Structural Characterization of the Merkel Cell Polyomavirus Tumor Antigens

DISSERTATION

zur Erlangung des Doktorgrades (Dr. rer. nat.)

an der Fakultät für Mathematik, Informatik und Naturwissenschaften

Fachbereich Chemie

der Universität Hamburg

vorgelegt von

Veronika Brinschwitz

Hamburg, 2023

Reviewer 1: Prof. Dr. Nicole Fischer

Reviewer 2: Prof. Dr. Wolfram Brune

Date of disputation: 08.12.23

Date of print release: 13.05.24

Examination Committee:

      Chair: Prof. Dr. Kay Grünewald

      Members: Prof. Dr. Nicole Fischer

            Prof. Dr. Arwen Pearson

            Prof. Dr. Adam Grundhoff

            Prof. Dr. Wolfram Brune

The experimental work and the authoring of this dissertation was carried out in Hamburg between February 2018 and June 2023 at the University Medical Center Hamburg-Eppendorf, Institute of Medical Microbiology, Virology and Hygiene under the supervision of Prof. Dr. Nicole Fischer.

For those who are willing to search.

I. Table of content

II. Abbreviations

Abbreviation	Meaning
°C	degree celsius
µg	microgram
µl	microliter
µm	micrometer
A	adenine
Å	angstrom
aa	amino acid
ALTO	alternative frame of the LT open reading frame
Amp	ampicillin
APS	ammonium persulfate
ATM	ataxia telangiectasia mutated
ATP	adenosine triphosphate
ATR	ATM- and Rad3-Related
BKPyV	BK Polyomavirus
bp	base pair
BPV	Bovine papillomavirus
BSA	bovine serum albumin
C	cytosine
CaCl$_2$	calcium chloride
cDNA	complementary DNA
ChIP	chromatin immunoprecipitation
CK20	cytokeratin 20
cryoEM SPA	transmission electron cryo-microscopy single particle analysis
d.p.t.	days post transfection/transduction
DAPI	4′,6-diamidino-2-phenylindole
DDR	DNA damage response
DMEM	dulbecco's modified eagle's medium
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
dNTP	deoxyribonucleotide triphosphate (dATP, dCTP, dGTP, dTTP)
DOC	2,5-Dimethoxy-4-chloroamphetamine
DPBS	dulbecco´s phosphate buffered saline

dsDNA	double-stranded DNA
DTT	1,4-dithiothreitol
E. coli	Escherichia coli
EBV	Epstein-Barr virus
EDTA	ethylenediamine tetra-acetic acid
ER	endoplasmic reticulum
et al.	and others (lat.: „et alteri")
EtOH	ethanol
EV	extracellular vesicle
FSC	Fourier shell correlation
FBS	fetal bovine serum
G	guanine
g	gram
GAG	glycosaminoglycan
GFP	green fluorescent protein
h	hour
HaPyV	Lymphoma and epithelioma-inducing hamster polyomavirus
HCl	Hydrochloric acid
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HPV	human papillomavirus
hPyV	human polyomavirus
hTERT	Human telomerase reverse transcriptase
IF	immunofluorescence
IFN	interferon
IPTG	isopropanyl-β-D-thiogalactopyranoside
ISG	interferon-stimulated gene
JCPyV	JC Polyomavirus
Kan	kanamycin
kbp	kilobase pairs
kDa	kilodalton
KSHV	Kaposi´s sarcoma-associated herpesvirus
L	liter
LANA	latency-associated nuclear antigen
LB	lysogeny broth (Luria-Bertani)
LSD	LT stabilizing domain

LT	large tumor antigen
M	molar
mA	milliampere
MCPyV	Merkel cell polyomavirus
MDa	mega Dalton
MEM	minimal essential medium
MeOH	methanol
mg	milligram
$MgCl_2$	magnesium(II) chloride
MHC-I	major histocompatibility complex-class I
min	minute
ml	milliliter
mM	millimolar
$MnCl_2$	mangan(II) chloride
MPyV	Murine polyomavirus
mRNA	messenger-RNA
mTOR	mammalian target of rapamycin
MUR	MCPyV unique region
Na_2CO_3	sodium carbonate
Na_3VO_4	sodium orthovandate
NaCl	sodium chloride
NaF	sodium fluoride
ng	nanogram
nM	nanomolar
nm	nanometer
OBD	origin binding domain
OD	optical density
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PD-1	programmed cell death
PEI	polyethyleneimine
pH	power of /potential for hydrogen
pM	picomolar
pmol	picomole

PMSF	phenylmethylsulfonyl fluoride
PP2A	protein phosphatase 2A
PP4C	protein phosphatase 4 catalytic subunit
pRb	retinoblastoma protein
PTM	post-translational modification
PVDF	polyvinylidendifluoride membrane
RacPyV	Raccoon polyomavirus
rev	reverse
R_H	hydrodynamic radius
RNA	ribonucleic acid
rpm	revolutions per minute
RT	room temperature
s	second
SD	standard deviation
SDS	sodium dodecyl sulfate
SDS-PAGE	SDS-polyacrylamide-gel electrophoresis
SP3	superfamily 3
ssDNA	single-stranded DNA
sT	small tumor antigen
SV40	Simian virus 40
T	thymine
TAE	TRIS-acetate-EDTA
Taq	thermus aquaticus
TBS	TRIS buffered saline
TBS-T	TBS with Tween-20
TEM	transmission electron microscopy
TEMED	tetramethylethylendiamin
TEMED	N,N,N',N'-tetramethyl ethylenediamine
tLT	truncated LT
TRIS	Tris-(hydroxymethyl)aminomethane
TSPyV	Trichodysplasia spinulosa polyomavirus
UKE	Universitätsklinikum Hamburg-Eppendorf
UV	ultraviolet
V	volt
v/v	volume per volume

VZV	Varicella zoster virus
w/v	weight per volume
WB	western blot
x g	times relative centrifugal force
X-Gal	5-bromo-4-chloroindol-3-yl-β-D-galactopyranoside

1. Zusammenfassung

Das Merkelzellpolyomavirus (MCPyV) ist eines von 14 humanen Polyomaviren (hPyV) und stellt, zusammen mit JCPyV, BKPyV und Trichodysplasia spinulosa Polyomavirus, eines von vier klinisch relevanten hPyVs dar. Obwohl es im Großteil der Bevölkerung asymptomatisch persistiert, kann es in seltenen Fällen, wie zum Beispiel bei immunsupprimierten oder älteren Menschen, zur Entstehung des sehr aggressiven Hautkrebses, dem Merkelzellkarzinom (MCC), führen. MCPyV ist das einzige hPyV, welches Tumore im eigenen Wirt hervorruft. Dabei spielen das kleine Tumorantigen (sT) und eine tumorspezifische C-terminal verkürzte Variante des großen Tumorantigens (tLT) eine entscheidende Rolle. So konnte zum Beispiel gezeigt werden, dass die sT Expression in Rattenfibroblasten zu deren *in vitro* Transformation führt und die transgene sT Expression in p53-defizienten Mäusen MCC ähnliche Tumore hervorruft. Die Zellzyklusdysregulation durch die Inaktivierung des LT LxCxE-motivgebundenen pRbs stellt einen weiteren wichtigen Faktor in der Tumorgenese dar. Neben verkürzenden und tumorspezifischen LT Mutationen, die zum Verlust der *Origin binding domain* (OBD) und der Helikase führen und eine viralen Replikationsdefizienz zur Folge haben, ist das virale Genom monoklonal und tumorspezifisch in MCC Zellen integriert. Während des permissiven Lebenszyklus ist das LT ohne Trunkierungsmutationen an der viralen Replikation beteiligt. Es bindet den viralen *origin of replication* (ori) mittels OBD und entwindet anschließend das virale Genom mit Hilfe der Helikase.

Das MCC ist ein sehr aggressiver Tumor mit einer vom initialen Tumorstadium abhängigen 5-Jahres Überlebensrate von 62-25%. Vorhandene Behandlungsmöglichkeiten haben sich durch die Einführung der Immuntherapie bereits deutlich verbessert, jedoch zeigen PD-1/PD-L1 Immuncheckpoint Inhibitoren bei einem erheblichen Anteil der MCC Patienten keine Wirkung. Da die T Antigene eine entscheidende Rolle im viralen Lebenszyklus und in der Entstehung von MCC spielen, ist deren atomare Struktur als Ansatzpunkt für neue Behandlungsmethoden von hoher Bedeutung. Bislang konnte jedoch nur eine Teilstruktur eines MCPyV T Antigens, die LT OBD, gelöst werden und auch darüber hinaus ist keine vollständige Proteinstruktur eines PyV LTs bekannt. Lediglich die Struktur von SV40 sT ist verfügbar. Neben dem unmittelbaren therapeutischen Nutzen können mit Hilfe der Proteinstruktur auch bisher unbekannte T Antigenfunktionen und neue Protein-Protein Interaktionen aufgedeckt werden. Dies ist insbesondere für LT von Interesse, da die MCPyV LT *Merkel unique regions* (MURs) keine Sequenzähnlichkeit zu anderen PyV LTs aufweisen und deren Funktion größtenteils unbekannt ist. Im Rahmen dieser Arbeit wurden daher die MCPyV T Antigene, im speziellen LT, biochemisch und strukturell analysiert. Da aufgrund der geringen Proteinkonzentration und einer hohen Polydispersität vorangegangene *in vitro*

Kristallisationsansätze von *E. coli*-exprimiertem tLT nicht erfolgreich waren, wurde in dieser Arbeit *in cellulo* Kristallisation verwendet. Die kristalline Matrix der katalytischen Domäne der Vielzeller-spezifischen Kinase PAK4 und deren endogener Inhibitor Inka1 wurde als Kristallisationshilfe der MCPyV T Antigene in verschiedenen eukaryotischen Zelllinien verwendet. *In cellulo* Co-Kristalle von tLT, sowie sT und trunkierten sT Versionen, konnten beobachtet werden, zeigten jedoch lediglich eine geringe bis keine Röntgenbeugung. Aufbauend auf vorangegangenen Erkenntnissen wurde danach eine transmissionselektronen-kryomikroskopische Einzelpartikelanalyse (cryoEM SPA) des LTs durchgeführt. Das Protein wurde in Insektenzellen exprimiert, chromatographisch aufgereinigt und die Polydispersität und Aggregation durch Additive verringert. Zusätzlich wurde das LT-DNA Bindeverhalten multipler Oligonukleotide untersucht. Es konnte gezeigt werden, dass der virale ori LT am effektivsten stabilisiert, die höchste Bindungsspezifität aufweist und dass die An- und Abwesenheit von DNA die Bildung von LT Hexameren nicht beeinträchtigt. Außerdem wurde eine LT Stabilisierung in phosphathaltigen Puffern festgestellt, die jedoch, zusammen mit den vorangegangenen Optimierungsschritten, nicht zu einer aggregatfreien und monodispersen Proteinlösung führt. Ungeachtet dieser Einschränkungen konnte mittels cryoEM SPA das bisher erste MCPyV LT Helikase Modell mit einer Auflösung von 6,8 Å beschrieben werden. Interessanterweise zeigt das Hexamer trotz der geringen Sequenzidentität von 33% eine ausgeprägte strukturelle Ähnlichkeit zu der SV40 LT Helikase. Gründe für die geringe Auflösung sind eine bevorzugte Partikelorientierung, Proteinaggregation und Polydispersität, die möglicherweise durch die, als unstrukturiert vorhergesagten, MURs hervorgerufen werden. Darüber hinaus sind die MURs möglicherweise auch ein Grund für das Fehlen der OBD und des LT N-Terminus in dem cryoEM SPA Modell.

Zusammenfassend zeigt diese Arbeit das erste MCPyV LT Helikase Modell mit einer Auflösung von 6,8 Å mittels cryoEM SPA. Trotz geringer Auflösung, beruhend auf Proteinaggregation, Polydispersität und einer bevorzugten Partikelorientierung, stellt das Modell einen ersten wichtigen Schritt zur Entwicklung dringend benötigter spezifischer antiviraler Therapiemöglichkeiten dar.

2. Abstract

The Merkel cell polyomavirus (MCPyV) is one of 14 human polyomaviruses (hPyV) and represents, together with JCPyV, BKPyV and Trichodysplasia spinulosa polyomavirus, one of four clinically relevant hPyVs. Although it persists asymptomatically in most of the population, in rare cases, such as in immunocompromised or elderly people, it can lead to the development of the very aggressive skin cancer, Merkel cell carcinoma (MCC). MCPyV is the only hPyV that causes tumors in its own host. Thereby, the small tumor antigen (sT) and a tumor-specific C-terminally truncated version of the large tumor antigen (tLT) play an essential role. For example, it has been shown that sT expression in rat fibroblasts leads to their *in vitro* transformation and that transgenic sT expression in p53-deficient mice provokes MCC-like tumors. Cell cycle dysregulation due to the inactivation of the LT LxCxE motif-bound pRb represents another important factor in tumorigenesis. In addition to truncating and tumor-specific LT mutations, leading to the loss of the origin binding domain (OBD) and the helicase which results in viral replication deficiency, the viral genome is monoclonal and tumor-specifically integrated into MCC cells. During the permissive life cycle, the full-length LT is involved in viral replication. It binds the viral origin of replication (ori) via its OBD and subsequently unwinds the viral genome by its helicase.

MCC is an aggressive tumor with a 5-year survival rate of 62-25%, depending on the initial tumor stage. The introduction of immunotherapy has already substantially improved existing treatment options, however, a significant proportion of the MCC patients do not respond to the PD-1/PD-L1 immune checkpoint inhibitors. Since the T antigens play a crucial role in the viral life cycle and the development of MCC, their atomic structure can serve as a starting point for the development of new treatment methods. However, only a single domain of a MCPyV T antigen, the LT OBD, could be solved, and no complete protein structure of any PyV LT is known. For sT, only the SV40 structure is available. In addition to serving as a drug target, the protein structure can reveal previously unknown T antigen functions and novel protein-protein interactions. This is of particular interest for LT, as the MCPyV LT Merkel unique regions (MURs) have no sequence similarity to other PyV LTs, and their function is largely unknown. Therefore, in this work, the MCPyV T antigens, specifically LT, were biochemically and structurally analyzed. Since previous *in vitro* crystallization approaches of *E. coli*-expressed tLT were not successful, ascribed to low protein concentration and high polydispersity, *in cellulo* crystallization was used in this work. The crystalline matrix of the catalytic domain of the metazoan-specific kinase PAK4 and its endogenous inhibitor Inka1 was used as a crystallization aid for the MCPyV T antigens in various eukaryotic cell lines. *In cellulo* co-crystals of tLT, as well as sT and truncated sT versions, could be observed, but showed little to no X-ray diffraction. Based on preceding

findings, transmission electron cryo-microscopy single particle analysis (cryoEM SPA) of LT was thereafter performed. The protein was expressed in insect cells, chromatographically purified, and the polydispersity and aggregation were reduced by additives. In addition, the LT-DNA binding behavior of multiple oligonucleotides was investigated. It could be shown that the viral ori stabilizes LT most effectively and reveals the highest binding specificity. Furthermore, it was demonstrated that the presence and absence of different oligonucleotides did not affect LT hexamer formation. In addition, LT was found to be stabilized in phosphate-containing buffers, which, together with the preceding optimization steps, did not result in an aggregate-free and monodisperse protein solution. Notwithstanding these limitations, the first MCPyV LT helicase model with a resolution of 6.8 Å has been generated by using cryoEM SPA. Interestingly, despite the low sequence identity of 33%, the hexamer shows a pronounced structural similarity to the SV40 LT helicase. Reasons for the intermediate resolution are preferential particle orientation, protein aggregation, and polydispersity, possibly caused by the MURs, which are predicted to be unstructured. In addition, the absence of the OBD and the LT N-terminus in the cryoEM SPA model may also be attributed to the MURs.

In summary, this work provides the first MCPyV LT helicase model with a resolution of 6.8 Å, using cryoEM SPA. Despite its intermediate resolution, based on protein aggregation, polydispersity, and preferential particle orientation, the model represents an important initial step towards the development of specific antiviral therapies.

3. Introduction

3.1. Polyomaviruses

Polyomaviruses (PyVs) are small non-enveloped DNA viruses with an icosahedral virion size of 40-45 nm, composed of 72 pentameric capsomers. They belong to the family of *Polyomaviridae*, are ubiquitously present, and infect mammals, birds, and fish. PyVs comprise a circular double-stranded genome with a size of approximately 5 kilobase pairs (kbp), encoding only 5-9 proteins. They show restricted host range and cell tropism. With the Murine polyomavirus (MPyV), the first representative was discovered in 1953 (Gross, 1953). It could be shown that it induces multiple tumors in newborn mice and hamsters from where the name poly (=many) and oma (=tumor) originates (Stewart et al., 1957, 1958). The second polyomavirus, Simian virus 40 (SV40), was discovered and isolated from latently infected rhesus monkey kidney cells as a contaminating agent of the Polio vaccine (Eddy et al., 1962; Sweet & Hilleman, 1960). It could be shown that SV40 induces tumors in newborn hamsters and transforms human cells *in vitro*, leading to serious public health concerns. However, long-term follow-up studies on affected subjects proved no increased cancer occurrence in these individuals, leading to the assumption that SV40 is not associated with human tumorigeneses (Ea et al., 2003; Mortimer et al., 1981). Independent of health concerns, MPyV and SV40 served as model organisms to study eukaryotic DNA replication, viral life cycle, and tumorigenesis, substantially contributing to today's knowledge about basic molecular processes (Howley & Livingston, 2009). The first human polyomaviruses (hPyVs), JC polyomavirus (JCPyV) and BK polyomavirus (BKPyV), were discovered in 1971 and named after their index case patients. JCPyV was recovered from the brain of an immune-suppressed patient suffering from progressive multifocal leukoencephalopathy (PML) (Padgett et al., 1971). BKPyV was isolated from the urine of a renal transplant patient and can cause polyomavirus-associated nephropathy (PVAN) or hemorrhagic cystitis in kidney and hematopoietic stem cell transplant patients, respectively. Until now, it is under debate whether these PyVs are associated with tumorigenesis such as colorectal, prostate, and bladder cancer (Taherkhani & Farshadpour, 2022; C. Shen et al., 2021; Shoraka et al., 2020). According to the virus taxonomy 2021 release, 117 existing PyVs are grouped in six genera, including 14 hPyVs (online available at http://ictv.global/taxonomy accessed on 24 February 2023). In Figure 1, the phylogenetic relationship of the hPyVs (blue) and their closest animal relatives (black) is shown, based on the primary structure of the large tumor antigen (LT). The recently discovered Quebec polyomavirus (Ondov et al., 2019) is not included.
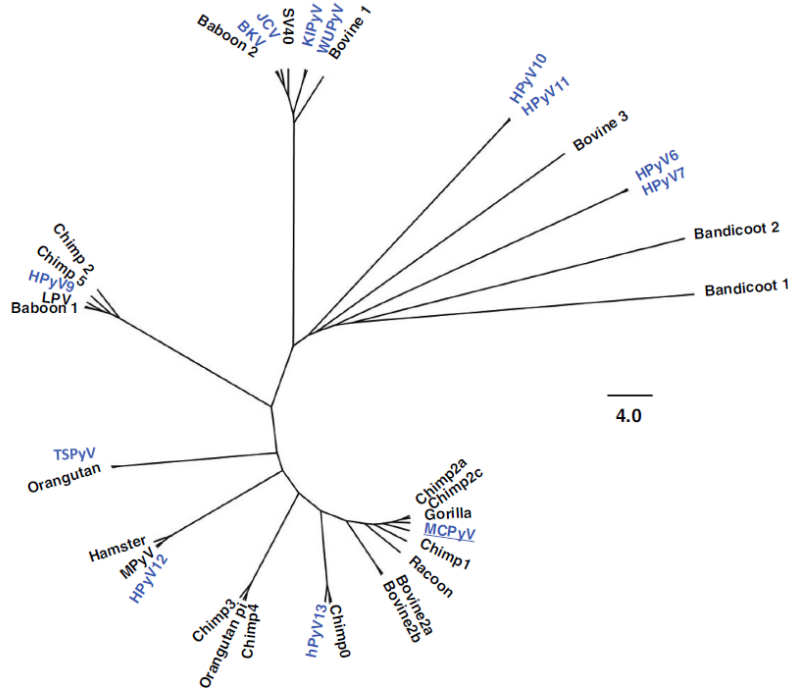
Figure 1: Phylogenetic tree of human polyomavirus large tumor antigen amino acid sequences and their closest animal relatives. Depicted are all human polyomaviruses known until 2015 (blue) and their closest animal relatives (black) (Grundhoff & Fischer, 2015).

The seroprevalence of PyVs in healthy adults varies between 25-90% (DeCaprio & Garcea, 2013; Kean et al., 2009), but the route of infection and transmission is frequently nuclear. Respiratory transmission was suggested for JCPyV and BKPyV since these viruses are often found in tonsillar tissue (Monaco et al., 1998; Goudsmit et al., 1982). PyV shedding in urine and feces argues for a fecal/urine-oral transmission (Vanchiere et al., 2009; Kitamura et al., 1990), while contaminated food and water is also suggested as source for PyV infections (Bofill-Mas et al., 2001). Skin-to-skin contact is proposed as a transmission pathway for cutaneous PyVs, such as MCPyV, which is frequently shed from the skin of healthy individuals (Schowalter et al., 2010). Primary asymptomatic infection occurs often in childhood, followed by a clinically inapparent life-long persistence. However, in immunocompromised individuals, PyVs can re-activate, causing pathologies such as trichodysplasia spinulosa (TS), PML, PVAN, or Merkel cell carcinoma (MCC). These complications are linked to Trichodysplasia spinulosa polyomavirus (TSPyV) (van der Meijden et al., 2010), BKPyV, JCPyV, or Merkel cell polyomavirus (MCPyV) (H. Feng et al., 2008), respectively, which are the only clinically relevant hPyVs known to date. Of these four hPyVs, MCPyV has the unique feature of causing tumorigenesis in its host, which will be discussed in more detail in the following paragraphs.

3.2. Merkel Cell Polyomavirus

Cancer is a serious health concern worldwide and infectious agents account for approximately 17% of all human tumors (Tempera & Lieberman, 2021). Next to *Helicobacter pylori*, causing gastric cancer and gastric mucosa-associated lymphoid tissue lymphoma (Hatakeyama, 2019), seven human tumor viruses have been identified and classified as carcinogenic agents. Among the most prominent cancer-causing viruses are human papillomavirus (HPV), which can be the causative agent of cervical carcinomas. Hepatitis B virus (HBV) and hepatitis C virus (HCV) are responsible for hepatocellular carcinoma. The Epstein-Barr virus (EBV) is associated with Burkitt´s and Hodgkin lymphoma and nasopharyngeal carcinoma. The human adult T-cell leukemia virus type 1 (HTLV-1) and the Kaposi´s sarcoma-associated herpesvirus (KSHV) also belong to the human tumor viruses and cause T-cell leukemia (HTLV-1) and Kaposi´s sarcoma primary effusion lymphoma and Castelman´s disease (KSHV) (Zapatka et al., 2020). The last representative, MCPyV, causes MCC and was discovered in 2008 as the fifth hPyV by digital transcriptome subtraction of sequenced MCC specimens (H. Feng et al., 2008). It is isolated from 80% of the skin of healthy subjects, indicating that it is a constituent of the human skin microbiome (Pastrana et al., 2012; Schowalter et al., 2010). A seropositivity for the viral capsid protein VP1 is detected in 42-80% of all tested individuals (Kean et al., 2009; Tolstov et al., 2009), whereas only 0.9-2% show antibodies against the early tumor antigens (T antigens) (Paulson et al., 2010). The high MCPyV seropositivity among humans could underlie multiple transmission routes, such as fecal-oral, based on detecting MCPyV DNA in the gut, gastrointestinal tract, and sewage (Campello et al., 2011; Loyo et al., 2010; Bofill-Mas et al., 2010). Krump and colleagues suggest cutaneous infection of UV-irradiated or abraded skin, consistent with a constant shedding of nuclease-resistant MCPyV DNA from human skin and its presence on environmental surfaces (Krump & You, 2021; Foulongne et al., 2011; Schowalter et al., 2010). Detection of MCPyV in the aerodigestive tract argues for a respiratory transmission route (Babakir-Mina et al., 2010; Kantola et al., 2009; Tolstov et al., 2009).

After first exposure in early childhood, MCPyV establishes a life-long persistence indicated by a rising seroprevalence with age (Nicol et al., 2013; T. Chen et al., 2011; Viscidi et al., 2011). In rare cases, preceding immunosuppression, UV-irradiation, or old age, the asymptomatic persistent infection results in MCC establishment. The ability to induce tumors in its host is shared by the raccoon polyomavirus (RacPyV), causing brain tumors (Dela Cruz et al., 2013), the lymphoma and epithelioma-inducing hamster polyomavirus (HaPyV) (Barthold et al., 1987; Graffi et al., 1969), and MPyV (Stewart et al., 1957). Despite the close relationship between MCPyV and RacPyV (Figure 1), the latter exists as an episome in transformed cells, contrary to the genomically integrated MCPyV genome, which will be discussed in the following (H. Feng et al., 2008).

3.2.1. Genome structure

The double-stranded MCPyV genome, which is present as episomes during the replicative viral life cycle and clonally integrated during MCC, comprises the early region, the late region, and the non-coding control region (NCCR) (Figure 2A). The former encodes the non-structural proteins small tumor antigen (sT), LT, 57K, and the alternative frame of the LT open reading frame (ALTO), which are mainly expressed in the early course of infection. Due to the restricted genome size of 5.4 kbp, sT, LT, and 57k are produced via alternative splicing. At the same time, ALTO is made by utilizing an alternate start codon and a +1 frameshift to the second LT exon (Carter et al., 2013). The major drivers of tumorigenesis are sT and LT, while LT holds an additional role in viral replication. Contrarily, the functions of 57K and ALTO are primarily unknown. A detailed description of the early gene products will be provided in 3.3.

The late coding region is arranged in anti-sense to the early region and expresses the structural proteins VP1 and VP2. They are expressed after viral replication initiation and form the viral capsid. Similar to other PyVs (Sullivan et al., 2009; Seo et al., 2008), MCPyV encodes a viral micro RNA (miRNA), miR-M1, which is located anti-sense to the early region within the second LT exon. It is expressed late during infection and is shown to autoregulate early gene expression and viral replication by using the RNA-induced silencing complex (RISC) (Theiss et al., 2015; Seo et al., 2009). It is suggested that miR-M1 is an essential viral tool for establishing persistence (Theiss et al., 2015). In addition, miR-M1-5p, having a two nucleotide shift compared to miR-M1, is detected in MCC cell lines (Lee et al., 2011). The NCCR separates the early and late region and harbors, next to bi-directional promoter elements for the early and late gene expression, the origin of replication (ori). Rearrangements of the NCCR, which are linked to enhanced pathogenesis and viral replication in BKPyV and JCPyV, could not be observed for MCPyV thus far (Gosert et al., 2010, 2008). The ori is essential for the initiation of viral replication and harbors ten GRGGC pentanucleotides (M1-10) (Kwun et al., 2009) (Figure 2D), differing from only six pentanucleotides in SV40, BKPyV, and JCPyV, and seven within the MPyV ori (C. Harrison et al., 2011). The minimal core ori of 71 bp is indispensable for replication and consists of the 5′ AT-rich region, the 1/2 site, the 3′ AT-rich region, and M7-8, which are located in the early enhancer. M1-6 are asymmetrically arranged as inverted repeats within Site 1/2. This arrangement allows a head-to-head formation of a LT double-hexamer, vital for DNA unwinding, which is shown intensively for SV40 (Cuesta et al., 2010; Joo et al., 1998; Dean et al., 1987). Interestingly, M3 provides an imperfect GRGGC sequence that is not bound to the origin-binding domain (OBD) in the crystal structure (Figure 5B), and which is shown to be irrelevant for viral replication, similar to M5-6 (C. Harrison et al., 2011; Kwun et al., 2009). In contrast, replication of SV40 and MPyV requires all four motifs located within the core ori (Joo et al., 1998;

Dean et al., 1987). M1-6 are flanked by two easily melted AT-rich regions, which prime ori unwinding via the LT helicase (Deb et al., 1986). While SV40 M5-6, located in site 1, are indispensable for viral replication, MCPyV M7, located in the analog early enhancer, is required for replication initiation (Kwun et al., 2009). Interestingly, M7 is naturally disrupted in the tumor-derived MCC350 viral strain (H. Feng et al., 2008). Another difference between MCPyV and SV40 reflects the DNA motif spacing, which is vital for a coordinated OBD binding (Parsons & Tegtmeyer, 1992). MCPyV ori comprises a two-nucleotide spacer between the inner and outer motif sets, while M2 and 3 overlap. In contrast, SV40 and related PyVs harbor one nucleotide between each pentanucleotide, making an ori interchange between viruses impossible (Shuda et al., 2008). Taken together, the ori differences between SV40, MPyV, and MCPyV hint towards a distinct and more complex MCPyV replication mechanism compared to other PyVs.
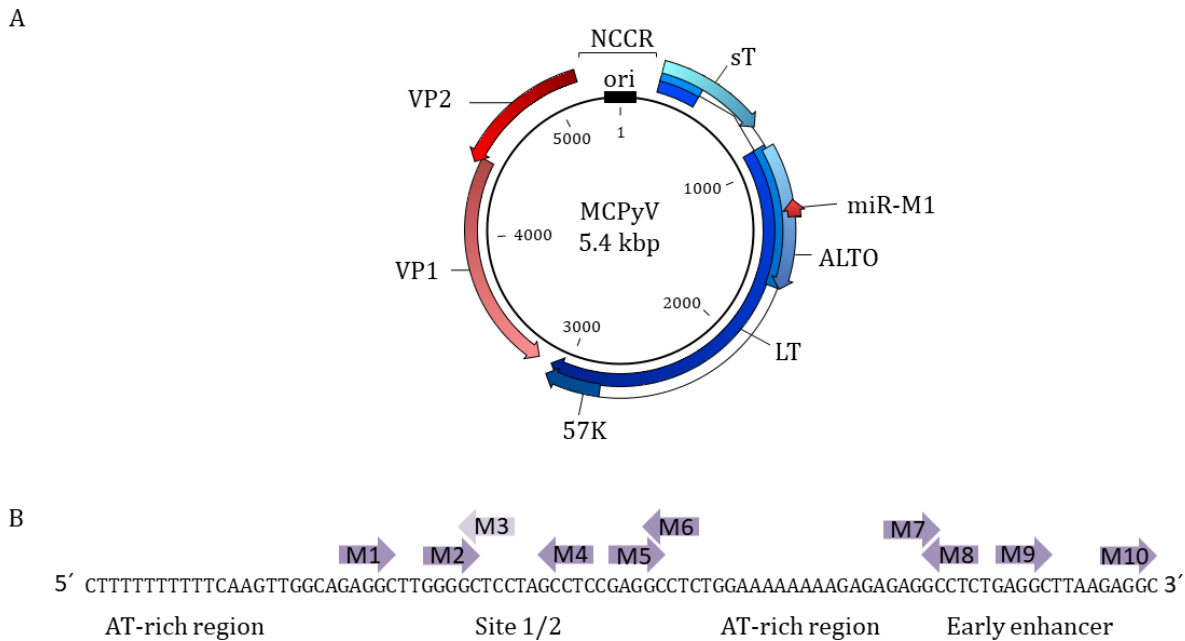
Figure 2: Structure of MCPyV genome. (A) The early region, encoding the small tumor antigen (sT, turquoise), large tumor antigen (LT, dark blue), 57K (blue), and the alternative frame of the LT open reading frame (ALTO, light blue) is depicted in sense direction. Additionally, the micro RNA (miR-M1, red) is located anti-sense to the second LT exon. The late region is arranged antisense to the early region and harbors the viral capsid proteins VP1 (light red) and VP2 (red). The early and late region flank the non-coding control region (NCCR, black), including the origin of replication (ori). The genome has a total size of 5.4 kilobase pairs (kbp). (B) MCPyV ori. The ori comprises a 5´ and 3´ AT-rich region, which flank site 1/2, and the early enhancer. It harbors nine complete GRGGC pentanucleotides (M1, 2, and 4-10, purple) and one imperfect GRGGC motif (M3, light purple). The arrows mark sense or antisense of the motifs.

3.2.2. MCPyV life cycle

Incomplete knowledge about the MCPyV host reservoir and the strict host tropism impedes a comprehensive examination of the viral life cycle and MCC establishment. Therefore, many insights into the viral biology are obtained by exploiting patient-derived MCC cell lines such as MKL-1 (Rosen et al., 1987), MS-1 (Guastafierro et al., 2013), and WaGa (Houben et al., 2010), which are challenging to grow *in vitro* and are defective in lytic replication (3.4.1).

Liu and colleagues suggest human dermal fibroblasts as host cells which support virus entry, transcription and translation as well as infectious particle production (W. Liu et al., 2016). However, these results remain to be verified by independent research groups. Succeeding studies screened various isolated animal dermal fibroblasts to identify a potential animal model to study the viral life cycle, showing that only chimpanzee dermal fibroblasts support the entire MCPyV viral life circle (W. Liu et al., 2018). This is not surprising since the chimpanzee PyV (ChPyV) is closely related to MCPyV (Figure 1). The use of MCPyV-mammalian chimeric viruses could therefore overcome the challenges of an animal model (W. Liu et al., 2018). Nonetheless, xenograft and transgenic mouse models were established to study MCPyV tumorigenesis and metastasis (Knips et al., 2017; Verhaegen et al., 2017; Guastafierro et al., 2013; Houben et al., 2012; Verhaegen et al., 2022). In addition, the semi-permissive *in vitro* model can be utilized to study viral replication and transcription. The human cell lines H1299, HEK293, or PFSK-1 cells are transfected with an *E. coli*-amplified, re-circularized MCPyV genome, permitting replication without the formation of infectious virions (Neumann et al., 2011).

MCPyV cell entry was examined by exploiting a recombinant, GFP-incorporating MCPyV pseudovirus, comprising VP1 and VP2. It was shown that MCPyV binds to sulfated glycosaminoglycans (GAGs) as an initial attachment receptor, similar to HPV16 (K. M. Johnson et al., 2009), followed by the interaction with sialylated glycans (Schowalter et al., 2011). These receptor-virus interactions are mediated by exposed apical VP1 loops (Bayer et al., 2020). Since sulfated and sialylated glycans are ubiquitously expressed on cell surfaces while *in vitro* and *in vivo* MCPyV propagation is still challenging, it is evident that viral entry is not the limiting factor. Additionally, thus far unknown prerequisites need to be fulfilled to support the viral life cycle. Alongside, Becker et al. suggest a bottleneck of virion trafficking from the endosome to the ER, which could hinder the viral life cycle (M. Becker et al., 2019).

After initial attachment, MCPyV enters the cell by caveolar/lipid raft-mediated endocytosis, similar to SV40 (Norkin et al., 2002; Anderson et al., 1996), and travels the endosomal pathway to the endoplasmic reticulum (ER). The release of viral DNA into the host nucleus, however, is still elusive. It was suggested that MCPyV depends on mitotic nuclear membrane breakdown, similar to HPV16 (M. Becker et al., 2019; Aydin et al., 2014; Pyeon et al., 2009). Once the histone-decorated viral DNA

has entered the nucleus, early gene transcription takes place, which can be activated by LT but is significantly controlled by cellular factors (Ajuh et al., 2018; Moens et al., 2015).

Viral replication is initiated by recognizing conserved GRGGC DNA motifs located within the ori of the viral NCCR via the LT OBD (C. Harrison et al., 2011; Kwun et al., 2009) (3.2.1). Thereafter, the helicase of the LT C-terminus assembles onto the viral ori as an active double-hexamer with a head-to-head orientation for DNA unwinding. Next to helicase activity, LT recruits cellular replication factors onto the DNA, such as Brd4, to enable host polymerase-mediated viral replication, as described in 3.3.13.3.1 (X. Wang et al., 2012). The interaction of LT and pRb via the LxCxE motif inhibits the pRb-mediated S phase entry control, which is beneficial for viral replication and which is mutated in various cancers (Borchert et al., 2014; Shuda et al., 2008). Viral replication is also supported via conserved iron-sulfur clusters of the sT C-terminus (Tsang et al., 2016) and by sT LSD-mediated LT stabilization (3.3.2) (Nwogu et al., 2020b; Kwun et al., 2013; H. Feng et al., 2011; Kwun et al., 2009). In addition, activation of the DDR plays a crucial role during viral replication. It is demonstrated that components of the DDR, such as ataxia telangiectasia and Rad3-related protein (ATR) and pChk2, colocalize with LT in viral replication centers. However, the exact molecular mechanism of how DDR contributes to viral replication remains unknown (Tsang et al., 2014).

Tumor-derived LT reveals truncating mutations within its C-terminus, leading to replication deficiency, which is supported by nucleotide mutations within the ori of, for example, the MCC cell line MCC350 (H. Feng et al., 2008; Shuda et al., 2008). Expression of the late genes, comprising VP1 and VP2, leads to the formation of new virions whose icosahedral capsid structure could be elucidated in 2020 and is shown in Figure 3A. The structure reveals a distinct disulfide pattern attributed to the Cys24, which is not conserved among PyVs. Additionally, MCPyV VP1 comprises a basic C-terminal extension, which lacks homology and which is essential for VP1 stability (Bayer et al., 2020). MCPyV does not express the N-terminal truncation of VP2, called VP3 (Schowalter & Buck, 2013). This contradicts SV40 and MPyV, where VP2 and VP3 contribute to the infectious entry process (Inoue & Tsai, 2011; Nakanishi et al., 2007; Daniels et al., 2006; Sahli et al., 1993). A role of VP2, located in the capsid interior, in the infectious cell entry could also be demonstrated for MCPyV but does not apply to all tested cell lines (Schowalter & Buck, 2013) (Figure 3C). The egress of newly formed virions completes the viral life cycle and is realized by cell lysis for SV40, BKPyV, and JCPyV. Alongside, hijacking extracellular vesicles to overcome cellular boundaries is reported for BKPyV and JCPyV and is also investigated for MCPyV, whose exact egress mechanism is still under debate (Handala et al., 2020; O'Hara et al., 2020; Morris-Love et al., 2019).

The fact that 60-70% of the human population is MCPyV positive hints towards a latent state of the virus. Despite the lack of knowledge about the site of viral persistence, Theiss and colleagues could

show that the viral miR-M1 supports the presence of the viral genome and LT in MCPyV genome transfected PFSK-1 cells over several months. Additionally, it controls early gene transcripts, suggesting an essential role in viral latency (2015). Along this line, SCF E3 ligase-mediated LT degradation is proposed as a novel MCPyV latency mechanism (Kwun et al., 2017).

Although profound knowledge of MCPyV has been accumulated over the past decade, important insights into the viral life cycle, such as the cellular reservoir, factors for virus entry, and egress mechanisms, are still missing, making it difficult to study MCPyV.
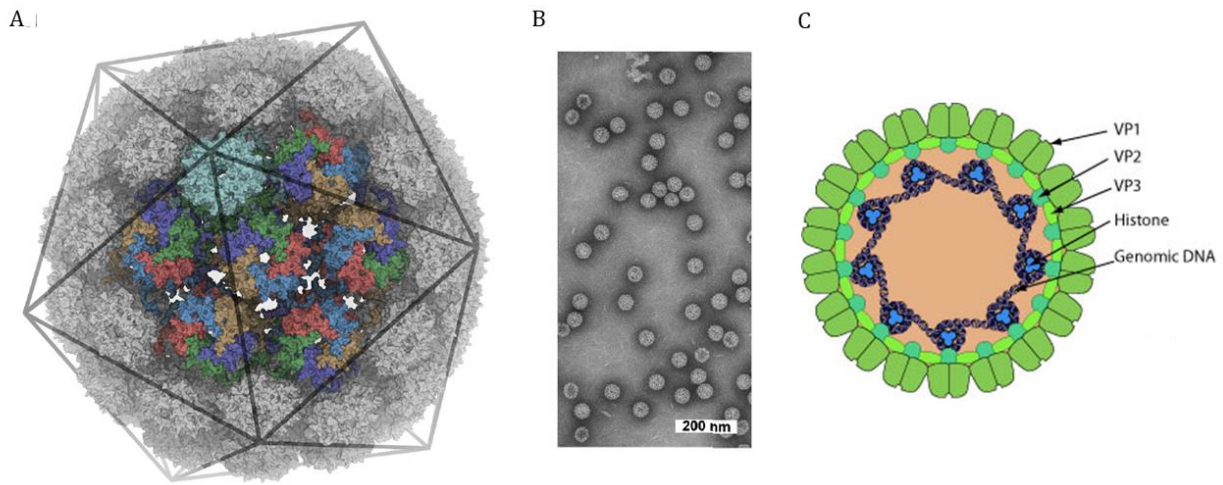
Figure 3: PyV virion and capsid structure. (A) Crystal structure of the icosahedral MCPyV capsid. It comprises 72 capsomers, built out of 360 VP1s. Six different VP1 conformations, α (cyan), α' (purple), α'' (green), β (red), β' (blue), and γ (brown), exist. The 12 α-VP1 pentamers are located at the capsomer vertices at a 5-fold symmetry axis and are surrounded by 60 α'-α''-β-β'-γ pentamers (Bayer et al., 2020). (B) Purified MCPyV viral-like particles (VLPs). Negative stain TEM of purified, DNA-filled VLPs composed of VP1 proteins (Bayer et al., 2020). (C) Schematic representation of a deltapolyomavirus cross-section. It comprises VP1 (green), VP2 (dark green), VP3 (light green), and histone (blue) encircling viral DNA (purple). Contrarily, MCPyV (alphapolyomavirus) does not comprise VP3 (from ViralZone, Swiss Institute of Bioinformatics).

3.3. The tumor antigens

MCPyV expresses four early gene products (Figure 4A), of which sT and LT were ascribed to the unique role of MCPyV to induce tumorigenesis in its host (Houben et al., 2012, 2010). On the contrary, the role of 57K and ALTO is still mainly unknown. Despite lacking sequence similarity, ALTO is related to the highly transforming middle tumor antigen (MT) of MPyV. It is expressed during replication, while a role in viral replication could not be demonstrated thus far (Carter et al., 2013). ALTO-encoding circular RNAs could be detected in MCC. However, their translation is negatively controlled by MCPyV-encoded miRNAs. It is suggested that ALTO functions as a transcriptional enhancer of host genes involved in MCPyV pathogenesis since a significant upregulation of proteins involved in the inhibition of NF-κB signaling is observed upon ALTO expression (Yang et al., 2021).

