, 6.8% ethylene glycol (v/v), 15% glycerol (v/v), 17% PEG10,000 (w/v) (Cryos screen, well H11) using 7.6 mg protein/ml. The structure of the published dimeric sVP40_{Δ43} WT (PDB-code 4LD8 [32]) was used as a template for molecular replacement. The 4LD8 structure could be placed confidently into the X-ray derived electron density map, which was further refined using the PHENIX Refinement suite as well as COOT. Both crystal structures belonged to space group 5 and had 1 protomer in the asymmetric unit with similar unit cell dimensions. The experimental data for each selected crystal is summarized in Table 4.

Table 4: Data collection and refinement statistics of sVP40_{Δ43} W95A and L117A. Values for the outer shell are given in parentheses.

Data collection	sVP40 _{Δ43} W95A	sVP40 _{Δ43} L117A
PDB ID	8AYT	8AYU
Space group	5	5
Unit cell parameters [Å] (a, b, c, α, β, γ)	64.75, 91.28, 48.93, 90.00, 96.71, 90.00	62.46, 90.70, 48.38, 90.00, 95.05, 90.00
Resolution range	48.60 – 1.90	48.19 – 2.00
Total number of reflections	84,504 (505)	67,419 (5,168)

Results

Number of unique reflections	22,067 (1,406)	17,832 (1,328)
I/ σ	10.7 (2.5)	11.5 (4.9)
Completeness [%]	99.1 (98.2)	98.3 (99.3)
Multiplicity	3.8 (3.6)	3.8 (3.9)
CC _{1/2} [%]	99.9 (91.2)	99.6 (95.6)
Refinement		
Resolution range	45.64 – 1.90	48.19 – 2.00
Rwork/Rfree [%]	16.70/19.87	17.00/19.50
Number of unique reflections	22,050	17,829
Number of protein residues	248	239
Average B factor [\AA^2]	49.41	39.87
R.m.s deviations:		
- Bond length [\AA]	0.007	0.007
- Bond angles [$^\circ$]	0.84	0.86
Ramachandran plot:		
- Favoured [%]	97.48	97.82
- Allowed [%]	2.52	2.18
- Outliers [%]	0.00	0.00
Rotamer outliers [%]	0.00	0.00
Clash score	2.13	1.65
Number of TLS groups	3	3

Both mutants formed dimers in the crystalline state with very similar packing, as the alignment with 4LD8 suggests (Figure 17). While a dimeric packing was expected for sVP40 Δ ₄₃ W95A, sVP40 Δ ₄₃ L117A still crystallized as a dimer, suggesting that the dimeric crystal packing is the most stable form in the crystalline state.

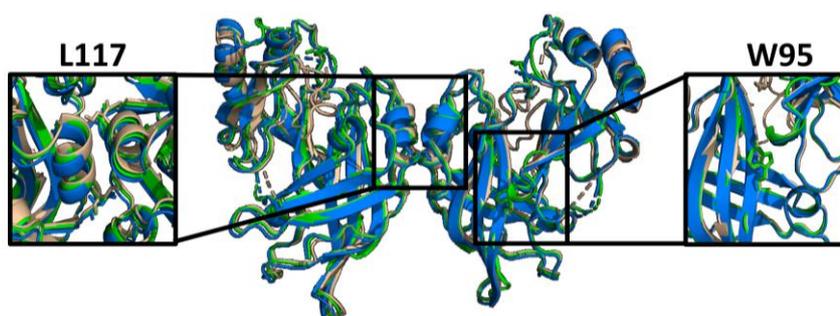


Figure 17: Alignment of sVP40 Δ ₄₃ W95A and L117A crystal structures with sVP40 Δ ₄₃ WT. Superposition of sVP40 Δ ₄₃ WT (wheat) with W95A (AW44) and A) L117A (AW47) with a close-up of respective residues of interest shown in boxes (L117 left-hand side and W95 on right-hand side).

Results

HDX-MS reveals major structural implications in solution for hot spot mutants

Because both mutants crystallized with the same structure as sVP40 Δ 43 WT, it was considered whether the crystal structures represented potential crystallographic artefacts, which seemed especially likely for sVP40 Δ 43 L117A. HDX-MS was performed as described in section 2.4.8.3 (experimental overview is given in the Supplemental Material, Table 22) using sVP40 Δ 43 WT, W95A, and L117A. Significant differences were observed for sVP40 Δ 43 L117A (Figure 18 A): D₂O uptake was especially increased for peptides 108 to 114, 176 to 180 and 248 to 260, which were mapped onto the crystal structure (Figure 18 B). As expected, residues spanning the dimeric interface showed highly increased exchange rates, suggesting that this protein is indeed a monomer showing an exposed dimeric interface. Interestingly, other areas were also affected, indicating an influence of the L117A mutation on secondary structure.

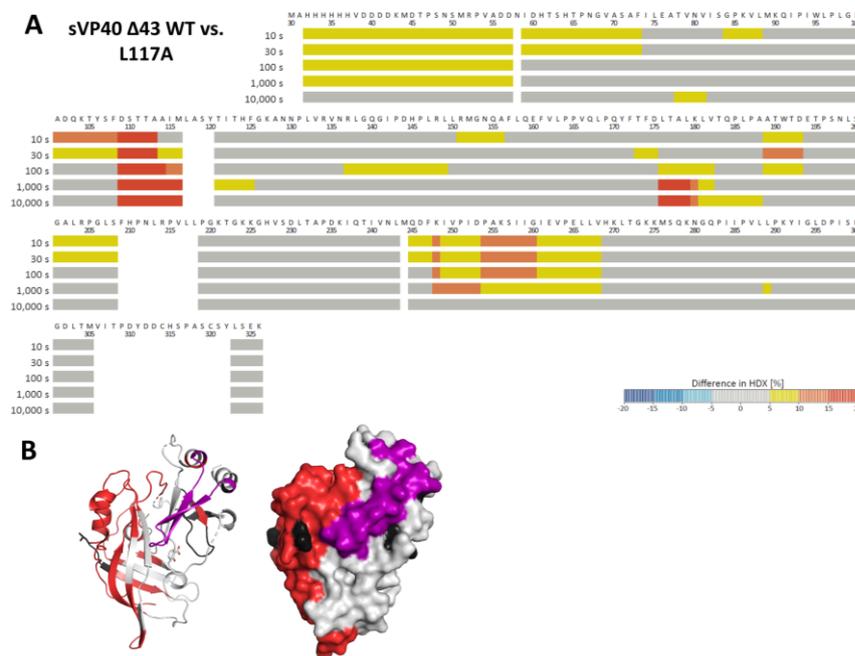


Figure 18: HDX-MS for sVP40 Δ 43 L117A compared to sVP40 Δ 43 WT. 25 μ M protein were incubated with D₂O-containing buffer. After 10, 30, 100, 1,000 or 10,000 s, each reaction was quenched, samples digested with pepsin and the peptides were ionized by ESI and analyzed via mass spectrometry. Differential deuterium uptake for plotted on the aa sequence (A) or mapped onto the sVP40 Δ 43 L117A crystal structure (B, cartoon representation right and surface left). Peptides with increased exchange rates are shown in red, peptides depicting bimodality in purple and residues shown in black were not covered.

In summary, these results support the notion that while sVP40 Δ 43 L117A crystallized as a dimer it was a monomer in solution, as based on the HDX-MS data. This result is in accordance with the SEC analysis and further stresses the high stability of the crystallographic packing of the dimer. This indicates that sVP40 Δ 43 L117A was clearly affected in its ability to homo-oligomerize and crystallographic dimers are only formed because of their stability. In contrast to what the crystallization data suggested, the protein was highly flexible in solution and even showed bimodality, indicating partial unfolding of the protein (Figure 61, supplemental material).

The comparison of sVP40 Δ 43 WT and sVP40 Δ 43 W95A also showed significant exchange rates for several peptides (Figure 19 A) with residues 248 to 260 exhibiting the strongest increase in deuterium uptake. Other affected areas included peptides 203 to 208 and 176 to 188 and also residues of the dimeric

Results

interface. Additionally, nearly the same peptides were affected as for sVP40 Δ 43 L117A and interestingly, also residues of the dimeric interface, although to a lesser extent than observed for sVP40 Δ 43 L117A.

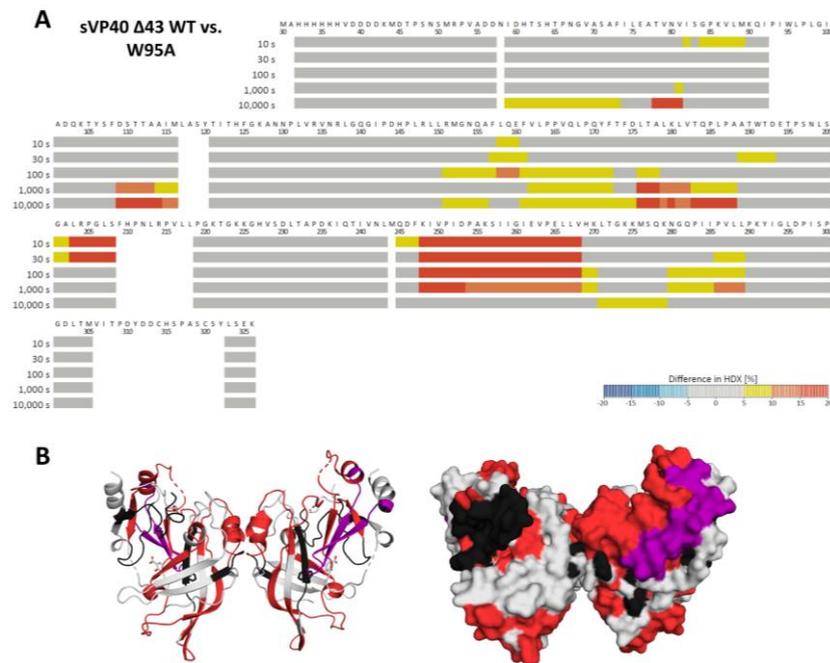


Figure 19: HDX-MS for sVP40 Δ 43 W95A compared to sVP40 Δ 43 WT. 25 μ M protein were incubated with D₂O-containing buffer. After 10, 30, 100, 1,000 or 10,000 s, each reaction was quenched, samples digested with pepsin and the peptides were ionized by ESI and analyzed via mass spectrometry. Differential deuterium uptake for plotted on the aa sequence (A) or mapped onto the sVP40 Δ 43 W95A crystal structure (B, cartoon representation right and surface left). Peptides with increased exchange rates are shown in red, peptides depicting bimodality in purple and residues shown in black were not covered.

Together, the crystal and SEC data suggest sVP40 Δ 43 W95A to be a dimer in both the crystalline and liquid state but especially the bimodality levels (Figure 61, supplemental material) of some peptides indicate a strong structural fluidity. The mapping of the relative H/D exchange and bimodality behavior of this mutant onto the crystal structure is shown in Figure 19 B. Bimodality is observed when a polypeptide chain is present in two or more different conformations that exhibit significant differences regarding their H/D uptake behavior, depicted as not one but two bell-shaped curves. Bimodal behavior potentially suggests more than one oligomeric species or (partial) unfolding of the protein. To investigate the oligomeric stability of both monomeric sVP40 Δ 43 L117A and dimeric sVP40 Δ 43 W95A, additional SEC runs of previously stored proteins were performed. When stored for several weeks at -80 °C, sVP40 Δ 43 W95A still eluted as dimers, indicating a high stability (Figure 62 A, supplemental material). Purified monomeric sVP40 Δ 43 L117A was stored at 4 °C and reapplied to the same SEC column after three, four and five weeks, where low amounts of higher-order oligomers, potentially octamers, could be observed in some batches (Figure 62 B, supplemental material). However, the presence of some minor octameric fractions observed for sVP40 Δ 43 L117A are unlikely to account for the observed phenotypes as bimodality was only observed for a relatively short peptide. A significant amount of octamers would be expected to result in more peptides exhibiting bimodality. In addition, the same peptide exhibiting bimodality was described for sVP40 Δ 43 W95A, which was shown to be stable in solution. Nevertheless, it could be assumed that both amino acids, L117 and W95, are crucial for the structural stability of their respective homo-oligomeric form. As the same peptides showed

Results

bimodality for both mutants, it is probable that these peptide are structurally highly dependent on other areas of the protein.

sVP40 Δ 43 W95A and L117A are less heat-stable than sVP40 Δ 43 WT

To further confirm the destabilizing effect of the L117A and W95A mutations seen in the HDX-MS experiments, a TSA was performed (2.4.7). As shown in Figure 20, melting temperatures significantly decreased from 49.3 ± 2.1 °C (sVP40 Δ 43 WT) to 42.2 ± 1.2 °C (sVP40 Δ 43 W95A) and to 46.9 ± 0.4 °C (sVP40 Δ 43 L117A), suggesting that the introduction of the point mutations resulted in lower protein stability.

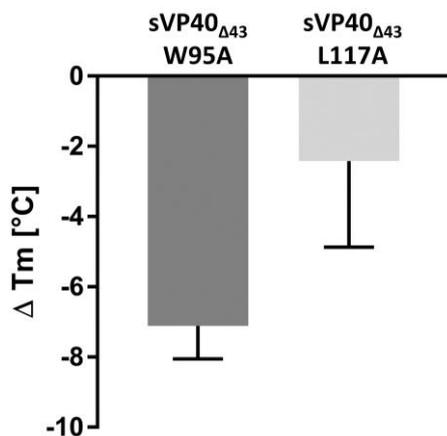


Figure 20: Thermal stability of sVP40 hot spot mutants. Means of denaturing temperatures from three independent experiments were compared to the sVP40 Δ 43 WT. Bars indicate the mean \pm SD of three independent experiments.

Taken together, sVP40 W95A showed increased budding activity and stronger inhibition of viral genome replication and transcription than sVP40 WT. The bacterially expressed sVP40 Δ 43 W95A formed dimers with increased flexibility in several regions and exhibited signs of bimodality, suggesting partial unfolding of the protein. This was also confirmed using TSA. For sVP40 L117A, the ability to dimerize was abolished when the protein was expressed in *E. coli*, which is in accordance to the loss-of-function regarding budding of VLPs and inhibition of viral RNA synthesis.

3.1.1.3 Discussion

Functional analyses of sVP40 L117A revealed a complete loss of its ability to inhibit vRNA synthesis and to induce the formation and release of VLPs, which is in line with the literature (Figure 14 and Figure 15) [32]. Using recombinant sVP40 Δ 43 L117A expressed in *E. coli*, a shift in the oligomerization pattern could be observed: while sVP40 Δ 43 WT eluted nearly exclusively as dimers, the sVP40 Δ 43 L117A mutant formed mainly monomers as seen in the SEC elution profile (Figure 16). However, it showed a nearly identical 3D structure compared to sVP40 Δ 43 WT (Figure 17). However, concluding from the functional assays, drastic structural and conformational changes were expected for sVP40 Δ 43 L117A compared to sVP40 Δ 43 WT. To understand those apparent differences, structural investigations of this mutant in solution using HDX-MS were performed as a follow-up experiment. As shown in Figure 18 A, For sVP40 Δ 43 L117A, a strong increase of H/D uptake could be observed for residues that support the structural integrity of the helical dimeric interface when compared to sVP40 Δ 43 WT. This shows that

Results

these residues are solvent-exposed and not shielded by the other dimeric half, as was suggested by the crystal structure. Additionally, further residues were greatly affected by the L117A mutation, pointing towards a central role of L117 for the overall stability of VP40. Combining the SEC results with crystal and HDX-MS data, it could be concluded that sVP40_{Δ43} L117A is a monomer in solution and the crystal structure represents an artefact, assuming the crystallographic packing of sVP40_{Δ43} WT is the most energy-efficient form in the crystalline state. Similar observations were made for the RdRp domain of nsp4 of Sindbis virus, which crystallized as a dimer but was shown to be a monomer in solution [298], for the bacterial RNA modifying enzyme Tgt [156], and also for Calmodulin. In the latter case the crystal structure showed highly rigid areas, whereas HDX-MS revealed the same areas to be flexible [248].

For sVP40 W95A, the finding of a stronger decrease of reporter gene activity in the MG assay (Figure 14 B) along with increased budding of VLPs induced by sVP40 W95A compared to sVP40 WT (Figure 15) pointed towards a potential gain-of-function mutation. On the other hand, its biochemical characterization showed that while the sVP40_{Δ43} W95A mutant was still able to form dimers in solution (Figure 16) and crystallized very similar to sVP40_{Δ43} WT (Figure 17), the recombinant protein was less heat-stable (Figure 20) suggesting decreased structural integrity. It also showed great differences in deuterium uptake as well as bimodal behavior (Figure 19 and Figure 61, supplemental material), indicating allosteric effects on the peptide in question, as the protein was shown to be stable even after weeks of storage (Figure 62 A, supplemental material). W95 was described to be the hot spot residue of one of the octamerization interfaces of zVP40 [32,115] and its mutation was associated with decreased formation of oligomers but not with a decrease of VP40's inhibitory function on viral RNA synthesis [136]. While the lack of influence of the W95A mutation on reporter gene activity is surprising, mutation of both W95 and E160 to alanine (WEA) completely abolished the inhibitory effect of zVP40 in a replication-deficient MG assay that tested only the viral transcription, suggesting that mutation of these residues did mainly affect transcription [136]. Additionally, WEA was reported to induce less iVLPs formation due to decreased transport of zVP40 to the plasma membrane [136]. W83 and N148 of MARV VP40 (mVP40) are the equivalent residues to zVP40 W95 and E160 and the W83A/N148A (WNA) double mutant of mVP40 has also been extensively characterized: increased reporter gene activity was described in a MG assay, suggesting that these residues are important for the inhibitory function of VP40 regarding viral genome replication and transcription [171]. Furthermore, mVP40 WNA showed no cluster formation at the plasma membrane and was unable to bud VLPs upon single expression of mVP40 [171,249], but still formed dimers in solution when expressed in *E. coli* [249]. zVP40 W95 was described to play a major role for the hexamers as the crystal structures showed a common interface of octamers and hexamers. Since hexamerization is correlated with budding activity and indirectly with membrane interaction, the results concerning the release of

Results

iVLPs and cluster formation at the plasma membrane made sense at the time. However, the hexameric crystal form was recently reported to be an artefact of crystallization, and the structure of EBOV and MARV particles revealed a series of dimers connected via their CTDs [318]. A comparison of the different dimeric structures of zVP40 and sVP40 WT revealed that the orientation of the W95 side chain is highly conserved and located on the outer rim of the pocket between the NTD and CTD with the aromatic ring pointing towards the solvent (Figure 21). Its aromatic nitrogen atom forms H-bonds to the backbone oxygen atom of K256, thereby stabilizing the loop between two β -sheets. Only W95 of PDB-ID 7JZT is slightly tilted and forms an additional H-bond to L289. Disruption of this interaction by mutating the tryptophan to alanine could possibly result in a destabilization of the NTD-CTD integrity which in turn could have implications for the formation of higher-order oligomers. This however does not explain the differences observed between zVP40 and sVP40.

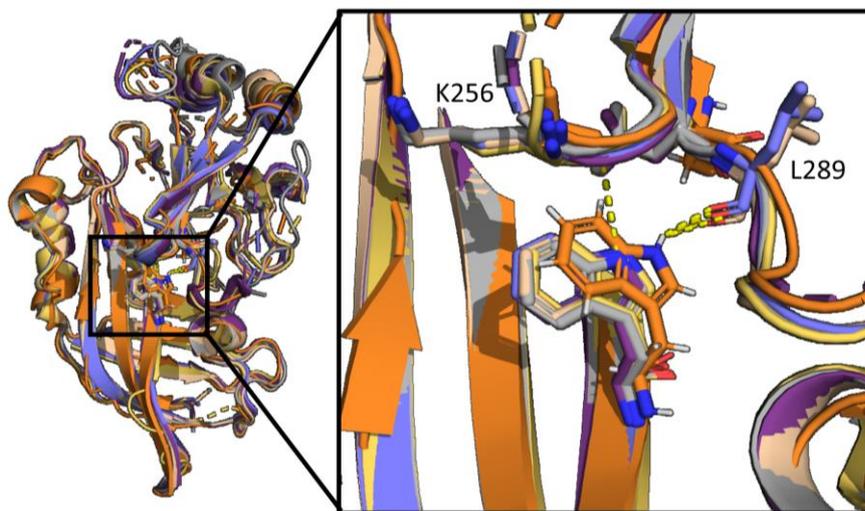


Figure 21: Position and orientation of residue W95 in different VP40 structures. Alignment of dimeric sVP40 (3TCQ and 4LD8; blue and dark purple) with monomeric (1ES6; wheat) and dimeric zVP40 (4LDB, 7JZT and 7JZJ; grey, orange and yellow) with W95, L289 and K256 shown as sticks.

The cell culture results of the present work are internally consistent (higher activity in both budding and MG assays) and the combined interpretation with the HDX-MS data suggests that in the case for W95A a more fluid and flexible sVP40 which is still able to dimerize might be advantageous for VP40's functions. A possible explanation is that the protein is stabilized by other viral or host proteins *in cellulo*. Also, it is feasible that species-specific differences are accountable or the double mutants zVP40 WEA and mVP40 WNA exhibit other behaviors in functional assays than the single mutant and not simply stronger effects, as would have been expected. Since most publications investigate the double mutants, a direct comparison is difficult. However, it is feasible that the investigation of budding by co-expression of VP40 and GP only does not allow proper comparison between the sVP40 W95A and the zVP40 W95A/WEA, where an iVLP setting was used, containing all viral proteins [136]. However, for mVP40 W83/WNA budding experiments only mVP40 was transfected with similar results to zVP40 WEA [171,249]. HDX-MS experiments have been performed on mVP40 W83A and WNA but

Results

the H/D exchange was not compared to mVP40, but to the absence or presence of liposomes [8,326], allowing no direct comparison with the present work.

In summary, this work provided a detailed characterization of two selected hot spot mutants of sVP40, expressed in both mammalian cells and *E. coli*. Regarding sVP40 L117A, the results were well in line with the literature for zVP40 (loss of budding and deficient inhibition of RNA synthesis in a MG assay). Concerning sVP40 W95A, the unexpected findings of increased activity in budding and MG assays raised several questions regarding W95A's mode of action and require further investigation. In addition, it became clear that there is a need to validate structural conformations in solution and not just in the crystalline state, which is important for future studies.

Results

3.1.2 Previously unresolved C-terminus of Sudan Ebolavirus VP40 exhibits functionally important post-translational redox modifications

3.1.2.1 Introduction

The end of the C-terminus of VP40 contains two cysteine residues whose positions are conserved among EBOV, RESTV, TAFV, BOMV and BDBV. sVP40 shares only one common cysteine residue (VP40 C314), while the second cysteine (sVP40 C320) is located further toward the C-terminal end and is separated by five amino acids from sVP40 C314 (Figure 22 A). Upon the mutation of the two cysteines in zVP40, the protein's ability to interact with lipid membranes was impacted, as was its ability to induce the formation of virus-like particles [159]. It was not yet clear whether the cysteines play a similar role in sVP40 and whether the second function of VP40, the regulation of transcription/replication, is also affected by the mutation of the C-terminal cysteines. Therefore, it was decided that the cysteines in sVP40 merit further investigation.

In this chapter, a crystal structure of sVP40 with a resolution of 1.53 Å is described that reveals previously unresolved residues, such as the intrasubunit hinge between the NTD and CTD of sVP40, as well as the near-complete C-terminal end, which give insight into previously uncharacterized structural regions, including the positions of the two sVP40 cysteines.

3.1.2.2 Results

High-resolution crystal structure elucidates missing residues

To characterize the cysteines in sVP40, we optimized the previously reported crystallization conditions for dimeric sVP40_{Δ43} [32] (section 2.4.4.2) and determined the position of more residues than in previously published structures of dimeric sVP40_{Δ43}, thus rendering the very end of the C-terminal domain visible in the electron density map. This crystal structure (AW122) originated from the crystal soaking experiments (section 3.3.2.2) and was soaked with 3-aminobenzamide. However, this fragment could not be placed in the structure (1.53 Å; PDB-code 8BX3). Crystallographic parameters are given in Table 5. The published crystal structure of dimeric sVP40_{Δ43} (PDB-code 4LD8) [32] was used for molecular replacement of our sVP40 structure. Several residues that were missing in 4LD8 could be confidently assigned to the sVP40 structure, as there was very clear electron density visible (Figure 63, supplemental material). Newly resolved residues involve the intrasubunit hinge, spanning residues 195-TPSNLS-200 between the NTD and CTD (Figure 22 B and C), which were too flexible in other sVP40 structures and thus not visible in the electron density map. The hinge formed a loop, which is stabilized through a hydrogen bond (H-bond) bridging D193 and T195. T195 also reaches R204 via H-bonds (Figure 22 B and C). S197 develops an H-bond to D310, as does N198. The latter is also connected to D310 and D312. Backbone atoms of L199 reach S197 and A202, further stabilizing the loop.

Results

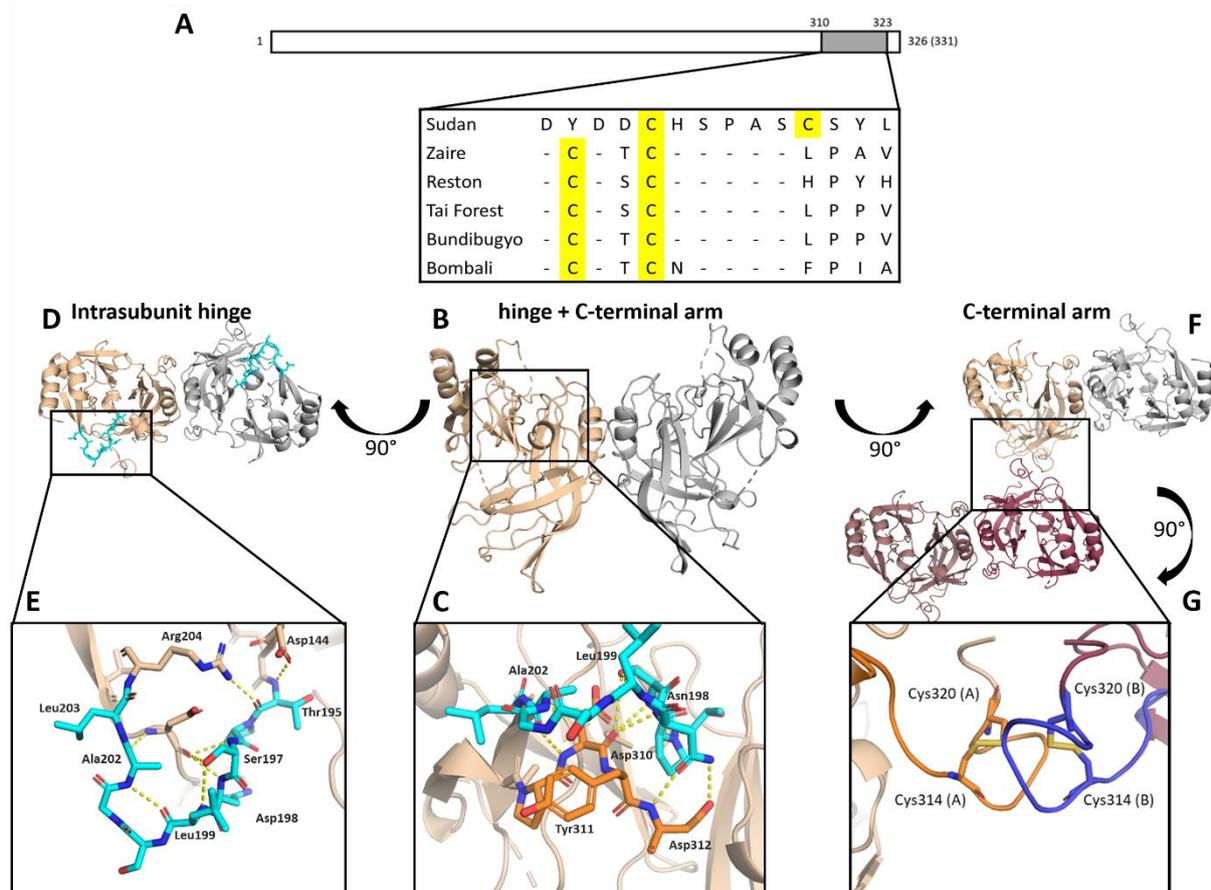


Figure 22: Sequence and structure of the last residues of the VP40-CTD. A) Schematic representation of the amino acid sequence of VP40 of different Ebola virus species showing residues 310–323 with cysteines highlighted in yellow. B and C) Top view of the intrasubunit hinge (cyan); D and E) interaction between the hinge region (cyan) and newly resolved residues in orange. F and G) Top view and close-up of the newly resolved C-terminal end with its cysteines, including the disulfide bridge in yellow of dimer 1 (wheat/orange) and dimer 2 (dark red/blue; right panel).

The hinge region's interaction partners are mainly part of the very end of the C-terminus, the most interesting residues of this crystal structure for the current study (Figure 22 D and E). The high resolution allows the assignment of amino acids 309-PDYDDCHSPASCSYL-323 to our sVP40 model, which pick up immediately after the last β -strand: the whole stretch forms a loop which was stabilized by a disulfide bridge between C314 and C320 (Figure 22 F and G). Interestingly, this turn is within close proximity to the C-terminal arm of the adjoining protomer in the crystal packing without forming any contacts (Figure 22 F).

Table 5: Data collection and refinement statistics of dimeric sVP40 Δ 43 WT. Values for the outer shell are given in parenthesis. *Final upload into the PDB is still pending, minor changes to refinement statistics are possible.

Data collection	sVP40 Δ 43 WT (AW122)*
PDB ID	8BX3
Space group	5
Unit cell parameters [Å] (a, b, c, α , β , γ)	61.95; 90.56; 48.13; 90.00; 93.92; 90.00
Resolution range	51.05 – 1.53

Results

Total number of reflections	124.804 (6.376)
Number of unique reflections	38.404 (1.870)
I/ σ	19.6 (4.5)
Completeness [%]	96.5 (93.5)
Multiplicity	3.2 (3.4)
CC _{1/2} [%]	99.9 (95.1)
Refinement	
Resolution range	51.05 – 1.53
Rwork/Rfree [%]	17.30/18.90
Number of unique reflections	38.404
Number of protein residues	261
Average B factor [\AA^2]	22.2
R.m.s deviations	
- Bond length [\AA]	0.009
- Bond angles [$^\circ$]	1.07
Ramachandran plot	
- Favoured [%]	98.46
- Allowed [%]	1.54
- Outliers [%]	0.00
Rotamer outliers [%]	1
Clash score	0.00

Cysteine mutants do not influence homo-oligomerization of sVP40

To determine whether the C-terminal cysteine residues were crucial for the structural integrity of the sVP40 $_{\Delta 43}$ dimer, we constructed bacterial expression plasmids encoding mutants of sVP40 in which the C-terminal cysteines were mutated to serines, resulting in the constructs pET46 sVP40 $_{\Delta 43}$ C314S, sVP40 $_{\Delta 43}$ C320S, and the double mutant sVP40 $_{\Delta 43}$ C314S-C320S (CCS). The mutant sVP40 $_{\Delta 43}$ proteins were expressed in *E. coli*, purified via IMAC and subjected to SEC (sections 2.4.1 and 2.4.2). For zVP40, Johnson *et al.* reported that bacterially expressed cysteine mutants preferentially formed octamers compared to zVP40 WT [159]. However, in SEC, sVP40 $_{\Delta 43}$ WT and all cysteine mutants eluted exclusively as dimers, suggesting that the mutation of cysteine residues did not impact the preferential formation of dimers (Figure 23 A).

Results

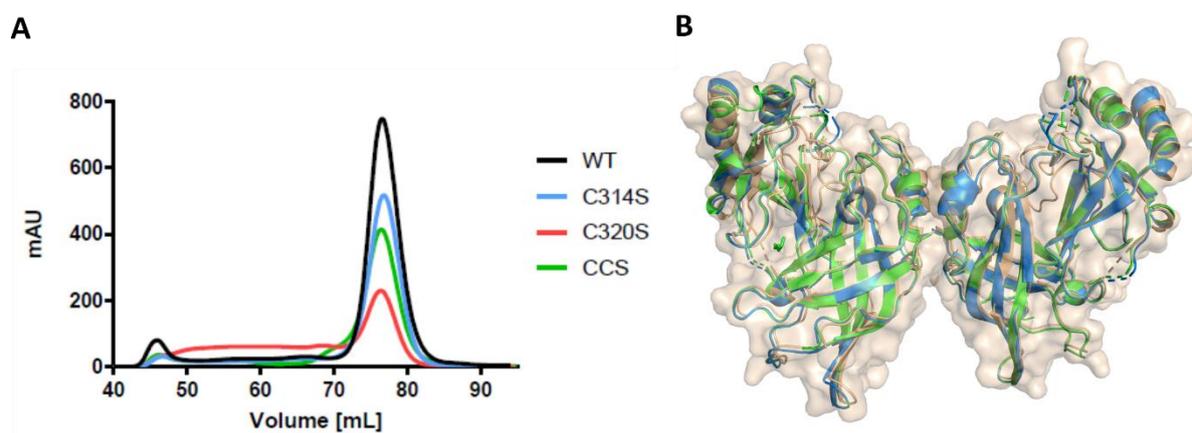


Figure 23: The mutation of sVP40 Δ 43 cysteines does not alter homo-oligomerization. A) Size-exclusion chromatography of sVP40 Δ 43 WT and cysteine mutants expressed in *E. coli* with all four recombinant proteins eluting as dimers. B) Alignment of crystal structures of dimeric sVP40 Δ 43 WT (wheat, shown in cartoon and surface representation) with C314S (blue) and CCS (green) showing the same crystal packing.

Nevertheless, an impact on the formation of the CTD or the crystal packing seemed possible, which is why crystallization screens of all three mutants of sVP40 Δ 43 were set up (section 2.4.4.2, kindly supported by the MarXtal facility, Marburg). All three proteins crystallized under a number of conditions, but crystals of sVP40 Δ 43 C320S showed only moderate stability and diffraction. sVP40 Δ 43 C314S (6.5 mg/ml; AW422) and CCS (6.8 mg/ml; AW424) were crystallized using the Morpheus crystallization screen A7 (30% (w/v) GOL_P4K (glycerol, PEG4000), 0.1 M Morpheus buffer 2 (HEPES and MOPS), pH 7.5, and 60 mM Morpheus divalents (MgCl₂; CaCl₂) (Table 6). Unfortunately, the electron density did not allow the assignment of residues 309–323 into either of the mutant models. This suggested that while the overall resolution of the crystal structure is high, the disruption of the disulfide bridge renders the C-terminal turn even more flexible, which could explain the poor electron density of this area. The alignment of both mutant datasets with the sVP40 Δ 43 WT showed a very similar structure (Figure 23 B), suggesting that the mutation of cysteine residues resulted in only minor structural changes.

Table 6: Data collection and refinement statistics of dimeric sVP40 Δ 43 C314S (AW422) and sVP40 Δ 43 CCS (AW424). Values for the outer shell are given in parenthesis.

Data collection	sVP40 Δ 43 C314S	sVP40 Δ 43 CCS
PDB ID	8B1O	8B1P
Space group	5	5
Unit cell parameters [Å] (a, b, c, α , β , γ)	64.25, 91.00, 48.43, 90.00, 96.06, 90.00	63.85, 91.07, 48.38, 90.00, 96.08, 90.00
Resolution range	48.16 – 1.70	48.11 – 1.60
Total number of reflections	113,966	135,900
Number of unique reflections	29,723	35,310

Results

I/ σ	10.1	19.0
Completeness [%]	98.5	98.4
Multiplicity	3.8	3.8
CC _{1/2} [%]	99.4	99.9
Refinement		
Resolution range	45.5 – 1.75	48.11 – 1.70
Rwork/Rfree [%]	18.61/21.58	17.68/20.56
Number of unique reflections	27,545	29,703
Number of protein residues	238	235
Average B factor [\AA^2]	42.37	41.94
R.m.s deviations:		
- Bond length [\AA]	0.006	0.009
- Bond angles [$^\circ$]	0.79	1.02
Ramachandran plot:		
- Favoured [%]	97.37	97.78
- Allowed [%]	2.63	2.22
- Outliers [%]	0.00	0.00
Clash score	4.99	3.02
Number of TLS groups	6	3

C-terminal cysteine residues influence the formation of VLPs and viral genome replication/transcription. Next, we investigated the biological relevance of cysteine residues 314 and 320 of sVP40 in MG and VLP assays (sections 2.3.1 and 2.3.2). We first investigated whether the mutation of the cysteine residues affected the ability of VP40 to downregulate viral genome replication and transcription in a MG assay [231]. We used a Zaire-based MG assay [231], which was supplemented with plasmids encoding sVP40 WT or cysteine mutants. The presence of sVP40 WT decreased reporter activity to approximately 50% in comparison to the control (Figure 24 A). The mutation of cysteine 314 of sVP40 significantly reduced the inhibitory activity of sVP40, and the inhibitory effect was also reduced in sVP40 C320S. Interestingly, the inhibitory effect of the double mutant sVP40 CCS was even higher than that of sVP40 WT. To also analyze the function of the C-terminal cysteine residues of zVP40, cysteine-to-serine mutants were constructed equivalent to the mutants described for sVP40 and employed in the MG assay (Figure 24 B). Upon expression of zVP40 WT, we observed the expected downregulation of minigenome activity as previously reported [142]. Similar to the effect of cysteine mutants of sVP40, zVP40 C311S or zVP40 C314S could not reduce minigenome activity as zVP40 WT could. As observed with sVP40 CCS, zVP40 CCS suppressed reporter gene activity as strongly as zVP40 WT (Figure 24 B).

Results

Taken together, the inhibitory activity of VP40 on reporter gene activity was attenuated by single mutations of the two cysteines in the C-terminus of the molecule, whereas the mutation of both cysteines resulted in proteins that behaved similarly to VP40 WT. These results suggested that cysteine residues at the very C-terminus of VP40 are used to regulate the inhibitory function of VP40 for viral genome replication and transcription.

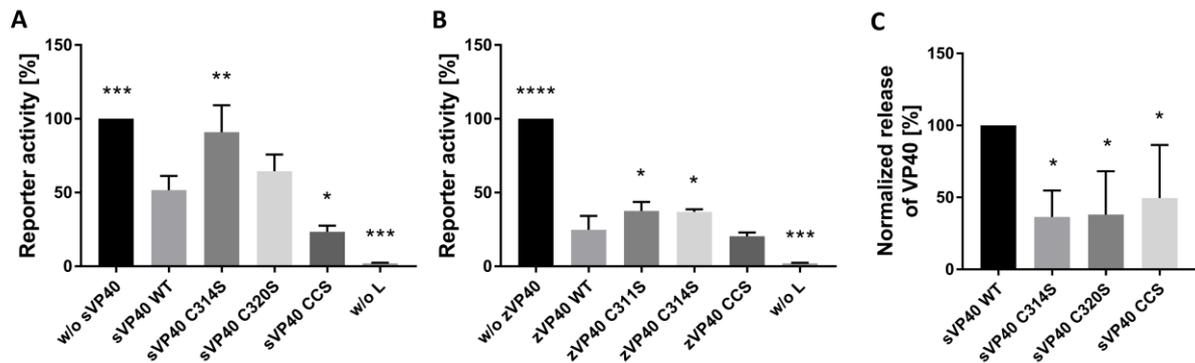


Figure 24: sVP40 cysteine mutants influence reporter gene activity in a MG assay and result in a decrease in VLP formation. A and B) Cysteine-to-serine mutations of s/z VP40 influenced viral transcription/replication in a Zaire-based MG system: HEK293 cells were transfected with pCAGGS VP40 WT or cysteine mutants along with the MG assay components. Cells were lysed 48 h pt, reporter gene activity was measured and normalized to the sample without VP40 (set to 100%). For the negative control, no polymerase L was added to the transfection mix. C) For VLP assays, cells were transfected with pCAGGS GP and sVP40 WT or cysteine-to-serine mutants. Supernatants were collected, and cells lysed 72 h pt WB analyses of cellular and released sVP40 were performed using rabbit α -VP40 and mouse α -tubulin primary and goat α -mouse and α -rabbit IRDye® 680 as secondary antibodies. Released VP40 was normalized to the amount of sVP40 in lysates, and sVP40 WT was set to 100%. Bars indicate the mean \pm SD of at least three independent experiments, asterisks indicate statistical significance (One-way ANOVA) compared to sVP40 WT as follows: * $P < 0.05$, ** $P < 0.005$, *** $P < 0.005$ and **** $P < 0.0001$.

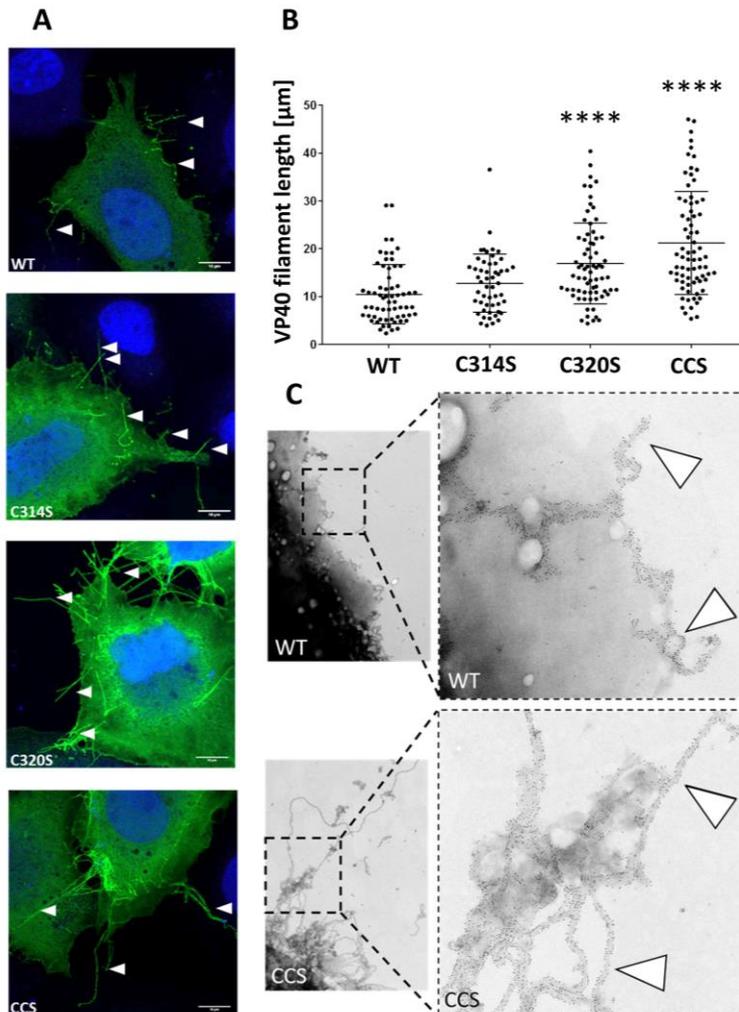
Because VP40 is the main viral factor that induces budding of EBOV particles and VLPs, we investigated whether the release of VLPs was affected by the mutation of the cysteines. HEK293 cells were transfected with plasmids encoding sVP40 WT, C314S, C320S or the double mutant CCS together with GP, the latter to increase the budding efficiency of VLPs [245]. Cell lysates and supernatants were collected at 24 h pt and amounts of cellular and released sVP40 were analyzed by WB analysis (Figure 24 C). sVP40 was visualized using sVP40-specific antibodies and typically exhibited two bands, the lower one probably resulting from an alternative translational start codon, as was reported for zVP40 [157]. The mutation of the cysteine residues to serine resulted in a significant decrease in sVP40 budding efficiency (Figure 24 C) for all three mutants.

We next investigated the intracellular distribution of sVP40 WT and the cysteine variants by immunofluorescence analyses in HuH7 cells (section 2.3.9). HuH7 cells were chosen because of their very flat morphology, which is beneficial for imaging analyses. Ectopic expression of zVP40 results in homogeneously distributed signals in the cytoplasm and prominent VP40-positive clusters at the plasma membrane [136,305]. This pattern was also observed for sVP40 WT and sVP40 C314S, whereas both sVP40 C320S and sVP40 CCS showed significantly more clusters and longer filaments at the plasma membrane than sVP40 WT (Figure 25 A). To quantify this observation, the length of sVP40-positive

Results

filaments was measured (Figure 25 B). Indeed, sVP40 C320S- and sVP40 CCS-positive filaments at the plasma membrane were significantly longer than the filaments positive for sVP40 WT or sVP40 C314S.

Figure 25: sVP40 cysteine mutation C320S results in the formation of longer filaments. A) Representative images of immunofluorescence analyses: HuH7 cells were transfected with sVP40 WT or sVP40 C314S/C320S/CCS. Cells at 24 h pt were fixed, and VP40 was stained as described



using a primary rabbit α -sVP40 and a secondary goat α -rabbit antibody. Nuclei were visualized with DAPI. Confocal images were acquired using a Leica TCS SP5 microscope. sVP40-positive filaments are indicated by white arrowheads. Scale bars are 10 μm . (B) Quantification of filament length: sVP40-positive filaments were measured and quantified using ImageJ 1.52p as described (filaments $n=55-76$). The Kruskal–Wallis test was performed, followed by Dunn’s multiple comparison test. Bars indicate the mean \pm SD, and asterisks indicate statistical significance as follows: **** $P<0.0001$. (C) Electron microscopy of whole mounted cells confirms the presence of long filaments in CCS expressing cells. The filaments exhibit a width of 30 to 50 nm, can be up to several dozen microns long, and are VP40 positive, as shown by immunogold labelling against sVP40 (white arrowheads).

Using electron microscopy analyses performed with sVP40-expressing HuH7 cells grown on transmission electron microscopy (TEM) grids (section 2.3.10), sVP40

WT and sVP40 CCS were detected with anti-Sudan VP40-specific antibodies and a secondary antibody conjugated to colloidal gold (section 2.3.11; Figure 25 C). At the plasma membrane of the sVP40-expressing cells, filamentous structures with a diameter of 30 to 50 nm were detected with specific sVP40 labelling. Generally, longer filaments could be observed in cells transfected with sVP40 CCS compared to sVP40 WT, which was in line with the immunofluorescence analyses.

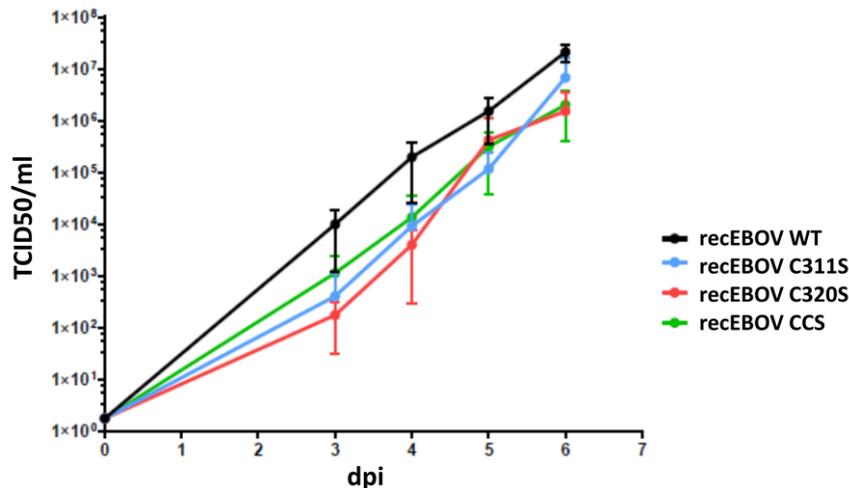
zVP40 cysteines are critical for the efficient release and morphology of recombinant EBOV

Thus far, we have shown that the C-terminal cysteines of VP40 are involved in suppressing viral genome replication/transcription and budding of sVP40-induced VLPs. To further evaluate the biological relevance of the two VP40 cysteines, we generated recombinant (Zaire) EBOV carrying single and double mutations of the cysteines in VP40 (recEBOV C311S/C314S/CCS) and compared the growth

Results

curves of the recombinant viruses in VeroE6 cells to recEBOV WT (sections 2.3.7 and 2.3.8). The growth of all mutant viruses was attenuated compared to EBOV WT (Figure 26). These results suggested that the cysteine residues were important for efficient growth and matched well with the impairment in budding seen in transfection studies for sVP40 as well as the effects seen in the MG assay for sVP40 and zVP40.

Figure 26: RecoEBOV carrying VP40 cysteine mutations exhibits slower growth kinetics than the WT. VeroE6 cells were infected with recEBOV WT or VP40 cysteine mutants (MOI 0.01). A) RecEBOV propagation is affected upon VP40 cysteine mutation. Cell culture supernatants were collected at 0, 3, 4, 5 and 6 dpi and analyzed via TCID50 assays. RecEBOV WT (black) is compared with recEBOV C311S (blue), reEBOV C314S (red), and recEBOV CCS (green) showing the mean \pm SD of three independent experiments.



RecEBOV WT (black) is compared with recEBOV C311S (blue), reEBOV C314S (red), and recEBOV CCS (green) showing the mean \pm SD of three independent experiments.

Next, we investigated the distribution of zVP40 WT and zVP40 mutants in HuH7 cells infected with recEBOVs. Cells were seeded on cover slips at approx. 40% confluency and

infected with recEBOV WT, C311S, C314S or CCS at an MOI of 2 (section 2.3.3.2). Cells were fixed 2 dpi and stained with an α -zVP40 antibody. The distribution of VP40 and filament formation were similar for the recEBOV C311S- and recEBOV CCS-infected cells and resembled the distribution of filament formation of cells ectopically expressing sVP40 C320S and sVP40 CCS in transfection studies (Figure 25 A). While longer filaments could also be observed for sVP40 C320S, zVP40 showed this phenotype for the C311S mutant (Figure 27 A). Since the mutation of zVP40 cysteines in recombinant EBOV led to the formation of longer filaments at the plasma membrane of infected cells, we wondered whether the length of viral particles might also be influenced by the obvious changes in plasma membrane dynamics, which had been described in previous studies analyzing zVP40 cysteine mutants [159]. Therefore, HuH7 cells were infected with recEBOV WT, C311S, C314S or CCS, and the supernatants were harvested at 2 dpi. Viral particles were collected via ultracentrifugation of supernatants of infected cells, and the pellets were fixed in 4% PFA for 48 h. The viral particles were then subjected to negative staining, and their length was measured by transmission electron microscopy (section 2.3.10). First, we noticed that the mean length of the recombinant mutant viruses was significantly longer than that of the WT virus (Figure 27 B). The increased average length in mutant particles is caused by an increase in the relative frequency of longer viral particles; for instance, while on average less than 7% of recEBOV WT particles are longer than 1.3 μ m, this number is more than doubled for the three

Results

recombinant mutants, and this difference is statistically significant for both recEBOV C314S and recEBOV CCS (Figure 27 C).

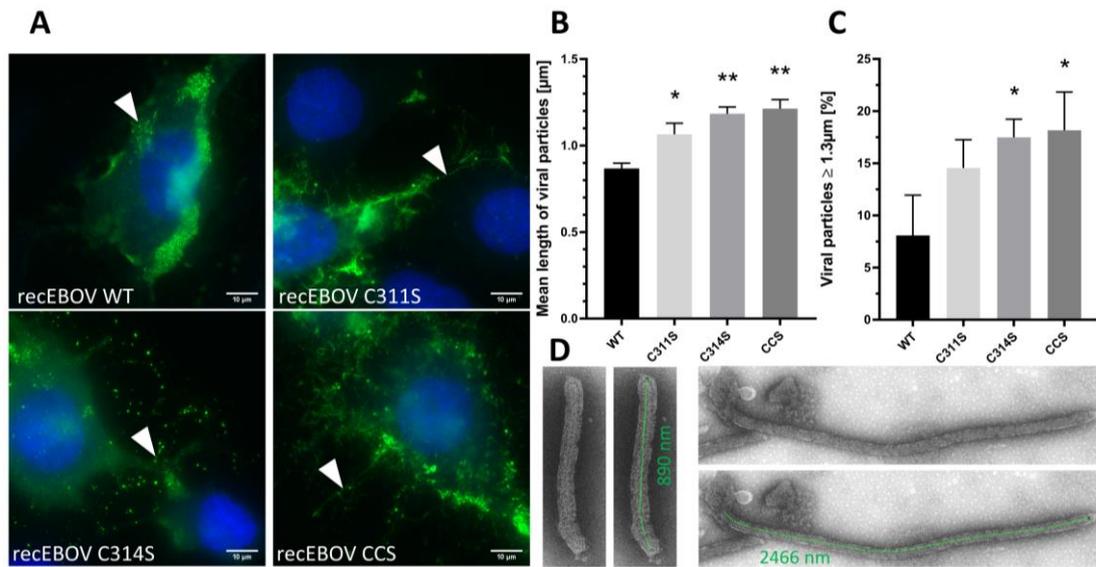


Figure 27: The mutation of cysteines in VP40 of recEBOV leads to the formation of longer filaments in infected cells compared to recEBOV WT-infected cells. A) Cellular distribution of VP40 in cells infected with recEBOV carrying cysteine mutations. HuH7 cells were infected with recEBOV WT/C311S/C314S/CCS at an MOI of 2 and fixed 48 h post infection. α VP40-specific antibodies were used for staining as well as DAPI. B and C) Measurement of the lengths of negatively stained viral particles collected from cell culture supernatant using electron microscopy. Left: Mean length with bars indicating the mean \pm SEM of at least three independent experiments. Right: Percentage of viral particles with a length $\geq 1.3 \mu\text{m}$. Bars indicate the mean \pm SEM for graph B and the mean \pm SD for graph C of at least three independent experiments, the statistical analysis was performed using Tukey's multiple comparisons test (average length) and Dunn's multiple comparisons test (percentage of viral particles $\geq 1.3 \mu\text{m}$). Asterisks indicate statistical significance as follows: * $P < 0.05$ and ** $P < 0.005$. D) Representative EM negative staining pictures of recEBOV WT (left) and recEBOV CCS (right). Green lines indicate the measured length, which is written in green as well.

These data suggest that the expression of cysteine mutants of VP40 causes the formation of longer viral particles, which is in line with what we observed in immunofluorescence analyses of cells expressing cysteine mutants of sVP40 where long VP40-induced filaments were formed in the case of C320S and CCS (Figure 25).

Cysteines in sVP40 are post-translationally modified

Next, it was investigated whether the cysteine residues were targets of posttranslational redox modifications (redox PTMs), such as nitrosylation and glutathionylation. As a first step, we performed a biotin-switch assay (section 2.4.10) using recombinant sVP40 $_{\Delta 43}$. For this assay, the disulfide bridge of sVP40 was reduced using DTT, and the resulting free thiol groups were nitrosylated with nitrosocysteine (Cys-NO). All non-nitrosylated thiol groups were then blocked using iodoacetamide (IAA). Nitroso groups were eventually reduced by sodium ascorbate (Na-Asc), and the free thiol groups reacted with Iodoacetyl-PEG2-biotin. Bound biotin was detected via WB analysis using biotin-specific antibodies and represents nitrosylated cysteine residues in VP40. As a control, the same bands were stained with VP40-specific antibodies afterward. As shown in Figure 28, biotinylated VP40 can be

Results

observed for all samples treated with both Cys-NO and Na-Asc, suggesting that both cysteine residues of VP40 are nitrosylated (Figure 28 A and B). Negative controls represent samples without treatment with either Cys-NO or Na-Asc. To verify these results, we ectopically expressed sVP40 WT and GP in HEK293 cells and collected released VLPs through ultracentrifugation of the supernatant (section 2.3.2). Pellets were analyzed using denaturing mass spectrometry (section 2.4.8.1). Unfortunately, the coverage of trypsin-digested peptide fragments of sVP40 was too low to reach statistical significance of the results. We therefore performed the same experiment with zVP40 and reached sufficient peptide coverage of the C-terminal end of zVP40. As shown in Figure 28 C, cysteine residues 311 and 314 of zVP40 are subjected to nitrosylation as well as glutathionylation. In some cases, sulfide and sulfo groups could also be detected, albeit not in all samples, analyzed via mass spectrometry. These results show that both sVP40 and zVP40 are modified by several redox PTMs, such as nitrosylation and glutathionylation.

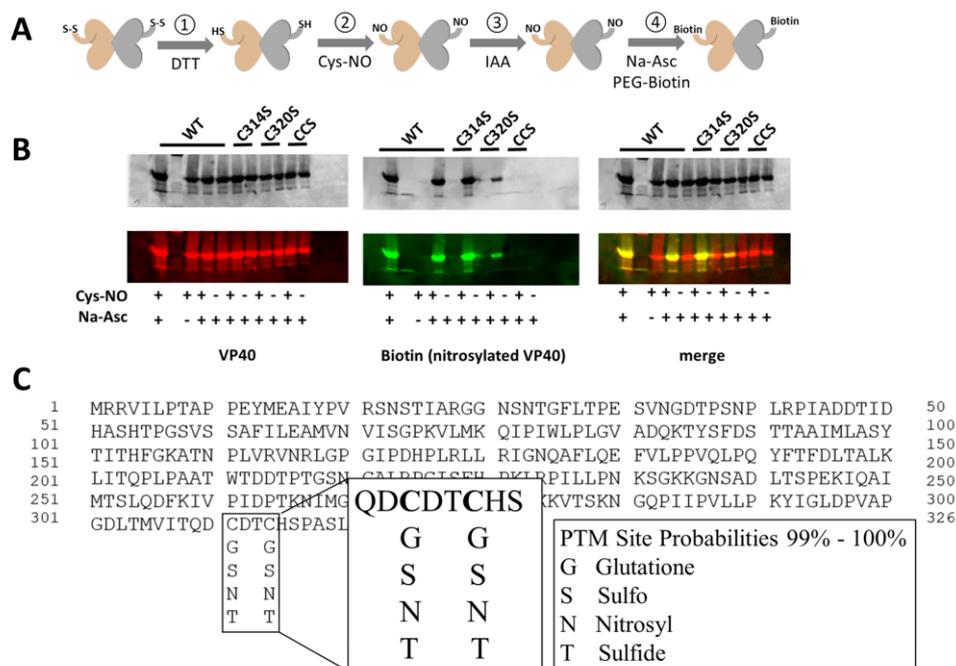


Figure 28: Post-translational modification of VP40. A) Overview of the biotin-switch experiment to detect nitrosylation events in vitro. Recombinant sVP40_{Δ43} was reduced with DTT and incubated with Cys-NO, followed by treatment with IAA. NO groups were then removed using Na-Asc and labeled with PEG2-Biotin. B) Representative WB analysis was performed with 10% sucrose as a blocking agent, mouse α-biotin and goat α-mouse IRDye® 800 antibodies for staining of nitrosylated/biotinylated groups and rabbit α-sVP40 and goat α-rabbit IRDye® 680 antibodies for sVP40-specific staining. Negative controls were not treated with Cys-NO or with Na-Asc. C) PTM view of zVP40 after qualitative protein identification via mass spectrometry.

VP40 is a direct substrate of human thioredoxin

Mass spectrometry results show that VP40 is likely to undergo oxidation when post-translationally modified, but also reduction which yields free thiol groups that can be oxidized again. As mentioned earlier, most VP40 proteins of the Ebola virus different strains exhibit a CXXC motif. Interestingly, this motif is a common recognition pattern among enzymes belonging to the redox homeostasis and their

Results

substrates. zVP40 was therefore investigated as a potential substrate of various redox systems such as the human thioredoxin or the glutathione system. Even though sVP40 does not have the same motif (but rather a CX₅C), it was included in the analysis, as well as sVP40_{Δ43} CCS and zVP40_{Δ43} CCS as negative controls. As a first step, the oxidation state of VP40 expressed in *E. coli* was assessed. Using a 5,5'-Dithiobis(2-Nitrobenzoic Acid) (DTNB) reduction assay [260], which can be used to estimate the number of free thiol groups per molecule VP40 upon reduction of DTNB at 412 nm (see section 2.4.11.1), it was determined that most thiol groups of VP40 (both s/zVP40_{Δ43} WT and CCS) are oxidized and therefore a potential substrate of reducing enzymes. Next, all four recombinant proteins were tested in a thioredoxin reductase assay (see section 2.4.11.2). Both TrxR and Trx are able to reduce suitable oxidized proteins, thereby being oxidized themselves. Oxidized hTrx can then be reduced by TrxR, which in turn can be regenerated by NADPH (Figure 29 A) [338]. For this assay, human hTrxR was isolated from placental tissue and recombinant human Trx was purified from *E. coli* [100,117]. As a first step, VP40_{Δ43} in varying concentrations was mixed with 0,25 U/ml hTrxR and 100 μM NADPH monitoring a slight drop in absorbance at 340 nm caused by the depletion of NADPH during hTrxR reduction. The NADPH consumption of hTrxR during reduction of an oxidized target protein can be monitored using a Tecan infinite M200 multiplate reader at 340 nm [338]. No additional change in absorbance that could be attributed to reduction of VP40 cysteines. This suggested that oxidized VP40 is not a direct substrate of hTrxR. Next, the same reaction setup was used but 10 μM hTrx were added as an additional last step. As expected, a decrease of absorbance at 340 nm was observed in the absence of VP40_{Δ43}, indicating that hTrx was successfully reduced by hTrxR which in turn used NADPH. Moreover, an even further decrease was observed when oxidized VP40_{Δ43} was added to the reaction mixture. This concentration dependent drop of OD was detected for both zVP40_{Δ43} WT and sVP40_{Δ43} WT but not for their respective cysteine double mutants (Figure 29 B and C). These results showed that the oxidized cysteine residues of both zVP40_{Δ43} and sVP40_{Δ43} are direct substrates of hTrx.

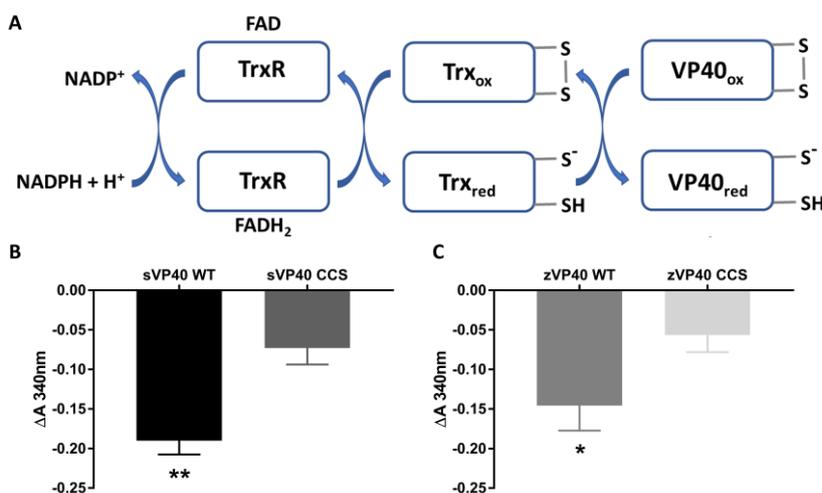


Figure 29: VP40 reduction by the thioredoxin reduction system. A) Overview of thioredoxin reductase system. 40 μM of oxidized sVP40_{Δ43} WT/CCS (B) and zVP40_{Δ43} WT/CCS (C) were mixed with NADPH and TrxR in assay buffer. After monitoring the baseline at 340 nm, Trx was added to the reaction mix and the change in absorbance detected. Bars indicate the mean ± SD of at least three independent experiment and asterisks indicate statistical significance (unpaired t-test) as follows: *P < 0.05 and **P < 0.005.

Results

Next, sVP40 Δ 43 was also investigated as a potential target of human glutathione reductase (GR) or its product glutathione in its reduced version (GSH). The assay (see section 2.4.11.3) followed the same principle as the hTrx/hTrxR assay. In this experiment, sVP40 Δ 43 was neither metabolized by GR nor by GSH (data not shown). Therefore, the specificity of this hTrx-VP40 reaction could be confirmed.

In summary, our data indicates that VP40 cysteine residues are oxidized via post-translational modifications or via the formation of intramolecular disulfide bridges and can be reduced via the Trx system.

3.1.2.3 Discussion

VP40 is a central player in the replication cycle of filoviruses and is responsible for the budding and release of viral particles. In addition, VP40 inhibits viral polymerase activity, and it is currently unclear how the latter activity is exerted or what purpose the negative regulation of transcription/replication might serve. While budding activity of VP40 occurs at the plasma membrane, inhibition of transcription/replication probably occurs in inclusion bodies near the nucleus, the sites of viral replication and transcription [141].

Our study provided new structural information of two regions of sVP40: the intrasubunit hinge between the N- and C-terminal domains of sVP40, spanning residues 195–200 [76], and the C-terminal end of VP40. Several of our sVP40 structures showed either the previously unresolved residues in both the hinge and C-terminal ends or none in either of them, suggesting that the C-terminal arm and the hinge region stabilize each other (data not shown). The C-terminal arm is stabilized by a disulfide bridge between C314 and C320, which are the only cysteines of sVP40. It is therefore not surprising that the C-terminal arms were not visible in either the sVP40 Δ 43 C314S or the sVP40 Δ 43 CCS crystal structure, while the overall crystal packing of the dimeric structure was not altered by the cysteine mutations.

When comparing the VP40 structures deposited in the PDB, it is interesting to note that the C-terminal ends are not visible in any of the published sVP40 Δ 43 crystal structures and adopt different conformations in most zVP40 structures (Figure 30). While residues 311–321 of 1ES6 (Figure 30 C) and 311–322 of 7JZT (Figure 30 D) are located in the pocket between the CTD and NTD, the same amino acids of 4LDB (Figure 30 B) adopt a more open conformation. However, the overall crystal structure of all datasets is the same, further stressing the flexibility of the C-terminal end and the role of the cysteines for the local conformation. The pliability of the C-terminus is potentially necessary for the various functions of VP40 and possibly determines the redox state of the cysteines or vice versa, i.e., an open conformation could be associated with enhanced post-translational modifications. This is in accordance with the publication by Johnson *et al.*, who reported that an increase in the flexibility of the zVP40 C-terminal arm supported interaction with the plasma membrane [159].

Results

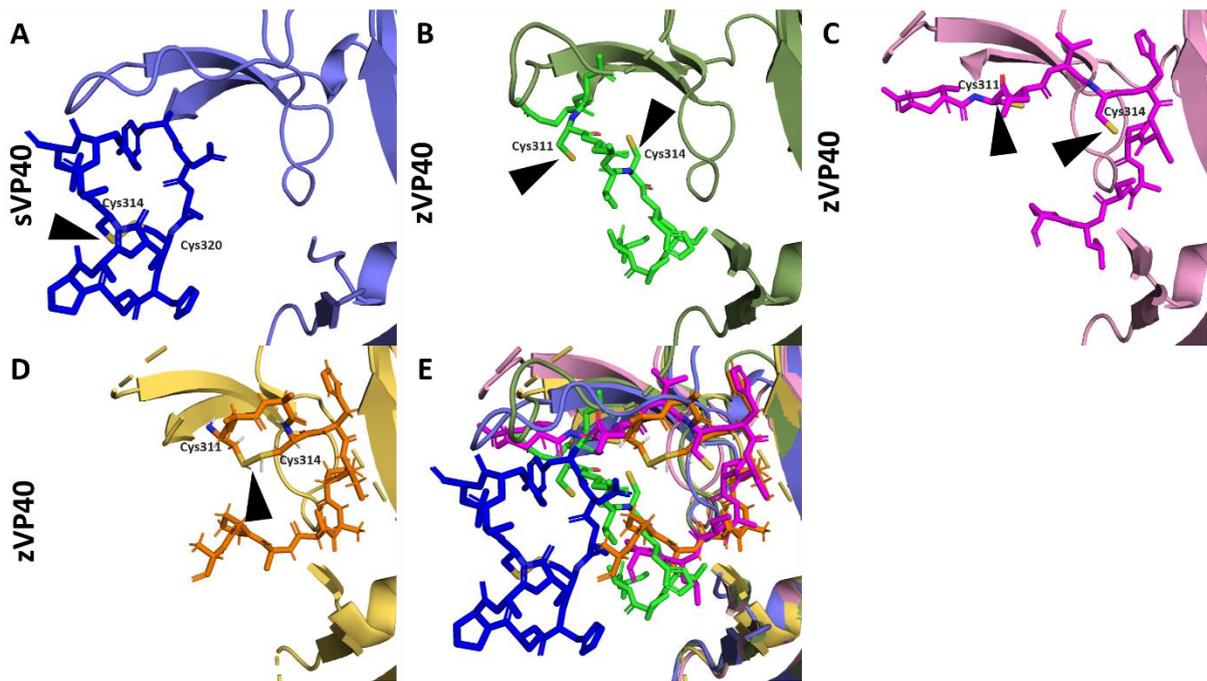


Figure 30: Comparison of the asymmetric unit of sVP40_{Δ43} with one protomer of each zVP40_{Δ43} crystal structure. A) Dimeric sVP40_{Δ43} (blue); B) zVP40_{Δ43} 4LDB (green); C) zVP40_{Δ43} 1ES6 (pink); D) zVP40_{Δ43} 7JZJ (orange) and E) alignment of all four structures with residues of the C-terminal end highlighted in each structure as sticks. Single cysteine residues (B and C) or disulfide bridges (A and D) are indicated by black arrowheads.

The mutation of one or both cysteines of VP40 had multiple effects on various steps of the viral replication cycle, which has been observed for SUDV and EBOV, suggesting that the function of C-terminal cysteines is conserved between the two EBOV species. While VP40 WT inhibited transcription/replication as expected, the mutation of single cysteines in VP40 largely abolished this effect (Figure 24 A and B). Surprisingly, the phenotype of the double mutant VP40 CCS was similar to VP40 WT, suggesting that either both or no cysteines are needed for this particular inhibitory function of VP40. Published data indicate that the octamerization of VP40 is responsible for the inhibitory effect on transcription/replication. The experimental conditions we chose, however, did not reveal sVP40 octamers. Instead, sVP40 WT and the cysteine mutants eluted as dimers from size exclusion chromatography. Previous publications showed that the mutation of either of the C-terminal cysteines of zVP40 shifted the balance between octameric and dimeric VP40 in the direction of the octamers (16). As mentioned earlier, octamerization was reported to be necessary for the suppression of viral polymerase activity [136], and an increase in octamers would suggest a stronger inhibition of viral genome replication and transcription. As an increase in sVP40 octamerization was not detected in our study, it seems likely that the mutation of the cysteine residues does not regulate polymerase inhibition by VP40 via oligomerization of the protein.

The membrane-interaction activities of VP40 resulting in filament growth and budding of particles were also influenced by the cysteine mutations. Here, cysteine-to-serine mutation affected the length of VP40-positive filaments at the plasma membrane, which increased in the order sVP40

Results

WT<C314S<C320S<CCS and was negatively correlated with the budding of virus-like particles (Figure 25). The important role of zVP40 cysteines was confirmed by analyzing recombinant EBOV viruses that encoded VP40 with mutations in one or both cysteines. While the successful rescue of the recombinant mutant viruses indicated that the role of the cysteines is not essential for the viral replication cycle, those viruses grew to lower titers than recEBOV WT (Figure 26). This correlated with the formation of elongated particles containing more than one nucleocapsid and/or surplus membranous material at the particle ends, which is in line with previous reports showing a correlation between the occurrence of elongated or aberrant forms of Marburg virus particles and lower infectious titers [323] (Figure 27 B). The presented data suggest that the mutation of the cysteine residues disturbed the interplay of VP40 with the lipids of the plasma membrane, which determines the ability to induce membrane curvature and facilitate budding. Johnson *et al.* postulated that when the cysteines in zVP40 are mutated, a change in lipid binding may occur that could be responsible for the observed formation of longer VLPs that was also detected in our study [159].

Our analyses revealed that both cysteines of VP40 are subjected to nitrosylation (Figure 28 B) and glutathionylation (Figure 28 C). As we observed various PTMs for both cysteines as well as the disulfide bridge in the VP40 crystal structures, it seems likely that a balance between reduced thiols, probably present as reactive thiolates *in cellulo*, and oxidized thiols (intramolecular disulfide bridges or PTMs) exists, depending on the redox environment. As PTMs expand the chemical repertoire of amino acids, it is not surprising that viruses also exploit redox modifications caused by reactive oxygen or nitrogen species (ROS and RNS, respectively) for their own benefit [144]. Additionally, the formation of ROS and RNS is a hallmark of several viral infections, such as Dengue virus [9,121], chikungunya virus [162], hepatitis C virus [150,151] and several coronaviruses [54,75,329]. It has been shown that glutathionylation modulates the functions of viral proteins such as the NS5 protein of Dengue virus [269], NS5B of Hepatitis C [182], and nsP2 of chikungunya virus [270]. For the latter, conformational changes in the proteins upon glutathionylation are suggested to induce modified function. Interestingly, glutathionylation of one cysteine in the protease of HIV increases activity, while glutathionylation of another cysteine leads to a decrease in protease activity [73,72]. It is currently impossible to pinpoint the role of redox PTMs for the observed phenotypes of the cysteine mutations of VP40. Potentially, redox PTMs induce additional minor conformational changes, act as cysteine–cysteine interaction inhibitors or have their own specific function. Additionally, the differences between nitrosylation, glutathionylation and other redox PTMs are unclear as of yet. Nevertheless, the identification of hTrx as a direct reducing agent of both sVP40 and zVP40 has to the best of our knowledge never been reported for a viral protein before. Especially, sVP40 with its uncommon CXXC motif is of interest, as it also broadens the understanding of substrate specificity of hTrx. The fact, that

Results

the interaction was specific to hTrx only stresses the importance of our findings. However, to understand the biological relevance of the PTMs, further investigations are necessary.

In conclusion, the presented results suggest the regulation of EBOV genome replication/transcription and budding through the two cysteines in VP40, potentially mediated by posttranslational redox modifications such as nitrosylation and/or glutathionylation. Further analyses regarding conformational changes in the local environment induced by the mutation of cysteines or the effect of redox PTMs as well as a deeper understanding of the hTrx-VP40 interaction are necessary to address the molecular mechanism determining the mutants phenotypes.

3.2 PartII: Targeting VP40's homo-oligomerization via interface-mimicking peptides

3.2.1 Introduction

The use of peptides as interface-mimicking PPI inhibitors is considered attractive due to their high affinity towards the target protein and target specificity. Identification of the target amino acids of interest is usually based on structural information and validated via characterization of mutated residues. The binding mode is therefore known from the start [68,77,166]. Inhibitory peptides can be classified into the three categories native, analogue and heterologue. Native peptides exhibit the identical amino acid sequence as in the target protein, while analogous peptides carry modifications and usually display improved drug properties in comparison to their native versions. This category represents the majority of the used peptide inhibitors. The last category comprises peptides, that were not specifically designed for the target protein, but rather discovered through screening assays, as part of a random or preselected library of peptides [189]. According to a recent review, a total of 58 therapeutic peptides targeting PPIs were under clinical development for the past three years with 13 in phase I, 26 in phase II, 15 in phase III. Only two of the 58 peptides in this review were produced recombinantly, the rest was synthesized using solid-phase peptide synthesis (SPPS) [43].

3.2.2 Results

3.2.2.1 Design of interface-mimicking peptides

Structural determination of the different homo-oligomeric states of VP40 revealed anti-parallel interfaces that are conserved among Ebola virus species (see Figure 5 and Figure 58, supplemental material). Their relevance as well as their respective hot spot amino acids were the subject of numerous studies [32,136] and the present work (chapter 3.1.1). The published structures (4LD8 and 1H2C [32]) of zVP40 served as templates during the structure-based design of four interface-mimicking peptides, called core peptides (Figure 31 and Table 7). Peptide #1 mimics the amino acid sequence of the anti-parallel helices in the dimeric interface. Octameric VP40 exhibits several contact surfaces, the intradimeric (#2, residues 90 to 101) and interdimeric interface of the octamer. The latter is discontinuous and spans residues 138 to 153 (#3) and 166 to 177 (#4). For better comprehensibility, the peptides are colour-coded according to their target interface.

Table 7: Sequence of the four core interface-mimicking peptides. Each peptide is described with the residues in question and their position, as well as their target interface (dimer = purple and octamer = green) and the number of residues.

#	sequence	target interface	target oligomer	# residues
1	107-SFDSTTAAIMLA-118	dimeric	Dimer	12
2	90-KQIPIWLPLGVA-101	intradimeric	Octamer	12
3	138-LGPGIPDHPLRLLRIG-153	interdimeric	Octamer	16
4	166-VQLPQYFTFDLT-177	interdimeric	Octamer	12

Results

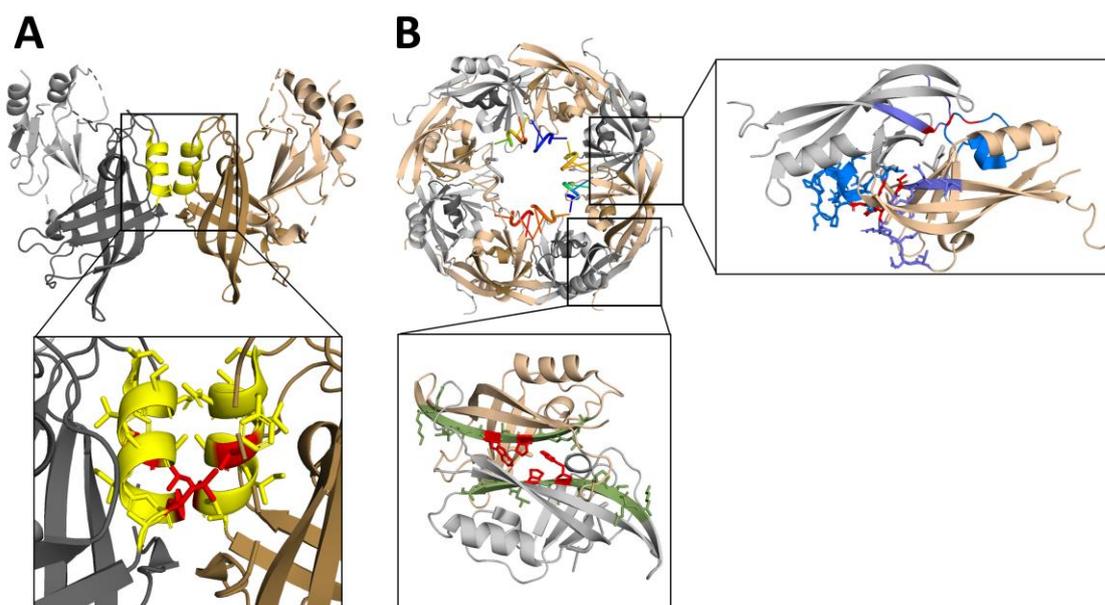


Figure 31: Structure-based design of inhibitory peptides. Interface-mimicking peptides of A) dimeric (yellow) and B) octameric VP40 (intradimeric interface coloured in green and inter-dimeric coloured in dark and light blue) with hot spot amino acids highlighted in red. Residues 59 to 64 are omitted in the dimeric structure, as they partially covered the interface of interest.

A peptide mimicking the C-terminal dimer-dimer interface in filaments was not considered as this discontinuous interface is stabilized by a few residues only, giving the dimer-dimer border the needed flexibility [32]. The four peptides were designated core peptides without any modifications. During this project, further peptides with different tags, varying lengths or protecting groups for the synthesis were designed (Table 8). However, due to their hydrophobic nature, numerous peptides could not be synthesized. Peptides SFDSTTAAIMLA and GVADQKTYSFDSTTAAIMLAS_Y were purchased by Pepmic, Jiangsu, China, whereas the other peptides were kindly provided by Torsten Steinmetzer and Dr. Thuy Van Lam van (Steinmetzer lab, Institute of Pharmaceutical Chemistry, Philipps-University of Marburg).

Table 8: Designed and synthesized peptides. Synthesized molecules are based on the four core peptides targeting different oligomers (dimer = purple and octamer = green) and carry different modifications such as various lengths, tags, mutations of selected residues (modifications underlined) or without N-acetylation (indicated by “non-ac”).

#	sequence	# residues	#	sequence	# residues	
1	SFDSTTAAIMLA	12	2	KQIPI <u>A</u> LPLGVA	12	
	<u>GVADQKTYS</u> SFDSTTAAIMLAS _Y	20		3	LGPGIPDHPLRLLRIG	16
	SFDST <u>ARR</u> IM <u>AA</u>	12			LGPGIPDHPLRLLRIG non-ac	16
	SFDST <u>AAR</u> IM <u>AA</u>	12		<u>flag</u> -LGPGIPDHPLRLLRIG	24	
2	KQIPIWLPLGVA	12	4	LG <u>A</u> GIPDHPLRLLRIG	16	
	KQIPIWLPLGVA non-ac	12		VQLPQYFTFDLT non-ac	12	
	<u>flag</u> -KQIPIWLPLGVA	20		<u>flag</u> -VQLPQYFTFDLT	20	

Results

3.2.2.2 Surface Plasmon Resonance identified promising peptides

As a preliminary screening method, surface plasmon resonance (2.4.9) was employed to test the different peptides with different ligands. As ligands, dimers and octamers of both zVP40_{Δ43} and sVP40_{Δ43} were tested. For better comprehensibility, the different ligands are colour-coded according to their species and oligomeric form (zVP40 dimer dark blue, zVP40 octamer light blue, sVP40 dimer orange, sVP40 octamer light orange). For the immobilisation of the VP40 proteins on the SPR sensor chip, the following conditions were used (see Table 9):

Table 9: Immobilisation conditions and final RU. The used protein concentration and buffer conditions are listed for each ligand (zVP40 dimer dark blue, zVP40 octamer light blue, sVP40 dimer orange, sVP40 octamer light orange). The immobilised RU for each replicate (n=3) are indicated as well. RU = response units.

Ligand	Protein conc.	Buffer	Immobilised RU		
zVP40 dimer	75 µg/ml	10 mM maleate pH 6.5	3008,2	5422,4	5537,3
zVP40 octamer	200 µg/ml	10 mM sodium acetate pH 4.0	8048,8	8048,8	8048,8
sVP40 dimer	75 µg/ml	10 mM sodium phosphate pH 6.0	3517,4	6393,2	5176,9
sVP40 octamer	150 µg/ml	10 mM sodium acetate pH 4,0	5723,7	8669,2	6754,1

Peptides were dissolved in DMSO and diluted in SPR running buffer to a final concentration of 500 µM for a single level binding screen (2.4.9.4). This screen served to identify promising interaction partners for a follow-up analysis and also confirm negative controls. Each peptide passed the sensor's surface with the immobilised VP40 attached and the resulting changes in the refractive index were converted into response units (RU). The RU of each analyte were normalized against the negative control and their respective molecular weight.

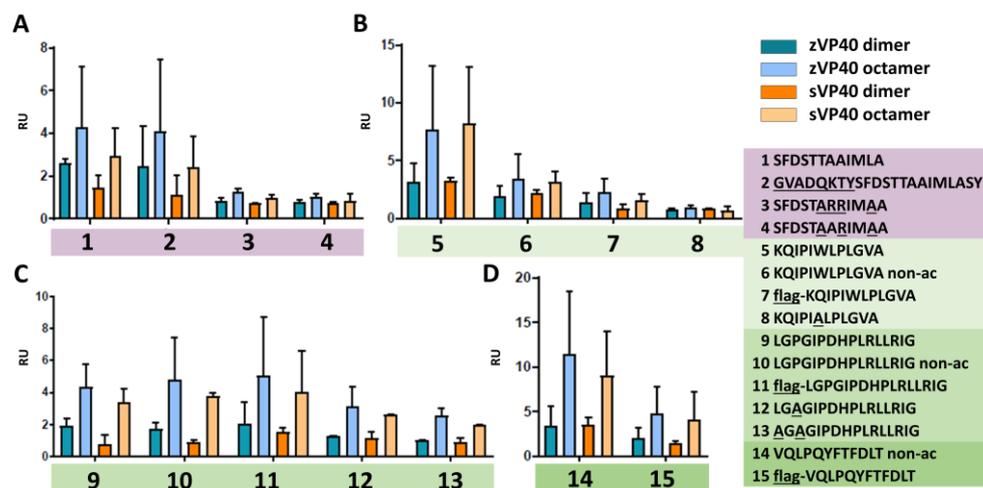


Figure 32: Identification of promising peptides using the SPR Single-Level Binding Screen. All peptides were dissolved in DMSO, further diluted in SPR running buffer to 500 µM and passed over chips coated with the VP40 and also non-coated chips. SRU from each measurement were normalized to the negative control and to the molecular weight of each peptide. Peptides are clustered according to their target interface: A) dimeric interface of the dimer, B) intradimeric interface of the

Results

octamer and C) and D) interdimeric interfaces (non-consecutive) of the octamer and coloured according to the ligand (zVP40 dimer dark blue, zVP40 octamer light blue, sVP40 dimer orange, sVP40 octamer light orange). Bars indicate the mean \pm SD of three independent experiments. RU = response units.

For each tested peptide, RU for octameric ligands were clearly higher than for dimeric ones which can be possibly explained by the higher number of immobilised VP40 molecules, indicated by the RU levels (Figure 32). Also, a clear difference between standard and modified peptides could be observed as molecules carrying point mutations exhibited lower RU, as seen for SFDSTTARRMAA and SFDSTTAARMAA (Figure 32 A), KQIPIALPLGVA (Figure 32 B) and AGAGIPDHPLRLLRIG (Figure 32 C). For core peptide 4 (VQLPQYFTFDLT), no negative control peptide could be synthesized. In addition, a drop in RU could be observed for peptides with flag-tags (flag-KQIPIWLPLGVA and flag-VQLPQYFTFDLT). Interestingly, all peptides (without modifications) seemed to bind to each ligand in general. This was surprising as the different interfaces of their respective homo-oligomer are buried and should therefore not have been accessible. To determine affinities for each analyte to each ligand, another SPR run was performed with additional analyte concentrations (31.25; 62.5; 125; 250 and 500 μ M). Unfortunately, neither affinity nor kinetic parameters could be properly determined for any experiment. Possibly, the steady-state could not be reached or the kinetics constants were outside the detection limits (see Table 23, supplemental material).

3.2.2.3 Peptides did not influence viral replication or release

Next, testing of the four core peptides in functional assays was performed. The peptides were either ectopically expressed from DNA plasmids, or synthetic peptides were transfected and the MG assay or the EBOV infection were performed under BSL2- and BSL4-conditions, respectively.

VP40's ability to downregulate viral genome replication and transcription is associated with VP40 octamer [136]. As three of the four core peptides mimic interfaces observed in the octameric ring stage of VP40, it was expected that their binding to VP40 resulted in a loss of inhibitory function of VP40. Therefore, the peptides were tested in a MG assay. In addition, peptide 1, mimicking the interface of the dimer, was expected to inhibit dimer formation of VP40 and consequently octamerization [32]. Oligonucleotides encoding the four core peptides were fused to a sequence encoding the flag-tag and a sequence encoding a GGGGS-linker, cloned into the pCAGGS vector and the resulting plasmid transfected into HEK293 cells along with the other MG components and a plasmid encoding zVP40 (see section 2.3.1). Peptides containing mutations of the hot spot amino acids served as negative controls and were also tested in the same setting. As shown in Figure 33, no statistically significant differences could be observed for any of the peptides.

Results

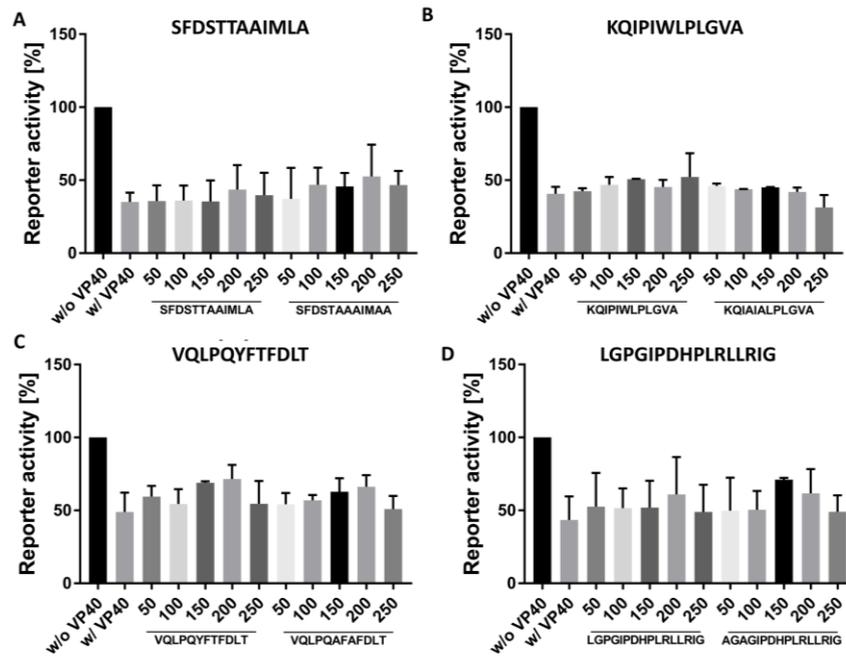


Figure 33: Co-expression of zVP40 and the core peptides in a MG setting. HEK293 cells at 50% confluency were trans-fected with the MG components as well as VP40 and increasing amounts of pCAGGS SFDSTTAAIMLA (A), KQIPIWLPLGVA (B), VQLPQYFFDLT (C) and LGPGIPDHPLRLLRIG (D). Cells were harvested 48 h pt and the supernatant used for luciferase measure-ments. The positive control without VP40 was set to 100%. Bars indicate the mean ± SD of three independent experiment.

For infection studies, HuH7 cells were transfected with selected peptides using the Chariot-transfection reagent, described in section 2.2.3. Cells were then infected with EBOV with an MOI of 0.1. Samples of the cell culture supernatant were taken at 0 and 2 dpi and viral titers assessed via TCID₅₀ assay. While none of the peptides induced a statistically significant decrease in viral titers (Figure 34 A), the strongest impact could be observed for the peptide with the amino acid sequence KQIPIWLPLGVA. A follow-up experiment using increased concentrations of this peptide was therefore performed, but showed no significant effects either (data not shown).

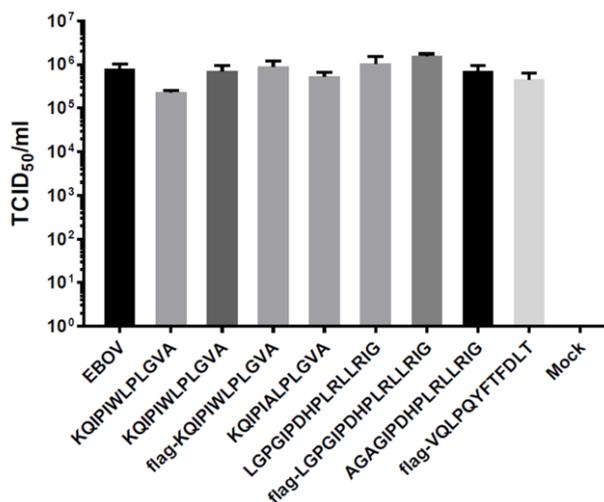


Figure 34: Effect of transfected peptides on EBOV titers. HuH7 cells were transfected with various peptides to a final concentration of 50 µM and afterwards infected with EBOV (MOI 0.1). Supernatants were taken at 0 dpi and 2 dpi and viral titers assessed via TCID₅₀. Bars indicate the mean ± SD of viral titers of three independent experiment at 2 dpi normalized to 0 dpi.

3.2.2.4 Co-expression of peptides with VP40 in *E. coli* did not in alter homo-oligomerization

As the peptides exhibited no effect in functional cell culture assays, their potential inhibition on VP40's homo-oligomerization process was assessed using recombinant VP40 expressed in *E. coli* and served

Results

as a proof of concept. A parallel approach was employed, using co-expressed peptides or a sVP40 Δ 43 - peptide fusion protein.

Only zVP40 was used in this analysis as this protein forms similar amounts of dimers and octamers in *E. coli* (see section 3.3.2.2, Figure 41). zVP40 was cloned into the first multiple cloning site (MCS) of pETDuet-1 (kindly provided by the Klebe lab, Institute of Pharmaceutical Chemistry, Philipps-University Marburg), whereas flag-tagged peptides were cloned into the second MCS. All plasmids with both zVP40 and flag-peptides as well as the control plasmid with an empty second MCS were transformed into Rosetta2 cells and expressed in LB medium overnight at RT (2.4.1). Purification via IMAC was performed, directly followed by SEC (2.4.2). The area under the curve (AUC) was measured using the Unicorn 5 software and the ratio of dimer/octamer of VP40 co-expressed with peptides (#1 to #4) was normalized to zVP40 without a co-expressed peptide (w/o) which was set to 100% (Figure 35). The dimer/octamer ratio was increased upon co-expression with all four peptides but standard deviations were very high for peptide #3 and #4. Co-expression of peptides #1 and #2 resulted consistently in a favoured formation of dimers, albeit not statistically significant.

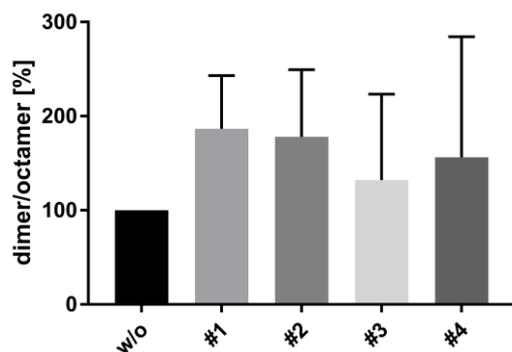


Figure 35: Dimer/octamer ratio of VP40 co-expressed with different peptides. VP40 in pETDuet-1 with or without one of the peptides cloned into the 2. MCS was expressed and purified according to the standard protocol. Size-exclusion chromatography served to identify the ratios of the different homo-oligomers by calculating the AUC of each major oligomeric peak and comparing the dimer/octamer ratio to VP40 without any peptide (set to 100%). Bars indicate the mean \pm SD of three independent experiment.

Next, it was investigated whether the expression of a fusion protein of sVP40 and a selected peptide influenced the oligomerization of sVP40. As a proof of concept, only peptide #1 targeting the dimeric interface was selected. Peptide #1 showed the most promising results in co-expression analyses (Figure 35) and was considered suitable to potentially disrupt oligomerization of sVP40, which forms nearly exclusively dimers. The pET25b+ construct of sVP40 Δ 43 with the C-terminal peptide #1, separated by a flag-tagged GGGGS⁵-linker was purchased from BIOCAT and transformed into Rosetta2 cells. As shown in Figure 36, sVP40 Δ 43 formed significantly more monomers when co-expressed as a fusion protein. As a negative control we introduced a stop codon after the linker via site-directed mutagenesis and performed the same experiment. The SEC elution profile looked very similar to the chromatogram of sVP40 Δ 43-linker in pET25b+ without stop codon, suggesting that the fusion of the linker to sVP40 Δ 43 already introduced a disruption of dimerization (Figure 64, supplemental material).

Results

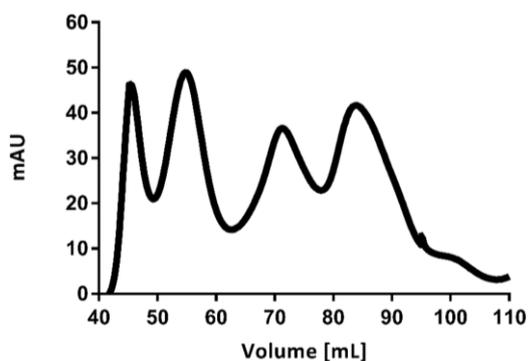


Figure 36: SEC profile of the sVP40_{Δ43}-pep1 fusion protein. The sVP40_{Δ43}-pep1 construct was cloned into pET25b, transformed into Rosetta2 cells, expressed and purified according to standard protocols. Size-exclusion chromatography served to identify the different homo-oligomers of VP40. Representative profile of three independently performed experiments.

3.2.2.5 Co-crystallization or crystal soaking did not yield structural information of peptide binding

The previous data suggested that the various peptides did bind to VP40 but affinities towards the site of interest were either too low to inhibit oligomerization or the peptides were bound to alternative residues of the protein. Structure determination was therefore attempted and the peptides were either co-crystallized with dimeric sVP40_{Δ43} or used for soaking (overnight or 5 sec) with pre-crystallized sVP40_{Δ43} dimer crystals (0.1 mM protein with and 10 mM peptide). Datasets were only included and analyzed using the PanDDA suite if the overall quality and statistics were according to the selected standards such as resolution < 2.5 Å (Table 10). As this tool requires a minimum of 40 datasets, the PanDDA run was performed including the datasets analyzed in section 3.3.2. PanDDA identified one eventful dataset of a co-crystal of sVP40_{Δ43} with peptide VQLPQYFTFDLT non-ac, which was manually evaluated but showed no positive electron density. In addition, all datasets were compared regarding their unit cell constants, as the binding of a peptide would result in an increase of unit cell parameters. However, no dataset merited a manual inspection following this criterion. As shown in the last column of Table 10, no bound peptide could be observed in any of the structures. Alternative methods need to be employed to confirm binding of the peptides to sVP40_{Δ43} also to determine the interaction site.

Table 10: Overview of sVP40_{Δ43} WT-peptide crystals. Peptides were dissolved in DMSO and either diluted in crystallization buffer for co-crystallization or diluted in VP40 buffer for soaking experiments. Crystals were flash-frozen in liquid nitrogen and subjected to X-ray diffraction. Collected datasets were processed using XDS, analyzed using the PanDDA pipeline and visually inspected. The table contains the number of harvested crystals and datasets with diffraction >3 Å for both soaked and co-crystallized crystals for each peptide. X indicates no bound peptide.

	Peptide	# crystals total	Co-crystallization		soaking		result
			# crystals	# datasets	# crystals	# datasets	
2	KQIPIWLPLGVA	11	8	4	1	0	X
	KQIPIWLPLGVA non-ac	7	4	1	3	1	X
	flag-KQIPIWLPLGVA	5	3	1	2	0	X
3	LGPGIPDHPLRLLRIG	7	3	2	4	2	X
	flag-LGPGIPDHPLRLLRIG	4	3	3	1	0	X

Results

4	VQLPQYFTFDLT non-ac	5	3	1	2	2	X
	flag-VQLPQYFTFDLT	3	1	0	2	1	X

3.2.3 Discussion

Structure-based design of potentially inhibiting peptides targeting proteins has been extensively applied in the past [12,190]. For the present work, type 1 peptidomimetics were used exclusively. These peptides mimic the secondary structure and resemble the parent amino acid sequence stretch very closely [11,261]. Crystal structures of both dimeric and octameric zVP40 (PDB-codes 4LDB and 1H2C, respectively) were used for the design and four core peptides and modifications thereof derived from the interface sequence. The single-level binding screen using the SPR technology showed an increased affinity for the core peptides in contrast to their negative controls, carrying mutations of the respective hot spot residues. These results suggested that the designed peptides indeed bound to various forms of VP40 (sVP40 Δ 43 and zVP40 Δ 43 as both dimer and octamer) and served as proof-of-concept that the hot spot residues play a major role in the stability of the respective interface. Determination of affinity and kinetic constants however, was not successful in most cases. Other techniques, such as microscale thermophoresis (MST) [275] or isothermal titration calorimetry (ITC) [186] might have been more suitable. Next, the peptides were tested in functional cell culture assays to analyse their inhibitory activity on viral RNA synthesis as a consequence of the disturbed oligomerization of VP40. Because the process of octamerization was the most likely to be affected by the peptides, a MG assay was chosen, as the downregulation of reporter gene activity is associated with octameric VP40. No significant effect of treatment or co-expression with the peptides could be observed in either assay. This data indicates that the selected peptides were not able to prevent VP40 oligomerization with the chosen methods. While the peptides were in general able to bind to VP40 they were not able to inhibit VP40's oligomerization neither when expressed in *E. coli* nor upon treatment of mammalian cells. The variety of selected methods and peptide delivery options suggests that the affinity towards the target protein was probably too low in any of the experiments. As already mentioned in the introductory section, a number of reasons might be accountable for this, which will be discussed in detail.

The four core peptides were designed with only 12 to 16 residues, but were also synthesized with different tags and extensions to induce proper folding. For co-expressed peptides, a GGGGS linker was added for the same reason. Hartlieb *et al.* used synthetic peptides exclusively with a peptide length of 24 amino acids [126]. In addition, in a similar paper by our working group, the VP30-PP2A interaction was disturbed by a peptide mimicking the B56-binding LxxIxE motif on NP. While the sequence of interest is very short and only spans six residues, they expanded the peptide to 34 amino acids and

Results

cloned four copies of it into a YFP-containing mammalian expression vector [178]. Nevertheless, a number of peptides in antiviral research were of similar length such as a hexapeptide blocking entry of Severe acute respiratory syndrome coronavirus 1 (SARS-CoV 1) [295], a 12-mer that disrupts Human immunodeficiency viruses (HIV) capsid assembly [336], 12-mer peptides binding to HIV Vif protein [333], two 15-mers inhibiting HIV integrase and reverse transcriptase [113] and a 10-mer inhibiting dimerization of the HIV reverse transcriptase [228]. Combining the results from the present work with the literature results, it seems unlikely that peptide length was the major issue as the longest peptide contained 24 amino acids. As mentioned earlier, peptide stability is a widely discussed issue [77]. Various options for structural modifications are available that typically modify other parameters such as solubility, permeability or ADME properties. As peptides are prone to degradation by exopeptidases *in vivo*, adding protective groups to the N- or C-termini via N-acetylation or C-amidation was reported to decrease the level of proteolysis. Most synthetic peptides used for the present work carried these modifications [77,116,258,274]. Interestingly, for two peptides without N-acetylation (KQIPIWLPLGVA non-ac and LGPGIPDHPLRLLRIG non-ac) relatively high RU values could be determined in the single level binding screen (especially for octameric VP40; see section 2.4.9.4 and Figure 32), however, affinities could not be determined. The exception was VQLPQYFTFDLT non-ac which showed promising results in both the single level binding screen and affinity determination with K_D values in the high μM to low mM range but was never tested in a cell culture setting. A different approach is the substitution of L-amino acids with D-amino acids to increase stability as proteases are less likely to recognize and bind to their target [56,307]. Even a peptide consisting of only D-amino acids was reported to be highly resistant to protease degradation but exhibiting high affinities towards the HIV target protein [322]. Cyclization, the introduction of disulfide bonds, or the synthesis of stapled peptides are additional methods with the off-trade of a reduced flexibility but nevertheless highly efficient [78,306,337]. Judging the improvements of peptide efficacy in *in vitro* and *in vivo* studies, it seems advisable to invest in synthesis strategies other than N-acetylation and C-amidation which were used for the present work. However, considering the variety of performed experiments and with special focus on the expression of a sVP40 $_{\Delta 43}$ -pep1 (SFDSTTAAIMLA) fusion protein, it seems likely that protein degradation was a problem but not the ultimate one because it has been shown that the peptides do bind to the proteins of interest. Rather, low affinities might be accountable for the lack of impact in the different assays using both functional cell culture assays and recombinantly expressed proteins, as the peptides were apparently not able to outcompete other VP40 molecules. A crucial next step therefore would be to produce a peptide array on the basis of the four core peptides, and screen for ideal sequence composition and length. Such a peptide array approach was conducted to identify ACE2-binding peptides, designed to inhibit SARS-CoV-2 entry [209].

Results

Another possible explanation for the lack of efficacy is the expression of only one peptide at a time. Babé *et al.* showed that only a two-stranded peptide was able to block assembly of HIV proteases [12]. Also, four copies in the same mammalian expression vector were needed to raise the peptide concentration sufficiently to inhibit the EBOV PP2A-VP30 interaction [178]. Another option to improve interface-mimicking peptides is the engineering of a protein of interest containing a cysteine instead of an amino acid close to the interface without being involved in its interactions or stability. The synthetic or co-expressed peptide also carries a thiol group and can therefore be tethered to the protein via formation of a disulfide group, thus greatly improving affinity. This approach was conducted for studies by Buck *et al.* and served as a starting point for a screening of thiol-containing small molecules [41,42].

In spite of the discussed possibilities to improve the peptide's properties, the focus of the present work shifted towards the development of small molecules as potential drug candidates. Fragments are more likely to be suitable as oral treatment and are both easier to synthesize and store than peptides. These properties are considered to be highly important for the advancement of anti-EBOV drugs being distributed to especially rural areas of affected countries [214].

3.3 Part III: Fragment-based lead discovery

Fragment-based drug design (FBDD) or fragment-based lead design (FBLD) has become a widely used approach to identify molecular binders that can be developed into promising inhibitors. The screening process usually necessitates only a relatively small library of fragments with low molecular weights because the hit rate increases as the size of the molecules decreases. In addition, fragments offer potential binding to small pockets that are hard to reach by conventional (larger) inhibitors. On the other hand, the molecules need to be large enough to form interactions with the target protein, typically exhibiting affinities in the μM to mM range. To generate lead structures, fragments (or lead-like molecules) must be optimized regarding size and functional groups via fragment growing, linking or merging. The low chemical complexity of fragments offers the exploitation of a large chemical space to achieve sufficient extension. However, the selected screening method needs to be sensitive enough to detect these usually weak binders and specific enough to distinguish low binders from false-positive ones [85,95,170,183,235].

To identify VP40-inhibiting compounds, two parallel approaches of FBLD were employed for the present work. While disulfide tethering (DT) enables site-specific binding of thiol-containing fragments to a native or engineered cysteine close to the location of interest, crystal soaking detects low-affinity binders to the protein of interest without guiding the ligand to a specific binding site. The latter approach offers the advantage of immediate structural information which gives insight in not just the exact binding location but also the binding mode of the fragment. DT on the other hand is a less laborious approach and downstream modification and enlargement of potential fragment hits are more intensive due to removal of the thiol group.

As potential binders identified via crystal soaking are not necessarily able to inhibit homo-oligomerization of VP40, the present two-pronged approach offers the additional opportunity to steer the binders to an oligomerization interface of interest thus doubling the chances of developing a functional inhibitor.

3.3.1 Site-directed ligand discovery via disulfide tethering

3.3.1.1 Introduction

The novel technique DT (see section 2.4.5) was designed to guide molecules to a specific site on the surface of the protein (site-directed ligand discovery, SDLD) [87] to circumvent typical issues during the FBDD process. These include the binding of fragments to random sites of the protein of interest, the potentially low structural accessibility of binding sites and the need for highly specific (high-throughput) assays and/or high compound concentrations that are necessary to detect weak binders, the latter often resulting in assay artifacts. The DT approach exploits the formation of disulfide bridges

Results

between a native or engineered cysteine at or near the site of interest of the protein and a ligand. The ligands originate from a library of thiol-containing molecules and are incubated with the protein under reducing conditions. This results in an equilibrium between thiol oxidation and reduction and binding of compounds is favoured when further interactions occur, additional to the disulfide bridge. The method can therefore detect weak binders because the protein-compound interaction is greatly increased due to the covalent disulfide bonds which can be detected by a mass shift using mass spectrometry [87,88,123]. This method was successfully employed in numerous studies [42,86,128,132,167,286,334].

For the present work, DT was used to identify compounds binding to a cysteine that was introduced via point mutagenesis replacing asparagine 67 of sVP40 (sVP40 $_{\Delta 43}$ N67C), close to the dimeric interface to possibly weaken sVP40 dimerization (Figure 37). The principle of DT was already described in section 2.4.5. Briefly, sVP40 $_{\Delta 43}$ N67C was incubated with thiol-containing fragments under reducing conditions, desalted and analyzed via IP-MS.

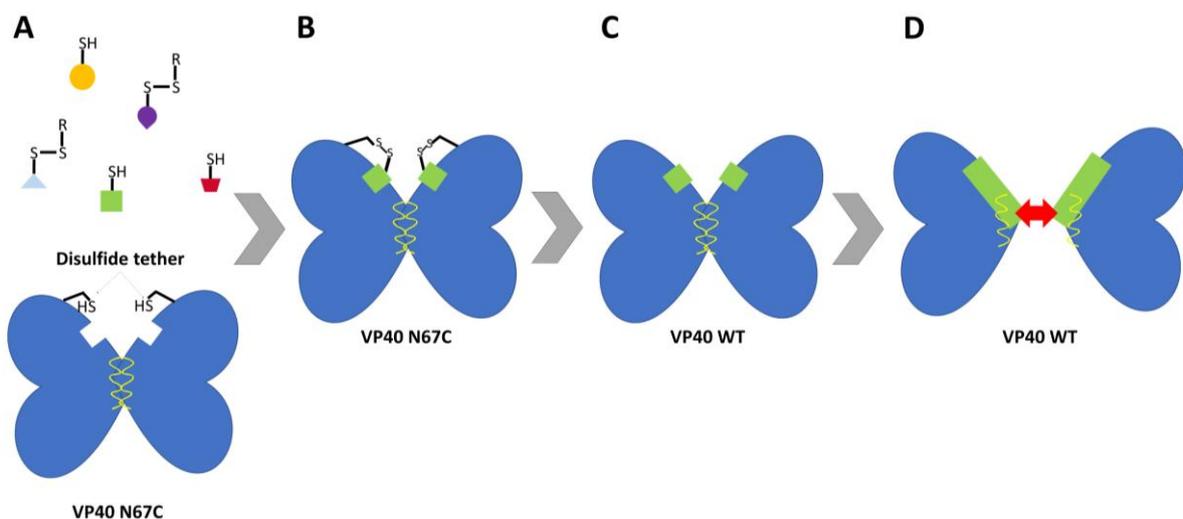


Figure 37: Principle of lead discovery via DT. A) Recombinant sVP40 $_{\Delta 43}$ N67C is incubated with thiol-containing fragments and DTT. B) Due to reducing conditions, fragments with other interactions than the disulfide bridge are bound to sVP40 $_{\Delta 43}$ N67C favourably and can be detected via IP-MS, such as the molecule depicted in green. C) The identified lead-like structure is tested on VP40 WT and D) is able to disrupt dimerization after optimization of binding upon fragment extension.