57K shares the amino acids 1-332 and 718-817 with LT, while most of the LT OBD and helicase are absent. Since 57K is highly conserved among MCPyV strains, it is thought to be involved in viral survival and propagation, whereas a role in viral replication could be excluded (Kwun et al., 2009). In the following, however, the focus is set on sT and LT, while 57K and ALTO will not be addressed further since their function is still largely elusive. Additionally, an overview of solved PyV T antigen structures is given in Table S 5.
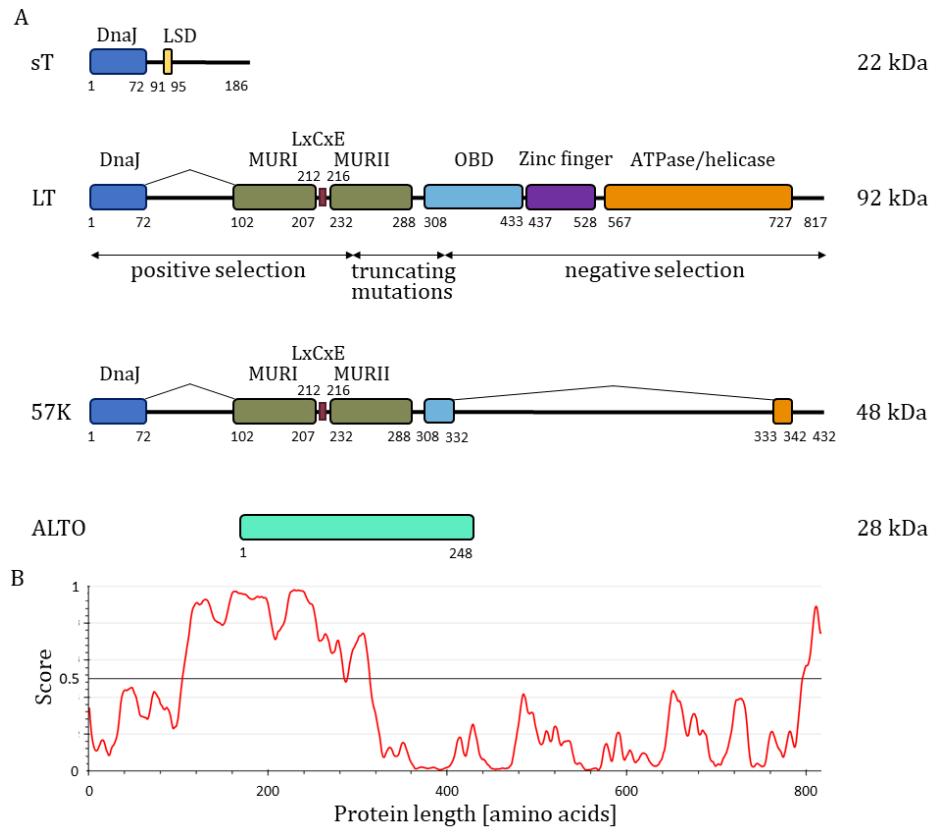
Figure 4: Domain organization of the MCPyV early region gene products. (A) The early region encodes sT, LT, and 57K, which are produced via alternative splicing and share the N-terminal DnaJ domain (blue). Next to the DnaJ, sT harbors a LT stabilizing domain (LSD, yellow). LT and 57K also share the Merkel unique region I and II (MURI and II, green), which flank the retinoblastoma protein (pRb) binding site LxCxE (red). The replication activity of LT is given by the origin-binding domain (OBD, light blue) and the C-terminal helicase, which comprises a zinc finger (purple) and an ATPase domain (orange). 57K is spliced a second time within the LT OBD, thereby losing replication activity, and shares the last 100 amino acids with LT. ALTO (turquoise) reveals a +1 frame shift relative to the second LT exon and utilizes an alternate start codon. The amino acids are shown below each construct, except those of the LxCxE motif. On the right side, the molecular weight of the proteins is depicted. (B) Prediction of the intrinsically disordered region of LT. Predictions above a threshold of 0.5 are considered intrinsically disordered. The prediction was performed with IUPred3 (Erdős et al., 2021).

3.3.1. Large tumor antigen

The multi-domain protein LT is, with a size of 92 kDa, the largest early gene product. Its first 72 amino acids harbor the DnaJ domain, which binds heat shock proteins 70 (Hsc70). This interaction is mediated via the HPDKGG motif of amino acids 42-47 and is required for efficient viral replication in cell culture (Kwun et al., 2009; Shuda et al., 2008). Additionally, abrogation of the DnaJ-Hsc70 interaction leads to the loss of the growth-promoting ability of LT by inhibiting retinoblastoma protein (pRb) binding (Houben et al., 2015). The concerted interaction between LT, Hsc70, and pRb, dissociating the DNA-bound Rb-E2F complex, which inhibits the transition into S-phase, was also described for SV40 LT (Sullivan et al., 2000; Stubdal et al., 1997). The highly conserved interaction between pRb and LT in PyVs is mediated by the LxCxE motif of amino acids 212-216 (Shuda et al., 2008). The co-crystal structure of the SV40 LT DnaJ domain (blue) and the central pocket domain of pRb (gray), which is the minimal region for E2F binding and essential for growth suppression, was published by Kim et al. (2001) (Figure 5A). The LxCxE motif (bright green) is located within the extended loop region and accounts for two-thirds of the binding interface with pRb. LT sequestration of pRb upregulates the E2F-mediated transcription of cyclin E and survivin, amongst others, leading to increased cell proliferation (Arora et al., 2012). Likewise, various other tumor virus proteins, such as the HPV E7, the KSHV latency-associated nuclear antigen (LANA), and the adenovirus protein E1A, also target the pRb/E2F pathway to induce tumorigenesis (Mesri et al., 2014; X. Liu & Marmorstein, 2007). The pRb-binding motif is embedded in the serin-rich Merkel unique region I and II (MURI and II). They are predicted to be unstructured and show no sequence similarity to any known PyV (Nwogu et al., 2020a; E. M. Johnson, 2010) (Figure 4B). Their function is not entirely understood, but it could be shown that approximately 70% of human cancer-associated proteins harbor *in silico* predicted intrinsically disordered domains (IDRs), suggesting an essential role in tumorigenesis (Iakoucheva et al., 2002). Further, multiple Skp, Cullin, F-box containing complex (SCF) E3 ligases-binding sites could be identified. They contribute to the proposed autoregulatory function of the MURs, which is suggested to be part of the viral latency mechanism (Kwun et al., 2017, 2013). However, controversial results of LT stability upon deleting the MURs exist. Nwogu et al. report increased wild-type LT stability and cell proliferation in hTERT-immortalized BJ human foreskin fibroblasts (BJ-hTERT), evoked by the removal of SCF E3 ligases binding sites (2020), while reduced truncated LT (tLT) levels and unaffected growth promotion are observed by Houben et al. in MCC cell lines (2015). Further, MURI binds the cytosolic vacuolar sorting protein Vam6p and translocates it to the nucleus with the aid of the LT nucleus localization signal (NLS) of amino acids 277-280 (X. Liu et al., 2011; Nakamura et al., 2010). The LT-dependent nuclear sequestration of Vam6p disrupts lysosomal clustering, repressing viral replication and virion production (H. Feng et al., 2011). A second binding partner of

the LT N-terminus (amino acids 1-404) is the bromodomain protein 4 (Brd4), which positively influences viral replication by recruiting replication factor C (RFC) to the replication sites (X. Wang et al., 2012). Downstream of MURII, two phosphorylation sites, T297 and T299, also contribute to viral replication by modifying the LT-ori affinity, thereby showing antagonistic effects (Diaz et al., 2014). Interestingly, phosphorylation also regulates SV40 replication by enabling cooperative double-hexamer interactions, which is not observed for MCPyV LT thus far (Barbaro et al., 2000; Weisshart et al., 1999). Along this line, it could be shown that MCPyV LT-ori interaction was enhanced by binding the cellular ubiquitin-specific processing protease 7 (Usp7) via multiple P/A/E-x-x-S motifs, present predominantly in the MURs (Czech-Sioli et al., 2019). A similar mechanism was described for the EBV DNA-binding protein, Epstein-Barr nuclear antigen 1 (EBNA1), whose binding to the family of repeats (FR) element of the ori is stimulated by Usp7 (Sarkari et al., 2009; Holowaty et al., 2003).

Interaction between the viral ori and LT is mediated by the OBD of amino acids 308-433, whose structure was published in 2011 (Figure 5B). It is the only known structure of a MCPyV LT domain and reveals an asymmetrical binding to the ori, similar to MPyV (C. Harrison et al., 2011, 2013) but contrary to SV40 (Bochkareva et al., 2006). It primes viral replication initiation by recognizing the GRGGC DNA motif and coordinating the helicase onto the ori. For SV40, it could be shown that the replication-essential loading of the replication protein A (RPA) onto single-stranded DNA (ssDNA) is mediated by the interaction with the OBD (X. Jiang et al., 2006). This interaction was structurally investigated by NMR chemical shift perturbation and is also conceivable for MCPyV OBD (Arunkumar et al., 2005).

The LT C-terminus harbors a zinc finger and an ATPase domain, which constitute the helicase, elaborated in detail in the superfamily 3 helicases paragraph. Since extensive studies of the MCPyV helicase are lacking, data from SV40 helicase will be quoted in the following. The most comprehensive structure of any PyV LT comprises the OBD-helicase of SV40 bound to the viral ori (Figure 5C). The helicase-OBD (blue and red) aligns as a dimer onto the DNA (black), which may reflect the initial state of assembly of an active hexameric helicase. Surprisingly, only one OBD interacts with a GRGGC motif, while the other interacts with a CpG dinucleotide within the minor groove of the ori (Y. P. Chang et al., 2013). Similar to the OBD, the LT helicase recruits important cellular replication factors onto the DNA to initiate replication. For example, the co-crystal structure of the helicase bound to the regulatory subunit of the DNA polymerase α-primase could be elucidated by Zhou et al. in (2012) (Figure 5D). The interaction between the helicase and the topoisomerase I was also reported (Khopde & Simmons, 2008). Despite many similarities, p53 binding differs between the MCPyV and SV40 helicases. The latter binds directly to the tumor suppressor protein p53 and inhibits p53-mediated transcription, resulting in the evasion of senescence and apoptosis (Pipas and Levine 2001; Symonds

et al. 1994; Jian et al. 1993, Bergonetti et al 1992). Its co-crystal structure could be elucidated by Lilyestrom and colleagues (2006). Similar interactions are reported for the oncoproteins HPV E6 and EBV BZLF1, contrary to MCPyV and MPyV, which lack direct p53 interaction (Martinez-Zapien et al., 2016; Cheng et al., 2013; Sato et al., 2009; Dilworth, 1990). However, a significant reduction of p53-dependent transcription by full-length MCPyV LT was reported (Borchert et al., 2014). Contrarily, p53 activation was observed upon LT-pRb binding, which is antagonized by the co-expression of sT and will be described in more detail below (Park et al., 2019; Cheng et al., 2013).

Next to its helicase activity, MCPyV LT C-terminus evokes cellular growth inhibition. It induces a DNA damage response (DDR), which is crucial for viral replication (Tsang et al., 2014), but also leads to p53 activation and cell cycle arrest (Cheng et al., 2013; J. Li et al., 2013). This is contrary to SV40 LT, which induces DDR but simultaneously inhibits p53 to support viral replication (J. Li et al., 2013; Hein et al., 2009). It could be demonstrated that the DDR-mediating Ataxia telangiectasia mutated (ATM) kinase phosphorylates the recently identified chromatin remodeling T antigen binding partner, KRAB-associated protein 1 (Kap1) (Siebels et al., 2020; Czech-Sioli et al., 2019). This leads to cellular senescence, reflecting a host defense mechanism against viral replication (Siebels et al., 2020). Kap1-mediated viral replication inhibition is also described for other DNA tumor viruses such as KSHV and EBV (X. Li et al., 2017; Bentz et al., 2015; Sun et al., 2014; Zhang et al., 2014). A critical role within the growth-inhibiting features of the LT C-terminus can be ascribed to the last 100 amino acids of LT. They are shared by the third exon of 57K and show cell growth retardation in a MCC cell line and in SV40 immortalized human fibroblasts (Cheng et al., 2013). This observation is supported by the fact that the ATM also phosphorylates LT S816, promoting growth inhibition and apoptosis (J. Li et al., 2015). To summarize, full-length LT reveals growth-promoting properties in its N-terminus and growth-inhibiting properties in its C-terminus. This seems beneficial to balance viral propagation and uncontrolled cell growth, which would lead to the abrogation of viral production.

However, during the establishment of MCC, the growth-inhibiting features necessitate the negative selection of the LT C-terminus. This negative selection is located within the region of amino acids 243-329 (Borchert et al., 2014). The loss of the helicase and OBD results in replication deficiency which is required for tumorigenesis (Shuda et al., 2008). Intriguingly, also SV40 and MPyV show enhanced tumorigenesis by implementing replication deficiency (Manos & Gluzman, 1984; Prives et al., 1983; Small et al., 1982; Lania et al., 1981; Israel et al., 1980). It was proposed that the helicase activity could lead to unlicensed replication of the integrated viral genome, resulting in the aforementioned DDR activation and cell death. However, observed DDR reduction by a mutated helicase could have originated from reduced protein expression and needs to be further investigated (J. Li et al., 2013). In contrast, the N-terminus of LT, including the pRb-binding motif, is always preserved, while the

presence of the NLS varies between tumors (Houben et al., 2015; Borchert et al., 2014; Fischer et al., 2010; Shuda et al., 2008). To test the tumorigenic potential of tLT, it was co-expressed with sT, resulting in a higher transformation capacity of primary rat epithelial cells in contrast to sT co-expressed with full-length LT (Borchert et al., 2014). The importance of tLT in tumorigenesis was further demonstrated by showing that the survival of a MCC cell line is dependent on the expression of both, sT and tLT (Houben et al., 2010). Further, an increased binding affinity of tLT to pRb compared to full-length LT was reported, highlighting the role of tLT in tumorigenesis (Borchert et al., 2014).
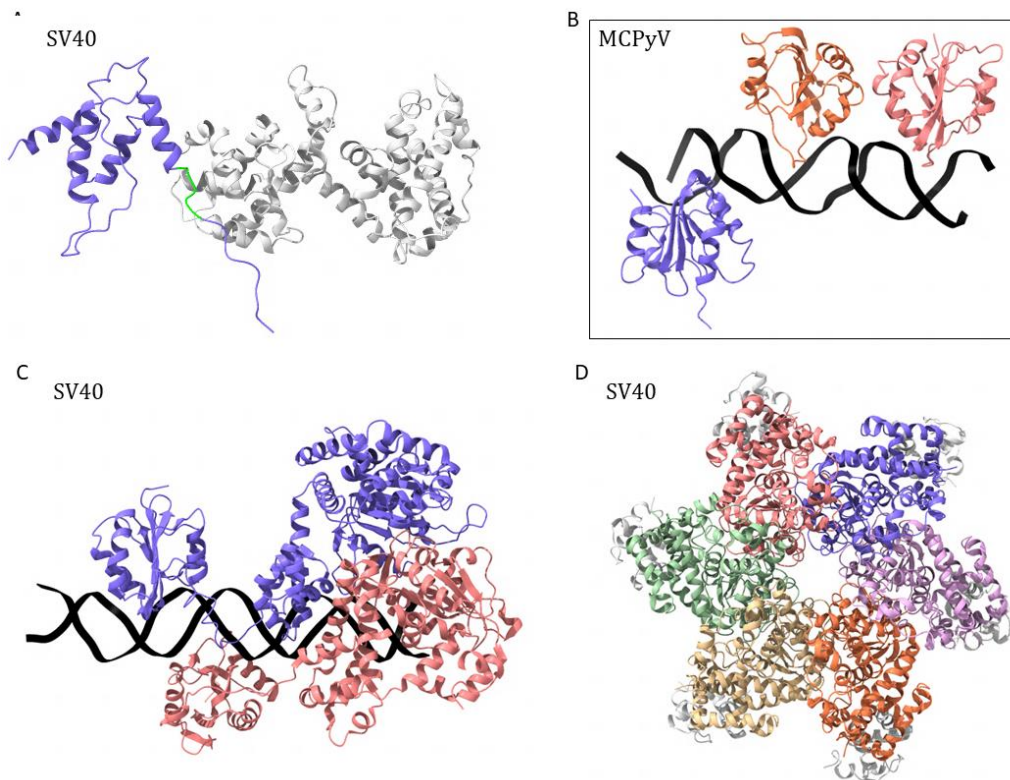
Figure 5: Crystal structures of PyV LT domains. (A) SV40 LT DnaJ domain (blue) with LxCxE motif (bright green) bound to the pRb AB Box (gray). PDB: 1GH6. (B) MCPyV LT OBD (blue, red, and orange) bound to the viral ori (black). PDB: 3QFQ. (C) Dimer of the SV40 LT OBD-helicase (blue and red) bound to the viral ori (black). PDB: 4GDF. (D) Hexamer of the SV40 LT helicase (blue, red, green, beige, orange, and purple) bound to six regulatory subunits of the DNA polymerase α-primase (gray). PDB: 4E2I.

The superfamily 3 helicases

The PyV LT helicases belong to the superfamily 3 (SF3), which are members of the ATPase associated with a variety of cellular activities (AAA+) superfamily. Helicases use ATP hydrolysis to translocate along DNA and RNA while unwinding. A plethora of distinct helicases exists, which are sorted into six superfamilies and are involved in replication, nucleic acid repair, recombination, and transcription (Tuteja & Tuteja, 2004; Patel & Picha, 2000). SF3 helicases were initially identified in small DNA and RNA viruses, such as SV40, papillomaviruses, and adeno-associated viruses (AAV) (Iyer et al., 2004;

Neuwald et al., 1999; Gorbalenya et al., 1990). These helicases are composed of an N-terminal zinc finger, which is flexibly linked to a C-terminal ATPase. They hexamerize to form a complex of two stacked rings with a central pore, expanding towards the N-terminus (Sanders et al., 2007; D. Li et al., 2003). For maximum activity, the helicase requires the concerted action of a DNA-binding domain, zinc finger, and ATPase.

Two working models exist on how hexameric helicases, comprising SF3-6, assemble onto nucleic acid. In one, monomers are successively loaded onto the DNA until a closed hexamer is formed, which is proposed for the papillomavirus E1 (Enemark et al., 2002). In the other, an open multimeric ring encircles DNA with the help of a ring breaker, as shown for the SF6-associated mini chromosome maintenance complex (MCM2-7) (Frigola et al., 2017; Samel et al., 2014; Davey & O'Donnell, 2003). An essential role in helicase oligomerization holds the zinc finger since it could be shown that mutation of the zinc-binding residues disables hexamerization (D. Li et al., 2003; Loeber et al., 1991). Mutating residues within the ATPase domain also reduces oligomerization, while the primary function of this domain is ATP-binding and hydrolysis, providing energy for the helicase activity (Wendler et al., 2012; Greenleaf et al., 2008).

After hexamer loading, a cooperative double-hexamer is formed in a head-to-head orientation for ori melting (Cuesta et al., 2010; Smelkova & Borowiec, 1998). The initiation of DNA unwinding follows the dissociation of the double-hexamer into single hexamers, which bidirectionally unwind the genome, as shown for SV40 (Yardimci et al., 2012). Thereby, double-hexamer disruption is mediated by hyper-phosphorylation of LT (Sowd & Fanning, 2012; Weisshart et al., 1999). Fork progression is explained by the steric exclusion model in which one hexamer migrates along one ssDNA while the other is sterically excluded from the central pore, as proposed for E1 and Mcm2-7 (Sowd & Fanning, 2012; Fu et al., 2011; Enemark & Joshua-Tor, 2006; Kaplan et al., 2003). Recently, this model was refined via the first high-resolution structure of Mcm2-7 on a replication fork (Yuan et al., 2020). The refined mechanism is called dam-and-division and includes dsDNA helicase entering. The DNA is divided at the N-terminus of the zinc finger, followed by releasing the lagging ssDNA in front of the zinc finger.

The ATPase core consists of a five-stranded beta sheet, flanked by multiple alpha helices (Figure 6A). It shares five conserved motifs, called walker A (A, yellow), walker B (B, yellow), B´ (purple), C (light yellow), and an arginine finger (R, blue) (Koonin, 1993). Motif C is unique to SF3 and is called sensor 1, while A, B, and R are conserved within all SFs. Motif B´ and adjacent residues of the β-hairpin loop are involved in ssDNA binding and couple ATP hydrolysis to nucleic acid unwinding, as shown for AAV2 Rep40 (Yoon-Robarts et al., 2004). ATP-binding and hydrolysis are realized in the interphase of two adjacent helicase monomers. Thereby, motif B and A, also called phosphate binding-loop (P-

loop), coordinate the water-activating magnesium ion and the β and γ phosphate of ATP, respectively (Saraste et al., 1990; Walker et al., 1982). Sensor 1, located in the β4 strand and separating motif A and B, assists ATP hydrolysis by coordinating the nucleophilic attacking water via its conserved polar asparagine, interacting with the γ-phosphate. Mutation of the SV40 LT motif B or the sensor 1 abolishes ATPase activity, similar to mutation of R (Greenleaf et al., 2008). R links ATP hydrolysis to locomotion and acts in *trans*. It senses the γ-phosphorylation state of the *cis*-bound nucleotide and unhitches upon ATP hydrolysis, causing the switch from a closed (nucleotide bound) to an open (nucleotide unbound) helicase. For SV40, next to R, a non-canonical positive lysin charge contributes to the formation of the ATP-binding pocket (Gai, Zhao, et al., 2004). Its mutation dampens ATP hydrolysis and oligomerization, similar to R (Greenleaf et al., 2008).

In the scenario of sequential ATP hydrolysis, whereby three different binding states, ATP, ADP, or empty, exist simultaneously within the same hexamer, bipartite ATP-binding leads to transient dimer formation and hexameric asymmetry, which is observed in many ATPases (Guo et al., 2019; Stinson et al., 2015; Enemark & Joshua-Tor, 2006; Singleton et al., 2000). Interestingly, crystal structures of the SV40 helicase reveal a six-fold symmetry, suggesting a concerted ATP hydrolysis rather than a sequential one, which is supported by the observed all-or-none ATP occupancy (Gai et al., 2010; Gai, Zhao, et al., 2004; D. Li et al., 2003; Huang et al., 1998). Along this line, DNA-helicase interaction, generally mediated by the β-hairpin loops intruding the central pore, differs between SF3 members. For E1, these loops are oriented in a spiral staircase, and their height corresponds to the catalytic state of the ATP-binding pocket. During one ATPase cycle, each loop continuously binds the backbone of one base, thereby pulling the ssDNA through the channel (Sanders et al., 2007; Enemark & Joshua-Tor, 2006). Consequently, one hexamer cycle translocates six nucleotides through the channel by hydrolyzing six ATPs and releasing six ADPs, which is called coordinated escort mechanism. The direction of sequential ATP hydrolysis around the hexameric ring is thereby suggested to determine the unwinding polarity, which is 3´-5´ for SF3 and SF6, by a clockwise ATPase site firing (Thomsen & Berger, 2009). Contrarily, the β-hairpin loops of the SV40 helicase show a planar arrangement, suggesting mechanistic differences in the ATPase locomotion within the SF3 family (Lilyestrom et al., 2006; Gai, Li, et al., 2004; D. Li et al., 2003).

Another difference between SF3 members originates within the unwinding initiation. The loading of E1 onto nucleic acid relies on the helper protein E2, and ori melting is mediated by a transient E1 double trimer, inserting its β hairpins into the DNA to pry (Schuck & Stenlund, 2011; X. Liu et al., 2007). Contrarily, LT recognizes and binds the ori independent of helper proteins, while its melting mechanism is still largely elusive (Valle et al., 2006). It could be shown via transmission electron cryo-microscopy single particle analysis (CryoEM SPA) that LT binds dsDNA, suggesting a squeeze-to-open

mechanism for ori melting (Cuesta et al., 2010; Gai et al., 2010). The helicase central pore varies between 7-15 Å, depending on the occupancy of the ATPase active site. Its diameter is therefore too small to fit dsDNA without its deformation (Gai et al., 2010; Gai, Zhao, et al., 2004). Additionally, conformational rearrangements within the double-hexameric DnaJ domain and OBD were suggested to be crucial for replication initiation (Cuesta et al., 2010).
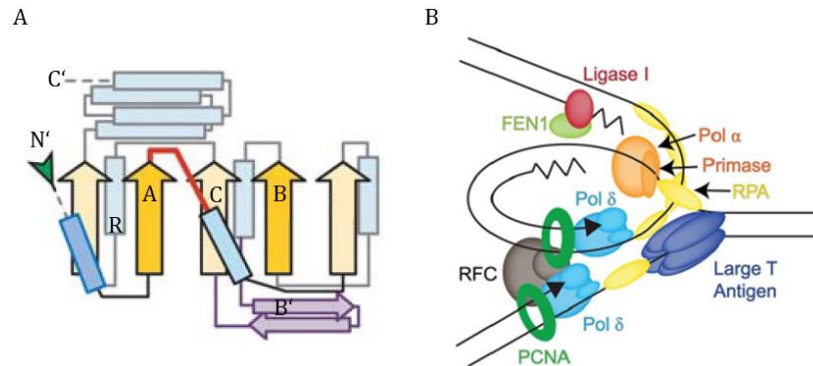

Figure 6: Superfamily 3 helicases. (A) Topology diagram of the SV40 LT ATPase. β-strands (yellow and purple) are shown as arrows and α-helices (blue) as rectangles. α-helices located below the β-sheet plane are colored in faint blue and those above the plane in bright blue. Conserved motifs, comprising the Walker A (A), Walker B (B), Sensor 1 (C), B', and arginine finger (R), are marked within the respective secondary structure. The β-hairpin loop, involved in ssDNA interaction, is shown in purple. The P loop, which is part of A and involved in ATP-binding and hydrolysis, is marked in red. Broken lines indicate non-resolved or missing secondary structures of the PDB file (1N25). The N- and C-terminus are marked with N´ and C´. Modified from (Iyer et al., 2004). (B) Replisome of SV40. LT (blue), polymerase δ (Pol δ, light blue), replication protein A (RPA, yellow), DNA polymerase α/primase (Pol α, orange), replication factor C (RFC, gray), proliferating cell nuclear antigen (PCNA, green), ligase 1 (red), and flap endonuclease-1 (FEN1, light green) are involved in viral DNA (black) replication. Modified from (Sowd & Fanning, 2012).

SV40 is one of the most extensively studied PyVs since it serves as a model system to elucidate the molecular mechanisms of eukaryotic replication. The SV40 LT helicase remarkably resembles the eukaryotic replication initiating helicase MCM2-7 in dimension, architecture, and loading onto the ori (N. Li et al., 2015; Cuesta et al., 2010; Evrin et al., 2009; Remus et al., 2009; Gai, Zhao, et al., 2004). Both belong to the AAA+ protein family, while MCM2-7 harbors additional DNA interaction loops (Yuan et al., 2020; O'Donnell & Li, 2018). The atomic structure of the MCM2-7 double-hexamer was published in 2015, while for the SV40 LT double-hexamer, only low-resolution information exists (N. Li et al., 2015; Cuesta et al., 2010).

Despite substantial similarities, the helicases differ in their composition. SF3 helicases build a mono-hexamer contrary to MCM2-7, consisting of six unique subunits. The sequence identity between SV40 or MCPyV LT and MCM 2-7 is low, ranging between 6-9% (Figure S 1). Nevertheless, the protein machinery of viral replication resembles that of host chromosomal replication (Figure 6B). Both utilize RPA, DNA polymerase α/primase, replication factor C (RFC), proliferating cell nuclear antigen (PCNA), and DNA polymerase γ (Sowd & Fanning, 2012; Stillman, 2008). The DNA polymerase α/primase interacts with LT/MCM 2-7 and is loaded onto the ori, synthesizing RNA primers (You et

al., 2013; Zhou et al., 2012). RFC displaces the DNA polymerase α/primase and places the ring-shaped PCNA onto the DNA, which recruits the DNA polymerase γ to synthesize progenitor DNA.

On the other hand, the SV40 replisome excludes fundamental host components such as the DNA polymerase ε, which is involved in lagging strand synthesis in eukaryotes and is substituted by the DNA polymerase γ in viral replication (Zlotkin et al., 1996). Additionally, the process of helicase loading onto the DNA substantially differs between viral and eukaryotic replication. Interestingly, Fanning and Sowd suggest that SV40 replication utilizes the host genome maintenance machinery rather than the replication machinery (Sowd & Fanning, 2012). Evidence is provided by the fact that SV40 and MCPyV activate the DDR for viral replication, leading to the presence of many DDR active proteins, such as ATR, and the phosphorylated γH2AX histone and checkpoint kinase 2 (pChk2), within viral replication centers (Tsang et al., 2014; X. Zhao et al., 2008). Similar results are reported for E1 (Spriggs & Laimins, 2017).

In summary, SF3 helicases are important viral enzymes involved in replication. Although there have been many milestones in understanding the molecular mechanism of SF3 helicases in the past decades, fundamental insights into helicase loading, ori melting, helicase progression, and unwinding termination are still pending, highlighting the urgency of further structural examination.

3.3.2. Small tumor antigen

Beside the above discussed LT, sT is the second essential player in viral transformation and replication. The N-terminal DnaJ domain of the 22 kDa sized sT is shared with LT, as depicted in Figure 4A. Interestingly, for sT, the function of this domain is unknown, and the mutation of the Hsc70-binding motif did not interfere with replication or transformation (Shuda et al., 2011; Kwun et al., 2009). The function of sT in cell transformation is not fully understood, but it could be shown that sT alone is sufficient to transform rodent fibroblasts and that its knockdown abolishes MCC cell growth (Kwun et al., 2013; Shuda et al., 2011). This is contrary to SV40 and MPyV sT, which only enhance the transformative activity of LT and MT, respectively (Bikel et al., 1987; Noda et al., 1987). Therefore, MCPyV sT possesses a unique role among PyVs as the primary driver of tumorigenesis. The sT-induced preservation of hyperphosphorylated eukaryotic translation initiation factor 4E-binding protein 1 (4E-BP1) is thereby of great importance. 4E-BP1 is an mTOR kinase target and regulates the cap-dependent translation. However, the exact mechanism of the 4E-BP1 regulation is still elusive. Further, sT-targeted 4E-BP1 is suggested to increase viral replication (Shuda et al., 2011). ST harbors a centrally located LT stabilizing domain (LSD) of amino acids 91-95. Over the last decade, several functions could be assigned to this positively charged domain. For example, LSD stabilizes LT by interacting with the tumor suppressor SCF E3 ligase F-box and WD repeat domain-containing 7 (Fbw7), which targets LT and causes its proteasomal degradation. The sT-mediated inhibition of

Fbw7 leads to increased LT levels that enhance viral replication, assigning sT an indirect role in viral replication (Nwogu et al., 2020b; Kwun et al., 2013; H. Feng et al., 2011; Kwun et al., 2009). LT-mediated viral replication is also enhanced by conserved iron-sulfur clusters residing in the sT C-terminus (Tsang et al., 2016). Next to LT stabilization, LSD-mediated Fbw7 sequestration is involved in cell transformation. It could be shown that upon LSD mutation, sT loses its ability to transform rodent cells *in vitro* and in a mouse model (Verhaegen et al., 2017; Kwun et al., 2013). Additionally, its inhibition affects Fbw7-targeted cell cycle regulators such as cyclin E and c-Myc, contributing to transformation (Kwun et al., 2013).

The tumor suppressor protein phosphatase 2A (PP2A) is a second sT binding partner, regulating pathways such as c-Myc and mitogen-activated protein kinase (MAPK). This interaction was identified by Shuda et al. via immunoprecipitation (2011) and is also known for other PyVs, such as BKPyV, JCPyV, MPyV, and SV40 (Guergnon et al., 2011; Bollag et al., 2010; Rundell et al., 1981). The co-crystal structure of SV40 sT, bound to the PP2A A subunit, was elucidated by Cho et al. in (2007) (Figure 7). It is the only PyV sT structure available thus far, highlighting the necessity for MCPyV sT structure elucidation. It could be shown that MCPyV sT binds specifically to the PP2A subunit Aα, Aβ, and C, thereby hindering the holoenzyme formation. Interestingly, PP2A interaction is independent of the transforming capacity of sT *in vitro* and *in vivo*, contrary to other PyVs (Verhaegen et al., 2017; Kwun et al., 2015; Griffiths et al., 2013; Shuda et al., 2011; Arroyo & Hahn, 2005).

ST interferes with the nuclear factor kappa-light-chain-enhancer of activated B cells (NF-κB)-mediated transcription for immune evasion, which was, among others, linked to the establishment of MCPyV persistence. NF-κB are transcription factors regulating cell functions such as immunity, proliferation, and inflammation. They are manipulated by many DNA tumor viruses, such as SV40, HPV, and KSHV (Field et al., 2003; Spitkovsky et al., 2002; Sontag et al., 1997). Interestingly, opposing results about the NF-κB modulation of MCPyV exist. Berrios et al. report sT-mediated NF-κB activation in human diploid lung fibroblasts (2016), while its inhibition was shown by Griffiths et al. upon inflammatory stimuli (2013). Thereby, sT inhibits the IκB kinase (IKK)α/IKKβ-mediated phosphorylation of IκB, preventing the activation of cytoplasmic NF-κB and its translocation to the nucleus. Phosphorylation inhibition is thereby regulated by the interaction of sT with the protein phosphatase 4 catalytic subunit (PP4C) and PP2A Aβ. Additionally to PP4C, the regulatory subunit of PP4 (PP4R1) mediates an interaction between sT and the non-catalytic regulatory NF-κB essential modulator (NEMO), which is also part of the IKK. The interaction between sT, NEMO, and the phosphatase subunits is provided by the sT amino acids 101-103. Intriguingly, the PP4R1-mediated inactivation of NEMO appears to be unique to MCPyV (Abdul-Sada et al., 2017; Griffiths et al., 2013). In line with this, transcription of the senescence-associated secretory phenotype (SASP) genes via

non-canonical NF-κB signaling, promoting cell proliferation, is accomplished by the sT LSD and seems to be specific for MCPyV sT (J. Zhao et al., 2020).

Next to the modulation of the NF-κB pathway, sT-PP4C mediates the destabilization of the microtubule, enhancing cell motility and mobility. This is suggested to contribute to the pronounced metastatic potential of MCC (Schrama et al., 2012; Lemos et al., 2010). The effect is induced by the PP4C-arbitrated dephosphorylation of the microtubule-associated protein stathmin, which is overexpressed in many human cancers (Knight et al., 2015; Rana et al., 2008). Additionally, the PP4C-mediated formation of migratory filopodia is reported to drive cell motility (Stakaitytė et al., 2018). The involvement of sT in MCC metastasis was further demonstrated by Knips et al. (2017). They showed that implantation of MCC cells into severe combined immunodeficiency (SCID) mice results in metastases with increased sT expression. Furthermore, genes involved in angiogenesis and extracellular matrix disassembly were upregulated in tumors with high sT expression. Next, the sT-mediated upregulation of the cellular sheddases A disintegrin and metalloproteinase domain-containing protein 10 (ADAM10) and ADAM 17 causes cell adhesion disruption, also contributing to enhanced cell motility (Nwogu et al., 2018). Many of these observations have been suggested to result from the sT-mediated p38 MAPK pathway activation, revealing a major pathway for MCC cell motility and migration, thereby contributing to metastasis (Dobson et al., 2020).

Interestingly, sT, similar to LT, induces a DDR by ATM activation and subsequent phosphorylation of the histone H2AX (J. H. Wu et al., 2019). As mentioned above, sT antagonizes p53 activation by LT-pRb binding. Thereby, sT interacts with the MycL-EP400 complex (Cheng et al., 2017) and targets the p53-degrading E3 ligase mouse double minute (MDM) 2 and 4 (Park et al., 2019). Strikingly, p53 is mutated in virus-negative MCC (vnMCC), while it is expressed as wild-type protein in virus-positive MCC (vpMCC) (Knepper et al., 2019). The sT-mediated p53 inhibition, therefore, may represent an alternative mechanism to disable p53, which is not fulfilled by MCPyV LT (Park et al., 2019; Cheng et al., 2013).

Next to manipulating cellular tumor suppressor genes, immune evasion is an important tool of oncoviruses to establish persistence and carcinogenesis. For example, upregulation of the immune checkpoint protein programmed death-ligand 1 (PD-L1) to inhibit T cell activation is observed for EBV, KSHV, and HPV and is also successfully targeted in MCC therapy (Hong et al., 2019; Gilardini Montani et al., 2018; Host et al., 2017) (3.4.2). Furthermore, sT enhances the expression of the cell surface receptor CD47 to evade phagocytosis (Schlemeyer et al., 2022). Alongside, it is reported that sT subverts the type I interferon (IFN 1) response by suppressing interferon-stimulated genes (ISGs) (Ohnezeit et al., in revision). However, although significant progress has been made in understanding

the role of sT in immune evasion, viral replication, and tumorigenesis, the atomic structure of MCPyV sT could provide structure-function information, substantially advancing the research on sT.
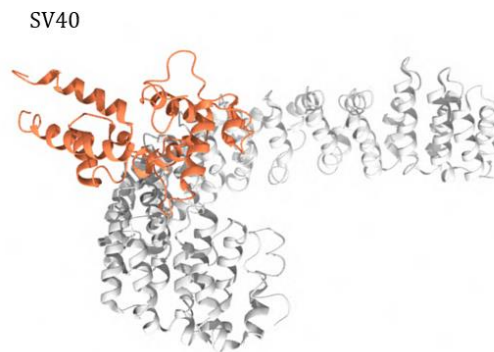
SV40

Figure 7: Crystal structure of SV40 sT. The T antigen (orange) is bound to the PP2A A subunit (gray). PDB: 2PF4.

3.4. Merkel cell carcinoma

MCPyV is the causative agent of 80% of all MCCs, which are rare but aggressive neuroendocrine skin tumors that occur mostly in elderly, fair-skinned, and immunosuppressed patients (Agelli & Clegg, 2003). The remaining 20% arise from UV-irradiation (H. Feng et al., 2008). Interestingly, vnMCC is phenotypically indistinguishable from vpMCC, while their genomes differ substantially. VnMCCs display typical UV-derived mutations and show an overall high mutational burden of 10.09 ± 2.32 mutations per mega base, including those in the well-known tumor suppressors p53 and pRb (Goh et al., 2015; Harms et al., 2015; Erstad & Cusack, 2014). Contrarily, vpMCCs show only 0.40 ± 0.09 mutations per mega base, and the inhibition of tumor suppressors is realized by the viral tumor antigens sT and LT, as described in 3.3 (Harms et al., 2015).

MCC was first described in 1972 and is named after its initially presumed cell of origin, the Merkel cell (MC). These cells function as mechanoreceptors and are located in the basal layer of the epithelium close to nerve ends (Halata et al., 2003). The erroneous assumption arose from similar features between MCs and MCC, such as the MCC marker cytokeratin-20 (CK20) (Toker, 1972) (Figure 8B). At present, however, high concordance exists that MCs are not the cell of origin. Their post-mitotic state fails to support viral genome integration and cellular transformation, and it could be shown that sT cannot stimulate proliferation and tumor formation in mature MCs (Shuda et al., 2015). Moreover, their epidermal location does not coincide with MCC´s dermal or subcuticular occurrence. Current debates about the MCC cell of origin include pro-/pre-B-cells (zur Hausen et al., 2013), dermal fibroblasts (Sunshine et al., 2018), and MC progenitors of the hair follicle (Verhaegen et al., 2022; Kervarrec et al., 2020). Interestingly, distinct origin cells of vpMCC and vnMCC, namely dermal fibroblasts and epidermal keratinocytes, respectively, are also considered (Sunshine et al., 2018), while others suggest a common epithelial origin (Gravemeyer et al., 2021).

MCC is predominantly found on UV-exposed body regions such as the head, neck, and upper limbs. However, instances are also reported at unusual locations, such as the lymph nodes and the nasal vestibule (Fazio et al., 2022; Khatchaturian & Zohrabian, 2018). It arises as fast-growing, firm, and painless nodules with a red-to-blue shade (Figure 8A). The non-specific appearance often leads to misinterpretation of MCC as benign inflammatory lesions, resulting in delayed clinical diagnosis. Therefore, the acronym AEIOU was introduced in 2008 to simplify MCC recognition. It describes the clinical characteristics as **A**symptomatic, **E**xpanding rapidly, **I**mmunosuppressed, **O**lder than 50 years of age, and **U**V-exposed (Heath et al., 2008). Next to morphological features, immunohistochemical staining against neuroendocrine and cytoskeletal markers, such as synaptophysin and CK20, respectively, is commonly used for MCC diagnosis. Additionally, the presence of viral DNA and T-antigens discriminate between vpMCC and vnMCC, even though LT detection fails in some vpMCCs (Bhatia et al., 2010; Busam et al., 2009; Shuda et al., 2009).

MCC is rare, with an incidence between 0.24-2.5 per 100.000 person-year (Scampa et al., 2022). The highest prevalence occurs in Australian men, with an incidence of 3.9 per 100.000 person-year (Garbutcheon-Singh et al., 2020). Most of these MCCs were scored virus-negative, linking them to high UV-irradiation (Youlden et al., 2014; Garneski et al., 2009). In contrast, MCC prevalence in North America in 2013 was 0.7 per 100.000 person-year, comprising 69% vpMCCs (Paulson et al., 2018; Garneski et al., 2009). In addition to geographical location, gender, and skin pigmentation, age is a significant risk factor, as reflected by a MCC incidence of 0.1 per 100.000 person-year among the age group of 40-44 years, increasing to 9.8 per 100.000 person-year of patients above the age of 85 (Paulson et al., 2018). Despite its rare occurrence, its overall mortality rate of 30% is the highest among all skin cancer types (Samimi, 2019; J. C. Becker et al., 2018; Eisemann et al., 2016). One reason for the high mortality rate is its high metastatic potential, linked to a low five-year survival rate. This rate decreases from 62% with only local tumors to 25% with metastases (Agelli et al., 2010; Albores-Saavedra et al., 2010). Intriguingly, MCC cases have tripled in the last two decades and are prognosed to continuously rise, owing to improved diagnostics, demographic aging, and increased use of immunosuppressants (Jacobs et al., 2021). This underscores the importance of antiviral drug target development, improving MCC therapy, which will be discussed in 3.4.2.
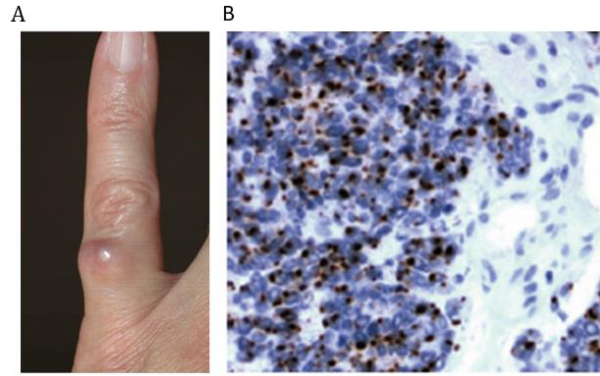
Figure 8: Merkel cell carcinoma. (A) Photograph of a MCC at a human pinky (Harms et al., 2018). (B) Histological sample of anti-CK20 immunostained MCC (20x magnification) (Y. Chang & Moore, 2012).

3.4.1. Viral tumorigenesis

VpMCC is induced and maintained by the T antigens of MCPyV. The viral genome is monoclonal integrated into the host genome, suggesting that tumorigeneses precedes the integration event (Fischer et al., 2010; H. Feng et al., 2008). Thereby, integration varies within patients but occurs predominantly within euchromatin with distinct integration patterns. Integration via double-strand breaks and non-homologous end joining (NHEJ) results in linear patterns, while microhomology-mediated end joining (MMEJ), also reported for HPV (Leeman et al., 2019), causes Z-pattern integration with amplified cellular DNA (Czech-Sioli et al., 2020; Starrett et al., 2020). A second prerequisite for tumorigenesis is the insertion of premature LT stop mutations, resulting in tLT, which lacks its growth-inhibiting C-terminus (3.3.1) and causes viral replication deficiency (Shuda et al., 2008). Interestingly, LT stop mutations vary between patients, while the pRb binding LxCxE motif is always preserved (3.3.1). It is suggested that stalling of a bidirectional replication fork during theta amplification can induce these aberrant viral genomes (Czech-Sioli et al., 2020). Intriguingly, mechanistic similarities of tumorigenesis exist between MCPyV and HPV, which also shows genome integration and gene deletions within the replication-responsible proteins E1 and E2 (Cricca et al., 2009; Arias-Pulido et al., 2006; Hopman et al., 2004; Klaes et al., 1999). The third prerequisite is a constitutive expression of tLT and sT, while VP1 is not or faulty expressed due to deletions or mutations (Houben et al., 2010; Pastrana et al., 2009; Kassem et al., 2008). ST and tLT deregulate the cell cycle and stimulate cell proliferation by interacting with tumor suppressors such as pRb (Shuda et al., 2014; Houben et al., 2012). However, the role of tLT in tumorigenesis has yet to be fully understood since it could be shown that sT alone is sufficient to transform rodent fibroblasts (Kwun et al., 2013; Shuda et al., 2011). Nonetheless, tumor maintenance and progression *in vitro* and *in vivo* are reported to depend on tLT (Houben et al., 2012, 2010). Further information on the tumorigenic potential of the T antigens is delineated in more detail in 3.3.