3.3.1.2 Results

Generation of recombinant VP40 with engineered cysteines

VP40 carries two cysteines at its C-terminal end which adopts a very flexible conformation. While the importance of these cysteines was described in section 3.1.2, the suitability of these residues for DT remains questionable due to their structural flexibility. An additional cysteine residue was therefore introduced at position 67 into sVP40 $_{\Delta 43}$ which replaced asparagine. This residue was chosen as it is in the vicinity of the dimeric interface of dimeric sVP40 but not too close, to avoid the formation of a disulfide bond with the other dimer half. In addition, a drastic change in the local environment seemed unlikely due to the similar properties of Asn compared to Cys. Thus, two constructs were generated:

Results

sVP40 $_{\Delta 43}$ N67C and sVP40 $_{\Delta 43}$ N67C-CCA, the latter carrying additional mutations of the native cysteines (C314A and C320A) to avoid unwanted intramolecular disulfide bridges. Both proteins were expressed in Rosetta2 cells and lysed according to the standard VP40 protocol (see section 2.4.1 and 2.4.2). However, after the initial purification via IMAC, the proteins precipitated during the subsequent concentration step using Amicon spin concentrators. Therefore, a total elution volume of 5 mL was chosen without further concentration and the eluted protein was treated with DTT (final concentration of 5 mM). This was necessary, as the protein formed major aggregations during the following SEC step. As shown in Figure 38, both sVP40 $_{\Delta 43}$ N67C and sVP40 $_{\Delta 43}$ N67C-CCA eluted as dimers.

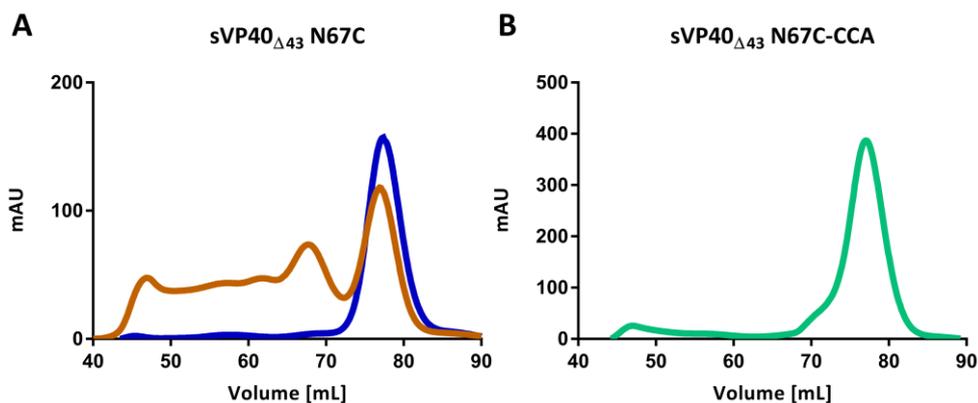


Figure 38: Size-exclusion profile of engineered sVP40 $_{\Delta 43}$ mutants. Recombinant sVP40 $_{\Delta 43}$ N67C (A) and N67C-CCA (B) in the pET46 vector were generated using site-directed mutagenesis, expressed in Rosetta2 cells and purified via IMAC and SEC using standard protocols with the exception of DTT treatment. A) sVP40 $_{\Delta 43}$ N67C without (brown) and with DTT (blue) treatment and B) sVP40 $_{\Delta 43}$ N67C-CCA after DTT treatment only.

Screening of thiol-containing fragments using IP-MS

30 μ M of dimeric sVP40 $_{\Delta 43}$ N67C were used for the initial screening together with fragments (single or in combinations of two or three fragments) and DTT (see section 2.4.5). The library of thiol-containing molecules was kindly provided by the Klebe lab, Institute of Pharmaceutical Chemistry, Philipps-University Marburg. After the incubation period, the proteins were desalted and analyzed via IP-MS, resulting in mass spectra with protein peaks as an output. Fragment hits were identified by a mass shift of unmodified protein plus fragment mass and were categorized in three classes based on mass spectra: +++ for a clear mass shift with a higher peak of modified protein than for unmodified, ++ for clear mass shift with a lower peak of modified protein than for unmodified and + for a detectable but ambiguous mass shift.

For the initial screening sVP40 $_{\Delta 43}$ N67C was used. This mutant still carried the two native cysteines that were characterised in Part I on this work. Figure 39 shows mass spectra of both unmodified (A) and one example of modified (B) sVP40 $_{\Delta 43}$ N67C, with results summarized in Table 11. Mass spectra of other identified hits are shown in the supplement (Figure 65 to Figure 68). Further detected mass peaks

Results

likely refer to modified protein species, including protonation, adducts with small ions such as Na⁺ or small organic molecules such as acetonitrile, and fragmentation products.

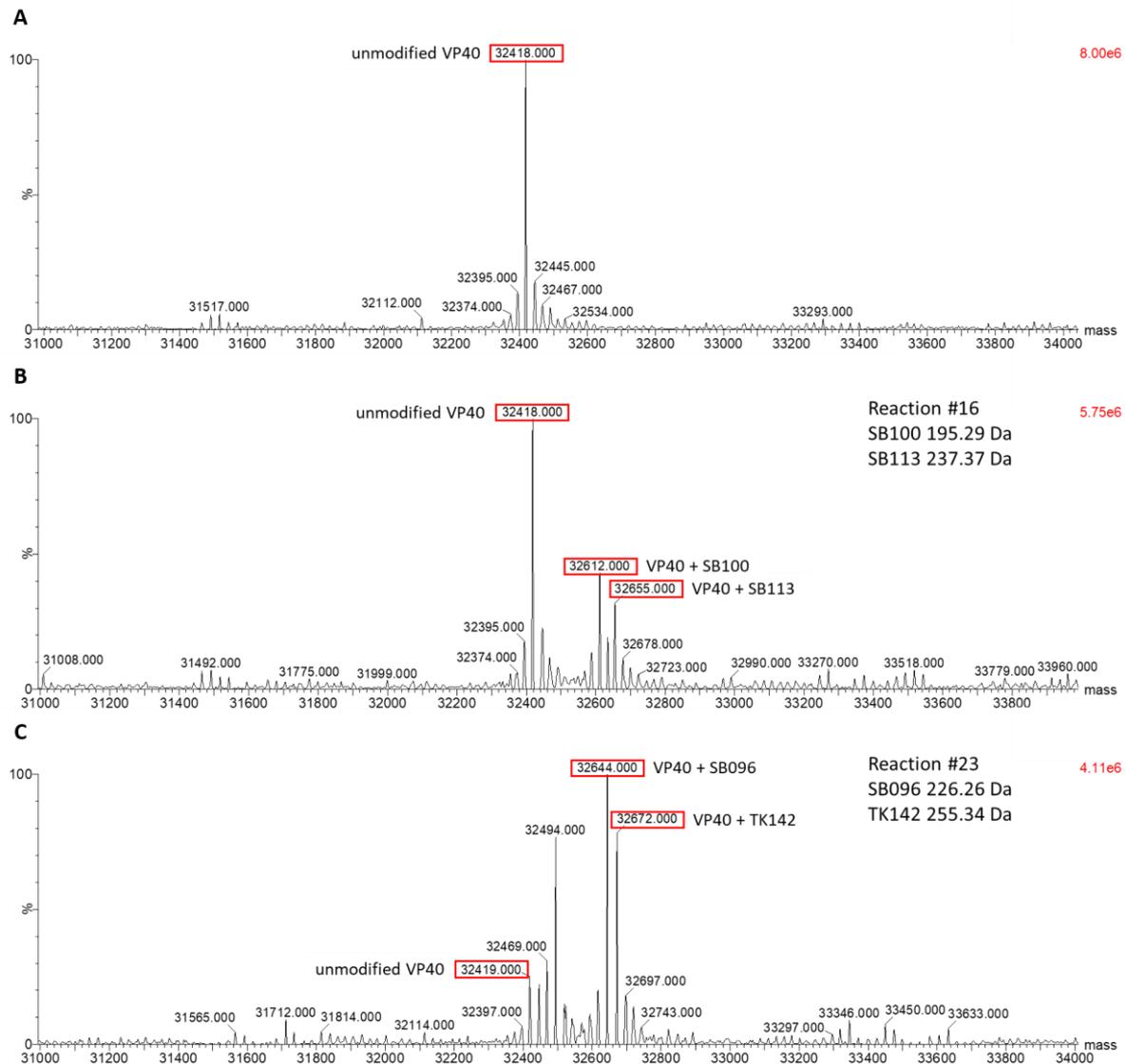


Figure 39: IP-MS spectra of sVP40_{Δ43} N67C. Recombinant sVP40_{Δ43} N67C was incubated with fragments or DMSO only under reducing conditions according to 2.4.5.1 and analyzed via IP-MS. A) Untreated protein served as a negative control to determine the native mass, B) sVP40_{Δ43} N67C incubated with SB100 (195.29 Da) and SB113 (237.37 Da) (reaction #16) and C) sVP40_{Δ43} N67C incubated with SB096 (226.26 Da) and TK142 (255.34 Da) (reaction #23).

Table 11: DT screening hits for sVP40_{Δ43} N67C. Only fragments are listed that were categorized as +, ++ or +++ per reaction (indicated by #; performed after the protocol outlined in Table 1). Reactions (#) without a fragment hit are not listed.

#	Fragment	MW	Rating	Fragment	MW	Rating
1	SB115	231.32	+++	TK103	265.38	+
2	SB124	307.15	+			
3	TK169	251.35	+++			
4	TK163	269.37	++			

Results

5	TK148	237.32	+++	TK098	273.4	+
6	SB095	181.26	+++	TK109	262.38	+++
7	TK106	276.4	+++			
8	TK149	211.28	++	TK170	251.35	++
	SB227	285.41	++			
9	TK096	251.37	+			
10	SB101	213.28	++	TK095	252.31	++
	SB101	213.28	+	TK100	264.35	+
11	TK162	243.39	+			
12	SB229	280.17	++	TK099	313.47	+
15	TK101	266.34	++			
16	SB100	195.29	++	SB113	237.37	+
18	TK147	197.26	+++	TK097	239.34	+++
19	TK158	239.34	+			
22	TK110/TK140	225.31	+++			
23	SB096	226.2	+++	TK142	255.34	++
25	TK141	229.37	+++	TK107	259.37	++
27	TK174	239.34	+			

In summary, 11 fragments were categorized as +++, 11 fragments were categorized as ++ and 11 fragments were categorized as +.

To determine which of the three cysteines was involved in the interactions, a reduced number of fragments was tested on both sVP40_{Δ43} WT and sVP40_{Δ43} N67C-CCA. Surprisingly, no mass shift could be observed for sVP40_{Δ43} WT which stresses the flexibility of the C-terminal loop and the few interaction sites this region had to offer. A mass spectrum of unmodified sVP40_{Δ43} WT is given in Figure 69, supplemental material. Figure 40 shows mass spectra of both unmodified sVP40_{Δ43} N67C-CCA (A) and three examples of modified sVP40_{Δ43} N67C-CCA (B and C) and Table 12 summarizes the results. Mass spectra of other identified hits are shown in the supplemental material (Figure 70 and Figure 71).

Results

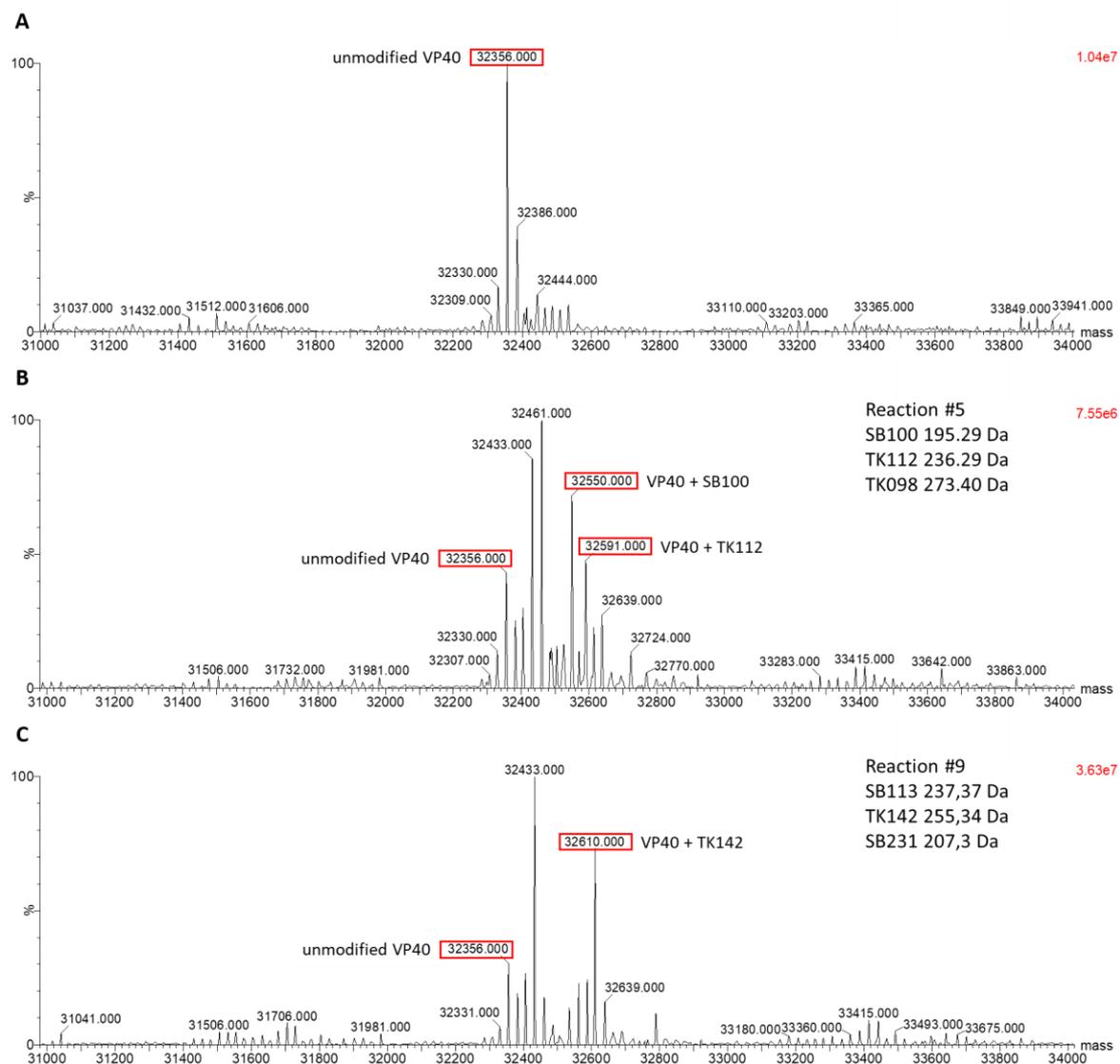


Figure 40: IP-MS spectra of sVP40 Δ 43 N67C-CCA. Recombinant sVP40 Δ 43 N67C-CCA was incubated with fragments or DMSO only under reducing conditions according to 2.4.5.1 and analyzed via IP-MS. A) Untreated protein served as a negative control to determine the native mass. B) sVP40 Δ 43 N67C-CCA incubated with SB100 (195.29 Da), TK112 (236.29 Da) and TK098 (273.40 Da) (reaction #5) and C) VP40 incubated with SB113 (237.37 Da), TK142 (255.34 Da) and SB231 (207.30 Da) (reaction #9).

Table 12: DT screening hits for sVP40 Δ 43 N67C-CCA. Only fragments are listed that were categorized as +, ++ or +++ per reaction (indicated by #; performed after the protocol outlined in Table 2). Reactions (#) without a fragment hit are not listed.

#	Fragment	MW	Rating	Fragment	MW	Rating	Fragment	MW	Rating
1	SB124	307.15	++						
2	SB101	213.28	++	TK169	251.35	+	SB227	285.41	+
3	SB101	213.28	++	TK148	237.32	++	SB229	280.17	+
4	TK163	269.37	+						
5	SB100	195.29	+++	TK112	236.29	+++			
6	TK096	251.37	++	TK106	276.4	++			

Results

7	TK147	197.26	+	SB096	226.26	+	TK109	262.38	+
8	TK095	252.31	+						
9	TK142	255.34	+++						

In summary, for sVP40_{Δ43} N67C-CCA 2 fragments were categorized as +++, 6 fragments were categorized as ++ and 8 fragments were categorized as +. When comparing the first and third screening (using sVP40_{Δ43} N67C and sVP40_{Δ43} N67C-CCA, respectively), it could be noted that only few fragments were categorized in the same category (+, ++ or +++) Fragments with similar categories and at least ++ in both screenings include SB100, SB101, TK106 and TK142 (Table 13).

Table 13: Comparison of DT screening hits between sVP40_{Δ43} N67C and sVP40_{Δ43} N67C-CCA. Only fragments hits are listed that were included in both screenings.

Fragment	N67C	N67C-CCA
SB096	+++	+
SB100	++	+++
SB101	++	++
SB124	+	++
SB227	+++	+
TK095	++	+
TK106	+++	++
TK109	+++	+
TK142	++	+++
TK148	+++	+
TK163	++	+
TK169	+++	+

Optimization of screening conditions for identified fragment hits

Next, further DT screening conditions were tested to determine ideal concentrations for the future setup of crystallization screenings. TK142 were used as well as SB100. The same categories were applied for the screening protocol, as shown in Table 14:

Table 14: DT optimization results for TK142 and SB100 with sVP40_{Δ43} N67C-CCA. Only fragments hits are listed that were included in both screenings. Recombinant sVP40_{Δ43} N67C-CCA was incubated with various concentrations of DTT as well as TK142 or SB100 according to the protocol outlined in 2.4.5.2.

Fragment [μM]	DTT [μM]	SB100	TK142
100	250	-	+
200	250	-	++
500	250	-	++
1000	250	-	-
100	500	-	+
200	500	-	++

Results

500	500	-	+
1000	500	-	-
100	750	-	+
200	750	-	+
500	750	-	-
1000	750	-	-

Surprisingly, binding of SB100 could not be reproduced and no fragment concentration yielded a mass shift which suggested potential degradation during freeze-thaw cycles. For TK142, the following pattern could be observed: no mass shifts were observed for fragment concentrations of 1000 μM and DTT concentrations of 750 μM seemed less successful. However, no clear correlation between DTT and fragment concentrations could be deduced from this optimization round. Additionally, reproduction of the previous screening conditions did not result in the exact same mass shift and ratio of modified:unmodified protein. Nevertheless, TK142 was bound to VP40 in most cases, confirming it as a promising fragment. Judging from the optimization results, ideal concentrations include 250 μM DTT and 500 μM fragment for potential future co-crystallization attempts.

3.3.1.3 Discussion

DT is a useful tool for site-directed ligand discovery. For the present work, an additional cysteine was introduced into sVP40 $_{\Delta 43}$ at position 67 (sVP40 $_{\Delta 43}$ N67C) in order to add a thiol anchor in close proximity to the dimeric interface. Purification of the recombinant sVP40 $_{\Delta 43}$ N67C via SEC in the absence of reducing agents resulted in aggregation, therefore necessitating the incubation with DTT prior to SEC. However, the formation of alternative oligomeric forms of sVP40 would have been advantageous as it has been previously reported that a similar approach resulted in an alternate, covalently linked dimer with an exposed original dimeric interface due to the new connection of the monomers via the engineered disulfide bridge [242]. However, purification of sVP40 $_{\Delta 43}$ N67C resulted in the formation of aggregates in the absence of DTT, suggesting that the position of the introduced cysteine was not ideal for the formation of an alternative dimer, similar to the work by Nguyen *et al.* [241]. Nevertheless, covalently bound dimers represent a promising option for further testing, such as the identification or validation of interface-binding compounds. However, for screening purposes, the sVP40 $_{\Delta 43}$ N67C and sVP40 $_{\Delta 43}$ N67C-CCA mutants were highly stable after a single treatment with DTT (5 mM for 30 min). During SEC, the DTT was washed out but dimers were stable even in the absence of DTT, suggesting that once dimers connected through the original dimeric interface, they did not undergo further conformational changes. Hence, it could be concluded that the introduced cysteine at position 67 was not involved in a protomer-protomer disulfide bridge but was free and accessible for fragment binding. For the first screening, sVP40 $_{\Delta 43}$ N67C was used due to the fact that the native cysteines C314 and C320 were shown to be important for both the negative regulation of viral genome replication and

Results

transcription as well as budding and the morphology of viral particles (Figure 24). Therefore, blockade of the two residues might inhibit the function of VP40. Additionally, neither C314S nor C320S underwent major conformational changes upon dimerization (Figure 23). Fragment binding to either of the three cysteines was therefore considered promising. The promising fragments from the first screening were tested in a subsequent screening using sVP40_{Δ43} WT and sVP40_{Δ43} N67C-CCA. Interestingly, no mass shift could be observed for the DT using sVP40_{Δ43} WT. This was however not surprising, as only fragments bound to VP40 not just via the disulfide bridge but also via additional interactions were expected to form stable complexes. The crystal structure of the sVP40_{Δ43} cysteine mutants (C314S and CCS) did not include positive electron density for the C-terminal loop, suggesting that the disruption of the disulfide bridge increased its flexibility (see section 3.1.2.2). A similar phenomenon could potentially be observed for the treatment with DTT, therefore rendering the C-terminus even less ordered when the disulfide bridge is reduced. Additionally, we expected this domain to adopt several conformations after comparing the high-resolution crystal structure described in section 3.1.2 (PDB-code 8BX3) to the different structures available for zVP40 (Figure 30). Hence, the C-terminal loop was considered to be too flexible to form stable complexes with fragments apart from the disulfide bridge, suggesting that all fragments were exclusively bound to the introduced C67. As degradation due to repeated freeze-thaw cycles as well as fragment-fragment interactions could not be excluded, results of both DT round #1 (sVP40_{Δ43} N67C) and #3 (sVP40_{Δ43} N67C-CCA) were evaluated together, even though follow-up experiments were performed with sVP40_{Δ43} N67C-CCA, exclusively. According to Table 12, fragments SB100 and TK142 were considered most promising for sVP40_{Δ43} N67C-CCA and were classified similarly for sVP40_{Δ43} N67C (DT round #1) (Table 11). However, experiments for optimized screening conditions were largely inconclusive. Irregularities during optimization were also observed in another study, using the same library [241]. To confirm the interaction of sVP40_{Δ43} N67C-CCA with TK142 and also identify the binding mode and orientation, structure determination of the complex is needed. This could be potentially facilitated by using irreversible reactions by incubating sVP40_{Δ43} with fragments in the absence of DTT. Nevertheless, DT was shown to be a promising screening technique to identify sVP40 ligands.

3.3.2 Lead discovery via crystal soaking

3.3.2.1 Introduction

The development of drugs against various diseases often requires the inhibition of the functions of intracellular proteins. This is achieved by generating small molecules that can easily cross the plasma membrane and bind to the target protein, thereby disabling its function. The identification of new inhibitors using soaking of crystals of the protein of interest with libraries of low molecular weight compounds (fragments, approx. 150 to 300 Da) is becoming increasingly popular in the FBLD process.

Results

Crystal soaking is a highly sensitive method and therefore suitable to detect even weak binders. In addition, crystal soaking provides structural information, e.g., how and where the fragment binds to the protein. Due to the low number of atoms and functional groups of employed fragments, they represent promising starting points for the development of lead structures, therefore requiring only a small library for screening. For crystal soaking experiments, preformed crystals of the protein of interest are soaked in a solution of single fragments or fragment cocktails and analyzed via X-ray diffraction. The presence of a binder can be detected by positive electron density and the molecule in question can be added to the structural model. The molecule then serves as a starting point for fragment extension, typically resulting in a potent binder after several synthesis steps. In spite of all these advantages, crystal soaking has limitations. First of all, the protein of interest must be able to crystallize easily exhibiting reasonably high diffraction in a highly reproducible fashion. In addition, the protein crystals must not be too sensitive to environmental factors such as pH changes. Other limiting steps in the process include synchrotron beam time, data collection and processing as well as interpretation, which represent considerable bottlenecks for the development process [61,71].

3.3.2.2 Results

Expression and Purification of recombinant VP40

Expression and purification of recombinant VP40 was performed as described in section 2.4.1 and 2.4.2.1, and, with modifications, as described by Bornholdt *et al.* [32]. Establishment and optimization of the protocols were performed at the Sapphire lab at the SCRIPPS Research Institute, La Jolla, California. In contrast to the published protocol, the purification steps following IMAC could be simplified to a single SEC step using a HiLoad 16/60 column which separated the different homooligomers efficiently. Expression and purification of zVP40 and sVP40 Δ 43 WT was performed according to section 2.4.1 and 2.4.2. Yields of both proteins were similar and zVP40 WT eluted as both octamers (56 ml) and dimers (70 ml), whereas sVP40 Δ 43 WT showed nearly exclusively dimers (76 ml), eluting later than zVP40 Δ 43 WT dimers due to the truncation of the first N-terminal 43 aa (Figure 41). Full-length sVP40 WT was also expressed and purified and showed a similar elution pattern as sVP40 Δ 43 WT but yields were significantly lower (data not shown). Truncated zVP40 (zVP40 Δ 43 WT) was also expressed, purified and used for crystallization attempts (see next section).

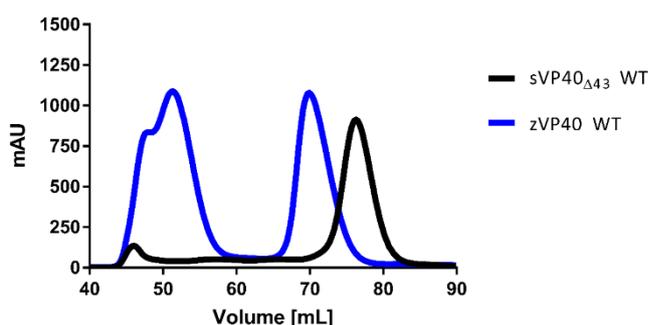


Figure 41: SEC profile of zVP40 full length WT and sVP40 Δ 43 WT. VP40 in pET46 was expressed in Rosetta2 cells, lysed using a microfluidizer and purified via IMAC and SEC. zVP40 eluted as octamers (51 ml) and dimers (70 ml) to near equal amounts, whereas sVP40 eluted mainly as dimers (76 ml).

Results

Peaks with an elution volume of approx. 51 ml could be identified as octameric VP40, while an elution volume of 70 and 76 ml represented dimeric zVP40 and sVP40, full-length and $\Delta 43$, respectively. As SEC experiments are not suitable to properly correlate specific homo-oligomerization states to certain elution volumes, in addition, SEC-MALS analyses were performed at the Sapphire lab (data not shown). This method uses analytical size-exclusion chromatography coupled to multi-angle light scattering (MALS) and can measure exact masses of proteins in solution and is thus used to clearly assign SEC elution volumes to different homo-oligomeric states [300]. In summary, the published expression and purification protocol could be successfully simplified and optimized.

Crystallization of VP40

First attempts to crystallize sVP40 $\Delta 43$ WT and zVP40 $\Delta 43$ WT were performed at the Sapphire lab, La Jolla. Further experiments were mainly performed with sVP40 $\Delta 43$ WT as this protein seemed to form crystals much more readily (Figure 42). Using the in-house x-ray beam mar345 at the Institute of Pharmaceutical Chemistry, Philipps-University Marburg, the ideal and most reproducible crystallization conditions could be determined and were as follows: 1 μ L of sVP40 $\Delta 43$ WT dimer (7 mg/ml) in 25 mM Tris, 300 mM NaCl, pH 8 was mixed with 1 μ L 100 mM HEPES, pH 7.2, 40 mM MgCl₂ and 10% PEG400 as a hanging drop over a 500 μ L reservoir of the same buffer. Crystallization plates were stored at 18 °C and crystals usually grew overnight. Figure 42 B shows typical pictures of dimeric sVP40 $\Delta 43$ WT crystals diffracting up to 1.5 Å in space group C2.

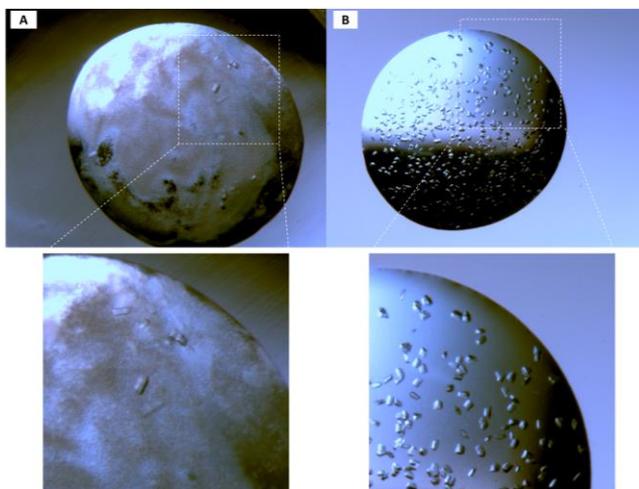


Figure 42: Crystals of zVP40 $\Delta 43$ and sVP40 $\Delta 43$ dimer. A) zVP40 $\Delta 43$ WT was crystallized in 0.5 μ L drops using various crystallization screen while B) sVP40 $\Delta 43$ WT was crystallized manually in drops of 2 μ L.

Optimization of crystal soaking conditions

As the library of fragments to be soaked with sVP40 $\Delta 43$ WT crystals were dissolved in DMSO, suitable DMSO concentrations and soaking times were tested. For this, pre-formed crystals were soaked in various concentrations of DMSO, diluted in crystallization buffer and afterwards soaked in crystallization cryo buffer, containing 10% ethylene glycol. Crystals were assessed visually and via diffraction patterns, as shown in Table 15.

Results

Table 15: DMSO soaking of dimeric sVP40_{Δ43} WT crystals. Crystals of dimeric sVP40_{Δ43} WT were soaked in varying concentration of DMSO for 12 h, flash-frozen in liquid nitrogen and analyzed via X-ray diffraction.

% DMSO	Soaking times	Resolution [Å]	% DMSO	Soaking times	Resolution [Å]
0%	2.5 h	3	5%	12 h	2
0%	12 h	3	10%	12 h	2.2
0%	12 h	1.7	10%	12 h	1.8
1%	2.5 h	1.7	15%	2.5 h	3.5
1%	12 h	2.2	15%	12 h	2.3
1%	12 h	2.4	20%	2.5 h	2.8
2%	2.5 h	2.9	20%	12 h	2.3
2%	12 h	2	20%	12 h	1.8
5%	12 h	1.8			

Diffractions varied greatly among crystals but soaking in 20% DMSO for 12 h did not significantly decrease crystal quality. Thus, it was concluded that soaking conditions include a final concentration of 10% DMSO with maximum soaking times of approx. 12 to 16 h. As the dissolved fragments were available as 1 M stocks in DMSO, a 1:10 dilution gave a final concentration of 100 mM, which was compatible with most published crystal soaking protocols [17,18,130,218,304]. Because ice rings were observed in most datasets, follow-up soakings were performed with 20% ethylene glycol instead of 10%.

Soaking of sVP40_{Δ43} WT crystals with fragment from the Jena library

Crystal soaking was used to identify low molecular weight binders and simultaneously gain structural information about their position and orientation. For this, the Jena FragXtal Screen was used, kindly provided by the Klebe lab (Institute of Pharmaceutical Chemistry, Philipps-University Marburg). This library consists of 96 fragments (J1 to J96, see section 8.13.2), mostly adhering to the Astex' rule of three (Ro3) [148,276]. This rule is based on Lipinski's rule of five (Ro5) [200], describing the chemical properties for oral availability of drug-like molecules or drug candidates. For Astex' Ro3, fragments obey limitations such as molecular weight ≤ 300 Da and number of hydrogen bond donors and acceptors ≤ 3 [64]. This library was used for several screenings [114,130,198,296].

All 96 fragments of the Jena FragXtal Screen were dissolved in DMSO and diluted 1:10 in crystallization (cryo) buffer, resulting in a final concentration of 100 mM. Crystals of dimeric sVP40_{Δ43} WT were soaked overnight in each fragment solution, afterwards soaked in crystallization cryo buffer (including the respective fragment), flash-frozen in liquid nitrogen and shipped to the beamline. If crystals dissolved, showed visible cracks or low to no diffraction, soaking was repeated with shorter soaking

Results

times (1 h, 5 min, 5 sec) until the obtained dataset met the quality standards. Most datasets had to be reprocessed manually using XDS. To identify binders, the PanDDA suite was applied to all 140 acquired datasets (including repeats) and diffraction maps with so-called events were manually evaluated. Crystals belonging to the ground state (without fragment soaking) were also included (AW107, AW108 and AW118). Crystals of a different space group (space group 1 instead of 5; AW151, AW180 and AW198) could not be included in the analysis and were refined manually. Afterwards, the automated refinement pipeline by Schiebel *et al.* [220,276] was applied to all 140 datasets with the kind help of Prof. Dr. Andreas Heine (Institute of Pharmaceutical Chemistry, Philipps-University Marburg). This pipeline performs 11 steps (see section 2.4.4.8) and for each dataset the refinement step with the lowest R_{free} was selected for manual evaluation. If the respective map revealed unexplained electron density, a manual run of MR and 1 to >10 refinement cycles were performed. This procedure is outlined in Figure 12 B. Table 16 shows the dataset name, soaking times and resolution as well as the results from the PanDDA analysis, the selected step of the autoprocessing pipeline as well as the manual refinement strategy.

Table 16: Overview of crystal soaking datasets and refinement strategy. For each fragment of the Jena FragXtal library, the crystal with its respective resolution, soaking times as well as the refinement strategy are listed. AW107, AW108 and AW118 are negative controls that were soaked with DMSO only.

Jena ID	Xtal ID	Soaking times	Resolution [Å]	PanDDA	Autoprocessing	Refinement
1	AW49	ON	1.73	–	11	1 cycle
1	AW50	ON	2.1	–	6	–
2	AW51	ON	2.4	–	5	–
3	AW52	ON	1.7	–	10	1 cycle
3	AW125	1 h	1.9	1 event	6	–
4	AW54	ON	1.8	–	5	1 cycle
4	AW181	ON	1.9	–	5	1 cycle
5	AW55	ON	1.9	–	10	–
6	AW57	ON	2.2	–	4	–
7	AW58	ON	2.2	–	5	–
7	AW192	ON	2.2	–	5	–
8	AW59	ON	2	–	11	–
9	AW60	ON	1.8	–	10	–
10	AW61	ON	1.9	–	5	–
10	AW158	1 h	1.9	–	5	–
10	AW159	1 h	1.65	–	10	5 cycles
11	AW126	1 h	1.9	–	5	–
11	AW182	ON	1.6	–	9	1 cycle
12	AW244	5 sec	2.2	–	5	1 cycle
13	AW63	ON	1.75	–	10	–
14	AW179	5 min	2.5	1 event	5	3 cycles
14	AW226	5 sec	2.3	–	5	–

Results

15	AW65	ON	1.8	1 event	–	> 10 cycles
15	AW161	1 h	1.75	1 event	5	4 cycles
15	AW162	1 h	1.7	1 event	9	4 cycles
15	AW163	1 h	1.9	1 event	5	1 cycle
16	AW183	ON	1.65	–	10	2 cycles
16	AW254	5 sec	1.85	–	11	–
17	AW67	ON	1.7	–	7	–
18	AW68	ON	1.7	–	10	–
19	AW128	1 h	2.4	–	6	–
20	AW69	ON	1.6	–	11	–
21	AW129	1 h	2.4	2 events	6	4 cycles
22	AW130	1 h	1.7	–	6	–
23	AW131	1 h	2	–	5	–
23	AW164	1 h	1.8	–	10	–
23	AW70	ON	2,5	–	6	–
24	AW132	1 h	1.95	–	5	1 cycle
25	AW72	ON	1,6	–	10	2 cycles
26	AW133	1 h	1.9	–	10	1 cycle
27	AW74	ON	1.9	–	5	–
28	AW134	1 h	1.65	–	10	1 cycle
28	AW165	1 h	1.8	–	10	–
29	AW135	1 h	1.7	–	11	1 cycle
29	AW166	1 h	1.9	–	5	–
30	AW167	1 h	1.65	–	5	1 cycle
30	AW77	ON	1.7	–	5	–
31	AW137	ON	1.9	–	5	–
31	AW184	ON	2.3	–	10	–
32	AW136	1 h	1.7	–	10	–
33	AW79	ON	1.7	–	10	–
34	AW185	ON	2	–	5	–
35	AW81	ON	2	–	5	–
36	AW82	ON	1.8	2 events	6	3 cycles
38	AW187	ON	1.85	–	10	–
39	AW85	ON	1.4	1 event	10	5 cycles
40	AW86	ON	1.75	–	7	–
41	AW87	ON	1.9	–	5	–
42	AW88	ON	2	–	6	1 cycle
43	AW89	ON	1.9	–	5	2 cycles
44	AW200	5 min	2	–	11	–
45	AW188	ON	2.1	3 events	10	7 cycles
46	AW93	ON	1.8	–	5	–
46	AW228	5 sec	3.2	–	8	–
47	AW94	ON	1.95	–	5	–
48	AW229	5 sec	2	–	5	–
49	AW169	1 h	1.75	–	10	–

Results

51	AW96	ON	1.85	–	5	2 cycles
52	AW138	ON	1.8	1 event	5	5 cycles
53	AW139	ON	1.85	–	10	–
54	AW140	ON	2.1	–	5	–
54	AW213	5 sec	1.85	–	10	–
54	AW257	5 sec	1.61	1 event	5	–
55	AW214	5 sec	2	–	9	–
55	AW258	ON	1.7	–	11	1 cycle
55	AW259	ON	1.6	–	10	1 cycle
56	AW141	ON	1.7	1 event	–	4 cycles
56	AW215	5 sec	1.65	–	10	–
57	AW142	ON	1.75	1 event	11	10 cycles
57	AW216	5 sec	1.85	–	11	2 cycles
57	AW261	ON	2.1	–	6	–
58	AW143	ON	1.85	–	11	–
58	AW230	5 sec	2.2	X	7	–
59	AW190	ON	1.9	–	10	–
60	AW232	5 sec	3	–	5	–
61	AW145	ON	1.9	–	10	–
61	AW234	5 sec	1.75	1 event	10	1 cycle
61	AW262	ON	2.65	–	7	–
62	AW146	ON	2.2	–	6	–
63	AW147	ON	2.1	–	5	–
64	AW148	ON	1.75	–	10	–
65	AW235	5 sec	1.85	–	10	–
66	AW150	ON	2	–	5	–
67	AW151	ON	1.9	X	5	1 cycle
68	AW191	ON	1.8	–	10	–
69	AW97	ON	1.9	–	6	–
69	AW98	ON	2.2	–	6	–
70	AW99	ON	1.8	–	5	2 cycles
70	AW194	ON	1.95	–	5	–
71	AW173	1 h	1.7	1 event	11	5 cycles
72	AW174	1 h	1.9	–	10	–
73	AW248	5 sec	1.65	–	11	–
74	AW102	ON	2.2	–	5	–
75	AW220	5 sec	1.9	–	5	1 cycle
75	AW249	5 sec	2.8	–	8	–
77	AW193	ON	1.95	–	10	2 cycles
78	AW175	1 h	2.3	1 event	9	2 cycles
79	AW250	5 sec	1.85	–	10	–
80	AW105	ON	1.9	–	10	2 cycles
80	AW180	ON	2.65	X	5	–
81	AW154	ON	1.7	–	9	2 cycles

Results

82	AW155	ON	1.9	–	5	–
83	AW156	ON	1.9	–	5	–
84	AW119	1 h	1.6	2 events	10	9 cycles
84	AW120	1 h	1.75	1 event	6	6 cycles
84	AW121	1 h	2.5	3 events	5	8 cycles
85	AW122	ON	1.5	1 event	10	> 10 cycles
86	AW106	ON	1.85	–	11	1 cycle
86	AW195	ON	1.5	2 events	5	1 cycle
87	AW109	ON	1.65	–	6	1 cycle
88	AW176	1 h	2.55	–	5	–
89	AW123	ON	1.9	1 event	6	8 cycles
90	AW111	ON	1.65	–	9	3 cycles
90	AW112	ON	2.2	–	6	–
90	AW113	ON	2.1	1 event	11	1 cycle
90	AW196	ON	1.75	–	11	2 cycles
91	AW198	ON	2.9	X	7	–
92	AW221	5 sec	2.3	–	7	–
92	AW222	5 sec	1.85	–	5	–
92	AW251	5 sec	1.75	–	10	–
93	AW115	ON	2.1	–	5	–
93	AW223	5 sec	1.65	–	5	–
93	AW252	5 sec	2.2	–	7	–
94	AW224	5 sec	1.8	–	10	–
95	AW117	ON	1.8	–	X	–
95	AW186	ON	2	–	5	–
96	AW124	ON	1.6	1 event	11	5 cycles
–	AW107	ON	2.1	1 event	11	–
–	AW108	ON	1.7	–	10	–
–	AW118	ON	1.8	–	10	–

For fragments #37, #50 and #76 no dataset with a diffraction $\leq 3.5 \text{ \AA}$ could be obtained and these molecules were therefore considered disruptive. AW151 (fragment #67), AW180 (fragment #80), AW198 (fragment #94) and AW230 (fragment #58) belong to space group 1 and were therefore excluded from the PanDDA run (indicated by an X). For fragments #80 and #58 other datasets (AW105 and AW143) in space group 5 were available, suggesting an error during data processing or crystal damages. Crystal quality of AW117 (fragment #95) did not meet the set standards but was included in the analysis by mistake and could not be processed using the autorefine pipeline. Soaking with fragment #95 was repeated with AW224. AW122 was already described in section 3.1.2.

It has to be noted that the quality of the crystals could be classified into two categories, 1) Molecular Replacement using 4LD8 explained the model fairly well, with no major positive electron density, mostly accounted for by water molecules and 2) a consistently higher quality of the crystals as the

Results

electron density maps allowed for the addition of further, previously unresolved amino acids to the VP40 models. Usually all or just a subset of residues 196-PSNLSGA-202 forming the linker between N- and C-terminal domain and amino acids 309-PDYDDCHSPASCSYL-323 forming the very end of the C-terminus could be assigned to the structure. This process improved the statistics vastly, leaving no undefined blobs.

Salicylic acid (SA, fragment #15; AW65) was identified as a crystallographic binder due to the clearly visible electron density after phasing and subsequent refinement cycles (Table 17). This fragment was mixed with N, N-diethylethanimidamide in a 1:1 ratio but the second molecule could not be identified. Soaking with SA was repeated three times more with reproducible detection of the molecule in the same location and with the same orientation (AW161, AW162 and AW163). SA forms two H-bonds, one to the backbone of L158 (NTD) as well as to the backbone of R214 (CTD) (Figure 43). Part of the side chain of R214 did not have clear electron density and was deleted in the resulting pdb-model. The aromatic ring of SA points further into the pocket.

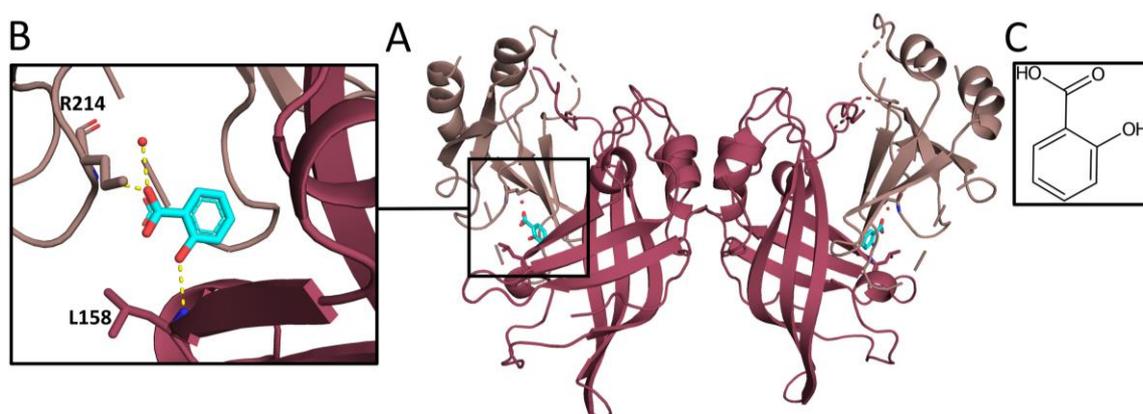


Figure 43: Structure of dimeric sVP40 Δ 43 WT soaked with SA. Overall (A) and close-up (B) cartoon representation of dimeric sVP40 Δ 43 WT in complex with bound SA in the pocket between the NTD and CTD with H-bonds to L158 and R214. C) Chemical structure of SA.

To confirm the binding location and orientation, SA was also mixed with sVP40 Δ 43 WT prior to the setup of crystallization drops. Figure 72 (Supplemental Material) shows the SA-sVP40 Δ 43 WT co-crystallization complex with the same binding mode as observed for the soaking experiments (Table 17).

Table 17: Data collection and refinement statistics of dimeric sVP40 Δ 43 WT in complex with SA. Values for the outer shell are given in parentheses.

Data collection	sVP40 Δ 43 WT + SA soaking (AW65)	sVP40 Δ 43 WT + SA co-crystallization (AW568)
PDB ID	8B2U	8B1S
Space group	5	5
Unit cell parameters [Å] (a, b, c,	62.30; 90.62; 48.10	62.10; 90.62; 48.19

Results

α, β, γ)	90.00; 94.02; 90.00	90.00; 94.25; 90.00
Resolution range	47.99 – 1.80	48.06 – 1.6
Total number of reflections	81,399 (4,329)	233,261 (9,242)
Number of unique reflections	24,102 (1,315)	33,883 (1,474)
I/ σ	6.4 (0.7)	21.8 (3.8)
Completeness [%]	97.9 (90.4)	97.7 (86.7)
Multiplicity	3.4 (3.3)	6.9 (6.3)
CC _{1/2} [%]	99.5 (42.4)	99.9 (97.2)
Refinement		
Resolution range	47.99 – 1.80	36.14 – 1.60
Rwork/Rfree [%]	20.81/23.02	18.89/20.70
Number of unique reflections	23,796	33,841
Number of protein residues	248	244
Average B factor [\AA^2]	43.51	36.95
R.m.s deviations:		
- Bond length [\AA]	0.007	0.004
- Bond angles [$^\circ$]	0.91	0.73
Ramachandran plot:		
- Favoured [%]	97.88	98.26
- Allowed [%]	2.12	1.74
- Outliers [%]	0.00	0.00
Rotamer outliers [%]	0.00	0.00
Clash score	2.15	3.19
Number of TLS groups	5	5

Salicylic acid binds to sVP40 Δ 43 WT in solution

While small molecules usually exhibit a low affinity towards the target protein, they are not easily detected and need a sensitive method such as crystal soaking. SA was also one of the smallest molecules of this library, further strengthening the fact that it would have been missed by most other methods. Nevertheless, attempts were made to confirm its binding to sVP40 Δ 43 WT not just in the crystalline state but also in solution. Hence, a thermal shift assay using different concentrations of sodium SA (NaSA) was performed. To confirm that NaSA binds as well to sVP40 Δ 43 WT, the fragment was also soaked into VP40 crystals using the same protocol as for the screening. Again, the molecule could be detected directly after performing MR (data not shown). As shown in Figure 44, incubation

Results

with NaSA led to a decreased denaturing temperature of dimeric sVP40_{Δ43} WT, confirming its binding to VP40 also in solution but also pointing towards a potential destabilizing effect.

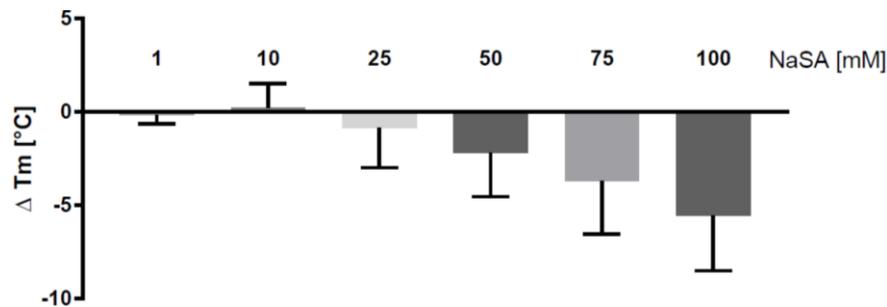


Figure 44: Thermal stability of dimeric sVP40_{Δ43} WT with NaSA. Means of denaturing temperatures upon incubation with increasing concentrations (1 to 100 mM) of NaSA from three independent experiments were compared to sVP40_{Δ43} WT.

To further characterise SA as a ligand in solution and identify further residues affected by its binding, hydrogen-deuterium exchange mass spectrometry (HDX-MS) was performed (Figure 45). Peptide coverage was 94.6% with a total of 137 peptides and 6.26 redundancy and only a few sequence stretches not covered (Table 22 and Figure 73, supplemental material).

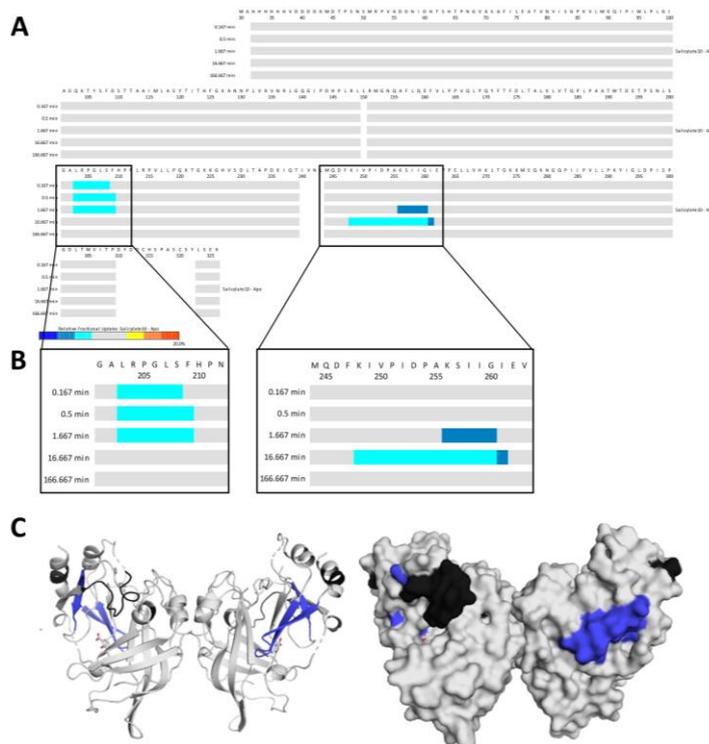


Figure 45: Differences in D₂O uptake between dimeric sVP40_{Δ43} WT in its apo form and in complex with NaSA. 25 μM protein were incubated with D₂O-containing buffer. After 10, 30, 100, 1,000 or 10,000 s, each reaction was quenched, digested with pepsin and the peptides were ionized by ESI and analyzed via mass spectrometry. A and B) Differential deuterium uptake is observed for peptides 202 to 208 and 248 to 260. C) Peptides with significantly decreased H/D uptake were mapped onto the WT crystal structure (cartoon representation right and surface left). Peptides with decreased exchange rates are shown in blue and residues shown in black were not covered.

The relative HDX of VP40 in the apo form and in combination with NaSA is shown in Figure 45 A. The results from this comparison are further shown in Figure 45 B, emphasizing the difference in D₂O-incorporation of NaSA-bound sVP40_{Δ43} WT versus apo- sVP40_{Δ43} WT. Sequence stretches 203 to 209 and 248 to 261 showed decreased HDX (> 5% less HDX) upon incubation of the protein with NaSA, the respective areas are colored in blue on the surface of the sVP40_{Δ43} WT structure (Figure 45 C). Interestingly, the pocket with interaction residues L158 and R214 did not exhibit changed HDX.

Results

Salicylic acid is not able to inhibit sVP40

To investigate whether SA was able to inhibit sVP40's functions, we next performed functional assays in cell culture systems. As SA is not soluble in cell culture medium, we used NaSA as a more hydrophilic alternative for further experiments. For octamerization, the C-terminal domain of VP40 has to flip up to allow the N-terminal domains of two monomers to interact with each other [32]. As SA forms hydrogen bonds to both termini of sVP40, it is possible that this interaction stabilized the protein in its dimeric form, thus leading to higher levels of reporter gene activity using a Zaire-based MG assay that monitors viral RNA synthesis activity. Cells were transfected with the MG components and treated with increasing concentrations of NaSA and the reporter gene activity was measured 48 h pt. Interestingly, an even lower reporter gene activity could be observed for the treatment of 1 mM NaSA and slightly higher levels for 10 μ M compared to the untreated sample with sVP40 (Figure 46 A), suggesting a dose-dependent influence on VP40's effect on viral genome replication and transcription. While treatment with NaSA seemed to be well tolerated by the cells, the lower levels of reporter gene activity might potentially be due to a decrease in cell culture medium pH. It is possible that this effect overlaps with a potential inhibition of sVP40 upon treatment with 1 mM or 100 μ M NaSA.

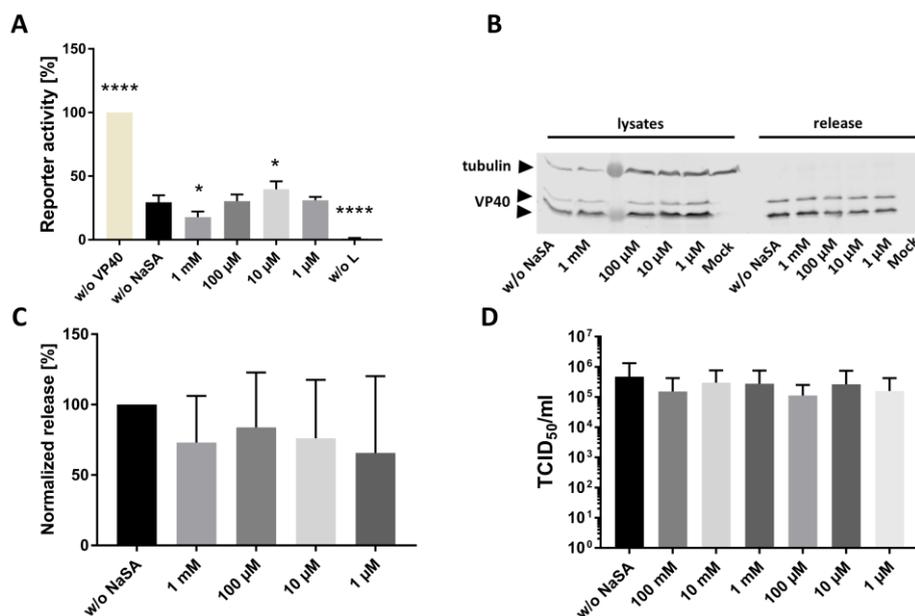


Figure 46: NaSA displays only minor effects on sVP40 in functional cell culture assays. A) Treatment with NaSA influences reporter gene activity in a MG assay. HEK293 cells were transfected with pCAGGS sVP40 WT and the MG assay components. Cell culture medium was changed 4 h pt and NaSA added. Cells were lysed 48 h pt, reporter gene activity was measured and normalized to the sample without sVP40 (set to 100%). For the negative control, no polymerase L was added to the transfection mix. B and C) Effect of NaSA-treatment on VP40-release. B) Cells were transfected with pCAGGS zGP and sVP40 WT. Cell culture medium was changed 4 h pt and NaSA added. Supernatants were collected and cells lysed 24 h pt and WB analyses of cellular and released sVP40 were performed using rabbit α -sVP40 and mouse α -tubulin primary and donkey α -mouse and goat α -rabbit IRDye® 680 as secondary antibodies. C) Quantification of VLPs. Released sVP40 was normalized to the amount of sVP40 in lysates and sVP40 WT w/o NaSA set to 100%. D) Effect of NaSA on EBOV titers. HuH7 cells were infected with EBOV (MOI 0.1) and afterwards treated with NaSA in varying concentrations. Supernatants were taken at 0 dpi and 2 dpi and viral titers assessed via TCID₅₀. Bars indicate titers at 2 dpi normalized to 0 dpi. Bars indicate the mean \pm SD of at least three independent experiments, asterisks indicate statistical significance (One-way ANOVA) as follows: *P<0.05 and ****P<0.0001 compared to the positive control without NaSA.

Results

Next, the effect of NaSA on the release of sVP40 was investigated. Upon ectopic intracellular expression, VP40 is able to form virus-like particles which is strengthened by co-expression of GP [245]. However, treatment with NaSA revealed no effect on the budding activity of sVP40 and the release of virus-like particles into the supernatant (Figure 46 B and C). Because these assays depict only individual steps of the viral replication cycle, it is possible that a potential effect of SA/NaSA could only be detected in the presence of all viral proteins. Therefore, testing of SA was performed in infection studies under BSL4 conditions. For this, VeroE6 cells were infected with Zaire EBOV and treated with increasing concentrations of NaSA. Viral titers in the supernatant were assessed at 2 dpi. As in the VLP-assay, no statistically significant effects could be observed (Figure 46 C). Taken together, NaSA seemed to exert no inhibition on release of virus(-like) particles and showed only minor effects on viral RNA synthesis. However, in particular due to the small size of this fragment and a probably low affinity binding, no major implications of SA treatment were expected.

Characterisation of residues involved in the interaction with sVP40

As a next step, it was decided to investigate the druggability of the pocket binding site of SA and further characterize the residues that are involved in the interaction with SA. For zVP40 it was reported that mutagenesis of residues K212, L213 and R214 diminishes the amount of zVP40 released into the supernatant [220]. However, none of the residues of interest were characterized for sVP40. Thus, the mutants sVP40 L158A, L213A and R214A (henceforth called pocket mutants) were generated in the mammalian expression vector pCAGGS-sVP40 and MG as well as VLP assays were performed. Mutation of sVP40 R214 to alanine resulted in a clear abolishment of sVP40's inhibitory function on viral genome replication and transcription (Figure 47). In addition, a slight but not significant effect could be observed for sVP40 L158A.

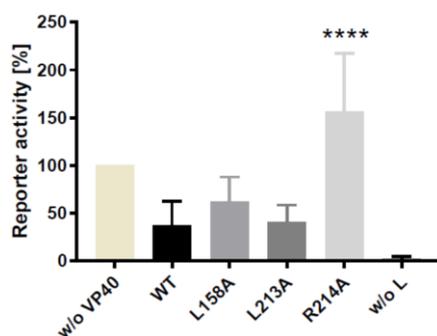


Figure 47: Influence of sVP40 pocket mutants on reporter gene activity of a MG assay. HEK293 cells were transfected with pCAGGS sVP40 WT or mutants along with the MG assay components. Cells were lysed 48 h pt, reporter gene activity was measured and normalized to the sample without VP40 (set to 100%). For the negative control, no polymerase L was added to the transfection mix. Bars indicate the mean \pm SD of at least three independent experiments, asterisks indicate statistical significance (One-way ANOVA) as follows: *** $P < 0.0005$ and **** $P < 0.0001$ compared to sVP40 WT.

To investigate the influence of the mutants of L158A, L213A and R214A on the release of VP40-induced VLPs, a budding assay was performed, which showed a significant decrease in released VP40 for all three pocket mutants with the strongest effect observed for sVP40 L213A (Figure 48).

Results

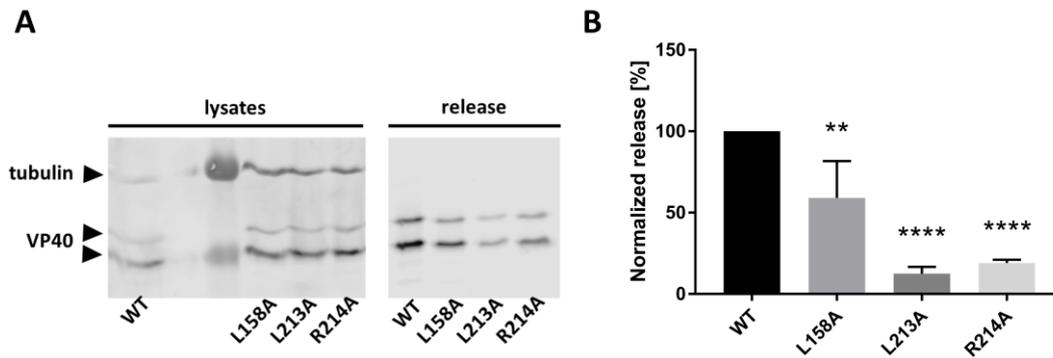
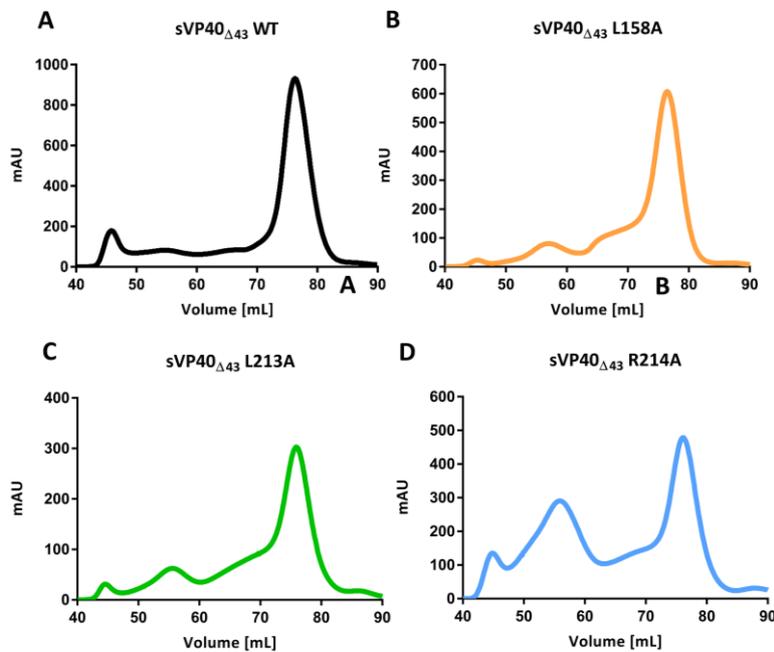


Figure 48: Effect of sVP40 pocket mutants on sVP40-release. A) Cells were transfected with pCAGGS zGP and sVP40 WT or mutants. Supernatants were collected and cells lysed 24 h pt and WB analyses of cellular and released VP40 were performed using rabbit α -sVP40 and mouse α -tubulin primary and donkey α -mouse and goat α -rabbit IRDye[®] 680 as secondary antibodies. B) Quantification of VLPs. Released sVP40 was normalized to the amount of sVP40 in lysates and sVP40 WT set to 100%. Bars indicate the mean \pm SD of three independent experiments, asterisks indicate statistical significance (One-way ANOVA) as follows: ** $P < 0.005$ and **** $P < 0.0001$.

Next, a potential effect of the pocket mutations on the homo-oligomerization ability of sVP40 was investigated. Therefore, the pocket mutations were introduced into the pET46 sVP40 $_{\Delta 43}$ plasmid via site-directed mutagenesis. These constructs were then transformed into Rosetta2 cells and the recombinant proteins sVP40 $_{\Delta 43}$ L158A, L213A and R214A were expressed and purified according to the standard protocols (see sections 2.4.1 and 2.4.2). After IMAC, suitable fractions were pooled, concentrated and further purified via SEC. As shown in Figure 49, all three pocket mutants formed dimers and octamers in varying ratios.



Usually, sVP40 $_{\Delta 43}$ WT forms predominantly dimers with only a small octamer fraction but the percentage of the octamer fraction was increased for all three mutants suggesting a potential effect

Results

on dimer stability upon mutation of L158, L213 or R214. While residue L158 has never been studied before, L213 and R214 of zVP40 were investigated by McCarthy *et al.* [220] who observed similar findings regarding the formation of VLPs. In addition, they observed a decreased plasma membrane association of the mutants in immune fluorescence analyses most strongly pronounced for the zVP40 L213 mutant. The zVP40 L213 mutant was also very prominent regarding VP40's oligomerization behavior with an increase in higher-order oligomers. The latter is again in line with the present findings for sVP40 L213A and R214A.

Taken together, these results indicate that targeting these residues offers a very promising approach for tackling several steps in the viral replication cycle. As (Na)SA interacts with both sVP40 L158 and R214, it seems likely that these molecules could principally also inhibit or reduce sVP40's release and impair its ability to downregulate viral genome replication and transcription, both processes that are considered crucial in the viral replication cycle.

It can be concluded that crystal soaking is a laborious but powerful screening technique and well worth the effort. While the identified binders exhibit usually a low affinity towards the target protein, these fragment hits would probably be missed by using conventional screening methods. In addition, the availability of structural information is of great advantage regarding future fragment extension studies. The present results suggest that the pocket occupied by SA is highly druggable. However, to actually block the residues and the pocket, the affinity of the compound needed to be greatly increased. This called for fragment extension to create a larger molecule with SA as a scaffold which forms more interactions to VP40, thus strengthening its binding. The next sections therefore address fragment extension, including the generation of a virtual databank of SA-derivatives, molecular docking to VP40, their synthesis and further soaking experiments.

Design of a databank with salicylic acid as a scaffold

The previous chapter showed that while SA was identified as a crystallographic binder, its affinity towards sVP40 was too low to have an impact on the protein's homo-oligomerization aptitude or functions such as the downregulation of viral genome replication and transcription as well as budding. Therefore, this section describes the structure-guided hit-to-lead development of SA to create a compound with improved affinity, which was to be achieved via fragment extension. This is a process of enlarging a molecule of interest by applying medicinal chemistry techniques so it eventually forms more interactions to the target molecule. While all sorts of interactions are advantageous, H-bonds and salt bridges are favourable compared to van-der-Waals forces. The work of the current chapter was performed in collaboration with the Kolb and the Diederich lab, both groups belonging to the Institute of Pharmaceutical Chemistry, Philipps-University Marburg.

Results

In the case of SA, its location in a pocket between the NTD and CTD of VP40 allowed the hypothesis that fragment's extension further into the pocket would stabilize the molecule in its position and potentially improve the affinity towards sVP40. An *in silico* strategy was selected which included the generation of a databank of molecules with SA as a scaffold, the preparation of 3D conformations, the docking to VP40 and eventually the inspection of docking poses to identify promising fragments.

The generation of the databank was performed using the online tool PINGUI, designed by the Kolb lab. Commercially available SA-derivatives were chosen as templates:

- Chlorine-based derivatives: 4-chloro-SA, 5-chloro-SA and 5-(chlorosulfonyl) SA
- Iodine-based derivatives: 4-iodo-SA, 5-iodo-SA and 3,5-diiodo-SA
- Bromine-based derivatives: 3,5-dibromo-SA and 5-bromo-SA
- Amino-based derivatives: 4-amino-SA and 3-amino-SA
- Hydroxy-based derivatives: 2,3,4-trihydroxybenzoic acid

All suitable reactions along with the appropriate reaction partners (acids, alcohols, amines, boronic acids, carbonyls and halides) were applied to the above-mentioned templates. Possible types of reactions include Williamson_ether, Suzuki, Negishi, Mitsunobu_phenol, Heck_terminal_vinyl, Heck_non-terminal_vinyl, Stille, Grignard_carbonyl, Grignard_alcohol, Schotten-Baumann_amide, Buchwald-Hartwig, urea and thiourea. Creating molecules by fusing the SA-derivatives to another reaction partner by the different reactions rendered 834,929 fragments. However, this number included redundant molecules as a result of substitution reactions, for example. Hence, the databank was purged with the kind help of Kai Schmidt, Master student from the Kolb lab, which resulted in a purged databank consisting of 542,545 molecules.

Docking of salicylic acid derivatives

As a first step, the various conformations the different molecules could adopt, had to be generated, which was performed with the kind help of Maria Giovanna Papadopoulos (Kolb lab). In addition, sVP40 as a receptor molecule had to be prepared. A large part of the sidechain of sVP40 R214 could not be assigned confidently to the electron density map of the crystal dataset, which necessitated the deletion in the structural model. As this sidechain could be crucial for successful docking, modelling of four different conformations of this sidechain based on the Coot tools "auto-fit rotamer" and "real-space refine zone" as well as MOE (molecular operating environment) tool was carried out. This process resulted in four different receptor proteins for the following docking analysis (Figure 50).

Results

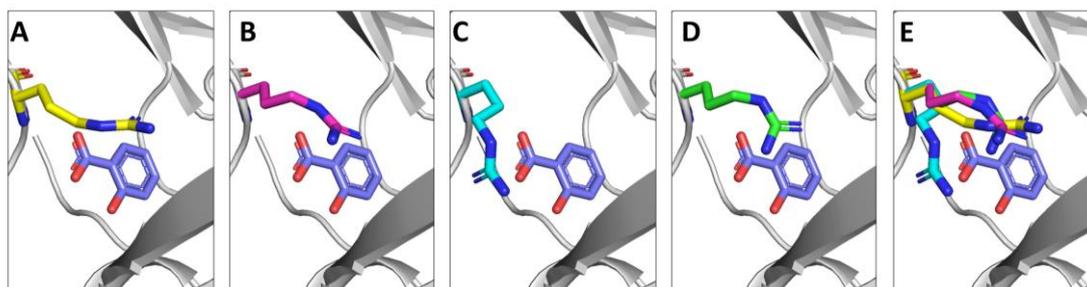


Figure 50: sVP40 R214 sidechain conformations. A) Intermediate, B) MOE-calculated, C) “up” or D) “down” orientation of the R214 sidechain used for docking and E) alignment of all conformations.

For this, the OpenEye software Hybrid was chosen. This suite allows the superposition of the docked fragments to the initial fragment hit. Program settings involved the max. generation of 50 conformers for each compound considering stereoisomers and number of possible docking poses set to 1. The thousand best docking poses of each of the four separate analyses were exported and visually inspected using the UCSF Chimera program according to different criteria such as occupation of the pocket, number of formed interactions between the fragment and sVP40, the introduction of clashes and the molecular weight of the fragment. Again, the above-mentioned criteria were employed with the additional aspect of interaction with the sVP40 R214 sidechain and superposition with the original fragment. Of the 4,000 molecules, 31 were chosen for further evaluation considering compound availability and synthesis feasibility (Table 24, supplemental material).

Soaking of Salicylic acid derivatives

To test the selected molecules for their ability to bind to sVP40_{Δ43} WT, the highly robust crystal soaking protocol was performed again. Synthesis of these compounds was kindly performed by Lennart Laube and Michael Daude, Diederich lab, Institute of Pharmaceutical Chemistry, Philipps-University Marburg. Of the chosen 31 molecules (Table 24, supplemental material), originating from the docking analysis and manual evaluation, building blocks needed for synthesis for compounds 1_030, 1_855 and 3_061 were not commercially available, 4_369 only in very small quantities. Product 2_017 was the only one commercially available and compounds 3_295 and 3_550 were considered most promising for initial synthesis due to the reaction (Schotten-Baumann_amides). Hence, the latter three were the prime candidates for initial testing (Figure 51).

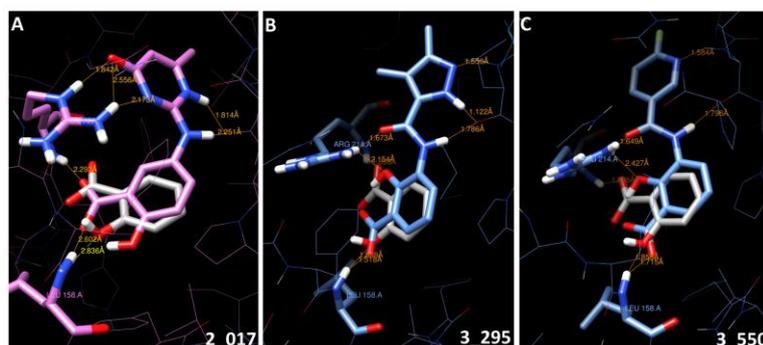
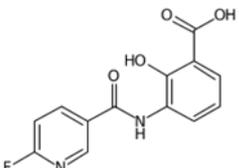
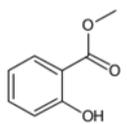
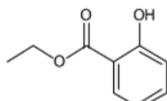
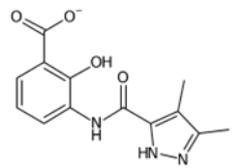
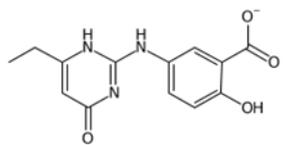
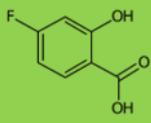
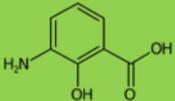
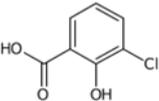
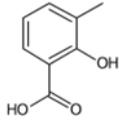
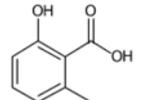
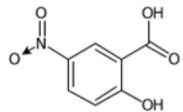
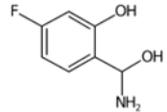
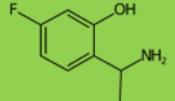
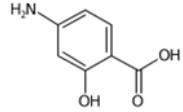
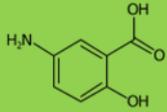


Figure 51: Docking poses of three exemplary molecules. Depending on the sidechain conformation sVP40 is coloured in either purple (2) or blue (3) and shown as lines with the exception of R214 and L158 (shown as sticks). SA is included in the model to visualize the superposition of the original with the extended fragment: compound 2_017 (A), 3_295 (B) and 3_550 (C). Graph generated using UCSF Chimera.

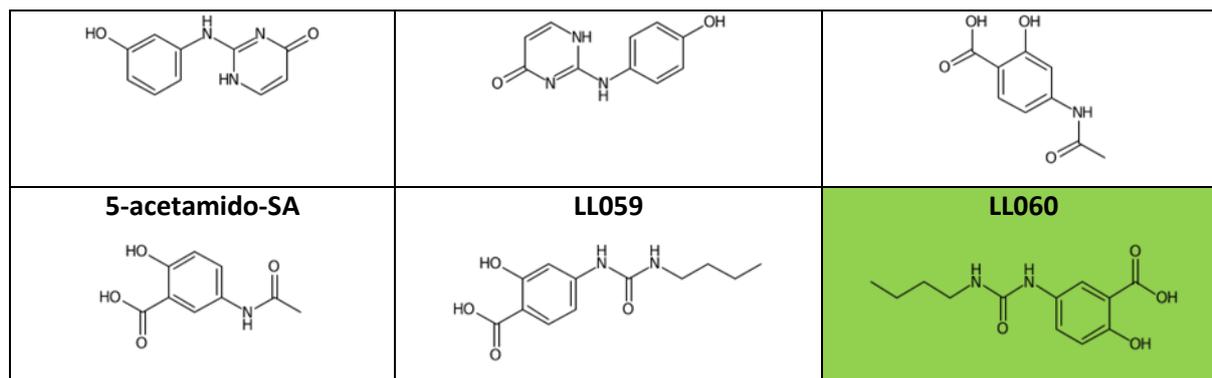
Results

In parallel, alternative SA-derivatives – both commercially purchased and synthesized by the Diederich lab – were tested via crystal soaking to answer the following questions: **1)** Are the hydroxyl and/or the carboxyl group crucial for the binding of SA to sVP40? **2)** Can the SA-derivatives adopt a different conformation or position than the original fragment? **3)** Which modifications at which C-atom of the aromatic ring are tolerated by sVP40? As during the initial soaking screening, compounds were dissolved in DMSO and diluted in crystallization cryo buffer to 100 mM. Fragments were considered disruptive when diffraction patterns of 3.5 Å or above were reached even when soaking times were reduced to only five seconds. Table 18 summarizes the tested molecules in chronological order of testing.

Table 18: Crystal soaking with SA-derivatives. Compound ID or IUPAC name as well as the structure are given for each synthesized molecule. Successfully soaked molecules are highlighted in green.

<p>3_550</p> 	<p>methyl-salicylate</p> 	<p>ethyl-salicylate</p> 
<p>3_295</p> 	<p>2_017</p> 	<p>4-fluoro-SA</p> 
<p>3-amino-SA</p> 	<p>3-chloro-SA</p> 	<p>3-methyl-SA</p> 
<p>6-methyl-SA</p> 	<p>5-nitro-SA</p> 	<p>4-fluoro-2-hydroxymethylbenzamide</p> 
<p>4-fluoro-2-hydroxybenzamide</p> 	<p>4-amino-SA</p> 	<p>5-amino-SA</p> 
<p>2-(3-hydroxyanilino)-1H-pyrimidin-6-one</p>	<p>2-(4-hydroxyanilino)-1H-pyrimidin-6-one</p>	<p>4-acetamido-SA</p>

Results



Location and binding mode of the successfully soaked fragments are shown in Figure 52 and parameters of the crystal data are given in Table 19.

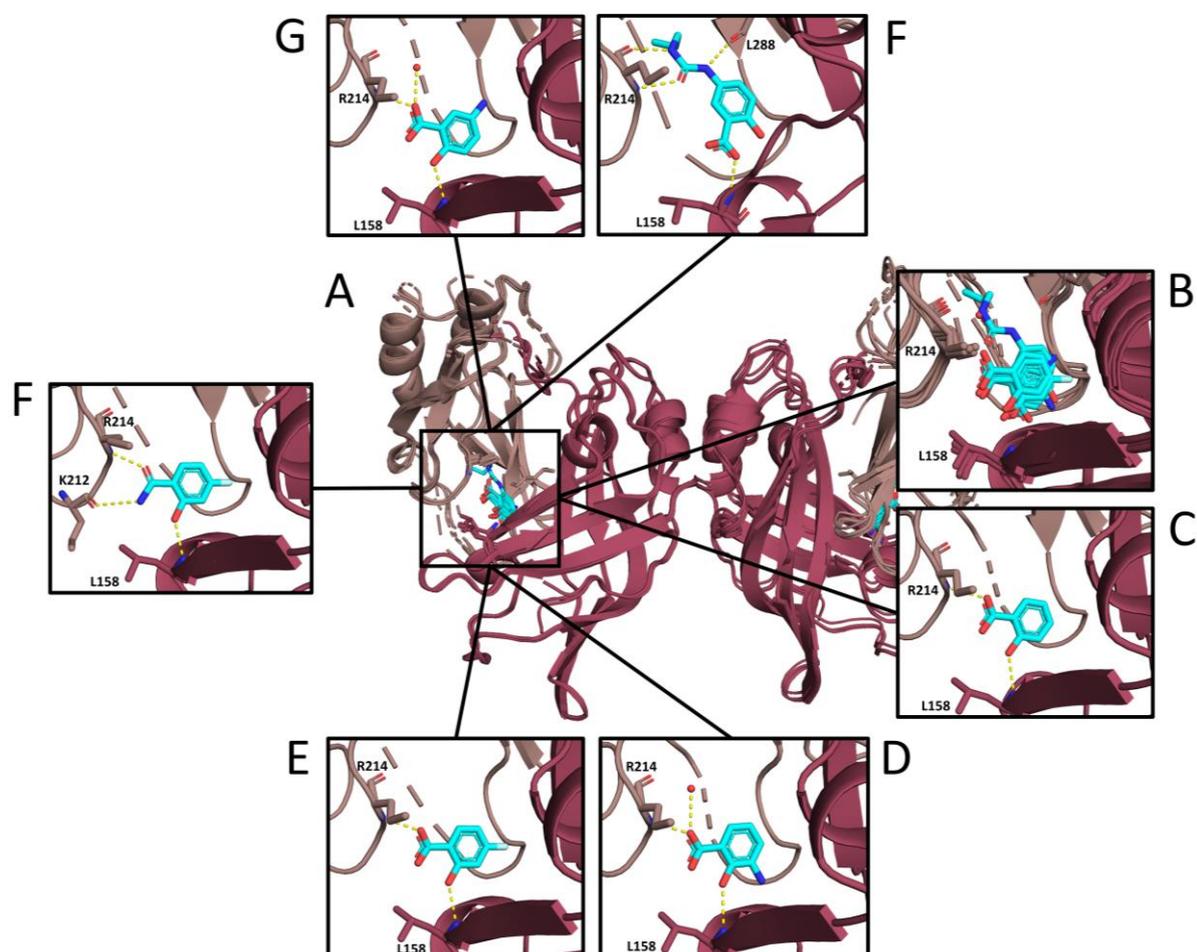


Figure 52: Structure of dimeric sVP40 Δ 43 WT with soaked SA-derivatives. Overall (A) and close-up (B) of all superimposed fragments bound to dimeric sVP40 Δ 43 WT. C) SA, D) 3-amino-SA, E) 4-fluoro-SA, F) 4-fluoro-2-hydroxybenzamide, G) 5-amino-SA and H) LL060.

Results

Table 19: Data collection and refinement statistics of dimeric sVP40_{Δ43} WT in complex with SA-derivatives. AW529 = 3-amino-SA, AW526 = 4-fluoro-SA, AW549 = 4-fluoro-2-hydroxybenzamide, AW588 = 5-amino-SA and AW633 = LL060. Values for the outer shell are given in parenthesis. *Deposition into the PDB not yet started, statistics are given for the last refinement.

Data collection	AW529	AW526	AW549	AW588	AW633*
PDB ID	8B13	8B1Q	8AZQ	8B1M	-
Space group	5	5	5	5	5
Unit cell parameters [Å] (a, b, c; α, β, γ)	62.78; 90.72;	62.99; 90.84;	63.06; 90.77;	62.60; 90.80;	62.79; 91.17;
Resolution range	48.35; 90.00;	48.15; 90.00;	48.17; 90.00;	48.28; 90.00;	48.31; 90.00;
	94.35; 90.00	94.98; 90.00	95.08; 90.00	94.47; 90.00	94.64; 90.00
Total number of reflections	48.21 – 1.91	47.96 – 1.85	47.98 – 1.90	48.14 – 1.60	48.15 – 1.96
	73,265	86,225	142,897	236,011	132,996
	(4,222)	(5,128)	(8,648)	(11,263)	(8,608)
Number of unique reflections	20,775	22,616	21,041	34,606	19,346
	(1,355)	(1,376)	(1,335)	(1,669)	(1,277)
I/σ	15.4 (8.4)	10.1 (1.6)	13.8 (5.8)	14.9 (4.2)	13.4 (5.1)
Completeness [%]	98.4 (96.2)	98.7 (97.9)	99.1 (98.2)	98.6 (96.4)	98.7 (92.6)
Multiplicity	3.5 (3.1)	3.8 (3.7)	6.8 (6.5)	6.8 (6.7)	6.9 (6.7)
CC _{1/2} [%]	99.8 (98.4)	99.8 (72.5)	99.3 (98.1)	99.5 (90.5)	99.7 (96.5)
Refinement					
Resolution range	45.38 – 1.91	45.42 – 1.85	47.98 – 1.90	48.14 – 1.61	45.58 – 1.96
Rwork/Rfree [%]	16.76/20.28	19.13/21.72	17.95/20.87	18.51/20.39	17.51/20.24
Number of unique reflections	20,651	22,586	21,023	34,301	19,285
Number of protein residues	242	233	234	240	243
Average B factor [Å ²]	36.07	48.55	38.56	34.13	39.83
R.m.s deviations:					
- Bond length [Å ²]	0.009	0.009	0.008	0.009	0.013
- Bond angles [°]	1.05	1.06	0.90	1.04	1.15
Ramachandran plot:					
- Favoured [%]	98.26	98.65	97.75	98.68	98.27
- Allowed [%]	1.74	1.06	2.25	1.32	1.73
- Outliers [%]	0.000	0.00	0.00	0.00	0.00

Results

Rotamer outliers [%]	0.00	0.00	0.00	0.48	0.00
Clash score	1.86	2.25	2.48	2.12	2.16
Number of TLS groups	6	4	6	3	5

Ethyl- and methyl-SA as well as 4-fluoro-2-hydroxymethylbenzamide could not bind, suggesting that the carboxyl group is needed for interaction with sVP40 R214. However, amide groups are also accepted, as seen for 4-fluoro-2-hydroxybenzamide. For compound 3_550 and compound 3_295, the docking analysis suggested a promising pose with switched carboxyl and hydroxyl groups at C1 and C2, respectively. However, this interpretation was ruled out after 3-amino- and 4-fluoro-SA were soaked successfully. The fact that compound 2_017 did not bind to sVP40 Δ 43 WT in the crystalline state could potentially be explained by the R214 sidechain that was not visible in any of the crystal structures but was modelled in four different orientations for the docking analysis. For compound 2_017, the sidechain orientation is shown in Figure 51 A. It was concluded that this particular sidechain simulation might not be a correct representation of the structural model. 3-methyl-SA did bind, whereas 3-chloro-SA did not, indicating that extension at C3 was possible but sVP40 Δ 43 WT does not tolerate every modification. Other smaller derivatives include 5-nitro-SA, 6-methyl-SA, 4-amino-SA and 5-amino-SA but only the latter (5-amino-SA) could be detected in the electron density.

In most cases, SA-derivatives with just small modifications allowed straightforward interpretations about the extension ability of the C-atom in questions or the importance of specific functional groups. The results can be summarized as follows: The orientation and position of SA and its derivatives were consistent among all soaked structures and both functional groups were crucial. However, the carboxyl group could be switched for an amide group but the interactions were the same, as observed for 4-fluoro-2-hydroxybenzamide. In addition, sVP40 Δ 43 WT tolerates modifications at C3, C4 and C5 in general but was demanding regarding their chemical nature as 3-amino-SA did bind, but 3-chloro-SA did not, for example. C5 was therefore considered to be the most promising position for future fragment extension processes. Soaking of LL060 revealed a rotation of the fragment with the carboxyl group interacting with sVP40 L158. Two further H-bonds were formed by the fragment's tail to sVP40 R214 and sVP40 L288. As the fragment LL060 exhibited more H-bonds than the original fragment SA, the last part of this chapter therefore describes the characterization of this fragment.

Characterization of LL060

In contrast to the treatment with NaSA, LL060 did not exhibit a significant effect on reporter gene activity using a MG assay (Figure 53 A). Significant effects, however, could be observed in budding

Results

assays (Figure 53 B), suggesting that in contrast to the experiments with NaSA, LL060 affects only budding and not viral RNA synthesis. This might be due to blockage of sVP40 residues important for budding such as L158, L213 and/or R214. The fact that the highest concentration did not impair VLP release and lower concentrations show nearly the same effects points towards a potential self-interaction of the compound or low solubility of LL060 in cell culture medium. Nevertheless, a significant inhibition of VLP formation using concentrations 1 μ M LL060 indicates that this compound represents a highly promising lead-like molecule for the structure-guided drug design process of the present work.

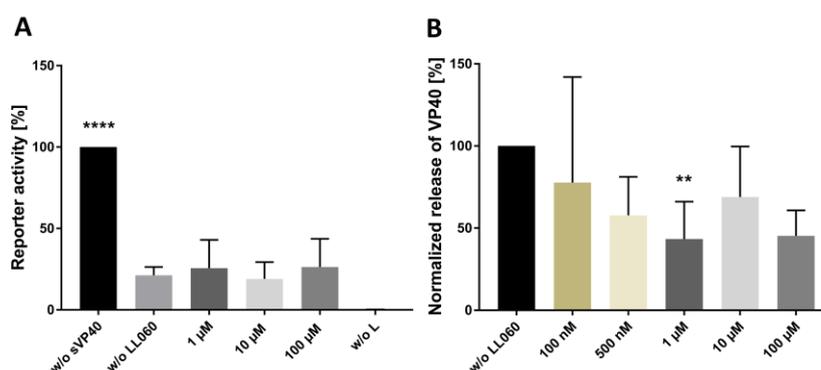


Figure 53: LL060 disrupts sVP40's ability to bud VLPs but not its effect on viral RNA synthesis. A) Treatment with LL060 influences reporter gene activity in a MG assay. HEK293 cells were transfected with pCAGGS sVP40 WT and the MG assay components. Cell culture medium was changed 4 h pt and LL060 added. Cells were lysed 48 h pt, reporter gene activity was measured and normalized to the sample without sVP40 (set to 100%). For the negative control, no polymerase L was added to the transfection mix. B) Effect of LL060-treatment on VP40-release. B) Cells were transfected with pCAGGS zGP and sVP40 WT. Cell culture medium was changed 4 h pt and LL060 added. Supernatants were collected and cells lysed 24 h pt and WB analyses of cellular and released sVP40 were performed using rabbit α -sVP40 and mouse α -tubulin primary and donkey α -mouse and goat α -rabbit IRDye[®] 680 as secondary antibodies. C) Quantification of VLPs: Released sVP40 was normalized to the amount of sVP40 in lysates and sVP40 WT w/o LL060 set to 100%. Bars indicate the mean \pm SD of at least three experiments, asterisks indicate statistical significance (One-way ANOVA) compared to sVP40 WT without LL060 as follows: ** $P < 0.005$ and **** $P < 0.0001$.

In conclusion, crystal soaking was a highly successful method for the identification of a lead-like molecule as well as follow-up lead structures.

3.3.2.3 Discussion

SA was bound to sVP40 _{Δ 43} WT in the pocket between the NTD and CTD and formed interactions with L158 and R214 in the initial soaking experiment. Due to the restricted flexibility of proteins in the crystalline state, soaking potentially introduces a bias into the interaction mode [82,325]. Location, orientation and the formation of interactions of SA were therefore confirmed via co-crystallization. In addition, binding in solution was investigated using TSA and HDX-MS. According to Konerman *et al.*, ligand binding in HDX-MS can result in three different scenarios (0, 1 and 2): type 0 describes unaltered exchange rates upon ligand binding, type 1 decreased rate due to increased intramolecular H-bonds and type 2 increased rate [176]. For the present work, a type 1 scenario with the decrease of D₂O incorporation was observed for peptides 203 to 209 and 248 to 261, suggesting that these residues are

Results

shielded against the solvent upon NaSA binding. The crystal structure of sVP40_{Δ43} WT in complex with SA shows slightly decreased B-factors for these peptides compared to the published apo-structure (Figure 54) which points towards a potential stabilizing effect upon binding of SA.

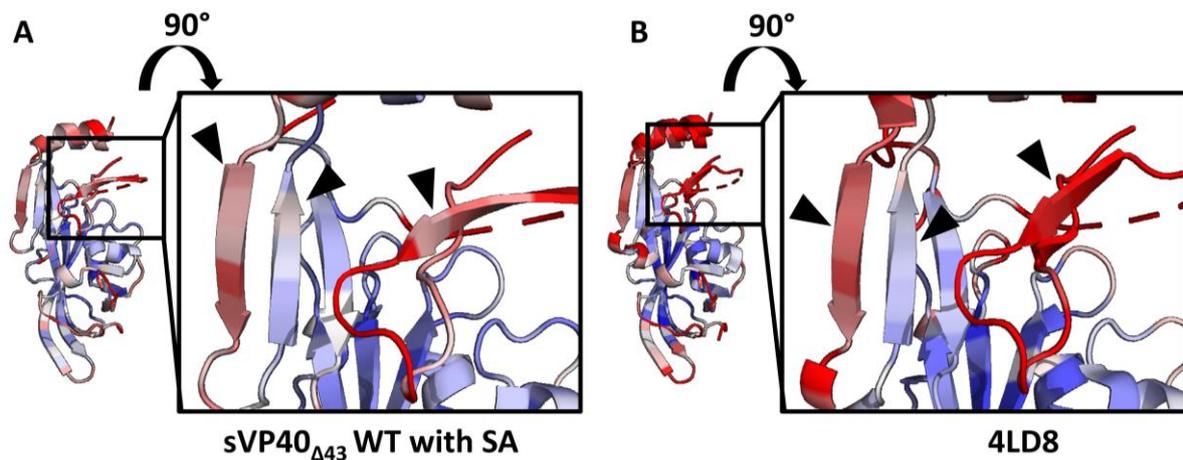


Figure 54: B-factors analysis of 8B2U and 4LD8 crystal structures. B-factors are plotted on the crystal structures of sVP40_{Δ43} WT in complex with SA (A, AW65, 8B2U) and in its apo form (B, 4LD8) and shown in a blue–white–red color scale with blue indicating the lowest (stabilization) and red the highest values (destabilization). For each crystal, an overview is given on the left side on each panel and a close-up on the right side, with arrowheads pointing towards peptides 203-209 and 248-261.

Interestingly, no significantly altered exchange rates were observed for residues L158 and R214 but similar phenomena were reported previously. Glasgow *et al.* compared the ligand-bound state with the apo-state of LacI via HDX-MS and observed decreased H/D exchange near but not directly at the ligand binding pocket. Other areas such as the domain interface were affected as well and the authors proposed water-mediated interaction networks as a possible reason for the results [112]. This could also be the case in the present work: in the original sVP40_{Δ43} WT-SA crystal structure, R214 forms an H-bond to F209 via a water molecule and sidechains of L207 interact with sidechains of I261 via van-der-Waals forces. It is also likely that peptide 203 to 209 indirectly interacts with peptide 248 to 261 via the poorly resolved amino acids 297 to 300 which are not included in the structural model. As the pocket usually harbours only few water molecules it seems likely that a compound influencing the pocket residues subsequently influences water-mediated bonds, possibly resulting in a chain reaction towards the C-terminal core. Next to orthosteric differences, decreased rates of D₂O uptake of allosteric sites is a well-reported phenomenon [122,169,199,292]. Interestingly, peptide 248 to 260 was affected in the SA-sVP40_{Δ43} WT complex as well as during HDX-MS experiments using sVP40_{Δ43} W95A and L117A (Figure 18 and Figure 19), albeit with an increased H/D exchange. This suggests, that this area is structurally highly dependent on the interaction of the pocket, where W95 is also located.

In TSA, binding of NaSA was correlated with a decrease of the melting temperature, suggesting a destabilizing effect of the ligand which has been reported previously for other protein-ligand complexes [29,62,211]. As the binding of NaSA/SA did not result in an altered crystal packing or

Results

increased H/D exchange rates for the dimer interface, it seems likely that other destabilizing factors play a role. Possibly, the destabilization events observed via TSA are caused by the interaction of SA with the amino acid backbone of VP40, which could potentially affect other peptides as well. Peptides outside of the pocket were also influenced during HDX-MS experiments and showed decreased H/D exchange rates, which could be interpreted as conformational tightening. In an attempt to explain the effects of a changed local environment of the pocket, a characterization of the pocket mutants, including the direct interaction partners sVP40 L158 and sVP40 R214A, as well as the neighbouring residue L213A was performed. It could be shown that all three residues were important for budding (Figure 48). Testing equivalent mutants of zVP40, McCarthy *et al.* proposed major conformational rearrangements and the formation of higher-order oligomers [220] which is in line with our results: it could be shown that while all three mutants still formed predominantly dimers, the amount of octamers was increased compared to sVP40_{Δ43} WT, suggesting decreased stability of the dimers (Figure 49). While an increase of octameric activity in a MG assay was therefore expected, the opposite was observed: for sVP40 R214A, downregulation of viral genome replication and transcription was completely abolished, whereas no significant effects could be seen for sVP40 L158A and L213A (Figure 47), suggesting a crucial but yet unidentified role for sVP40 R214. zVP40 L213 was reported to penetrate the plasma membrane, along with I293, L295 and V298 [5]. Hence, it is possible that similar implications are true for both sVP40 R214 and potentially also for L158 as this residue is also located on the outer rims of the pocket between the NTD and CTD.

Combining the structural information from both HDX-MS and crystal soaking, SA is highly promising regarding the orthosteric as well as the allosteric binding sites, all shown to be important for various functions of sVP40 WT, suggesting that sufficiently potent binders could inhibit these functions efficiently.

A search of SA as a crystallographic binder in the PDB revealed numerous hits [20,96,105,226,227,263,317,319,320,339], also including screenings using the same library [27,114]. SA-derivatives such as 2-hydroxy-4-aminobenzoic acid P-hydroxybenzoic acid are bound to the SARS-CoV-2 NSP3 macrodomain in the crystalline state (5RUE and 5RTJ, respectively) [65].

VP40 is not an enzymatic protein and therefore does not contain active sites in deep pockets but exhibits a relatively flat surface with the exception of the pocket between the N- and C-terminal domain and another shallower tunnel formed by residues 67 to 70, 102 to -105 and 273 to 280. The SA-binding pocket contains several hydrophobic residues W95, L96, P97, L98 and A156, F157, L158, F161, V162 in the NTD and L213 and V287, L288, L289, Y292, I293 in the CTD. Both viral and host proteins were characterized as interaction partners, such as NP [330] or proteins of the COPII transport system [332], but the involved interaction residues were not always determined and no structural

Results

studies are available. Therefore, it can be proposed that the pocket serves as a potential binding site for an interaction partner. This was suggested by Urata *et al.* who characterized the 292-YIGL-295 motif for zVP40. These residues are located deep in the pocket and were shown to be critical for both VLP production and the regulation of viral RNA synthesis [308]. This further increases the attractiveness of the pocket but also raises questions about the accessibility. In this regard, it needs to be mentioned that accessibility is given due to the flexibility of the R214 sidechain, which could not be completely modelled in any of the sVP40 Δ ₄₃ WT structures of the present work. Generating the symmetry mates of the crystal packing showed that the pocket in question is still solvent-exposed and not blocked by other protomers in the crystal whereas other areas are mainly obstructed (Figure 55).

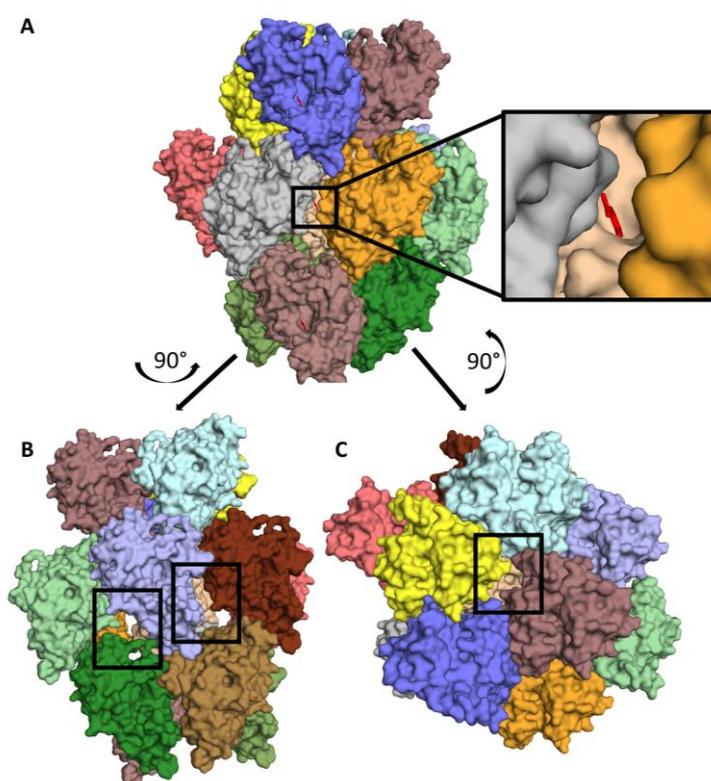


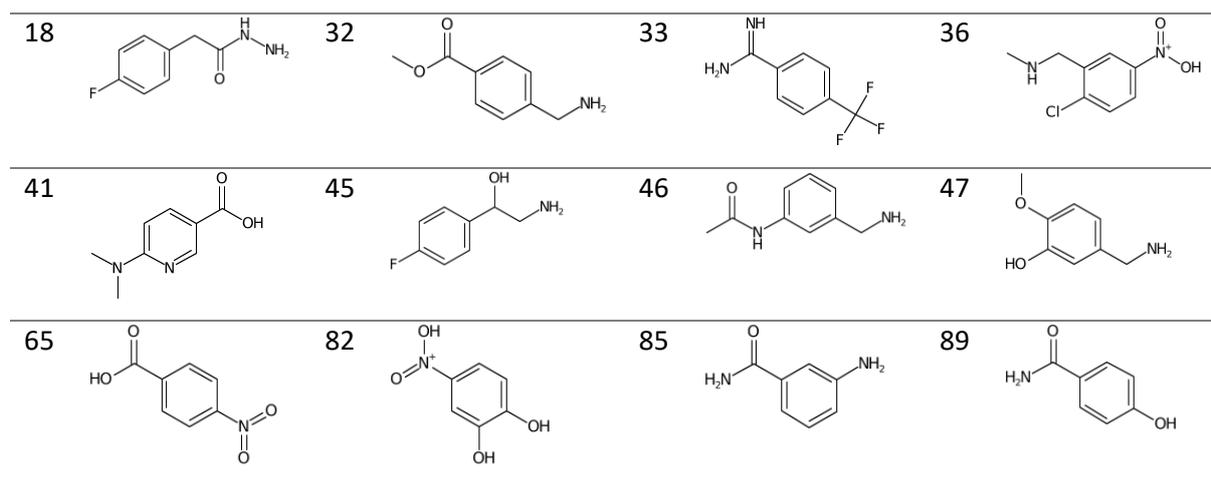
Figure 55: Crystal packing of sVP40 Δ ₄₃ WT in complex with SA. Various symmetry mates are shown in different colours centred around the original protomer (wheat). A) Front view of the original protomer (wheat) with a close-up of the SA-binding pocket. B) Side view and C) top view depicting further access sites to the original protomer in boxes.

With SA as the only bound fragment, the hit rate of this screening is surprisingly low compared to previous publications using the same fragment library [114,130,296]. Fragments of the same library that are chemically similar did not bind, suggesting that the functional groups are crucial for successful soaking (Table 20).

Table 20: Fragments of the Jena FragXtal Screen with similar building blocks than SA.

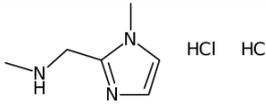
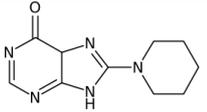
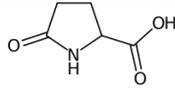
ID	Structure	ID	Structure	ID	Structure	ID	Structure
1		10		12		17	

Results



Also, it seems conceivable that some bound fragments were missed due to the soaking conditions. In the present work, a constant fragment concentration of 100 mM was chosen but soaking times varied. However, an alternative soaking protocol with lower concentrations in combination with longer soaking times or a stepwise increase in fragment concentration might have provided a higher hit rate, [233] especially with regards to the disruptive fragments. For these molecules (#37, #50 and #76; see Table 21), no dataset compliant with the selected quality standards could be achieved, even for soaking times of only seconds. Docking studies or co-crystallization of these fragments with sVP40 are potential alternatives to elucidate if, how and where they might possibly bind.

Table 21: Disruptive fragments of the FragXtal screen. Jena fragments #37, #50 and #76 were considered disruptive, as the sVP40_{Δ43} WT crystals showed low to no diffraction when incubated with one of those fragments.

ID	Structure	ID	Structure	ID	Structure
37		50		76	

The necessary soaking time is ligand-dependent and some fragments might need several hours to days to be soaked into the crystal to a high occupancy [114,273,277,281]. However, this is often the case if the binding site is already occupied by another molecule, possibly buffer ingredients [273], which can be probably excluded for this study, as the pocket was empty except for a few water molecules. In addition, the fact that PanDDA detected 23 events (repeats of AW65 and AW107 not included), but only one with reliable electron density, was surprising as this suite was originally designed to detect bound fragments with a low occupancy. However, the soaking times of the different crystal datasets varied immensely, ranging from seconds to 16 h. This potentially introduced a bias into the analysis and could potentially explain the difference between observed and actual binding sites. Employing the Autorefinement pipeline by Schiebel *et al.* on all 140 datasets was meant to identify further fragments

Results

that were potentially missed in the previous analysis using PanDDA. However, this analysis revealed that approx. 50% of datasets were of superior quality than 4LD8 which was used for molecular replacement and therefore necessitating the addition of further amino acids to the model. The positive electron density belonging to missing residues was typically filled with water molecules by the refinement pipeline, thereby introducing a bias into these models. The manual correction of these models was highly time-consuming without resulting in further identified fragments. It could therefore be concluded that sVP40 is a highly demanding protein. The fact that VP40 has a flat surface and offers only one major pocket was already discussed in the first part of this section but is surely another factor regarding the low hit rate. In hindsight, performing a pre-screening such as TSA or SPR using a larger library and then validating the identified hits using crystal soaking would have been less laborious. Another approach to increase the hit rate consists of introducing point mutations and thereby engineering pockets, generating dimerization-incompetent mutants (such as sVP40_{Δ43} L117A, see section 3.1.1), changed crystallization conditions to achieve an alternative crystal packing or the formation of a covalently-linked dimer exposing the interface (as discussed in section 3.3.1.3).

As discussed in the previous section, the obtained results suggest a high druggability of the pocket and a compound that binds with a higher affinity than SA could very likely result in a stabilization of the dimer, thereby inhibiting octamerization. While the formation of filaments is probably not hindered, the ability to induce membrane curvature and bud viral particles would most likely be reduced due to the decreased protein fluidity. However, when considering TSA results, destabilization of the dimer is also possible. Other possibilities include impaired functions due to the blockade of the crucial residues L158, L213 and R214 and the potential hindered interaction with other viral or host proteins due to the occupied pocket, although the latter is purely speculative. An overview of the possible inhibition modes is depicted in Figure 56 at the end of this section. However, to achieve any of the mentioned effects, affinity of the compounds needs to be greatly increased.

Fragment extension without structural knowledge (using for example TSA, SPR, ITC, MST or others) implies methods such as methyl walks where methyl groups are attached to C-atoms or switched with functional groups to identify potential positions for the attachment of further atoms. However, this requires extensive work to synthesize a multitude of derivatives. With a structure-guided approach, the efforts are greatly reduced as possible extension sites are easily located and validated. *In silico* methods represent efficient tools for fragment extension, including the design of derivatives and prediction of binding modes. Depending on the selected programs, other factors can be estimated such as feasibility of chemical synthesis. For the present work, the online tool PINGUI [60] was used which allows the selection of reactions and comprises commercially available building blocks only. While *in silico* docking usually provides a useful tool to screen a large number of molecules for fragment

Results

extension, this approach proved to be unsuitable for the early stages of fragment extension for the present work. As shown by further soaking experiments, fragments 3-amino-SA, 4-fluoro-SA, 4-fluoro-2-hydroxy-benzamide and 5-amino-SA were bound to VP40 in the exact same location and orientation than the original fragment SA and a flip of functional groups was therefore considered unlikely. However, the output of docking analyses showed flipped molecules to a high percentage, suggesting that this particular software judged these poses favourable. Of course, this bias might have been introduced by the selected parameters. An alternative docking suite or adjusted parameters such as an increased number of generated poses might help to avoid this bias. Additionally, the four different sidechains of R214 were revealed to be of only minor importance, as all of the extended molecules interacted with the backbone of R214, exclusively. Nevertheless, this approach will most likely be used during later steps of fragment extension, using LL060 as a new template. The alternative approach to test SA analogues helped to confirm the binding site, the interaction partners as well as the orientation of the fragment. In retrospect, this analysis revealed to be most successful. The atom-by-atom extension had the additional advantage of taking synthesis feasibility into constant account.

While performing a crystal soaking screening is very laborious, once the bottleneck of beamtime is overcome, it serves as a powerful technique to detect crystallographic binders. While a final compound could not yet be identified, substantial effort has been made in establishing reproducible expression of recombinant VP40 in *E. coli*, purification, crystallization as well as a robust crystal soaking and the identification of SA as a crystallographic binder. This enabled us to employ structure-guided hit-to-lead development via fragment extension. Soaking of different SA-derivatives allowed us to draw useful conclusions towards possible modifications, resulting in the synthesis of the promising compound LL060, taking the project closer towards the design of a sVP40 inhibitor that is able to impair the protein's functions through various potential mechanisms. A possible model of inhibition is shown in Figure 56.

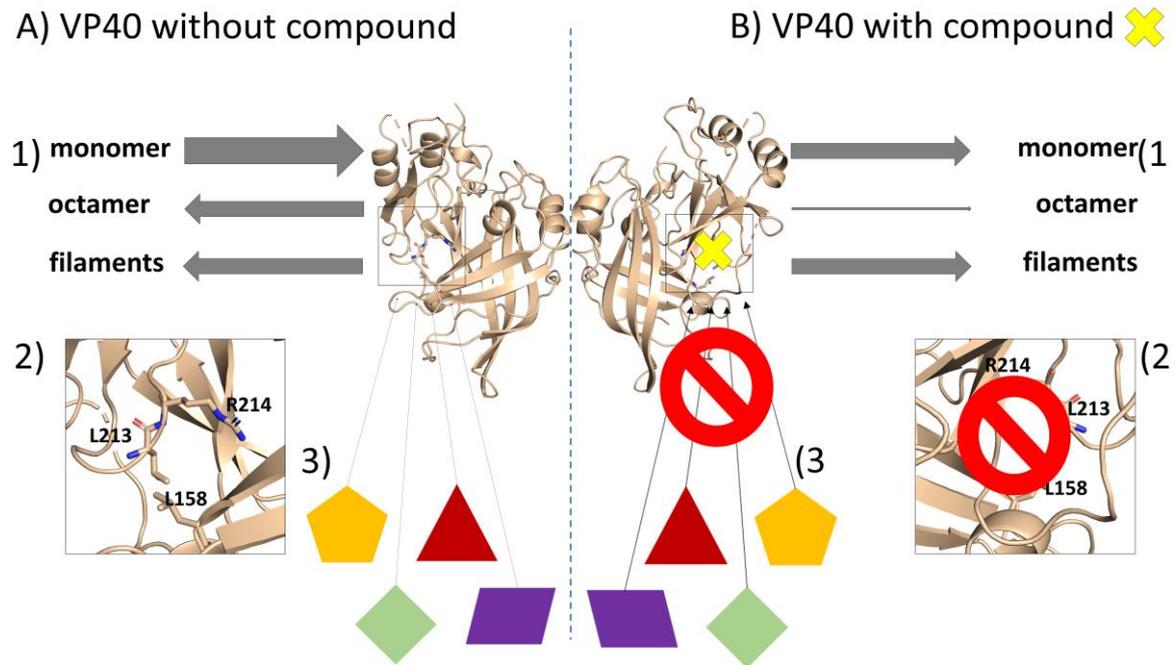


Figure 56: Model of VP40 inhibition upon binding of a compound in the pocket between the NTD and CTD. The left side of the panel depicts the apo state without compound, the right side with compound (indicated by a yellow X): 1) Direct homo-oligomerization: stabilization of dimeric VP40 potentially decreases octamerization because the CTD cannot flip up which potentially increases the formation of filaments or destabilizes the dimer, therefore increasing the monomeric fraction of the protein; 2) Impairment of VP40's functions due to blocked amino acids L158, L213 and R214; 3) Upon occupation of the pocket, other viral or host proteins are unable to bind to and interact with VP40.