Another critical requirement for MCC development is immunosuppression, which could be linked to acute viremia preceding MCC. It is shown that MCC patients reveal significantly more antibodies against MCPyV compared to the control group, which is suggested to be beneficial for genome integration (Carter et al., 2009). Since viral integration results in the loss of progenitor production, it entails a viral "dead-end". This allows the hypothesis that MCC is a rare complication during MCPyV infection rather than a survival strategy. The lack of *in vitro* and *in vivo* transformation models impedes the elucidation of the exact tumorigenic mechanism. Remedy could provide the recently published MCC mouse model, in which sT, tLT, and atonal bHLH transcription factor 1 (Atoh1) are expressed in keratin 5 (Krt5) expressing epidermal cells, forming tumors with a convincing MCC phenotype within hair follicles for the first time. However, genomic p53 depletion is required, which is not observed in human MCC (Verhaegen et al., 2022). Krump and You suggest that inadequate MCPyV restriction may contribute to the formation of MCC (2021), arguing towards a distinct MCC cell of origin.

The orchestrated prerequisite of viral integration, LT stop mutations, and immunosuppression could hold one reason for the rare occurrence of MCC, despite the high MCPyV prevalence within the human population. However, the exact mechanism and the reciprocal influence of the single prerequisites remain to be determined.

3.4.2. Therapy

The aggressive occurrence of MCC necessitates an efficient and specific antiviral treatment, which is hitherto unavailable. Additionally, 30-40% recurrence rates within the first two years after diagnosis complicate the therapy (van Veenendaal et al., 2018; Allen et al., 2005). Interestingly, despite the substantial difference in the mutational burden of vpMCC and vnMCC, their clinical appearance is similar, resulting in an identical therapy. The standard treatment of primary tumors includes spacious resection with subsequent radiotherapy and chemotherapy (Andruska et al., 2021; Bichakjian et al., 2018; Strom et al., 2016). This procedure is effective in 60% of all cases, while only 10% can be cured completely (Cowey et al., 2017; P. Nghiem et al., 2017; Tai et al., 2000). Exclusive treatment with radiotherapy is suggested if radical surgery is not applicable since similar relapse rates compared to combinatory therapy are reported in early-stage MCC (Dubois et al., 2021). The benefit of chemotherapy, on the other hand, is under debate. No unequivocal increase in overall survival could be observed, it does not induce durable responses, and it is accompanied with high toxicity (Cowey et al., 2017; P. Nghiem et al., 2017; Voog et al., 1999).

The introduction of immunotherapy changed MCC treatment profoundly since it could be shown that MCC is highly immunogenic. For example, MCC outcome is improved in tumors with infiltrated CD8+ T cells (Paulson et al., 2011). Further, downregulation of the major histocompatibility complex-class

I (MHC-I) receptors and upregulation of the immune checkpoint protein PD-L1 to evade immune surveillance is reported for MCC (Paulson et al., 2014; Afanasiev et al., 2013). The cell surface receptor programmed cell death protein 1 (PD-1) is expressed by activated T cells and binds PD-L1, which is presented by tumor cells or antigen-presenting cells. This interaction inhibits kinase signaling involved in T cell proliferation and survival and prevents overstimulation of the immune system. The administration of immune checkpoint inhibitors (ICIs), such as anit-PD-1 antibodies, pembrolizumab and nivolumab, and the anti-PD-L1 antibody, avelumab, inhibit the PD-L1-PD-1 interaction and mediate T-cell reactivation and antiviral activity of the adaptive immune system (Topalian et al., 2017; P. T. Nghiem et al., 2016; Kaufman et al., 2016). Durable response rates in 50-70% of MCC cases with metastasis upon treatment with ICIs are reported (J. C. Becker et al., 2019). Despite the success of immunotherapy, a significant portion of patients do not respond to it, highlighting the need for additional therapeutics, such as dimethyl fumarate (DMF). DMF shows anti-neoplastic activity in vnMCC and has been demonstrated to inhibit melanoma cell invasion and metastasis by preventing the nuclear entry of NF-εB (Gambichler et al., 2023; Takeda et al., 2020). Another promising approach would be the direct targeting of the T antigens, representing the main drivers of tumorigenesis in vpMCCs. Unfortunately, no full-length structure of the MCPyV sT or LT could be published thus far, making a structure-based development of antiviral therapeutics challenging. It underscores the importance of uncovering the T antigen structures

4. Aim of the study

MCPyV is one of four clinically relevant hPyV and is the only hPyV associated with tumorigenesis, causing MCC. This rare, aggressive, and highly metastatic skin cancer occurs predominantly in the immunosuppressed and elderly population. MCPyV pathogenesis is mediated by its oncoproteins, sT and tLT, which have been shown to promote cell transformation (Richards et al., 2015; Shuda et al., 2015; Borchert et al., 2014). Full-length LT is essential for initiating viral DNA transcription and replication during the viral life cycle, which is mainly driven by its C-terminal OBD and helicase. Next to tumorigenesis, full-length LT is responsible for the viral replication by its C-terminal OBD and helicase. Interestingly, in 80% of all MCCs, the viral genome is monoclonal integrated, and LT undergoes tumor-specific C-terminal truncations. The loss of the OBD and the helicase leads to viral replication deficiency (H. Feng et al., 2008; Shuda et al., 2008). The lack of specific antiviral treatment and the low 5-year survival rate urges the development of new drug targets (Agelli et al., 2010; Albores-Saavedra et al., 2010). Since the T antigens are key players in the viral life cycle and clinical manifestation of MCPyV, they serve as ideal drug targets for treating MCC and as preventive treatment for risk groups such as solid-organ transplant recipients. A common strategy for antiviral drug development is the structure-based inhibitor design, requiring the atomic structure of the T antigens. However, no full-length LT structure of any PyV is known, and only the structure of SV40 sT could be elucidated (Cho et al., 2007). For MCPyV, solely the LT OBD is available (C. Harrison et al., 2011), while the LT helicase could be elucidated for SV40 and JCPyV (Bonafoux et al., 2016; D. Li et al., 2003). Next to antiviral treatment development, atomic protein structures serve in understanding the protein function, molecular mechanism of action, and protein-protein interaction. This is particularly interesting since the absence of a fully permissive replication system and suitable *in vivo* and *in vitro* transformation models impede the contextualized examination of the T antigens and their underlying mechanisms. Therefore, the aim of this study is the structural characterization of the MCPyV oncoproteins, sT and LT, using two protein structure elucidation methods. X-ray crystallography of *in cellulo* grown crystals of tLT244, sT, and 122tsT will be implemented as *in vitro* crystallization of tLT244 was unsuccessful due to low protein concentration and high polydispersity. These challenges will be addressed by *in cellulo* crystallization since laborious and error prone high yield protein purification and the screening of *in vitro* crystallization conditions can be circumvented. CryoEM SPA will be applied on full-length LT as it can be analyzed under near physiological conditions and does not require crystallization, which its flexible MURs may impede. Additionally, LT forms large complexes, benefiting cryoEM SPA by increasing the signal-to-noise ratio. To improve protein

conditions, the biochemical characterization of LT will precede its TEM analysis to give insights into its stability, DNA-binding behavior, and oligomeric state.

5. Material and methods

5.1. Devices

Table 1: Devices

Device	Manufacturer
Accu-jet® pro	Brand, Wertheim, Germany
Agarose gel electrophoresis chamber	PeqLab, Erlangen, Germany
ÄKTApurifier 10	Cytiva, Marlborough, USA
ÄKTATM pure	Cytiva, Marlborough, USA
Cell counting chamber Neubauer	Hartenstein, Würzburg, Germnay
Cell incubator	Binder, Tuttlingen, Germany
Centrifuges: Eppendorf 5810R Rotor A-4-81, 5424R Rotor F 45-24-11, 5417R Rotor F 45-30-11; Megafuge 3.0 R; Sigma 3-18K	Eppendorf, Hamburg, Germany; Heraeus, Hanau, Germany; Sigma, Osterode am Harz, Germany
Clean bench Hera Safe	Heraeus, Hanau, Germany
Digital Sonifier 250-D	Branson, Danbury, USA
Eclipse TS100	Nikon GmbH, Düsseldorf, Germany
Electroblotter CTI	Hartenstein, Würzburg, Germany
Electron microscope CM120 with LaB6 source and MultiScan 794 CCD camera	Philipps/FEI, Eindhoven, Netherlands
Electron microscope Talos Arctica 200kV with field emission gun source and Falcon 3EC direct detector	Thermo Fisher Scientific, Waltham, USA
Electron microscope Talos L120C 120kV with LaB6 thermionic source with CETA camera	Thermo Fisher Scientific, Waltham, USA
Electrophoresis chamber Mini Protean ® Tetra System	Biorad, Munich, Germany
Film cassette	Hartenstein, Würzburg, Germany
Fluorescence microscope	Zeiss, Jena, Germany
Fraction collector Frac-950	Cytiva, Marlborough, USA
Freezer -20 °C	Liebherr, Hamburg, Germany
GloQube glow discharger	Quorum Technologies Ltd, Laughton, UK
Hi 2211 pH/ORP meter	Hanna instruments, Woonsocket, USA
High Precision Cell Quartz SUPRASIL® cuvette	Hellma Analytics, Müllheim, Germany
Illuminator Intensilight C-HGFI	Nikon GmbH, Düsseldorf, Germany
Image QuantTM 800 Westernblot imaging system	Amersham Bioscience/Cytiva, Marlborough, USA
Incubator Multitron pro	Infors HT; Einsbach, Germany

Incubator shaker Certomat ® BS-1; Thermomixer 5436, Thermomixer compact	Satorius, Göttingen, Germany; Eppendorf, Germany
Microscope Axiovert 25	Zeiss, Jena, Germany
Microscope Leica DMI 6000 TCS SP5 Tandem	Leica, Wetzlar, Germany
NanoDrop® ND-1000	PeqLab, Erlangen, Germany
PCR bench	LFT Labortechnik, Wasserburg, Germany
Photometer Ultrospec 3000 pro	Amersham Bioscience/Cytiva, Marlborough, USA
Pipettes: 2, 10, 20, 100, 200, 1000 µl Research Plus	Eppendorf, Hamburg, Germany
Plate reader Infinite M2000	Tecan, Männedorf, Switzerland
Powersupply Power Pac 1000, Power Pac Universal	Biorad, Munich, Germany
Prometheus NT.48 fluorometer	NanoTemper Technologies, Munich, Germany
Qubit® 2.0 Fluorometer	Invitrogen, Darmstadt, Germany
Rotilabo-Block-Heater H250	Roth, Karlsruhe, Germany
Sample loop 1ml	Cytiva, Marlborough, USA
SpectroSize 300	Xtal Concepts, Hamburg, Germany
Superloop 10 ml, 50 ml	Cytiva, Marlborough, USA
Thermocycler ep gradients	Eppendorf, Hamburg, Germany
Thermomixer compact	Eppendorf, Hamburg, Germany
Tweezers Nr 5	Manufactures D´Outils Dumont SA, Montignez, Switzerland
Ultramicrotome Leica Ultracut EM UCT	Leica Microsystems, Vienna, Austria
UV-Transilluminator and Detector	Polaroid, Dreieich-Sprendlingen, Germany
Vitrobot Mark IV plunge freezer	FEI/Thermo Fisher Scientific, Waltham, USA
Vortex-Genie2	Scientific Industries, Bohemia, USA
Water bath GFL Typ 1013	GFL, Burgwedel, Germany
Weighing scales	Sartorius, Göttingen, Germany
X-ray film developer Curix 60	Agfa, Cologne, Germany

5.2. Disposables and reagents

Pipette tips, serological pipettes, cell culture dishes and flasks, reagent flasks and falcons were obtained from Falcon (Becton and Dickinson Europe; Le Pont de Claix, France), Sarstedt (Nürnberg, Germany), Greiner (Kremsmünster, Österreich) and Eppendorf (Hamburg, Germany).

Table 2: Disposables and reagents

Name	Manufacturer
10x FastDigest® Green buffer	Thermo Fisher Scientific, Waltham, USA
2% Uranyl acetate solution in H_2O	Electron Microscopy Sciences, Hatfield, USA
30% Acrylamide/Bis-acrylamide solution	Sigma, Munich, Germany
6x DNA loading dye	Thermo Fisher Scientific, Waltham, USA
ABTS™ solution	Roche Diagnostics, Mannheim, Germany
Alkaline Phosphatase	Roche Diagnostics, Mannheim, Germany
Amicon® Ultra Centrifugal Filter Units 100K, 50k, 30k, 10k, 3k	Merk Millipore, Darmstadt, Germany
Ampicillin Sodium Salt	Sigma, Munich, Germany
AMP-PNP	Sigma, Munich, Germany
Aqua	B. Braun, Melsungen, Germany
Benzonase Nuclease	Sigma, Munich, Germany
Biotin	Thermo Fisher Scientific, Waltham, USA
Bromophenol blue	Serva, Heidelberg, Germany
BSA	GE Healthcare, Freiburg, Germany
C-flat™ holey carbon-coated copper grid; 2/2 hole diameter/hole spacing; 400 mesh	Quantifoil Micro Tools, Großlöbichau, Germany
Color Protein Marker	New England Biolabs, Frankfurt am Main; Germany
cOmplete™ Protease Inhibitor Cocktail, Mini	Roche Diagnostics, Mannheim, Germany
Coverslips	A. Hartenstein, Würzburg, Germany
Cryo autogrid box	SubAngstrom, New York City, USA
Cryo autogrid C-clip	Thermo Fisher Scientific, Waltham, USA
Cryo autogrid C-clip rings	Thermo Fisher Scientific, Waltham, USA
Distilled Water DNase/ RNase free	Life Technologies, Carlsbad, USA
DNA loading buffer	Thermo Fisher Scientific, Waltham, USA
dNTPs (dATP, dCTP, dGTP, dTTP)	Invitrogen, Karlsruhe, Germany
DTT	Sigma, Munich, Germany
EM grid with continuous carbon film 300 mesh, Cu	Electron Microscopy Sciences, Hatfield, USA

Epon 812	Roth, Karlsruhe, Germany
FastDigest® restriction enzymes	Thermo Fisher Scientific, Waltham, USA
Gel filtration protein standard 1.35-670 kDa	BioRad, Biorad, Munich, Germany
GeneRuler™ 1kb DNA Ladder	Thermo Fisher Scientific, Waltham, USA
Gentamycin sulfate	Serva, Heidelberg, Germany
IPTG	Thermo Fisher Scientific, Waltham, USA
Kanamycin sulfate	Roth, Karlsruhe, Germany
Laboratory film	Parafilm, Neenah, USA
Maltose	Sigma, Munich, Germany
Maxima SYBR Green/ROX®qPCR Master Mix (2x)	Thermo Fisher Scientific, Waltham, USA
Paraformaldehyde (16%)	Electron Microscopy Sciences, Hatfield, USA
PEI	Polyscience Inc., Washington, USA
Restriction enzymes	New England Biolabs, Frankfurt am Main; Germany
Roti® GelStain	Roth, Karlsruhe, Germany
Roti® -PVDF-Membrane (0,45 µm)	Roth, Karlsruhe, Germany
Skimmed milk powder Sucofin®	Tsi GmbH & Co. KG, Zeven, Germany
Slide-A-Lyzer® Dialysis Cassette G2 Dialysis Unit, 20K 3 ml, 2K 0.5 ml	Thermo Fisher Scientific, Waltham, USA
Sonicated Sperm DNA	Invitrogen, Karlsruhe, Germany
Specimen slide	Glaswarenfabrik Karl Hecht, Sondheim, Germany
Steritop Filter Units (0.22 µm), Sterilfilter Millex-GP (0,22 µm); Whatman GD, X 25 Syringe Filter (0,45 µm)	Merck Millipore, Darmstadt, Germany; GE Healthcare, Freiburg, Germany
T4-DNA-Ligase (400U/ml); T4-DNA-Ligase (2000 000U/ml)	Invitrogen, Karlsruhe, Germany; NewEngland Biolabs GmbH, Frankfurt am Main, Germany
Tetracycline hydrochloride	Roth, Karlsruhe, Germany
TEV protease, recombinant 10 U/µl	Serva, Heidelberg, Germany
Ultra Pure™ Agarose	Invitrogen, Karlsruhe, Germany
Vectashield mounting medium with DAPI	Vektor Laboratories Inc, Burlingame, USA
Whatman 3MM paper	A. Hartenstein, Würzburg, Germany
X-Gal	Thermo Fisher Scientific, Waltham, USA
X-ray films	Fuji-Film Europe, Düsseldorf, Germany
X-tremeGENE™ HP DNA Transfection Reagent	Roche Diagnostics, Mannheim, Germany
µ-Dish 35 mM, high Grid-500, ibiTreat	Ibidi, Gräfelfing, Germany

5.3.Columns and matrices for protein purification

Table 3: Columns and matrices

Name	Manufacturer
HiLoad Superdex 200 10/300 GL	Cytiva, Marlborough, USA
HisTrap HP 1ml, 5 ml	Cytiva, Marlborough, USA
HiTrap Heparin HP 1 ml	Cytiva, Marlborough, USA
HiTrap™ Q FF 5 ml	Cytiva, Marlborough, USA
MBP amylose	New England Biolabs, Frankfurt am Main, Germany
MBPTrap 1ml, 5 ml	Cytiva, Marlborough, USA
NiNTA agarose	Qiagen, Hilden, Germany
Polypropylene column 1ml, 5 ml	Qiagen, Hilden, Germany
Protein G Sepharose	Cytiva, Marlborough, USA
StrepTactin Sepharose	Iba Lifesciences GmbH, Göttingen, Germany
Superose 6 Increase 10/300 GL	Cytiva, Marlborough, USA

5.4. Buffers

Chemicals were obtained from Roth (Karlsruhe, Germany), Sigma (Munich, Germany), Fluka (Neu-Ulm, Germany) or Merk (Darmstadt, Germany). Buffers were prepared in double distilled water, autoclaved or sterile filtered.

Table 4: Buffers for various applications

Name	Buffer
5x annealing buffer	500 mM TRIS-HCl 50 mM $MgCl_2$ 17 mM DTT
Blocking buffer (IF)	1x PBS 1% Triton X-100 0.5% Tween 20 3% BSA
Coomassie destaining solution	50% (v/v) H_2O 40% (v/v) Ethanol 10% (v/v) Acetic acid (96%)
Coomassie staining solution	50% (v/v) H_2O 40% (v/v) Ethanol 10% (v/v) Acetic acid (96%) 0.25% (w/v) Coomassie Brilliant Blue G-250
Fixing solution (DPI-ELISA)	100 mM Na_2CO_3

	pH 9.7
PBS, pH 7.5	10 mM Na_2HPO_4 140 mM NaCl 25 mM KCl 1 mM $CaCl_2$ 0.5 mM $MgCl_2$
PBS-Tween	1x PBS 0.05% (v/v) Tween 20
PDBB (DPI-ELISA)	4 mM HEPES, pH 7.9 100 mM KCl 8% Glycerol 0.2% BSA 5.5 mM DTT 1 pmol/100 µl unbiotinylated ori 2 mM $MgCl_2$ 4 mM ADP
Permeabilization buffer (IF)	1x PBS 1% Triton X-100 0.1% sodium citrate
Running gel	4-12% (v/v) Acryl-Bisacrylamide 375 mM TRIS-HCl, pH 8.8 0.1% (w/v) SDS 0.1% (w/v) APS 0.04% (v/v) TEMED
SDS running buffer	25 mM Tris 0.192 M glycine 3.5 mM SDS
Stacking gel	4% (v/v) Acryl-Bisacrylamide 125 mM TRIS-HCl, pH 6.8 0.1% (w/v) SDS 0.1% (w/v) APS 0.1% (v/v) TEMED
TAE, pH 7.8	2 M TRIS-HCl 0.2 M sodium acetate 1 mM EDTA
TBS, pH 7.5	50 mM TRIS-HCl 150 mM NaCl
TBS-T	1x TBS 0.05% (v/v) Tween 20
TFB1, pH 5.8	15% (v/v) Glycerol 10 mM $CaCl_2$ 30 mM Potassium acetate 100 mM $RbCl_2$ 50 mM $MnCl_2$
TFB2	15% (v/v) Glycerol

	10 mM MOPS 75 mM $CaCl_2$ 10 mM $RbCl_2$
WB blocking solution	5% (w/v) milk in PBS-T or TBS-T
WB Buffer A, pH 10.4	0.3 M Tris 10% (v/v) methanol
WB Buffer B, pH 10.4	0.025 M Tris 10% (v/v) methanol
WB Buffer C, pH 9.4	0.025 M Tris 0.025 M aminohexanoic acid 10% (v/v) methanol

Table 5: Buffer for protein purification

Name	Buffer
B1	50 mM TRIS-HCl 30 mM Na_2HPO_4 250 M NaCl 5% Glycerol 1 mM DTT 2 mM $MgCl_2$ 2 mM AMP-PnP 0.01% (w/v) Triton X-100 pH 8.0
EB1	50 mM TRIS-HCl 30 mM Na_2HPO_4 250 M NaCl 5% Glycerol 1 mM DTT 10 mM Maltose pH 8.0
EB2	50 mM TRIS-HCl 250 M NaCl 5% Glycerol 1 mM DTT 10 mM Maltose pH 8.5
EB3	50 mM TRIS-HCl 1 M NaCl 5% Glycerol 1 mM DTT pH 8.5
EB4	50 mM TRIS-HCl 250 M NaCl 5% Glycerol 1 mM DTT

	pH 8.5
EB5	50 mM TRIS-HCl 30 mM Na$_2$HPO$_4$ 250 M NaCl 5% Glycerol 1 mM DTT pH 8.0
IEXB1	50 mM TRIS-HCl 100 mM NaCl 5% Glycerol 1 mM DTT pH 8.5
LB1	50 mM TRIS-HCl 30 mM Na$_2$HPO$_4$ 1 M NaCl 5% Glycerol 1 mM DTT 0.1% (v/v) Triton X-100 1 cOmplete™, EDTA-free Protease Inhibitor Cocktail/50 ml pH 8.0
LB2	50 mM TRIS-HCl 30 mM Na$_2$HPO$_4$ 250 M NaCl 5% Glycerol 1 mM DTT 0.1% (v/v) Triton X-100 1 cOmplete™, EDTA-free Protease Inhibitor Cocktail/50 ml pH 8.0
LB3	50 mM TRIS-HCl 250 M NaCl 5% Glycerol 1 mM DTT 0.1% (v/v) Triton X-100 1 cOmplete™, EDTA-free Protease Inhibitor Cocktail/50 ml pH 8.5
LB4	50 mM TRIS-HCl 1 M NaCl 10% Glycerol 1 mM DTT 10U/ml Benzonase 0.1% (v/v) Triton X-100 1 cOmplete™, EDTA-free Protease Inhibitor Cocktail/50 ml pH 8.5

LB5	50 mM TRIS-HCl 1 M NaCl 10% Glycerol 1 mM DTT 10U/ml Benzonase 1% (v/v) Nonidet P-40 1 cOmplete™, EDTA-free Protease Inhibitor Cocktail/50 ml pH 8.5
LB6a	10 mM TRIS-HCl 10 mM KCl 0.1 mM EDTA 2.5 mM DTT 1 mM PMSF 1 mM NaF 1 mM Na_3VO_4 2 mM β-glycerolphosphate 1 cOmplete™, EDTA-free Protease Inhibitor Cocktail/50 ml pH 8.5
LB6b	50 mM TRIS-HCl 420 mM NaCl 1 mM EDTA 2.5 mM DTT 1 mM PMSF 1 mM NaF 1 mM Na_3VO_4 2 mM β-glycerolphosphate 10% Glycerol 1 cOmplete™, EDTA-free Protease Inhibitor Cocktail/50 ml pH 8.5
LB7	50 mM TRIS-HCl 1 M NaCl 5% Glycerol 1 mM DTT 10U/ml Benzonase 0.1% (v/v) Triton X-100 1 cOmplete™, EDTA-free Protease Inhibitor Cocktail/50 ml pH 8.5
RIPA buffer	50 mM TRIS-HCl pH 7.5 150 mM NaCl 1% NP40 0.5% DOC 0.1% SDS 2 mM β-glycerolphosphate 1 mM Na3VO4 0.4 mM PMSF

1 mM NaF
1 mM EGTA
Protease inhibitor 1 tablet/50mL

5.5.Commercial systems

Table 6: Commercial systems

Kit	Manufacturer
AccuStart™ II GelTrack PCR SuperMix	Quantabio, Beverly, USA
Biorad Protein Assay	Biorad, Munich, Germany
DNA Plasmid Purification Kit NucleoBond PC100	Macherey-Nagel, Düren, Germany
DNeasy® Blood & Tissue Kit	Quiagen, Hilden, Germany
femto LUCENT™ PLUS-HRP Kit	Biosciences, St. Louis, USA
NEBuilder HiFi DNA Assembly Master Mix	New England Biolabs, Frankfurt am Main; Germany
NucleoSpin® Gel and PCR Clean-up	Macherey-Nagel, Düren, Germany
PeqGold Plasmid Miniprep Kit	Peqlab, Erlangen, Germany
Phusion® High-Fidelity PCR Kit	New England Biolabs, Frankfurt am Main, Germany
Qubit® Protein Assay Kit	Invitrogen, Darmstadt, Germany
Western Lightning® Plus-ECL	PerkinElmer, Waltham, USA

5.6.Growth media and additives

Table 7: Growth media and additives

Name	Manufacturer
DMEM + Glutamax	Life Technologies, Carlsbad, USA
DPBS	Life Technologies, Carlsbad, USA
ESF921 Insect Cell Culture Medium (for High Five™)	Expression Systems, Davis, USA
Ex-Cell® 420 Serum-free Medium (for Sf9)	Sigma, Munich, Germany
FBS	Life Technologies, Carlsbad, USA
Gelatine solution Type b 2% in H_2O	Sigma, Munich, Germany
HEPES 1 M	Life Technologies, Carlsbad, USA
LB agar	Roth, Karlsruhe, Germany
LB medium	Roth, Karlsruhe, Germany
L-glutamine, 200 mM	Life Technologies, Carlsbad, USA
OptiMEM™ I Reduced Serum Medium	Life Technologies, Carlsbad, USA

Pen/Strep (10 U/µl penicillin/10 µg/µl streptomycin)	Life Technologies, Carlsbad, USA
Sodium pyruvate 100 mM	Thermo Fisher Scientific, Waltham, USA
Trypsin + EDTA 0.05%	Life Technologies, Carlsbad, USA

5.7. Plasmids

Table 8: Plasmids

Name	Description
pFastBac1-LT-His-StrepII	cDNA of MCPyV LT 1-817 with mutated splice sites to inhibit 57K expression; C-terminal his, and StrepII tag, which are cleavable via TEV recognition site; polyhedrin promoter; amp resistance;
pFastBac1-LT-His-StrepII-MBP	cDNA of MCPyV LT aa 1-817 with mutated splice sites to inhibit 57K expression; C-terminal his, StrepII, and MBP tag, which are cleavable via TEV recognition site; polyhedrin promoter; amp resistance;
pFastBac1-ori	MCPyV ori with a size of 66 bp, including M1-6; polyhedrin promoter; amp resistance;
pFastBac1-tLT244-iBox-PAK4cat	cDNA of MCPyV LT aa 1-244; C-terminally linked to the interaction site of the PAK4 inhibitor (iBox) aa 166-203 and the catalytic domain of PAK4 (PAK4cat) aa 278-591; polyhedrin promoter; amp resistance;
pFastBacHT-His-sT-iBox-PAK4cat	cDNA of MCPyV sT aa 1-186; C-terminally linked to the interaction site of the PAK4 inhibitor (iBox) aa 166-203 and the catalytic center of PAK4 (PAK4cat) aa 278-591, which are separated by an EcoRI restriction site; N-terminal his tag; polyhedrin promoter; amp resistance;
pXJ40-122tsT-Inka1	PAK4 inhibitor (Inka1) N-terminally linked to 122tsT aa 122-186; CMV promoter; amp resistance;
pXJ40-Flag-PAK4cat	Catalytic domain of PAK4 (PAK4cat) aa 278-591; N-terminal Flag tag; CMV promoter; amp resistance; kindly provided by Yohendran Baskaran (Institute of Molecular and Cell Biology, Singapore);
pXJ40-HA-Inka1	PAK4 inhibitor (Inka1); N-terminal HA tag; CMV promoter; amp resistance; kindly provided by Yohendran Baskaran (Institute of Molecular and Cell Biology, Singapore);

pXJ40-Inka1-122tsT	PAK4 inhibitor (Inka1) C-terminally linked to 122tsT aa 122-186; CMV promoter; amp resistance;
pXJ40-sT-Inka1	PAK4 inhibitor (Inka1) N-terminally linked to sT; N-terminal HA tag; CMV promoter; amp resistance;
pXJ40-tLT244-Inka1	PAK4 inhibitor (Inka1) N-terminally linked to tLT244 aa 1-244; N-terminal HA tag CMV promoter; amp resistance;

5.8.Oligonucleotides

All oligonucleotides were purchased from Eurofins Genomics or Sigma.

Table 9: Oligonucleotides

Name	Sequence
MCPyV ori	GGCAGAGGCTTGGGGCTCCTAGCCTCCGAGGCCTCTGGAAAAAAAAGAGAGAG GCCTCTGAGGCTTAAGAGGCTTAATTA
MCPyV ori -	CGGGATCCTCTTTTTTTTTCCAGAGGCC
MCPyV VP1	TCAGACTCCTACAGTTCTTCAATTTAGTAATACTCTAACTACTGTCCTTTTAGA TGAGAATGGAGTGGGCCCTCTATGCA
MCyV ori scr	GCTCTCTCGTCACACCTGCATCACCAGGAGACGTGACTGCTCACACCTTGGTAT TCTCACACGTCGTATGGTCGTCTATG
MCyV viral miR-M1 promoter	GCCTGAGGAGCCCCCCTCCTCAAGAAGCTCGCCCCGGCAGCCCCCGTCTTCCTCT GCCGAGGAGGCCTCGTCATCTCAGT
pUC/M13 -	AGCGGATAACAATTTCACACAGG
pUC/M13 +	CCCAGTCACGACGTTGTAAAACG
TP53INP1 promoter	TGCTGGGAGGGGGCCGCTGCTGGGCCCCAGATAGAGGGGTAGAGGCCGCTCAA TCACGGAGCCAGGCCCCACCTGCTGTA

Table 10: Primers for cloning

Name	Sequence	Purpose
p753 Gib+	CGAATCTAGAGCCTGCAGTCATGGAT TTAGTCCTAAATAG	Generation of the plasmid pFastBac1-LT-His-StrepII
p753 Gib-	CTGGTACCGCATGCCTCGATTGAGAA AAAGTACCAGAATC	
p753 Gib+	CGAATCTAGAGCCTGCAGTCATGGAT TTAGTCCTAAATAG	Generation of the plasmid pFastBac1-LT-His-StrepII-MBP
p753 Gib-	CTGGTACCGCATGCCTCGATTGAGAA AAAGTACCAGAATC	
p771 Gib+	CACCCGCAGTTCGAAAAACCCGGGAA AATCGAAGAAGGTAAA	

Material and methods

p771 Gib-	CTAGTACTTCTCGACAAGCTTTATGTC TGCGCGTCTTTCAGG	
p646 Gib+	CCGTCCCACCATCGGGCGCGGATCCAC CGGTATGGATTTAGTCCTAAATAGGA AAGAAAGAGAGG	Generation of the plasmid pFastBac1-tLT244-iBox-PAK4cat
p660 Gib-	GTCCTCTGCTTCCATCTCGAGTCTGGG CTGCCGGGGCGAG	
p653 Gib+	CCTGTATTTTCAGGGCGCCATACCGGT GGATCCATGGATTTAGTCCTAAATAG G	Generation of the plasmid pFastBacHT-His-sT-iBox-PAK4cat
p653 Gib-	GATCCTCTAGTACTTCTCGACCCCGGG AAGCTTTCATCTGGTGCGGTTCTG	
p721 Gib+	GCCAGACTACGCAGGATCCATTAGCTG TAAGTTGTCTCG	Generation of the plasmid pXJ40-122tsT-Inka1
Inka1_sT Gib-	CGAGCGCTGTGCATAAGCTTGAAAAG GTGCAGATGCAGTAAGC	
Inka1_tsT Gib+ tsT_pXJ Gib-	GATGTCAGCTACCTCGGTACCATTAGC TGTAAGTTGTCTC GTTCTGCTTTAATAagatctCTAGAAAA GGTGCAGATGC	Generation of the plasmid pXJ40-Inka1-122tsT
p697 Gib+	GACTCACTATAGGGCGAATTCCACCAT GGATTTAGTCCTAAATAG	Generation of the plasmid pXJ40-sT-Inka1
p693 Gib-	CCGAGCGCTGTGCATAAGCGGATCCA AAAGGTGCAGATGCAGTAAG	
p697 Gib+	GACTCACTATAGGGCGAATTCCACCAT GGATTTAGTCCTAAATAG	Generation of the plasmid pXJ40-tLT244-Inka1
p697 Gib-	CCGAGCGCTGTGCATCGAAGCGGATCC TCTGGGCTGCCGGGGCGAGC	

5.9. Antibodies

Table 11: Primary and secondary antibodies

Target	Name	Concentration	Manufacturer
Anti-His tag		1:10,000 (WB)	Biolegend, San Diego, USA
Flag tag	ANTI-FLAG® M2 Monoclonal Antibody, mouse	1:3000 (WB) 1:500 (IF)	Sigma, Munich, Germany
Ha tag	Anti-HA-Peroxidase, rat (3F10)	1:1000 (WB) 1:200 (IF)	Roche Diagnostics, Mannheim, Germany
MCPyV LT	CM2B4, mouse (sc-136172)	1:1000 (WB, DPI-ELISA) 1:500 (IF)	Santa Cruz Biotechnology, Heidelberg, Germany
MCPyV sT/tLT	2T2, mouse	1:1 (WB)	self-made (supernatant from 2T2 hybridoma cell line)

39

Mouse IgG-Fluorescein Isothiocyanate (FITC)		1:1000 (IF)	Santa Cruz Biotechnology, Heidelberg, Germany
Mouse IgG-HRP	Peroxidase-AffiniPure goat Anti-Mouse IgG	1:10,000 (WB) 1:5000 (DPI-ELISA)	Santa Cruz Biotechnology, Heidelberg, Germany
Mouse IgG-TRITC		1:500 (IF)	Santa Cruz Biotechnology, Heidelberg, Germany
Rat IgG-Rhodamine Red	Rhodamine Red™-X-conjugated AffiniPure Goat Anti-Rat IgG	1:2000 (IF)	Jackson ImmunoResearch Laboratories Inc, West Grove, USA

5.10. Bacteria strains

Table 12: Bacteria strains

Name	Manufacturer
E. coli TOP10	Invitrogen, Karlsruhe, Germany
MAX Efficiency® DH10Bac™ Competent cells	New England Biolabs, Frankfurt am Main; Germany

5.11. Eukaryotic cell lines

Table 13: Mammalian cell lines

Name	Description	Manufacturer
2T2 hybridoma cell line	Hybrid of murine immunized B-cells and myeloma cells (SP2/mIL-6), required for the sT antibody 2T2 production	Kindly provided by Christopher Buck (NIH, MD, USA)
HEK293	Human embryonic kidney cell line, transformed via the integration of the first 4344 bp of the AdV5 genome (ATCC no: CRL-1573)	

Table 14: Insect cell lines

Name	Description	Manufacturer
High Five™ (BTI-Tn-5B1-4)	Immortalized ovarian cells of the cabbage looper *Trichoplusia ni*	Kindly provided by Lars Redecke (Lübeck, Germany)
Sf9	Clonal isolate of immortalized ovarian cells of the *Spodoptera frugiperda* SF21 cell strain (IPLB-Sf21-AE) (ATCC no: CRL-1711)	Thermo Fisher Scientific, Waltham, USA

5.12. Data processing

Table 15: Software

Name	Implementation	Manufacturer
AlphaFold2	Protein molecular structure prediction	(Jumper et al., 2021)
CLC DNA Workbench 20	Sequence data processing	Quiagen, Venlo, Netherlands
Clustal Omega	Multiple sequence alignment	(Goujon et al., 2010)
EPU software	cryoEM SPA data acquisition	Thermo Fisher Scientific, Waltham, USA
ExPASy	Protein characteristics prediction	(Gasteiger et al., 2003)
GraphPad Prism 5	Data visualization and statistics	GraphPad Software, San Diego, USA
ImageJ	Image processing and analysis	(Schneider et al., 2012)
IUPred3	Prediction of intrinsically disordered regions	(Erdős et al., 2021)
JPred 4	Protein secondary structure prediction	(Drozdetskiy et al., 2015)
Microsoft Office	Text and numerical data processing	Microsoft, Redmond, USA
Relion 3.1	EM data processing	(Scheres, 2012)
ResMap	cryoEM SPA local resolution estimation of the model	(Kucukelbir et al., 2014)
UCSF ChimeraX	Protein structure analysis and visualization	(Pettersen et al., 2021)

5.13. Procaryotic cell culture

Bacteria were used for plasmid amplification and the generation of bacmids. They were maintained in LB medium or on agar plates. The medium and agar were prepared by using the indicated manufacturer´s instructions and autoclaved for 20 min at 121 °C and 1.4 bar. Plasmid-specific antibiotics were added according to Table 16. All bacteria were cultivated overnight at 37 °C. Suspension cultures were additionally stirred at 180 rpm.

Table 16: Antibiotic concentrations used in procaryotic cell culture

Name	Final concentration [mg/l]
Ampicillin	100
Kanamycin	50
Tetracycline	10
Gentamycin	7

5.13.1. Production of chemically competent bacteria

The introduction of exogenous DNA into *E. coli* prerequisites a permeable cell wall, which is attained by the treatment with buffer TFB1 and TFB2 (Table 4).

A single colony of *E. coli* TOP10 or DH10Bac was cultivated in 20 ml LB medium overnight at 37 °C, followed by dilution with fresh LB medium at a ratio of 1:100. The culture was complemented with 4 mM $MgSO_4$ and 10 mM KCl and incubated at 37 °C until it reached an OD_{600} between 0.3-0.5. The cells were stored on ice for 15 min and pelleted at 6000 xg for 5 min at 4 °C. Subsequently, the pellet was resuspended in 60 ml TFB1, incubated on ice for 1 h, and centrifuged at 6000 xg for 5 min at 4 °C. The final pellet was resuspended in 8 ml TFB2 and aliquots of 100 µl were immediately flash-frozen with liquid nitrogen and stored at -80 °C.

5.13.2. Transformation of chemically competent bacteria

Chemically competent cells are treated with heat to increase the uptake of exogenous DNA by inducing the fusion of the inner and outer cell membrane, leading to pore formation in the cell wall.

E. coli cells (100 µl) were thawed on ice and mixed with the entire ligation batch, or with 100 ng of plasmid DNA, and incubated for 30 min at 4 °C. A heat shock was performed for 45 sec at 42 °C, followed by diluting the mixture with 500 µl LB medium (TOP10) or 900 µl SOC medium (DH10Bac). The transformants were incubated for 1 h (TOP10) or 4 h (DH10Bac) at 37 °C and plated on LB agar, which was supplemented with antibiotics (Table 16).

5.13.3. Blue-white selection

The successful insertion of DNA fragments into a vector, including the lacZ gene, is examined via blue-white selection. The lacZ gene, encoding for β-galactosidase, is disrupted during DNA insertion, leading to the loss of the chromogenic substrate X-Gal (5-Bromo-4-chloro-3-indoxyl-β-D-galactopyranosid) conversion into its blue product. IPTG controls the β-galactosidase expression.

DH10bac transformants (30 µl) were plated onto LB agar plates, which were complemented with gentamycin, tetracycline, and kanamycin as well as coated with 20 µl IPTG (0.1 M) and 20 µl X-Gal (40 mg/ml). The plates were incubated overnight at 37 °C, and white clones were picked for baculovirus production.

5.14. Eukaryotic cell culture

5.14.1. Cultivation and storage of mammalian cells

Mammalian cells (Table 13) were cultured at 37 °C, 5% CO_2, and 95% humidity in polystyrene dishes. Different media and specific supplements were used, which are summarized in Table 17.

Table 17 Cell culture media and additives used for different mammalian cell lines

Cell line	Cell culture medium	Supplements
HEK293	DMEM (Dulbecco´s Modified Eagle Medium)	10% FCS, 1% P/S
2T2 hybridoma cell line	RPMI 1640, 1x HFCS	10% FCS, 1% P/S

All cells were passaged twice a week at a ratio of 1:10. They were washed with DPBS and detached with trypsin/EDTA. After 5 min incubation at 37 °C, the cells were suspended in supplemented DMEM and re-seeded at the desired concentration.

For cryopreservation, 5-10 x 10^6 cells were pelleted at 200 xg for 5 min and resuspended in sterile filtered FBS supplemented with 10% DMSO. The cells were stored in cryotubes and gently cooled down to -80 °C, using a Mr. Frosty™ freezing container. After 24 h, the cells were transferred to -196 °C for long term storage.

Cells were thawed at 37 °C, washed in fresh medium to remove DMSO, and transferred to a 10 cm polystyrene dish with 10 ml pre-warmed medium. The medium was changed after 12 h to remove cell debris.

5.14.2. Cultivation and storage of insect cells

Sf9 and High Five™ insect cells were grown either as suspension or adherent culture at 27 °C in polystyrene flasks or Erlenmeyer glass flasks, respectively. Suspension cultures were shaken at 100 rpm. Different media with distinct supplement concentrations were used, which are summarized in Table 18.

Table 18 Cell culture media and additives used for different insect cell lines

Cell line	Cell culture medium	Supplements
Sf9	Insect-XPRESS™ Protein-free Insect Cell Medium + L-glutamine	0.25% P/S
High Five™	ESF921	1% P/S

Cells were passaged twice (High Five™) or once (Sf9) a week by reseeding them to a concentration of 5×10^5 cells/ml into pre-warmed, fresh medium with additives.

For cryopreservation, 10^7 cells were pelleted at 200 xg for 5 min. The cells were resuspended in 1.5 ml sterile filtered freezing medium, comprising 46.25% fresh medium, 46.25% cell culture supernatant, and 7.5% DMSO, followed by the transfer into a cryotube. Gentle freezing to -80 °C was secured by using a Mr. Frosty™ freezing container. After 24 h, the cells were transferred to -196 °C for long-term storage.

Cells were thawed by incubation at RT, washed with fresh medium, and transferred to a T75 flask with 10 ml pre-warmed medium. After 1 h, the medium was changed to remove cell debris.

5.14.3. Transfection of eukaryotic cells

The non-viral introduction of nucleic acid into eukaryotic cells is called transfection. Thereby, negatively charged nucleic acids form complexes with specific chemicals, such as polyethyleneimine (PEI) or X-tremeGENE™ HP, to improve their uptake into the cell.