4 Summary discussion and outlook

The different sub-projects of the present work are all intertwined regarding the interpretation of results. The characterization of residues directly responsible for homo-oligomerization or indirectly linked to the functions of the various homo-oligomers showed that the shapeshifting process of VP40 is more complex than anticipated: This was clearly shown for sVP40 W95A and L117A. When expressed in mammalian cells, sVP40 W95A showed a stronger inhibition of viral RNA synthesis and induction of the formation of VLPs than the WT but was less stable in solution measured by TSA and SEC. In contrast, sVP40 L117A completely lost its ability to dimerize and also showed fully abolished functions concerning the downregulation of viral genome replication and transcription as well as budding. Because biochemical assays showed this mutant to be a monomer, which, however, crystallized as a dimer, this mutant would not have been useful for the identification of fragments binding to the dimeric interface via crystal soaking. This notion strengthens our two-pronged approach of performing screenings with proteins in solution and in the crystalline state but also the parallel approach of recombinant proteins expressed in both *E. coli* and mammalian cells. Nevertheless, targeting the homo-oligomerization process seems highly suitable for the development of antiviral drugs, as already little changes such as point mutations have tremendous effect on VP40's functions, as shown for sVP40 L117A, W95A as well as L158A, L213A and R214A. The observed differences between zVP40 and sVP40 expressed in *E. coli* could potentially be attributed to the experimental conditions, but a lower degree of octamerization for sVP40 compared to zVP40 seems to be possible. The higher amount of transfected sVP40- compared to zVP40-plasmid in a MG setting to reach similar levels of reporter gene activity points in this direction. Also, sVP40 $_{\Delta 43}$ L213A and R214A mutants behaved in a similar fashion as reported for the same zVP40 mutants [220]. Therefore, it can be concluded that homo-oligomerization is as crucial for sVP40 as it is for zVP40, especially with regards to the functional assays in cell culture experiments. Hence, inhibition of the VP40 self-assembly process represents a promising strategy to disrupt the viral replication cycle, regardless of the species (EBOV or SUDV).

Peptides did not represent a suitable approach for the present work but the peptide-based studies demonstrated a very strong affinity between VP40 protomers, further stressing the challenge to target self-assembly. This chapter also suggested that allosteric sites are potentially more promising targets for the development of antivirals.

The cysteine residues characterized in section 3.1.2 were shown to not influence homo-oligomerization directly but probably allosterically as their mutation led to decreased budding efficiency and changed behaviour regarding the regulation of viral RNA synthesis. In addition, they were shown to be redox-regulated, which also represents a potential drug target. Cysteines are also typical targets of covalent inhibitors, which was also briefly discussed in section 3.3.1.3. The knowledge

from chapter 3.1.2 and the principle of covalent inhibitors was exploited during the screening process via DT, which revealed to be promising for the identification of a lead-like molecule.

Crystal soaking was the most successful approach for the present study with one initial hit and five follow-up structures. Due to the robust soaking conditions, this method was chosen for the identification of extended fragments. While this approach was laborious and time-consuming, it allowed structure-guided design of larger molecules from the very beginning. Also, new fragments were not synthesized in batches, so usually two to three fragments at a time were ready for testing. In addition, TSA was performed to confirm binding in solution. The combined results of binders and non-binders allowed a thorough characterization of possible fragment orientations and locations. This structure-guided drug design process is considered highly promising. Once a sufficient number of fragments with a reasonable molecular weight and number of interactions are identified, medium-throughput screens such as SPR will be used to not just determine binding but also affinity and kinetics. This method was already established for chapter 3.2. The interdigitation of the different subprojects is depicted in Figure 57.

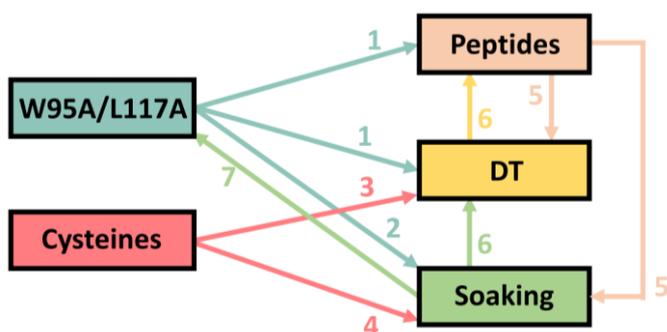


Figure 57: Cross-project results for the present work. The present work covered five different subprojects that influenced interpretation of other sections. 1) Complexity of homo-oligomerization, 2) Knowledge about HDX-MS behaviour, 3) Relevance of cysteines, 4) Relevance of allosteric sites as drug target, 5) Unsuitability of peptides, 6) Co-crystallization success towards low-affinity binders and 7) Characterization of additional important residues.

In summary, this work presents the basis for the design of a potent antiviral compound targeting the Ebola virus matrix protein VP40, by delivering extensive data on hot spot residues and other amino acids, such as cysteines or pocket mutants as well as providing detailed protocols for various highly robust screening methods and LL060 as a potent lead compound.

5 Summary

Viruses of the genus Ebola virus cause severe fevers with unusually high case fatality rates, and as of today, no licensed antiviral drugs are available. The Ebola virus matrix protein VP40 plays a major role during budding of new viral particles and is also involved in the regulation of viral genome replication and transcription. VP40 exists in three different homo-oligomeric forms, namely dimers, octamers and polymeric filaments. Dimers are transported towards the plasma membrane where the interaction with lipids triggers the formation of VP40-filaments. Filaments represent a flexible row of dimers which enable budding of viral particles and line the inner layer of the new virions. Upon binding of cellular RNA, VP40 dimers turn into ring-shaped octamers that exert an inhibitory effect on viral RNA synthesis. For the present work, the resolution of the VP40 crystal structure was improved and a structure-guided drug design was employed with the aim to impair VP40's essential homo-oligomerization. The study investigated VP40 of two Ebola virus strains Zaire (zVP40) and Sudan (sVP40).

Residues L117 and W95 are so-called hot spot amino acids of the dimeric and octameric interface of sVP40 that were characterized by mutational analyses. As expected, both sVP40 wildtype (WT) and W95A formed dimers nearly exclusively whereas sVP40 L117A formed mainly monomers when expressed in *E. coli*. Surprisingly though, the structure of both sVP40 $_{\Delta 43}$ mutant oligomers was similar to sVP40 WT. As a dimeric crystal packing was surprising for monomeric sVP40 L117A, crystallographic artefacts were considered which prompted the structural analysis of the mutants in solution. Using hydrogen-deuterium exchange mass spectrometry and thermal shift assays, it could be shown that both mutants exhibited increased fluidity and decreased stability in solution and it could be confirmed that sVP40 L117A was indeed a monomer. *In cellulo*, the ability of sVP40 L117A to form virus-like particles (VLPs) and inhibit viral genome replication and transcription was completely abolished, whereas sVP40 W95A exhibited a gain of function as this protein released more VP40-VLPs into the cellular supernatant and also showed a stronger inhibition of viral RNA synthesis. This data suggests that targeting homo-oligomerization is a promising strategy to impair VP40 functionality, but demands interdisciplinary methods, especially regarding structure determination.

Based on a high-resolution crystal structure of sVP40 $_{\Delta 43}$ WT, the structure of the C-terminus of sVP40 could be analyzed. The C-terminus contains the only two cysteines of the molecule which were oxidized and formed a disulfide bridge in the crystal. When VP40 was expressed in mammalian cells and released into the supernatant, the cysteines were also oxidized by post-translational modifications such as glutathionylation and nitrosylation. *In vitro*, VP40 could be reduced again by human thioredoxin. While the overall structure and oligomeric state of sVP40 was preserved, mutation of the cysteines resulted in altered phenotypes with regard to VP40's ability to regulate viral RNA synthesis and to induce budding and particle formation.

In an attempt to disrupt homo-oligomerization directly, interface-mimicking peptides were designed and tested in both functional cell culture and biochemical assays. While their binding to VP40 could be demonstrated, the peptides were unable to influence VP40's functions or self-assembly.

Further, the dimeric interface of VP40 should be targeted with small molecules. To this end, disulfide tethering was established as an alternative approach using a sVP40 variant with a cysteine residue near the dimeric interface (N67C). This method exploits the formation of a covalent disulfide bridge between the introduced cysteine at position 67 and thiol-containing fragments. Upon incubation under reducing conditions, only fragments with additional interactions to VP40 were bound favourably and could be detected via intact protein mass spectrometry, yielding several fragment hits. While no structural information of one of the sVP40 N67C-fragment complexes could be determined to assess binding mode and location, this strategy proved to be highly successful in identifying promising lead-like molecules.

Using a library of 96 preselected fragments and crystal soaking, salicylic acid (SA) was identified as a crystallographic binder of VP40. The binding to VP40 could be confirmed in solution. As expected, the weak binding resulted in only minor effects on VP40's function in RNA synthesis and budding. The characterization of residues of VP40 involved in the interaction with SA (L158 and R214) suggested that the binding pocket between the N- and the C-terminal domain is a highly vulnerable target site as mutation of these residues resulted in a decreased ability to regulate viral genome replication and transcription for sVP40 R214A as well as decreased budding for both sVP40 L158A and R214A. This prompted the testing of SA derivatives and the identification of four further crystallographic binders (3-amino-SA; 4-fluoro-SA, 4-fluoro-2-hydroxybenzamide and 5-amino-SA). Further structure-guided drug design led to the design, synthesis and testing of LL060, a compound that was also able to impair the formation of VP40-VLPs.

In summary, a drug design process from scratch to target the function of Ebola virus VP40 is described. Here, we characterized and validated VP40 homo-oligomerization as a target and identified several lead-like molecules originating from a site-directed ligand discovery screening. The highly promising lead compound LL060, which was identified via crystal soaking represents the starting point for the development of a potent Ebola virus inhibitor.

7 Zusammenfassung

Viren aus dem Genus Ebola-Virus der Familie der *Filoviridae* verursachen schweres Fieber mit ungewöhnlich hoher Sterblichkeitsrate, und bis heute gibt es keine zugelassenen antiviralen Medikamente. Das virale Matrixprotein VP40 spielt eine wichtige Rolle bei der Knospung neuer Viruspartikel und ist auch an der Regulierung der viralen Genomreplikation und Transkription beteiligt. VP40 existiert in drei verschiedenen homo-oligomeren Formen, nämlich als Dimere, Oktamere und polymere Filamente. VP40-Dimere werden zur Plasmamembran transportiert, wo die Interaktion mit Lipiden die Bildung von VP40-Filamenten auslöst. Filamente stellen eine flexible Reihe von Dimeren dar, die die Knospung von Viruspartikeln ermöglichen und die innere Schicht der neuen Virionen auskleiden. Nach der Bindung von zellulärer RNA verwandeln sich die VP40-Dimere im Cytosol in ringförmige Oktamere, die eine hemmende Wirkung auf die virale RNA-Synthese ausüben. Für die vorliegende Arbeit wurde die Auflösung der Kristallstruktur von VP40 optimiert und ein strukturgesteuertes Wirkstoffdesign mit dem Ziel eingesetzt, die für VP40 essentielle Homo-Oligomerisierung zu beeinträchtigen. In der Studie wurde VP40 der beiden Ebolavirus-Stämme Zaire (zVP40) und Sudan (sVP40) untersucht.

Die VP40-Reste L117 und W95 sind sogenannte Hot-Spot-Aminosäuren der dimeren und oktameren Schnittstelle von sVP40, die durch Mutationsanalysen charakterisiert wurden. Wie erwartet bildeten sowohl sVP40 Wildtyp (WT) als auch W95A fast ausschließlich Dimere, während sVP40 L117A bei der Expression in *E. coli* hauptsächlich Monomere bildete. Überraschenderweise war die Struktur der beiden sVP40-Mutanten jedoch ähnlich wie die von sVP40 WT. Da die monomere sVP40 L117A eine dimere Kristallpackung aufwies, wurden kristallografische Artefakte in Betracht gezogen, was die Strukturanalyse der Mutanten in Lösung zur Folge hatte. Mit Hilfe der Wasserstoff-Deuterium-Austausch-Massenspektrometrie und der thermischen Verschiebung konnte gezeigt werden, dass beide Mutanten in Lösung eine erhöhte Flexibilität und eine verringerte Stabilität aufweisen. Zudem konnte bestätigt werden, dass sVP40 L117A tatsächlich ein Monomer ist. *In cellulo* war die Fähigkeit von sVP40 L117A virusähnliche Partikel (VLPs) zu bilden und die virale Genomreplikation und -transkription zu hemmen, fast vollständig aufgehoben, während sVP40 W95A einen Funktionsgewinn zeigte, da dieses Protein mehr VP40-VLPs in den Zellüberstand freisetzte und auch eine stärkere Hemmung der viralen RNA-Synthese zeigte. Diese Daten deuten darauf hin, dass die gezielte Homo-Oligomerisierung eine vielversprechende Strategie zur Beeinträchtigung der VP40-Funktionalität ist, die jedoch interdisziplinäre Methoden erfordert, insbesondere hinsichtlich der Strukturbestimmung.

Basierend auf einer hochauflösenden Kristallstruktur von dimerischem sVP40 WT konnte die Struktur des C-Terminus von sVP40 analysiert werden. Der C-Terminus enthält die einzigen beiden Cysteine des Moleküls, die im Kristall eine Disulfidbrücke bilden. Wenn VP40 in Säugetierzellen exprimiert und in

Zusammenfassung

den Überstand freigesetzt wurde, wurden die Cysteine auch durch posttranslationale Modifikationen wie Glutathionylierung und Nitrosylierung oxidiert. *In vitro* konnte VP40 durch humanes Thioredoxin wiederum reduziert werden. Während die Gesamtstruktur und der oligomere Zustand von sVP40 erhalten blieben, führte die Mutation der Cysteine zu veränderten Phänotypen im Hinblick auf die Fähigkeit von VP40, die virale RNA-Synthese zu regulieren und die Knospung und Partikelbildung zu induzieren.

In einem Versuch, die Homo-Oligomerisierung direkt zu unterbrechen, wurden Peptide, die die Aminosäuresequenz der verschiedenen Schnittstellen nachahmen, entwickelt und sowohl in funktionellen als auch in biochemischen Versuchen getestet. Obwohl ihre Bindung an VP40 gezeigt werden konnte, waren die Peptide nicht in der Lage, die Funktionen oder die Selbstorganisation von VP40 zu beeinflussen.

Darüber hinaus sollte die dimere Schnittstelle von VP40 mit kleinen Molekülen angegangen werden. Zu diesem Zweck wurde das Disulfid-*Tethering* als alternativer Ansatz unter Verwendung einer sVP40-Variante mit einem durch Mutagenese eingefügten Cysteinrest in der Nähe der dimeren Schnittstelle (N67C) etabliert. Diese Methode nutzt die Bildung einer kovalenten Disulfidbrücke zwischen dem eingeführten Cystein und thiolhaltigen Fragmenten. Nach der Inkubation unter reduzierenden Bedingungen wurden nur Fragmente mit zusätzlichen Wechselwirkungen zu VP40 positiv gebunden und konnten mittels Massenspektrometrie des intakten Proteins nachgewiesen werden, wobei sich mehrere Fragmenttreffer ergaben. Obwohl keine strukturellen Informationen über einen der sVP40 N67C-Fragmentkomplexe vorlagen, um den Bindungsmodus und den Bindungsort zu bestimmen, erwies sich diese Strategie als äußerst erfolgreich bei der Identifizierung vielversprechender leitstrukturähnlicher Moleküle.

Mit Hilfe einer Bibliothek von 96 vorselektierten Fragmenten wurde Salicylsäure (SA) mittels *crystal soaking* als kristallographischer Binder von VP40 identifiziert. Die Bindung an VP40 konnte in Lösung bestätigt werden. Die schwache Bindung hatte erwartungsgemäß nur geringe Auswirkungen auf die Funktion von VP40 bei der RNA-Synthese und beim *Budding*. Die Charakterisierung von VP40-Resten, die an der Interaktion mit SA beteiligt sind (L158 und R214), deutete darauf hin, dass die Bindungstasche zwischen der N- und der C-terminalen Domäne ein sehr empfindlicher Zielort ist, da die Mutation dieser Reste zu einer verminderten Fähigkeit zur Regulierung der viralen Genomreplikation und Transkription für sVP40 R214A sowie zu einer verminderten Knospung sowohl für sVP40 L158A als auch R214A führte. Dies veranlasste das Testen von SA-Derivaten und die Identifizierung von vier weiteren kristallographischen Liganden (3-Amino-SA, 4-Fluor-SA, 4-Fluor-2-Hydroxybenzamid und 5-Amino-SA). Ein weiterer strukturgeführter Wirkstoffentwurf führte zur

Zusammenfassung

Entwicklung, Synthese und Prüfung von LL060, einer Verbindung, die erstmals die Bildung von VP40-VLPs beeinträchtigen konnte.

Zusammenfassend wird das strukturbasierte Verfahren zur Entwicklung von Arzneimitteln von Grund auf beschrieben, das auf die Funktion des Ebola-Virus VP40 abzielt. Die Homo-Oligomerisierung von VP40 wurde als vielversprechendes Ziel charakterisiert und validiert und mehrere Leitmoleküle identifiziert, die aus zwei verschiedenen Liganden-Screenings hervorgingen. Die vielversprechende Leitverbindung LL060, die durch *crystal soaking* entwickelt wurde, bildet den Ausgangspunkt für die Entwicklung eines Ebola-Virus-Inhibitors.

8 Materials

8.1 Cells and viruses

8.1.1 Mammalian cells

HEK293 (Human embryonic kidney cells)	Institute for Virology, Marburg
HuH7 (Human hepatoma cells)	Institute for Virology, Marburg
VeroE6 (African green monkey kidney epithelial cells)	Institute for Virology, Marburg

8.1.2 Viruses

Zaire Mayinga Ebola virus	Institute for Virology, Marburg
---------------------------	---------------------------------

8.1.3 *E. coli*

Rosetta2	F ⁻ <i>ompT hsdS_B(r_B⁻ m_B⁻) gal dcm pRARE2 (Cam^R)</i>	Sapphire lab, La Jolla Institute for Immunology, San Diego, USA
XL1 Blue	<i>recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac [F⁺ proAB lacIqZΔM15 Tn10 (Tet^r)]</i>	Stratagene, Heidelberg
XL10 Gold	<i>TetrD(mcrA)183 D(mcrCB-hsdSMR-mrr)173 endA1 supE44 thi-1 recA1 gyrA96 relA1 lac Hte [F⁺ proAB lacIqZDM15 Tn10 (Tetr) Amy Camr]</i>	Agilent technologies, Ratingen

8.2 Vectors

8.2.1 Mammalian expression vectors

pAMP_recEBOV_WT	Institute for Virology, Marburg
pANDY_3E-5E (Renilla luciferase minigenome)	Institute for Virology, Marburg
pCAGGS	Institute for Virology, Marburg
pCAGGS_GP	Institute for Virology, Marburg
pCAGGS_L	Institute for Virology, Marburg
pCAGGS_NP	Institute for Virology, Marburg
pCAGGS_T7	Institute for Virology, Marburg
pCAGGS_VP30	Institute for Virology, Marburg
pCAGGS_VP35	Institute for Virology, Marburg
pCAGGS_zVP40_WT	Institute for Virology, Marburg
pKAN_recEBOV_fragment-1	Institute for Virology, Marburg
pKAN_recEBOV_fragment-2	Institute for Virology, Marburg
pKAN_recEBOV_fragment-3	Institute for Virology, Marburg

Plasmid name	DNA template	Primer	Strategy
pCAGGS_VP40	DNA string Life technologies	#5871; #5156	PCR; NotI, SacI
pCAGGS_mcherry-linker	DNA string Life technologies	#4875, #4840	PCR; XhoI, NheI
pCAGGS_myc-sVP40	pCAGGS_sVP40_WT	#5155; #5156	PCR; NotI, SacI
pCAGGS_sVP40_W95A	pCAGGS_sVP40_WT	#4908	Mutagenesis
pCAGGS_sVP40_L117A	pCAGGS_sVP40_WT	#4910	Mutagenesis
pCAGGS_sVP40_C314S	pCAGGS_sVP40_WT	#5424	Mutagenesis
pCAGGS_sVP40_C320S	pCAGGS_sVP40_WT	#5425	Mutagenesis
pCAGGS_sVP40_CCS	pCAGGS_sVP40_WT	#5424, #5425	Mutagenesis
pCAGGS_zVP40_C311S	pCAGGS_zVP40_WT	#4946	Mutagenesis

Materials

pCAGGS_zVP40_C314S	pCAGGS_zVP40_WT	#4947	Mutagenesis
pCAGGS_zVP40_CCS	pCAGGS_zVP40_WT	#4955	Mutagenesis
pCAGGS_flag-GGGGS	pCAGGS_mcherry-linker	#4907, #4840	PCR; XhoI, NheI
pCAGGS_flag-GGGGS-inhpep1	pCAGGS_sVP40_WT	#4913, #4914	PCR; SacI, NheI
pCAGGS_flag-GGGGS-inhpep1_NK	pCAGGS_sVP40_WT	#4915, #4916	PCR; SacI, NheI
pCAGGS_flag-GGGGS-inhpep2	pCAGGS_zVP40_WT	#4869, #4832	PCR; SacI, NheI
pCAGGS_flag-GGGGS-inhpep2_NK	pCAGGS_zVP40_WT	#4870, #4833	PCR; SacI, NheI
pCAGGS_flag-GGGGS-inhpep3	pCAGGS_zVP40_WT	#4871, #4834	PCR; SacI, NheI
pCAGGS_flag-GGGGS-inhpep3_NK	pCAGGS_zVP40_WT	#4872, #4835	PCR; SacI, NheI
pCAGGS_flag-GGGGS-inhpep4	pCAGGS_zVP40_WT	#4873, #4836	PCR; SacI, NheI
pCAGGS_flag-GGGGS-inhpep4_NK	pCAGGS_zVP40_WT	#4874, #4837	PCR; SacI, NheI
pCAGGS_sVP40_L158A	pCAGGS_sVP40_WT	#5950	Mutagenesis
pCAGGS_sVP40_L213A	pCAGGS_sVP40_WT	#5952	Mutagenesis
pCAGGS_sVP40_R214A	pCAGGS_sVP40_WT	#5947	Mutagenesis
pCAGGS_sVP40_L288A	pCAGGS_sVP40_WT	#7212	Mutagenesis
pKAN_recEBOV_fragment-1-VP40-W95A	pKAN_recEBOV_fragment-1	#4956	Mutagenesis

8.2.2 Bacterial expression vectors

pET46EK/LIC	Saphire lab, La Jolla Institute for Immunology, San Diego, USA
pETDuet-1	Klebe lab, Institute for Pharmaceutical Chemistry, Marburg
pET46_sVP40 _{Δ43} _WT	Saphire lab, La Jolla Institute for Immunology, San Diego, USA
pET46_sVP40 _{Δ43} _WT	Saphire lab, La Jolla Institute for Immunology, San Diego, USA
pET25_sVP40 _{Δ43} _WT-linker-pep1	BIOCAT

Plasmid name	DNA template	Primer	Strategy
pET46_sVP40 _{Δ43} _W95A	pET46_sVP40 _{Δ43} _WT	#4908	Mutagenesis
pET46_sVP40 _{Δ43} _L117A	pET46_sVP40 _{Δ43} _WT	#4910	Mutagenesis
pET46_sVP40 _{Δ43} _C314S	pET46_sVP40 _{Δ43} _WT	#5424	Mutagenesis
pET46_sVP40 _{Δ43} _C320S	pET46_sVP40 _{Δ43} _WT	#5425	Mutagenesis
pET46_sVP40 _{Δ43} _CCS	pET46_sVP40 _{Δ43} _WT	#5424, #5425	Mutagenesis
pETDuet_zVP40_WT	pCAGGS_zVP40_WT	#5042, #4691	PCR; BamHI + SacI
pETDuet_zVP40_WT	pETDuet_zVP40_WT	#5079	Mutagenesis
pETDuet_zVP40_WT _{Δ43} (1.MCS)_inhpep1 (2.MCS)	pET46_zVP40 _{Δ43} _WT	#5306, #5307	PCR; NdeI + XhoI
pETDuet_zVP40_WT _{Δ43} (1.MCS)_inhpep2 (2.MCS)	pET46_zVP40 _{Δ43} _WT	#5308, #5309	PCR; NdeI + XhoI
pETDuet_zVP40_WT _{Δ43} (1.MCS)_inhpep3 (2.MCS)	pET46_zVP40 _{Δ43} _WT	#5310, #5311	PCR; NdeI + XhoI
pETDuet_zVP40_WT _{Δ43} (1.MCS)_inhpep4 (2.MCS)	pET46_zVP40 _{Δ43} _WT	#5312, #5313	PCR; NdeI + XhoI
pET25_sVP40 _{Δ43} _WT-linker	pET46_sVP40 _{Δ43} _WT-linker-pep1	#7344	Mutagenesis
pET46_sVP40 _{Δ43} _N67C	pET46_sVP40 _{Δ43} _WT	#5344	Mutagenesis

Materials

pET46_sVP40 _{Δ43} _N67C-CCA	pET46_sVP40 _{Δ43} _WT	#5344; #5694; #5695	Mutagenesis
pET46_sVP40 _{Δ43} _L158A	pET46_sVP40 _{Δ43} _WT	#5950	Mutagenesis
pET46_sVP40 _{Δ43} _L213A	pET46_sVP40 _{Δ43} _WT	#7096	Mutagenesis
pET46_sVP40 _{Δ43} _R214A	pET46_sVP40 _{Δ43} _WT	#5451	Mutagenesis

8.3 Oligonucleotides and DNA

8.3.1 Primers

8.3.1.1 PCR primers

#3456_Seq_EBOV_5255_rev	CATGATATTTTTGGTTGGATC
#3457_Seq_EBOV_7004_rev	GACTCTGACCACTGATGTTT
#3458_Seq_EBOV_8301_rev	TCTTAATCTTCATCAAGTTTATCA
#3459_Seq_EBOV_9958_rev	GATGAAGAATGCCTCTTGG
#3461_Seq_EBOV_12948_rev	GTTCTTTAATCATTGGCAACG
#3462_Seq_EBOV_14901_rev	CAACTGTGCAAGTTATTTGG
#3463_Seq_EBOV_16799_rev	CACATTGTGTCTCGATCGAT
#3466_Seq_EBOV_4420_rev	CAGCCGAGGTAGGTTTT
#3467_Seq_EBOV_6145_rev	CCAAGTGGGATGGAAAATG
#3468_Seq_EBOV_10733_rev	TTGCCGATTTGTCGGGT
#3469_Seq_EBOV_12111_rev	CATCATGAACAAACCATGTT
#3471_Seq_EBOV_13929	CACAGGCACTTGTAACTTT
#3472_Seq_EBOV_15910	ATTGGGTCTGTAGATGAATT
#3473_Seq_EBOV_17753	GCGGATATAACTTAAATGTAA
#3474_Seq_EBOV_18795_rev	TGTAAGAAGGTATTTCTGAT
#4691_EBOV_hisVP40_aa44-326_SacI_r	GcagcaGAGCTCTTACTTCTCAATCAC
#4840_pCAGGS_mCheLink_SacINhel_r	tgctgctagcagctcGCTGCTGCCACCACCTCCGC
#4869_VP40mChe_KQIPIWLPLGVA_fwd	GGTGGTGGCAGCAGCgagctcAAGCAGATCCCCAT CTGGCTGCCTCTGGGAGTTGCTTAAGCTAGCgagca
#4832_VP40mChe_KQIPIWLPLGVA_rev	tgctgcGCTAGCTTAAGCAACTCCCAGAGGCAGCCA GATGGGGATCTGCTTgagctcGCTGCTGCCACCACC
#4870_VP40mChe_KQIPIALALGVA_fwd	GGTGGTGGCAGCAGCgagctcAAGCAGATCCCCATC GCTCTGGCTCTGGGAGTTGCTTAAGCTAGCgagca
#4833_VP40mChe_KQIPIALALGVA_rev	tgctgcGCTAGCTTAAGCAACTCCCAGAGCCAGAGCG ATGGGGATCTGCTTgagctcGCTGCTGCCACCACC
#4871_VP40mCheLPGIPDHPLRLLRIG_f	GGTGGTGGCAGCAGCgagctcCTCGGACCTGGAATCC CTGATCACCTCTGCGGCTGCTGAGAATCGGATAAGCTA GCgagca
#4834_VP40mCheLPGIPDHPLRLLRIG_r	tgctgcGCTAGCTTATCCGATTCTCAGCAGCCGAGAG GGTGATCAGGGATTCCAGGTCCGAGgagctcGCTGCTGCC ACCACC
#4872_VP40mCheLAPAIPDHPLRLLRIG_f	GGTGGTGGCAGCAGCgagctcCTCGCTCCTGCTATCCCT GATCACCTCTGCGGCTGCTGAGAATCGGATAAGCTAGC gagca
#4835_VP40mCheLAPAIPDHPLRLLRIG_r	tgctgcGCTAGCTTATCCGATTCTCAGCAGCCGAGAGG GTGATCAGGGATAGCAGGAGCGAGgagctcGCTGCTGCC ACCACC
#4873_VP40mChe_VQLPQYFTFDLT_fwd	GGTGGTGGCAGCAGCgagctcGTTACAGTGCCCCAGT ACTTCACCTTCGACCTGACATAAGCTAGCgagca
#4836_VP40mChe_VQLPQYFTFDLT_rev	tgctgcGCTAGCTTATGTCAGGTGGAAGGTGAAGTACT GGGGCAGCTGAACgagctcGCTGCTGCCACCACC

Materials

#4874_VP40mChe_VQLPQAFADLT_fwd	GGTGGTGGCAGCAGCgagctcGTTTCAGCTGCCCCAGGC TTTCGCTTTTCGACCTGACATAAGCTAGCgagca
#4837_VP40mChe_VQLPQAFADLT_rev	tgctgcGCTAGCTTATGTCAGGTGCGAAAGCGAAAGCCT GGGGCAGCTGAACgagctcGCTGCTGCCACCACC
#4907_pCAGGS_XhoIKpnI_flag_linker_f	GcagcaCTCGAGGGTACCATGgattacaaggatgacgac gataagGGCGGAGGCGGATCTG
#4913_SacI_SFDSTTAAIMLA_NheI_f	gcagcagagctcAGCTTTGATTCTACCACCGCCGCCAT CATGCTGGCCTAAGCTAGC
#4914_SacI_SFDSTTAAIMLA_NheI_r	gcagcaGCTAGCTTAGGCCAGCATGATGGCGGCGGT GGTAGAATCAAAGCTgagctc
#4915_SacI_SFAATAAAIMAA_NheI_f	gcagcagagctcAGCTTTGCAGCAACGCAGGCCGCCATCAT GGCAGCTTAAGCTAGC
#4916_SacI_SFAATAAAIMAA_NheI_r	gcagcaGCTAGCTTAAGCTGCCATGATGGCGGCTGCGTT GCTGCAAAGCTgagctc
#4691_EBOV_hisVP40_aa44-326_SacI_r	GcagcaGAGCTCTTACTTCTCAATCAC
5305_pETDuet_zflagVP40_XhoI_r	tgctgcCTCGAGTTACTTCTCAATCAC
5306_flagZ40_intrf1_36as_Nde	gcagcaCATATGGATTACAAGGATGACGACGATAAG TGGCTTCTCTAGGTGTC
5307_Z40_intrf1_36as_XhoI	tgctgcCTCGAGTTAATTGGTTGCCCTGCGAA
5308_flagZ40_intrf2_36as_Nde	gcagcaCATATGGATTACAAGGATGACGACGATAAGATGG TGAATGTCATATCG
5309_z40_intrf2_36as_XhoI	tgctgcCTCGAGTTAGGCCGTAGTTGAGTCAAA
5310_flagZ40_intrf3_40as_Nde	gcagcaCATATGGATTACAAGGATGACGACGATAAGGGCA AGGCAACCAATCCA
5311_Z40_intrf3_40as_XhoI	tgctgcCTCGAGTTATGGCGGAAGAACGAACTC
5312_flagZ40_intrf4_36as_Nde	gcagcaCATATGGATTACAAGGATGACGACGATAAGAACC AGGCTTCTCTCCAG
#5042_pETDuet-1_zVP40_BamHI_f	gcagcaGGATCCAGGCGGGTTATATTGCCT
#5155_pCAGGS_sGu_NotI_mycVP40_f	GcagcaGCGGCCGCTAAATGGAACAAAACTCATCTCA GAAGAGGATCTGAGAAGGGTCACTGTGCCG
#5156_pCAGGS_sGu_SacI_mycVP40_r	tgctgcGAGCTCTCACTTTTCACTGAGATA
#5871_pCAGGS_sGu_NotI_VP40_f	GcagcaGCGGCCGCTAAATGAGAAGGGTCACTGTGCCG
#6054_Seq_EBOV_812 rev	TCACTCCATCACGCTTCTTG
#6055_Seq_EBOV_639	TACAGGCCTTTGAAGCAGGT
#6227_Seq_EBOV_247	CACAAGTTCCACCTCAGAC
#6235_Seq_EBOV_2579	GAGGCTATGAATGAAGAGAA
#6238_Seq_EBOV_3477	GGTCTAAAGCCAGTTTATGA
#6241_Seq_EBOV_4372	GTAACCCAAACCAAAAGTGA
#6246_Seq_EBOV_6055	TATTGCAGTTACCTCGTGAT
#6249_Seq_EBOV_6938	ATTCGCAGTGAAGAGTTGTC
#6253_Seq_EBOV_8198	GCCAATGTGATTCTAACTCC
#6258_Seq_EBOV_9894	TAATGCGGAGGTCTGATAAG
#6261_Seq_EBOV_10862	ACGGATTGTTGAGCAGTATT
#6264_Seq_EBOV_11866	TGCCAACAGTTCTTAGATGA
#6267_Seq_EBOV_12853	TCTTGGTACAGTGTTACTTC
#6270_Seq_EBOV_13832	TCTTCCCCTTAGAGACTGAC
#6273_Seq_EBOV_14822	GGTGGAGCCTATGGTTTAGT
#6276_Seq_EBOV_15734	CTCTAAAAGGAGGACTCAAT
#6279_Seq_EBOV_16719	GTCGTTCCAGTCCTTTCTAA
#6282_Seq_EBOV_17692	CCATACTGGCTAAGTCATTT
#7044_EBOV_ab_2269	AAATTCAAGTACATGCAGAGCAAG

Materials

#7045_EBOV_bis_3717	TTAACGCCTATTTGCATTAAGACT
#7046_EBOV_ab_5758	TTGTAAACGTTGAGCAAAATTGTT
#7047_EBOV_bis_6148	CCTCTAATGGACGGATTACTAAAT
#7048_EBOV_ab_10477	GATGTGCATGATTGCCCTAATATA
#7049_EBOV_bis_11355	TTTTTGGGCGATATTAGATTGTAT
#7175_EBOV_ab_3561	GAGAATAGATTTGTTACATTGGAT
#7176_EBOV_bis_4374	CAATGAAGCCAATGTTTGTACTAC

8.3.1.2 Primers for mutagenesis

#4908_SUDAN-VP40_Muta_W95A:	gtcctcatgaacaaatccctattgcgttgccactcgga
#4910_SUDAN-VP40_Muta_L117A:	ctcaacaacagcagcaattatggccgcatcttatacgatcacc
#4946_Muta_EBOV_VP40_C311S:	catggtaatcacacaggatagtgacacgtgtcattctc
#4947_Muta_EBOV_VP40_C314S:	cacaggattgtgacacagtcattctcctgcaagt
#4955_Muta_EBOV_VP40_C311SC314S	ggtaatcacacaggatagtgacacagtcattctcctgcaa
#4956_Muta_EBOV_VP40_W95A	ccaaagtctaataagcaattccaattgcgcttctctaggtgt
#5079_pETDuet_frameshift_Insert	atcatcaccacagcgcaggatccaggcgg
#5344_Muta_sVP40_Asn67Cys:	tgatcacacaagtcatacccatgaggagtgccctca
#5424_Muta_SUBOV_VP40_C314S:	ctggcaggtgaatgactatcatcataatctggtgttatg
#5425_Muta_SUBOV_VP40_C320S:	gattgtcattcacctgccagtagctcttatctcagt
#5451_Muta_pET46_sVP40_R214A:	tcatttcacccaaagctggcaccgttctactccagg
#5694_Muta_SUBOV_VP40_C314A:	ctggcaggtgaatgagcatcatcataatctggtgttatgacca
#5695_Muta_SUBOV_VP40_C320A:	gattgtcattcacctgccagtgccctcttatctcagtgaaaagt
#5947_Muta_pCAGGS_sVP40_R214A:	cctggaagtagaacgggtgccagcttgggtgaaatga
#5950_Muta_pCAGGS_sVP40_L158A:	ggtagcacaactcttgagcgaagcctggttcccat
#5952_Muta_pCAGGS_sVP40_L213A:	tggaagtagaacgggtctcgccttgggtgaaatgaaagc
#7096_Muta_pET46_sVP40_L213A:	ctctcatttcacccgaacgcgagaccgttctacttcc
#7212_Muta_pCAGGS_sVP40_L288A:	gacagcctataattcctgtcgcactcccaaaatacattgggc
#7344_Muta_pET25_sVP40_GS_taa	ggtggtggcagctaaagcgagctcagc

8.3.1.3 Primer for sequencing

#1233_pCAGGS-forw	CCTTCTTCTTTTCTACAG
#1234_pCAGGS-rev	CCTTTATTAGCCAGAAGTCAG
#3456_Seq_EBOV_5255_rev	CATGATATTTTTGGTTGGATC
#4911_Seq-Primer_SUDAN-VP40_f	AGAAGGGTCACTGTGCCG
#4912_Seq-Primer_SUDAN-VP40_r	CTTTTCACTGAGATAAGA
#4933_T7_promotor	TAATACGACTCACTATAGGG
#4934_T7_terminator	TGCTAGTTATTGCTCAGCGGT
#5017_EBOV_VP40_qPCR_fwd	TTGAAGCTATGGTGAATG
#5018_EBOV_VP40_qPCR_rev	GCAGTGGTTGGGTGATCA
#5040_SUBOV_VP40_qPCR_fwd	AACCCCTCGTCAGAGTG
#5041_SUBOV_VP40_qPCR_rev	CACAATTGTCTGGATTTT
#6055_Seq_EBOV_639	TACAGGCCTTTGAAGCAGGT
#6227_Seq_EBOV_247	CACAAGTTCCACCTCAGAC
#6238_Seq_EBOV_3477	GGTCTAAAGCCAGTTTATGA
#6246_Seq_EBOV_6055	TATTGCAGTTACCTCGTGAT
#6249_Seq_EBOV_6938	ATTCGAGTGAAGAGTTGTC
#6253_Seq_EBOV_8198	GCCAATGTGATTCTAACTCC
#6258_Seq_EBOV_9894	TAATGCGGAGGTCTGATAAG
#6261_Seq_EBOV_10862	ACGGATTGTTGAGCAGTATT

Materials

#6264_Seq_EBOV_11866	TGCCAACAGTTCCTTAGATGA
#6267_Seq_EBOV_12853	TCTTGGTACAGTGTTACTTC
#6270_Seq_EBOV_13832	TCTTCCCCTTAGAGACTGAC
#6273_Seq_EBOV_14822	GGTGGAGCCTATGGTTTAGT
#6276_Seq_EBOV_15734	CTCTAAAAGGAGGACTCAAT
#6279_Seq_EBOV_16719	GTCGTTCCAGTCCTTTCTAA
#6282_Seq_EBOV_17692	CCATACTGGCTAAGTCATTT
#7044_EBOV_ab_2269	AAATTCAAGTACATGCAGAGCAAG
#7046_EBOV_ab_5758	TTGTAAACGTTGAGCAAATTGTT
#7048_EBOV_ab_10477	GATGTGCATGATTGCCCTAATATA
#7175_EBOV_ab_3561	GAGAATAGATTTGTTACATTGGAT
#7176_EBOV_bis_4374	CAATGAAGCCAATGTTTGTACTAC

8.3.2 Nucleic acids and nucleotides

O'GeneRuler 1kb Plus DNA Ladder	Thermo Fisher Scientific
dNTPs	Thermo Fisher Scientific

8.4 Consumables

2 ml cryotubes	Corning®
25 cm ² , 75 cm ² , 175 cm ² cell culture flasks	Greiner bio-one
ACQUITY UPLC BEH C18 Column, 130Å, 1.7 µm, 1 mm X 100 mm	Waters
ACQUITY UPLC BEH C18 VanGuard Pre-column, 130Å, 1.7 µm, 2.1 mm X 5 mm	Waters
Blotting paper	Whatman
Centrifuge tubes, Ultra-Clear™ for SW41, SW60 and SW32	Beckmann
Chromabond C18WP spin columns	Macherey-Nagel
CM5 sensor chip	Cytiva
cOmplete™ Protease Inhibitor Cocktail	Sigma-Aldrich
Coverslips, Ø 12 mm	Menzel
Crystallization plates (pre-greased)	Molecular dimensions
Cuvettes	Roth
Guard column 2 mm x 2 cm	IDEX
MRC 2 well plates	Swissci
MicroAMP® Fast 96-well reaction plate (0.1 ml)	Applied Biosystems
MicroAMP™ Optical Adhesive Film	Applied Biosystems
Nitrocellulose	Whatman
Object slide 76 x 22 mm	Menzel
Parafilm	Pechiney Plastic
PCR-tubes, 0.2 ml	Biozym
Petri dishes	Sarstedt
Phosphotungstic acid	SERVA FEINBIOCHEMICA GmbH & Co. KG
Pierce™ Protein Concentrator PES, 10K MWCO, 5-20 mL	Fisher Scientific
Pipette tips 0.1-1 µl, 10-100 µl, 100-1000 µl TipOne (with and without filter)	Starlab
Polypropylene reaction tubes 15 / 50 ml	Greiner bio-one
Reaction tube 1.5 ml	Sarstedt

Materials

Reaction tube 2 ml	Eppendorf
Reaction tubes (screw top)	Sarstedt
Spin concentrators Amicon Ultra-15 (3 – 30 kDa MWCO)	Fisher Scientific
Stericup™ Sterile Vacuum Filter Units	Merck Millipore
Zeba Spin 7K MWCO Desalting columns	Fisher Scientific

8.5 Chemicals

4',6-Diamidino-2-phenylindol (DAPI)	Sigma-Aldrich
5,5'-dithiobis(2-nitrobenzoic acid) (DTNB) (Ellman's reagent)	Roth
Acetic acid, glacial	Merck
Acetone	Honeywell
Acetonitrile	Thermo Scientific
Acrylamide, 40%	Roth
Agarose	VWR Chemicals
Ammonium persulfate	Biorad
Ampicillin	Serva
Bacto agar	Beckton Dickinson
Bromphenol blue	Sigma Aldrich
Chariot™	Active Motif
Coomassie R250	Thermo Scientific
Deuterium oxide	Deutero
Dimethylsulfoxide	Roth
Dithiothreitol DTT	Roth
Dulbecco's Modified Eagle Medium (DMEM)	Gibco
Ethylenediaminetetraacetic acid (EDTA)	Roth
Ethanol	Roth
Ethidium bromide	Roth
Ethylene glycol	Hampton Research
EZ-link™ iodoacetyl-PEG2 biotin	Thermo Fisher Scientific
Fetal calf serum (FCS)	PAN Biotech GmbH
Fluorprep	BioMérieux
Formic acid	AppliChem
GelGreen Nucleic Acid Stain 10000X Water	Merck
[Glu1]-Fibrinopeptide B standard	Waters
Glycerol	Roth
Glycin	Roth
[Glu1]-Fibrinopeptide B Standard	Waters
Glutamine	Invitrogen
GSH	Sigma Aldrich
GSSG	Sigma Aldrich
Guanidine hydrochloride	Carl Roth
HEPES	Sigma Aldrich
Hydrochloric acid	Carl Roth
Imidazole	Roth
Iodoacetamide	Sigma Aldrich
Isopropyl-beta-D-thiogalactopyranoside (IPTG)	Roth
Isopropanol	Roth

Materials

2-Mercaptoethanol	Sigma Aldrich
Methanol	Roth
Magnesium chloride (MgCl ₂)	Roth
Monosodium citrate	Carl Roth
Ni-NTA agarose	Qiagen
Nicotinamide adenine dinucleotide phosphate	Biomol
Neocuproine	Sigma Aldrich
OPTI-MEM®	Gibco
PageRuler™ Plus Prestained Protein Ladder	Thermo Fisher Scientific
Paraformaldehyde	Roth
PEG400	Hampton Research
Penicillin/Streptomycin	Invitrogen
Phosphoric Acid	Roth
Phosphotungstic acid	Serva
Porcine Pepsin	Sigma Aldrich
POROS 20 AL Aldehyde activated resin	Thermo Scientific
Potassium chloride	Roth
Potassium dihydrogen phosphate	Roth
Sequencing Grade Modified Trypsin	Serva
Skim milk powder	Saliter
Sodium ascorbate	Sigma Aldrich
Sodium azide (NaN ₃)	Merck
Sodium chloride	Roth
Sodium chloride	VWR
Sodium cyanoborohydride	Sigma Aldrich
Sodium dodecyl sulfate (SDS)	Merck
Sodium sulfate	Roth
Sucrose	Roth
Surfactant P20	Cytiva
Sypro Orange Protein Gel Stain	Sigma Aldrich
Tetramethylethylenediamine (TEMED)	BioRad
TransIT®-LT1 Transfection Reagent	Mirus Bio
Trichloroacetic acid (TCA)	Roth
Tris	Roth
Triton X-100	Roth
Trypsin/EDTA	Gibco
Tryptone/Peptone	Roth
Tween20	Sigma Aldrich
Yeast Extract	Roth

8.6 Buffers, media and gels

8.6.1 Buffers and solutions

Ampicillin	100 mg ad 1 ml dH ₂ O
Biotin Switch nitrosation buffer	50 mM Tris 1 mM EDTA 0.2 mM neocuproine; pH 7.4
Biotin Switch blocking buffer	50 mM Tris 8 M urea 1 mM EDTA

Materials

	0.01 mM neocuproine; pH 8
Biotin Switch labeling buffer	50 mM Tris 4 M urea 1 mM EDTA 0.01 mM neocuproine; pH 8
Blocking buffer for Western blots	10% (w/v) skim milk powder in PBS _{def}
Blocking buffer for IFA	2% BSA 0,2% Tween® 20 5% glycerin 0.05% NaN ₃ in PBS _{def}
Coomassie staining solution	0.25 g Coomassie Brilliant Blue ad 500 ml dH ₂ O
Coomassie destaining solution	8% acetic acid in H ₂ O
DTNB buffer	100 mM potassium phosphate 2 mM EDTA, pH 7.4
Fixing solution for IFA	4% Paraformaldehyd in PBS _{def}
GR assay buffer	20.5 mM KH ₂ PO ₄ 26.5 mM K ₂ HPO ₄ 1 mM EDTA 200 mM potassium chloride, pH 6.9
HDX-MS quench buffer	400 mM KH ₂ PO ₄ /H ₃ PO ₄ 2 M guanidine-HCl, pH 2.2
PBS _{def} (autoclaved)	0.2 g KCl 0.2 g KH ₂ PO ₄ 1.15 g Na ₂ HPO ₄ 8 g NaCl ad 1000 mL dH ₂ O
Permeabilization buffer for IFA	0.1% Triton X-100 in PBS _{def}
Ponceau staining solution	2 g Ponceau 3 g TCA ad 100 ml dH ₂ O
Quenching solution for IFA (autoclaved)	100 mM glycin in PBS _{def}
Sample buffer (4x)	10 ml Mercaptoethanol 200 mg Bromphenolblue 20 ml glycerine 10 ml 1 mol/l Tris-HCl, pH 6.8 4 g SDS ad 50 ml dH ₂ O
SDS-PAGE running buffer	144 g glycin 30 g Tris-Base 10 g SDS ad 100 mL dH ₂ O
SDS-PAGE Stacking gel buffer	0.4% SDS 1 M Tris/HCl, pH 6.8
SDS-PAGE separation gel buffer	0.4% SDS 1.5 M Tris/HCl, pH 8.8
SPR buffer	10 mM PBS pH 7.4 0,005% Surfactant P20 5% DMSO

Materials

Sucrose solution (20%)	20% sucrose (w/v) in TNE buffer
TAE buffer (50x), pH 8.0	100 mL 0.5 M EDTA, pH 8.0 57,1 g acetic acid 242 g Tris-Base ad 1000 ml dH ₂ O
TNE buffer (1x, autoclaved)	1 mM EDTA, pH 8.0 150 mM NaCl 10 mM Tris/HCl, pH 7.5
Transfer buffer (1x) for Western blots	100 mL ethanol 144 mg glycine 300 mg Tris ad 1000 mL dH ₂ O
TrxR assay buffer	100 mM potassium phosphate 2 mM EDTA, pH 7.4
VP40 buffer	25 mM Tris 300 mM NaCl, pH 8
VP40 lysis buffer	10 mM imidazole in VP40 buffer
VP40 wash buffer	20 mM imidazole in VP40 buffer
VP40 elution buffer	250 mM imidazole in VP40 buffer
VP40 crystallization buffer	100 mM HEPES, pH 7.2 40 mM MgCl ₂
VP40 crystallization cryo buffer	20% (v/v) ethylene glycol in VP40 crystallization buffer
Wash buffer for Western blots	0.1% Tween [®] 20 in PBS _{def}

8.6.2 Media

8.6.2.1 Bacterial media

LB medium	10 g NaCl 5 g yeast extract 10 g peptone/tryptone ad 1000 ml dH ₂ O
LB agar	3.75 g Bacto agar ad 250 ml LB medium
NZY ⁺ medium	5 g NaCl 5 g yeast extract 10 g NZ amine ad 1000 ml dH ₂ O
SOB medium	20 g peptone 5 g yeast extract 0.58 g NaCl 0.19 g KCl 10 ml 1 mol/l MgCl ₂ 10 ml 1 mol/l MgSO ₄ ad 1000 ml dH ₂ O

8.6.2.2 Media for mammalian cells

DMEM +++ (3/10%)	500 mL DMEM 5 mL L-Glutamin (200 mM) 5 mL Penicillin-Streptomycin (5000 U/ml) 15/50 mL FCS
------------------	---

Materials

DMEM 4% PFA	4% paraformaldehyde in DMEM
-------------	-----------------------------

8.6.3 Gels

4% Stacking gel (Western blot, for four gels)	1.25 mL SDS-PAGE stacking gel buffer 750 μ L acrylamide 2.9 mL H ₂ O 50 μ L APS 10 μ L TEMED
12% Separation gel (Western blot, for four gels)	2.5 mL SDS-PAGE separation gel buffer 3.3 mL acrylamide 4 mL H ₂ O 100 μ L APS 10 μ L TEMED
1% Agarose gel	1% (w/v) agarose in TAE buffer

8.7 Antibodies

8.7.1 Primary antibodies

Chicken α -NP	Institute for Virology, Marburg
Chicken α -VP40 Zaire	Institute for Virology, Marburg
Goat α - Zaire serum	Institute for Virology, Marburg
Goat α - Sudan serum	Institute for Virology, Marburg
Mouse α -biotin	Santa Cruz Biotechnology
Mouse α -his	Thermo Fisher Scientific
Mouse α -tubulin	Sigma Aldrich
Mouse α -VP40 (Zaire)	Sino Biological
Mouse α -VP40 (Zaire) 2C4	Institute for Virology, Marburg
rabbit α -VP40 (Sudan)	Sino Biological
Mouse α -flag	Sigma Aldrich
Rabbit α -myc	Santa Cruz Biotechnology

8.7.2 Secondary antibodies

Donkey α -chicken 800	Li-Cor
Donkey α -mouse 680	Li-Cor
goat α -mouse 488	Li-Cor
goat α -mouse 800	Li-Cor
goat α -mouse 680	Invitrogen
goat α -rabbit 488	Li-Cor
goat α -rabbit 680	Li-Cor

8.8 Enzymes

Alkalische Phosphatase, Calf Intestinal Phosphatase (CIP) (10 U/ μ l)	New England Biolabs
<i>DpnI</i>	Agilent
Restriction enzymes	New England Biolabs
T4 DNA Ligase (4 U / μ l)	New England Biolabs

Materials

8.9 Kits

Crystallization Screens	Hampton Research
Beetle-Juice Kit	PJK GMBH
CellTiter 96® Aqueous One Solution Cell Proliferation Assay	Promega
E.Z.N.A.® Cycle Pure Kit	Omega Bio-Tek Inc.
E.Z.N.A.® Plasmid DNA Mini Kit	Omega Bio-Tek Inc.
E.Z.N.A.® Fast Filter Plasmid DANN Maxi Kit	Omega Bio-Tek Inc.
E.Z.N.A.® Gel Extraction Kit	Omega Bio-Tek Inc.
QIAamp Viral RNA Mini Kit	Qiagen
QuikChange Multi Site-Directed Mutagenesis Kit	Agilent
QuikChange Lightning Multi Site-Directed Mutagenesis Kit	Agilent
Renilla-Juice Kit	PJK GMBH
Transcriptor One-Step RT-PCR Kit	Roche
Z-Competent <i>E. coli</i> Transformation Kit™ and Buffer Set	Zymo Research

8.10 Peptides

Sequence	mol weight calculated [g/mol]	Supplier
Ac-SFDSTTAAIMLA-amide	1227	Pepmic
Ac-GVADQKTYSFDDSTTAAIMLASIY-amide	2341	Pepmic
Ac-SFDSTARRIMAA-amid	1594.61	Steinmetzer lab
Ac-SFDSTAARIMAA-amid	1395,48	Steinmetzer lab
Ac-KQIPIWLPLGVA-amide	1489.75	Steinmetzer lab
KQIPIWLPLGVA non-ac	1447.71	Steinmetzer lab
Ac-flag-KQIPIWLPLGVA-amide	2712.76	Steinmetzer lab
Ac-KQIPIALPLGVA-amid	1489,75	Steinmetzer lab
Ac-LGPGIPDHPLRLLRIG-amide	1993.19	Steinmetzer lab
LGPGIPDHPLRLLRIG non-ac	1951.15	Steinmetzer lab
Ac-flag-LGPGIPDHPLRLLRIG-amide	3102.18	Steinmetzer lab
Ac-LGAGIPDHPLRLLRIG-amid	3102.18	Steinmetzer lab
Ac-AGAGIPDHPLRLLRIG-amid	2039.09	Steinmetzer lab
VQLPQYFTFDLT non-ac	1470.69	Steinmetzer lab
Ac-flag-VQLPQYFTFDLT-amide	2735.74	Steinmetzer lab

8.11 Equipment

Äkta	Cytiva
ACQUITY H-Class HPLC-system (with MassPrep column)	Waters
ACQUITY UPLC BEH C18 1.7 µm 1 x 100 mm column	Waters
ACQUITY UPLC BEH C18 VanGuard Pre-column, 130 Å, 1.7 µm, 2.1 mm x 5 mm	Waters
ACQUITY UPLC M-Class System with HDX Technology	Waters
Aurora column (25cm x 75µm) C18 RP column	IonOpticks
Biacore 3000 system	Cytiva
Centrifuge 5430-R (tabletop)	Eppendorf
Centrifuge Avanti J-25	Beckman Coulter

Materials

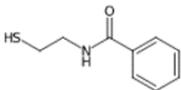
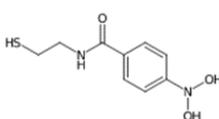
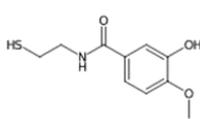
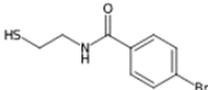
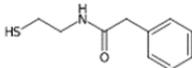
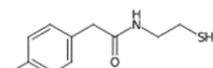
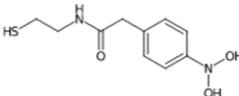
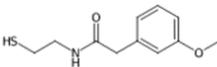
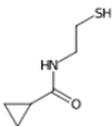
Centrifuge Avanti J-26 XP	Beckman Coulter
ChemiDoc™ XRS+	Bio-Rad
Chromatography column HiLoad 16/60	Cytiva
Confocal microscope Leica TCS SP5	Leica
Cryo caps (plain)	Molecular dimensions
Crystallization loops	Molecular dimensions
DNA gel chambers	bsb11
Electrophoresis chamber	Bio-Rad
Eppendorf Reference® Pipetten (0,1-2,5 µl, 0,5-10 µl, 10-100 µl, 100-1000 µl)	Eppendorf
Fisherbrand™ Safeview BLUE Light Transilluminator 15.4 x 15.4cm	Thermo Fisher Scientific
G2-Si HDMS mass spectrometer with ion mobility separation	Waters
Heating block Neoblock 1	neoLab®
Horizontal shaker Gio Gyrotory® Shaker	New Brunswick Scientific CO., Inc.
Horizontal shaker TR	Infors AG
Incubator HERAcell150	Thermo Fisher Scientific
Heraeus Multifuge 3S-R	Heraeus
Honeybee 963 crystallization robot (MarXtal)	Digilab
JEM-1400 transmission electron microscope	JEOL
MAR345 Image Plate Detector	MAR-Incoatec
Microfluidizer	Microfluidics, Bange lab, Philipps-University Marburg
Mounted cryo loops 20 micron	Hampton Research
Fluorescence microscope DMI6000 CS	Leica
Luminometer Centro LB 960	Berthold
NanoElute HPLC system	Bruker Daltonics
Odyssey CLx Infrared Imaging System	LI-COR Biosciences
Overhead shaker	Labortechnik Fröbel GmbH
Pipetting aid Pipetboy	Integra Bioscience
pH meter pHenomenal	VWR
Photometer WPA Lightwave II	Biochrom WPA
Robotic autosampler	LEAP Technologies
Rotor Ultracentrifuge SW32, SW41, SW60	Beckmann Coulter
Rotor JA 25.50 and JLA 8.1000	Beckmann Coulter
Rotor SW41, SW60 and SW32	Beckmann Coulter
Sample loop 5ml PEEK	Cytiva
Safety Cabinet	BDK
Sonifier Branson Ultrasonics S-450	BANDELIN electronic
Spectrophotometer NanoDrop Lite	Peqlab
StepOne™ Real-Time PCR System	Applied Biosystem
Stereomicroscope SZ60	Olympus
Synapt G2Si mass spectrometer	Waters
Tecan infinite M200 multiplate reader	Thermo Fisher
TemCam-F416 camera	TVIPS
Thermocycler Primus 25	Peqlab
TimsTOF Pro mass spectrometer	Bruker Daltonic
Trans-Blot Turbo Transfer System	Bio-Rad
Ultracentrifuge Optima™ L-100K / -80XP	Beckman Coulter

8.12 Software

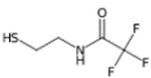
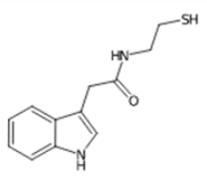
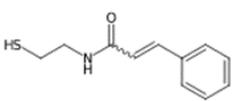
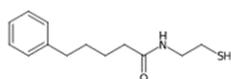
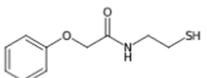
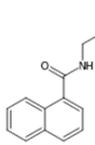
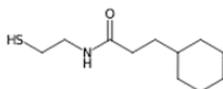
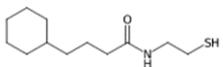
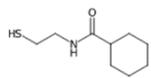
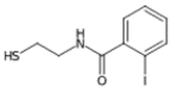
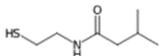
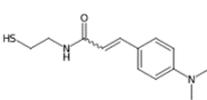
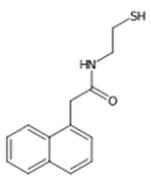
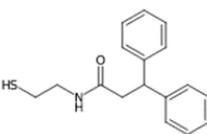
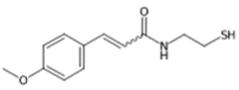
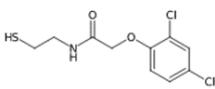
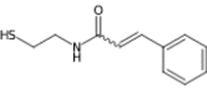
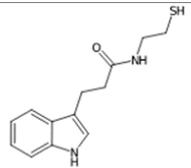
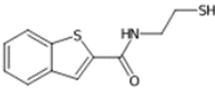
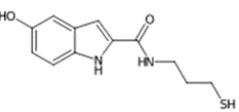
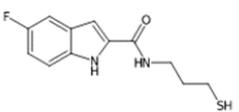
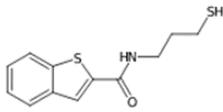
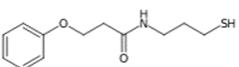
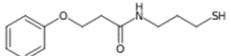
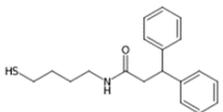
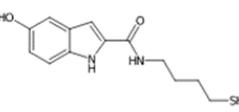
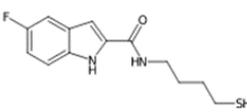
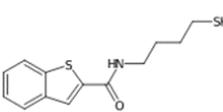
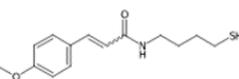
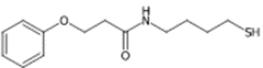
Adobe Illustrator CS6	Adobe Systems
autoPROC	https://www.globalphasing.com/autoproc [313]
BUSTER	https://www.globalphasing.com/buster [37,290]
CCP4i2	Potterton <i>et al.</i> 2018
BioEdit	http://www.mbio.ncsu.edu/BioEdit/bioedit.html
BLAST (Basic local alignment search tool, NCBI)	www.ncbi.nlm.nih.gov/BLAST/
Chimera 1.15	Pettersen <i>et al.</i> 2004
Crystallization plate documentation Imager	Rock Formulatrix
Coot 0.8.9.2	Emsley and Cowtan 2004
DA+ SLS Data acquisition GUI	Wojdyla <i>et al.</i> 2018
DynamX 3.0.0	Waters
ImageJ 1.52p	ImageJ Development Team
Inkscape 1.0.1	Inkscape, 2020
MassLynx instrument software with MaxEnt1 extension	Waters
Li-Cor Odyssey Image Studio 5.2	LI-COR Biosciences,
MikroWin2010	Labsis Laborsysteme GmbH
PEAKS	Bioinformatics Inc
Phenix	Liebschner <i>et al.</i> 2019
PLGS 3.0.1	Waters
Prism 8.0	GraphPad
ProteinLynx Global SERVER	PLGS, Waters
Proteome Discoverer 2.4	Thermo Scientific
Pymol 2.2.0	Schrödinger, LLC, 2020
Sequest search engine with Bionic node	Proteinmetrics
UCLA Diffraction Anisotropy Server	Strong <i>et al.</i> 2006
Unicorn 5.31	Cytiva
XDS	Kabsch 2010

8.13 Fragments

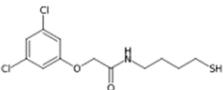
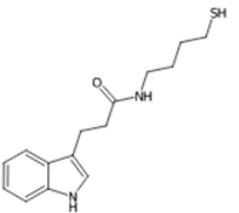
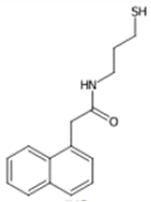
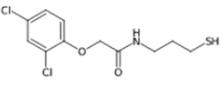
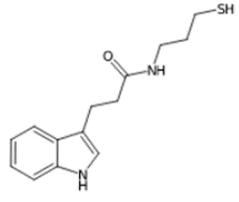
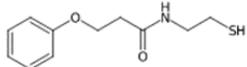
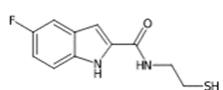
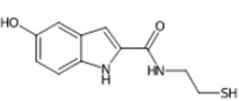
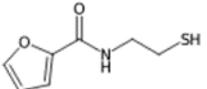
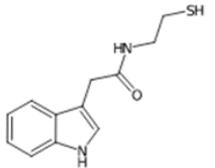
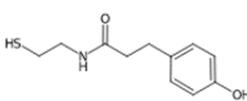
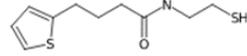
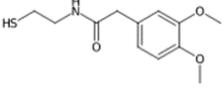
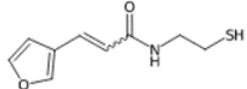
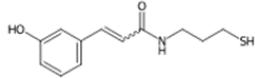
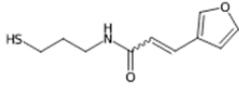
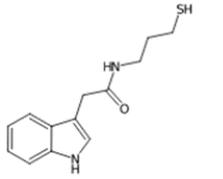
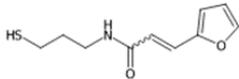
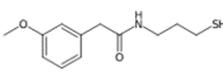
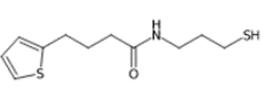
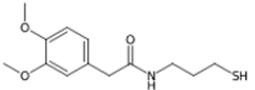
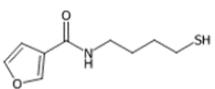
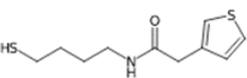
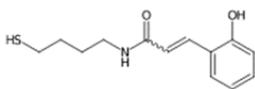
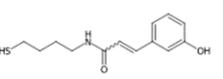
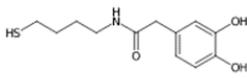
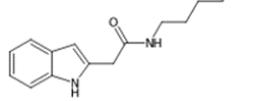
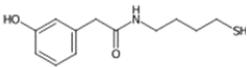
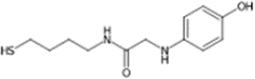
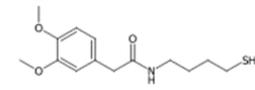
8.13.1 Thiol-containing fragments for DT

#	Structure	#	Structure	#	Structure
SB095 181.26		SB096 226.26		SB097 227.28	
SB099 260.15		SB100 195.29		SB101 213.28	
SB102 240.28		SB103 225.31		SB104 145.22	

Materials

SB106 173.16		SB107 234.32		SB109 207.3	
SB113 237.37		SB114 211.28		SB115 231.32	
SB119 215.36		SB120 229.39		SB123 187.31	
SB124 307.15		SB127 161.27		SB128 250.37	
SB226 245.35		SB227 285.41		SB228 237.32	
SB229 280.17		SB231 207.3		SB232 248.35	
SB233 237.34		TK094 250.32		TK095 252.31	
TK096 251.37		TK097 239.34		TK098 273.40	
TK099 313.47		TK100 264.35		TK101 266.34	
TK102 265.40		TK103 265.38		TK104 253.37	

Materials

TK105 308.23		TK106 276.4		TK107 259.37	
TK108 294.20		TK109 262.38		TK110 225.31	
TK111 238.29		TK112 236.29		TK135 171.22	
TK136 234.32		TK140 225.31		TK141 229.37	
TK142 255.34		TK147 197.26		TK148 237.32	
TK149 211.28		TK150 248.35		TK157 211.28	
TK158 239.34		TK162 243.39		TK163 269.37	
TK164 199.27		TK165 229.37		TK169 251.35	
TK170 251.35		TK171 255.34		TK173 262.38	
TK174 239.34		TK175 254.35		TK176 283.39	

8.13.2 Fragments for crystal soaking

The following fragments of the Jena FragXtal Screen were all kindly provided by the Klebe lab, Institute of Pharmaceutical Chemistry, Philipps-University Marburg.

ID	Structure	CAS-ID
1	Benzoic acid, 3-methyl-, hydrazide	13050-47-0
2	1H-Isoindol-3-amine	76644-74-1
3	Thiourea, N - (4- bromo- 2-chlorophenyl) -N '- methyl-	838813-44-8
4	1H - Azepine, hexahydro- 1- (6-methyl- 4- pyrimidinyl) -	1055196-50-3
5	4- Morpholineethanamine , β -ethyl- β - methyl-	914206-26-1
6	1, 3- Benzodioxole- 5-methanamine, N -cyclopentyl-	114413-77-3
7	1- Piperidineacetamide, N -1, 3-benzodioxol- 5- yl-	487006-82-6
8	1,3,5-Cycloheptatrien-1-amine, N-(1-methylethyl)-7-[(1-methylethyl)imino]-	3337-06-2
9	2H - Benzimidazol- 2- one, 1, 3-dihydro- 5- [(2-pyridinylmethyl) amino] -	876898-01-0
10	Ethanone, 2-amino-1-(4-bromophenyl)-,hydrochloride (1:1)	5467-72-1
11	Imidazo[2, 1-c] [1, 2, 4] triazine,3- (3, 4- difluorophenyl) - 1, 4, 6, 7- tetrahydro-	219929-40-5
12	Benzeneacetic acid, 3-chloro- α -hydroxy-	61008-98-8
13	1H -Pyrazole- 5- methanamine,1, 3- dimethyl-	499770-63-7
14	2- Furanmethanamine, N -methyl- 5- (3- pyridinyloxy) -	859850-77-4
15	Benzoic acid, 2- hydroxy- ,compd. with N ,N -diethylethanimidamide (1:1)	301843-27-6
16	Acetamide, 2-[(cyanomethyl) methylamino] -N- (6- methyl- 2- pyridinyl) -	1311649-76-9
17	Benzoic acid, 3-(dimethylamino) - , hydrazide	205927-64-6
18	Benzeneacetic acid, 4- fluoro-, hydrazide	34547-28-9
19	Pyrrolidine, 2- (2, 3- dihydro-1, 4- benzodioxin- 6- yl) -	524674-08-6
20	1H -Pyrazole- 4- carbonitrile, 5-amino- 3- (1- pyrrolidinyl) -	1119391-48-8
21	1-Piperidinecarboximid amide,hydriodide (1:1)	102392-91-6
22	3- Isoxazolecarboxamide , 5-methyl-N - (2, 2, 2-trifluoroethyl) -	878595-52-9
23	1- Piperidineacetamide, 4-methyl-N - [3- (1- methylethyl)-5- isoxazolyl] -	923192-35-2
24	2- Thiophenebutanamide, γ -oxo-N - [1- (3- pyridinyl) ethyl] -	923810-83-7
25	3- Pyridinecarbonitrile, 4, 6-dimethyl- 2- [[3- (4-morpholinyl) propyl] amino]-	831183-48-3
26	2- Thiazolamine, 4- methyl- 5-(1- methyl- 1H - imidazol- 2- yl) -	929975-79-1

Materials

27	1, 3, 5- Triazine- 2, 4- diamine, 6- [1- (hexahydro- 1H- azepin- 1-yl) ethyl] - N2, N2- dimethyl-	923208-22-4
28	3-Pyridineethanamine	20173-24-4
29	3-Thiophenecarboximidamide,hydrochloride (1:1)	51528-02-0
30	2- Furanmethanamine, 5-(trifluoromethyl)-	868755-68-4
31	Methanone, 1- piperidiny- 3-piperidiny-	845885-83-8
32	Benzoic acid, 4-(aminomethyl)-, methylester, hydrochloride (1:1)	6232-11-7
33	Benzenecarboximidami de, 4-(trifluoromethyl)-	38980-96-0
34	Guanidine, N - (4- ethoxy- 8-methyl- 2- quinazoliny)-	923803-92-3
35	1H-Indole-3-ethanamine, N-[(1-methyl-1H-pyrrol-2-yl)methyl]-	289487-79-2
36	Benzenemethanamine, 2-chloro-N - methyl- 5- nitro-	1177318-36-3
37	1H - Imidazole- 2-methanamine, N , 1- dimethyl-	1049717-38-5
38	Imidazo[1,2-a]pyridine-2-acetamide, N-phenyl-	1090352-33-2
39	Cyclohexanecarboxamide, N-[(tetrahydro-2-furanyl)methyl]-	401597-32-8
40	4(3H)-Quinazolinone, 2-[[[1-cyclopropylethyl)methylamino]methyl]-	1090518-66-3
41	3-Pyridinecarboxylic acid, 6-(dimethylamino)-	82846-28-4
42	3-Furancarboxamide, 2,5-dimethyl-N-4-pyridiny-	501442-73-5
43	1H-Pyrazole-4-acetamide,1,3,5-trimethyl-N-2-pyridiny-	1171575-61-3
44	1-Isoquinolineacetamide, 2-acetyl-1,2-dihydro-N,Ndimethyl-	1090496-96-0
45	Benzenemethanol, α -(aminomethyl)-4-fluoro-	456-05-3
46	Acetamide, N-[3-(aminomethyl)phenyl]-	96783-68-5
47	Phenol, 5-(aminomethyl)-2-methoxy-, hydrochloride (1:1)	89702-89-6
48	4-Pyridinemethanamine, N-(2,3-dihydro-1,4-benzodioxin-6-yl)-	774553-49-0
49	Carbamimidothioic acid, (2-chlorophenyl)methyl ester	14122-38-4
50	6H-Purin-6-one, 1,9-dihydro-8-(1-piperidiny)-	312923-57-2
51	Furo[2,3-d]pyrimidine-3(4H)-propanamine, 4-iminoN,N,5,6-tetramethyl-	929868-45-1
52	1H-Imidazole-1-acetamide, N-(4-methylcyclohexyl)-	1118862-09-1
53	(1S,2S,3S,4R,5R)-2-amino-4-(phenylthio)-6,8-dioxabicyclo[3.2.1]octan-3-o	1212574-75-8
54	3-Oxetanamine, 3-[[5-(1,1-dimethylethyl)-3-isoxazolyl]methyl]-	1371533-72-0
55	1,2-Cyclopentane-1,2-diol, 3-amino-4-(3-methylphenyl)-,(1R,2S,3R)-	1421058-66-3
56	1,4-Oxazepin-6-ol, hexahydro-, (6R)-	1022915-33-8
57	No name assigned in Scifinder	not available
58	1, 2, 4- Oxadiazole, 3- (1-methylethyl) - 5- (2-pyrrolidiny) -	1185170-55-1
59	3H - Oxazolo[3, 4-a] pyrazin- 3-one, hexahydro-	1020349-31-8
60	1,2-Pyrrolidinedicarboxylic acid, 1-(phenylmethyl) ester,(2S)-	1148-11-4

Materials

61	3-Pyridinecarboxamide	98-92-0
62	L-Phenylalanine, methyl ester, hydrochloride (1:1)	7524-50-7
63	1H-Purine-2,6-dione, 3,7-dihydro-3,7-dimethyl-	83-67-0
64	6H-Purin-6-one, 2-amino-1,9-dihydro-	73-40-5
65	Benzoic acid, 4-nitro-	62-23-7
66	2(3H)-Benzoxazolone, 5-chloro-	95-25-0
67	Pentanoic acid, 5-amino-	660-88-8
68	Imidazo[4,5-d]imidazole-2,5(1H,3H)-dione, tetrahydro-	496-46-8
69	4- Pyrimidinecarboxylic acid,5- amino- 1, 2, 3, 6- tetrahydro-2, 6- dioxo-	7164-43-4
70	L-Histidine, 1-methyl-	332-80-9
71	1H-Purine-2,6-dione, 3,7-dihydro-1,3,7-trimethyl-	58-08-2
72	L-Homoserine, O-[(aminoiminomethyl)amino]-	543-38-4
73	1H - Thieno[3, 4-d]imidazole- 6-pentanoic acid, 2- amino-3a, 4, 6, 6a-tetrahydro- ,(3aR , 6S , 6aS) -	13395-35-2
74	1,2-Benzenediol, 4-[(1R)-2-amino-1-hydroxyethyl]-	51-41-2
75	Propanamide, 2-methyl-N-[4-nitro-3-(trifluoromethyl)phenyl]-	13311-84-7
76	L-Proline, 5-oxo-	98-79-3
77	D-Arginine	157-06-2
78	Glycine, glycyglycyl-	556-33-2
79	3H-1,2,3-Triazolo[4,5-d]pyrimidine-5,7(4H,6H)-dione	1468-26-4
80	Benzenesulfonamide, 4-amino-N-(5-methyl-3-isoxazolyl)-	723-46-6
81	1-Propanaminium, 3-carboxy-2-hydroxy-N,N,N-trimethyl-,inner salt, (2R)-	541-15-1
82	1,2-Benzenediol, 4-nitro-	3316-09-4
83	Benzenamine, 3- (5- oxazolyl) -	157837-31-5
84	1H-Indol-5-ol	1H-Indol-5-ol
85	Benzamide, 3-amino-	3544-24-9
86	Cyclohexanecarboxylic acid,4-(aminomethyl)-, trans-	1197-18-8
87	Methionine, N -(aminocarbonyl) -	30411-84-8
88	Butanedioic acid, 1-(2,2-dimethylhydrazide)	1596-84-5
89	Benzamide, 4-hydroxy-	619-57-8
90	8-Quinolinol, 5-nitro-	4008-48-4
91	L-Proline, 1-acetyl-	68-95-1
92	1,2-Benzenediol, 3,5-dinitro-	7659-29-2
93	2,3-Quinoxalinedione, 1,4-dihydro-6,7-dinitro-	2379-57-9
94	2-Thiophenesulfonamide, 5-chloro-	53595-66-7

Materials

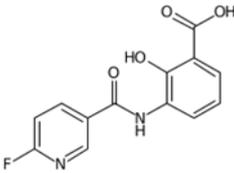
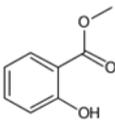
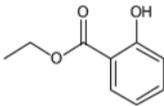
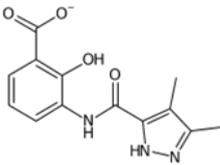
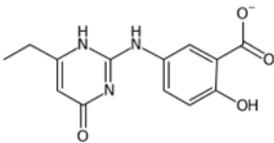
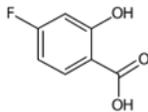
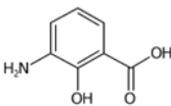
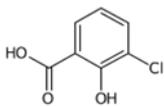
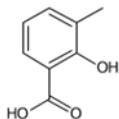
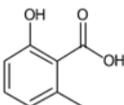
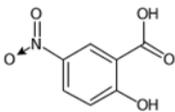
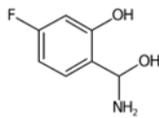
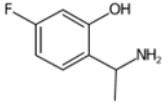
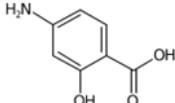
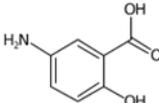
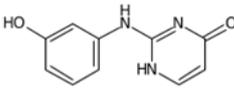
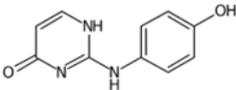
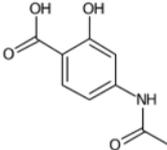
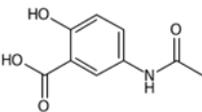
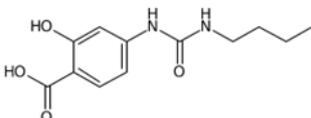
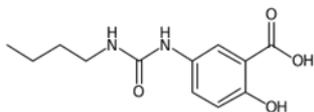
95 | 3- Furancarboxamide,tetrahydro-

871677-92-8

96 | Bicyclo[2.2.1] heptane- 1-carboxamide, N - hydroxy-

256954-29-7

The following molecules were all kindly provided by the Diederich lab, Institute of Pharmaceutical Chemistry, Philipps-University Marburg.

<p>3_550</p> 	<p>Methyl salicylate</p> 	<p>Ethyl salicylate</p> 
<p>3_295</p> 	<p>2_017</p> 	<p>4-fluoro-SA</p> 
<p>3-amino-SA</p> 	<p>3-chloro-SA</p> 	<p>3-methyl-SA</p> 
<p>6-methyl-SA</p> 	<p>5-nitro-SA</p> 	<p>4-fluoro-2-hydroxymethylbenzamide</p> 
<p>4-fluoro-2-hydroxybenzamide</p> 	<p>4-Amino-SA</p> 	<p>5-amino-SA</p> 
<p>2-(3-hydroxyanilino)-1H-pyrimidin-6-one</p> 	<p>2-(4-hydroxyanilino)-1H-pyrimidin-6-one</p> 	<p>4-acetamido-SA</p> 
<p>5-acetamido-SA</p> 	<p>LL059</p> 	<p>LL060</p> 

9 References

1. (22.01.2022) 2019 Biological License Application Approvals. Accessed 22 Aug. 2022
2. Adu-Gyamfi E, Digman M A, Gratton E et al (2012) Single-particle tracking demonstrates that actin coordinates the movement of the Ebola virus matrix protein. *Biophys J* 103(9): L41-3
3. Adu-Gyamfi E, Johnson K A, Fraser M E et al (2015) Host Cell Plasma Membrane Phosphatidylserine Regulates the Assembly and Budding of Ebola Virus. *J Virol* 89(18): 9440–9453
4. Adu-Gyamfi E, Soni S P, Jee C S et al (2014) A loop region in the N-terminal domain of Ebola virus VP40 is important in viral assembly, budding, and egress. *Viruses* 6(10): 3837–3854
5. Adu-Gyamfi E, Soni S P, Xue Y et al (2013) The Ebola virus matrix protein penetrates into the plasma membrane: a key step in viral protein 40 (VP40) oligomerization and viral egress. *J Biol Chem* 288(8): 5779–5789
6. Albariño C G, Wiggleton Guerrero L, Chakrabarti A K et al (2018) Transcriptional analysis of viral mRNAs reveals common transcription patterns in cells infected by five different filoviruses. *PLoS One* 13(8): e0201827
7. Alvarez C P, Lasala F, Carrillo J et al (2002) C-type lectins DC-SIGN and L-SIGN mediate cellular entry by Ebola virus in cis and in trans. *J Virol* 76(13): 6841–6844
8. Amiar S, Husby M L, Wijesinghe K J et al (2021) Lipid-specific oligomerization of the Marburg virus matrix protein VP40 is regulated by two distinct interfaces for virion assembly. *J Biol Chem* 296: 100796
9. Anand S K, Tikoo S K (2013) Viruses as modulators of mitochondrial functions. *Adv Virol* 2013: 738794
10. Arkin M R, Tang Y, Wells J A (2014) Small-molecule inhibitors of protein-protein interactions: progressing toward the reality. *Chem Biol* 21(9): 1102–1114
11. Azzarito V, Long K, Murphy N S et al (2013) Inhibition of α -helix-mediated protein-protein interactions using designed molecules. *Nat Chem* 5(3): 161–173
12. Babé L M, Rosé J, Craik C S (1992) Synthetic "interface" peptides alter dimeric assembly of the HIV 1 and 2 proteases. *Protein Sci* 1(10): 1244–1253
13. Baize S, Leroy E M, Georges-Courbot M C et al (1999) Defective humoral responses and extensive intravascular apoptosis are associated with fatal outcome in Ebola virus-infected patients. *Nat Med* 5(4): 423–426
14. Balmith M, Soliman M E S (2017) VP40 of the Ebola Virus as a Target for EboV Therapy: Comprehensive Conformational and Inhibitor Binding Landscape from Accelerated Molecular Dynamics. *Cell Biochem Biophys* 75(1): 65–78
15. Banadyga L, Hoenen T, Ambroggio X et al (2017) Ebola virus VP24 interacts with NP to facilitate nucleocapsid assembly and genome packaging. *Sci Rep* 7(1): 7698
16. Barry H, Mutua G, Kibuuka H et al (2021) Safety and immunogenicity of 2-dose heterologous Ad26.ZEBOV, MVA-BN-Filo Ebola vaccination in healthy and HIV-infected adults. *PLoS Med* 18(10): e1003813
17. Bauman J D, Patel D, Arnold E (2012) Fragment screening and HIV therapeutics. *Top Curr Chem* 317: 181–200
18. Bauman J D, Patel D, Baker S F et al (2013) Crystallographic fragment screening and structure-based optimization yields a new class of influenza endonuclease inhibitors. *ACS Chem Biol* 8(11): 2501–2508
19. Bausch D G, Towner J S, Dowell S F et al (2007) Assessment of the risk of Ebola virus transmission from bodily fluids and fomites. *J Infect Dis* 196 Suppl 2: S142-7
20. Bavi N, Clark M D, Contreras G F et al (2021) Cryo-EM Structure of dolphin Prestin: Inhibited II (Sulfate +Salicylate) state

Materials

21. Baz-Martínez M, El Motiam A, Ruibal P et al (2016) Regulation of Ebola virus VP40 matrix protein by SUMO. *Sci Rep* 6: 37258
22. Belli L S, Carlis L de, Rondinara G F et al (1994) Sandimmun-Neoral in liver transplantation: a remarkable improvement in long-term immunosuppression. *Transplant Proc* 26(5): 2983–2984
23. Beniac D R, Booth T F (2017) Structure of the Ebola virus glycoprotein spike within the virion envelope at 11 Å resolution. *Sci Rep* 7: 46374
24. Beniac D R, Melito P L, Devarenes S L et al (2012) The organisation of Ebola virus reveals a capacity for extensive, modular polyploidy. *PLoS One* 7(1): e29608
25. Bennett R P, Finch C L, Postnikova E N et al (2020) A Novel Ebola Virus VP40 Matrix Protein-Based Screening for Identification of Novel Candidate Medical Countermeasures. *Viruses* 13(1)
26. Bergendahl L T, Marsh J A (2017) Functional determinants of protein assembly into homomeric complexes. *Sci Rep* 7(1): 4932
27. Bertoletti N, Marchais-Oberwinkler S, Heine A et al (2018) 17beta-hydroxysteroid dehydrogenase 14 variant T205 in complex with Salicylic acid
28. Bharat T A M, Noda T, Riches J D et al (2012) Structural dissection of Ebola virus and its assembly determinants using cryo-electron tomography. *Proc Natl Acad Sci U S A* 109(11): 4275–4280
29. Bhusal R P, Patel K, Kwai B X C et al (2017) Development of NMR and thermal shift assays for the evaluation of Mycobacterium tuberculosis isocitrate lyase inhibitors. *Medchemcomm* 8(11): 2155–2163
30. Billioux B J, Smith B, Nath A (2016) Neurological Complications of Ebola Virus Infection. *Neurotherapeutics* 13(3): 461–470
31. Bond N G, Grant D S, Himmelfarb S T et al (2021) Post-Ebola Syndrome Presents With Multiple Overlapping Symptom Clusters: Evidence From an Ongoing Cohort Study in Eastern Sierra Leone. *Clin Infect Dis* 73(6): 1046–1054
32. Bornholdt Z A, Noda T, Abelson D M et al (2013) Structural rearrangement of ebola virus VP40 begets multiple functions in the virus life cycle. *Cell* 154(4): 763–774
33. Bradfute S B, Braun D R, Shamblin J D et al (2007) Lymphocyte death in a mouse model of Ebola virus infection. *J Infect Dis* 196 Suppl 2: S296-304
34. BRAGG W L (1966) The Diffraction of Short Electromagnetic Waves by a Crystal**Proceedings of the Cambridge Philosophical Society, 17, 43–57 (1913). Communicated by Professor Sir J. J. Thomson. Read 11 November 1912. In: X-ray and Neutron Diffraction. Elsevier, pp. 109–125
35. Brauburger K, Boehmann Y, Krähling V et al (2016) Transcriptional Regulation in Ebola Virus: Effects of Gene Border Structure and Regulatory Elements on Gene Expression and Polymerase Scanning Behavior. *J Virol* 90(4): 1898–1909
36. Bray M, Geisbert T W (2005) Ebola virus: the role of macrophages and dendritic cells in the pathogenesis of Ebola hemorrhagic fever. *Int J Biochem Cell Biol* 37(8): 1560–1566
37. Bricogne G., Blanc E., Brandl M., Flensburg C., Keller P., Paciorek W. (2017) BUSTER. Global Phasing Ltd.
38. Broadhurst M J, Brooks T J G, Pollock N R (2016) Diagnosis of Ebola Virus Disease: Past, Present, and Future. *Clin Microbiol Rev* 29(4): 773–793
39. Broadhurst M J, Kelly J D, Miller A et al (2015) ReEBOV Antigen Rapid Test kit for point-of-care and laboratory-based testing for Ebola virus disease: a field validation study. *Lancet* 386(9996): 867–874
40. Brunton B, Rogers K, Phillips E K et al (2019) TIM-1 serves as a receptor for Ebola virus in vivo, enhancing viremia and pathogenesis. *PLoS Negl Trop Dis* 13(6): e0006983
41. Buck E, Bourne H, Wells J A (2005) Site-specific disulfide capture of agonist and antagonist peptides on the C5a receptor. *J Biol Chem* 280(6): 4009–4012

Materials

42. Buck E, Wells J A (2005) Disulfide trapping to localize small-molecule agonists and antagonists for a G protein-coupled receptor. *Proc Natl Acad Sci U S A* 102(8): 2719–2724
43. Cabri W, Cantelmi P, Corbisiero D et al (2021) Therapeutic Peptides Targeting PPI in Clinical Development: Overview, Mechanism of Action and Perspectives. *Front Mol Biosci* 8: 697586
44. Caffrey D R, Somaroo S, Hughes J D et al (2004) Are protein-protein interfaces more conserved in sequence than the rest of the protein surface? *Protein Sci* 13(1): 190–202
45. Camarasa M-J, Velázquez S, San-Félix A et al (2006) Dimerization inhibitors of HIV-1 reverse transcriptase, protease and integrase: a single mode of inhibition for the three HIV enzymes? *Antiviral Res* 71(2-3): 260–267
46. Campobasso N, Huddler D (2015) Hydrogen deuterium mass spectrometry in drug discovery. *Bioorg Med Chem Lett* 25(18): 3771–3776
47. Cantoni D, Hamlet A, Michaelis M et al (2016) Risks Posed by Reston, the Forgotten Ebolavirus. *mSphere* 1(6)
48. Cardinale D, Guaitoli G, Tondi D et al (2011) Protein-protein interface-binding peptides inhibit the cancer therapy target human thymidylate synthase. *Proc Natl Acad Sci U S A* 108(34): E542–9
49. Carette J E, Raaben M, Wong A C et al (2011) Ebola virus entry requires the cholesterol transporter Niemann-Pick C1. *Nature* 477(7364): 340–343
50. Carlberg I, Mannervik B (1985) [59] Glutathione reductase. In: *Glutamate, Glutamine, Glutathione, and Related Compounds*. Elsevier, pp. 484–490
51. Centers for Disease Control and Prevention (February 22, 2022) History of Ebola Virus Disease (EVD) Outbreaks. https://www.cdc.gov/vhf/ebola/history/chronology.html?CDC_AA_refVal=https%3A%2F%2Fwww.cdc.gov%2Fvhf%2Febola%2Foutbreaks%2Fhistory%2Fchronology.html. Accessed 26 Mar. 2022
52. Chandramouly G, McDevitt S, Sullivan K et al (2015) Small-Molecule Disruption of RAD52 Rings as a Mechanism for Precision Medicine in BRCA-Deficient Cancers. *Chem Biol* 22(11): 1491–1504
53. Chandran K, Sullivan N J, Felbor U et al (2005) Endosomal proteolysis of the Ebola virus glycoprotein is necessary for infection. *Science* 308(5728): 1643–1645
54. Chang R, Mamun A, Dominic A et al (2020) SARS-CoV-2 Mediated Endothelial Dysfunction. *Front Physiol* 11: 605908
55. Chayen N E, Saridakis E (2008) Protein crystallization: from purified protein to diffraction-quality crystal. *Nat Methods* 5(2): 147–153
56. Chen S, Gfeller D, Buth S A et al (2013) Improving binding affinity and stability of peptide ligands by substituting glycines with D-amino acids. *Chembiochem* 14(11): 1316–1322
57. Cheng H H, Xu L, Kumagai H et al (1999) Oligomerization state influences the degradation rate of 3-hydroxy-3-methylglutaryl-CoA reductase. *J Biol Chem* 274(24): 17171–17178
58. Cherpillod P, Schibler M, Vieille G et al (2016) Ebola virus disease diagnosis by real-time RT-PCR: A comparative study of 11 different procedures. *J Clin Virol* 77: 9–14
59. Chertow D S, Kleine C, Edwards J K et al (2014) Ebola virus disease in West Africa--clinical manifestations and management. *N Engl J Med* 371(22): 2054–2057
60. Chevillard F, Rimmer H, Betti C et al (2018) Binding-Site Compatible Fragment Growing Applied to the Design of β 2-Adrenergic Receptor Ligands. *J Med Chem* 61(3): 1118–1129
61. Chilingaryan Z, Yin Z, Oakley A J (2012) Fragment-based screening by protein crystallography: successes and pitfalls. *Int J Mol Sci* 13(10): 12857–12879
62. Cimmperman P, Baranauskiene L, Jachimovičiūtė S et al (2008) A quantitative model of thermal stabilization and destabilization of proteins by ligands. *Biophys J* 95(7): 3222–3231
63. Clark D V, Kibuuka H, Millard M et al (2015) Long-term sequelae after Ebola virus disease in Bundibugyo, Uganda: a retrospective cohort study. *Lancet Infect Dis* 15(8): 905–912

Materials

64. Congreve M, Carr R, Murray C et al (2003) A 'Rule of Three' for fragment-based lead discovery? *Drug Discov Today* 8(19): 876–877
65. Correy G J, Young I D, Thompson M C et al (2020) PanDDA analysis group deposition -- Crystal structure of SARS-CoV-2 NSP3 macrodomain in complex with ZINC000000000922
66. Corti D, Misasi J, Mulangu S et al (2016) Protective monotherapy against lethal Ebola virus infection by a potently neutralizing antibody. *Science* 351(6279): 1339–1342
67. Cowtan K (2001) Phase Problem in X-ray Crystallography, and Its Solution. In: John Wiley & Sons L (ed.) eLS. Wiley
68. Craik D J, Fairlie D P, Liras S et al (2013) The future of peptide-based drugs. *Chem Biol Drug Des* 81(1): 136–147
69. Das D K, Govindan R, Nikić-Spiegel I et al (2018) Direct Visualization of the Conformational Dynamics of Single Influenza Hemagglutinin Trimers. *Cell* 174(4): 926-937.e12
70. Davey R T, Dodd L, Proschan M A et al (2016) A Randomized, Controlled Trial of ZMapp for Ebola Virus Infection. *N Engl J Med* 375(15): 1448–1456
71. Davies T G, Tickle I J (2012) Fragment screening using X-ray crystallography. *Top Curr Chem* 317: 33–59
72. Davis D A, Brown C A, Newcomb F M et al (2003) Reversible oxidative modification as a mechanism for regulating retroviral protease dimerization and activation. *J Virol* 77(5): 3319–3325
73. Davis D A, Yusa K, Gillim L A et al (1999) Conserved cysteines of the human immunodeficiency virus type 1 protease are involved in regulation of polyprotein processing and viral maturation of immature virions. *J Virol* 73(2): 1156–1164
74. Deen G F, Broutet N, Xu W et al (2017) Ebola RNA Persistence in Semen of Ebola Virus Disease Survivors - Final Report. *N Engl J Med* 377(15): 1428–1437
75. Delgado-Roche L, Mesta F (2020) Oxidative Stress as Key Player in Severe Acute Respiratory Syndrome Coronavirus (SARS-CoV) Infection. *Arch Med Res* 51(5): 384–387
76. Dessen A, Volchkov V, Dolnik O et al (2000) Crystal structure of the matrix protein VP40 from Ebola virus. *EMBO J* 19(16): 4228–4236
77. Di L (2015) Strategic approaches to optimizing peptide ADME properties. *AAPS J* 17(1): 134–143
78. Diao L, Meibohm B (2013) Pharmacokinetics and pharmacokinetic-pharmacodynamic correlations of therapeutic peptides. *Clin Pharmacokinet* 52(10): 855–868
79. Dunning J, Kennedy S B, Antierens A et al (2016) Experimental Treatment of Ebola Virus Disease with Brincidofovir. *PLoS One* 11(9): e0162199
80. Dunning J, Sahr F, Rojek A et al (2016) Experimental Treatment of Ebola Virus Disease with TKM-130803: A Single-Arm Phase 2 Clinical Trial. *PLoS Med* 13(4): e1001997
81. (1978) Ebola haemorrhagic fever in Zaire, 1976. *Bull World Health Organ* 56(2): 271–293
82. Ehrmann F R, Stojko J, Metz A et al (2017) Soaking suggests "alternative facts": Only co-crystallization discloses major ligand-induced interface rearrangements of a homodimeric tRNA-binding protein indicating a novel mode-of-inhibition. *PLoS One* 12(4): e0175723
83. Elliott L H, Kiley M P, McCormick J B (1985) Descriptive analysis of Ebola virus proteins. *Virology* 147(1): 169–176
84. Emsley P, Cowtan K (2004) Coot. *Acta Crystallogr D Biol Crystallogr* 60(Pt 12 Pt 1): 2126–2132
85. Erlanson D A (2012) Introduction to fragment-based drug discovery. *Top Curr Chem* 317: 1–32
86. Erlanson D A, Arndt J W, Cancilla M T et al (2011) Discovery of a potent and highly selective PDK1 inhibitor via fragment-based drug discovery. *Bioorg Med Chem Lett* 21(10): 3078–3083
87. Erlanson D A, Braisted A C, Raphael D R et al (2000) Site-directed ligand discovery. *Proc Natl Acad Sci U S A* 97(17): 9367–9372

Materials

88. Erlanson D A, Hansen S K (2004) Making drugs on proteins: site-directed ligand discovery for fragment-based lead assembly. *Current Opinion in Chemical Biology* 8(4): 399–406
89. European Medicines Agency (29.05.2020) New vaccine for prevention of Ebola virus disease recommended for approval in the European Union
90. Evans P R (2011) An introduction to data reduction: space-group determination, scaling and intensity statistics. *Acta Crystallogr D Biol Crystallogr* 67(Pt 4): 282–292
91. Ewer K, Rampling T, Venkatraman N et al (2016) A Monovalent Chimpanzee Adenovirus Ebola Vaccine Boosted with MVA. *N Engl J Med* 374(17): 1635–1646
92. Feldmann H, Bugany H, Mahner F et al (1996) Filovirus-induced endothelial leakage triggered by infected monocytes/macrophages. *J Virol* 70(4): 2208–2214
93. Feldmann H, Geisbert T W (2011) Ebola haemorrhagic fever. *The Lancet* 377(9768): 849–862
94. Feldmann H, Volchkov V E, Volchkova V A et al (2001) Biosynthesis and role of filoviral glycoproteins. *J Gen Virol* 82(Pt 12): 2839–2848
95. Feyfant E, Cross J B, Paris K et al (2011) Fragment-based drug design. *Methods Mol Biol* 685: 241–252
96. Filippakopoulos P, Picaud S, Felletar I et al (2015) Crystal Structure of the fifth bromodomain of human PB1 in complex with salicylic acid
97. Florescu D F, Kalil A C, Hewlett A L et al (2015) Administration of Brincidofovir and Convalescent Plasma in a Patient With Ebola Virus Disease. *Clin Infect Dis* 61(6): 969–973
98. Formenty P, Hatz C, Le Guenno B et al (1999) Human infection due to Ebola virus, subtype Côte d'Ivoire. *J Infect Dis* 179 Suppl 1: S48-53
99. Forrester M T, Foster M W, Benhar M et al (2009) Detection of protein S-nitrosylation with the biotin-switch technique. *Free Radic Biol Med* 46(2): 119–126
100. Fritz-Wolf K, Kehr S, Stumpf M et al (2011) Crystal structure of the human thioredoxin reductase-thioredoxin complex. *Nat Commun* 2: 383
101. Gabizon R, Friedler A (2014) Allosteric modulation of protein oligomerization: an emerging approach to drug design. *Front Chem* 2: 9
102. Galant N J, Bugyei-Twum A, Rakhit R et al (2016) Substoichiometric inhibition of transthyretin misfolding by immune-targeting sparsely populated misfolding intermediates: a potential diagnostic and therapeutic for TTR amyloidoses. *Sci Rep* 6: 25080
103. García M, Cooper A, Shi W et al (2012) Productive replication of Ebola virus is regulated by the c-Abl1 tyrosine kinase. *Sci Transl Med* 4(123): 123ra24
104. Gaudinski M R, Coates E E, Novik L et al (2019) Safety, tolerability, pharmacokinetics, and immunogenicity of the therapeutic monoclonal antibody mAb114 targeting Ebola virus glycoprotein (VRC 608). *Lancet* 393(10174): 889–898
105. Gavira J A, Mantilla M A, Fernandez M et al (2020) Ligand binding domain of the *P. putida* receptor PcaY_PP in complex with salicylic acid
106. Geisbert T W, Hensley L E, Gibb T R et al (2000) Apoptosis induced in vitro and in vivo during infection by Ebola and Marburg viruses. *Lab Invest* 80(2): 171–186
107. Geisbert T W, Hensley L E, Larsen T et al (2003) Pathogenesis of Ebola hemorrhagic fever in cynomolgus macaques: evidence that dendritic cells are early and sustained targets of infection. *Am J Pathol* 163(6): 2347–2370
108. Geisbert T W, Jahrling P B (1995) Differentiation of filoviruses by electron microscopy. *Virus Res* 39(2-3): 129–150
109. Geisbert T W, Strong J E, Feldmann H (2015) Considerations in the Use of Nonhuman Primate Models of Ebola Virus and Marburg Virus Infection. *J Infect Dis* 212 Suppl 2: S91-7

Materials

110. Geromanos S J, Vissers J P C, Silva J C et al (2009) The detection, correlation, and comparison of peptide precursor and product ions from data independent LC-MS with data dependant LC-MS/MS. *Proteomics* 9(6): 1683–1695
111. Gignoux E, Azman A S, Smet M de et al (2016) Effect of Artesunate-Amodiaquine on Mortality Related to Ebola Virus Disease. *N Engl J Med* 374(1): 23–32
112. Glasgow A, Hobbs H T, Perry Z R et al (2021) Ligand-induced changes in dynamics mediate long-range allostery in the lac repressor
113. Gleenberg I O, Herschhorn A, Hizi A (2007) Inhibition of the activities of reverse transcriptase and integrase of human immunodeficiency virus type-1 by peptides derived from the homologous viral protein R (Vpr). *Journal of Molecular Biology* 369(5): 1230–1243
114. Glöckner S, Heine A, Klebe G (2020) A Proof-of-Concept Fragment Screening of a Hit-Validated 96-Compounds Library against Human Carbonic Anhydrase II. *Biomolecules* 10(4)
115. Gomis-Rüth F X, Dessen A, Timmins J et al (2003) The matrix protein VP40 from Ebola virus octamerizes into pore-like structures with specific RNA binding properties. *Structure* 11(4): 423–433
116. Greenfield N J, Stafford W F, Hitchcock-DeGregori S E (1994) The effect of N-terminal acetylation on the structure of an N-terminal tropomyosin peptide and alpha alpha-tropomyosin. *Protein Sci* 3(3): 402–410
117. Gromer S, Merkle H, Schirmer R H et al (2002) Human Placenta Thioredoxin Reductase: Preparation and Inhibitor Studies. In: *Protein Sensors and Reactive Oxygen Species - Part A: Selenoproteins and Thioredoxin*. Elsevier, pp. 382–394
118. Groseth A, Marzi A, Hoenen T et al (2012) The Ebola virus glycoprotein contributes to but is not sufficient for virulence in vivo. *PLoS Pathog* 8(8): e1002847
119. Gsell P-S, Camacho A, Kucharski A J et al (2017) Ring vaccination with rVSV-ZEBOV under expanded access in response to an outbreak of Ebola virus disease in Guinea, 2016. *Lancet Infect Dis* 17(12): 1276–1284
120. Guedj J, Piorkowski G, Jacquot F et al (2018) Antiviral efficacy of favipiravir against Ebola virus. *PLoS Med* 15(3): e1002535
121. Gullberg R C, Jordan Steel J, Moon S L et al (2015) Oxidative stress influences positive strand RNA virus genome synthesis and capping. *Virology* 475: 219–229
122. Guo C, Steinberg L K, Cheng M et al (2020) Site-Specific Siderocalin Binding to Ferric and Ferric-Free Enterobactin As Revealed by Mass Spectrometry. *ACS Chem Biol* 15(5): 1154–1160
123. Hallenbeck K K, Davies J L, Merron C et al (2018) A Liquid Chromatography/Mass Spectrometry Method for Screening Disulfide Tethering Fragments. *SLAS Discov* 23(2): 183–192
124. Han Z, Madara J J, Liu Y et al (2015) ALIX Rescues Budding of a Double PTAP/PPEY L-Domain Deletion Mutant of Ebola VP40: A Role for ALIX in Ebola Virus Egress. *J Infect Dis* 212 Suppl 2: S138-45
125. Han Z, Sagum C A, Takizawa F et al (2017) Ubiquitin Ligase WWP1 Interacts with Ebola Virus VP40 To Regulate Egress. *J Virol* 91(20)
126. Hartlieb B, Modrof J, Mühlberger E et al (2003) Oligomerization of Ebola virus VP30 is essential for viral transcription and can be inhibited by a synthetic peptide. *J Biol Chem* 278(43): 41830–41836
127. Hartlieb B, Weissenhorn W (2006) Filovirus assembly and budding. *Virology* 344(1): 64–70
128. Harvey E P, Hauseman Z J, Cohen D T et al (2020) Identification of a Covalent Molecular Inhibitor of Anti-apoptotic BFL-1 by Disulfide Tethering. *Cell Chem Biol* 27(6): 647-656.e6
129. Hashimoto K, Panchenko A R (2010) Mechanisms of protein oligomerization, the critical role of insertions and deletions in maintaining different oligomeric states. *Proc Natl Acad Sci U S A* 107(47): 20352–20357

Materials

130. Hassaan E, Eriksson P-O, Geschwindner S et al (2020) Fragments as Novel Starting Points for tRNA-Guanine Transglycosylase Inhibitors Found by Alternative Screening Strategies. *ChemMedChem* 15(3): 324–337
131. Hayouka Z, Rosenbluh J, Levin A et al (2007) Inhibiting HIV-1 integrase by shifting its oligomerization equilibrium. *Proc Natl Acad Sci U S A* 104(20): 8316–8321
132. Heikal A, Nakatani Y, Jiao W et al (2018) 'Tethering' fragment-based drug discovery to identify inhibitors of the essential respiratory membrane protein type II NADH dehydrogenase. *Bioorg Med Chem Lett* 28(13): 2239–2243
133. Henao-Restrepo A M, Camacho A, Longini I M et al (2017) Efficacy and effectiveness of an rVSV-vectored vaccine in preventing Ebola virus disease. *The Lancet* 389(10068): 505–518
134. Henao-Restrepo A M, Longini I M, Egger M et al (2015) Efficacy and effectiveness of an rVSV-vectored vaccine expressing Ebola surface glycoprotein. *Lancet* 386(9996): 857–866
135. Higgs E S, Gayedyu-Dennis D, Fischer I W A et al (2021) PREVAIL IV. *Clin Infect Dis* 73(10): 1849–1856
136. Hoenen T, Biedenkopf N, Zielecki F et al (2010) Oligomerization of Ebola virus VP40 is essential for particle morphogenesis and regulation of viral transcription. *J Virol* 84(14): 7053–7063
137. Hoenen T, Groseth A (2017) *Ebolaviruses*. Springer New York, New York, NY
138. Hoenen T, Groseth A, Kok-Mercado F de et al (2011) Minigenomes, transcription and replication competent virus-like particles and beyond. *Antiviral Res* 91(2): 195–208
139. Hoenen T, Groseth A, Kolesnikova L et al (2006) Infection of naive target cells with virus-like particles. *J Virol* 80(14): 7260–7264
140. Hoenen T, Groseth A, Rosenke K et al (2016) Nanopore Sequencing as a Rapidly Deployable Ebola Outbreak Tool. *Emerg Infect Dis* 22(2): 331–334
141. Hoenen T, Shabman R S, Groseth A et al (2012) Inclusion bodies are a site of ebolavirus replication. *J Virol* 86(21): 11779–11788
142. Hoenen T, Volchkov V, Kolesnikova L et al (2005) VP40 octamers are essential for Ebola virus replication. *J Virol* 79(3): 1898–1905
143. Hong P, Koza S, Bouvier E S P (2012) Size-Exclusion Chromatography for the Analysis of Protein Biotherapeutics and their Aggregates. *J Liq Chromatogr Relat Technol* 35(20): 2923–2950
144. Hu J, Zhang L, Liu X (2020) Role of Post-translational Modifications in Influenza A Virus Life Cycle and Host Innate Immune Response. *Front Microbiol* 11: 517461
145. Huang J Y, Louis F J, Dixon M G et al (2016) Notes from the Field: Baseline Assessment of the Use of Ebola Rapid Diagnostic Tests--Forécariah, Guinea, October–November 2015. *MMWR Morb Mortal Wkly Rep* 65(12): 328–329
146. Huang Y, Xu L, Sun Y et al (2002) The assembly of Ebola virus nucleocapsid requires virion-associated proteins 35 and 24 and posttranslational modification of nucleoprotein. *Mol Cell* 10(2): 307–316
147. Huang Y, Zhu Y, Yang M et al (2014) Nucleoprotein-based indirect enzyme-linked immunosorbent assay (indirect ELISA) for detecting antibodies specific to Ebola virus and Marburg virus. *Virology* 463(2): 372–380
148. Huschmann F U, Linnik J, Sparta K et al (2016) Structures of endothiapepsin-fragment complexes from crystallographic fragment screening using a novel, diverse and affordable 96-compound fragment library. *Acta Crystallogr F Struct Biol Commun* 72(Pt 5): 346–355
149. Irie T, Licata J M, Harty R N (2005) Functional characterization of Ebola virus L-domains using VSV recombinants. *Virology* 336(2): 291–298
150. Ivanov A V, Valuev-Elliston V T, Ivanova O N et al (2016) Oxidative Stress during HIV Infection. *Oxid Med Cell Longev* 2016: 8910396