Transfection of mammalian cells with polyethyleneimine

HEK293 cells were seeded one day pre-transfection to obtain a confluency of 60-90% (Table 19). The desired DNA amount (Table 19), diluted with the hundredfold quantity of DMEM without additives, was mixed with the tenfold amount of PEI, incubated for 30 min and added drop by drop to the cells. Prior to transfection, the cell medium was substituted with DMEM without supplements. After 12 h, the medium was changed to DMEM with supplements (Table 17).

Table 19: DNA amount and cell seeding density used for transfection

Format	Seeded cell density	Transfected amount [µg]	DNA
12-well plate	2.5×10^5	1	
µ-Dish 35 mM/6-well plate	6.0×10^5	1.5	
T25 flask	1.4×10^6	4.5	
10 cm dish	4.4×10^6	10	

Transfection of insect cells with X-tremeGENE™ HP

For the generation of recombinant baculovirus (P1), Sf9 cells were seeded at a density of 5×10^5 cells in a 12-well format, filled with 1 ml supplemented medium (Table 18). After 1 h incubation at 27 °C, the cells were transfected with a transfection mixture composed of 1,5 µg DNA, 100 µl OptiMEM™, and 8 µl X-tremGENE™. The mixture was incubated for 30 min at RT and added dropwise to the cells.

5.14.4. Generation and amplification of recombinant baculovirus

The baculovirus expression system is used to express recombinant proteins in insect cells. In this work, the Bac-to-Bac® Baculovirus expression system of Invitrogen was used, summarized in Figure 9.
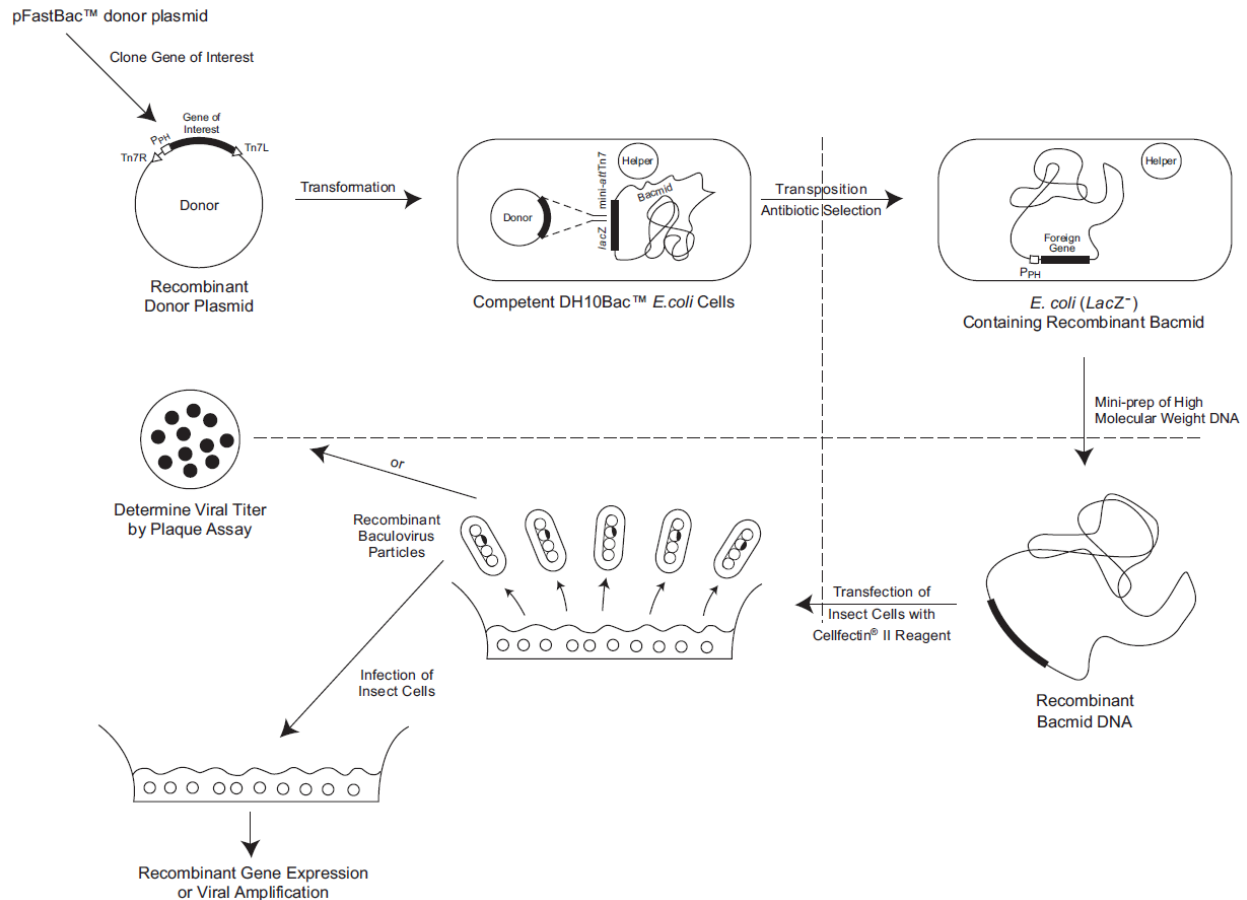
Figure 9: Bac-to-Bac® Baculovirus expression system. The donor plasmid pFastBac, containing the gene of interest under the control of the polyhedrin promoter (P_{PH}), which is flanked by the right and left arm of the Tn7 transposon (Tn7R,L), is transformed into DH10Bac™. These cells carry a baculovirus shuttle vector (bacmid), which harbors a *lacZ* gene with an inserted transposon attachment site (mini-*att*Tn7) and a transposase encoded helper plasmid. A site-specific transposition of the gene cassette of interest causes the *lacZ* disruption, which can be monitored via blue-white selection. Thereafter, the recombinant bacmid is isolated and used to transfect Sf9 cells for recombinant baculovirus production. These viral particles can be utilized to determine the viral titer via plaque assay or to transduce insect cells for either virus amplification or for recombinant protein production.

The donor plasmid (pFastBac, gentamycin resistance), containing the gene of interest under the control of the late polyhedrin promoter and a mini-Tn7 element, was transformed into DH10bac™ *E. coli* cells (5.13.2). These cells are equipped with the transposase encoded helper plasmid (tetracycline resistance) and a baculovirus shuttle vector (bacmid, kanamycin resistance). The bacmid harbors the *lacZ* gene, in which a mini-*att*Tn7 target site is inserted. During recombination, the gene cassette of interest is transposed into the mini-*att*Tn7 site and disrupts the *lacZ* gene. The successful transposition was monitored via blue-white selection (5.13.3). Bacmid DNA was extracted from white colonies via alkaline lysis (5.15.9) with the following modification. A volume of 300 μl of the buffer S1-3 were used to lyse the cells, to precipitate nucleic acid, and to resolubilize plasmid DNA, respectively. After centrifugation at 11,000 xg for 10 min at 4 °C, plasmid DNA was precipitated with 700 μl isopropanol (100%) and centrifuged with the above stated conditions. The pellet was washed

with 200 µl EtOH (70%), centrifuged, and resuspended in 50 µl TE buffer. A control PCR verified the successful transposition of the gene of interest by using the insert site flanking primers, pUC/M13 forward and reverse. The PCR product was analyzed via agarose gel electrophoresis (5.15.4). The production of the first baculovirus generation (P1) was carried out by transfecting 5×10^5 Sf9 cells in a 12-well plate with bacmid DNA and X-tremeGENE™ (5.14.3). After 4-5 days incubation at 27 °C, the supernatant was harvested and centrifuged at 200 xg for 5 min at RT. 500 µl of P1 were used to transduce 20×10^6 Sf9 cells in a T175 flask, filled with 20 ml supplemented medium (Table 18), for virus amplification (P2). The supernatant was harvested after 4-5 days of incubation at 27 °C by centrifugation, using the above stated conditions. The virus stock was supplemented with 2% FCS, sterile filtered, and stored for a maximum of two months at 4 °C.

5.14.5. Protein expression in insect cells

Protein expression in insect cells is generally used for proteins exhibiting posttranslational modifications (PTMs) in combination with high yield expression. Nevertheless, several PTMs differ between mammalian and insect cells such as N-glycosylation (X. Shi & Jarvis, 2007), which is less complex in insect expressed proteins. High Five™ cells, featuring mostly higher protein expression in comparison to Sf9 cells, were used in combination with the Bac-to-Bac® Baculovirus expression system to express the majority of proteins used in this work.

Cells were seeded at a density of 5×10^5 cells/ml in the required volume as suspension culture (5.14.2), followed by their transduction with 0.5% (v/v) recombinant baculovirus P2. The culture was incubated for 3-4 days at 100 rpm and 27 °C, prior to harvesting at 1,000 xg for 10 min and 4 °C. Thereafter, the pellet was used for protein purification.

5.14.6. Production of 2T2 antibody

The T antigen antibody 2T2 was produced in the 2T2 hybridoma cell line, provided by Christopher Buck (NIH, Bethesda, USA). These cells are myeloma cells, fused with sT-immunized murine B lymphocytes, which show infinite cell growth and produce 2T2 antibody. The cell suspension was grown in complete RPMI medium, supplemented with 1x HFCS, in T175 flasks at 37 °C and constantly expanded until reaching 10^8 cells. Cells were centrifuged at 500 xg for 5 min and resuspended in 50 ml 1x hybridoma-SFM for monoclonal antibody selection. After the medium turned yellow, cells were centrifuged, the supernatant aliquoted and supplemented with NaN_3 (1:1000). The aliquots were stored at 4 °C and used for up to 6 months. For subsequent antibody productions, hybridoma cells were frozen as described in 5.14.1 with freezing medium composed of 50% conditioned supernatant, 40% FCS, and 10% DMSO.

5.15. Molecular biology techniques

5.15.1. Polymerase chain reaction

The polymerase chain reaction (PCR) allows DNA amplification and was used for cloning experiments as well as for baculoviral bacmid verification (5.14.4). Specific primers (Table 10), polymerases, PCR mixtures (Table 20 and Table 22) and temperature profiles (Table 21 and Table 23) were used. The Phusion® High-Fidelity PCR Kit, equipped with a 3´→5´exonuclease active polymerase, was utilized for all cloning experiments due to its high precision, whereas the AccuStart™ II SuperMix, comprising a hot-start Taq-Polymerase with a lower accuracy, was used to verify the insertion of the gene cassette of interest into the baculoviral bacmid backbone. The annealing temperature depends on the primer length and content, while the elongation time results from the DNA fragment length. PCR products were generally subjected to agarose gel electrophoresis for analysis (5.15.4).

Table 20 Phusion PCR reaction mix

Component	Amount
Phusion polymerase	1 U
Template DNA	100 ng
Primer forward	10 pmol
Primer reverse	10 pmol
dNTP mix	10 nmol
5x Phusion HF buffer	10 µl
DMSO	3%
DNAse/RNAse freeH$_2$O	ad 50 µl

Table 21 Phusion PCR temperature profile

Step	Time [min]	Temperature [°C]	Cycles
Initiation	00:30	98	1
Denaturation	00:10	98	35
Annealing	00:30	45-72	35
Elongation	00:30/kb	72	35
Termination	05:00	72	1
Storage	∞	4	1

Table 22 PCR reaction mixture for bacmid verification

Component	Amount
AccuStart™ II PCR SuperMix	12.5 µl
Template DNA	100 ng
pUC/M13 forward	10 pmol
pUC/M13 reverse	10 pmol
DNAse/RNAse free H_2O	ad 25 µl

Table 23 PCR temperature profile for bacmid verification

Step	Time [min]	Temperature [°C]	Cycles
Initiation	10:00	93	1
Denaturation	00:45	94	30
Annealing	00:45	55	30
Elongation	00:30/kb	72	30
Termination	07:00	72	1
Storage	∞	4	1

5.15.2. Oligonucleotide annealing

Single-stranded DNA oligonucleotides were annealed to obtain double-stranded DNA oligonucleotides, used for DPI-ELISA (5.16.15), nanoDSF (5.16.14), negative stain TEM (5.18.1), and cryoEM (5.18.3). Two complementary single-stranded oligonucleotides, with a concentration of 100 pmol/µl, were mixed with 5x annealing buffer (Table 4) to the desired concentration of 1 pmol/µl and heated to 95 °C for 5 min. Thereafter, the temperature was successively decreased by 5°C per minute to 25 °C. The annealed oligonucleotides were stored at 4 °C until usage.

5.15.3. Restriction enzyme digest

Restriction endonucleases cleave enzyme specific at primarily palindromic recognition sites within the DNA and produce either sticky (single-stranded DNA overhangs) or blunt (no DNA overhangs) ends. In this work, the method was used to verify the correctness of, for example, newly synthesized plasmids and to prepare the insert and vector during cloning experiments for ligation (5.15.6).

The reaction mixture, which is summarized in Table 24, was incubated at 37 °C for 30 min and stored at 4 °C, prior subjection to agarose gel electrophoresis (5.15.4).

Table 24 Restriction enzyme digest reaction mixture

Component	Amount
DNA	1 µg or entire purified PCR product (5.15.5)
FastDigest® restriction enzyme	1 µl
10x FastDigest® Green buffer	2 µl or 3 µl
DNAse/RNAse free H_2O	ad 20 or 30 µl

5.15.4. Agarose gel electrophoresis

The agarose gel electrophoresis allows horizontal separation of DNA fragments according to their size in an electric field, based on their negative charged sugar-phosphate backbone. Smaller fragments migrate faster in comparison to larger ones, depending on the agarose concentration (0.8-2% (w/v)), the applied voltage (90-100 V), and the size and conformation of the examined DNA. In this work, the method was used to separate and visualize restriction digests (5.15.3) and PCR products (5.15.1). Agarose was boiled in 1x TAE buffer and complemented with 0.005% of the DNA intercalating agent Roti® GelStain. The mixture was polymerized in a gel chamber at RT and covered with 1x TAE buffer. Next to the sample, which was mixed with 6x DNA loading buffer, a size standard (GeneRuler™ 1kb DNA Ladder) was applied onto the gel, which was run at 90-100 V for 50 min. The separated DNA was visualized with UV light using a GelDoc-station.

5.15.5. DNA purification out of agarose gels and other reaction mixtures

DNA was purified via the NucleoSpin® Gel and PCR Clean-up Kit to reduce potential interference of components from preceded reactions, such as restriction enzyme digests (5.15.3), with follow-up reactions. The isolation of DNA out of an agarose gel was carried out with the same kit, following manufacturer´s instruction.

5.15.6. Ligation

Ligation is used to form phosphodiester bonds between complementary digested insert and vector DNA ends. The method was applied during cloning experiments and a molecular ratio of 1:5 between the vector and insert DNA was implemented. All experiments were carried out with the T4-DNA ligase and the reaction mixture shown in Table 25. The sample was incubated overnight at 16 °C and used to transform chemically competent *E. coli* cells (5.13.2).

Table 25 Ligation reaction mixture

Component	Amount
Vector DNA	100 ng
Insert DNA	500 ng
T4-DNA ligase	0.4 U
10x ligase buffer	1 µl
DNAse/RNAse freeH$_2$O	ad 10 µl

5.15.7. Gibson assembly

Gibson assembly is a cloning technique, in which multiple DNA inserts can be combined simultaneously into a vector. This is independent of producing complementary DNA ends by restriction enzymes. Therefore, the inserts require 15-20 nt long overlaps with the adjacent vector/insert, which is accomplished by using 30-40 nt long primers with approximately 20 nt complementary to the insert and the adjacent vector/insert. The PCR-produced inserts are cloned into the vector by using a Gibson assembly reaction mixture. This mixture comprises a 3´-overhang producing 5´-exonuclease, which is a polymerase filling in any gaps, and a ligase, which covalently joins the DNA fragments.

A PCR was carried out to equip the inserts with 5´-overlaps (5.15.1), followed by the analysis and purification of the insert via agarose gel electrophoresis (5.15.4) and PCR clean-up (5.15.5), respectively. The insert was mixed with linearized vector in a 2:1 ratio and complemented with 10 µl NEBuilder HiFi DNA Assembly Master Mix in a total volume of 20 µl. The reaction mixture was incubated for 30 min at 50 °C. Subsequently, 2 µl of the sample were used for *E. coli* transformation (5.13.2).

5.15.8. DNA sequencing

Plasmid sequences were verified by Sanger Sequencing with help of the company Seqlab (Göttingen, Germany). Therefore, 30 pmol primer, which bound close to the sequence of interest, was mixed with 0.9-1.2 µg plasmid in a total volume of 15 µl and sent to Seqlab overnight. The obtained sequence was analyzed via CLC DNA Workbench 8.0.1.

5.15.9. Isolation of plasmid DNA

Plasmid DNA needs to be isolated out of the prokaryotic producer cells to be utilized for further experiments, such as transfection (5.14.3) or PCR (5.15.1). A single colony of plated *E. coli* cells was used to inoculate 2 ml LB medium, supplemented with antibiotics (Table 16), which was incubated overnight at 37 °C while shaking. *E. coli* DH10Bac cells were used for baculovirus bacmid isolation

(5.14.4) and *E. coli* TOP10 cells for all other applications. The PeqGold Plasmid Miniprep Kit was used for small scale isolation out of 2 ml overnight cultures, whereas the DNA Plasmid Purification Kit NucleoBond® PC100 was used for large scale isolation out of 200 ml overnight cultures. Both kits were used according to the manufacturer's instructions. The final DNA concentration was determined using the NanoDrop2000 (5.15.10).

5.15.10. Determination of DNA and protein concentration

The concentration of nucleic acid and proteins was determined by measuring the absorption at 260 and 280 nm, respectively, using the NanoDrop2000. Following the manufacturer´s instructions, 1.5 µl sample was measured against the sample buffer. The sample was considered pure if it exhibited a 260/230 ratio of ≥2.0-2.2 and a 260/280 ratio of ≥1.8 for DNA and ≤0.6 for protein samples.

The Qubit® 2.0 Fluorometer provides an alternative method for nucleic acids and protein concentration measurements. It exhibits a higher accuracy, in comparison to NanoDrop2000 measurements, and was used to determine protein concentrations with help of the Qubit® Protein Assay Kit, following the manufacturer's instructions.

5.16. Protein biochemical techniques

5.16.1. Protein concentration via filtration

Protein concentration via filtration is facilitated by volume reduction with help of Amicon® Ultra centrifugal filter units (0.5 ml, 4 ml and 15 ml) with a molecular weight cut-off (MWCO) of 10-100K. The filter was equilibrated with protein buffer, followed by sample filtration steps at 2,500 xg (4 ml and 15 ml filter units) or 14,000 xg (0.5 ml filter unit) for 5 min at 4 °C. In between centrifugation steps, the sample was carefully inverted to avoid spatial concentration differences. The protein solution could be concentrated to a final volume of 50 µl.

5.16.2. Sodium dodecyl sulfate-polyacrylamide gel electrophoresis

Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) is a method to separate denaturized proteins according to their molecular weight in an electrical field. The anionic detergent SDS denaturizes the proteins and applies a negative charge to the sample. This enables all proteins to migrate towards the anode at a velocity proportional to their size. The protein mixture was supplemented with 4x SDS loading buffer and incubated for 5 min at 95 °C, prior to loading onto a discontinuous 8-15% polyacrylamide gel. The composition of the gel, comprising a stacking gel to focus and a resolving gel to separate the sample, as well as the composition of all buffers are summarized in Table 4. For size estimation, a protein size standard was added. The electrophoresis

was carried out in SDS running buffer at 180 V for 40-70 min, depending on the protein size and gel concentration.

5.16.3. Coomassie staining

Proteins, which are separated via SDS-PAGE, can be visualized by Coomassie staining. The resolving gel was incubated with Coomassie staining solution for 15 min while shaking. Thereafter, the gel was destained with destaining solution, which was exchanged repeatedly until the protein bands appeared in a pronounced contrast to the gel. The composition of the used solutions is summarized in Table 4.

5.16.4. Western Blot

Western Blot (WB) facilitates specific immunodetection of proteins. The sample, separated by SDS-PAGE, was blotted via a semi-dry blotting system onto a polyvinylidene difluoride (PVDF) membrane. Therefore, the membrane was activated in 100% methanol, and Whatman paper was soaked in WB buffer A-C. The acrylamide gel, the PVDF membrane, and the Whatman paper were arranged according to Figure 10 and run at 400 mA for 40 min. The blotted membrane was blocked for 30 min in WB blocking solution to inhibit unspecific antibody binding, followed by the addition of primary antibody. The antibody was diluted in WB blocking solution and incubated overnight at 4 °C. Thereafter, the membrane was washed three times with PBS-T or TBS-T for 5 min. The secondary antibody, conjugated to horseradish peroxidase (HPR) and diluted in WB blocking solution, was added onto the membrane, and incubated for 1 h at room temperature. The antibody dilution specifications can be found in Table 11. The washing step was repeated three times for 10 min to remove unbound antibody, followed by chemiluminescence imaging. Depending on the protein concentration, femto LUCENT™ PLUS-HPR or Western Lightning® Plus-ECL reagent was used for signal development according to the manufacturer's instructions. These solutions allowed the oxidation of luminol in the presence of peroxide by the HPR, resulting in light emission, which was detected via X-ray films and the Curix 60 or Image Quant™ 800. All used buffers are summarized in (Table 4).
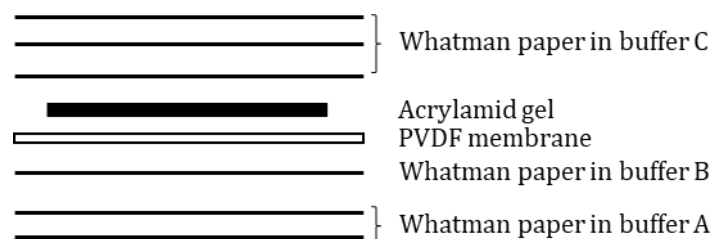
Figure 10: Scheme of Western Blot assembly

5.16.5. Cell lysis

Cell lysis is used to disrupt cell membranes, which is a prerequisite for the extraction of intracellular proteins. Different cell lysis methods exist, such as French press, freeze-thawing, or high-pressure homogenization. These methods represent physical lysis methods, while detergents are used in the chemical lysis. In this work, the physical method ultrasonication, in combination with detergents, was predominantly used to disrupt insect cell membranes, followed by protein purification. Cells were harvested at 1,000 xg for 10 min at 4 °C and resuspended in lysis buffer, containing 0.1% Triton X-100, cOmplete™ Protease Inhibitor Cocktail, and 1 mM DTT. The sample was incubated for 15 min-1.5 h at 4 °C and sonified. Therefore, an amplitude of 30% was applied six times for 10 sec, with an interruption of 10 sec between each cycle. The cell debris was separated by centrifugation at 20,000 xg for 15 min at 4 °C, followed by ultrafiltration of the soluble fraction. The supernatant was thereafter subjected to chromatography.

5.16.6. Affinity chromatography

Affinity chromatography is used to purify recombinant proteins by utilizing specific binding interactions between an immobilized ligand and its binding partner, called affinity tag. The hexa-histidin (6xHis), the strep tag (StrepII) and the maltose binding protein (MBP) were used for the fusion protein production. Their advantages, eluents, and the suitable resin are summarized in Table 26.

Table 26: Overview of affinity tags used in this work

Affinity tag	Resin	Eluent	Size [aa]	Advantage
His	Nickel-nitrilotriacetic acid (NiNTA) agarose	250 mM imidazole	6	Small size
StrepII	StrepTactin Sepharose	50 mM biotin	8	Small size and high specificity
MBP	Amylose agarose (small scale purification) Dextrin Sepharose (Large scale purification)	10 mM maltose	370	High specificity and enhancement of recombinant protein solubility

Small scale purification

Cell culture volumes of 10-100 ml were used for gravity flow affinity chromatography with a resin to cell culture ratio of 1:100. The resin was transferred into a polypropylene column (1 ml or 5 ml), equipped with a filter and a valve, and washed with 5 column volume (CV) H_2O, followed by the equilibration with 5 CV lysis buffer without protease inhibitors. The cell lysate (5.16.5) was transferred to the column and incubated with the resin for 1-2 h at 4 °C, while shaking. The bound

recombinant protein was washed with 10 CV wash buffer and eluted with 1 CV per fraction of elution buffer. All steps were carried out on ice.

Large scale purification

Cell culture volumes of ≥250 ml were used for automized affinity chromatography using the Äkta Pure or Purifier 10 and prepacked columns with a CV of 1 ml and 5 ml. The column was washed with 5 CV H_2O, followed by equilibration with 5 CV lysis buffer without protease inhibitors at a flow rate of 0.7 ml/min (1 ml CV column) or 4 ml/min (5 CV column). The cell lysate was applied onto the column at a flow rate of 0.2 ml/min, followed by a 3 CV wash step with wash buffer and a flow rate of 0.5 ml/min (1 CV column) or 2 ml/min (5 CV column). A buffer exchange was carried out with 4 CV at a flow rate of 0.1 ml/min. Thereafter, the recombinant protein was eluted with 5 CV elution buffer and a flow rate of 0.1 ml/min (1 CV column) or 0.25 ml/min (5 CV column). The eluate was spectroscopically detected at a wavelength of 280 nm and 260 nm. The purification was carried out with degassed and sterile filtered liquids and at 4 °C. The collected fractions were analyzed via Coomassie staining (5.16.3) and/or WB (5.16.4).

5.16.7. 2T2 immunoaffinity chromatography

Immunoaffinity chromatography is used to purify proteins via immobilized antibodies. Unconcentrated, self-made MCPyV T antigen antibody 2T2 (5.14.6) was coupled to 4% cross-linked agarose-bound protein G in a ratio of 100:1 on a rotary shaker overnight at 4 °C. The resin was washed with 100 CV EB4 (Table 5), centrifuged at 500 xg for 5 min, and resuspended in 4 CV EB4. LT, pre-purified by affinity chromatography, was incubated with the resin at a ratio of approximately 2:1 (μg:CV) on a rotary shaker for 2 h at 4 °C. The immobilized LT was washed with 60 CV EB4, centrifuged, and analyzed via WB.

5.16.8. Ion exchange chromatography

Ion exchange chromatography is a protein purification method utilizing electrostatic interactions between charged proteins and the oppositely charged immobile phase. Elution is facilitated by increasing the ionic strength. The sample was loaded onto the strong anion exchange column HiTrap Q FF (5 ml CV) with 2 CV in low salt buffer (100 mM). Thereafter, the bound protein was washed with 8 CV low salt buffer, followed by a gradient elution with 10 CV with high salt buffer (1 M). The eluate was monitored spectroscopically at 280 nm and analyzed via SDS-PAGE (5.16.3). Elution was performed at a flow rate of 2.5 ml/min and the chromatography was carried out at 4 °C with an Äkta Purifier 10.

5.16.9. Adsorption chromatography

Adsorption chromatography facilitates protein fractionation upon the adsorption of proteins onto a stationary phase. In this work, the negatively charged linear polysaccharide heparin was used as an adsorbent, which binds, among others, specifically DNA-binding proteins by mimicking the polyanionic structure of nucleic acid. The sample was loaded onto a HiTrap Heparin HP (1ml CV) and washed with 5 CV binding buffer (300 mM NaCl), followed by a gradient elution over 19 CV with eluent (1 M NaCl). The elution was performed at a flow rate of 0.5 ml/min and monitored spectroscopically at 280 nm and 260 nm. The chromatography was carried out at 4 °C with an Äkta Purifier 10.

5.16.10. Size exclusion chromatography

Size exclusion chromatography (SEC) fractionates proteins according to their hydrodynamic radius (R_H) by passing a porous matrix. Thereby, small proteins exhibit longer retention times since they penetrate the matrix deeper, in comparison to larger proteins. The Äkta system Pure and Purifier 10 were used together with the column Superose™ 6 Increase (24 ml CV). This column exhibits an exclusion limit of 4×10^7 kDa and a separation range of 5-5,000 kDa, making it ideal for the separation of large protein complexes. The sample, with a final volume of 0.5 ml, was centrifuged at 17,000 xg for 10 min at 4 °C, prior application. Gel filtrations were carried out at a flow rate of 0.25 ml/min at 4 °C and with degassed and sterile filtered buffers. Protein elution and nucleic acid contamination was spectroscopically monitored by measuring the absorbance at 280 nm and 260 nm. The collected fractions were analyzed via Coomassie staining (5.16.3) and/or WB (5.16.4).

5.16.11. Proteolytic cleavage of affinity tags

Proteolytic cleavage is used to remove affinity tags from the recombinant protein via a specific cleavage site. In this work, the recombinant TEV protease with its recognition sequence ENLYFQS was used according to the manufacturer's instructions. Therefore, the affinity purified protein of interest was concentrated with an Amicon® Ultra centrifugal filter unit (5.16.1) to approximately 300 µl. The concentrated protein was supplemented with TEV protease and manufacturer supplied buffer to a concentration of 8 U/20 µg recombinant protein. The reaction was carried out in a final volume of 500 µl overnight at 4 °C, while shaking. Thereafter, the sample was applied onto the SEC column to remove the affinity tags and the TEV protease.

5.16.12. Dialysis

Dialysis facilitates buffer exchange and the removal of small molecular weight compounds from protein solutions by passive and selective diffusion through a cellulose membrane. The sample was transferred to a dialysis cassette with a MWCO of 20K and incubated in the appropriate buffer at a ratio of 1:100 (v/v) overnight at 4 °C, while stirring.

5.16.13. Thermal stability assay

The thermal stability assay is used to assess the thermal stability of proteins as a function of different buffers, additives, and pH. The fluorescent dye Sypro orange, emitting light at a wavelength of 570 nm by binding to hydrophobic protein patches, was used to detect protein unfolding during incremental temperature increase. In this work, an optimized assay for large protein complexes, called ProteoPlex (Chari et al., 2015), was implemented in cooperation with the research group of Dr. Ashwin Chari, located at the Max Planck Institute for Biophysical Chemistry in Göttingen, Germany. Thereby, parameters such as the enthalpy, entropy, and cooperativity of unfolding were included in the analysis to consider multiple protein unfolding states in the final melting curve. The assay was carried out in a 96-well format with a total volume of 20 µl/reaction. One well contained 2 µl Sypro orange, 16 µl protein solution (0.2 mg/ml), and 2 µl of the additive under investigation (1M) (Table S4). After mixing, the plate was transferred to a real-time PCR machine and equilibrated to 20 °C. A successive heating of 1 °C steps to a final temperature of 100 °C, with equilibration times of 30 sec, secured a controlled unfolding transition. The fluorescence was measured after every equilibration cycle at 570 nm. The underlying analysis is described in detail in the above stated publication.

5.16.14. Nano differential scanning fluorimetry

Nano differential scanning fluorimetry (nanoDSF) is applied to analyze the thermal stability of proteins in the presence of various buffers and additives by detecting the intrinsic fluorescence of tryptophane and tyrosine as a function of temperature. 10 µl protein solution, with a concentration of 0.1 mg/ml, was applied to a capillary and placed into a NanoTemper Prometheus NT.48. Temperature was successively increased from 20 °C-95 °C in 1 °C steps per minute, while protein unfolding was measured at 330 nm and 350 nm. The first deviation of the 330/350 nm ratio was normalized, depicting the melting temperature as inflection point.

5.16.15. DNA-protein interaction enzyme-linked immunosorbent assay

The DNA-protein interaction enzyme-linked immunosorbent assay (DPI-ELISA) is used to detect direct and specific protein-DNA interaction. 1 pmol of biotinylated and annealed (5.15.2) oligonucleotides were immobilized for 1 h at RT onto an avidin-coated 96-well plate, which was fixed with fixing buffer (Table 4) overnight at 4 °C. The plate was washed three times with PBS-T, blocked

with 5% milk in PBS-T for 30 min at RT, and incubated with 1 μg purified LT for 1 h at RT. Following 3 washing steps, 100 μl/well of the LT-antibody Cm2B4 (Table 11) were incubated for 2 h, while shaking. Wells were washed three times, prior to the incubation of 100 μl/well of the secondary antibody (Table 11) for 1 h, while shaking. The wells were washed 6 times and 100 μl/well of ABTS substrate at RT was added. After 30 min incubation under light exclusion, DNA-protein interaction was calorimetrically detected at 405 nm with the plate reader Infinite M2000.

5.16.16. Immunofluorescence staining

Immunofluorescence (IF) staining allows the detection and localization of macromolecules inside fixed cells via fluorophore conjugated antibodies. A 24-well plate format, equipped with sterile glass coverslips, was coated with 2% gelatine solution for 5 min and washed with PBS. HEK293 cells were seeded at a concentration of 10^5 cells/well and transfected with PEI one day post seeding (5.14.3). The cells were incubated 2-4 days, washed with cold PBS, and fixed with 4% (v/v) paraformaldehyde (PFA) in PBS for 30 min at 4 °C. All subsequent steps were implemented at room temperature. The sample was washed twice with PBS and incubated with permeabilization buffer for 10 min to allow antibodies to penetrate the cell. The wash step was repeated, followed by incubation with blocking buffer for 30 min to inhibit unspecific antibody binding. The coverslips were moistened with primary antibody, which was diluted in blocking solution for 1.5 h in a humid chamber to prevent evaporation. Unbound antibody was removed by washing with PBS three times and secondary antibody was added to the sample for 1 h under light exclusion. All utilized buffers as well as antibodies and their dilutions are summarized in Table 4 and Table 11, respectively. After a final PBS wash step, the coverslips were mounted onto microscope slides with Vectashield, containing the cell nucleus dye DAPI. The sample was stored at 4 °C for microscopic analysis using the confocal laser scanning microscope Leica TCS SP5 and a 63x oil immersion objective.

5.16.17. Dynamic light scattering

Dynamic light scattering (DLS) is a non-invasive method to assess the R_H of macromolecules by measuring the Brownian motion. The scattered light of laser-illuminated particles is detected over time, and the resulting intensity fluctuation can be correlated to the particle size, considering viscosity and temperature. The sample, with a concentration of 0.2 mg/ml, was centrifuged at 17,000 xg for 10 min at 4 °C, and 13 μl were placed into a quartz glass SUPRASIL® cuvette. The measurement was carried out on a Xtal Concepts SpectroSize 300 with 10 sec acquisition time and 20 iterations at 20 °C.

5.17. *In cellulo* crystallization

The formation of endogenous protein crystals in cells is a long-known phenomenon first described in the Brazilian nut seed in 1855 (Hartig, 1855). They hold a plethora of functions such as protein storage (Papassideri, 2007), protection (Coulibaly et al., 2007), and solid-state catalysis (Veenhuis et al., 2003; Tsukada et al., 1966). Besides many benefits of native crystalline proteins, non-native *in cellulo* crystals can be the causative agent of diseases such as congenital cataracts (Pande et al., 2001). Given the vast amount of native and non-native *in cellulo* protein crystals, it is evident that scientists aim to synthesize heterologously expressed *in cellulo* protein crystals for structure analysis.

The elucidation of the atomic structure of proteins applying *in vitro* crystallization is laborious and often unsuccessful, as no comprehensive theory of crystallization conditions exists. They need to be determined empirically for each protein anew. In addition, proteins must be expressed in vast amounts, isolated, and purified before crystallization occurs. The necessity of high protein concentration and high purity often leads to crystallization failure, as the protein solubility restricts its concentration. To increase protein solubility, buffer composition, including pH, ionic strength, and additives, such as reducing or stabilizing agents, must be screened and optimized. Using *in cellulo* crystals for structure analysis holds the potential to circumvent these challenges.

Polyhedrin, expressed in the cytoplasm of insect cells, was the first atomic structure elucidated from *in cellulo*-derived crystals (Coulibaly et al., 2007). Long time, *in cellulo*-derived crystals were difficult to use for structure elucidation because their size ranged between 0.75 - 3500 μm^3 (Schönherr et al., 2018). Applying conventional X-ray crystallography, they were too small to provide sufficient diffraction signals within the tolerable dose limit of >30 MGy at cryo conditions and 1 MGy at room temperature (Southworth-Davies et al., 2007). Only the establishment of high-brilliance synchrotron radiation, enabling microfocus beams, the advent of X-ray free-electron laser (XFEL), and the development of serial femtosecond crystallography (SFX) applying the "diffraction-before-destruction" principle (Chapman et al., 2011; Barty et al., 2012), gave rise to the analysis of *in cellulo* grown microcrystals. *In cellulo* crystals could be observed in expression systems with bacterial (Sawaya et al., 2014), yeast (Vonck & van Bruggen, 1992), insect (Schönherr et al., 2015), and mammalian origin (Hasegawa et al., 2011). Likewise, protein crystals emerged in various cellular organelles, such as the peroxisome (Schönherr et al., 2015), the nucleus (Tsutsui et al., 2015) or the ER (Hasegawa et al., 2014). However, despite the versatile occurrence of *in cellulo* crystals, the physicochemical specifications and the molecular mechanism driving *in cellulo* crystallization remain elusive. A promising hypothesis to enhance intracellular protein crystallization suggests the co-expression of the protein of interest with a proteinogenic crystallization matrix (Baskaran et al., 2015). The concept of crystalline molecular flasks has been used for a long time to analyze solid-state

chemical reactions and is reviewed by Inokuma and colleagues (Inokuma et al., 2011). To apply this concept to protein crystallography, the catalytic site of the metazoan-specific serine/threonine p21-activated kinase 4 (PAK4cat) and its endogenous inhibitor Inka1 have been co-crystallized and used as a potential proteinogenic crystalline molecular flask (Baskaran et al., 2015). Upon their overexpression in various mammalian cell lines, such as HEK293, U2OS, and COS-7, PAK4cat and Inka1 build intracellular, up to 50 μm, needle-like crystals with a hollow tube spanning 80 Å. Their molecular structure could be solved to a resolution of 2.95 Å by transferring the fully integrated cells, harboring the protein crystals, directly to synchrotron radiation. PAK4:Inka1 was the first human protein structure to be solved intracellularly. The void volume of the crystal is spacious enough to accommodate assorted proteins with a size of maximum 30 kDa, such as the green fluorescent protein (GFP), the cell division control protein 42 homolog, or the actin modulating cofilin, without impairing *in cellulo* crystal growth. As proof of principle, PAK4cat and fusion proteins of Inka1, or its N-terminally truncated version iBox, were co-expressed with the above stated proteins. Figure 11 shows a schematic representation of the potential packing arrangements of GFP-Inka1:PAK4cat. The cross-section of the crystal is built out of twelve PAK4cat units providing all crystal contacts, whereas Inka1, located close to the PAK4 C-terminus, is exposed to the solvent channel without any crystal contacts. It features a perfect target to clonally engineer proteins of interest to its C- or N-terminus. The protein of interest aligns inside the crystal cavity, which may enhance its arrangement in a single state. It remains uncovered whether the proposed concept of a proteinogenic crystalline molecular flask can be applied for protein structure analysis since the published *in cellulo* fusion crystals of Baskaran and colleagues were not analyzed by synchrotron radiation so far.
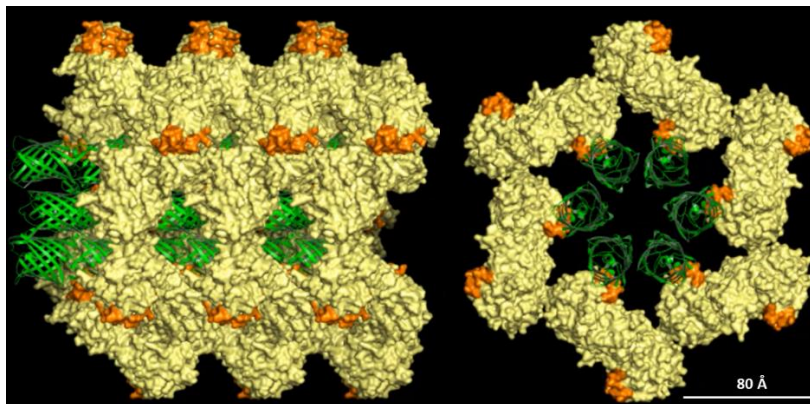
Figure 11: Modeling of GFP into the crystal lattice of Inka1:PAK4cat. GFP (green), fused to Inka1 (orange), is aligned inside the central cavity of *in cellulo* crystals which are produced by co-expressing PAK4cat (yellow) and Inka1 in mammalian cells. The crystal cross-section is built out of twelve PAK4cat and Inka1 protein subunits whereas Inka1 does not contribute to crystal contacts. The central cavity spans a diameter of 80 Å and theoretically could incorporate six GFP molecules. The accommodation of a protein of interest inside the crystalline matrix hypothetically enhances its crystallization propensity making it suitable for structure analysis. Modified from (Baskaran et al., 2015).

Intracellularly synthesized protein crystals can either be isolated from the expression system prior to analysis or subjected directly to X-ray radiation with the cellular reservoir. It depends on their sensitivity to environmental changes and ability to withstand harsh buffer conditions (Redecke et al., 2013; Gati et al., 2014; Boudes et al., 2016). Taken together, *in cellulo* crystallization constitutes a promising alternative to *in vitro* crystallization if universally applicable *in cellulo* crystallization protocols will be developed in the future.

5.17.1. *In cellulo* crystal production in insect cells

The production of *in cellulo* crystals in insect cells was predominantly established for large scale crystal production, used for crystal purification. The cultivation of insect cells is cost-effective and can be realized as suspension and adherent cell culture, which is advantageous over HEK293 and COS7 cells. Test expressions, comprising a volume of 0.5-4 ml, were carried out with Sf9 cells at 27 °C. Their enhanced adherence, compared to High Five™ cells, simplified the monitoring of crystal growth. Large scale expressions were performed in High Five™ cells as suspension cell culture, comprising a volume of 20-100 ml, at 100 rpm and 4 °C. The crystalline matrix was provided by the expression of the protein constructs PAK4 and its inhibitor Inka1. To facilitate co-crystallization of the crystalline matrix and the protein of interest, the constructs were co-transduced with 0.5% (v/v) of each recombinant baculovirus P2, 1 h post seeding. Further analysis via WB was realized 2-10 days post transduction.

5.17.2. In cellulo crystal production in mammalian cells

HEK293 and COS7 cells were mainly used for *in cellulo* crystallization since they were most successful in crystal production, compared to insect cells, and provide authentic PTMs of mammalian and viral proteins. Cells were seeded and co-transfected (5.14.3) with the protein of interest and the crystalline matrix providing PAK4 and Inka1 constructs. The transfected cells were incubated for 3-6 days at 37 °C, prior analysis via WB (5.16.4), IF (5.16.16), synchrotron X-ray radiation (5.19.1), and transmission electron microscopy (TEM) of ultra-thin sections of resin embedded cells (5.18.2).

5.18. Electron microscopy

5.18.1. Negative stain transmission electron microscopy

Negative stain TEM is used to visualize structural details of macromolecules or viruses, which are embedded in opaque heavy metal salts. The sample is adsorbed on an electron transparent sample support and can be resolved to up to 15 Å. A 300-mesh copper grid, coated with a continuous carbon layer (Table 2), was glow-discharged for 30 sec at 25 mA with a Quorum GloQube glow discharger to obtain a negatively charged surface. The grid was next incubated with 4 µl purified protein solution

with a concentration of 0.12 mg/ml for 30 sec, followed by blotting with Whatman paper to enhance sample adsorption and to remove excess liquid. The sample was washed twice with H_2O, prior staining with a 2% uranyl acetate solution. To secure exhaustive sample staining, the protein was incubated a second time with the heavy metal salt for 30 sec. Excess liquid was removed by Whatman paper between each step and the entire staining was conducted at room temperature. The sample was air-dried for 15 min and analyzed with the electron microscope Talos L120C (FEI/Fischer Scientific), equipped with a LaB6 cathode and a CETA camera and operated at 120 kV. For single particle analysis (SPA) of negatively stained proteins, approximately 100 micrographs were recorded at a nominal magnification of 92k x and a pixel size of 1.53 Å with a defocus range of -0.3 to -4.7 μm, which were analyzed with Relion 3.1.

5.18.2. Transmission electron microscopy of ultra-thin sections of resin-embedded cells

TEM of ultra-thin sections of resin-embedded cells allows to analyze subcellular structures within their native environment. The restriction of low electron penetration depth is overcome by slicing resin embedded cells into ultra-thin sections by ultramicrotomy. In this work, the method was used to analyze the crystalline structure of *in cellulo* protein crystals. HEK293 cells were grown in μ-Dish 35 mM ibiTreat and transfected with the respective plasmids (5.14.3). 3 d.p.t., the cells were fixed with 2% (v/v) formaldehyde and 2.5% (v/v) glutaraldehyde in PBS for 60 min, followed by five washing steps in PBS for 2 min on ice. The specimen was stained with 1% (w/v) osmium tetroxide in PBS for 30 min and washed twice in PBS and H_2O for 10 min. A second staining step with 1% (w/v) uranyl acetate for 30 min was carried out, followed by two washing steps in H_2O for 10 min. The sample was dehydrated with 50%, 70%, 90% (v/v), and 100% EtOH, each step for 10 min with two additional dehydration steps in 100% EtOH for 10 min. Thereafter, the cells were embedded in 50% (v/v) Epon 812 epoxy resin (Table 2) in EtOH for 30 min. Embedding was repeated with 70% (v/v) Epon 812 epoxy resin in EtOH for 90 min and with 100% Epon 812 epoxy resin overnight, for 3 h, 5 h, overnight, and again for 5 h. Polymerization was conducted in 100% Epon 812 epoxy resin at 60 °C overnight and areas of interest were recovered with a 6 mM disc punch. A four-sided, flat-top pyramid with approximately 30° angles was formed with a scalpel and transferred onto an ultramicrotome to remove the ibidi polymer culture dish layer and to slice 50 nm sections. The slices were placed on copper grides (200 mesh), post-contrasted with saturated uranyl acetated in 70% (v/v) EtOH, washed extensively with H_2O, dried on filter paper, and imaged on the transmission electron microscope CM120 (Philips/FEI). The microscope, equipped with a LaB6 cathode, was operated at 80 kV. Micrographs were recorded using a MultiScan 794 CCD camera.

5.18.3. Transmission electron cryo-microscopy single particle analysis

CryoEM SPA is used to gain high-resolution information of macromolecular structures. Electrons, emitted by a field emission gun, will be highly accelerated in a vacuum to interact with the specimen. The elastically scattered electrons are detected on a direct electron detector, producing 2D projection images of the sample, which are used to construct a 3D structure of the sample. The imaging is carried out under cryogenic conditions by imaging frozen hydrated specimens, which are cooled by liquid nitrogen. Since many biological samples are radiation sensitive, the imaging has to be carried out at low electron dose (L. A. Baker & Rubinstein, 2010). This results in poor signal-to-noise ratios in the corresponding micrographs, leading to reduced contrast. By averaging large numbers of identical projection images, the challenge of poor signal-to-noise ratio can be counteracted. Specimens were previously vitrified by plunge freezing the sample in an ethane-propane mixture using a Vitrobot Mark IV (FEI/Thermo Fischer Scientific). 4 μl of purified LT, with a protein concentration of 0.8 mg/ml, were added on a C-flat™ holey carbon copper grid (Protochips) and plunge-frozen using blotting conditions of 100% humidity, 4 sec blot time, and a blot force of 0. Thereafter, the sample was constantly kept at liquid nitrogen temperature to prevent ice crystal formation and devitrification of the sample. The data set was recorded on a Talos Arctica 200kV electron microscope, equipped with a Falcon 3EC direct detector (FEI/Thermo Fischer Scientific). Images were recorded with a total dose of 42 e-/$Å^2$, at a magnification of 150k x with the EPU software (Table 15), using a defocus range of -0.6 to -2.9 μm (Table 27). The SPA was carried out with Relion 3.1 (Scheres, 2012)(Figure 12). 37 frames were motion corrected to form one micrograph, followed by CTF correction, using CTFFIND4.1 (Rohou & Grigorieff, 2015). 21 micrographs were selected to define parameters for automated reference-free auto-picking based on a Laplacian-of-Gaussian (LoG) filter with a diameter of 200 x 100 Å. The parameters were applied onto the entire dataset, and the picked particles were classified in 2D, according to their shape and structural features, applying a soft circular mask diameter of 160 Å. By aligning particles with the same orientation and conformation, the signal-to-noise ratio was drastically increased and randomly distributed noise could be eliminated. Broken particles and ice crystals were discarded, while the remaining particles were used for a subsequent 2D classification. This procedure was iteratively performed until the quality of the alignments did not improve anymore. These particles (18,442 particles) were utilized for an initial 3D model, out of which a 3D reference with no applied symmetry was built. The 3D reference was utilized for 3D alignment and the reconstruction of five 3D classes. The class with the highest particle distribution and the best alignment parameters (11,981 particles) was selected for auto-refinement to improve the resolution. For gold-standard Fourier Shell Correlation (FSC) calculations of resolution estimates and to avoid overfitting, the dataset was split into two independent subsets and their

consistency as a function of special frequency was compared (van Heel & Schatz, 2005). The final model was not submitted to postprocessing since overfitting was observed. Similarly, particle polishing was not applied.
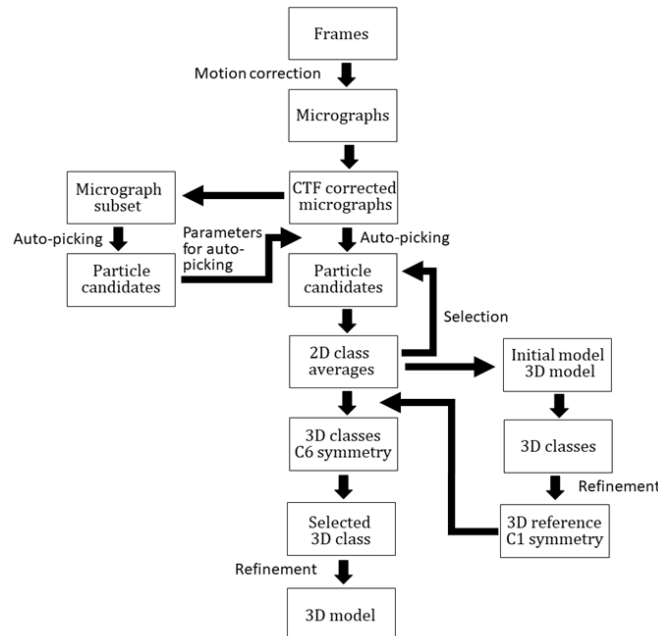
Figure 12: The Relion 3.1 workflow applied for the cryoEM SPA.

5.19. Synchrotron radiation-based X-ray techniques

5.19.1. X-ray scattering of *in cellulo* protein crystals

Protein X-ray crystallography is used to elucidate the atomic structure of protein crystals by exposing them to X-ray radiation. The monochromatic, focused, and coherent X-ray beam interacts with electrons of the protein atoms. Since the interaction of X-rays with electrons is low, a periodic lattice is vital to amplify the signal. X-rays are scattered by the crystal plane with a reflection angle θ. If the path length difference, dependent on the crystal plane distance and the reflection angle, is equal to an integer multiple of the wavelength, constructive interference occurs. This is described by the Bragg´s law and is a prerequisite to detect scattering events (Bragg & Bragg, 1913). The elastically scattered photons are recorded as diffraction patterns in reciprocal space and used to calculate the electron density map in real space by its Fourier transformation. Thereby, phases and structure factors are important, of which the latter can be obtained by the recorded scattering intensity, depending on the scattering atom. However, it is difficult to calculate the phase of the incoming beam from the diffraction data, known as the phase problem (R. W. Harrison, 1993). The phase problem can be solved, for example, by isomorphous replacement (Green et al., 1954) or molecular replacement (Rossmann & Blow, 1962). Despite the long tradition of X-ray crystallography, sample crystallization

is still a "trial and error" procedure and reflects the main bottle neck of this method. As a consequence, *in cellulo* grown crystals (5.17) were analyzed via X-ray radiation in this work. The crystals, grown in T25 flasks-cultured HEK293 and COS7 cells, were trypsinized 4 d.p.t.. 2 µl crystal-cell suspension were mixed with 4 µl 37.5 % polyethylene glycol (PEG) 400 and mounted on a nylon micro mesh, comprising 400 10 x 10 µm pores. Thereafter, the crystals were examined via the synchrotron X-ray source of PETRAIII, EMBL P14, Hamburg. The analysis was performed under cryo-conditions with a microbeam (5 x 10 µm) with 12.7 photons/sec and a wavelength of 0.98 Å.

6. Results

6.1. *In cellulo* crystallization of MCPyV sT and LT

Solving the structure of MCPyV oncoproteins is of great interest as it constitutes an important step towards developing specific antiviral treatment against MCC and discloses unknown structure-function relationships helping to reveal how MCPyV T antigens contribute to the establishment of MCC. Hitherto, no full-length LT structure is known and only single domains of related PyV LTs, such as the DnaJ, OBD, and helicase have been solved (Meinke et al., 2014; D. Li et al., 2003; H.-Y. Kim et al., 2001). For MCPyV LT, only the OBD could be crystallized, which shows the intricacy of solving its full-length structure (C. Harrison et al., 2011). The LT MURI/II domains, which constitute one-third of the protein and are predicted to be unstructured, pose a major challenge in solving their structure.

The truncated version of LT (tLT244), comprising the first 244 amino acids without the C-terminal helicase and OBD, derived from the MCC cell line 12 (MCCL12), was tried to be crystallized in a previous dissertation work (Borchert, 2013). The protein was expressed in *E. coli,* purified via affinity chromatography (NiNTA) and SEC, followed by *in vitro* crystallizing experiments. Unfortunately, the approach failed due to the need for a monodispersed and sufficiently concentrated LT solution. Similarly, there is no full-length structure available for MCPyV sT. In this work, *in cellulo* crystallization (5.17) was applied to solve the structure of the MCPyV T antigens, which is an essential prerequisite to develop antiviral drug targets.

6.1.1. MCPyV T antigen constructs screened for *in cellulo* crystallization

The concept of a proteinogenic crystalline molecular flask was applied to increase the propensity of sT and LT to crystallize (Baskaran et al., 2015) (5.17). The catalytic site of PAK4 and its endogenous inhibitor Inka1 spontaneously form intracellular protein crystals, which can harbor proteins of interest in their central solvent channel to stabilize them in a single state. The size restriction of the cavity of 80 Å may lead to an erroneous alignment of the T antigens, impairing the crystal packaging. Therefore, Baskaran and colleagues proposed an upper molecular weight limit for proteins residing inside the channel of 30 kDa, which was applied to design the T antigen constructs (Table 8 and Table 9). The MCCL12 LT variant tLT244 (27 kDa), sT (22 kDa), and 122tsT (8 kDa) were fused to Inka1 (31 kDa) and co-expressed with PAK4cat (35 kDa) (Figure 13A). The possibility of inadequate crystal packaging arising from the protein size was further reduced by designing the protein 122tsT. ST was N-terminally truncated, comprising the amino acids 122-186. With the tool Jpred 4, this protein part was predicted to harbor mainly alpha helices and beta sheets, which provide protein stability.

Additionally, the N-terminal DnaJ domain, shared with LT, and the LSD domain, flanked by unstructured regions, were omitted.

An approach to increase the solvent channel capacity was the substitution of Inka1, protruding into the cavity, with its truncated version iBOX (Figure 11). This peptide comprises only the C-terminal PAK4cat interacting site, residing in the region of the Inka1 amino acids 166-203 with a molecular weight of 4 kDa. The HA, His, and FLAG tag, fused C-terminally or N-terminally to the protein constructs (Figure 13A), were used to validate the protein expression by WB (Figure 13B and C) and IF (Figure 14). Next to the co-expression of Inka1 fusion proteins with PAK4cat, T antigens were expressed as complete fusion proteins, coupled to iBox and PAK4cat (Figure 13A). It was anticipated to increase crystal formation by providing all crystal components on one construct.

Next to the principle of the proteinogenic crystalline molecular flask, crystallization propensity may be enhanced by expressing the T antigens in different cellular compartments, thereby increasing their local concentration, or exploiting different cellular environments such as pH, protein composition or a proteinogenic crystallization starter such as crystalline oxidases in peroxisomes (Veenhuis et al., 2003). For example, intracellular crystals of exogenously expressed firefly luciferase could be observed in the peroxisome of Sf9 cells (Schönherr et al., 2015), and crystals of cathepsin B were built in the ER of the same cell system (Koopmann et al., 2012). Various N-terminal LT truncations, comprising the OBD, zinc finger, and helicase (tLT328), the zinc finger and helicase (tLT430), or only the helicase (tLT544) as well as sT and 122tsT, were sorted in different cell compartments. The constructs were C-terminally tagged with the peroxisomal target sequence serine-lysine-leucine (SKL) or fused to the ER signal peptide lysine-aspartic acid-glutamic acid-leucine (KDEL), accompanied with an N-terminal ER signal sequence of 22 amino acids. Unmodified tLT244 and sT were expressed to screen crystal formation in the nucleus as their nuclear localization was published in different human cell lines (Borchert et al., 2014; Tsang et al., 2016).

All constructs expressed in mammalian cells were cloned in the pJX vector under the control of a CMV promoter (Table 8). In contrast, proteins expressed in insect cells were cloned into the pFastBac1 vector, harboring the polyhedrin promoter (Table 8). A tabular overview of all engineered constructs is given in Table S 2.

6.1.2. Co-expression of a crystalline matrix and MCPyV oncoprotein variants result in *in cellulo* crystals in mammalian cells

The protein constructs used for the proteinogenic crystalline molecular flask approach were either transfected in a 1:1 ratio into HEK293 cells or COS7 monkey kidney fibroblasts (5.14.3) or transduced via recombinant baculoviruses into Sf9 insect cells by using 0.5% (v/v) of each virus stock (5.14.4). COS7 cells were utilized in this work since the proof of principle to exploit Inka1:PAK4cat as crystalline matrix was performed in these cells (Baskaran et al., 2015). Next to COS7, HEK293 cells were chosen to circumvent the presence of the polyomavirus SV40 LT, resembling the MCPyV LT, which was utilized to immortalize COS7 cells. The insect cell line Sf9 provided a robust and cheaper production system compared to mammalian cells and is generally used as a model cell line for *in cellulo* crystallization (Redecke et al., 2013; Schönherr et al., 2015). However, an altered post translational N-glycosylation, summarized in (X. Shi & Jarvis, 2007), could be a disadvantage for highly post-translational modified proteins, which applies to many proteins of viral origin.

Crystal formation was observed 2-5 d.p.t., depending on the protein construct and the cell line. All constructs which successfully built *in cellulo* crystals (Figure 13A, construct 2-7) were expressed in HEK293 cells and detected 3 d.p.t. via WB, using the FLAG antibody to visualize PAK4cat, the HA antibody for Inka1 and 122tsT, the 2T2 antibody for sT and the Cm2B4 antibody for tLT244 (Figure 13B). Interestingly, the 122tsT-Inka1 variants (lanes 6 and 7) do not appear at the same size of approximately 41 kDa. As all proteins, besides PAK4cat, run slightly higher than their theoretical molecular weight shown in Figure 13A, it was anticipated that 122tsT-Inka1 (lane 6) may have undergone degradation. Thus, it was excluded from further analysis. *In cellulo* crystals of the protein constructs 2-7, grown in HEK293 cells and analyzed 3 d.p.t., are exhibited in Figure 13D. All crystals formed acicular structures whereas the number per cell and the size ranging from 10 – 100 µm was protein-dependent. However, crystals of the individual proteins did not differ between COS7 and HEK293 cells, which is why crystal formation in COS7 cells is not shown. All crystals arose predominantly as bundles or as single, elongated spicules (micrographs 2, 3, 4, and 5), except for those incorporating 122tsT. These crystals were highly abundant per cell and frequently grew radially next to bundle formation (micrographs 6 and 7). Non-impaired cell integrity by punctuating the plasma membrane, but all eventually led to cell detachment. Since Inka1 does not contribute to any crystal contacts (Baskaran et al., 2015), but functions as a linker between PAK4cat and the protein of interest, crystal formation could be observed in the absence of the inhibitor (micrograph 2). Table S 2 summarizes all constructs screened for *in cellulo* crystallization and evaluates their ability to build *in cellulo* crystals in mammalian and insect cells.
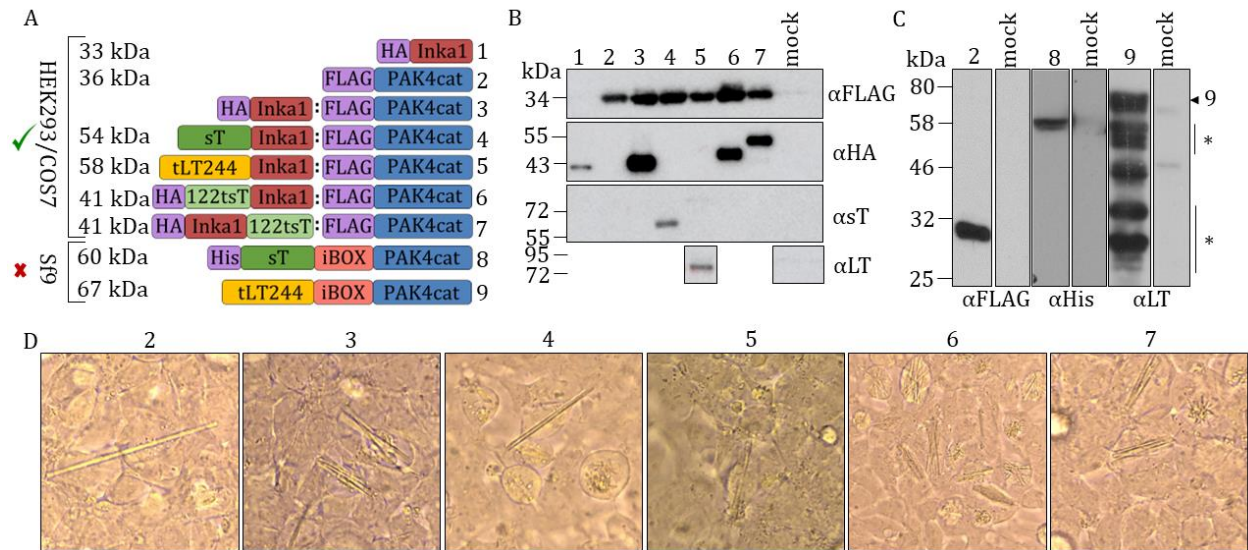
Figure 13: Formation of *in cellulo* crystals of various protein constructs. (A) Co-expression (colon) of the catalytic site of PAK4 (PAK4cat, blue) with sT (green), truncated sT of amino acids 122-186, (122tsT, light green) or truncated LT of amino acids 1-244 (tLT244, yellow) fused N- or C-terminally to Inka1 (red). The HA, FLAG and His tag (purple) were used for protein detection by IF and WB. Additionally, sT and tLT244 were expressed as fusion proteins of the PAK4cat interacting domain iBOX (light red), residing in Inka1, and PAK4cat. The molecular weight, the expression system (COS7/HEK293: mammalian origin; Sf9: insect origin) and crystal formation (successful: green check; unsuccessful: red cross) are exhibited on the left of the constructs. The construct numbers being used in (B), (C) and (D) are assigned on the right. (B) HEK293 cells were co-transfected with 500 ng DNA of the constructs in a 6-well format. Successful protein expression was shown by WB using the FLAG, HA, sT (2T2) and LT (Cm2B4) antibody for PAK4cat, Inka1 and 122tsT, sT and LT detection, respectively, 3 d.p.t.. (C) Sf9 insect cells were transfected with 1.5 µg recombinant bacmid DNA for baculovirus production in a 12-well format. Successful protein expression was shown by WB using the FLAG, His and LT (Cm2B4) antibody for PAK4cat, sT and tLT244 detection, respectively, 4 d.p.t.. The stars mark potential degradation products. (D) Bright field light microscopy (40x objective) shows *in cellulo* crystals of construct 2-7, 3 d.p.t..

The expression of PAK4cat alone as well as the T antigen-iBox-PAK4cat fusion proteins (Figure 13A, 2, 8 and 9), whose expression was verified via WB, using the FLAG, His and LT (Cm2B4) antibody for the detection of PAK4cat, sT-iBox-PAK4cat, and tLT244-iBox-PAK4cat, respectively, did not result in *in cellulo* crystals in insect cells (Figure 13C). In addition, the expression of the constructs, being sorted into the peroxisome, ER, or nucleus, also being performed in insect cells, did not yield *in cellulo* crystals after up to 8 d.p.t. (data not shown). These results demonstrate that insect cells were not a suitable expression system to produce *in cellulo* crystals of the MCPyV T antigens. Moreover, the expression of GFP-iBox-PAK4, which served as a model construct since expression and crystal formation could be easily monitored, did not result in *in cellulo* crystals in mammalian cells (data not shown). Therefore, the expression of the entire crystalline matrix together with the protein of interest as multi-domain protein was not considered further.

In summary, the approach of the proteinogenic crystalline molecular flask, by co-expressing the T antigen-Inka1 with PAK4cat, led to successful crystal formation in mammalian cells. In contrast, the expression of the crystalline matrix as fusion proteins did not result in intracellular crystals in any expression system. In addition, neither the proteinogenic crystalline molecular flask principle nor the

utilization of different cellular compartments induced *in cellulo* crystals in insect cells, which is why all subsequent analyses were performed in mammalian cells.

6.1.3. *In cellulo* crystals harbor T antigens but exhibit morphological heterogeneity

Since PAK4cat built crystals without the inhibitor Inka1, shown in Figure 13C, the presence of the T antigens in *in cellulo* crystals was verified by immunofluorescence staining (5.16.16). The T antigens fused to HA-Inka1 were co-expressed with PAK4cat in COS7 cells, which grew on gelatin-coated glass coverslips, in a 24-well plate. The cells were fixed 3 d.p.t., stained against LT or HA, and analyzed with confocal microscopy. Figure 14 shows the presence of tLT244 (A), sT (B), 122tsT (C), and Inka1 alone (D) in the intracellular PAK4cat crystals. In contrast, the sT variants were detected indirectly via the HA tag since no sT antibody was available for IF experiments. The antibodies did not penetrate the crystals and stained only their exterior, resembling the positive control Inka1 alone, whose atomic structure had already been solved, providing first indications towards a dense crystal packaging and the suitability of the engineered crystals for high-resolution structure analyses.

Summarized, crystals which arose from the expression of the T antigens and the crystalline matrix incorporated the protein of interest and were not solely built out of PAK4cat subunits. Furthermore, the crystal staining of the positive control and the T antigen harboring crystals was comparable, indicating their suitability for structure analysis.
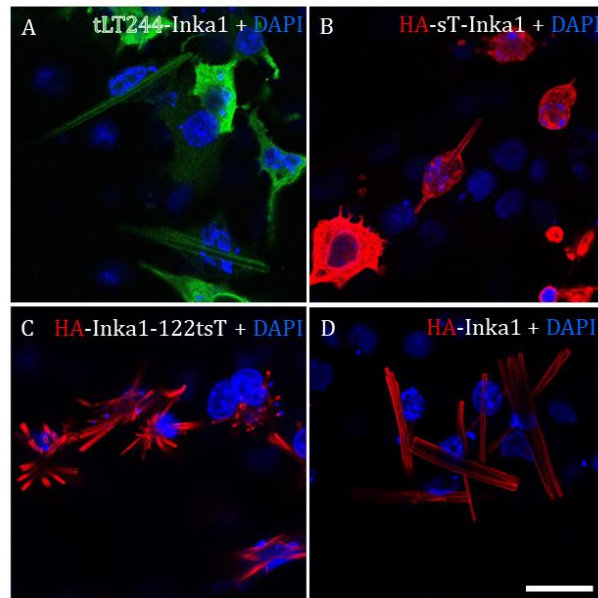
Figure 14: Immunofluorescence staining of T antigen variants incorporated in intracellular crystals. COS7 cells were co-transfected with HA-Inka1 fused to tLT244 (A), sT (B), 122tsT (C) or without fusion protein (D) and PAK4cat in a DNA ratio of 1:5. Cells were fixed 3 d.p.t. and stained with HA (red) or LT (Cm2B4, green) antibody and DAPI (blue) and analyzed with confocal microscopy. Scale bar: 30 μm.

After the presence of the T antigens in *in cellulo* crystals was verified, their crystallinity was analyzed. Ultra-thin sections of resin-embedded cells incorporating recombinant crystals were examined via TEM. HEK293 cells were grown in ibiTreat 35 mM μ-Dishes and co-transfected with sT-Inka1 or Inka1 alone and PAK4cat, which were analyzed 5 d.p.t. or 2 d.p.t., respectively. TEM analysis was restricted to sT-Inka1:PAK4cat due to its higher abundance and higher molecular weight, compared to the tLT244 and 122tsT construct, respectively. Thereby, high abundance abridged detection of crystal harboring cells and higher molecular weight increased the probability that the protein of interest affects crystal packaging, which could be revealed by TEM. Figure 15A shows the cross-section of a sT-Inka1:PAK4 hexagonal crystal with a visible crystal lattice, showing the solvent channels built by twelve PAK4cat units (Figure 11). The longitudinal section of these crystals (Figure 15B) represents the characteristic acicular shape, confirming the observation by light microscopy and immunofluorescence staining (Figure 13 and Figure 14). They are packed in bundles with a crystal width of approximately 75-250 nm. The cross-section of an Inka1:PAK4cat crystal alone is shown in Figure 15D. These crystals resemble the hexagonal shape seen in micrograph A, suggesting that sT is suitable for high-resolution structure analysis. However, next to crystals with a visible crystal lattice and defined shape, many crystals of sT-Inka1:PAK4cat showed frayed edges and an incomplete or missing hexagonal shape, indicating insufficient crystal packaging upon T antigen incorporation (Figure 15C).

It can be stated that the incorporation of sT into the Inka1:PAK4cat crystal matrix led to a crystal heterogeneity, including those with hexagonal shapes and visible crystal lattice, resembling the positive control as well as crystals missing hexagonality and with frayed edges.
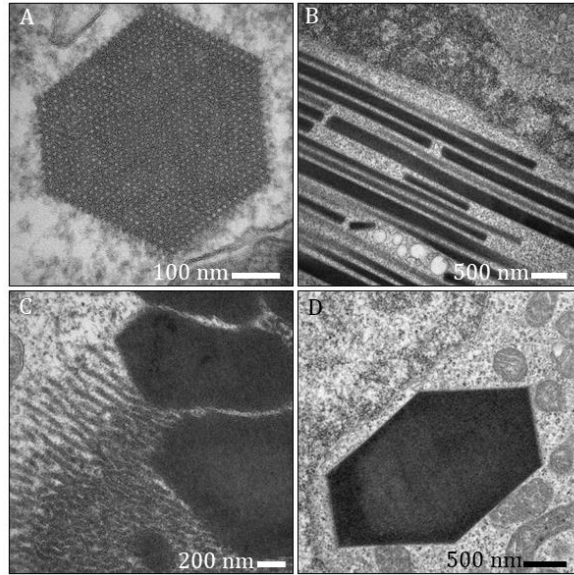
Figure 15: Transmission electron microscopy of *in cellulo* crystals of Inka1:PAK4cat variants. HEK293 cells were co-transfected with sT-Inka1 (A-C) or Inka1 alone (D) and PAK4cat in a DNA ratio of 1:5 and analyzed 5 d.p.t. and 2 d.p.t., respectively. Cells were stained with 1% osmium tetroxide and 1% uranyl acetate, followed by Epon 812 epoxy resin embedding and 50 nm sectioning. Cross (A) and longitudinal (B) section of sT-Inka1:PAK4cat crystals, showing the crystal lattice and an acicular shape. (C) Cross-section of sT-Inka1:PAK4cat crystals without hexagonal shape and with fraying edges. (D) Cross-section of the positive control Inka1:PAK4 crystals.

The crystals were examined with the synchrotron X-ray source of PETRAIII, EMBL P14, Hamburg to obtain diffraction patterns for the molecular structure elucidation of the T antigens. HEK293 and COS7 cells, containing crystals described in 6.1.2, were grown in T25 flasks and trypsinized 4 d.p.t.. Single cells were mounted onto a nylon loop, using PEG400 as cryo protectant and analyzed with a 12.7 keV micro beam (5 x 10 µm) with a wavelength of 0.98 Å. None of the T antigens containing crystals provided sufficient diffraction signals for analysis. The crystals did not diffract or diffracted only to very low resolution, associated with a very low crystal-hit rate.

In parallel to *in cellulo* diffraction, attempts were made to use isolated crystals for diffraction to increase the crystal concentration and the signal-to-noise ratio. However, crystals could not be recovered after chemical cell lysis (buffer composition is shown in Table S 3) or sonication (15% amplitude, 30 s) and were thus not used for synchrotron analysis.

In summary, the production of *in cellulo* crystals of T antigen variants was successful by utilizing the proteinogenic crystalline molecular flask approach. The presence of the T antigens within the crystals could be verified and the crystal lattice of sT containing crystals was revealed even though a crystal heterogeneity was observed, including those with missing hexagonality and frayed edges. Despite the successful crystal production and their promising characterization, no diffraction data could be collected to obtain high-resolution structure information of the MCPyV T antigens.

6.2. Expression, purification, characterization, and optimization of MCPyV LT for structure analysis with TEM

The lack of intracellular crystal diffraction, protein polydispersity, and high protein precipitation (results shown in the following) made the approach of *in cellulo* and *in vitro* crystallization unsuitable for structure elucidation of the MCPyV T antigens. In the following, the focus is set on TEM to gain structural insights into LT. The advantage of cryoEM SPA is the small sample volume and low protein concentration required. Additionally, crystallization is dispensable, heterogeneity is tolerated to an extent, and large protein complexes, such as the full-length LT hexamer with a size of 552 kDa, are beneficial to increase the signal-to-noise ratio during data acquisition. ST was not further analyzed due to its small size of 22 kDa. The structure of the full-length LT is of great interest, as no comprehensive LT structure of any PyV is available. In addition, domains of the full-length protein may stabilize its unstructured MURI/II regions. These regions hold no similarity to any protein and may be involved in the exceptional characteristic of MCPyV to induce carcinogenesis in its host. Taken together, the structure of the full-length LT might reveal intradomain interactions and extensive structure-function relationships, leading to a better understanding of the function of LT and its involvement in MCC. In addition, the atomic protein structure would provide a starting point for targeted antiviral treatment development.

6.2.1. Relative domain position and unstructured regions of LT cannot be predicted by *in silico* analysis

First insights into the full-length MCPyV LT were obtained by sequence alignments with SV40 and MPyV LT, greatly varying in the presence of an N-terminal unstructured region (Figure 4B), and AlphaFold2 prediction. LT sequence alignment of all three PyV revealed a sequence identity between 30-40%, verifying protein homology with greater proximity between MCPyV and MPyV. Single domains, such as the OBD and helicase, showed higher sequence identity of up to 52% (MCPyV:SV40) and 62% (MCPyV:MPyV), respectively. The unique regions of MCPyV and MPyV revealed sequence identity below 30% and were not considered homologous. The DnaJ domain also showed low sequence identity between 28-31% (Figure 16A). To assess whether the high sequence similarity of single PyV domains was linked to structure homology, MCPyV LT OBD, the only published crystal structure of MCPyV LT, was aligned to the SV40 equivalent (Figure 16B). The domains showed high structure similarity with a root mean square deviation (RMSD) of 1.01 Å, providing evidence that SV40 LT can serve as a model protein for MCPyV LT buffer optimization and structure analysis. AlphaFold2 prediction resulted in compact DnaJ, OBD, and helicase domains with high local prediction confidence of pLDDTs (predicted local distance difference tests) up to 98.1, reflecting

accurate backbone prediction and a frequent correct side chain orientation. However, the MURI/II and the C-terminus were predicted to be ribbon-like with pLDDTs below 50, considering these parts as unpredictable and disordered in isolation (Figure 16C and D). Further, low confidence in the relative domain positions of the predicted LT was demonstrated by the predicted aligned error (PAE) plot, assessing the confidence in the relative position of residue pairs measured in Å (Figure 16E). These results emphasized the necessity of experimental data for comprehensive structure analysis.
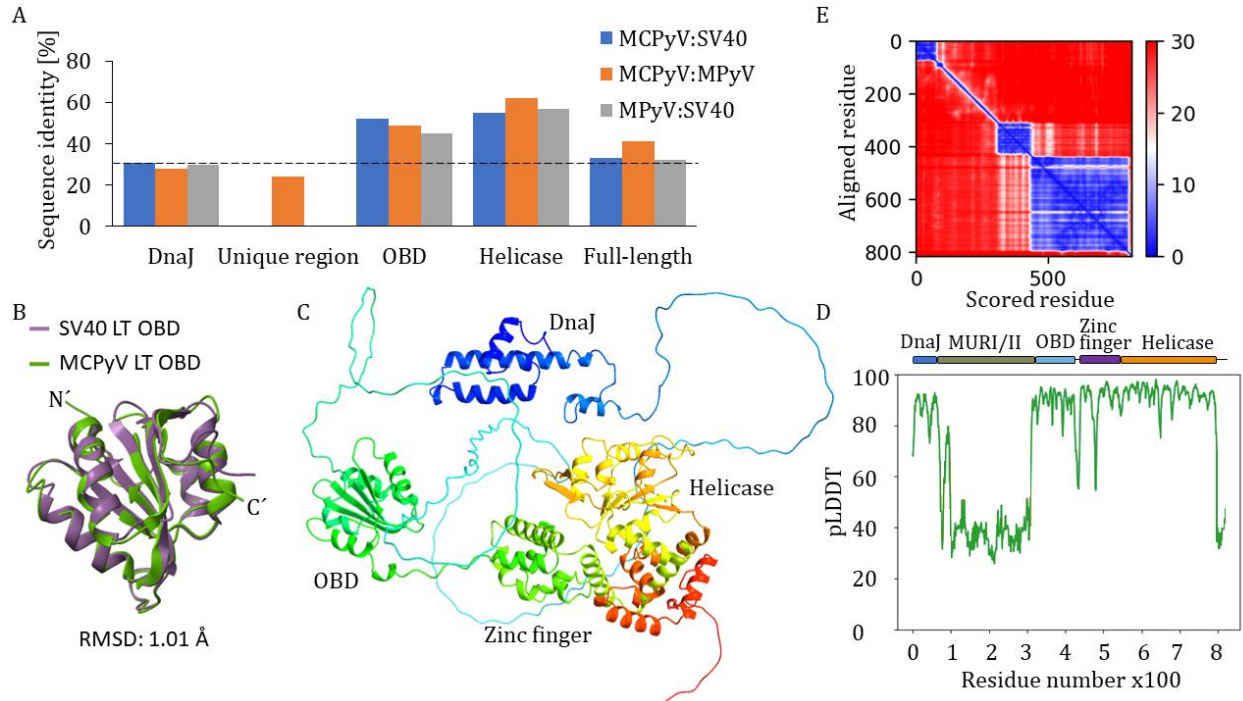
Figure 16: *In silico* analyses of MCPyV LT. (A) Pairwise sequence comparison of MCPyV, MPyV, and SV40 LT. Multiple sequence alignment of full-length LT and all single domains (DnaJ, unique region, OBD, and helicase) was performed in CLC. Sequence identity was color-coded in blue (MCPyV:SV40), orange (MCPyV:MPyV), and grey (MPyV:SV40). Low sequence homology was considered below 30% sequence identity (dashed line). (B) Structure alignment of MCPyV and SV40 LT OBD. The MCPyV crystal structure (green, PDB: 3QFQ) was aligned to the SV40 crystal structure (purple, PDB: 2ITL), and the associated root mean square deviation (RMSD) was depicted. N- and C-terminus were marked with N´ and C´, respectively, and the alignment was performed and visualized by ChimeraX. (C) AlphaFold2 prediction of LT. The protein was rainbow color-coded from blue (N-terminus) to red (C-terminus), and domain names were depicted next to the associated structures. (D) Predicted local distance difference test (pLDDT) plot of (C). PLDDTs of all LT residues, depicted on the x-axis, can be correlated to the LT domains (DnaJ: blue; MURI/II: green; OBD: light blue; Zinc finger: purple; Helicase: orange), which were arranged above the plot. (E) Predicted aligned error (PAE) plot of (C). The expected position error of each residue, given in Å, was color-coded in blue (high position confidence) and red (low position confidence).

6.2.2. Determination of optimal LT expression and cell lysis conditions

Full-length LT was expressed via the baculovirus expression system in insect cells (5.14.5), which provides most posttranslational modifications administered in mammalian cells, but shows higher expression rates, and is economically favored compared to mammalian expression systems. LT, with a size of 92 kDa and an isoelectric point (pI) of 6.9, was C-terminally coupled to a His, StrepII, and MBP tag (Table 8 and Table 9), which were cleaved prior to SEC via a TEV cleavage site (Figure 17A). MBP was utilized to increase the protein solubility, as vast amounts of the protein, tagged solely with

His and StrepII (Table 8 and Table 9), were found in the cell pellet (Figure S 2A and B). Moreover, despite verified LT expression (Figure S 2C), His and StrepII tag could not be detected by WB (results not shown) and the protein failed to be purified via NiNTA or StrepTactin affinity chromatography. LT was found in the flowthrough and wash fractions, which is shown exemplarily for the StrepTactin affinity chromatography in Figure S 2B.

The expression endpoint and the P2 baculovirus transduction volume were determined to maximize protein yield. LT-His-StrepII-MBP expression was observed over seven days after transducing High Five™ cells with 0.2%, 0.5%, or 1% (v/v) P2 baculovirus supernatant. The pellet was resuspended in lysis buffer (Table 5), sonified, and the resulting supernatant and pellet were analyzed for soluble protein expression by SDS-PAGE and verified by WB. The use of 0.5% (v/v) P2 baculovirus supernatant provided the highest soluble protein expression, which was generally low and comparable to cellular protein expression (Figure 17B). LT was detected on day 2 at the expected size of approximately 137 kDa. It had its maximum expression between day 2-3 and only a small fraction could be found in the insoluble pellet. Therefore, all following LT expressions were carried out with 0.5% (v/v) P2 baculovirus supernatant and harvested 3 d.p.t..
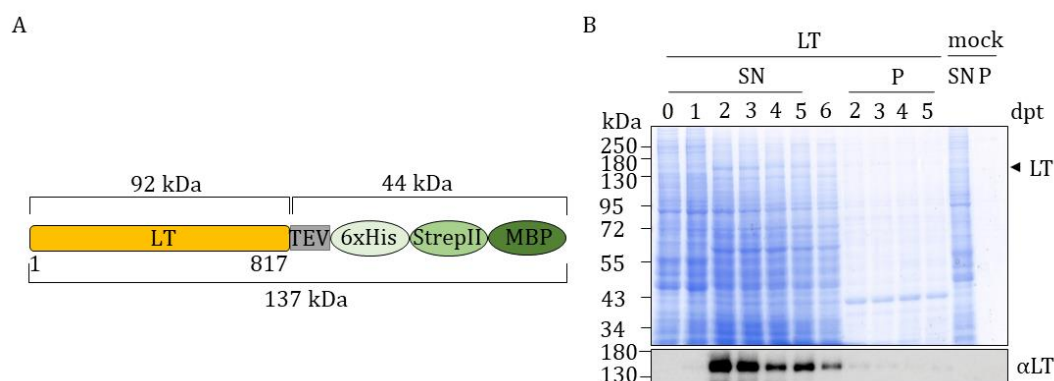
Figure 17: LT test expression to determine the day of harvest and the P2 baculovirus volume for high LT yield. (A) LT construct used for expression and purification. LT (yellow) was C-terminally tagged with 6xHis (light green), StrepII (green) and MBP (dark green), which were cleavable via a TEV cleavage site (grey). The amino acid length is depicted under the construct and the molecular weight of LT with and without tags is indicated in kDa. (B) SDS-PAGE and WB of LT test expression using 0.5% (v/v) P2 baculovirus supernatant. 10^6 transduced (LT) or untransduced (mock) High Five™ cells were incubated up to 6 days post-transduction (dpt), lysed in 200 µl RIPA buffer (Table 5), and sonified. The pellet (P) was dissolved in 200 µl RIPA buffer and same volumes of supernatant (SN) and P were analyzed via SDS-PAGE (8% gel) and Coomassie staining. LT expression was verified by WB using the Cm2B4 LT-antibody (αLT).

Thereafter, optimal cell lysis conditions were tested. LT is a DNA-binding protein, making purification without nucleic acid contamination challenging. It harbors two nuclear localization signals (NLS) in the MURII domain, leading to nuclear localization in the viral replicative life cycle (Borchert et al., 2014), which necessitates nucleus disintegration during cell lysis. Therefore, three different lysis approaches were applied, and protein yield, protein solubility, and nucleic acid contamination were examined. The production of nuclear extracts in two steps, by varying the salt concentration (LB6a

and b, Table 5) and excluding the cytosolic fraction for protein purification, was used as a first approach (1). Degradation of the cytoplasmic membrane using Triton X-100, high salt concentration, and sonication (LB5, Table 5) was adapted as the second lysis method (2). Nonidet P-40, also disintegrating the cytoplasmic membrane, and high salt concentration (LB4, Table 5) was applied as a third approach (3) (Figure 18). High Five™ cells were transduced with recombinant baculovirus, followed by LT recovery 3 d.p.t. via cell lysis and MBP gravity flow affinity chromatography. Lysis method 2 showed no insoluble LT protein and resulted in the highest protein yield after affinity chromatography with the lowest nucleic acid contamination, showing a 260/280 ratio of 0.67. LT purification was also successful with method 3, but insoluble LT emerged and nucleic acid contamination with a 260/280 ratio of 0.76 was elevated compared to method 2. The sole use of the nuclear extract of method 1 led to high protein aggregation, apparent above 250 kDa, insoluble LT, and only low protein amounts could be recovered from the resin.
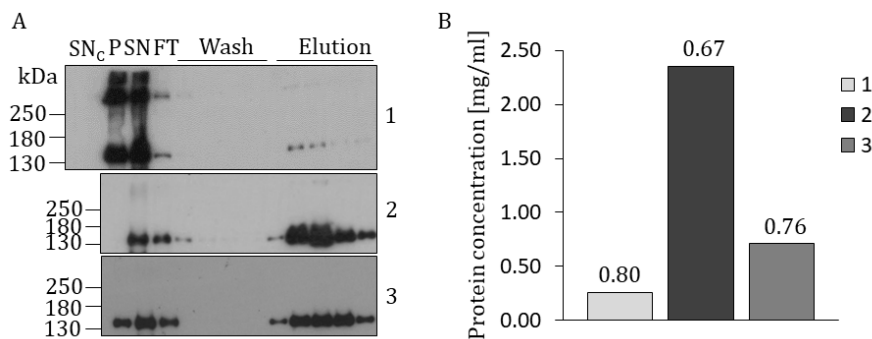
Figure 18: Determination of the optimal cell lysis method for LT recovery. (A) WB of LT-His-StrepII-MBP MBP gravity flow affinity chromatography, based on different lysis methods. 5 x 10^7 High Five™ cells were transduced with recombinant baculovirus and analyzed 3 d.p.t.. The lysis methods are exhibited on the right and include: (1) the production of nuclear extracts in two steps by increasing the salt concentration (LB6a and b, Table 5) and excluding the cytosolic fraction for protein purification; (2) Triton X-100, high salt concentration and sonication (LB5, Table 5); (3) Nonidet P-40 and high salt concentration (LB4, Table 5). The chromatography was carried out with 500 µl amylose resin and same volumes of the resulting cytosolic supernatant (SN_C), nucleic supernatant for (1), and combined supernatant for (2 and 3) (SN), pellet (P) which was resuspended in equal volumes as SN, flowthrough (FT), wash, and elution fractions were applied to an SDS-PAGE (8% gel). LT was visualized via WB, using the LT-antibody Cm2B4. (B) Protein concentration of combined elution fractions of (A). The eluate was concentrated via a 100 K amicon filter to approximately 110 µl and measured by NanoDrop. The 260/280 ratio is depicted above each column.