Materials

151. Ivanov A V, Valuev-Elliston V T, Tyurina D A et al (2017) Oxidative stress, a trigger of hepatitis C and B virus-induced liver carcinogenesis. *Oncotarget* 8(3): 3895–3932
152. Jacob S T, Crozier I, Fischer W A et al (2020) Ebola virus disease. *Nat Rev Dis Primers* 6(1)
153. Jacobs M, Rodger A, Bell D J et al (2016) Late Ebola virus relapse causing meningoencephalitis: a case report. *Lancet* 388(10043): 498–503
154. Jaffrey S R, Erdjument-Bromage H, Ferris C D et al (2001) Protein S-nitrosylation. *Nat Cell Biol* 3(2): 193–197
155. Jain S, Martynova E, Rizvanov A et al (2021) Structural and Functional Aspects of Ebola Virus Proteins. *Pathogens* 10(10)
156. Jakobi S, Nguyen T X P, Debaene F et al (2014) Hot-spot analysis to dissect the functional protein-protein interface of a tRNA-modifying enzyme. *Proteins* 82(10): 2713–2732
157. Jasenosky L D, Neumann G, Lukashevich I et al (2001) Ebola virus VP40-induced particle formation and association with the lipid bilayer. *J Virol* 75(11): 5205–5214
158. Johnson E, Jaax N, White J et al (1995) Lethal experimental infections of rhesus monkeys by aerosolized Ebola virus. *Int J Exp Pathol* 76(4): 227–236
159. Johnson K A, Bhattarai N, Budicini M R et al (2021) Cysteine Mutations in the Ebolavirus Matrix Protein VP40 Promote Phosphatidylserine Binding by Increasing the Flexibility of a Lipid-Binding Loop. *Viruses* 13(7)
160. Johnson K A, Pokhrel R, Budicini M R et al (2020) A Conserved Tryptophan in the Ebola Virus Matrix Protein C-Terminal Domain Is Required for Efficient Virus-Like Particle Formation. *Pathogens* 9(5)
161. Johnsson B, Löfås S, Lindquist G (1991) Immobilization of proteins to a carboxymethyl-dextran-modified gold surface for biospecific interaction analysis in surface plasmon resonance sensors. *Anal Biochem* 198(2): 268–277
162. Joubert P-E, Werneke S, La Calle C de et al (2012) Chikungunya-induced cell death is limited by ER and oxidative stress-induced autophagy. *Autophagy* 8(8): 1261–1263
163. Kabsch W (2010) XDS. *Acta Crystallogr D Biol Crystallogr* 66(Pt 2): 125–132
164. Karplus P A, Diederichs K (2012) Linking crystallographic model and data quality. *Science* 336(6084): 1030–1033
165. Karthick V, Nagasundaram N, Doss C G P et al (2016) Virtual screening of the inhibitors targeting at the viral protein 40 of Ebola virus. *Infect Dis Poverty* 5: 12
166. Kaspar A A, Reichert J M (2013) Future directions for peptide therapeutics development. *Drug Discov Today* 18(17-18): 807–817
167. Kathman S G, Xu Z, Statsyuk A V (2014) A fragment-based method to discover irreversible covalent inhibitors of cysteine proteases. *J Med Chem* 57(11): 4969–4974
168. Khan S, Fakhar Z, Ahmad A (2021) Targeting ebola virus VP40 protein through novel inhibitors: exploring the structural and dynamic perspectives on molecular landscapes. *J Mol Model* 27(2): 49
169. Kielkopf C S, Ghosh M, Anand G S et al (2019) HDX-MS reveals orthosteric and allosteric changes in apolipoprotein-D structural dynamics upon binding of progesterone. *Protein Sci* 28(2): 365–374
170. Kirsch P, Hartman A M, Hirsch A K H et al (2019) Concepts and Core Principles of Fragment-Based Drug Design. *Molecules* 24(23)
171. Koehler A, Pfeiffer S, Kolesnikova L et al (2018) Analysis of the multifunctionality of Marburg virus VP40. *J Gen Virol* 99(12): 1614–1620
172. Kolesnikova L, Berghöfer B, Bamberg S et al (2004) Multivesicular bodies as a platform for formation of the Marburg virus envelope. *J Virol* 78(22): 12277–12287

Materials

173. Kolesnikova L, Mühlberger E, Ryabchikova E et al (2000) Ultrastructural organization of recombinant Marburg virus nucleoprotein: comparison with Marburg virus inclusions. *J Virol* 74(8): 3899–3904
174. Kolesnikova L, Nanbo A, Becker S et al (2017) Inside the Cell: Assembly of Filoviruses. *Curr Top Microbiol Immunol* 411: 353–380
175. Konde M K, Baker D P, Traore F A et al (2017) Interferon β -1a for the treatment of Ebola virus disease: A historically controlled, single-arm proof-of-concept trial. *PLoS One* 12(2): e0169255
176. Konermann L, Rodriguez A D, Sowole M A (2014) Type 1 and Type 2 scenarios in hydrogen exchange mass spectrometry studies on protein-ligand complexes. *Analyst* 139(23): 6078–6087
177. Krähling V, Dolnik O, Kolesnikova L et al (2010) Establishment of fruit bat cells (*Rousettus aegyptiacus*) as a model system for the investigation of filoviral infection. *PLoS Negl Trop Dis* 4(8): e802
178. Kruse T, Biedenkopf N, Hertz E P T et al (2018) The Ebola Virus Nucleoprotein Recruits the Host PP2A-B56 Phosphatase to Activate Transcriptional Support Activity of VP30. *Mol Cell* 69(1): 136–145.e6
179. Ksiazek T G, Rollin P E, Jahrling P B et al (1992) Enzyme immunosorbent assay for Ebola virus antigens in tissues of infected primates. *J Clin Microbiol* 30(4): 947–950
180. Ksiazek T G, West C P, Rollin P E et al (1999) ELISA for the detection of antibodies to Ebola viruses. *J Infect Dis* 179 Suppl 1: S192–8
181. Kuhn J H, Radoshitzky S R, Guth A C et al (2006) Conserved receptor-binding domains of Lake Victoria marburgvirus and Zaire ebolavirus bind a common receptor. *J Biol Chem* 281(23): 15951–15958
182. Kukhanova M K, Tunitskaya V L, Smirnova O A et al (2019) Hepatitis C Virus RNA-Dependent RNA Polymerase Is Regulated by Cysteine S-Glutathionylation. *Oxid Med Cell Longev* 2019: 3196140
183. Kuo L C (2011) Fragment-based drug design - tools, practical approaches, and examples. Preface. *Methods Enzymol* 493: xxi–xxii
184. Kurpiers L A, Schulte-Herbrüggen B, Ejotre I et al (2015) Bushmeat and Emerging Infectious Diseases: Lessons from Africa. *Problematic Wildlife*: 507–551
185. Lado M, Walker N F, Baker P et al (2015) Clinical features of patients isolated for suspected Ebola virus disease at Connaught Hospital, Freetown, Sierra Leone: a retrospective cohort study. *The Lancet Infectious Diseases* 15(9): 1024–1033
186. Lai A L, Millet J K, Daniel S et al (2017) The SARS-CoV Fusion Peptide Forms an Extended Bipartite Fusion Platform that Perturbs Membrane Order in a Calcium-Dependent Manner. *Journal of Molecular Biology* 429(24): 3875–3892
187. Lamontagne F, Fowler R A, Adhikari N K et al (2018) Evidence-based guidelines for supportive care of patients with Ebola virus disease. *Lancet* 391(10121): 700–708
188. Landeras-Bueno S, Wasserman H, Oliveira G et al (2021) Cellular mRNA triggers structural transformation of Ebola virus matrix protein VP40 to its essential regulatory form. *Cell Rep* 35(2): 108986
189. Lau J L, Dunn M K (2018) Therapeutic peptides: Historical perspectives, current development trends, and future directions. *Bioorg Med Chem* 26(10): 2700–2707
190. Laudet B, Barette C, Dulery V et al (2007) Structure-based design of small peptide inhibitors of protein kinase CK2 subunit interaction. *Biochem J* 408(3): 363–373
191. Lee J E, Fusco M L, Hessel A J et al (2008) Structure of the Ebola virus glycoprotein bound to an antibody from a human survivor. *Nature* 454(7201): 177–182
192. Lee J S, Adhikari N K J, Kwon H Y et al (2019) Anti-Ebola therapy for patients with Ebola virus disease: a systematic review. *BMC Infect Dis* 19(1): 376

Materials

193. Leroy E M, Epelboin A, Mondonge V et al (2009) Human Ebola outbreak resulting from direct exposure to fruit bats in Luebo, Democratic Republic of Congo, 2007. *Vector Borne Zoonotic Dis* 9(6): 723–728
194. Leroy E M, Kumulungui B, Pourrut X et al (2005) Fruit bats as reservoirs of Ebola virus. *Nature* 438(7068): 575–576
195. Li X, Keskin O, Ma B et al (2004) Protein-protein interactions: hot spots and structurally conserved residues often locate in complemented pockets that pre-organized in the unbound states: implications for docking. *Journal of Molecular Biology* 344(3): 781–795
196. Licata J M, Simpson-Holley M, Wright N T et al (2003) Overlapping motifs (PTAP and PPEY) within the Ebola virus VP40 protein function independently as late budding domains: involvement of host proteins TSG101 and VPS-4. *J Virol* 77(3): 1812–1819
197. Liebschner D, Afonine P V, Baker M L et al (2019) Macromolecular structure determination using X-rays, neutrons and electrons. *Acta Crystallogr D Struct Biol* 75(Pt 10): 861–877
198. Lima G M A, Talibov V O, Jagudin E et al (2020) FragMAX: the fragment-screening platform at the MAX IV Laboratory. *Acta Crystallogr D Struct Biol* 76(Pt 8): 771–777
199. Lin Y, Gross M L (2022) Mass Spectrometry-Based Structural Proteomics for Metal Ion/Protein Binding Studies. *Biomolecules* 12(1)
200. Lipinski C A, Lombardo F, Dominy B W et al (2001) Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings 1PII of original article: S0169-409X(96)00423-1. The article was originally published in *Advanced Drug Delivery Reviews* 23 (1997) 3–25. 1. *Adv Drug Deliv Rev* 46(1-3): 3–26
201. Loughran H M, Han Z, Wrobel J E et al (2016) Quinoxaline-based inhibitors of Ebola and Marburg VP40 egress. *Bioorg Med Chem Lett* 26(15): 3429–3435
202. Lu J, Qu Y, Liu Y et al (2013) Host IQGAP1 and Ebola virus VP40 interactions facilitate virus-like particle egress. *J Virol* 87(13): 7777–7780
203. Lucht A, Grunow R, Otterbein C et al (2004) Production of monoclonal antibodies and development of an antigen capture ELISA directed against the envelope glycoprotein GP of Ebola virus. *Med Microbiol Immunol* 193(4): 181–187
204. Luczkowiak J, Lasala F, Mora-Rillo M et al (2018) Broad Neutralizing Activity Against Ebolaviruses Lacking the Mucin-Like Domain in Convalescent Plasma Specimens From Patients With Ebola Virus Disease. *J Infect Dis* 218(suppl_5): S574-S581
205. Lunev S, Butzloff S, Romero A R et al (2018) Oligomeric interfaces as a tool in drug discovery: Specific interference with activity of malate dehydrogenase of *Plasmodium falciparum* in vitro. *PLoS One* 13(4): e0195011
206. M Alam El-Din H, A Loutfy S, Fathy N et al (2016) Molecular docking based screening of compounds against VP40 from Ebola virus. *Bioinformation* 12(3): 192–196
207. Mabonga L, Kappo A P (2019) Protein-protein interaction modulators: advances, successes and remaining challenges. *Biophys Rev* 11(4): 559–581
208. Madelain V, Baize S, Jacquot F et al (2018) Ebola viral dynamics in nonhuman primates provides insights into virus immuno-pathogenesis and antiviral strategies. *Nat Commun* 9(1): 4013
209. Mahindra A, Tejada G, Rossi M et al (2021) Peptides derived from the SARS-CoV-2 receptor binding motif bind to ACE2 but do not block ACE2-mediated host cell entry or pro-inflammatory cytokine induction. *PLoS One* 16(11): e0260283
210. Maity D, Kumar S, AlHussein R et al (2020) Sub-stoichiometric inhibition of IAPP aggregation: a peptidomimetic approach to anti-amyloid agents. *RSC Chem Biol* 1(4): 225–232
211. Maji B, Gangopadhyay S A, Lee M et al (2019) A High-Throughput Platform to Identify Small-Molecule Inhibitors of CRISPR-Cas9. *Cell* 177(4): 1067-1079.e19

Materials

212. Marí Saéz A, Weiss S, Nowak K et al (2015) Investigating the zoonotic origin of the West African Ebola epidemic. *EMBO Mol Med* 7(1): 17–23
213. Markham A (2021) REGN-EB3. *Drugs* 81(1): 175–178
214. Martínez M J, Salim A M, Hurtado J C et al (2015) Ebola Virus Infection: Overview and Update on Prevention and Treatment. *Infect Dis Ther* 4(4): 365–390
215. Martini G A, Knauff H G, Schmidt H A et al (1968) Über eine bisher unbekannte, von Affen eingeschleppte Infektionskrankheit. *Dtsch Med Wochenschr* 93(12): 559–571
216. Masson G R, Jenkins M L, Burke J E (2017) An overview of hydrogen deuterium exchange mass spectrometry (HDX-MS) in drug discovery. *Expert Opin Drug Discov* 12(10): 981–994
217. Mate S E, Kugelman J R, Nyenswah T G et al (2015) Molecular Evidence of Sexual Transmission of Ebola Virus. *N Engl J Med* 373(25): 2448–2454
218. Maveyraud L, Mourey L (2020) Protein X-ray Crystallography and Drug Discovery. *Molecules* 25(5)
219. Mayer B J (2006) Protein-protein interactions in signaling cascades. *Methods Mol Biol* 332: 79–99
220. McCarthy S E, Johnson R F, Zhang Y-A et al (2007) Role for amino acids 212KLR214 of Ebola virus VP40 in assembly and budding. *J Virol* 81(20): 11452–11460
221. McCoy A J, Grosse-Kunstleve R W, Adams P D et al (2007) Phaser crystallographic software. *J Appl Crystallogr* 40(Pt 4): 658–674
222. McGann M (2012) FRED and HYBRID docking performance on standardized datasets. *J Comput Aided Mol Des* 26(8): 897–906
223. McPherson A, Gavira J A (2014) Introduction to protein crystallization. *Acta Crystallogr F Struct Biol Commun* 70(Pt 1): 2–20
224. Meng X-Y, Zhang H-X, Mezei M et al (2011) Molecular docking: a powerful approach for structure-based drug discovery. *Curr Comput Aided Drug Des* 7(2): 146–157
225. Moreira I S, Fernandes P A, Ramos M J (2007) Hot spots—a review of the protein-protein interface determinant amino-acid residues. *Proteins* 68(4): 803–812
226. Mori M, Villa S, Meneghetti F et al (2020) M. tuberculosis salicylate synthase MbtI in complex with salicylate and Mg²⁺
227. Morimoto Y, Uemura T (2015) Salicylate hydroxylase substrate complex
228. Morris M C, Depollier J, Mery J et al (2001) A peptide carrier for the delivery of biologically active proteins into mammalian cells. *Nat Biotechnol* 19(12): 1173–1176
229. Motsa B B, Stahelin R V (2021) Lipid-protein interactions in virus assembly and budding from the host cell plasma membrane. *Biochem Soc Trans* 49(4): 1633–1641
230. Mühlberger E (2007) Filovirus replication and transcription. *Future Virol* 2(2): 205–215
231. Mühlberger E, Weik M, Volchkov V E et al (1999) Comparison of the transcription and replication strategies of marburg virus and Ebola virus by using artificial replication systems. *J Virol* 73(3): 2333–2342
232. Mulangu S, Dodd LE, Davey R T, JR et al (2019) A Randomized, Controlled Trial of Ebola Virus Disease Therapeutics. *N Engl J Med*
233. Müller I (2017) Guidelines for the successful generation of protein-ligand complex crystals. *Acta Crystallogr D Struct Biol* 73(Pt 2): 79–92
234. Mullis K, Faloona F, Scharf S et al (1986) Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction. *Cold Spring Harb Symp Quant Biol* 51 Pt 1: 263–273
235. Murray C W, Rees D C (2009) The rise of fragment-based drug discovery. *Nat Chem* 1(3): 187–192
236. Nagarajan N, Yapp E K Y, Le N Q K et al (2019) In silico screening of sugar alcohol compounds to inhibit viral matrix protein VP40 of Ebola virus. *Mol Biol Rep* 46(3): 3315–3324

Materials

237. Nanbo A, Imai M, Watanabe S et al (2010) Ebola virus is internalized into host cells via macropinocytosis in a viral glycoprotein-dependent manner. *PLoS Pathog* 6(9): e1001121
238. Narang D, Lento C, J Wilson D (2020) HDX-MS: An Analytical Tool to Capture Protein Motion in Action. *Biomedicines* 8(7)
239. Neumann G, Ebihara H, Takada A et al (2005) Ebola virus VP40 late domains are not essential for viral replication in cell culture. *J Virol* 79(16): 10300–10307
240. Neumann G, Feldmann H, Watanabe S et al (2002) Reverse genetics demonstrates that proteolytic processing of the Ebola virus glycoprotein is not essential for replication in cell culture. *J Virol* 76(1): 406–410
241. Nguyen A (2020) Neue Ansätze zur Entwicklung von Modulatoren der homodimeren tRNA-Guanin-Transglycosylase. Philipps-Universität Marburg
242. Nguyen D, Xie X, Jakobi S et al (2021) Targeting a Cryptic Pocket in a Protein-Protein Contact by Disulfide-Induced Rupture of a Homodimeric Interface. *ACS Chem Biol* 16(6): 1090–1098
243. Nguyen H H, Park J, Kang S et al (2015) Surface plasmon resonance: a versatile technique for biosensor applications. *Sensors (Basel)* 15(5): 10481–10510
244. Niikura M, Ikegami T, Saijo M et al (2001) Detection of Ebola viral antigen by enzyme-linked immunosorbent assay using a novel monoclonal antibody to nucleoprotein. *J Clin Microbiol* 39(9): 3267–3271
245. Noda T, Sagara H, Suzuki E et al (2002) Ebola virus VP40 drives the formation of virus-like filamentous particles along with GP. *J Virol* 76(10): 4855–4865
246. Nordenstedt H, Bah E I, La Vega M-A d et al (2016) Ebola Virus in Breast Milk in an Ebola Virus-Positive Mother with Twin Babies, Guinea, 2015. *Emerg Infect Dis* 22(4): 759–760
247. Nordhoff A, Bücheler U S, Werner D et al (1993) Folding of the four domains and dimerization are impaired by the Gly446--Glu exchange in human glutathione reductase. Implications for the design of antiparasitic drugs. *Biochemistry* 32(15): 4060–4066
248. O'Brien D P, Durand D, Voegelé A et al (2017) Calmodulin fishing with a structurally disordered bait triggers CyaA catalysis. *PLoS Biol* 15(12): e2004486
249. Oda S-I, Noda T, Wijesinghe K J et al (2016) Crystal Structure of Marburg Virus VP40 Reveals a Broad, Basic Patch for Matrix Assembly and a Requirement of the N-Terminal Domain for Immunosuppression. *J Virol* 90(4): 1839–1848
250. Okumura A, Pitha P M, Harty R N (2008) ISG15 inhibits Ebola VP40 VLP budding in an L-domain-dependent manner by blocking Nedd4 ligase activity. *Proc Natl Acad Sci U S A* 105(10): 3974–3979
251. Osorio-Valeriano M, Altegoer F, Steinchen W et al (2019) ParB-type DNA Segregation Proteins Are CTP-Dependent Molecular Switches. *Cell* 179(7): 1512-1524.e15
252. Pascal K E, Dudgeon D, Trefry J C et al (2018) Development of Clinical-Stage Human Monoclonal Antibodies That Treat Advanced Ebola Virus Disease in Nonhuman Primates. *J Infect Dis* 218(suppl_5): S612-S626
253. Pearce N M, Krojer T, Bradley A R et al (2017) A multi-crystal method for extracting obscured crystallographic states from conventionally uninterpretable electron density. *Nat Commun* 8: 15123
254. Pollard A J, Launay O, Lelievre J-D et al (2021) Safety and immunogenicity of a two-dose heterologous Ad26.ZEBOV and MVA-BN-Filo Ebola vaccine regimen in adults in Europe (EBOVAC2). *The Lancet Infectious Diseases* 21(4): 493–506
255. Potterton L, Agirre J, Ballard C et al (2018) CCP4i2: the new graphical user interface to the CCP4 program suite. *Acta Crystallogr D Struct Biol* 74(Pt 2): 68–84

Materials

256. Pourrut X, Souris M, Towner J S et al (2009) Large serological survey showing cocirculation of Ebola and Marburg viruses in Gabonese bat populations, and a high seroprevalence of both viruses in *Rousettus aegyptiacus*. *BMC Infect Dis* 9: 159
257. Qiu X, Wong G, Audet J et al (2014) Reversion of advanced Ebola virus disease in nonhuman primates with ZMapp. *Nature* 514(7520): 47–53
258. Ree R, Varland S, Arnesen T (2018) Spotlight on protein N-terminal acetylation. *Exp Mol Med* 50(7): 1–13
259. Rhodes G (2006) *Crystallography made crystal clear*, 3rd ed. Elsevier, Amsterdam
260. Riddles P W, Blakeley R L, Zerner B (1983) [8] Reassessment of Ellman's reagent. In: *Enzyme Structure Part I*. Elsevier, pp. 49–60
261. Ripka A S, Rich D H (1998) Peptidomimetic design. *Current Opinion in Chemical Biology* 2(4): 441–452
262. Roels T H, Bloom A S, Buffington J et al (1999) Ebola hemorrhagic fever, Kikwit, Democratic Republic of the Congo, 1995: risk factors for patients without a reported exposure. *J Infect Dis* 179 Suppl 1: S92-7
263. Roske Y, Diehl A, Ball L et al (2018) The structural versatility of TasA in *B. subtilis* biofilm formation
264. Ruiz-Santaquiteria M, Sánchez-Murcia P A, Toro M A et al (2017) First example of peptides targeting the dimer interface of *Leishmania infantum* trypanothione reductase with potent in vitro antileishmanial activity. *Eur J Med Chem* 135: 49–59
265. Rupp B (2010) *Biomolecular crystallography*. Routledge, London
266. Ruthel G, Demmin G L, Kallstrom G et al (2005) Association of ebola virus matrix protein VP40 with microtubules. *J Virol* 79(8): 4709–4719
267. Ruyck J de, Brysbaert G, Blossey R et al (2016) Molecular docking as a popular tool in drug design, an in silico travel. *Adv Appl Bioinform Chem* 9: 1–11
268. Saeed M F, Kolokoltsov A A, Albrecht T et al (2010) Cellular entry of ebola virus involves uptake by a macropinocytosis-like mechanism and subsequent trafficking through early and late endosomes. *PLoS Pathog* 6(9): e1001110
269. Saisawang C, Kuadkitkan A, Auewarakul P et al (2018) Glutathionylation of dengue and Zika NS5 proteins affects guanylyltransferase and RNA dependent RNA polymerase activities. *PLoS One* 13(2): e0193133
270. Saisawang C, Kuadkitkan A, Smith D R et al (2017) Glutathionylation of chikungunya nsP2 protein affects protease activity. *Biochim Biophys Acta Gen Subj* 1861(2): 106–111
271. Sanchez A, Kiley M P (1987) Identification and analysis of Ebola virus messenger RNA. *Virology* 157(2): 414–420
272. Sanchez A, Kiley M P, Holloway B P et al (1993) Sequence analysis of the Ebola virus genome. *Virus Res* 29(3): 215–240
273. Sandner A, Ngo K, Sager C P et al (2021) Which Properties Allow Ligands to Open and Bind to the Transient Binding Pocket of Human Aldose Reductase? *Biomolecules* 11(12)
274. Sato A K, Viswanathan M, Kent R B et al (2006) Therapeutic peptides: technological advances driving peptides into development. *Curr Opin Biotechnol* 17(6): 638–642
275. Scala M C, Agamennone M, Pietrantoni A et al (2021) Discovery of a Novel Tetrapeptide against Influenza A Virus: Rational Design, Synthesis, Bioactivity Evaluation and Computational Studies. *Pharmaceuticals (Basel)* 14(10)
276. Schiebel J, Krimmer S G, Röwer K et al (2016) High-Throughput Crystallography: Reliable and Efficient Identification of Fragment Hits. *Structure* 24(8): 1398–1409

Materials

277. Schiebel J, Radeva N, Krimmer S G et al (2016) Six Biophysical Screening Methods Miss a Large Proportion of Crystallographically Discovered Fragment Hits: A Case Study. *ACS Chem Biol* 11(6): 1693–1701
278. Schneider C A, Rasband W S, Eliceiri K W (2012) NIH Image to ImageJ. *Nat Methods* 9(7): 671–675
279. Schnittler H J, Feldmann H (1999) Molecular pathogenesis of filovirus infections: role of macrophages and endothelial cells. *Curr Top Microbiol Immunol* 235: 175–204
280. Schudt G, Dolnik O, Kolesnikova L et al (2015) Transport of Ebolavirus Nucleocapsids Is Dependent on Actin Polymerization: Live-Cell Imaging Analysis of Ebolavirus-Infected Cells. *J Infect Dis* 212 Suppl 2: S160-6
281. Schuller M, Correy G J, Gahbauer S et al (2020) Fragment Binding to the Nsp3 Macrodomain of SARS-CoV-2 Identified Through Crystallographic Screening and Computational Docking. *bioRxiv*
282. Shahian T, Lee G M, Lazic A et al (2009) Inhibition of a viral enzyme by a small-molecule dimer disruptor. *Nat Chem Biol* 5(9): 640–646
283. Shandilya S M D, Nalam M N L, Nalivaika E A et al (2010) Crystal structure of the APOBEC3G catalytic domain reveals potential oligomerization interfaces. *Structure* 18(1): 28–38
284. Shantha J G, Crozier I, Yeh S (2017) An update on ocular complications of Ebola virus disease. *Curr Opin Ophthalmol* 28(6): 600–606
285. Siddiquee K, Zhang S, Guida W C et al (2007) Selective chemical probe inhibitor of Stat3, identified through structure-based virtual screening, induces antitumor activity. *Proc Natl Acad Sci U S A* 104(18): 7391–7396
286. Sijbesma E, Hallenbeck K K, Leysen S et al (2019) Site-Directed Fragment-Based Screening for the Discovery of Protein-Protein Interaction Stabilizers. *J Am Chem Soc* 141(8): 3524–3531
287. Silvestri L S, Ruthel G, Kallstrom G et al (2007) Involvement of vacuolar protein sorting pathway in Ebola virus release independent of TSG101 interaction. *J Infect Dis* 196 Suppl 2: S264-70
288. Sissoko D, Laouenan C, Folkesson E et al (2016) Experimental Treatment with Favipiravir for Ebola Virus Disease (the JIKI Trial). *PLoS Med* 13(3): e1001967
289. Sivapalasingam S, Kamal M, Slim R et al (2018) Safety, pharmacokinetics, and immunogenicity of a co-formulated cocktail of three human monoclonal antibodies targeting Ebola virus glycoprotein in healthy adults. *Lancet Infect Dis* 18(8): 884–893
290. Smart O S, Womack T O, Flensburg C et al (2012) Exploiting structure similarity in refinement: automated NCS and target-structure restraints in BUSTER. *Acta Crystallogr D Biol Crystallogr* 68(Pt 4): 368–380
291. Soriano-Teruel P M, García-Laínez G, Marco-Salvador M et al (2021) Identification of an ASC oligomerization inhibitor for the treatment of inflammatory diseases. *Cell Death Dis* 12(12): 1155
292. Sowole M A, Alexopoulos J A, Cheng Y-Q et al (2013) Activation of ClpP protease by ADEP antibiotics: insights from hydrogen exchange mass spectrometry. *Journal of Molecular Biology* 425(22): 4508–4519
293. Stahelin R V (2013) Surface plasmon resonance: a useful technique for cell biologists to characterize biomolecular interactions. *Mol Biol Cell* 24(7): 883–886
294. Ströher U, West E, Bugany H et al (2001) Infection and activation of monocytes by Marburg and Ebola viruses. *J Virol* 75(22): 11025–11033
295. Struck A-W, Axmann M, Pfefferle S et al (2012) A hexapeptide of the receptor-binding domain of SARS corona virus spike protein blocks viral entry into host cells via the human receptor ACE2. *Antiviral Res* 94(3): 288–296

Materials

296. Švecová L, Østergaard L H, Skálová T et al (2021) Crystallographic fragment screening-based study of a novel FAD-dependent oxidoreductase from *Chaetomium thermophilum*. *Acta Crystallogr D Struct Biol* 77(Pt 6): 755–775
297. Takamatsu Y, Kolesnikova L, Becker S (2018) Ebola virus proteins NP, VP35, and VP24 are essential and sufficient to mediate nucleocapsid transport. *Proc Natl Acad Sci U S A* 115(5): 1075–1080
298. Tan Y B, Lello L S, Liu X et al (2022) Crystal structures of alphavirus nonstructural protein 4 (nsP4) reveal an intrinsically dynamic RNA-dependent RNA polymerase fold. *Nucleic Acids Res* 50(2): 1000–1016
299. Tapia M D, Sow S O, Lyke K E et al (2016) Use of ChAd3-EBO-Z Ebola virus vaccine in Malian and US adults, and boosting of Malian adults with MVA-BN-Filo. *Lancet Infect Dis* 16(1): 31–42
300. Tarazona M, Saiz E (2003) Combination of SEC/MALS experimental procedures and theoretical analysis for studying the solution properties of macromolecules. *Journal of Biochemical and Biophysical Methods* 56(1-3): 95–116
301. Thabault L, Liberelle M, Frédérick R (2021) Targeting protein self-association in drug design. *Drug Discov Today* 26(5): 1148–1163
302. (2015) The ring vaccination trial. *BMJ* 351: h3740
303. Thorn K S, Bogan A A (2001) ASEdb: a database of alanine mutations and their effects on the free energy of binding in protein interactions. *Bioinformatics* 17(3): 284–285
304. Thorpe J H, Wall I D, Sinnamon R H et al (2020) Cocktailed fragment screening by X-ray crystallography of the antibacterial target undecaprenyl pyrophosphate synthase from *Acinetobacter baumannii*. *Acta Crystallogr F Struct Biol Commun* 76(Pt 1): 40–46
305. Timmins J, Scianimanico S, Schoehn G et al (2001) Vesicular release of ebola virus matrix protein VP40. *Virology* 283(1): 1–6
306. Tugyi R, Mezö G, Fellingner E et al (2005) The effect of cyclization on the enzymatic degradation of herpes simplex virus glycoprotein D derived epitope peptide. *J Pept Sci* 11(10): 642–649
307. Tugyi R, Uray K, Iván D et al (2005) Partial D-amino acid substitution: Improved enzymatic stability and preserved Ab recognition of a MUC2 epitope peptide. *Proc Natl Acad Sci U S A* 102(2): 413–418
308. Urata S, Ishikawa T, Yasuda J (2019) Roles of YIGL sequence of Ebola virus VP40 on genome replication and particle production. *J Gen Virol* 100(7): 1099–1111
309. Urata S, Omotuyi O I, Izumisawa A et al (2022) Identification of novel chemical compounds targeting filovirus VP40-mediated particle production. *Antiviral Res*: 105267
310. van Griensven J, Edwards T, Lamballerie X de et al (2016) Evaluation of Convalescent Plasma for Ebola Virus Disease in Guinea. *N Engl J Med* 374(1): 33–42
311. Venkatraman N, Ndiaye B P, Bowyer G et al (2019) Safety and Immunogenicity of a Heterologous Prime-Boost Ebola Virus Vaccine Regimen in Healthy Adults in the United Kingdom and Senegal. *J Infect Dis* 219(8): 1187–1197
312. Volchkov V E, Feldmann H, Volchkova V A et al (1998) Processing of the Ebola virus glycoprotein by the proprotein convertase furin. *Proc Natl Acad Sci U S A* 95(10): 5762–5767
313. Vonrhein C, Flensburg C, Keller P et al (2011) Data processing and analysis with the autoPROC toolbox. *Acta Crystallogr D Biol Crystallogr* 67(Pt 4): 293–302
314. Wahl-Jensen V, Kurz S, Feldmann F et al (2011) Ebola virion attachment and entry into human macrophages profoundly effects early cellular gene expression. *PLoS Negl Trop Dis* 5(10): e1359
315. Wales T E, Fadgen K E, Gerhardt G C et al (2008) High-speed and high-resolution UPLC separation at zero degrees Celsius. *Anal Chem* 80(17): 6815–6820

Materials

316. Walker P J, Siddell S G, Lefkowitz E J et al (2020) Changes to virus taxonomy and the Statutes ratified by the International Committee on Taxonomy of Viruses (2020). *Arch Virol* 165(11): 2737–2748
317. Walton W G, Redinbo M R, Dangl J L (2021) Crystal structure of the MarR family transcriptional regulator from *Variovorax paradoxus* bound to Salicylic acid
318. Wan W, Clarke M, Norris M J et al (2020) Ebola and Marburg virus matrix layers are locally ordered assemblies of VP40 dimers. *Elife* 9
319. Wang J L, Wang Z J (2021) Pyochelin synthetase, a dimeric nonribosomal peptide synthetase elongation module-after-condensation
320. Wang W, Withers J, Li H et al (2020) Structural Basis of Salicylic Acid Perception by Arabidopsis NPR Proteins
321. Warren T K, Jordan R, Lo M K et al (2016) Therapeutic efficacy of the small molecule GS-5734 against Ebola virus in rhesus monkeys. *Nature* 531(7594): 381–385
322. Welch B D, Francis J N, Redman J S et al (2010) Design of a potent D-peptide HIV-1 entry inhibitor with a strong barrier to resistance. *J Virol* 84(21): 11235–11244
323. Welsch S, Kolesnikova L, Krähling V et al (2010) Electron tomography reveals the steps in filovirus budding. *PLoS Pathog* 6(4): e1000875
324. Whelan S P J, Barr J N, Wertz G W (2004) Transcription and replication of nonsegmented negative-strand RNA viruses. *Curr Top Microbiol Immunol* 283: 61–119
325. Wiene-Schmidt B, Oebbeke M, Ngo K et al (2021) Two Methods, One Goal: Structural Differences between Cocrystallization and Crystal Soaking to Discover Ligand Binding Poses. *ChemMedChem* 16(1): 292–300
326. Wijesinghe K J, Urata S, Bhattarai N et al (2017) Detection of lipid-induced structural changes of the Marburg virus matrix protein VP40 using hydrogen/deuterium exchange-mass spectrometry. *J Biol Chem* 292(15): 6108–6122
327. Wojdyla J A, Kaminski J W, Panepucci E et al (2018) DA+ data acquisition and analysis software at the Swiss Light Source macromolecular crystallography beamlines. *J Synchrotron Radiat* 25(Pt 1): 293–303
328. Wojdyr M, Keegan R, Winter G et al (2013) DIMPLE - a pipeline for the rapid generation of difference maps from protein crystals with putatively bound ligands. *Acta Crystallogr A Found Crystallogr* 69(a1): s299-s299
329. Wood E, Hall K H, Tate W (2021) Role of mitochondria, oxidative stress and the response to antioxidants in myalgic encephalomyelitis/chronic fatigue syndrome. *Chronic Dis Transl Med* 7(1): 14–26
330. Wu L, Jin D, Wang D et al (2020) The two-stage interaction of Ebola virus VP40 with nucleoprotein results in a switch from viral RNA synthesis to virion assembly/budding. *Protein Cell* 13(2): 120–140
331. Xu Z, Jin B, Teng G et al (2016) Epidemiologic characteristics, clinical manifestations, and risk factors of 139 patients with Ebola virus disease in western Sierra Leone. *Am J Infect Control* 44(11): 1285–1290
332. Yamayoshi S, Noda T, Ebihara H et al (2008) Ebola virus matrix protein VP40 uses the COPII transport system for its intracellular transport. *Cell Host Microbe* 3(3): 168–177
333. Yang B, Gao L, Li L et al (2003) Potent suppression of viral infectivity by the peptides that inhibit multimerization of human immunodeficiency virus type 1 (HIV-1) Vif proteins. *J Biol Chem* 278(8): 6596–6602
334. Yang W, Fucini R V, Fahr B T et al (2009) Fragment-based discovery of nonpeptidic BACE-1 inhibitors using tethering. *Biochemistry* 48(21): 4488–4496

Materials

335. Yasuda J, Nakao M, Kawaoka Y et al (2003) Nedd4 regulates egress of Ebola virus-like particles from host cells. *J Virol* 77(18): 9987–9992
336. Zhang H, Zhao Q, Bhattacharya S et al (2008) A cell-penetrating helical peptide as a potential HIV-1 inhibitor. *Journal of Molecular Biology* 378(3): 565–580
337. Zheng M, Cong W, Peng H et al (2021) Stapled Peptides Targeting SARS-CoV-2 Spike Protein HR1 Inhibit the Fusion of Virus to Its Cell Receptor. *J Med Chem* 64(23): 17486–17495
338. Zhong L, Arnér E S, Holmgren A (2000) Structure and mechanism of mammalian thioredoxin reductase: the active site is a redox-active selenolthiol/selenenylsulfide formed from the conserved cysteine-selenocysteine sequence. *Proc Natl Acad Sci U S A* 97(11): 5854–5859
339. Zhu L, Yang F, Chen L et al (2008) Human serum albumin complexed with myristate, 3'-azido-3'-deoxythymidine (AZT) and salicylic acid

10 Appendix

10.1 Supplemental material

SUDV	MRRVTVPTAPPAYADIGYPM SMLPIKSSRAVSGIQKQEVLP GMDTPSN S MRPVADDNID	60
RESTV	MRRGVLP TAPPAYNDIAYPMSILPTRPSVIVNETKSDVLAVPGADVPSNSMRPVADDNID	60
BOMV	MRRTVIPTAPPDYTEALYPQRTNSIGSTSTSTNQYYPTDVPGGDTPSN SLRPVADDNID	60
EBOV	MRRVILPTAPPEYMEAIYPVRSNSTIARGGNSNTGFLTPE SVNGDTPSN LRPVADDNID	60
TAFV	MRR IILPTAPPEYMEAVYPMRTMNSGADNTASGPNYTTTGVM TNDTPSN SLRPVADDNID	60
BDBV	MRRAILPTAPPEYIEAVYPMRTVST SIN STASGPNFPAPDVMSDTPSN SLRPVADDNID	60
	*** :***** * : ** . * .*** :**:* **	
SUDV	HTSHTPNGVASAFILEATVNVISGPKVLMKQIP IWLPLGIADQK TYSFDSTTAAIMLASY	120
RESTV	HSSHTPSGVASAFILEATVNVISGTKVLMKQIP IWLPLGVADQKIYSFDSTTAAIMLASY	120
BOMV	HTNHTPDNVSSAFILEAMVNVISGPKVLMKQIP IWLPLGVADQKKYSFDSTTAAIMLASY	120
EBOV	HASHTPGSVSSAFILEAMVNVISGPKVLMKQIP IWLPLGVADQK TYSFDSTTAAIMLASY	120
TAFV	HPSHTPNSVASAFILEAMVNVISGPKVLMKQIP IWLPLGVSDQK TYSFDSTTAAIMLASY	120
BDBV	HPSHTPTS VSSAFILEAMVNVISGPKVLMKQIP IWLPLGVADQK TYSFDSTTAAIMLASY	120
	* .*** .*:***** ***** *****:*** *****	
SUDV	TITHFGKANPLVRVNR LGQIPDHPLRLLRMGNQAF LQEFVLP PVQLPQYFTFDLTALK	180
RESTV	TVTHFGKISNPLVRVNR LGPGIPDHPLRLLRLGNQAF LQEFVLP PVQLPQYFTFDLTALK	180
BOMV	TITHFGKTSNPLVRVNR LGPGIPDHPLRLLRLGNQAF LQEFVLP PVQLPQYFTFDLTALK	180
EBOV	TITHFGKATNPLVRVNR LGPGIPDHPLRLLRIGNQAF LQEFVLP PVQLPQYFTFDLTALK	180
TAFV	TITHFGKTSNPLVRINR LGPGIPDHPLRLLRIGNQAF LQEFVLP PVQLPQYFTFDLTALK	180
BDBV	TITHFGKTSNPLVRINR LGPGIPDHPLRLLRIGNQAF LQEFVLP PVQLPQYFTFDLTALK	180
	*:***** .*****:*** *****:*****:*****:*****	
SUDV	LVTQPLPAATWTDETPSNLSGALRPGLSFHPKLRPVLLPGKTGKKGHVSDLTAPDKIQTI	240
RESTV	LITQPLPAATWTDETPAGAVNALRPGLSLHFKLRPILLPGKTGKKGHASDLTSPDKIQTI	240
BOMV	LITQPLPAATWTDEVLLTSPNMLRPGLSFHPKLRPILLPGKPGKKGANLGLTAPDKIHAI	240
EBOV	LITQPLPAATWTDDTPTGSNGALRPGLSFHPKLRPILLPNKSGKKGNSADLTSPDKIQAI	240
TAFV	LITQPLPAATWTDETPAVSTGTLRPGLSFHPKLRPILLPGRAGKKGNSDLTSPDKIQAI	240
BDBV	LITQPLPAATWTDDTPTGPTGILRPGLSFHPKLRPILLPGKTGKRGSDDL TSPDKIQAI	240
	*:*****: . *****:*** *****:***: **:* .***:***:*	
SUDV	VNLMQDFKIVPIDPAKSIIGIEVPELLVHKL TGKKMSQKNGQPIIPVLLPKYIGLDPISP	300
RESTV	MNAIPDLKIVPIDPTKNIVGIEVPELLVQLR TGKKPQPKNGQPIIPVLLPKYVGLDPISP	300
BOMV	MDHLQDLKVVVIDSSRNIVGIEVPDALVQLR TGKKPHAKNGQSIIIPVLLPKYIGDPVSA	300
EBOV	M TSLQDFKIVPIDPTKNIMGIEVPETLVHKL TGKKVTSKNGQPIIPVLLPKYIGLDPVAP	300
TAFV	MNFLQDLKIVPIDPTKNIMGIEVPELLVHRL TGKKTTTKNGQPIIPILLPKYIGLDPISQ	300
BDBV	MNFLQDLKLVPIDPAKNIMGIEVPELLVHRL TGKKITTKNGQPIIPILLPKYIGMDPISQ	300
	: : *:*:**** :.:*:*****: **:*:***** **** *:*:*****:* **:*:	
SUDV	GDLT MVIITPDYDDCHSPASCSYLSEK----- 326	
RESTV	GDLT MVIITQDCDSCHSPASHPYHMDKQNSYQ 331	
BOMV	SDLT MVIITQDCDTCNSPASFP IASEK----- 326	
EBOV	GDLT MVIITQDCDTC HSPASLPAVIEK----- 326	
TAFV	GDLT MVIITQDCDSCHSPASLPPVNEK----- 326	
BDBV	GDLT MVIITQDCDTC HSPASLPPVSEK----- 326	
	.***** * * *:**** :*	

Figure 58: Comparison of the VP40 sequences. Alignment of SUDV, RESTV, BOMV, EBOV, TAFV and BDBV using Clustal Omega with the residues of the different homo-oligomeric interfaces highlighted in green, purple and brown (for the dimeric, octameric and hexameric interface, respectively).

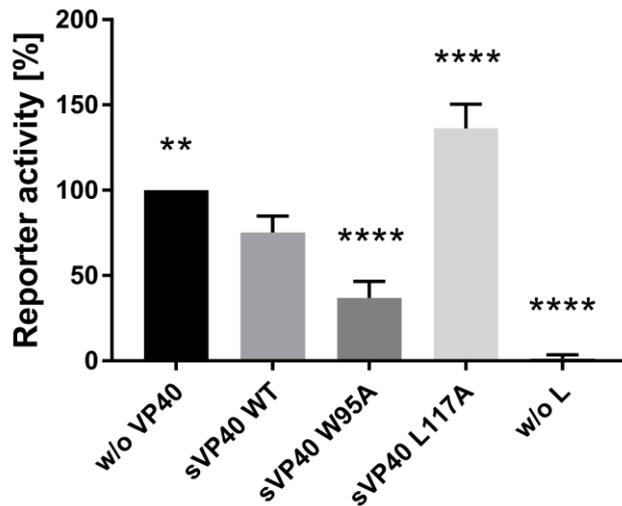


Figure 59: sVP40 WT and mutants influence reporter gene activity of a MG assay. HEK293 cells were transfected with 200 ng pCAGGS sVP40 WT or mutants along with the MG assay components. Cells were lysed 48 h pt, reporter gene activity was measured and normalized to the sample without VP40 (set to 100%). For the negative control, no polymerase L was added to the transfection mix. Bars indicate the mean \pm SD of five experiments, asterisks indicate statistical significance (One-way ANOVA) compared to sVP40 WT as follows: ** $P < 0.005$ and **** $P < 0.0001$.

Table 22: Overview of data obtained by HDX-MS. Each experiment is shown in a separate column.

	sVP40 Δ 43 WT +/- 10 mM NaSA	sVP40 Δ 43 WT, W95A	sVP40 Δ 43 WT, W95A, L117A
Conditions of H/D exchange	25 °C in 25 mM Tris-Cl pH 8.0, 300 mM NaCl; final D ₂ O = 90%	25 °C in 25 mM Tris-Cl pH 8.0, 300 mM NaCl; final D ₂ O = 90%	25 °C in 25 mM Tris-Cl pH 8.0, 300 mM NaCl; final D ₂ O = 90%
Timecourse of H/D exchange	10/30/100/1,000/10,000 s	10/30/100/1,000/10,000 s	10/30/100/1,000/10,000 s
Replicates	3, technical (separate H/D exchange reactions measured)	3, technical (separate H/D exchange reactions measured)	3, technical (separate H/D exchange reactions measured)
Number of Peptides	137	83	115
Sequence coverage [%]	94,6	89,9	90,2
Redundancy per amino acid	6,26	4,6	5,41
Significance criterium applied	5% difference in HDX	5% difference in HDX	5% difference in HDX

Appendix

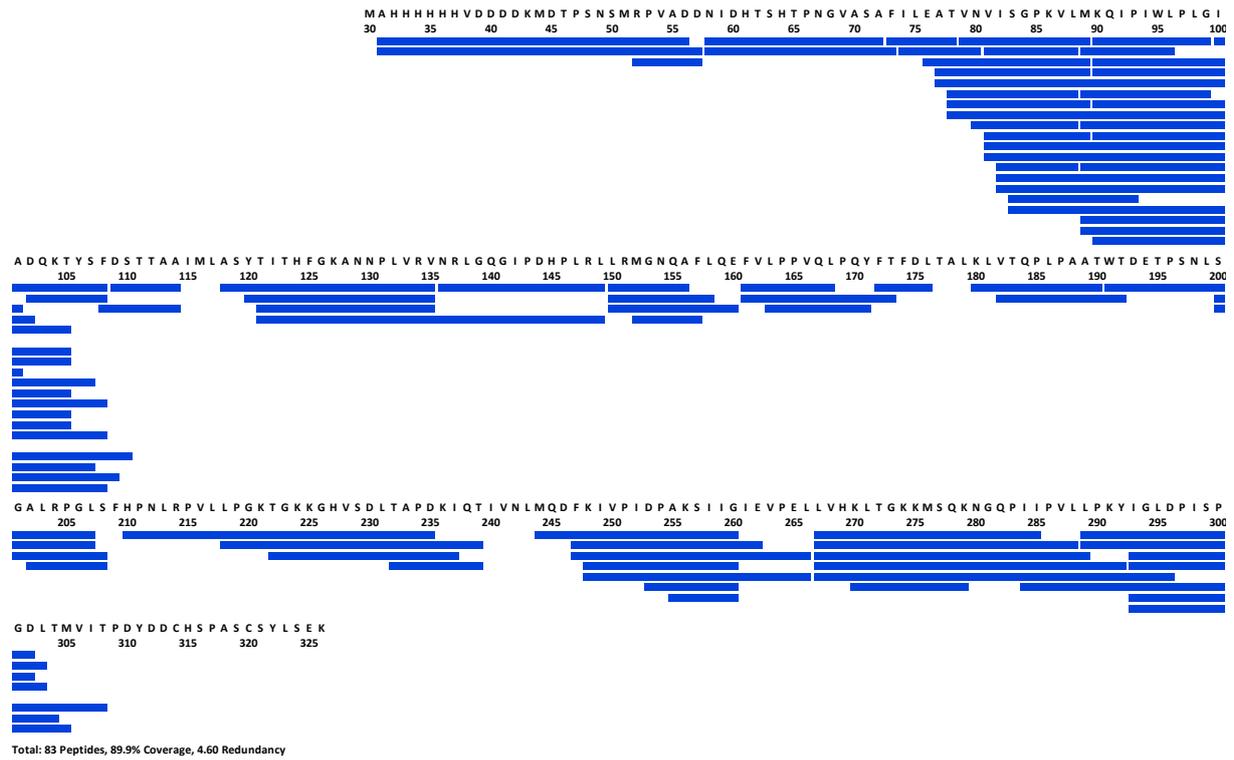


Figure 60: Peptide coverage map of HDX-MS experiments comparing sVP40 Δ 43 WT with W945A and L117A. Each blue bar represents a peptide that has been analyzed. Peptides 89-100 were not covered found in W95A but appear here because identified in other variants.

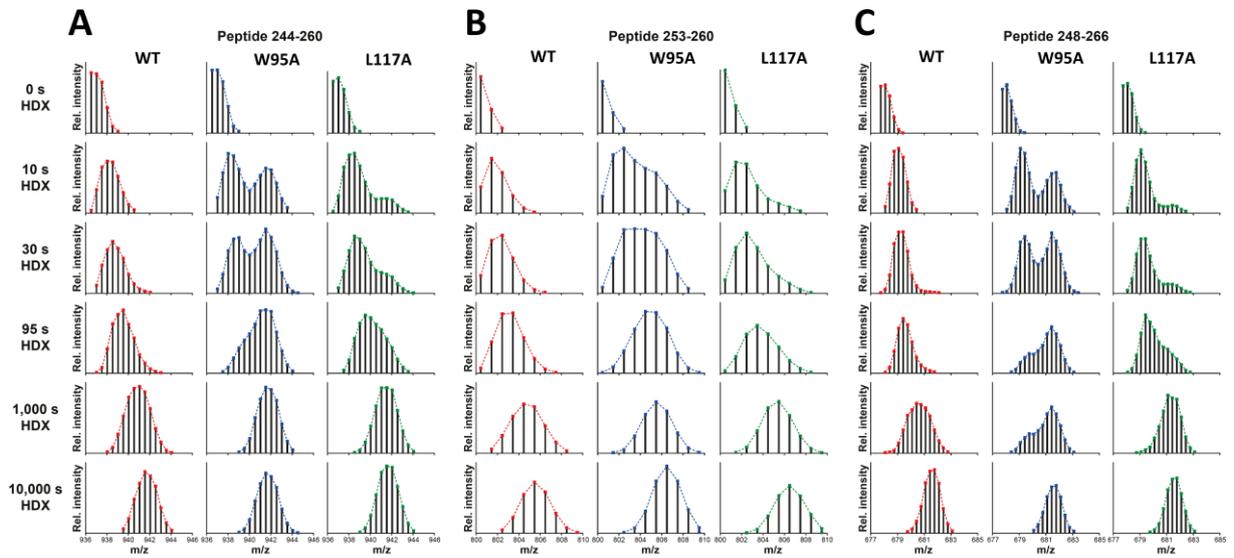
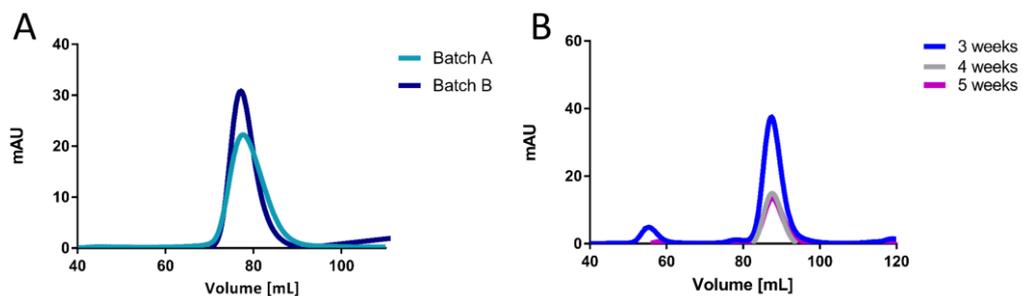


Figure 61: HDX-MS of sVP40 Δ 43 mutants depicting bimodality. Bimodality is observed for peptides A) 244-260, B) 252-260 and C) 248-266 for both sVP40 Δ 43 W95A and L117A.



Appendix

Figure 62: Stability of sVP40_{Δ43} W95A and L117A via SEC. A) Purified dimeric sVP40_{Δ43} W95A from two different expression/purification batches, stored at -80 °C (light and dark blue) and reapplied to the same SEC column after > 5 months. B) and purified monomeric sVP40_{Δ43} L117A after storage of three, four or five weeks at 4 °C.

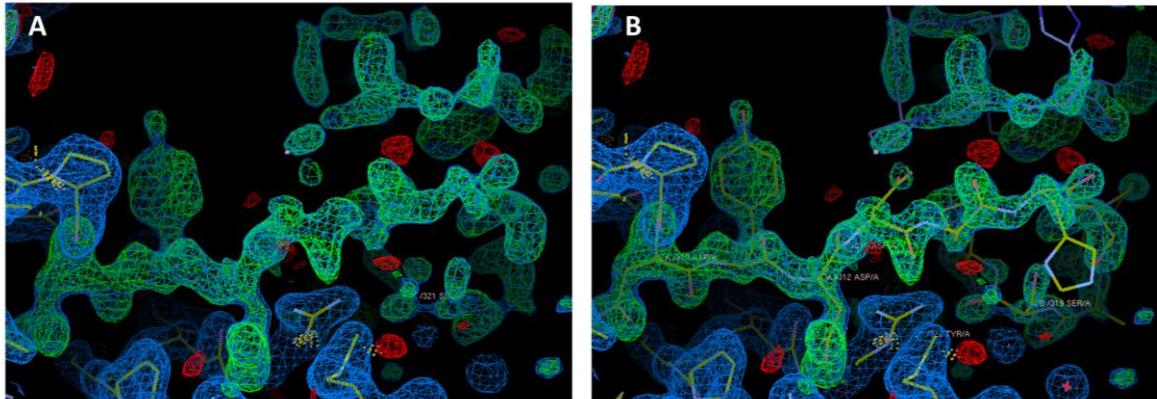


Figure 63: Electron density of dimeric sVP40_{Δ43} WT (AW122). Positive electron density could be observed for residues 310-DYDDCHSPASCSYL-323 before (A) and after (B) manual model building using coot.

Table 23: Determined affinities [mM] of various peptides to target proteins using SPR. All ten peptides were diluted to 31.25; 62.5; 125; 250 and 500 μM and passed over the ligand-covered sensor surface (dimer (dim.) and octamer (oct.) for both zVP40 and sVP40 WT). A 1:1 binding model was used to calculate K_D .

#	peptide	zVP40 dim.	zVP40 oct.	sVP40 dim.	sVP40 oct.
1	SFDSTTAAIMLA	-	-	-	0.07
	GVADQKTYSFDSTTAAIMLAS	-	-	-	-
2	KQIPIWLPLGVA	0.34	0.94	0.25	0.83
	KQIPIWLPLGVA non-ac	-	-	-	-
	flag-KQIPIWLPLGVA	-	-	-	-
3	LGPGIPDHPLRLLRIG	3.43	3.77	19.90	3.68
	LGPGIPDHPLRLLRIG non-ac	-	-	-	-
	flag-LGPGIPDHPLRLLRIG	1.39	1.40	0.38	1.38
4	VQLPQYFTFDLT non-ac	0.24	0.38	1.42	0.52
	flag-VQLPQYFTFDLT	0.91	0.24	0.50	0.13

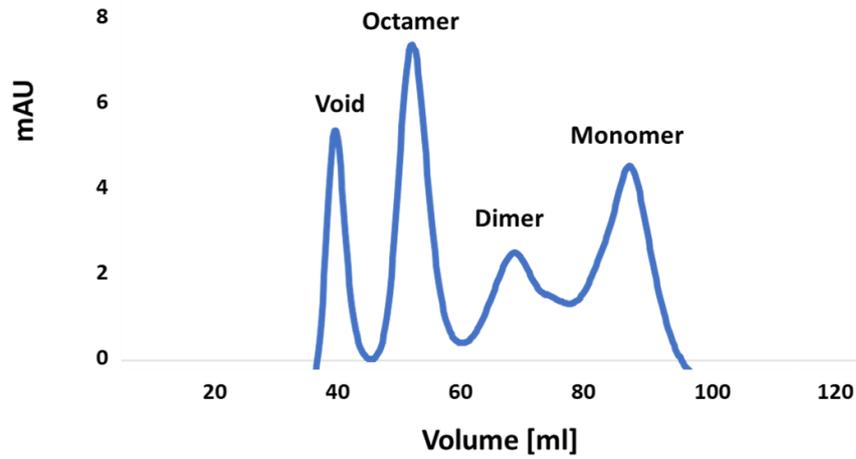


Figure 64: Size-exclusion profile of the sVP40-linker fusion protein. The sVP40-linker (without peptide 1) construct was cloned into pET25b, transformed into Rosetta2 cells, expressed and purified according to standard protocols. Size-exclusion chromatography served to identify the different homo-oligomers of sVP40. Representative profile of two independently performed experiments.