Summarized, LT was successfully expressed in High Five™ insect cells, obtaining the highest expression rate by using 0.5% (v/v) recombinant baculovirus and recovery 3 d.p.t.. The highest protein concentration and lowest nucleic acid contamination after affinity chromatography was accomplished by lysing the cells with Triton X-100 in high salt buffer and sonication. Additionally to lysis method 2 and 1 M NaCl, protein purification was realized with the same method and 250 mM NaCl to stabilize LT during purification by unspecific DNA binding, which was inhibited in the presence of high salt buffer (results not shown).

6.2.3. Affinity and size exclusion chromatography are applied for LT purification

Affinity chromatography using the MBP tag was applied as a first purification step to recover LT from the crude cell lysate. It showed improved binding specificity compared to NiNTA and is economically favored over StrepTactin, although the purification exploiting the StrepII tag was effective as well (results not shown). Figure 19A shows the successful purification and enrichment of LT-His-StrepII-MBP, which was exogenously expressed in High Five™ cells, lysed as described in 6.2.2 with the lysis buffer LB7 (Table 5), and purified by an Äkta-coupled MBPTrap (5.16.6). However, vast amounts of LT, present in the flowthrough and pellet, could not be recovered and many proteins co-eluted, requiring further purification steps. Additionally, LT attached irreversibly to the resin and could not be eluted without compromising its biological integrity (Figure 19B).

Following initial purification, the affinity tags were removed by TEV protease (5.16.11). Proteolytic cleavage was examined up to 38 h at 4 °C and 21 °C (Figure 19C). Comprehensive tag removal was observed after 6 h at 21 °C and 38 h at 4 °C. Since the majority of LT was cleaved after 16 h, all subsequent tag removals were carried out overnight at 4 °C to decrease purification time.
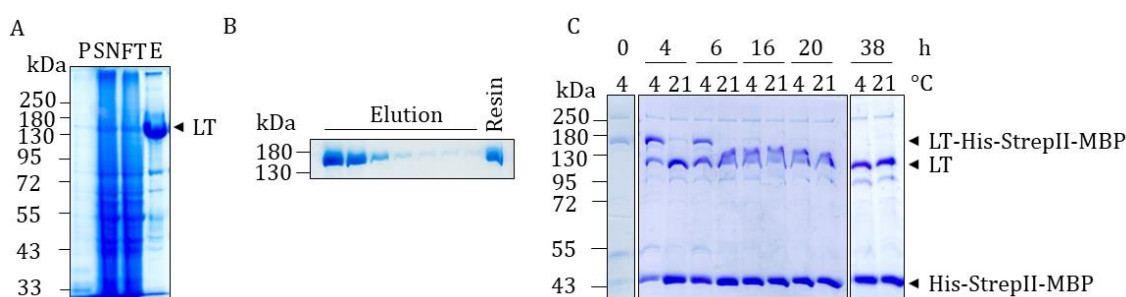
Figure 19: MBPTrap affinity chromatography of LT and its tag removal. (A) MBPTrap affinity chromatography of LT-His-StrepII-MBP. 2.5×10^8 High Five™ cells were transduced with recombinant baculovirus. Cells were lysed in LB7 (Table 5), 3 d.p.t., and purified with an MBPTrap on an ÄktaPurifier system with the elution buffer EB2 (Table 5). Similar volumes of resuspended pellet (P), supernatant (SN), flowthrough (FT), and pooled and concentrated elution fractions (E) were separated via an SDS-PAGE (8% gel) and Coomassie stained. (B) MBP gravity flow affinity chromatography of LT-His-StrepII-MBP. 3×10^8 High Five™ cells were transduced with recombinant baculovirus. Cells were lysed in LB3 (Table 5), 3 d.p.t., and purified using 1 ml amylose resin and the elution buffer EB2 (Table 5). Similar volumes of elution fractions and resin (1:1 ratio of resin:buffer) were separated via SDS-PAGE (8% gel) and Coomassie stained. (C) Similar volumes of LT, purified via affinity chromatography and incubated with TEV at 4°C or 21 °C, were taken at 4-38 h. Time 0 shows the elution fraction of the preceded affinity chromatography.

After initial affinity chromatography, various purification methods were examined to improve LT purity. SV40 LT was mainly purified by immunoaffinity chromatography, resulting in 95% pure protein with high homogeneity, showing ATPase activity, sequence-specific DNA binding, and p53 binding (Dixon & Nathans, 1985; Simanis & Lane, 1985). Since SV40 LT served as a model protein, immunoaffinity purification was also applied to MCPyV LT. Protein G was linked to the self-made T antigen binding monoclonal antibody 2T2 (5.14.6) and incubated with LT-His-StrepII-MBP, which was pre-purified via MBP gravity flow affinity chromatography (5.16.6). WB of the immunoaffinity chromatography showed successful 2T2-coupling onto Protein G (P$_G$-2T2), but the co-purification of

high molecular weight aggregate could not be excluded thoroughly, and only a minor fraction of LT-His-StrepII-MBP was captured by the antibody (P_G-2T2-LT) (Figure 20A). The majority of LT was found in the flowthrough and wash fractions, wherefore immunoaffinity chromatography using the 2T2 antibody was not suitable for MCPyV LT purification.

For anion exchange chromatography (5.16.8), the salt concentration of the protein buffer (EB2, Table 5) was reduced by overnight dialysis from 250 mM to 100 mM (IEXB1, Table 5), resulting in vast LT precipitation of 70% from 2.2 mg to 1.5 mg (Figure 20C, P). The remaining protein was applied to a HiTrap Q FF at a conductivity of 13 mS/cm. LT eluted at 36% and 62% EB3 (Table 5) and a conductivity of 30 mS/cm and 46 ms/cm, respectively (Figure 20B and C,2 and 3). Impurities could not be removed entirely, and most of the applied LT was not captured by the resin and was found in the flowthrough (Figure 20C, FT). LT instability in moderate salt buffer and low protein-resin interaction led to the fact that ion exchange chromatography as well as heparin adsorption chromatography (results not shown) were not further pursued. Therefore, structure analyzing methods, such as SEC, DLS, or EM were not applied to analyze potential differences in LT conformation of the elution fraction 1 and 2.
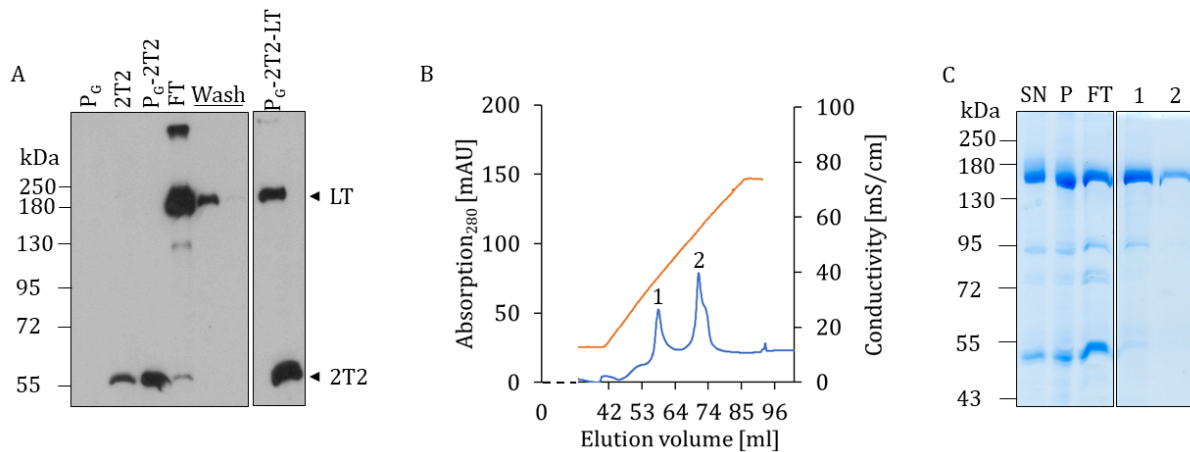
Figure 20: Challenging LT purification approaches. (A) Immunoaffinity chromatography of LT. 50 µl Protein G (P_G) was coupled to 5 ml of unconcentrated T antigen-antibody 2T2. Resin was incubated with 78 µg LT-His-StrepII-MBP, purified by MBP gravity flow affinity chromatography, for 2 h in EB4 (Table 5). LT-bound resin (P_G-2T2-LT) was washed with 2x 1 ml EB4 (Wash). LT was visualized via WB (8% gel), using the T antigen-antibody 2T2. Sample volumes added onto the gel: P_G: 5 µl; 2T2: 27 µl; P_G-2T2: 15 µl; Flowthrough (FT): 10% of total FT; Wash: 3% of each wash fraction; P_G-2T2-LT: 20% of LT-bound resin. (B) Anion exchange chromatography of LT. 2.2 mg LT-His-StrepII-MBP, purified by MBPTrap affinity chromatography, were dialyzed overnight from 250 mM (EB2, Table 5) to 100 mM NaCl (IEXB1, Table 5), and applied to an Äkta-coupled HiTrap Q FF column. LT was gradually eluted with 1 M NaCl (EB3, Table 5) over 50 ml with a flow rate of 0.25 ml/min. The absorption at 280 nm was measured (blue) and, together with the conductivity (orange), plotted against the elution volume. (C) SDS-PAGE of (B). Similar volumes of dialyzed LT (SN) and resuspended precipitated LT after dialysis (P) as well as similar volumes of the anion exchange chromatography flowthrough (FT) and the absorption peaks 1 and 2 were applied to an 8% gel and Coomassie stained.

Since immunoaffinity, anion exchange, and heparin chromatography failed for LT purification, the protein, previously purified via MBPTrap and proteolytically cleaved, was subjected to SEC to examine its oligomeric state and to increase purity. The sample was applied to a HiLoad Superdex 200 10/300 GL and eluted with EB4 (Table 5). The theoretical elution volume of 10 ml for the LT hexamer (552 kDa) and 31.1 ml for the LT monomer (92 kDa) was calculated by utilizing the SEC calibration chromatogram (Figure S 3). However, LT elution was observed at a volume of 7.42 ml, which was within the column void volume of 8 ml, indicating protein aggregation (Figure 21A). In addition, LT could not be comprehensively purified, and contaminations were still visible by Coomassie staining (Figure 21B). A 260/280 ratio of 0.64 and 1.4 was observed for LT after SEC, depending on the use of 1 M or 250 mM NaCl for cell lysis, respectively. Exploitation of a Superose 6 Increase 10/300 GL, beneficial for the separation of large protein complexes with a fractionation range of 5-5000 kDa, also showed LT elution within the void volume (results not shown). This observation excluded the concern of an insufficient separation capacity to fractionate spacious LT complexes such as double-hexamers (1.1 MDa), which are known for SP3 helicases and have been described previously for SV40 LT (Mastrangelo et al., 1989; Valle et al., 2000).
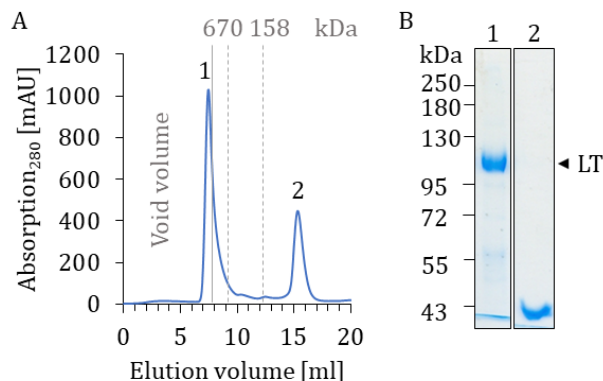
Figure 21: Size exclusion chromatography of LT. (A) 1.2 mg LT, purified by MBPTrap and proteolytically cleaved, was applied to a Superdex 200 10/300 GL and eluted in EB4 (Table 5) with a flow rate of 0.25 ml/min. The elution volume of thyroglobulin (670 kDa) and γ-globulin (158 kDa) from the column calibration (Figure S 3) are depicted with dashed lines (gray). The void volume is marked with a continuous line (gray). (B) Fractions of 280 nm absorption peaks (1 and 2) were analyzed by SDS-PAGE (8% gel) and Coomassie stained.

Summarized, MBPTrap affinity chromatography was utilized as initial LT purification step, although vast protein amounts were not captured by the resin or were irreversibly attached to the matrix. Poor LT-resin interaction and its instability in moderate salt buffer disqualified immunoaffinity and anion exchange chromatography for further LT purification and affinity tag segregation, which was realized by SEC, whereby protein aggregation was observed.

6.2.4. The viral ori, MgCl$_2$, ADP, and phosphate stabilize LT and reduce polydispersity

The R$_H$ of purified LT was assessed by DLS. For SV40 LT, the formation of various oligomers was observed in the absence of MgCl$_2$ and ATP. By adding these compounds, LT assembled into hexamers, or double-hexamers in the presence of the viral ori (Gai, Li, et al., 2004; D. Li et al., 2003; Reynisdóttir et al., 1993; Dean et al., 1992; Mastrangelo et al., 1989). Therefore, the addition of the viral ori (Figure 23C), MgCl$_2$, and ADP was examined to reduce MCPyV LT aggregation and polydispersity. ADP was substituted with ATP to mimic a post-hydrolytic state, diminishing ATPase activity. The inhibition of ATP hydrolysis was hypothesized to reduce conformational changes, leading to helicase activation, as shown for SV40 LT (Gai, Zhao, et al., 2004), which possibly increases polydispersity. In addition, hexamer formation and LT-DNA interaction is independent of ATP hydrolysis and require only ATP/ADP binding (Reynisdóttir et al., 1993; Borowiec & Hurwitz, 1988).

LT with and without additives was incubated in EB4 (Table 5) for 1 h at 37 °C to enhance protein-DNA interaction prior to examination. The mean R$_H$ calculated by averaging the autocorrelation functions (ACF) of LT without additives resulted in two distinct populations with an R$_H$ of 20 nm and 47 nm (Figure 22A, green). The addition of MgCl$_2$ and ADP changed the ratio of the populations to approximately 2:1, respectively (Figure 22B, green). The incubation of LT with MgCl$_2$, ADP, and the viral ori led to one dominant LT fraction with a peak maximum at 18 nm (Figure 22C, green). The second fraction was strongly reduced and, in parallel, a small population at 8 nm appeared. The mean R$_H$ calculated by averaging the radii originating from non-averaged ACFs reflected a sharper peak without shoulder upon the addition of MgCl$_2$, ADP, and ori, indicating reduced polydispersity (Figure 22C, blue). All peaks appeared in the size range of oligomers or aggregates whereas monomers, expected in the low to middle single-digit nm range, were not detected. Examination of the buffer with additives but without LT did not show signals in the stated size range (Figure 22D). Comparison of the radius plots also revealed that the addition of MgCl$_2$ and ADP reduced LT polydispersity, which was further enhanced by adding the viral ori. However, no condition reflected a monodisperse LT solution.
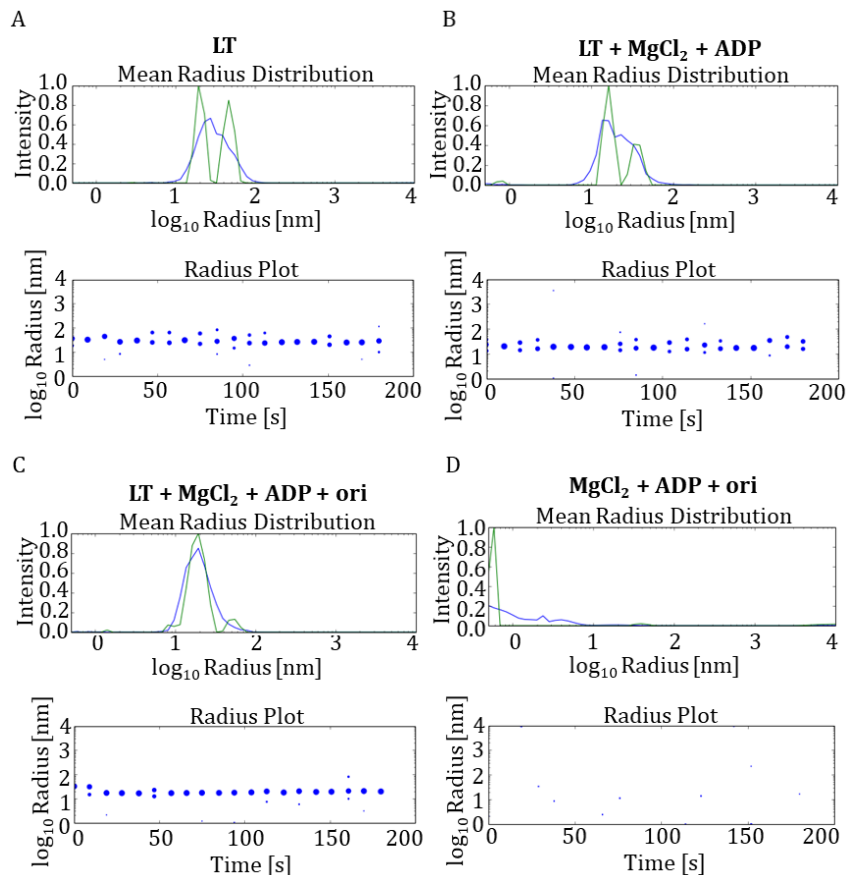
Figure 22: DLS of purified LT in various conditions. (A) 0.2 mg/ml LT, (B) 0.2 mg/ml LT supplemented with 2 mM MgCl$_2$ and 2 mM ADP, (C) 0.2 mg/ml LT supplemented with 2 mM MgCl$_2$, 2 mM ADP, and 0.6 pmol/µg LT ori (2.3 pmol), or (D) 2 mM MgCl$_2$, 2 mM ADP, and 0.6 pmol/µg LT ori without LT were incubated at 37 °C for 1 h and analyzed at 20 °C. The measurements were carried out in EB4 (Table 5) with 10 sec acquisition time and 20 iterations. The mean radius distribution plot shows the particle mean hydrodynamic radii (R$_H$) of the averaged autocorrelation function (ACF) (green) or averaged R$_H$ originating from non-averaged ACFs (blue). The radius plot depicts R$_H$ versus time, whereby the spot diameters are proportional to the scattered light intensities.

Next to specific LT-ori binding between the OBD and the pentanucleotide GRGGC (C. Harrison et al., 2011; Kwun et al., 2009), unpublished chromatin immunoprecipitation with subsequent DNA sequencing (ChIP-Seq), performed in our group, revealed specific binding between LT and certain cellular DNA promoters in various cell lines. In contrast, tLT244, lacking DNA interacting domains, did not bind to DNA (Figure S 4A). A motif calling on the LT-bound cellular promoters revealed the above stated LT-binding motif, though in varying number and arrangement (Figure S 4B). RNA-Seq analysis showed only 14% differential gene expression of those genes located in close proximity to the LT-bound promoters, excluding a direct impact of LT-cellular DNA binding on the expression of those genes (Figure S 4C). In addition, the viral microRNA (miR-M1) promoter, identified by Theiss et al., 2015 and located antisense in the early gene region, contains five GRGGC motifs and was therefore also a target for LT binding. An LT-miR-M1 interaction, however, has not been published thus far, but was observed in our group (data not shown). In this work, the binding behavior of LT to various DNA oligos was investigated to determine the oligo with the strongest effect on LT stability,

potentially leading to reduced aggregation and polydispersity. The viral ori, the viral miR-M1 promoter, and the cellular promoter of the p53 stabilizing tumor suppressor TP53INP1 (tumor protein p53-inducible nuclear protein 1) (Tomasini et al., 2003), which showed pronounced LT binding in the ChIP-Seq experiments, were analyzed and compared. Emphasis was placed on the specific DNA binding, thermal stability, and conformation of LT (Figure 23C).

DPI-ELISA (5.16.14) was utilized to test whether LT purification compromised LT-DNA binding and to verify specific LT-DNA interaction of the above stated, 80 bp long oligonucleotides (Figure 23C). Biotin-labeled oligonucleotides were immobilized onto an avidin-coated 96-well format and incubated with purified LT, whose binding was colorimetrically detected with the antibody Cm2B4. Absorption was normalized to the MCPyV capsid protein VP1 oligonucleotide, lacking GRGGC-binding motifs and having a GC content of 40%. DNA binding of purified LT could be detected, and it was shown that this binding was specific, revealing the strongest interaction between LT and the viral ori. Scrambled ori (ori scr), containing the same GC content but without GRGGC motifs, as a result of ori sequence rearrangement, showed approximately 0.7-fold LT binding enhancement, which disclosed an impact of the GC content on LT-DNA binding. No significant difference in the binding affinity between viral miR-M1, cellular DNA promoter, and ori scr could be observed (Figure 23A).

NanoDSF (5.16.14) was used to assess the thermal stability of LT in the presence of the ori and ori scr by measuring the change in fluorescence intensity of tryptophane and tyrosine upon successive temperature increase. Since the binding affinity of LT to ori scr, the cellular DNA promoter, and viral miR-M1 were comparable, the latter two were omitted for this analysis. Both, ori and ori scr, increased the thermal stability of LT by 5 °C and 2 °C, respectively (Figure 23B).

Lastly, the conformational state of LT, bound to the above stated oligonucleotides, was examined to determine the most suitable conformation for cryoEM SPA, favoring large protein complexes as they increase the signal-to-noise ratio. Since the binding of LT to cellular promoters did not result in pronounced expression changes of genes in their proximity (Figure S 4C), the function of this phenomenon is still elusive. Whether LT acts as a helicase or holds a different feature most likely aligns with a differed protein conformation compared to the expected LT-ori hexamer. Negative stain TEM SPA was used to examine LT in the presence of the viral ori, miR-M1 promoter, cellular TP53INP1 promoter, VP1, and without DNA. The final models comprised a maximum of 2400 (LT + ori) and a minimum of 336 (LT + VP1) particles, all revealing a hexamer shape with a small hexamer atop a larger one. Interestingly, LT in the absence of DNA also formed hexamers, indicating a DNA-independent hexamer formation (Figure 23D). The above described results suggested to supplement LT with ori for cryoEM SPA, showing the highest LT binding activity, the strongest LT thermal stabilization, and no obvious conformation difference, compared to the other tested oligonucleotides.
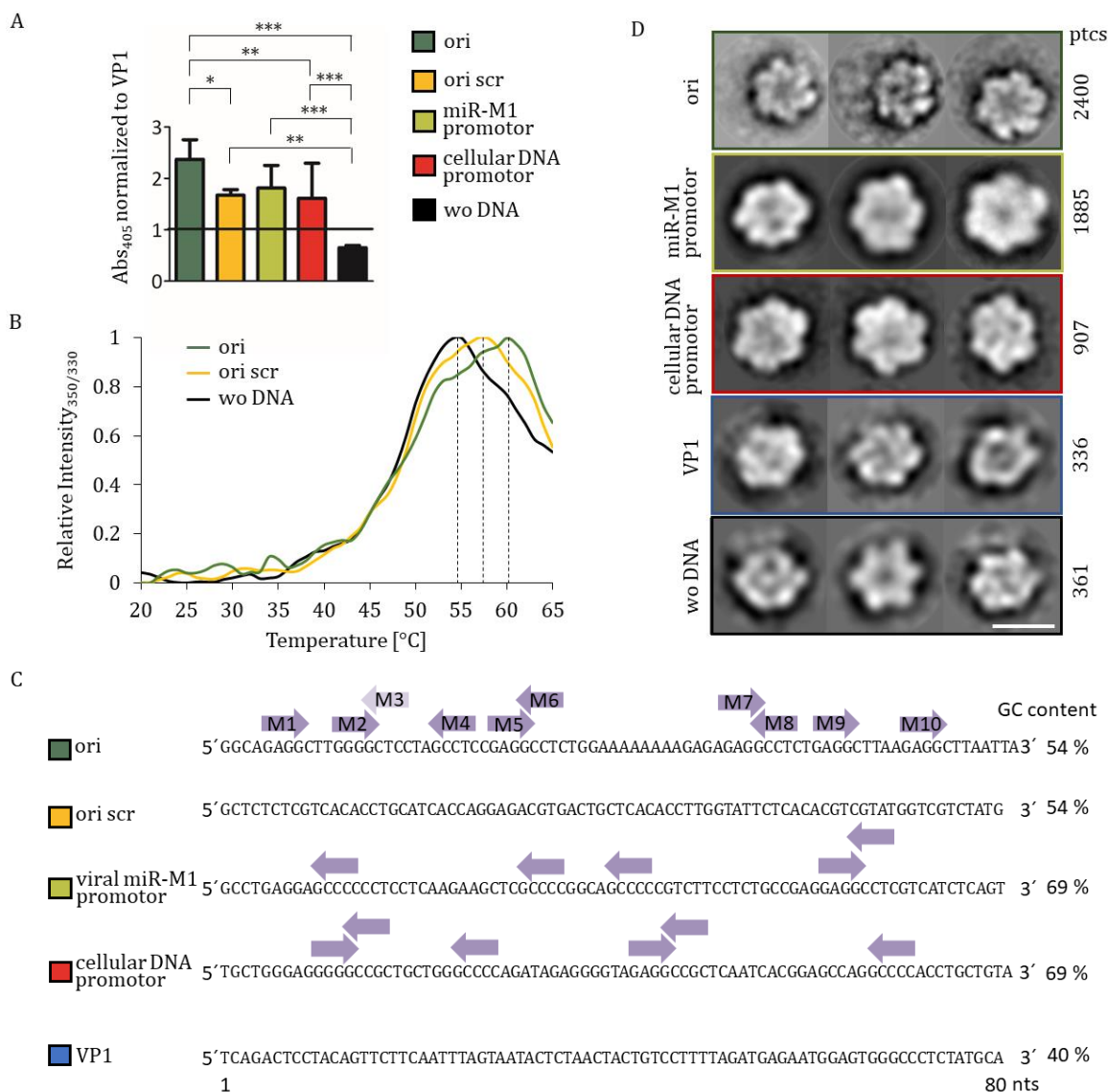
Figure 23: LT-DNA interaction. (A) DPI-ELISA of LT in the presence of oligonucleotides depicted in (C). 1 µg purified LT was incubated with 1 pmol biotinylated, immobilized oligonucleotides in PDBB (Table 4). LT-DNA binding was colorimetrically detected at 405 nm by using the antibody Cm2B4. The binding was normalized to VP1. Depicted is the standard deviation of n=6 and the statistical significance, calculated using a one-way ANOVA test with GraphPad Prism. (B) NanoDSF of LT in the presence of ori (dark green), ori scr (yellow), and without DNA (black) shown in (C). 2 µg of purified LT was incubated with 1.2 pmol of ori or ori scr in EB4 (Table 5) and successively heated from 20-95 °C. The normalized first derivation of the intensity ratio measured at 350 nm and 330 nm of the intrinsic fluorescence of tyrosine and tryptophane was measured and the melting temperature determined (dashed line) (C) Oligonucleotides examined for specific LT-DNA binding, thermal stability, and conformation of LT. All oligonucleotides had a length of 80 nucleotides (nts), and their GC content is depicted behind each sequence. Intact GRGGC motifs (M, dark purple) and motifs with one mismatch (M, light purple) are depicted above each oligonucleotide in sense and antisense (right-pointing arrow and left-pointing arrow, respectively). Exhibited is the viral ori (dark green), the viral scrambled ori (ori scr, yellow), having demolished GRGGC motifs (dark purple) but the same GC content compared to the wildtype ori, the viral miRNA promoter (miR-M1, light green), and the cellular promoter of TP53INP1 (tumor protein p53-inducible nuclear protein 1) (red). The oligonucleotide of the MCPyV capsid protein VP1 fraction (blue) was used as negative control. (D) Reference-free 2D averages of negative stain TEM SPA of LT in the presence of oligonucleotides depicted in (C). 0.5 µg LT in EB4 (Table 5) supplemented with 2 mM MgCl$_2$, 4 mM ADP, and 0.3 pmol DNA were incubated at 37 °C for 15 min and centrifuged prior to staining with uranyl acetate. LT was applied onto a glow-discharged carbon-coated copper grid and analyzed with a Talos L120C 120V electron microscope at 92k x nominal magnification. 2D classification was performed with the software Relion 3.1 without applying C6 symmetry. LT + ori (dark green), miR-M1 promoter (light green), cellular TP53INP promoter (red), MCPyV capsid protein VP1 (blue), and without DNA (black) were analyzed. The particles (ptcs) included in the final model are depicted on the right side. The scale bar represents 100 Å.

Definition of optimal LT conditions for cryoEM SPA was continued by buffer optimization, utilizing the ProteoPlex thermal stability assay, specifically designed for large protein complexes (5.16.13). LT was incubated with 88 different buffer conditions (Table S 3A), and its thermal stability was determined with help of the fluorescent dye Sypro orange and a successive increase in temperature. The most stabilizing buffer systems were sodium citrate (orange), succinic acid-phosphate-glycine (SPG, yellow), and phosphate (green) (Figure 24A). Only one condition was depicted exemplarily for sodium citrate, pH 6.0 (Tm 46.6 °C) and SPG, pH 8.0 (Tm 46.9 °C), showing an elevated LT stability of 1.7 °C and 2 °C, respectively, compared to the reference TRIS-HCl, pH 8.5 (black, Tm 44.9 °C). For the most prominent stabilizing agent, phosphate, all conditions were shown. LT stabilization could be observed over a broad pH range, inducing a thermal stability increase up to 3.5 °C at a pH of 6.4 (Tm 48.5 °C). The above stated buffer systems were tested for LT purification with varying pH, salt concentration, and resin (Figure 24B). Sodium citrate and many phosphate buffers, mostly with acidic or neutral pH, were not suitable for LT purification, leading to protein precipitation and low protein yield (red). Purification in TRIS-HCl, SPG, and TRIS-HCl, supplemented with phosphate, was successful, showing high protein yield after SEC (gray). The SEC elution profiles of LT in all successfully tested conditions were similar to that of TRIS-HCl, depicted in Figure 21A (results not shown). Negative stain TEM was applied to verify the above stated results and to assess the suitability of LT for cryoEM. Micrographs of LT in TRIS-HCl pH 8.5 (black) and phosphate pH 7.3 (green), both supplemented with ori, $MgCl_2$, and ADP, are exemplarily shown in Figure 24C. Unfortunately, all tested buffers revealed protein aggregation, and no uniform protein distribution could be observed across the grid. However, next to aggregates, hexamers were the dominant protein conformation as observed in Figure 23D already. Since phosphate showed LT stabilization at a broad pH range but did not reduce protein aggregation and led to pronounced protein precipitation at neutral or acidic pH, phosphate was applied only as a stabilizing supplement of the already employed TRIS-HCl buffer system.
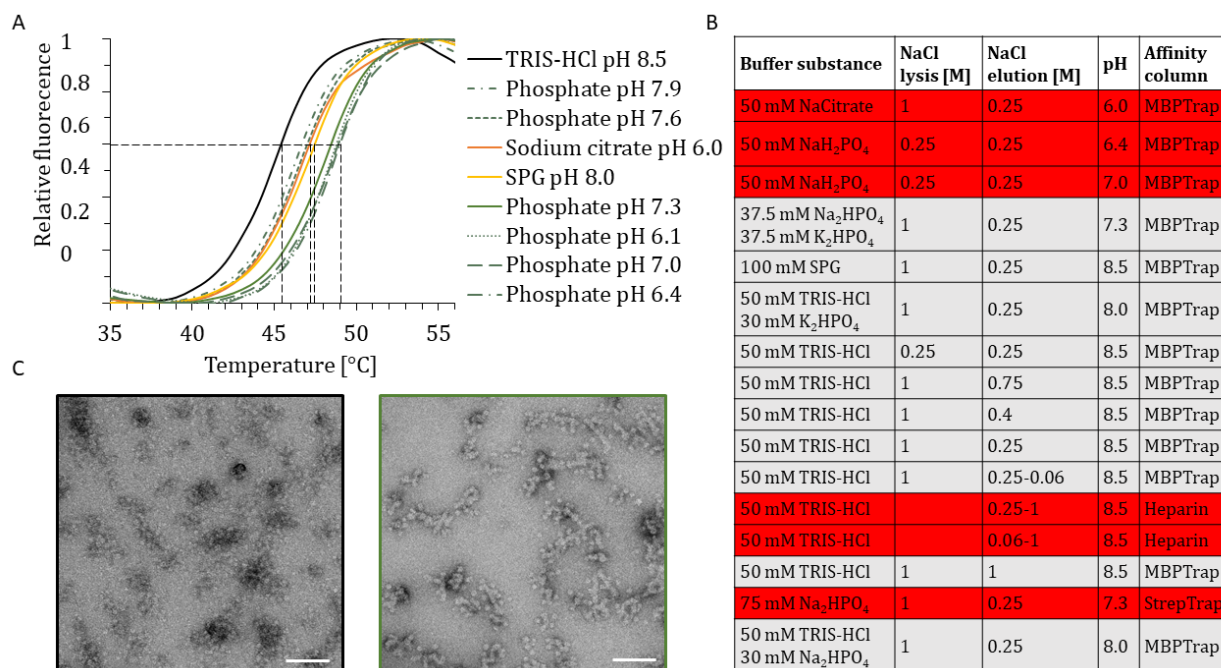
Figure 24: LT buffer optimization. (A) ProteoPlex thermal stability assay. Purified LT was incubated with Sypro orange and buffer, followed by successive heating to 100 °C in a CFX-connect real-time PCR machine. A custom-made buffer screen FEI 88 buffers/ligand screens with 1 M stock solutions was applied (Table S4). The experiments were carried out by Dr. Ashwin Charis' group, Max Planck Institute for Biophysical Chemistry, Göttingen, Germany. Melting temperatures (Tm) were determined at 0.5 relative fluorescence (black dashed lines) for phosphate-based buffers (green), sodium citrate pH 6.0 (orange), succinic acid-phosphate-glycine pH 8.0 (SPG, yellow), and the reference TRIS-HCl pH 8.5 (black). (B) Summary of tested LT purification conditions. Red highlighted conditions failed due to protein precipitation and low protein yield, those in gray led to successful purification and high protein yield after SEC. (C) Negative stain TEM of LT in various buffer conditions. 0.5 µg purified LT+2 mM MgCl$_2$+2 mM AMP-PnP+0.3 pmol ori in TRIS-HCl pH 8.5 (micrograph outlined in black) and phosphate pH 7.3 (micrograph outlined in green) were incubated at 37 °C for 15 min and centrifuged prior staining with uranyl acetate. LT was applied onto a negatively glow-discharged carbon-coated copper grid and recorded with a Talos L120C 120V electron microscope at 92k x magnification. The scale bar represents 100 nm.

Summarized, the experiments above showed a reduced LT polydispersity, by the addition of MgCl$_2$, ADP, and the viral ori, without establishing a monodisperse LT solution. Further, LT purification did not impair specific DNA interaction. The highest binding affinity was observed between LT and the viral ori, whereas no significant difference in LT-DNA interaction between the viral miR-M1, the cellular DNA promoter TP53INP1, and ori scr could be detected. In addition, protein-DNA interaction stabilized LT and increased its thermal stability up to 5 °C by binding the ori. Conformational studies of LT via negative stain TEM revealed a hexamer shape in the presence of all tested oligonucleotides as well as in the absence of DNA. Lastly, it could be shown that phosphate also increased the thermal stability of LT up to 3.6 °C. However, LT purification in phosphate buffer at neutral and acidic pH led to pronounced protein precipitation. As a result, one cryoEM SPA experiment was carried out in the presence of ADP, MgCl$_2$, and the viral ori in TRIS-HCl buffer, and a second was performed under the same conditions, but with phosphate as a stabilizing agent.

6.2.5. Co-expression of LT and ori shows only subtle reduction of LT aggregation

LT aggregation was not fully addressed by changing the buffer system from TRIS-HCl to phosphate or by the addition of $MgCl_2$, ADP, and ori post-purification, as delineated by SEC and negative stain TEM (Figure 21A and Figure 24C, respectively). However, since a pronounced shift in the thermal stability of LT in the presence of the viral ori could be demonstrated, the protein was co-expressed with the ori to stabilize LT immediately after expression. Aggregation, as a consequence of DNA deprivation during cell lysis with high salt buffer, could thus be prevented. The utilized ori sequence was trimmed to a minimum of 66 bp, including site 1/2 with the motif M1-M6, flanked by the 5´- and 3´-AT-rich region (Figure 25D). The early enhancer region M7-10, located downstream of the 3´-AT-rich region, was omitted to reduce LT binding outside site 1/2. This binding was demonstrated for SV40 LT, which could introduce polydispersity (Müller et al., 1987; Cowie & Kamen, 1986). The dispensability of the early enhancer region for LT-ori binding was demonstrated by crystallizing MCPyV LT OBD with an ori sequence comprising only M1-4 (C. Harrison et al., 2011). The AT-rich regions were included since the full-length LT was investigated, which may be stabilized by providing helicase-AT-rich region interaction sites, similar to SV40 LT (Chang et al., 2013). Further, it has been shown that the SV40 core ori flanking regions enhanced LT-ori binding affinity (Müller et al., 1987). LT and the trimmed ori were cloned into independent vectors and expressed and distributed, respectively, in High Five™ insect cells via the baculovirus expression system. Purification was carried out 3 d.p.t. without using high salt lysis buffer, which would disrupt LT-DNA interaction. Ori distribution was verified by extracting DNA out of co-transduced High Five™ cells with the DNeasy® Blood & Tissue Kit and PCR, using primers binding in the bacmid backbone and the ori (Figure 25A). Purification by MBPTrap and SEC were successful, showing a similar LT elution profile compared to LT alone (chromatogram not shown). SDS-PAGE showed concentrated LT, even though not all impurities could be removed (Figure 25B). Analysis by negative stain TEM revealed only subtle aggregation reduction and increased occurrence of single hexamers compared to LT purified with high salt buffer and subsequent ori addition (Figure 25C).

Summarized, LT, co-expressed with the trimmed ori, was successfully purified. However, impurities could not be completely removed and only a minor reduction of LT aggregation was observed. Hence, the co-expression of LT and ori was not further pursued.
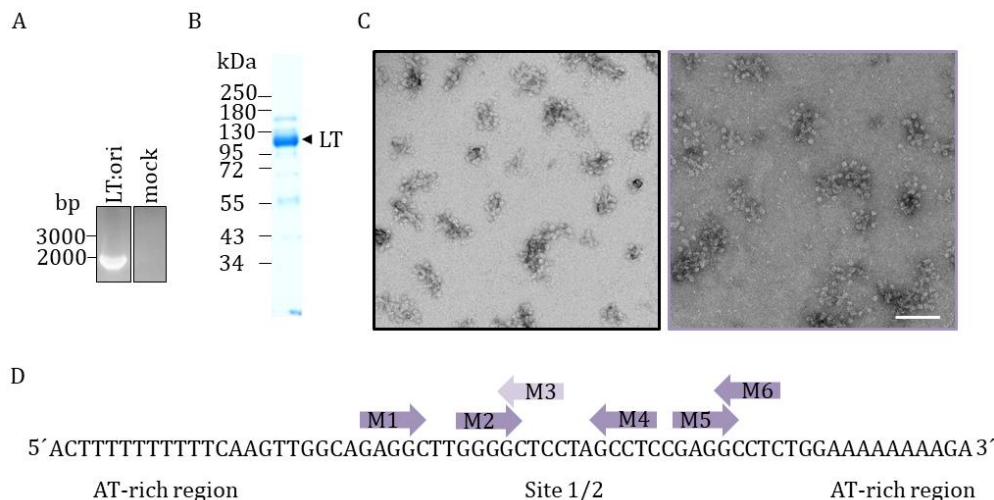
Figure 25: Co-expression of LT and ori. (A) Verification of ori distribution in co-transduced High Five™ cells. DNA was extracted out of 3.3 x 10^6 cells with the DNeasy® Blood & Tissue Kit, followed by PCR using the primer pUC/M13+ and MCPyV ori-, binding in the bacmid backbone and ori, respectively (Table 9). (B) Coomassie-stained SDS-PAGE gel of purified LT. 12.5 x 10^7 cells were co-transduced with LT and ori and purified 3 d.p.t.. Cell lysis was performed in LB2, MBPTrap in EB5 and EB1, and SEC in EB5 (Table 5). (C) Negative stain TEM of LT purified with and without ori. LT, co-transduced with ori and purified as described in (B) (purple), and LT, lysed in high salt buffer (LB1, Table 5), supplemented with 0.3 pmol ori post-purification (black), were analyzed in B1 (Table 5). Samples were incubated at 37 °C for 15 min and centrifuged prior to staining with uranyl acetate on a negatively glow-discharged carbon-coated copper grid. Micrographs were recorded on a Talos L120C 120V electron microscope at 92k x magnification. The scale bar represents 100 nm. (D) The trimmed ori sequence used for the LT:ori co-expression. The ori had a length of 66 bp, including site 1/2 and the 5´- and 3´-AT-rich region. Intact GRGGC motifs (M1, 2, 4-6, dark purple) and the M3 with one mismatch (light purple) are depicted above each oligonucleotide in sense and antisense (right-pointing arrow and left-pointing arrow, respectively).

6.3. Structural characterization of LT

6.3.1. CryoEM SPA of LT reveals helicase-zinc finger model

CryoEM SPA was performed after LT characterization and buffer optimization. Two cryoEM SPA experiments were carried out, both with ADP/AMP-PnP, MgCl$_2$, and the viral ori. Since the addition of phosphate showed increased LT thermal stability, but did not result in reduced protein aggregation, cyroEM SPA was performed in TRIS-HCl alone (A) and in TRIS-HCl supplemented with phosphate (B). Additionally, high salt lysis was omitted in A to increase LT stability via unspecific DNA binding during purification. The analysis of A resulted in a final 3D model, whereas approach B led to noisy 2D class averages which could not be further analyzed. Therefore, only the results of approach A are shown in the following. Purified LT with additives was incubated at 37 °C for 15 min and blotted on holey carbon copper grids at 4 °C, prior to vitrification. 1682 micrographs, each consisting of 37 frames, were recorded on a Talos Arctica 200kV electron microscope (Table 1) and analyzed with Relion 3.1 (Table 15). Parameters for data collection and cryoEM map reconstruction are summarized in (Table 27).

Table 27: Parameters for data collection and cryoEM map reconstruction.

Parameter	Value
Microscope	Talos Arctica 200kV with field emission gun
Detector	Falcon 3EC direct detector
Accelerating voltage	200 kV
Total dose	42 e-/Å^2
Exposure time	40 s
Pixel size	0.96 Å
Nominal negative defocus range	0.6-2.9 µm
Nominal magnification	150k x
Number of frames per micrograph	37
Number of micrographs	1682
Particles used for 2D classification	332,321
Particles used for 3D classification	18,442
Particles comprised in the final 3D model	11,981
Symmetry group	C6
Final resolution ($FSC_{0.143}$)	6.76 Å
Final resolution ($FSC_{0.5}$)	8.28 Å

Micrographs were corrected for beam-induced motion by MotionCor2 1.3 (Zheng et al., 2017). Exemplarily, a corrected micrograph with its corresponding contrast transfer function (CTF) fit of the 2D power spectrum, performed with CTFFIND4.1 (Rohou & Grigorieff, 2015), is shown in Figure 26A. Next to single particles, large protein aggregates and ice contaminations were observed. However, an average of 196 particles per micrograph and a total number of 332,321 particles were picked, based on reference-free auto-picking with Laplacian of Gaussian (LoG) filtering. Repeated 2D classification, thereby discarding low-quality particles, resulted in 16 distinct 2D classes, consisting of 18,442 particles (Figure 26B). These classes revealed a star-shape protein conformation in various views, predominantly close to top views. The conformation comprised two different-sized hexamers, arranged atop of each other, which were visible in the tilted classes. Unfortunately, no profile view was observed, and top and bottom views could not be distinguished. An initial 3D model was reconstructed out of the 2D classes and used as a reference for 3D classification. Further particle discarding and alignment refinement resulted in five 3D classes, whose 2D representations are depicted in Figure 26C. Classes were sorted after class distribution ranging from 63.8% (class 1) to 2.7% (class 5). 3D class comparison revealed a pronounced heterogeneity between the reconstructions. Class 1 (red frame), including most particles and showing high translation and

rotation accuracy of 3.28 ° and 1.53 Å, respectively, was used for structure refinement. B-factor sharpening, FSC calculation, and the application of C6 symmetry yielded the final model with a resolution of 6.8/8.3 Å ($FSC_{0.143}$/$FSC_{0.5}$) and dimensions of 140 x 80 Å, including 11,981 particles. The FSC curve, providing the final resolutions, is given in Figure S 5. The angular particle distribution of the model showed preferred orientation of the top views from the C-terminus, while profile views were underrepresented or absent (Figure 26D). Further, the local resolution ($FSC_{0.143}$) varied asymmetrically within the model and showed values up to 5.7 Å in its interior, which decreased towards its periphery. The C-terminus reached a resolution of only 7.9 Å and was, together with an N-terminal region of 7.4 Å, the least resolved part of the model (Figure 26E).
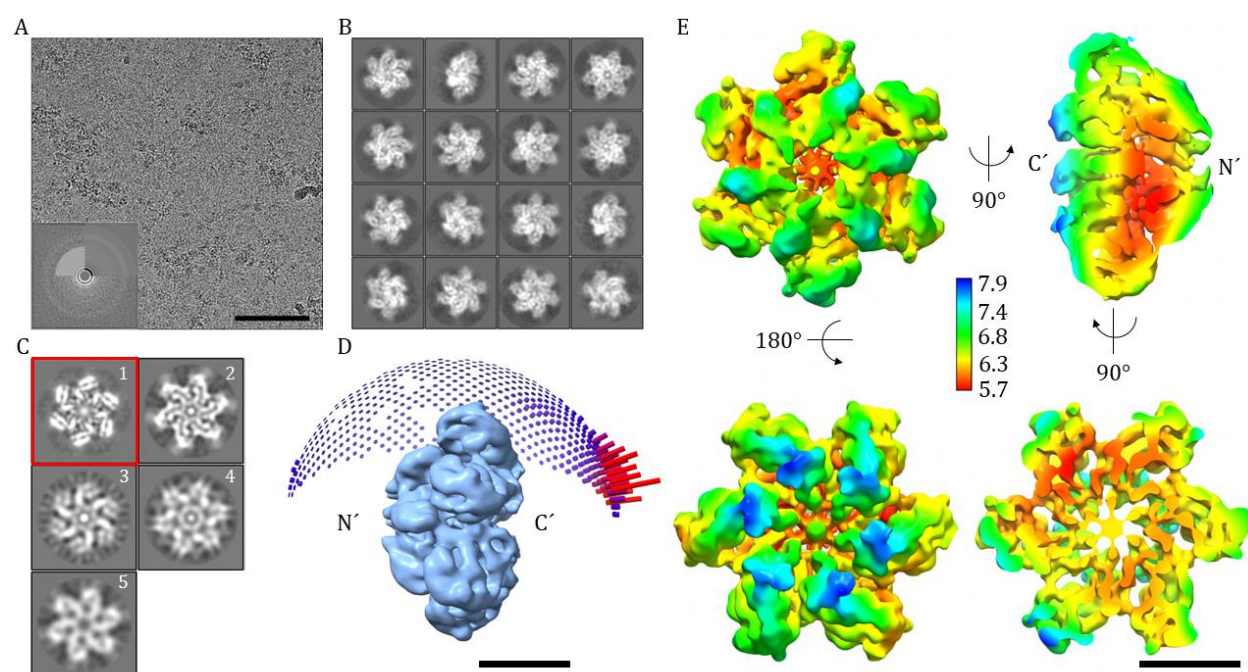
Figure 26: CryoEM SPA of LT. 0.8 mg/ml LT, lysed and purified in LB3 and EB4 (Table 5), respectively, was incubated with 4 mM ADP,2 mM $MgCl_2$, and 0.6 pmol/µg protein at 37°C for 15 min. The sample was applied onto holey carbon-coated copper grids, plotted, and plunge frozen. 1682 micrographs were recorded on a Talos Arctica 200kV using EPU software and subsequently analyzed with the software Relion 3.1. (A) Motion-corrected micrograph with its corresponding CTF-fitted power spectrum. Rotational average (upper right quadrant) and Tone rings approximation (upper left quadrant) were depicted within the Fourier spectrum of the image (lower quadrants). The scale bar represents 100 nm. (B) 2D averages used for 3D classification, comprising 332,321 particles. (C) 2D representations of 3D classes from the final classification iteration, comprising 18,442 particles. Classes are sorted according to the class distribution with 63.8%, 20.7%, 7.2%, 5.5%, and 2.7% for class 1-5, respectively. (D) Angular particle distribution of class 1 depicted in (C). The model represents the final 3D reconstruction with 11,981 particles and underwent 3D refinement, utilizing gold-standard Fourier shell correlation (FSC) calculation. Low angular particle distribution is depicted by blue points and high angular particle distribution with red columns. The C- and N-terminus are marked with C′ and N′, and the scale bar represents 50 Å. (E) Local resolution of the 3D reconstruction shown in (D). Resolution ($FSC_{0.143}$) of the final model ranged from 5.7 Å (red) to 7.9 Å (blue) and was depicted for the N-terminal top view (upper left), the cross-section of the profile view (upper right), and the C-terminal bottom view (lower right) with its corresponding cross-section (lower left). The C- and N-terminus are marked with C′ and N′. The estimation was performed with the software ResMap. The scale bar represents 50 Å.

Since no atomic structure of any MCPyV LT domain in its hexameric form exists, and only a dimeric form of the SV40 helicase-zinc finger-OBD is published (Y. P. Chang et al., 2013), AlphaFold2 prediction of the MCPyV LT$_{316-817}$ homo-hexamer was performed. The prediction, representing the OBD, zinc finger, and helicase, was compared to the cryoEM model to identify LT domains comprised in the model. The prediction showed overall high confidence with pLDDT values up to 97.2 (Figure 27A). However, residues of the DNA-interacting β-hairpin (aa679-683), the N- (aa327-328) and C-terminus (aa799-817), the zinc finger-OBD interdomain region (aa427-328), and a region within the zinc finger (aa479-481), showed low prediction confidence with pLDDT values below 60. Especially the C-terminus, with pLDDT values of 15, was not predictable and considered unstructured, as shown in Figure 4. Similar to other SF3 helicases, the center of the hexamer comprised a hollow tube where DNA unwinding takes place. Superposition of the LT cryoEM model and the LT prediction revealed a similar shape and dimension to the helicase-zinc finger (Figure 27B). Single alpha helices of the helicase-zinc finger could be fitted into the electron density map. Interestingly, also the helicase DNA-interacting β-hairpins (aa679-683), predicted with low confidence, could be aligned to the cryoEM SPA model and coupled the centrally located DNA to the helicase body. Hitherto, the obtained model is the first PyV hexamer helicase model with bound DNA. Unfortunately, the OBD and the C-terminus could not be aligned to the model. Further, the cryoEM model was superimposed to the crystal structure of the SV40 LT helicase-zinc finger hexamer in various nucleotide binding states (Gai, Zhao, et al., 2004) to verify the AlphaFold2 prediction-based results with an experimentally obtained protein structure (Figure 27C-E). High shape and dimension similarities were also observed for these superpositions, which enqueues with high structure similarity of SV40 and MCPyV LT OBD shown in Figure 16B. Analyzing the superposition with SV40 LT bound to ATP revealed an overlap between the cryoEM model and the helicase DNA-interacting loops (aa510-515), protruding towards the DNA density, similar to Figure 27B (Figure 27C). The distance between the DNA density and the helicase DNA-interacting loops increased for SV40 LT bound to ADP, reflecting the condition used for the cryoEM SPA (Figure 27D). Between the superposition of the ADP-bound and the apo state, no pronounced differences could be detected (Figure 27E).

Taken together, the results provided strong evidence that the model obtained by cryoEM SPA represents the MCPyV LT helicase and zinc finger, holding high similarities to the SV40 analog. For the first time, a model of the MCPyV LT helicase-zinc finger could be obtained, paving the road towards structure-based antiviral drug development. Additionally, it is the first PyV hexamer helicase model with bound DNA.
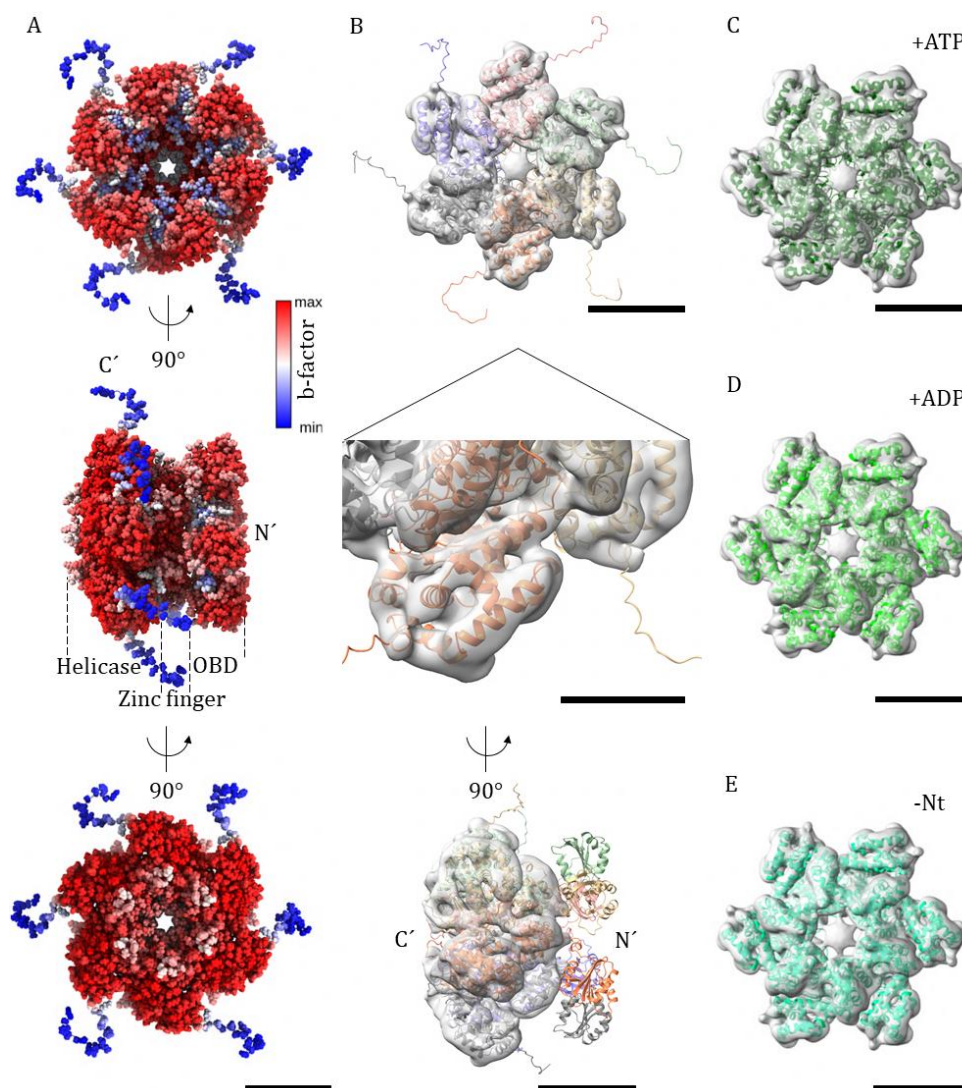
Figure 27 Structure superposition of the LT cryoEM SPA model. (A) AlphaFold2 prediction of the MCPyV LT$_{316-817}$ homo-hexamer. The atomic sphere representation included the OBD, zinc finger, and helicase (dashed lines), the N- and C-terminus are marked with N´ and C´, and pLDDT is color-coded, showing high and low prediction confidence in red and blue, respectively. Depicted is the N-terminal top (top), the profile (middle), and the C-terminal bottom view (bottom). The scale bar represents 50 Å. (B) Superposition of the cryoEM LT model (Figure 26) with the LT prediction of (A). Predicted LT$_{316-817}$ monomers were color-coded (green, red, blue, gray, orange, and yellow) and the cryoEM map is shown in transparent gray. Depicted is the profile-view without OBD (top), including a scale bar of 50 Å, and a magnified section (middle), showing a scale bar of 25 Å. The profile view, including the OBD, is shown at the bottom with a scale bar of 50 Å. The N- and C-terminus are marked with N´ and C´, respectively. (C-E) Superposition of the cryoEM LT model with the SV40 LT$_{251-627}$ homo-hexamer crystal structure bound to ATP (C), ADP (D), or nucleotide-free (E). The top views of SV40 LT bound to ATP (PDB:1SVM), ADP (PDB:1SVL), and without nucleotide (PDB:1SVO) are color coded in shades of green, the cryoEM model in transparent gray. Bound nucleotides are depicted in stick style. The scale bar represents 50 Å.

7. Discussion

MCPyV was discovered in 2008 and has a seroprevalence of 60-70% in healthy adults (Paulson et al., 2018; Garneski et al., 2009; Carter et al., 2009; Pastrana et al., 2009). Upon immune suppression or old age, it causes the rare but highly aggressive skin cancer, MCC (Shuda et al., 2008). It is the only hPyV driving tumorigenesis in its host. Approximately 80% of all MCC are viral-positive with a low mutational burden (H. Feng et al., 2008; Goh et al., 2015). VnMCCs are mainly induced by UV-irradiation, revealing a high rate of mutations, including those in the well-known tumor suppressors pRb and p53 (Goh et al., 2015; Harms et al., 2015; Erstad & Cusack, 2014). In vpMCC, the T antigens, sT and tLT, are the predominant factor of tumorigenesis (Wendzicki et al., 2015). Contrarily to other PyVs, MCPyV sT predominantly impacts tumorigenesis, as it was demonstrated to be sufficient to transform rodent fibroblasts and to induce MCC-like tumors in p53 deficient transgenic mice (Verhaegen et al., 2022; Shuda et al., 2011). However, tLT, which is solely expressed in MCC, is essential for tumor maintenance. It could be shown that the tLT-pRB interaction drives MCC tumor cells into the S-phase, causing their proliferation (Houben et al., 2012). Next to C-terminal LT truncations, viral genome integration is observed during MCC initiation (H. Feng et al., 2008; Shuda et al., 2008). However, the underlying mechanisms are still elusive. The full-length LT is essential for viral DNA replication and viral transcription (An et al., 2012). It recognizes the viral ori via its OBD and unwinds the genome via its C-terminal helicase (C. Harrison et al., 2011; Kwun et al., 2009). To date, only restricted knowledge of PyV T antigen structures is available since no full-length LT structure of any PyV is known, and only the structure of the SV40 sT could be solved (Cho et al., 2007). For MCPyV, solely its OBD could be elucidated (C. Harrison et al., 2011), although its full-length structure is of particular interest due to its MURs, which are absent in other PyVs such as SV40. The lack of specific antiviral treatment and the low 5-year survival rate of 62-25% urges the development of new drug targets (Agelli et al., 2010; Albores-Saavedra et al., 2010). Since the T antigens play a vital role in the viral life cycle and the clinical manifestation, they serve as ideal targets for therapeutic treatment of MCC and preventive treatment for solid-organ transplant recipients. Their atomic structure can thereby serve as starting point for drug development. Additionally, insights into the function and molecular mechanisms of action can be gained, which would further help to understand the development and maintenance of MCC.

In this work, two protein structure elucidation methods, X-ray crystallography of *in cellulo* grown crystals and cryoEM SPA, were used to approach the atomic structure of the MCPyV T antigens. A biochemical characterization of LT preceded its TEM analysis and gave insights into its stability, DNA-binding behavior, and oligomeric state. This data was used for the cryoEM SPA of LT, resulting in the

first model of the LT helicase hexamer with a resolution of 6.8 Å ($FSC_{0.143}$). The obtained results will be discussed in the following.

7.1. *In cellulo* crystallization

7.1.1. Potential post translational modifications alter the molecular weight of tLT244

In cellulo crystallization of the T antigens was used to circumvent the challenges of protein purification, concentration, and *in vitro* crystallization. This was necessary since *in vitro* crystallization of MCC derived tLT244 failed in a previous PhD thesis due to low protein concentration and polydispersity (Borchert, 2013). The propensity of *in cellulo* crystallization was increased by using a proteinogenic crystalline molecular flask, consisting of the catalytic site of the metazoan-specific kinase PAK4 and its endogenous inhibitor Inka1 (Baskaran et al., 2015; Inokuma et al., 2011). T antigen variants were linked to Inka1 and co-expressed with PAK4cat, using the baculovirus expression system. The successful protein expression of all constructs could be verified. However, especially Inka1-linked tLT244, comprising the DnaJ domain, the entire MURI and the first part of the MURII, exhibited an increased molecular weight displayed by WB. It was detected at a size of 72-95 kDa, which is approximately 20 kDa above the theoretical molecular weight of 58 kDa (Figure 13B). The increased size could be linked to bulky PTMs, such as ubiquitination, which has been described for MCPyV LT. The protein contains 82 predicted phosphorylation sites, including 15 minimum phosphodegron sites for SCF E3 ligases ubiquitination (Kwun et al., 2017). Those recognition sites are primarily located within the MURs and could therefore significantly affect the size of tLT244 by mediating the coupling of the 8.5 kDa sized ubiquitin. Alongside, LT C- and N-terminal interactions with the cellular ubiquitin degrading enzyme Usp7 verifies LT-bound ubiquitin (Czech-Sioli et al., 2019). A reduced stability of MCPyV LT ($t_{1/2}$ of 4-6 h) compared to SV40 ($t_{1/2}$ of >36 h), displaying predominantly acetylation rather than ubiquitination, also argues for a dominant MCPyV LT ubiquitination (Ortiz et al., 2021; Shimazu et al., 2006). A second bulky PTM comprises glycosylation, and it is known that S111 and S112 of SV40 LT are O-glycosylated (Medina et al., 1998). It is conceivable that MCPyV LT is altered by similar glycosylation since the content of serine (26%) and threonine (8%), the receptor amino acids of O-glycosylation, within the MURs is remarkably high. Combining the data from SV40 and MCPyV, it can be hypothesized that MCPyV LT exhibits ubiquitination and O-glycosylation as PTM, which increase its size by approximately 20 kDa.

7.1.2. Improving *in cellulo* crystal quality and growth efficiency

Single, elongated spicules, crystal bundles, or radially grown crystals with a size of 10-100 nm were observed in HEK293/COS7 cells, 3 d.p.t. (Figure 13D). The incorporation of tLT244, sT, and 122tsT into the crystalline matrix was verified by IF (Figure 14), and the crystallinity of sT was examined via TEM of ultra-thin sections of resin-embedded HEK293 cells (Figure 15). Although the dimensions, IF staining, and crystallinity of some of the T antigen-harboring crystals were comparable to those of Inka1:PAK4cat alone, no analyzable diffraction data could be recorded. It is hypothesized that the integration of the T antigens into the crystal cavity may reduce crystal quality and obstruct lattice formation, leading to low or no diffraction. A first indication for this hypothesis provided the morphology of the 122tsT integrated crystals. Brightfield and IF analysis revealed a radial growth, suggesting an altered nucleation process compared to the exclusively as single or in bundles grown crystals of the positive control (Figure 13D 3,6, and 7; Figure 14C and D). Those cluster crystals could indicate a low probability of 3D nucleation, leading to a 2D nuclear growth from another crystals surface (Hashizume et al., 2020). Lowering the degree of supersaturation was a proposed solution to increase crystal quality in *in vitro* crystallization. Reducing protein expression could also be adopted for the *in cellulo* scenario, which will be discussed below. Next to structure elucidation, the crystalline matrix of Inka1:PAK4cat was used to engineer a magnetic protein crystal by incorporating a ferritin subunit into the solvent channel (T. L. Li et al., 2019). Similar to 122tsT crystals, ferritin inclusion changed the crystal morphology, leading to shorter crystals with a tapered end together with a slower nucleation rate. It is conceivable that in addition to morphological and growth behavior, the crystal quality might also be affected by the incorporation of ferritin, which was not analyzed in this study. These observations provided further evidence that the incorporation of proteins into the Inka1:PAK4cat cavity can alter the crystal, potentially impeding diffraction. TEM analysis of sT-Inka1:PAK4cat crystals provided an additional indication for the reduction of crystal quality upon the integration of proteins of interest. Next to the positive control-resembling crystals with a hexagonal shape and a clear crystal lattice, crystals without hexagonality and with frayed edges were observed. This indicates insufficient crystal packaging which lowers diffraction. In *in vitro* crystallization experiments, crystal packaging is often improved by increasing the intermolecular force. This is realized by changing the pH, reducing the counter ion, or adding metal ions (Hashizume et al., 2020). However, the restricted possibility to vary the crystallization environment is a major disadvantage in *in cellulo* crystallization and makes it impossible to improve the crystal packaging of the T antigen crystals by changing the buffer conditions. Another approach is to reduce the crystal growth rate. It could be shown that high crystal growth rates lead to low crystal quality, as single molecules bind too fast to the growing crystal, resulting in a non-productive configuration, which disables ordered crystal

growth (Schmit & Dill, 2012). The protein expression rate is one of the few things, which is adjustable in *in cellulo* crystallization. Together with reducing the transfected DNA, substitution of the strong cytomegalovirus immediate-early (CMV) promoter with the weak human ubiquitin C (UBC) promoter could improve the crystal packaging of T antigen-incorporated Inka1:PAK4cat crystals (Qin et al., 2010). Despite all improvements, it is still unknown if the idea of a proteinogenic crystalline molecular flask is viable. Hitherto, neither the proof of principle-serving GFP-incorporated crystals, nor other examples of the system were analyzed via synchrotron radiation, wherefore a proof of principle is still pending (Baskaran et al., 2015). However, sT-incorporated crystals with an ordered lattice and perfect hexagonality were also observed, demonstrating that ordered crystal packaging is possible in the presence of T antigens. This suggests additional factors for the insufficient data collection such as a low crystal hit-rate. Koiwai and colleagues managed to overcome the low hit-rate by increasing crystal production of the human lysosomal neuraminidase-1 (NEU-1). They evaluated five mammalian cell lines, thereby detecting a 2-fold greater efficiency in HEK293FT cells over HeLa cells (2019). Additionally, they re-transfected NEU-1-overexpressing cells with the same protein construct, yielding a 6-fold crystal increase. In this work, HEK293, HeLa, COS7, and High Five™ insect cells were screened for crystal production, revealing the highest efficiency within HEK293 and COS7 cells. Re-transfection to improve growth efficiency was not validated and needs to be tested. Using truncated T antigens to increase the crystal growth efficiency, by reducing the possibility of spacious clashes of the protein of interest within the solvent channel of 80 Å, was not successful, as stated above. Attempts to express the crystalline matrix together with the T antigens as a multi-domain protein to maximize the possibility of protein interaction did not result in crystals in either mammalian or insect cells. Since iBox-PAK4cat alone and the co-expression of Inka1-T antigen:PAK4cat formed crystals (Table S 2), it is conceivable that the linker between the T antigen and the iBox was not flexible enough to enable crystal formation. Extending this linker to more than two amino acids could enable crystal formation, which are potentially increased in number. Another method to increase crystal formation was reported by Hasegawa et al., who connected crystal size and number with cell size (2019). Upon chemical induced cell enlargement by inhibiting cell division, using the topoisomerase II inhibitor dexrazoxane, they could show that the number per cell of NEU-1 crystals and the size of crystalline gamma D (CRYGD) and Charcot-Leyden crystals (CLCs) increased remarkably. However, a careful balancing between enhancing crystal growth to increase the crystal-hit-rate and decelerating crystal growth to improving crystal quality is required. In summary, a plethora of improvements in crystal quality and crystal abundance are available. However, whether the application of a crystalline matrix to improve *in cellulo* crystal production is constructive remains to be validated.

7.2.Limitations of the LT prediction by AlphaFold2

The demand to predict the sequence-structure function has been present for decades, which is illustrated by the low number of approximately 200,000 protein structures deposited in the PDB (June 2023). Alone the human genome encodes approximately 20,000 non-modified proteins, which increases to approximately 2,000,000 if alternative splicing, single amino acid morphisms, and PTMs are taken into account (Ponomarenko et al., 2016). Protein structure prediction was long time only possible via homology modeling or de novo modeling (Waterhouse et al., 2018; Rohl et al., 2004). These methods require sequence similarity between the template and the target protein or are only accurate for small proteins, ascribed to the fast conformational space and computational power limitations (Fiser, 2010; D. E. Kim et al., 2009). Since 2021, the neuronal network-based program AlphaFold2 can be used for accurate protein structure prediction (Jumper et al., 2021). For example, the structural coverage of the human proteome was increased from 48% to 76% by AlphaFold2 (McCafferty et al., 2023; Tunyasuvunakool et al., 2021), and more than 2,000,000 predicted protein structures are deposited in the AlphaFold2 Protein Structure Database (Varadi et al., 2022). However, neuronal *in silico* predictions have their limitations and cannot substitute experimental structure elucidating methods, such as cryoEM SPA, X-ray crystallography, or NMR spectroscopy. For example, AlphaFold2 is challenged by protein dynamics (Fowler & Williamson, 2022) and IDRs (Ruff & Pappu, 2021). The single domains of MCPyV LT could be predicted with high confidence, while the MURs depicted pLDDT values below 50 (Figure 16D). Low pLDDT values correlate often with IDRs, which was confirmed by the IDR prediction program IUPred3 (Figure 4B). It is known that 70% of all cancer associated proteins harbor IDRs (Iakoucheva et al., 2002). It is therefore hypothesized that the MURs, which are unique for the only human tumorigenic PyV, play a substantial role in tumorigenesis. However, due to their conformational heterogeneity, which is encoded in their amino acid sequence (Holehouse et al., 2017; Das et al., 2015), AlphaFold2 provided no information on their three-dimensional structure. It would be of enormous interest to analyze their context specific folding, depending on the binding partner and the cellular environment such as the pH (P. E. Wright & Dyson, 2009). An example for the adaptation of various conformations upon the binding of different proteins provides p53. 71% of its interactions are ascribed to the N- and C-terminal IDR. Seven core amino acids within the C-terminus were shown to transform into either a helix, a β-sheet, or a coil upon binding of four different proteins (Oldfield et al., 2008).

Further, the lack of predicted MURs conformations disables docking experiments to identify small molecules for the development of antiviral drugs. This is of particular interest as targeting the MURs would provide high specificity and reduced off-target effects. It is also impossible to conclude on the R_H of the protein, based on the size of the conformational cloud. However, a reliable R_H prediction

would have been beneficial for the DLS analysis (Figure 22), which will be discussed in 7.6.1. Unfortunately, it has been shown by small angle X-ray scattering (SAXS) that AlphaFold2 generally overestimates the R_H (Ruff & Pappu, 2021). Likewise, the relative position and orientation of all LT domains are unpredictable (Figure 16E). Consequently, interdomain interactions or insights on how the IDRs enable multivalent interactions cannot be concluded from the LT prediction. Interestingly, the predicted aligned error estimates between the helicase and the OBD are lower compared to those of the DnaJ-helicase and the DnaJ-OBD. The only crystal structure comprising the OBD and the helicase revealed the involvement of both domains in the ori recognition and assembly (Y. P. Chang et al., 2013) (Figure 5C). It is therefore not surprising that their cooperative action results in a higher confidence in their relative position and orientation to each other. This is contrary to the LT domains, which are separated by the MURs, acting as large flexible linker.

In summary, despite the high accuracy of single LT domains, the AlphaFold2 predicted structure still leaves questions unanswered. These include the three-dimensional structure of the IDRs, the relative positions of the LT domains, and the associated imprecision of the R_H. Nevertheless, valuable insights into the single LT domains could be gained. Additionally, the predicted $LT_{316-817}$ homo-hexamer structure impacted the helicase identification within the cryoEM SPA model (Figure 27A and B).

7.3. LT is potentially involved in liquid-liquid phase separation via its MURs

Despite the low confidence of the LT MURs structure prediction via AlphaFold2, making it impossible to follow the biochemical dogma "structure determines function", presumptions on their mechanism of action can be made. AlphaFold2 and IUPred3 defined the MURs as IDRs (7.2), which are often described to be key elements in driving liquid-liquid phase separation (LLPS) (Uversky, 2017). LLPS compartmentalizes proteins and nucleic acids spatiotemporally into liquid-like, membraneless biomolecular condensates with various functions. It is involved in transcription control (Sabari et al., 2018; Hnisz et al., 2017), genome organization (Gibson et al., 2019), and in the innate immune response (B. Wang et al., 2021; Du & Chen, 2018) by providing a reaction vessel, enrichment, localization, or inactivation (J. Li et al., 2022; Alberti, 2017). Prominent examples display nucleoli and promyelocytic leukemia (PML) bodies in the nucleus and cytoplasmic stress granules (Banani et al., 2017). Molecular condensation is driven by weak, multivalent interactions between macromolecules. Next to IDR mediated homotypic interactions, heterotypic interaction with nucleic acids and multivalent proteins can act as trigger to induce LLPS (Caragliano et al., 2022; Y. Lin et al., 2015; P. Li et al., 2012). In recent years, many studies showed that dsDNA viruses induce LLPS throughout their life cycle. For example, LLPS could be associated with the formation of KSHV LANA-nuclear bodies (Vladimirova et al., 2021). The EBV proteins EBNA2 and EBNALP, which are IDRs harboring

transcription factors, modulated the EBV-host interaction by inducing LLPS at superenhancer sites to control host gene expression (Peng et al., 2020). EBNA2-induced LLPS also regulated cancer-associated alternative RNA splicing patterns (Peng et al., 2021). Last, HSV-1 and HCMV replication centers reveal liquid-like characteristics, conferred by the HSV-1 ICP4 and the HCMV UL112-113, both harboring IDRs and binding DNA (Caragliano et al., 2022; Seyffert et al., 2021). Interestingly, the HPV E2 protein, holding selected functions of LT, such as viral ori recognition and transcription regulation, and which recruits the viral helicase E1 onto the genome, was reported to be involved in LLPS by p53-binding (Borkosky et al., 2022; McBride, 2013). The heterotypic condensation was modulated by the C-terminal DNA-binding domain of E2, which is structurally similar to the LT OBD with an RMSD value of 3.1 Å between the PDB structures 3QFQ and 2AYB. LT harbors next to the OBD the intrinsically disordered MURs, which are both protein domain types known to induce LLPS. It is therefore highly conceivable that LT does not only modulate LLPS, similar to E2, but induces LLPS by homotypic interactions to drive key functions of MCPyVs. Alongside, it is known that PTMs play an essential role in LLPS by affecting the driving forces of LLPS, including hydrophilic and hydrophobic forces, electrostatic interactions, and non-covalent bonds (J. Li et al., 2022). It was demonstrated that self-assembly of the fused in sarcoma (FUS) protein was restricted upon phosphorylation, resulting in LLPS disruption (Monahan et al., 2017; Shorter, 2017). Ubiquitination was also reported do influence LLPS by its multivalent binding potential, which has been studied in the formation of the proteasome (Yasuda et al., 2020; Dao et al., 2018). LT modification by PTMs such as ubiquitination and phosphorylation, which was partially discussed in 7.1.1 already, provided another indication that LT is involved in LLPS. A possible LLPS participation could be in the formation of viral replication centers, similar to HSV-1 and HCMV (Caragliano et al., 2022; Seyffert et al., 2021), while cancer association is also possible, resembling KSHV and EBV (Xie et al., 2023; Vladimirova et al., 2021; Peng et al., 2021, 2020). Intriguingly, LT resembling proteins of other dsDNA viruses such as the HPV E1, harboring a helicase domain and a DNA-binding domain (Bergvall et al., 2013) and the EBV EBNA1, comprising a DNA-binding domain (Frappier, 2012), have not been reported to be involved in LLPS, suggesting that the MURs may be the major factor of the hypothesized LLPS involvement of LT. However, this needs to be experimentally investigated in the future.

7.4. The salt concentration strongly influences LT stability

Protein purification often includes multiple steps of optimization since every protein behaves differently, which needs to be experimentally examined. The ability of LT to bind DNA requires special considerations since DNA stabilizes LT, demonstrated by the co-crystal of the DNA-bound LT OBD (C. Harrison et al., 2011). At the same time, unspecific DNA binding contaminates the purified protein. The second challenge are the LT MURs, which are predicted as IDRs, tending to self-interact and to aggregate (Figure 4B, 7.2, 7.3). Together, these characteristics make it extremely difficult to purify LT in a monodispersed, soluble, and pure form.

Cell lysis precedes protein purification and is of special interest since LT harbors two NLS signals and resides in the nucleus. Three different cell lysis methods were examined to determine high cell lysis efficiency, while maintaining sample integrity: (1) the removal of the cytosolic fraction prior nucleus disruption and 420 mM salt; (2) the combination of 1 M NaCl, 0.1% Triton X-100, and sonication; (3) 1 M NaCl with 1% Nonidet P-40; The combination of chemical and physical disruption (2) showed the highest protein yield and the lowest DNA contamination and was used for following experiments. Intriguingly, (1) provoked high molecular weight structures, pronounced protein precipitation, and irreversibly adsorbed to the resin (Figure 18). It is conceivable that the reduced salt concentration of (1), compared to (2) and (3), caused these effects. Salt inhibits DNA-protein interaction but stabilizes proteins via the salting-in phenomenon, depending on the protein of interest, the pH, and the salt ion (Hofmeister, 1888). 420 mM NaCl could be sufficient to disable LT-DNA interactions but inefficient in stabilizing the unbound LT, resulting in aggregation. For this reason, cell lysis (2) was also performed with 250 mM NaCl next to 1 M NaCl. The stabilizing unspecific DNA interactions are thus supported at the expense of LT purity. The high salt induced withdrawal of unspecific DNA, combined with a potential lack of LT stabilization by reducing the salt concentration to 250 mM for cryoEM SPA, possibly also accounts for the unsuccessful second cryoEM SPA (B) (6.3.1).

The elevated LT amount in (1) could be a result of transduction efficiency variations since the same batch could not be applied for all lysis screens. The experiments were too laborious to be combined in one day, and sample freezing led to protein damage (unpublished results).

LT purification method (3) was also successful but yielded a lower protein concentration. LT could be detected in the cell pellet, indicating that sole chemical lysis with Nonidet P-40 may not be sufficient to disintegrate the nucleus membrane, resulting in an incomplete LT release and reduced protein yield.

In summary, high salt lysis reduces unspecific LT-DNA interaction, while stabilizing the protein. However, withdrawing DNA with medium salt concentration or reducing the salt concentration after cell lysis potentially leads to LT aggregation.

7.5.Principles of LT aggregation and prevention approaches

Protein aggregation poses a challenge during purification, and laborious optimization steps are required to reduce aggregation. The tendency of LT to aggregate was already evident during the cell lysis screening, discussed in 7.4, in which high molecular weight structures could be detected via WB (Figure 18A). Alongside, LT eluted in the void volume of the SEC column, indicating aggregation (Figure 21A), which was verified via negative stain TEM (Figure 24C). Protein aggregation posed a major challenge during cryoEM SPA, wherefore it is essential to analyze its driving force. One approach to reduce aggregation was the co-expression of LT with the viral ori, which was not successful (Figure 25). The presence of the ori within the cell suspension was verified via PCR, 3 d.p.t. (Figure 25A), while it could not be detected via southern blot, utilizing the purified LT:ori complex (results not shown). Unfortunately, no positive control was available to confirm correct handling of the DNA detecting method. Next to the possibility that LT-ori interaction concomitant to protein expression affected LT aggregation only marginally, it is also conceivable that LT-ori interaction did not withstand purification, carried out at 250 mM NaCl. However, since LT, purified in 250 mM NaCl, always revealed strong nucleic acid contaminations with a 260/280 ratio of approximately 1.4 (6.2.3), the inhibition of LT-ori interaction by 250 mM salt is highly unlikely. An alternative explanation is a weak or a complete absence of LT-ori interaction within insect cells. To reduce protein heterogeneity, the GRGGC M1-6 and their flanking AT-rich regions were chosen as ori sequence, while M7-10 were omitted. However, Kwun and colleagues showed that M7, located downstream of the 3´ AT-rich region, profoundly increased LT-ori affinity (2009). It is conceivable that M7 cannot be omitted in the trimmed ori to ensure strong LT-ori interaction.

A second approach to reduce aggregation was to improve the buffer conditions, monitored by the increase of LTs thermal stability (Figure 24A). Intriguingly, LT was stabilized in a broad pH range in phosphate buffer, while its tendency to aggregate was not reduced (Figure 24C). It is conceivable that phosphate mimics the interaction of LT with the negatively charged DNA phosphate backbone, leading to increased LT thermal stability in the absence of DNA. However, DNA may displace stabilizing phosphate ions from the LT-DNA binding pocket, resulting in an unchanged aggregation behavior in the presence of DNA. Alongside, studies of SV40 indicated that LT interactions with the DNA phosphate backbone are more important than those with the DNA bases (SenGupta & Borowiec, 1992). The ability of phosphate to affect protein-DNA interaction was already demonstrated for the *E. coli* DNA-binding protein RecA, in which protein-dsDNA binding was impaired in phosphate buffer, while protein stability was increased (Metrick et al., 2013). However, LT still aggregates in the presence of phosphate and DNA (Figure 24C), indicating that aggregation does not predominantly originate from the DNA-binding domains, OBD and helicase, but from additional domains such as the

MURs. These regions, hypothesized to induce LLPS (7.3), could significantly impact aggregation via their IDRs. IDRs are prone to self-interact and to aggregate, which can result in neurodegenerative diseases such as Alzheimer´s disease (Tsoi et al., 2023). Mutations of LLPS involved amino acids, such as disorder promoting proline residues (Uversky, 2017), which are highly enriched in EBNA2 and EBNALP, were shown to disrupt droplet formation (Peng et al., 2020). The LT MURs have a proline content of 12.8%, which is significantly increased compared to the average proline occurrence rate of approximately 5% (X. Wu et al., 2019; Mittal & Jayaram, 2011). Additionally, serine is a disorder promoting residue (Uversky, 2017), also substantially increased within the LT MURs (25.9%), compared to the average occurrence rate (7-8%) (Mittal & Jayaram, 2011; Tourasse & Li, 2000). Consequently, proline and serine were suggested as auspicious targets to disrupt LLPS. IUPred3 predictions of *in silico* mutated proline (B) and serine (C) residues to alanine within the MURs revealed no pronounced IDR disruption (Figure 28). LT residue mutation for aggregation reduction is therefore not suggested as an experimental setup.
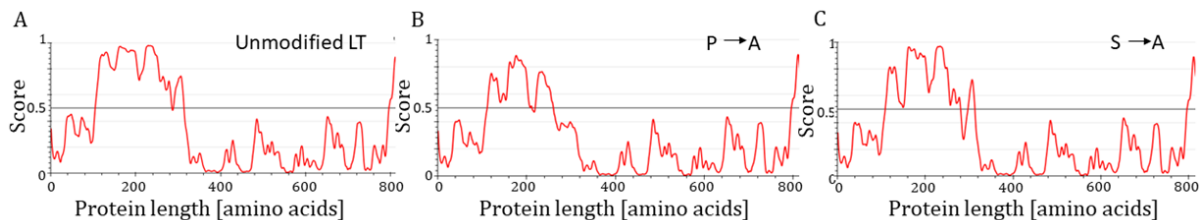
Figure 28: Intrinsically disordered region prediction of LT with mutations within the MURs. Predictions above a threshold of 0.5 are considered intrinsically disordered. Depicted are unmodified LT (A), LT with mutated proline to alanine (P→A) within the MURs (aa73-315) (B), and LT with mutated serine to alanine (S→A) within the MURs. The prediction was performed with IUPred3 (Erdős et al., 2021).

Further, IDR interactions are often electrostatically driven (Mitrea & Kriwacki, 2016) and thus can be influenced by PTMs, as stated in 7.3 already. It could be demonstrated that phosphorylation of the LLPS model protein FUS, mainly composed of IDRs, disrupt FUS induced LLPS and reduce protein aggregation (Monahan et al., 2017). Similar results have been observed by arginine methylation of the protein DEAD-Box helicase 4 (Ddx4) (Nott et al., 2015). Interestingly, the phosphorylation state of SV40 LT was also observed to influence interprotein cooperativity. Hypophosphorylation enables cooperative interactions between two hexamers to form an active double-hexamer, which undergoes remodeling to initiate DNA unwinding. Hyperphosphorylation disrupts the double-hexamer to license bidirectional unwinding (Sowd & Fanning, 2012). For MCPyV LT, phosphorylation was reported to modulate protein-DNA interaction (Diaz et al., 2014), emphasizing the fundamental role of electrostatic interactions in the regulation of PyV LT protein-protein and protein-DNA interaction. It is therefore suggested to alter the charge characteristics of MCPyV LT via PTM modulation to reduce its aggregation tendency. Since LT does not contain many RGG/RG motifs, reported to function as

methylation sites of Ddx4 (Nott et al., 2015), charge modulation via methylation may not be beneficial. However, LT harbors 82 predicted phosphorylation sites (Kwun et al., 2017), which can be used for phosphorylation alteration. For SV40 LT, several kinases were identified to specifically phosphorylate protein residues, such as the cdc2/CDK1 kinase, casein II kinase, and ATM kinase (Y. Shi et al., 2005; Götz et al., 1995; McVey et al., 1989). Similarly, ATM kinase was shown to specifically phosphorylate S816 in MCPyV LT (J. Li et al., 2015). An appropriate kinase for LT phosphorylation needs to be determined, preceding the examination of an aggregation reduction by disrupting self-interacting forces via phosphorylation. Conversely, preliminary experiments of LT dephosphorylation were performed and analyzed via negative stain TEM (Figure 29). LT treatment with alkaline phosphatase (AP) led to reduced aggregation and enhanced formation of single hexamers (green). These preliminary results suggest that the reduction of negative charge, conferred by phosphorylation of uncharged residues, such as serin, threonine, and tyrosine, introduced repulsive forces, leading to reduced protein aggregation. However, this phenomenon has to be verified via DLS or SEC. Additionally, it would be of interest to examine protein aggregation of *E. coli*-expressed LT, which profoundly differs in its PTMs. Bacterial LT expression was shown to be successful for SV40 LT, while losing its ability to activate bidirectional ori unwinding (Sowd & Fanning, 2012).
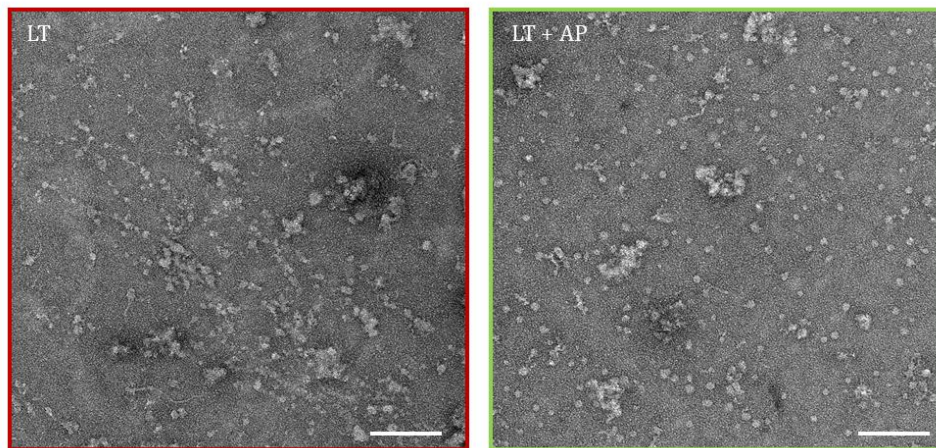
Figure 29: Negative stain TEM of LT with an altered phosphorylation state. 0.5 µg purified LT + 2 mM MgCl$_2$, 4 mM AMP-PnP, and 0.9 pmol ori in TRIS-HCl pH 8.5 (red), and supplemented with 0.07 U alkaline phosphatase (AP) (green) were incubated at 37 °C for 15 min and centrifuged prior staining with uranyl acetate. LT was applied onto a negatively glow-discharged carbon-coated copper grid and recorded with a Talos L120C 120V electron microscope at 92k x magnification. The scale bar represents 100 nm.