Appendix

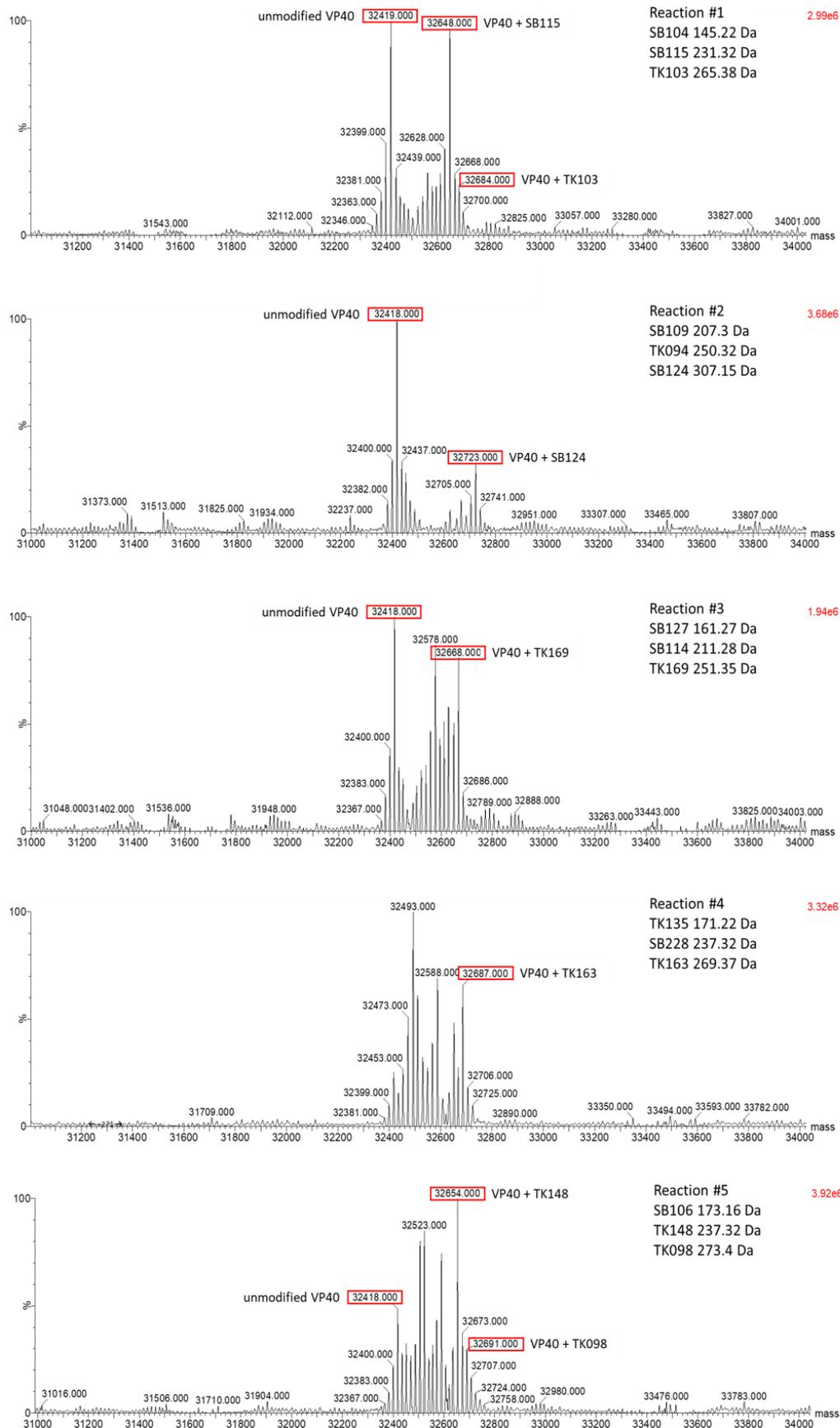


Figure 65: ESI mass spectra of DT hits for sVP40 Δ 43 N67C of reaction 1 to 5. 30 μ M of dimeric sVP40 Δ 43 N67C were incubated with a mixture of thiol-containing fragments and DTT. Afterwards, the mixtures were desalted and analyzed via IP-MS. Fragment hits were identified by a mass shift of unmodified protein plus fragment mass.

Appendix

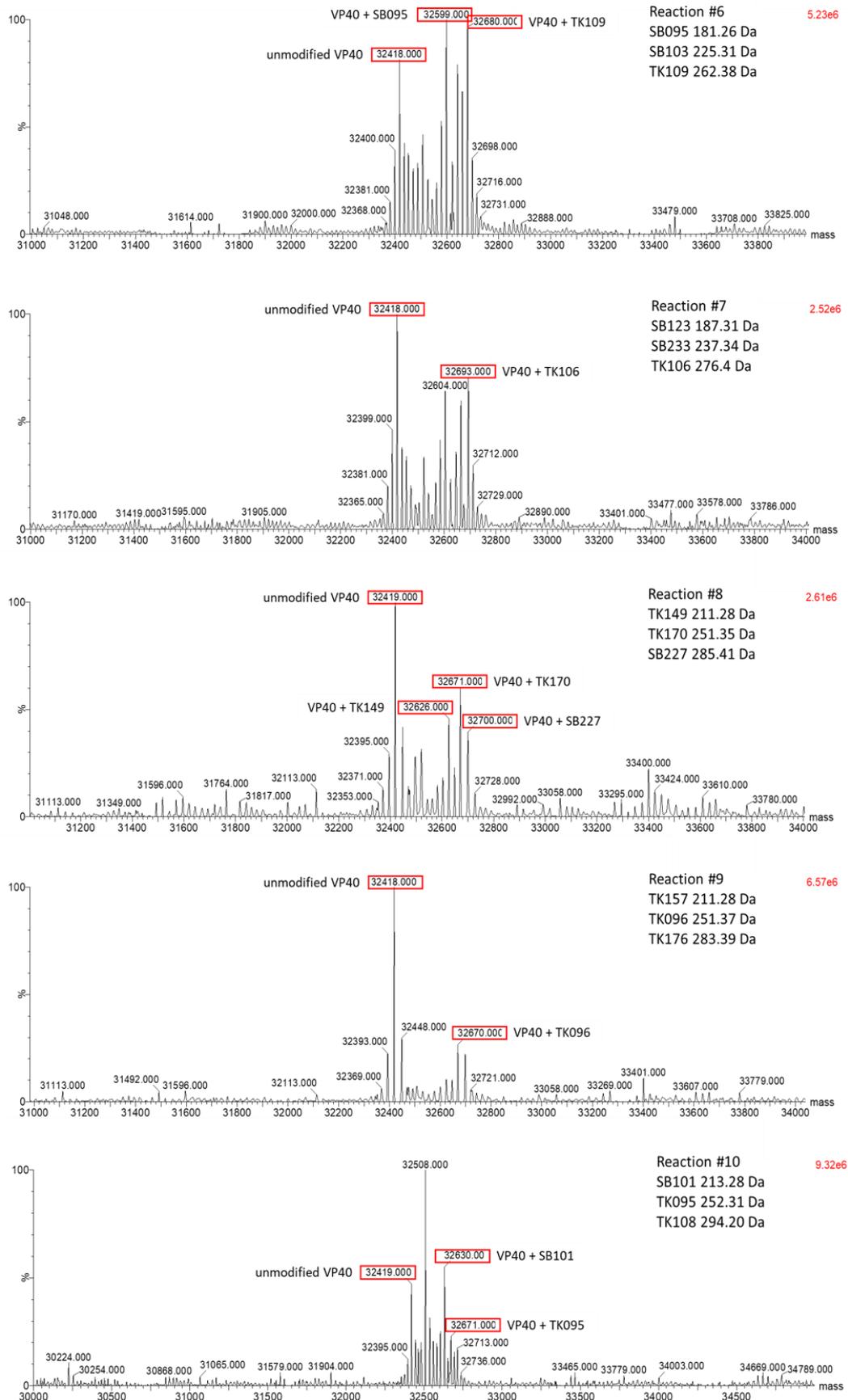


Figure 66: ESI mass spectra of DT hits for sVP40_{Δ43} N67C of reaction 6 to 10. 30 μM of dimeric sVP40_{Δ43} N67C were incubated with a mixture of thiol-containing fragments and DTT. Afterwards, the mixtures were desalted and analyzed via IP-MS. Fragment hits were identified by a mass shift of unmodified protein plus fragment mass.

Appendix

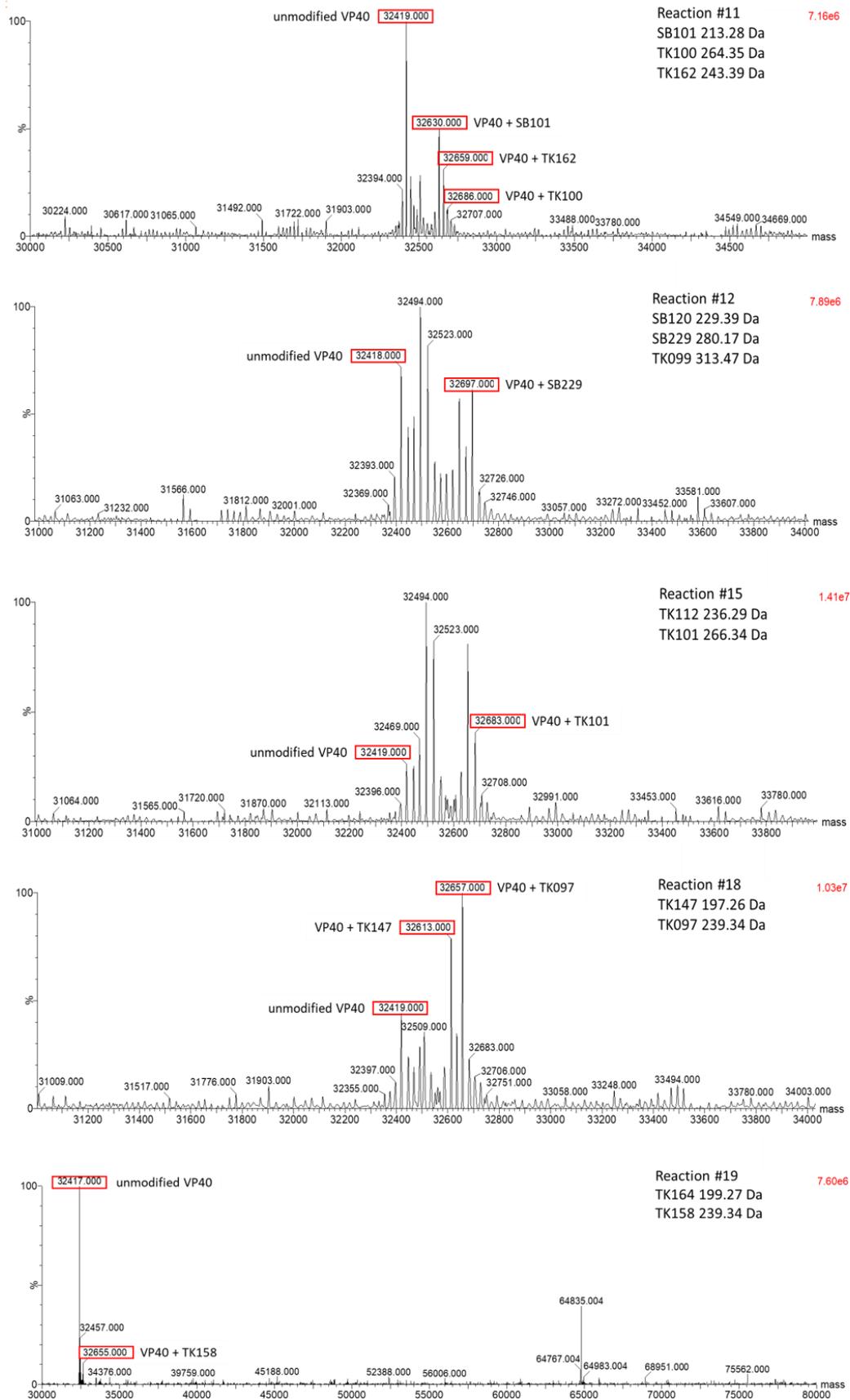


Figure 67: ESI mass spectra of DT hits for sVP40_{Δ43} N67C of reaction 11 to 19. 30 μM of dimeric sVP40_{Δ43} N67C were incubated with a mixture of thiol-containing fragments and DTT. Afterwards, the mixtures were desalted and analyzed via IP-MS. Fragment hits were identified by a mass shift of unmodified protein plus fragment mass.

Appendix

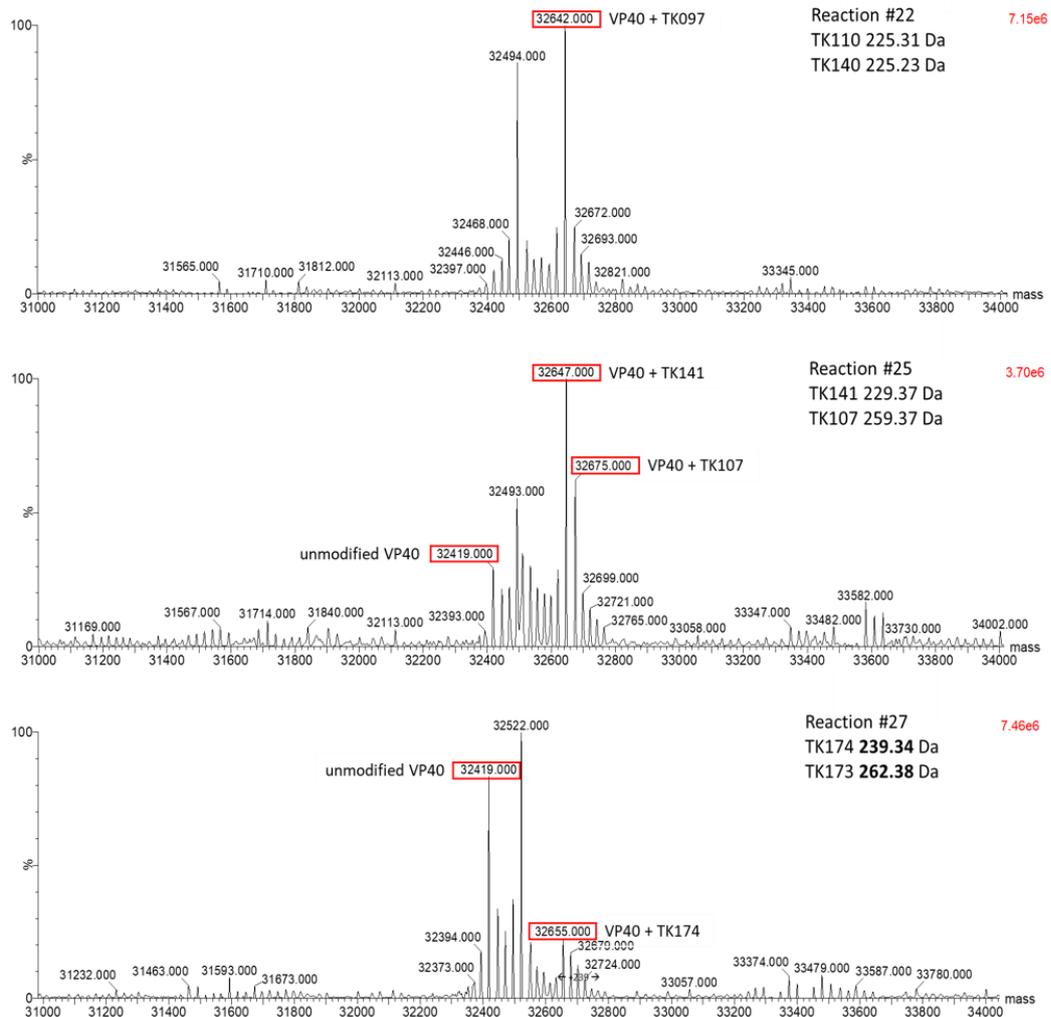


Figure 68: ESI mass spectra of DT hits for sVP40_{Δ43} N67C of reaction 22 to 27. 30 μM of dimeric sVP40_{Δ43} N67C were incubated with a mixture of thiol-containing fragments and DTT. Afterwards, the mixtures were desalted and analyzed via IP-MS. Fragment hits were identified by a mass shift of unmodified protein plus fragment mass.

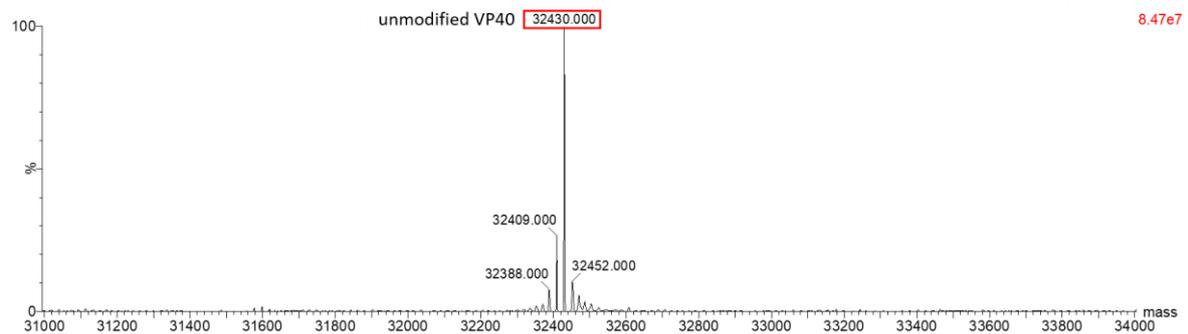


Figure 69: ESI mass spectra of DT hits for sVP40_{Δ43} WT negative control without fragments. 30 μM of dimeric sVP40_{Δ43} WT was incubated with a DMSO and DTT. Afterwards, the mixture was desalted and analyzed via IP-MS.

Appendix

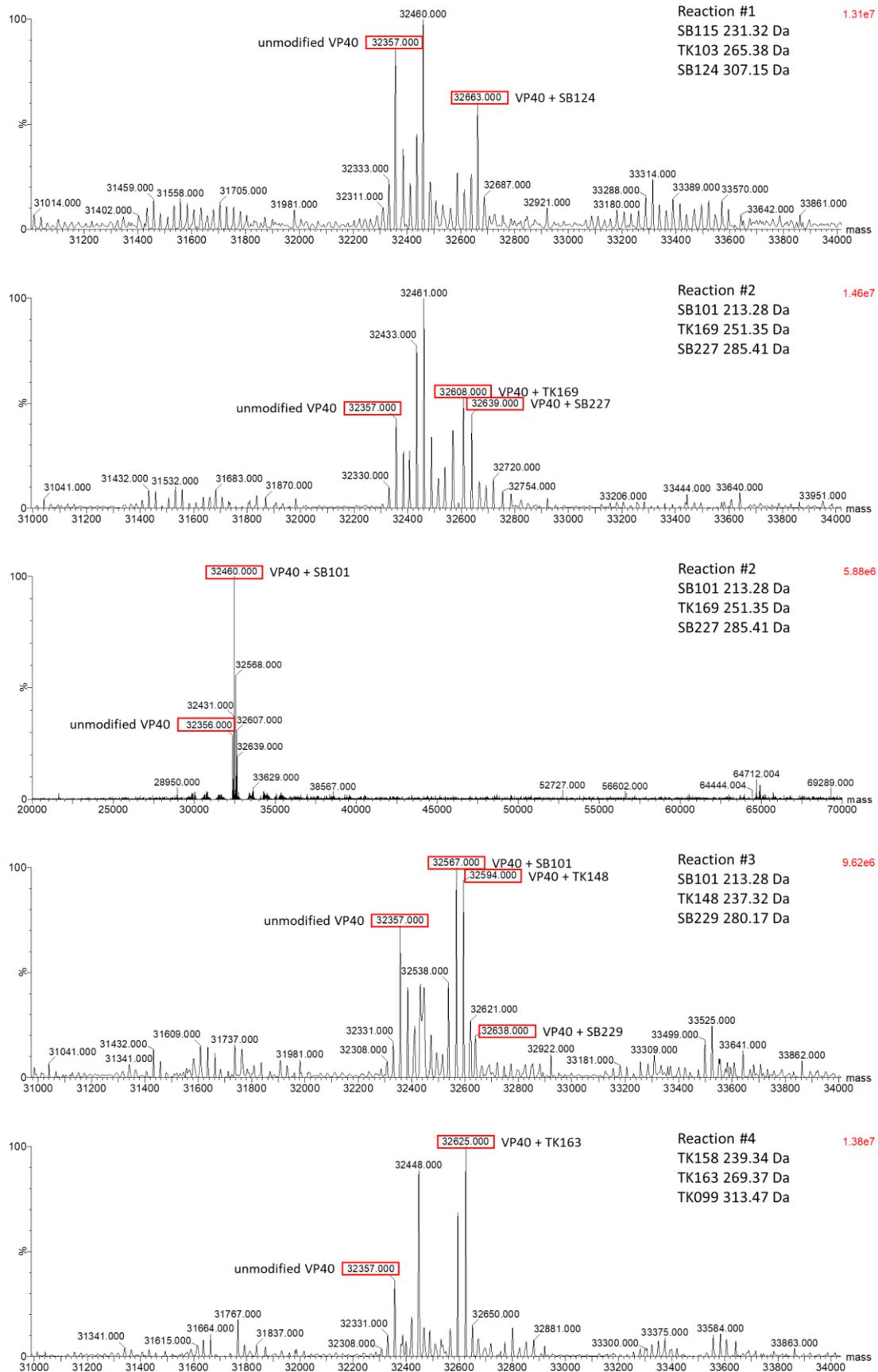


Figure 70: ESI mass spectra of DT hits for sVP40_{Δ43} N67C-CCA of reaction 1 to 4. 30 μM of dimeric sVP40_{Δ43} N67C-CCA were incubated with a mixture of thiol-containing fragments and DTT. Afterwards, the mixtures were desalted and analyzed via IP-MS. Fragment hits were identified by a mass shift of unmodified protein plus fragment mass.

Appendix

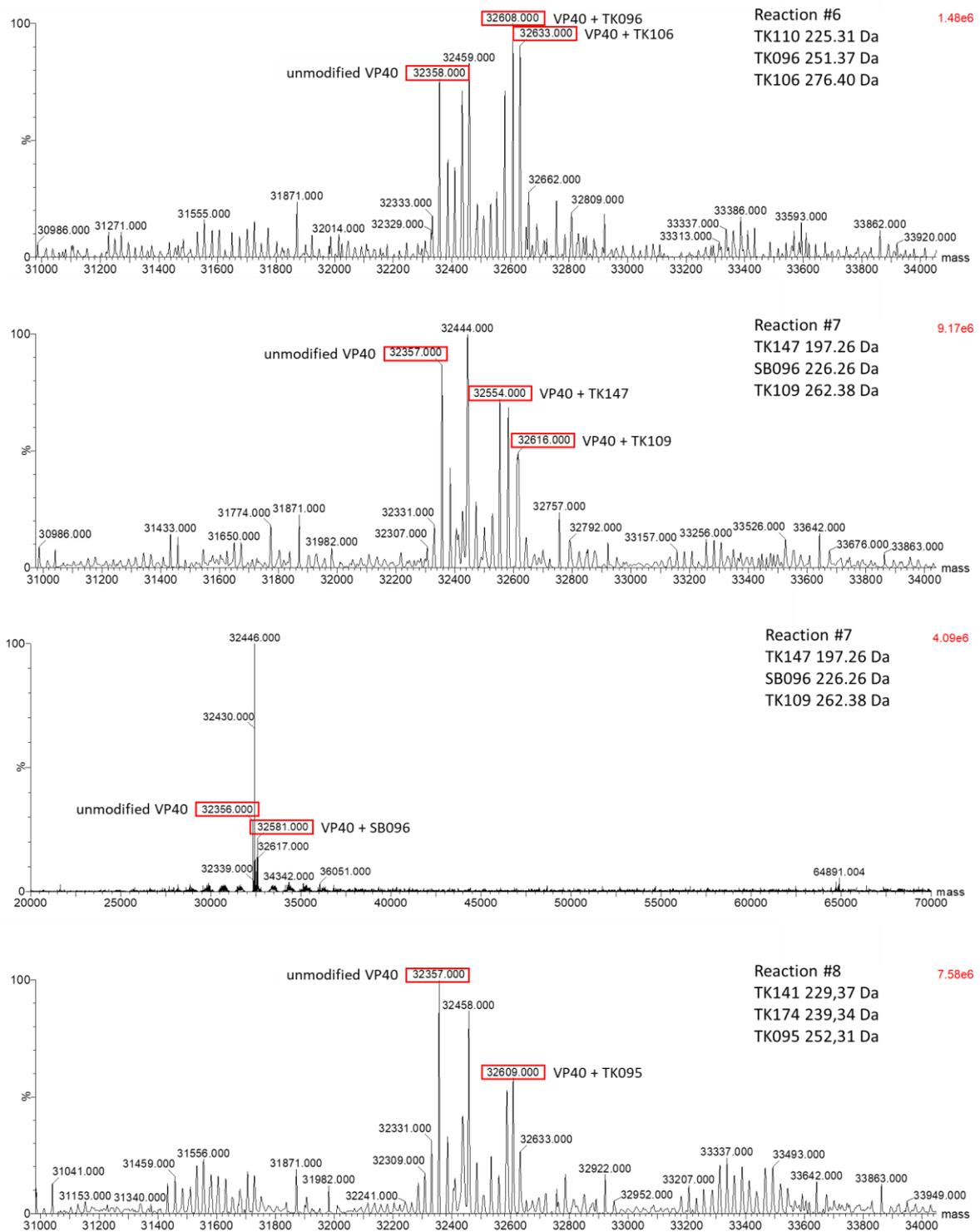


Figure 71: ESI mass spectra of DT hits for sVP40 $_{\Delta 43}$ N67C-CCA of reaction 6 to 8. 30 μ M of dimeric sVP40 $_{\Delta 43}$ N67C-CCA were incubated with a mixture of thiol-containing fragments and DTT. Afterwards, the mixtures were desalted and analyzed via IP-MS. Fragment hits were identified by a mass shift of unmodified protein plus fragment mass.

Appendix

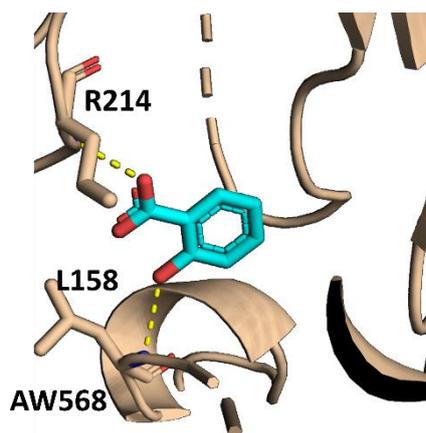


Figure 72: Structure of dimeric sVP40_{Δ43} WT with co-crystallized SA. Overall cartoon representation of dimeric sVP40_{Δ43} WT as a close-up of bound SA in the pocket between the NTD and CTD with H-bonds to L158 and R214

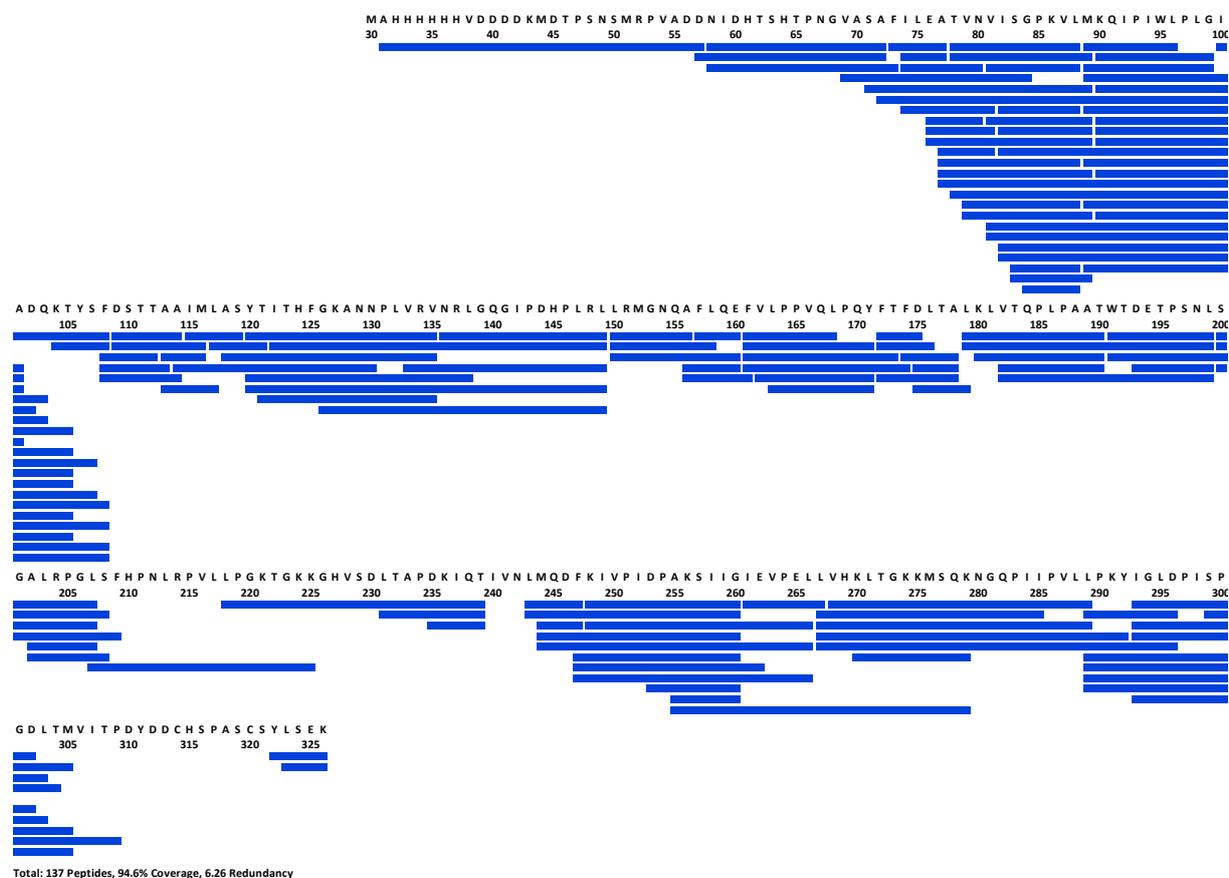


Figure 73: Peptide coverage map of HDX-MS experiments comparing sVP40_{Δ43} WT with and without SA. Each blue bar represents a peptide that has been analyzed.

Table 24: Selection of the most promising compounds after docking to sVP40. The first number represents the conformation of the sVP40 R214 sidechain, followed by the compound number for each respective docking analysis (1 to 1,000), the SA derivative used for the reaction, the type of reaction performed virtually to generate the final products shown as SMILES codes.

Cpd #	Scaffold	Reactions	SMILES of Product
1_010	5-chloro-SA	Negishi	<chem>O=C([O-])c1cc(-c2cc3c(Cl)ncnc3[nH]2)ccc1O</chem>
1_030	3,5-diiodo-SA	Buchwald-Hartwig	<chem>O=C([O-])c1cc(Nc2nc3scc(-c4cccs4)c3c(=O)[nH]2)cc(I)c1O</chem>
1_056	3-amino-SA	Buchwald-Hartwig	<chem>Cn1cnc2nc(Nc3cccc(C(=O)[O-])c3O)[nH]c(=O)c21</chem>

Appendix

1_102	3,5-diiodo-SA	Grignard_alcohol	<chem>c1cnc2c(C=C([C@H](c3cc(cc3O)C(=O)[O-])O)N2)c1Br</chem>
1_381	5-chloro-SA	Buchwald_Hartwig	<chem>Cc1ccc(-c2cc(=O)[nH]c(Nc3ccc(O)c(C(=O)[O-])c3)n2)cc1</chem>
1_428	5-chloro-SA	Buchwald-Hartwig	<chem>O=C([O-])c1cc(Nc2nc3ncc(Br)cc3c(=O)[nH]2)ccc1O</chem>
1_855	5-chloro-SA	Buchwald_Hartwig	<chem>Cn1ccc2c(=O)[nH]c(Nc3ccc(O)c(C(=O)[O-])c3)nc21</chem>
1_865	5-chloro-SA	Buchwald_Hartwig	<chem>c1c(cc(cc1N[C@@H]1CC[C@@H](C(=O)NC(=O)N)CC1)O)C(=O)[O-]</chem>
2_012	5-chloro-SA	Buchwald_Hartwig	<chem>O=C([O-])c1cc(Nc2ncc(CCl)c(=O)[nH]2)ccc1O</chem>
2_017	5-chloro-SA	Buchwald_Hartwig	<chem>CCc1cc(=O)nc(Nc2ccc(O)c(C(=O)[O-])c2)[nH]1</chem>
2_029	5-chloro-SA	Buchwald_Hartwig	<chem>Cc1sc2nc(Nc3ccc(O)c(C(=O)[O-])c3)[nH]c(=O)c2c1C</chem>
2_044	3,5-diiodo-SA	Buchwald_Hartwig	<chem>O=C([O-])c1cc(Nc2nc3nc(Cl)sc3c(=O)[nH]2)cc(l)c1O</chem>
2_093	3,5-diiodo-SA	Buchwald_Hartwig	<chem>Cc1sc2nc(Nc3cc(l)c(O)c(C(=O)[O-])c3)[nH]c(=O)c2c1C</chem>
2_122	3,5-diiodo-SA	Buchwald_Hartwig	<chem>Cn1cnc2nc(Nc3cc(l)c(O)c(C(=O)[O-])c3)[nH]c(=O)c21</chem>
2_270	3-amino-SA	Buchwald_Hartwig	<chem>c1cc(c(c(c1)NC1=CC(=O)N(C(=O)N1)C1CC1)O)C(=O)[O-]</chem>
2_328	5-chloro-SA	Buchwald-Hartwig	<chem>O=C([O-])c1cc(Nc2ccc3cc(Br)cnc3n2)ccc1O</chem>
2_342	3,5-diiodo-SA	Buchwald-Hartwig	<chem>Cc1nc(Nc2cc(l)c(O)c(C(=O)[O-])c2)[nH]c(=O)c1CCC(=O)[O-]</chem>
2_547	3,5-diiodo-SA	Buchwald_Hartwig	<chem>c1cc(cc2c1N=C(NC2=O)Nc1cc(c(c(c1)l)O)C(=O)[O-])n([o])[o]</chem>
3_009	5-chloro-SA	Buchwald_Hartwig	<chem>O=C([O-])c1cc(Nc2nc(-c3ccc(F)cc3)cc(=O)[nH]2)ccc1O</chem>
3_061	5-chloro-SA	Buchwald_Hartwig	<chem>O=C([O-])c1cc(Nc2nc3scc(-c4ccco4)c3c(=O)[nH]2)ccc1O</chem>
3_127	3,5-diiodo-SA	Grignard_alcohol	<chem>O=C([O-])c1cc(l)cc([C@H](O)c2cc3c(Br)ccnc3[nH]2)c1O</chem>
3_271	3,5-diiodo-SA	Buchwald_Hartwig	<chem>Cc1nc(-c2ccnc(Nc3cc(l)cc(C(=O)[O-])c3O)c2)n[nH]1</chem>
3_295	3-amino-SA	Schotten-Baumann_amides	<chem>Cc1n[nH]c(C(=O)Nc2cccc(C(=O)[O-])c2O)c1C</chem>
3_298	3-amino-SA	Buchwald-Hartwig	<chem>O=C([O-])c1cccc(Nc2nc(Cl)cc(=O)[nH]2)c1O</chem>
3_550	3-amino-SA	Schotten-Baumann_amides	<chem>O=C(Nc1cccc(C(=O)[O-])c1O)c1ccc(F)nc1</chem>
3_723	3,5-diiodo-SA	Buchwald-Hartwig	<chem>c1c(c(c(cc1l)NC1=NC(=C(C(=O)N1)n([o])[o])Cl)O)C(=O)[O-]</chem>
4_084	5-chloro-SA	Buchwald_Hartwig	<chem>O=C([O-])c1cc(Nc2cc3cccc3c(=O)[nH]2)ccc1O</chem>
4_188	5-chloro-SA	Buchwald_Hartwig	<chem>COC(=O)Cc1cnc(Nc2ccc(O)c(C(=O)[O-])c2)[nH]c1=O</chem>
4_369	5-chloro-SA	Grignard_alcohol	<chem>Cc1c([C@H](O)c2ccc(O)c(C(=O)[O-])c2)c(=O)[nH]c(=O)n1C1CCCCCCC1</chem>
4_373	5-chloro-SA	Buchwald_Hartwig	<chem>Cc1cc2[nH]c(Nc3ccc(O)c(C(=O)[O-])c3)cc(=O)n2n1</chem>
4_885	3,5-diiodo-SA	Buchwald_Hartwig	<chem>O=C([O-])c1cc(Nc2nc(=O)c3c(ncn3CCO)[nH]2)cc(l)c1O</chem>

10.2 Publications, presentations and poster

Publications

- 07/2021 Tscherne A, Schwarz JH, Rohde C, Kupke A, Kalodimou G, Limpinsel L, Okba NMA, Bošnjak B, Sandrock I, Odak I, Halwe S, Sauerhering L, Brosinski K, Liangliang N, Duell E, Jany S, Freudenstein A, Schmidt J, **Werner A-D**, Gellhorn Serra M, Klüver M, Guggemos W, Seilmaier M, Wendtner CM, Förster R, Haagmans BL, Becker S, Sutter G, Volz A. **Immunogenicity and efficacy of the COVID-19 candidate vector vaccine MVA-SARS-2-S in preclinical vaccination.** Proc Natl Acad Sci U S A. 2021 Jul 13;118(28):e2026207118. doi: 10.1073/pnas.2026207118.
- 07/2021 Halwe S, Kupke A, Vanshylla K, Liberta F, Gruell H, Zehner M, Rohde C, Krähling V, Gellhorn Serra M, Kreer C, Klüver M, Sauerhering L, Schmidt J, Cai Z, Han F, Young D, Yang G, Widera M, Koch M, **Werner A-D**, Kämper L, Becker N, Marlow MS, Eickmann M, Ciesek S, Schiele F, Klein F, Becker S. **Intranasal Administration of a Monoclonal Neutralizing Antibody Protects Mice against SARS-CoV-2 Infection.** Viruses. 2021 Jul 29;13(8):1498. doi: 10.3390/v13081498.
- 03/2020 Ashor AW, Shannon OM, **Werner A-D**, Scialo F, Gilliard CN, Cassel KS, Seal C J, Zheng D, Mathers J C, Siervo M. **Effects of inorganic nitrate and vitamin C co-supplementation on blood pressure and vascular function in younger and older healthy adults: a randomised double-blind crossover trial.** Clin Nutr. 2020;39:708–717. doi: 10.1016/j.clnu.2019.03.006.
- 03/2017 Ashor A, **Werner A-D**, Lara J, Willis N D, Mathers J C, Siervo M. **Effects of vitamin C supplementation on glycaemic control: a systematic review and meta-analysis of randomised controlled trials.** Eur J Clin Nutr 71, 1371–1380 (2017). doi: <https://doi.org/10.1038/ejcn.2017.24>.
- 10/2016 Afshar S, Kelly SB, Seymour K, Woodcock S, **Werner A-D**, Mathers JC. **The Effects of Bariatric Procedures on Bowel Habit.** Obes Surg. 2016 Oct;26(10):2348-54. doi: 10.1007/s11695-016-2100-9.

Publications in preparation

- /2022 **Werner A-D**, Krapoth N, Norris M, Heine A, Klebe G, Ollmann Saphire E, Becker S. **Development of a crystallographic screening to identify Ebola virus VP40 ligands.**
- /2022 **Werner A-D**, Steinchen W, Norris MJ, Ollmann Saphire E, Becker S. **Characterisation of a SUDV VP40 dimerization-deficient mutant.**
- /2022 **Werner A-D**, Schauflinger M, Norris M, Klüver M, Trodler A, Herwig A, Heine A, Klebe G, Ollmann Saphire E, Becker S. **Butterflies reaching out - the previously unresolved**

C-terminus of Sudan Ebolavirus VP40 reveals functionally important post-translational redox-modifications.

-/2022 Wilson E, Roth B, Cassart J-P, Limburg H, Schwerdtner M, Werner A-D, Harbig A, Böttcher-Friebertshäuser E, Stokes A. **pH inactivation of SARS-CoV-2 and SARS-CoV in virus spiked protein A eluates from a mAb purification process.**

-/2022 Dillenberger M, Werner A-D, Rahlfs S, Becker K, Fritz-Wolf K. **Structural Characterization and posttranslational redox regulation of *Plasmodium falciparum* hexokinase.**

Presentations and Posters

09/2022 **Development of Ebola virus VP40 inhibitors via structure-guided drug design.** Anke-Dorothee Werner, Lennart Laube, Maria Giovanna Papadopoulos, Nils Krapoth, Michael J. Norris, Wieland Steinchen, Andreas Heine, Gerhard Klebe, Erica Ollmann Saphire, Peter Kolb, Wibke Diederich, Stephan Becker

Presentation as invited speaker, Fillovirus meeting 2022, La Jolla, CA, USA

06/2022 **Development of Ebola virus VP40 inhibitors via structure-guided drug design.** Anke-Dorothee Werner, Maria Giovanna Papadopoulos, Nils Krapoth, Michael J. Norris, Andreas Heine, Gerhard Klebe, Erica Ollmann Saphire, Peter Kolb, Stephan Becker

Poster, Negative Strand Virus Meeting 2022, Braga, Portugal

06/2022 **Soaking of Ebola virus VP40 crystals with small molecules paves the way for rational drug design using *in silico* methods.** Anke-Dorothee Werner, Maria Giovanna Papadopoulos, Lennart Laube, Michael J. Norris, Michael Daude, Andreas Heine, Gerhard Klebe, Erica Ollmann Saphire, Wibke Diederich, Peter Kolb, Stephan Becker

Poster, DGI / DZIF 2022 Joint Annual Meeting 2022, Stuttgart

03/2022 **Development of Ebola virus VP40 inhibitors via structure-guided drug design.** Anke-Dorothee Werner, Maria Giovanna Papadopoulos, Lennart Laube, Michael J. Norris, Wibke Diederich, Andreas Heine, Gerhard Klebe, Erica Ollmann Saphire, Peter Kolb, Stephan Becker

Poster, Annual Meeting of German Society for Virology 2022, online

02/2022 **Development of Ebola virus VP40 inhibitors via structure-guided drug design.** Anke-Dorothee Werner, Stephan Becker

Presentation, DRUID Spring Symposium, online

11/2021 **Butterflies reaching out - the previously unresolved C-terminus of Sudan Ebolavirus VP40 reveals functionally important post-translational redox-modifications.** Anke-

- Dorothee Werner, Martin Schauflinger, Michael Norris, Michael Klüver, Anna Trodler, Astrid Herwig, Andreas Heine, Gerhard Klebe, Erica Ollmann Saphire, Stephan Becker
Presentation, 4th DRUID Retreat, online
- 10/2021 **Butterflies reaching out – the previously unresolved C-terminus of Sudan Ebolavirus VP40 reveals functionally important post-translational redox-modifications.** Anke-Dorothee Werner, Martin Schauflinger, Michael Norris, Michael Klüver, Anna Trodler, Astrid Herwig, Andreas Heine, Gerhard Klebe, Erica Ollmann Saphire, Stephan Becker
Cell Biology of Viral Infections, Kloster Schöntal 2021
- 09/2021 **Protein-Protein-Interaktionen des Ebolavirus Matrixproteins VP40 als Ziele antiviraler Strategien.** Anke-Dorothee Werner, Stephan Becker
Poster, DRUID Project Assessment, online
- 07/2021 **Soaking of Ebola virus VP40 crystals with small molecules paves the way for rational drug design using *in silico* methods.** Anke-Dorothee Werner, Maria Giovanna Papadopoulos, Michael J. Norris, Andreas Heine, Gerhard Klebe, Erica Ollmann Saphire, Peter Kolb, Stephan Becker
Presentation, Annual Meeting of American Society for Virology 2021, online
- 04/2021 **Soaking of Ebola virus VP40 crystals with small molecules paves the way for rational drug design using *in silico* methods.** Anke-Dorothee Werner, Maria Giovanna Papadopoulos, Michael J. Norris, Andreas Heine, Gerhard Klebe, Erica Ollmann Saphire, Peter Kolb, Stephan Becker
Poster, Chica and Heinz Schaller (CHS) Virology e-Symposium, online
- 03/2021 **Butterflies reaching out: The previously unresolved C-terminus of Sudan Ebolavirus VP40 reveals functionally important post-translational redox-modifications.** Anke-Dorothee Werner, Michael Norris, Anna Trodler, Michael Klüver, Andreas Heine, Gerhard Klebe, Erica Ollmann Saphire, Stephan Becker
Presentation, Annual Meeting of German Society for Virology 2021, online
- 10/2020 **Protein/protein interactions of Ebola virus proteins as targets for new antiviral strategies.** Anke-Dorothee Werner, Stephan Becker
Presentation, 3rd DRUID Retreat, online
- 03/2020 **Soaking of Ebola VP40 crystals to identify lead-like structures using a library of small molecules.** Anke-Dorothee Werner, Michael J. Norris, Alexander Metz, Erica Ollmann Saphire, Gerhard Klebe, Stephan Becker
Presentation, Annual Meeting of German Society for Virology 2020, cancelled
- 11/2019 **Screening for potential EBOV VP40 inhibitors via crystal soaking.** Anke-Dorothee Werner, Stephan Becker

Presentation, 2nd DRUID Retreat

03/2019 **Protein-protein interactions of the Ebola virus matrix protein VP40.** Anke-Dorothee Werner, Erica Ollmann Saphire, Stephan Becker

Poster, Annual Meeting of German Society for Virology 2019

11/2018 **Protein-protein interactions of Ebola VP40 as potential novel drug targets.** Anke-Dorothee Werner, Stephan Becker

Presentation, 1st DRUID Retreat

10.3 Curriculum vitae

10.4 Ehrenwörtliche Erklärung

„Ich erkläre ehrenwörtlich, dass ich die dem Fachbereich Medizin Marburg zur Promotionsprüfung eingereichte Arbeit mit dem Titel

Identification of lead molecules for the development of antivirals targeting the Ebola virus matrix
protein VP40

im Institut für Virologie unter Leitung von Prof. Dr. Stephan Becker ohne sonstige Hilfe selbst durchgeführt und bei der Abfassung der Arbeit keine anderen als die in der Dissertation aufgeführten Hilfsmittel benutzt habe. Ich habe bisher an keinem in- oder ausländischen Medizinischen Fachbereich ein Gesuch um Zulassung zur Promotion eingereicht, noch die vorliegende oder eine andere Arbeit als Dissertation vorgelegt. Ich versichere, dass ich sämtliche wörtlichen oder sinngemäßen Übernahmen und Zitate kenntlich gemacht habe.

Mit dem Einsatz von Software zur Erkennung von Plagiaten bin ich einverstanden.“

Ort, Datum

Anke-Dorothee Werner

„Die Hinweise zur Erkennung von Plagiaten habe ich zur Kenntnis genommen“

Ort, Datum

Prof. Dr. Stephan Becker

10.5 Acknowledgements

An dieser Stelle möchte ich mich bei allen danken, die mich in den letzten Jahren so toll unterstützt haben:

...allen voran natürlich bei Prof. Dr. **Stephan Becker**, der mir erstens die Doktorarbeit ermöglicht hat und mir so viel Freiheit bei der Gestaltung meines Themas gelassen, aber auch zwischendurch ein Machtwort gesprochen hat, wenn es zu chaotisch wurde. Vielen Dank für dein Vertrauen, deine Zeit und deine Diskussionsbereitschaft, danke einfach für alles!

... bei der Prüfungskommission

... beim LOEWE-Zentrum DRUID für die Finanzierung und die vielen tollen Symposien und Veranstaltungen

... der gesamten AG Becker: bei Dr. **Nadine Biedenkopf** für die Einarbeitung, das Korrekturlesen, die vielen Ratschläge zwischendurch und das immer offene Ohr; Dr. **Cornelius Rohde** für Korrekturlesen, so so so viele Tipps und Tricks bei allem und für Gossip; bei Dr. **Olga Dolnik** für die Einarbeitung ins BSL4; bei **Sabine Fischbach** für die unermüdliche Hilfe bei allem was Organisation betrifft; beim Schwabenlab für „angeregten Diskussionen“ und schöne Mittagspausen, hier vor allem danke an Dr. **Martin Schauflinger** und Dr. **Michael Klüver** für die gemeinsamen Projekte (Martin auch für's Korrekturlesen), sowie Lab Mate **Christopher Veeck** für gute Schredderstimmung im Labor, aber auch **Marcel Benz, Sebastian Pfeiffer & Sebastian Höck**; bei **Astrid Herwig** und **Martina Huxol** für Zellen und nochmal extra bei Astrid für ihre Hilfe und Unterstützung, sondern auch alles andere (vor allem in den letzten Wochen!); bei der Mittagspausengruppe Dr. **Cornelius Rohde** (ja, nochmal), Dr. **Alexandra Kupke, Isabel von Creytz, Lennart Kämper, Jana Schneider, Dr. Gesche Gerresheim**; bei der **Tiergruppe**; bei meinen studentischen und wissenschaftlichen Hilfskräfte **Anna Trodler** und **Nils Krapoth**, sowie den Praktikant*Innen **Johanna Dürrwald, Peter Sliwiak** und **Michael Neike** und Studierende für Abschlussarbeiten **Ann-Kristin Bank** für die tatkräftige Unterstützung im Labor; sowie bei allen anderen der **AG Becker, AG Biedenkopf** und **AG Krähling**; beim gesamten BSL4-Team Dr. **Markus Eickmann, Katharina Kowalski, Gotthard Ludwig, Michael Schmidt** und **Sebastian Schmidt** für viel Hilfe und Geduld; bei allen Ehemaligen, die viel zu früh das Institut verlassen haben: Dr. **Nelly Mostajo Henze, Dr. Katharina Grikscheit, Dr. Sandro Halwe**, sowie Dr. **Hannah Limburg** und **Jana Demper**

... bei Prof. Dr. **Erica Ollmann Saphire** und ihrer Arbeitsgruppe zuerst am Scripps Research Institute und dann am La Jolla Institute for Immunology, La Jolla, San Diego USA, für ihre unkomplizierte Art, ihre Unterstützung, Vertrauen und auch Gastfreundschaft. Danken möchte ich auch Dr. **Mike Norris** für

so viel Hilfe, die Einarbeitung, das Fluchen und die Diskussionen. Außerdem möchte ich **Sara Landeras-Bueno & Hal Wasserman** danken, sowie Amar Parvate, Sean Hui, Adrian Enriquez, Katie Fuller, Heather Callaway, Kathryn Hastie, Xiaoying Yu, Stephannie Harkins, Haoyang Li, Diptiben Parekh, Michelle Zandonatti and all the others for the wonderful time!

... bei meinen Kooperationspartnern aus Marburg: Prof. Dr. **Gerhard Klebe**, Prof. Dr. **Andreas Heine** und Prof. Dr. **Klaus Reuter**, Dr. **Alexander Metz**, sowie **Christian Sohn**, **Marina Gardony** und **Lea Huber** für das Bereitstellen des Kristallisationslabors, Expertise, Vektoren und Compounds sowie Hilfe bei meinen vielen Fragen; AG Kolb besonders Prof. Dr. **Peter Kolb**, **Maria Giovanna Papadopoulos** und **Kai Schmidt** für die tolle Unterstützung mit so viel Motivation; AG Diederich, vor allem Prof. Dr. **Wibke Diederich**, **Michael Daude** und **Lennart Laube** nicht nur für die Synthese, sondern auch für die tollen Diskussionen und eure Geduld, wenn ich von der Chemie mal wieder nichts verstehe, mit euch macht die Zusammenarbeit einfach total Spaß <3; AG Steinmetzer, vor allem Prof. Dr. **Torsten Steinmetzer** und Dr. **Thuy Van Lam van** für die Synthese der Peptide; Mass Spec Facility, vor allem Dr. **Uwe Linne**, **Tina Krieg** und Dr. **Wieland Steinchen** und allen anderen Mitarbeiter*Innen für die Messdaten, die unkomplizierte Kommunikation und eure Mühen, mir alles so gut es geht zu erklären; MarXtal Facility um **Ralf Pöschke** für die unkomplizierte Zusammenarbeit und vor allem die Spontanität!

... weiteren tollen Arbeitsgruppen aus Marburg: AG Schnare, Prof. Dr. **Markus Schnare** und **Sandra Stein**, für die Äkta, die Hilfe und Geduld; bei der AG Essen, die mich so toll eingebunden und mir in so vielen Aspekten geholfen haben allen voran natürlich Prof. Dr. **Lars-Oliver Essen**, Dr. **Viktoria Reithofer**, Dr. **Laura Werel**, best buddy **Lukas Korf**, Dr. **Maximilian Biermeier** und **Hans-Joachim Emmerich**. Wir hatten so tolle Messzeiten und ihr seid wie meine zweite Arbeitsgruppe <3; AG Bange, Prof. Dr. **Gert Bange** und Nils Maisch für die Benutzung des Microfluidizers und der unkomplizierten Stimmung

... bei meinen Kooperationspartnern aus Gießen, der AG Becker/Przyborski mit Prof. Dr. **Katja Becker**, Dr. **Jude Przyborski**, Dr. **Christina Brandstädter**, Dr. **Julia Hahn**, **Melissa Dillenberger** für SPR- und vor allem Redox-Studien. Es ist so schön, immer mal wieder zurück kommen zu können <3 Hierzu gehören natürlich auch Dr. **Susanne Schipper** und **Isabell Berneburg** für die moralische Unterstützung!

... bei Dr. **Nadine Biedenkopf**, **Lukas Korf**, **Isabel von Creytz**, Dr. **Cornelius Rohde** (zum 3. Mal, you famous), Dr. **Martin Schauflinger**, **Christopher Veeck**, **Marcel Benz** und Dr. **Susanne Schipper** für das Korrekturlesen dieser Arbeit.

... bei den Uddis Silks, Sili, Anni, Simi, Katja, Mohli, Matze, Stefan, Svenja, Manu, Bärbel, Max, sowie der Gießener Gang Isi, Kriss, Tommy, Sophia und Pauli für „normale“ Unterhaltungen abseits von Wissenschaft

... und am allerwichtigsten: bei Lukas und meiner Familie Mama, Papa, Oma, Sina, Hannes, Daniel, Sonni und vor allem Helen und Franz für einfach alles aber vor allem für die Geduld mit mir <3.