An alternative approach to reduce aggregation is the stabilization of LT via binding partners. PRb is known to interact with the LT LxCxE motif, flanked by the MURs (Shuda et al., 2008) (Figure 4A). This interaction could potentially stabilize the unstructured MURs by inducing specific folding upon pRb binding. The existence of an SV40 LT structure bound to the pocket domain of pRb supports the hypothesis (Figure 5A). In the context of a bachelor thesis, which was supervised within this project,

the presumption was challenged by co-expressing MCPyV tLT244 with His-pRb AB Box. The pRb AB Box is the minimal region for E2F interaction and was used to crystallize SV40 LT (H.-Y. Kim et al., 2001). The two proteins were co-expressed in High Five™ insect cells via the baculovirus expression system. Although a pronounced fraction of the proteins could be detected in the pellet, the tLT244:pRb AB Box complex was successfully purified via NiNTA (results not shown). Similar results were reported in a PhD thesis, in which the complex formation of pRb AB Box was examined with *E. coli*-expressed tLT244 (Borchert, 2013). Alongside, Czech-Sioli and colleagues could show that Usp7 binds to LT via multiple P/A/E-x-x-S motifs, predominantly located in the MURs. They further reported an enhanced LT-ori affinity upon LT-Usp7 interaction (Czech-Sioli et al., 2019). This could be of additional value since the LT-stabilizing ori was added for cryoEM SPA (6.3.1). Next steps require the analysis of LT bound to pRb AB Box or Usp7 via DLS, SEC, or negative stain EM to assess whether these protein-protein interactions reduce LT aggregation. Further, it has to be noted that only tLT244 was used for the above stated co-expression experiments, while for cryoEM SPA the stabilization of full-length LT is of interest. It is therefore suggested to repeat the co-expression experiments with full-length LT.

Lastly, the deletion of the MURs could reduce LT aggregation. Unfortunately, contradictory results have been reported on the LT stability upon the deletion of the MURs. Houben and colleagues stated reduced protein stability and growth restriction of MCC cells (2015), while increased LT stability and proliferation of the human BJ-hTERT fibroblasts have been demonstrated by others (Nwogu et al., 2020a). Additionally, also the MURs-harboring chimeric SV40 LT showed increased expression in MCC cells and reduced expression in BJ-hTERT cells. Since the MURs are unique to MCPyV and show no sequence similarity to any other protein, it is conceivable that they are involved in tumorigenesis. The deletion of the MURs would prevent important structural findings about the unique mechanism of action of the MCPyV LT and is therefore no suitable strategy.

In summary, LT aggregation could neither be resolved by the co-expression of LT:ori, nor by the addition of stabilizing phosphate ions. It is hypothesized that the IDRs of the MURs largely account for the LT aggregation, suggesting altering the charge distribution of those regions to disrupt homotypic interactions. Further, the co-expression of the MURs binding partner Usp7, or the LxCxE motif binding protein pRb, could stabilize LT, potentially leading to reduced aggregation.

7.6.Structural characterization of LT

7.6.1. DLS analysis indicates the presence of LT hexamers and aggregates

DLS measurements of LT in the presence and absence of additives revealed that ADP, $MgCl_2$, and ori reduced polydispersity, showing a dominant protein population of 18 nm (Figure 22). The measured R_H argues against the existence of LT monomers in solution, while it is difficult to conclude on the size of the observed oligomer. R_H is calculated for globular proteins, which presumably does not account for LT. Although the SV40 LT OBD-helicase, the most complete PyV LT hexamer structure available, reveals near globular dimensions of 120x110 Å, MCPyV LT additionally harbors the MURs. These regions comprise approximately one quarter of the protein (Figure 4A) and are predicted to be unstructured (Figure 4B), which is accompanied with the ability to attain compact or expanded conformations. This characteristic most likely prevents a globular shape of the MCPyV LT, impeding precise R_H determination. The peak at 47 nm may reflect aggregate or higher order oligomers such as double-hexamers of 1.1 MDa. This species could be visualized via EM for SV40 LT (Cuesta et al., 2010; Valle et al., 2006, 2000), but was not observed via EM in this work, nor reported in literature for MCPyV LT. The addition of the ori, ADP, and $MgCl_2$, provoking SV40 LT double-hexamer formation (Valle et al., 2000; Mastrangelo et al., 1989), reduced the fraction of potential double-hexamers in the DLS analysis. This indicates that the R_H of 47 nm reflects aggregates rather than double-hexamers.

Taking the negative stain TEM analysis into account (Figure 23D and Figure 24C), in which the predominant species were hexamers, it is presumable that the measured peak of 18 nm represents LT hexamers. Extensive biochemical analyses of the SV40 LT oligomerization reveal hexamer formation in the presence of nucleotides and $MgCl_2$ (D. Li et al., 2003; Reynisdóttir et al., 1993; Dean et al., 1992; Mastrangelo et al., 1989). Several SV40 LT helicase structures in the presence and absence of nucleotides and Mg^{2+} (Table S 5) provide structural explanations for this phenomenon. The zinc finger, which is the main oligomerization component of the helicase via hydrophobic interactions, in addition to polar and charged residues at the interface of two AAA+ subdomains, mediate primarily hexamerization (D. Li et al., 2003). This illustrates, why PyV LT helicases hexamerize in the absence of nucleotides and DNA, also observed in this DLS analysis (Gai, Zhao, et al., 2004; D. Li et al., 2003). However, these interactions are not the only intermonomer interactions. Coordinated six-fold DNA-β-hairpin interactions may also contribute to hexamer stabilization. Further, the ATP-binding pocket contributes to oligomerization via the *trans* acting arginine finger (SV40 LT: R540) and additional five charged residues (SV40 LT: K418, K419, D502, R498, and D499), shifting the SV40 LT monomer:hexamer 1:2 ratio to only hexamers (Gai, Zhao, et al., 2004; D. Li et al., 2003). Intriguingly, no MCPyV monomers were observed via SEC (6.2.3Figure 21) and DLS. It is hypothesized that

monomeric MCPyV LT is not stable, attributed to the MURs, resulting in aggregates. However, the addition of nucleotides, MgCl$_2$, and DNA mediates auxiliary intermonomer interactions, enabling the disengagement of LT from aggregates into ordered hexamers.

7.6.2. Flexible regions within LT possibly prevent the visualization of the full-length protein

CryoEM SPA was applied on LT since X-ray crystallography of *in cellulo* grown crystals of various T antigen constructs was not successful. The final model of 6.8 Å (FSC$_{0.143}$) revealed the MCPyV LT helicase, which was determined by superposing the model to the AlphaFold2 predicted MCPyV LT$_{316-817}$ homo-hexamer and the crystal structure of the SV40 LT helicase (Figure 27). Although the full-length LT was examined, the N-terminus and the OBD were missing in the model, possibly originating from an increased flexibility of these regions. This hypothesis was supported by the results of the superposition of the AlphaFold2 prediction and the MCPyV LT helicase. Regions with pLDDTs below 60 (blue), indicating low prediction confidence and high flexibility, are mostly not covered by the model (Figure 30A and B). These regions include the C-terminus (magenta arrow), which is also predicted as IDR via IUPred (Figure 4B) and the linker between the zinc finger and the helicase (green arrow). It is conceivable that this linker together with the MURs provide such high flexibility that the average of single particles did not result in a defined structure of the LT N-terminus. Interestingly, Langston and colleagues managed to visualize the SV40 LT$_{131-627}$, comprising the OBD and the helicase. The final model has a length of 110 Å and a resolution of 5.6 Å (Langston et al., 2022). The size of the OBD-helicase linker is comparable between the two PyVs, while SV40 lacks the MURs. The absence of the N-terminus in the MCPyV LT cryoEM SPA model is therefore hypothesized to be largely linked to the MURs. In any case, an exclusion of the MCPyV LT OBD by the application of a too small particle mask during 2D classification can be excluded since the applied diameter of 160 Å exceeded the length of the SV40 LT helicase-OBD by 50 Å. This should be sufficient to account for differences introduced by the MURs.

The only AlphaFold2 predicted region with low confidence, which could be superimposed to the cryoEM SPA model, was the DNA-interacting β-hairpin (yellow arrow). For SV40 LT, it could be shown that these regions move longitudinally along the hexameric helicase channel to translocate DNA (J. Shen et al., 2005; Gai, Li, et al., 2004). Further, it was demonstrated that this movement is linked to the occupancy of the ATP-binding pocket. ATP-binding resulted in a contracted channel with a diameter of 14 Å, while the apo state showed an expanded channel width of 22 Å (Gai, Li, et al., 2004). To stabilize the flexible predicted MCPyV LT β-hairpins, ADP and ori were supplemented for the cryoEM SPA, resulting in a density, which could be aligned to the predicted β-hairpins (Figure 27B and Figure 30A and B). Further, the presumed density of the MCPyV β-hairpins overlapped most

precisely with those of SV40 LT bound to ATP, although ADP was supplemented for the cryoEM SPA (Figure 27C). It is important to note that the SV40 LT helicases, displayed in Figure 27C-E, were crystallized without DNA, contrary to the MCPyV LT. It is conceivable that the DNA-interacting β-hairpins behave differently in the presence of DNA, resulting in a contracted channel, similar to SV40 LT bound to ATP. The superposition of the model to SV40 LT bound to ADP revealed an overlap of only two DNA-binding hairpins (Figure 27D). The central channel of SV40:ADP showed an asymmetry with a diameter of 17x20 Å, which possibly originates from three different conformations of the ADP bound monomers within the hexamer (Gai, Li, et al., 2004). Such a channel asymmetry could not be detected in the MCPyV model, possibly resulting from the applied C6 symmetry.
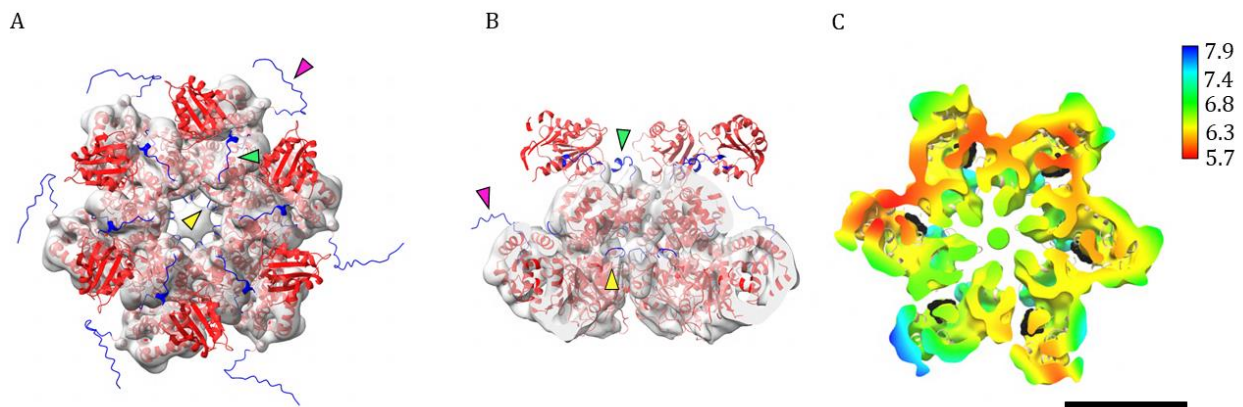
Figure 30: Superposition of the LT model to its predicted flexible regions and SV40 LT:ADP. (A) Top view of the LT cryoEM SPA model (gray) superimposed to the MCPyV $LT_{316-817}$ homo-hexamer AlphaFold2 prediction. PLDDT values are depicted above 60 (red) and below 60 (blue). Regions below 60 include the C-terminus (magenta arrow), the OBD-zinc finger linker (green arrow), and the DNA-binding β-hairpin (yellow arrow). (B) Cropped profile view of (A). (C) Local resolution ($FSC_{0.143}$) of the LT cryoEM SPA model, superimposed to the SV40 LT helicase bound to ADP. The cross-section of the N-terminal view is shown, and the resolution is color coded, ranging from 5.7 Å (red) to 7.9 Å (blue). The SV40 LT helicase hexamer (beige, PDB: 1SVL) is depicted in ribbon representation, while ADP is shown as surface representation (black). The scale bar represents 50 Å.

7.6.3. Asymmetric anisotropic resolution of the LT helicase model suggests mechanistic similarities to HPV E1

The MCPyV LT helicase model obtained by cryoEM SPA revealed an anisotropic resolution with higher resolution in its interior compared to its exterior (Figure 26E and Figure 30C). Particularly, the resolution decreased towards the C-terminus, which was expected since the C-terminus is predicted as IDR, as aforementioned. Lower resolution at the protein-solvent interface is also comprehensible since less stabilizing protein-protein interactions exist. However, an asymmetric resolution anisotropy within a hexamer of a six-fold symmetry was not expected. The crystal structure of the bovine papillomavirus (BPV) E1 ATPase, bound to ssDNA, deviated from a proper six-fold symmetry and showed varying modes of Mg^{2+}-ADP coordination of the single E1 subunits (Enemark & Joshua-Tor, 2006). While the walker A and B motif from one subunit were consistently structured within the

entire hexamer, the nucleotide-coordinating residues from the adjacent subunit adopted different conformations, reflecting the ATP, ADP and apo state. This correlated with a staircase arrangement of the β-hairpins. The ATP state revealed the most interactions, while the apo state showed no interactions of nucleotide-coordinating residues between two adjacent subunits. Furthermore, it could be demonstrated for several hexameric ATPases that the arginine finger serves as bridge between two adjacent monomers, resulting in an asymmetry by the formation of a compact dimer within the hexamer (Guo et al., 2019). It is conceivable that different modes of Mg^{2+}-nucleotide coordination also exist within the MCPyV LT ATPase. Pronounced interactions between two adjacent subunits, mediated by the ATP-binding state, could stabilize this region, resulting in a higher local resolution. Interestingly, by superimposing the SV40 LT helicase bound to ADP (black) to the MCPyV LT cryoEM SPA model, increased resolution can be observed in proximity to the ATP-binding pocket (Figure 30C). Only one LT monomer showed intermediate resolutions between 6.3-6.8 Å in this area, possibly reflecting the apo state. This would be contradictory to SV40 LT, which showed overall reduced distances between the monomers in the presence of ATP, compared to the apo state, but was suggested to follow an all-or-none ATP-binding mode and a concerted nucleotide hydrolysis, linked to a planar arrangement of the DNA-binding β-hairpins (Gai, Li, et al., 2004). The argument that the differences between SV40 LT and BPV E1 occur due to the absence of DNA in the SV40 LT structure can be refuted by the preservation of the asymmetry in the DNA-free BPV E1 structure (Sanders et al., 2007). It is also conceivable that the different arrangements reflect distinct stages during DNA unwinding. The planar arrangement of SV40:dsDNA could reflect the pre-initiation complex, while the staircase arrangement of E1 represents the complex during DNA unwinding elongation (Gai et al., 2016). However, the asymmetric anisotropic resolution of the MCPyV LT helicase model provides first indications that an asymmetric hexamer is reasonable in the presence of dsDNA, arguing against the hypothesize of different distinct stages of the SV40 LT and BPV E1 crystal structures, and suggests mechanistic ATPase similarities between the papillomavirus E1 and MCPyV LT rather than SV40 and MCPyV.

7.6.4. Approaches to increase the resolution of the MCPyV LT helicase model

The resolution of the final MCPyV LT helicase model reached 6.8 Å ($FSC_{0.143}$), which is insufficient to elucidate its structure, routinely realized for resolutions below 4 Å. Moreover, this resolution is potentially slightly overestimated since explicit secondary structures such as alpha helices and beta sheets, which become visible at approximately 8 Å (M. L. Baker et al., 2007), were challenging to detect. The model revealed preferred orientation of the helicase C-terminus, while profile views were underrepresented (Figure 26D). The limited angular representations result in anisotropic resolution, which can prevent structure elucidation. Preferred particle orientation is often associated with

protein adsorption to the air-water interface (AWI), in which hydrophobic protein patches interact with hydrophobic air (B. Li et al., 2021; J. Chen et al., 2019; Noble, Wei, et al., 2018). Next to orientation preference, proteins can denature upon AWI adsorption, shown for the yeast fatty acid synthase (D'Imprima et al., 2018). It has been demonstrated that 90% of all examined proteins associate with the AWI and that up to 94% of all particles of one data set are excluded from the final model (D'Imprima & Kühlbrandt, 2021; Noble, Dandey, et al., 2018). Together with the observed preferred orientation of LT, these facts provide additional indications that LT suffers from the interaction with the AWI. Based on Brownian motion, 100 kDa proteins need approximately 0.1 ms or less to reach the AWI, leaving more than 1000 opportunities to adopt a preferred orientation or denature before freezing (B. Li et al., 2021; Naydenova & Russo, 2017). To overcome this challenge for LT and to increase resolution, the time between protein blotting and freezing should be held at a minimum to decrease protein exposition to air. Promising results were reported by utilizing a microfluidic spray-plunge freezing machine (X. Feng et al., 2017), a surface acoustic wave-based microfluid dispenser (Ashtiani et al., 2018), and by delivering picolitre volumes onto a self-blotting nanowire grid, called Spotiton (Noble, Wei, et al., 2018; Wei et al., 2018; Razinkov et al., 2016). Another method to reduce protein-WAI interaction is to supplement the sample with detergents, masking the WAI with their hydrophobic part. Chen at al. could show that the addition of the zwitterionic detergent CHAPSO prevents WAI adsorption of a bacterial RNA polymerase (2019). Further, tilting the specimen can help to gain information of rare angular representations. However, the contrast is reduced since the electron path length through the ice increases (Tan et al., 2017). A fourth method to overcome preferred orientation and protein unfolding is the implementation of physical supports such as very thin continuous layers of amorphous carbon or a crystalline monomolecular carbon layer, graphene, providing a protein interaction surface (D'Imprima et al., 2018). Thereby, proteins can be evenly spread onto amorphous carbon while higher background is provoked. Contrarily, graphene is almost electron-transparent, while its hydrophobicity challenges protein adsorption (D'Imprima et al., 2018). Continuous carbon support was utilized in the cryoEM SPA approach B, in which phosphate-supplemented buffer and high salt cell lysis were applied, resulting in noisy LT 2D representations, which were not further analyzed. Although the analysis was not successful, LT was evenly distributed on the grid compared to holey carbon (results not shown), suggesting the use of continuous carbon grids for subsequent experiments. An additional advantage of continuous carbon grids is the effect of sample concentration on the grid. This may allow the omission of sample concentration via a semipermeable membrane, which was used in this work and which can damage the protein, as shown for the fatty acid synthase (Joppe et al., 2020; D'Imprima et al., 2019).

A second challenge during LT structure analysis was its pronounced heterogeneity (Figure 26C), impairing high-resolution by reducing the number of available particles for the final model. One explanation for this phenomenon could be LT-bound heterogeneous DNA fragments. Unspecific DNA-interaction was already observed for SV40 LT (Wessel et al., 1992; H. J. Lin et al., 1992; Scheffner et al., 1989) and is also conceivable for MCPyV LT, especially as pronounced nucleic acid contamination was measured after LT purification without high salt cell lysis (6.2.3). CryoEM SPA was performed with purified LT, lacking high salt cell lysis to leave unspecific DNA-protein interactions intact for LT stabilization during purification. The ori was supplemented only after purification and it remains to be verified whether the unspecific DNA was substituted by the LT-specific ori. However, the DPI-ELISA and nanoDSF analyzes (Figure 23 A and B), demonstrating increased thermal stability and specific binding upon ori addition, indicate the substitution of unspecific DNA with the viral ori. Nonetheless, an incomplete substitution cannot be excluded.

A factor which further limits resolution are salt concentrations above 100 mM. Their ions scatter the incoming electrons stronger compared to water, reducing image contrast (D'Imprima & Kühlbrandt, 2021). Unfortunately, LT is unstable below 250 mM NaCl, as demonstrated in the anion exchange chromatography (Figure 20B and C), eliminating the option of salt reduction for contrast enhancement.

In summary, LT potentially interacts with the AWI, resulting in preferred orientation. The restricted number of angular representations as well as protein heterogeneity potentially mediated by unspecific LT-DNA interactions, led to reduced resolution. Preferred orientation could be addressed by decreasing the time of grid preparation, by supplementing specific detergents, or by using continuous carbon grids. Unfortunately, restricting unspecific LT-DNA interactions by increasing the salt concentration to reduce protein heterogeneity would result in additional resolution reduction.

7.6.5. MCPyV LT helicase as target for antiviral drug development

Helicases are ubiquitously utilized for all aspects of nucleic acid metabolism and play an essential role in MCPyV replication. Hitherto, no specific antiviral treatment is available against MCPyV, which could be altered by utilizing the LT helicase as drug target. This domain displays a perfect drug target since it is indispensable for the MCPyV life cycle, and it seems to be structurally well conserved across multiple PyVs (Table S 5). Non-structure-based efforts to identify inhibitors against PyV LT ATPase were reported for BKPyV, JCPyV, and SV40, providing a proof of principle for the druggability of PyV LT helicases (Randhawa et al., 2014; Seguin et al., 2012; C. M. Wright et al., 2009). Bonafoux et al. reported the first structure-based design of an ATP-competitive inhibitor of the JCPyV LT helicase. Therefore, the co-crystal structure of the JCPyV LT helicase, bound to preliminary inhibitors, were solved and utilized to improve their biochemical activity, resulting in the final compound

triazolopyridine. They further demonstrated that this compound was also active against BKPyV (Bonafoux et al., 2016). This is not surprising since the sequence alignment of the helicase segment, harboring ATP-coordinating residues (red), revealed high conservation between PyVs (Figure 31). It is conceivable that previously identified ATPase inhibitors against other clinically relevant PyVs, such as BKPyV and JCPyV, could be utilized as starting point to develop compounds against MCPyV LT helicase. However, it has to be noticed that the JCPyV lysine at position 420 (light red) (position 419 in SV40), interacting with the inhibitory triazolopyridine in the JCPyV crystal structure (Bonafoux et al., 2016) and with the ATP ribose in the SV40:ATP crystal structure (Gai, Zhao, et al., 2004), is not conserved between PyVs. In MCPyV, it is substituted with the hydrophobic tyrosine (aa568), and it remains to be validated whether this substitution affects the activity of the known ATPase inhibitors against MCPyV.

Figure 31: Sequence alignment of the helicase ATP-binding segment of MCPyV, SV40, BKPyV, and JCPyV. The amino acid number of the MCPyV LT is marked above the alignment and those which coordinate ATP-binding are marked in red (conserved and highly similar) and light red (not conserved). Identical (asterisk), highly similar (colon), and weakly similar (period) amino acids are marked below the alignment, which was performed with Clustal Omega. ATP-binding residues were extracted from (An et al., 2015).

Another approach to impair MCPyV LT helicase could be the inhibition of the helicase-primase complex, as it was implemented for the herpesviruses HSV and varicella zoster virus (VZV) (Chono et al., 2010). Although PyVs utilize the cellular primase for viral replication, contrary to herpes viruses, targeting this complex could be of interest since a structure-based approach is possible by applying the crystal structure of the SV40 helicase-primase as starting point for a structure-based antiviral drug approach (Zhou et al., 2012).

7.7. Conclusion

In this work, the first MCPyV LT helicase hexamer model with a resolution of 6.8 Å has been presented. The model revealed a high structure similarity to the SV40 LT helicase, despite low sequence similarity. Although the full-length LT was analyzed, only its helicase could be visualized, potentially due to the MURs. However, since they are unique to the MCPyV and provide an essential difference to other clinically relevant hPyVs, such as BKPyV and JCPyV, their structure analysis is fundamental to understand its molecular mechanism of action and its involvement in tumorigenesis. It is therefore beneficial to stabilize the LT MURs with known binding partners, such as Usp7 or pRb, rather than deleting them for subsequent cryoEM SPA analyses. Additionally, the resolution has to be increased to be able to use the LT helicase for structure-based antiviral drug target development. Reducing preferred orientation by decreasing the grid preparation time represents one option.

A second finding of this work was the unsuitability of the PAK4cat:Inka1 proteinogenic crystallization aid for *in cellulo* crystallization of the MCPyV oncoproteins. However, it remains elusive whether this approach is not applicable for structure elucidation at all, or whether the MCPyV oncoproteins are not suitable due to their size or other unknown factors. Since tLT244 *in vitro* crystallization was also unsuccessful (Borchert, 2013), crystallization of MCPyV LT should not be considered further. However, *in vitro* crystallization of the SV40 sT was successful and has not been reported for the MCPyV sT thus far (Cho et al., 2007). Therefore, *in vitro* crystallization could hold a promising approach for MCPyV sT structure elucidation since its analysis via cryoEM SPA may result in low signal-to-noise ratios, ascribed to its small size of 22 kDa.

8. Literature

Abdul-Sada, H., Müller, M., Mehta, R., Toth, R., Arthur, J. S. C., Whitehouse, A., & Macdonald, A. (2017). The PP4R1 sub-unit of protein phosphatase PP4 is essential for inhibition of NF-κB by merkel polyomavirus small tumour antigen. *Oncotarget*, *8*(15), 25418–25432. https://doi.org/10.18632/oncotarget.15836

Afanasiev, O. K., Yelistratova, L., Miller, N., Nagase, K., Paulson, K., Iyer, J. G., Ibrani, D., Koelle, D. M., & Nghiem, P. (2013). Merkel polyomavirus-specific T cells fluctuate with merkel cell carcinoma burden and express therapeutically targetable PD-1 and Tim-3 exhaustion markers. *Clinical Cancer Research: An Official Journal of the American Association for Cancer Research*, *19*(19), 5351–5360. https://doi.org/10.1158/1078-0432.CCR-13-0035

Agelli, M., & Clegg, L. X. (2003). Epidemiology of primary Merkel cell carcinoma in the United States. *Journal of the American Academy of Dermatology*, *49*(5), 832–841. https://doi.org/10.1016/s0190-9622(03)02108-x

Agelli, M., Clegg, L. X., Becker, J. C., & Rollison, D. E. (2010). The etiology and epidemiology of merkel cell carcinoma. *Current Problems in Cancer*, *34*(1), 14–37. https://doi.org/10.1016/j.currproblcancer.2010.01.001

Ajuh, E. T., Wu, Z., Kraus, E., Weissbach, F. H., Bethge, T., Gosert, R., Fischer, N., & Hirsch, H. H. (2018). Novel Human Polyomavirus Noncoding Control Regions Differ in Bidirectional Gene Expression according to Host Cell, Large T-Antigen Expression, and Clinically Occurring Rearrangements. *Journal of Virology*, *92*(7), e02231-17. https://doi.org/10.1128/JVI.02231-17

Alberti, S. (2017). The wisdom of crowds: Regulating cell function through condensed states of living matter. *Journal of Cell Science*, *130*(17), 2789–2796. https://doi.org/10.1242/jcs.200295

Albores-Saavedra, J., Batich, K., Chable-Montero, F., Sagy, N., Schwartz, A. M., & Henson, D. E. (2010). Merkel cell carcinoma demographics, morphology, and survival based on 3870 cases: A population based study. *Journal of Cutaneous Pathology*, *37*(1), 20–27. https://doi.org/10.1111/j.1600-0560.2009.01370.x

Allen, P. J., Bowne, W. B., Jaques, D. P., Brennan, M. F., Busam, K., & Coit, D. G. (2005). Merkel Cell Carcinoma: Prognosis and Treatment of Patients From a Single Institution. *Journal of Clinical Oncology*, *23*(10), 2300–2309. https://doi.org/10.1200/JCO.2005.02.329

An, P., Brodsky, J. L., & Pipas, J. M. (2015). The conserved core enzymatic activities and the distinct dynamics of polyomavirus large T antigens. *Archives of Biochemistry and Biophysics*, *573*, 23–31. https://doi.org/10.1016/j.abb.2015.02.033

An, P., Sáenz Robles, M. T., & Pipas, J. M. (2012). Large T antigens of polyomaviruses: Amazing molecular machines. *Annual Review of Microbiology*, *66*, 213–236. https://doi.org/10.1146/annurev-micro-092611-150154

Anderson, H. A., Chen, Y., & Norkin, L. C. (1996). Bound simian virus 40 translocates to caveolin-enriched membrane domains, and its entry is inhibited by drugs that selectively disrupt caveolae. *Molecular Biology of the Cell*, *7*(11), 1825–1834. https://doi.org/10.1091/mbc.7.11.1825

Andruska, N., Mahapatra, L., Brenneman, R. J., Rich, J. T., Baumann, B. C., Compton, L., Thorstad, W. L., & Daly, M. D. (2021). Reduced Wide Local Excision Margins are Associated with Increased Risk of Relapse and Death from Merkel Cell Carcinoma. *Annals of Surgical Oncology*, *28*(6), 3312–3319. https://doi.org/10.1245/s10434-020-09145-7

Arias-Pulido, H., Peyton, C. L., Joste, N. E., Vargas, H., & Wheeler, C. M. (2006). Human Papillomavirus Type 16 Integration in Cervical Carcinoma In Situ and in Invasive Cervical Cancer. *Journal of Clinical Microbiology*, *44*(5), 1755–1762. https://doi.org/10.1128/JCM.44.5.1755-1762.2006

Arora, R., Shuda, M., Guastafierro, A., Feng, H., Toptan, T., Tolstov, Y., Normolle, D., Vollmer, L. L., Vogt, A., Dömling, A., Brodsky, J. L., Chang, Y., & Moore, P. S. (2012). Survivin Is a Therapeutic Target

in Merkel Cell Carcinoma. *Science Translational Medicine*, *4*(133), 133ra56-133ra56. https://doi.org/10.1126/scitranslmed.3003713

Arroyo, J. D., & Hahn, W. C. (2005). Involvement of PP2A in viral and cellular transformation. *Oncogene*, *24*(52), Article 52. https://doi.org/10.1038/sj.onc.1209038

Arunkumar, A. I., Klimovich, V., Jiang, X., Ott, R. D., Mizoue, L., Fanning, E., & Chazin, W. J. (2005). Insights into hRPA32 C-terminal domain–mediated assembly of the simian virus 40 replisome. *Nature Structural & Molecular Biology*, *12*(4), 332–339. https://doi.org/10.1038/nsmbXX

Ashtiani, D., Venugopal, H., Belousoff, M., Spicer, B., Mak, J., Neild, A., & de Marco, A. (2018). Delivery of femtolitre droplets using surface acoustic wave based atomisation for cryo-EM grid preparation. *Journal of Structural Biology*, *203*(2), 94–101. https://doi.org/10.1016/j.jsb.2018.03.012

Aydin, I., Weber, S., Snijder, B., Ventayol, P. S., Kühbacher, A., Becker, M., Day, P. M., Schiller, J. T., Kann, M., Pelkmans, L., Helenius, A., & Schelhaas, M. (2014). Large Scale RNAi Reveals the Requirement of Nuclear Envelope Breakdown for Nuclear Import of Human Papillomaviruses. *PLOS Pathogens*, *10*(5), e1004162. https://doi.org/10.1371/journal.ppat.1004162

Babakir-Mina, M., Ciccozzi, M., Lo Presti, A., Greco, F., Perno, C. F., & Ciotti, M. (2010). Identification of Merkel cell polyomavirus in the lower respiratory tract of Italian patients. *Journal of Medical Virology*, *82*(3), 505–509. https://doi.org/10.1002/jmv.21711

Baker, L. A., & Rubinstein, J. L. (2010). Radiation damage in electron cryomicroscopy. *Methods in Enzymology*, *481*, 371–388. https://doi.org/10.1016/S0076-6879(10)81015-8

Baker, M. L., Ju, T., & Chiu, W. (2007). Identification of Secondary Structure Elements in Intermediate Resolution Density Maps. *Structure (London, England : 1993)*, *15*(1), 7–19. https://doi.org/10.1016/j.str.2006.11.008

Banani, S. F., Lee, H. O., Hyman, A. A., & Rosen, M. K. (2017). Biomolecular condensates: Organizers of cellular biochemistry. *Nature Reviews Molecular Cell Biology*, *18*(5), Article 5. https://doi.org/10.1038/nrm.2017.7

Barbaro, B. A., Sreekumar, K. R., Winters, D. R., Prack, A. E., & Bullock, P. A. (2000). Phosphorylation of simian virus 40 T antigen on Thr 124 selectively promotes double-hexamer formation on subfragments of the viral core origin. *Journal of Virology*, *74*(18), 8601–8613. https://doi.org/10.1128/jvi.74.18.8601-8613.2000

Bargonetti, J., Reynisdóttir, I., Friedman, P. N., & Prives, C. (1992). Site-specific binding of wild-type p53 to cellular DNA is inhibited by SV40 T antigen and mutant p53. *Genes & Development*, *6*(10), 1886–1898. https://doi.org/10.1101/gad.6.10.1886

Barthold, S. W., Bhatt, P. N., & Johnson, E. A. (1987). Further evidence for papovavirus as the probable etiology of transmissible lymphoma of Syrian hamsters. *Laboratory Animal Science*, *37*(3), 283–288.

Barty, A., Caleman, C., Aquila, A., Timneanu, N., Lomb, L., White, T. A., Andreasson, J., Arnlund, D., Bajt, S., Barends, T. R. M., Barthelmess, M., Bogan, M. J., Bostedt, C., Bozek, J. D., Coffee, R., Coppola, N., Davidsson, J., DePonte, D. P., Doak, R. B., … Chapman, H. N. (2012). Self-terminating diffraction gates femtosecond X-ray nanocrystallography measurements. *Nature Photonics*, *6*(1), Article 1. https://doi.org/10.1038/nphoton.2011.297

Baskaran, Y., Ang, K. C., Anekal, P. V., Chan, W. L., Grimes, J. M., Manser, E., & Robinson, R. C. (2015). An in cellulo-derived structure of PAK4 in complex with its inhibitor Inka1. *Nature Communications*, *6*. https://doi.org/10.1038/ncomms9681

Bayer, N. J., Januliene, D., Zocher, G., Stehle, T., Moeller, A., & Blaum, B. S. (2020). Structure of Merkel Cell Polyomavirus Capsid and Interaction with Its Glycosaminoglycan Attachment Receptor. *Journal of Virology*, *94*(20), e01664-19. https://doi.org/10.1128/JVI.01664-19

Becker, J. C., Eigentler, T., Frerich, B., Gambichler, T., Grabbe, S., Höller, U., Klumpp, B., Loquai, C., Krause-Bergmann, A., Müller-Richter, U., Pföhler, C., Schneider-Burrus, S., Stang, A., Terheyden, P., Ugurel, S., Veith, J., & Mauch, C. (2019). S2k-Leitlinie Merkelzellkarzinom (MZK,

MCC, neuroendokrines Karzinom der Haut)—Update 2018. *Journal Der Deutschen Dermatologischen Gesellschaft = Journal of the German Society of Dermatology: JDDG*, *17*(5), 562–577. https://doi.org/10.1111/ddg.13841_g

Becker, J. C., Stang, A., Hausen, A. Z., Fischer, N., DeCaprio, J. A., Tothill, R. W., Lyngaa, R., Hansen, U. K., Ritter, C., Nghiem, P., Bichakjian, C. K., Ugurel, S., & Schrama, D. (2018). Epidemiology, biology and therapy of Merkel cell carcinoma: Conclusions from the EU project IMMOMEC. *Cancer Immunology, Immunotherapy: CII*, *67*(3), 341–351. https://doi.org/10.1007/s00262-017-2099-3

Becker, M., Dominguez, M., Greune, L., Soria-Martinez, L., Pfleiderer, M. M., Schowalter, R., Buck, C. B., Blaum, B. S., Schmidt, M. A., & Schelhaas, M. (2019). Infectious Entry of Merkel Cell Polyomavirus. *Journal of Virology*, *93*(6), e02004-18. https://doi.org/10.1128/JVI.02004-18

Bentz, G. L., Moss, C. R., Whitehurst, C. B., Moody, C. A., & Pagano, J. S. (2015). LMP1-Induced Sumoylation Influences the Maintenance of Epstein-Barr Virus Latency through KAP1. *Journal of Virology*, *89*(15), 7465–7477. https://doi.org/10.1128/JVI.00711-15

Bergvall, M., Melendy, T., & Archambault, J. (2013). THE E1 PROTEINS. *Virology*, *445*(0), 10.1016/j.virol.2013.07.020. https://doi.org/10.1016/j.virol.2013.07.020

Berjanskii, M. V., Riley, M. I., Xie, A., Semenchenko, V., Folk, W. R., & Van Doren, S. R. (2000). NMR structure of the N-terminal J domain of murine polyomavirus T antigens. Implications for DnaJ-like domains and for mutations of T antigens. *The Journal of Biological Chemistry*, *275*(46), 36094–36103. https://doi.org/10.1074/jbc.M006572200

Bernardes, N. E., Takeda, A. A. S., Dreyer, T. R., Freitas, F. Z., Bertolini, M. C., & Fontes, M. R. M. (2015). Structure of Importin-α from a Filamentous Fungus in Complex with a Classical Nuclear Localization Signal. *PloS One*, *10*(6), e0128687. https://doi.org/10.1371/journal.pone.0128687

Berrios, C., Padi, M., Keibler, M. A., Park, D. E., Molla, V., Cheng, J., Lee, S. M., Stephanopoulos, G., Quackenbush, J., & DeCaprio, J. A. (2016). Merkel Cell Polyomavirus Small T Antigen Promotes Pro-Glycolytic Metabolic Perturbations Required for Transformation. *PLoS Pathogens*, *12*(11), e1006020. https://doi.org/10.1371/journal.ppat.1006020

Bhatia, K., Goedert, J. J., Modali, R., Preiss, L., & Ayers, L. W. (2010). Immunological detection of viral large T antigen identifies a subset of Merkel cell carcinoma tumors with higher viral abundance and better clinical outcome. *International Journal of Cancer. Journal International Du Cancer*, *127*(6), 1493–1496. https://doi.org/10.1002/ijc.25136

Bichakjian, C. K., Olencki, T., Aasi, S. Z., Alam, M., Andersen, J. S., Blitzblau, R., Bowen, G. M., Contreras, C. M., Daniels, G. A., Decker, R., Farma, J. M., Fisher, K., Gastman, B., Ghosh, K., Grekin, R. C., Grossman, K., Ho, A. L., Lewis, K. D., Loss, M., … Engh, A. M. (2018). Merkel Cell Carcinoma, Version 1.2018, NCCN Clinical Practice Guidelines in Oncology. *Journal of the National Comprehensive Cancer Network: JNCCN*, *16*(6), 742–774. https://doi.org/10.6004/jnccn.2018.0055

Bikel, I., Montano, X., Agha, M. E., Brown, M., McCormack, M., Boltax, J., & Livingston, D. M. (1987). SV40 small t antigen enhances the transformation activity of limiting concentrations of SV40 large T antigen. *Cell*, *48*(2), 321–330. https://doi.org/10.1016/0092-8674(87)90435-1

Bochkareva, E., Martynowski, D., Seitova, A., & Bochkarev, A. (2006). Structure of the origin-binding domain of simian virus 40 large T antigen bound to DNA. *The EMBO Journal*, *25*(24), 5961–5969. https://doi.org/10.1038/sj.emboj.7601452

Bofill-Mas, S., Formiga-Cruz, M., Clemente-Casares, P., Calafell, F., & Girones, R. (2001). Potential transmission of human polyomaviruses through the gastrointestinal tract after exposure to virions or viral DNA. *Journal of Virology*, *75*(21), 10290–10299. https://doi.org/10.1128/JVI.75.21.10290-10299.2001

Bofill-Mas, S., Rodriguez-Manzano, J., Calgua, B., Carratala, A., & Girones, R. (2010). Newly described human polyomaviruses Merkel cell, KI and WU are present in urban sewage and may

represent potential environmental contaminants. *Virology Journal*, *7*, 141. https://doi.org/10.1186/1743-422X-7-141

Bollag, B., Hofstetter, C. A., Reviriego-Mendoza, M. M., & Frisque, R. J. (2010). JC Virus Small t Antigen Binds Phosphatase PP2A and Rb Family Proteins and Is Required for Efficient Viral DNA Replication Activity. *PLOS ONE*, *5*(5), e10606. https://doi.org/10.1371/journal.pone.0010606

Bonafoux, D., Nanthakumar, S., Bandarage, U. K., Memmott, C., Lowe, D., Aronov, A. M., Bhisetti, G. R., Bonanno, K. C., Coll, J., Leeman, J., Lepre, C. A., Lu, F., Perola, E., Rijnbrand, R., Taylor, W. P., Wilson, D., Zhou, Y., Zwahlen, J., & ter Haar, E. (2016). Fragment-Based Discovery of Dual JC Virus and BK Virus Helicase Inhibitors. *Journal of Medicinal Chemistry*, *59*(15), 7138–7151. https://doi.org/10.1021/acs.jmedchem.6b00486

Borchert, S. (2013). *Strukturelle und biochemische Analyse der Funktion des Merkelzellpolyomavirus Large Tumor Antigens*. Universität Hamburg.

Borchert, S., Czech-Sioli, M., Neumann, F., Schmidt, C., Wimmer, P., Dobner, T., Grundhoff, A., & Fischer, N. (2014). High-affinity Rb binding, p53 inhibition, subcellular localization, and transformation by wild-type or tumor-derived shortened Merkel cell polyomavirus large T antigens. *Journal of Virology*, *88*(6), 3144–3160. https://doi.org/10.1128/JVI.02916-13

Borkosky, S. S., Fassolari, M., Campos-León, K., Rossi, A. H., Salgueiro, M., Pascuale, C. A., Martínez, R. P., Gaston, K., & de Prat Gay, G. (2022). Biomolecular Condensation of the Human Papillomavirus E2 Master Regulator with p53: Implications in Viral Replication. *Journal of Molecular Biology*, 167889. https://doi.org/10.1016/j.jmb.2022.167889

Borowiec, J. A., & Hurwitz, J. (1988). ATP stimulates the binding of simian virus 40 (SV40) large tumor antigen to the SV40 origin of replication. *Proceedings of the National Academy of Sciences*, *85*(1), 64–68. https://doi.org/10.1073/pnas.85.1.64

Boudes, M., Garriga, D., Fryga, A., Caradoc-Davies, T., & Coulibaly, F. (2016). A pipeline for structure determination of in vivo-grown crystals using in cellulo diffraction. *Acta Crystallographica Section D: Structural Biology*, *72*(4), Article 4. https://doi.org/10.1107/S2059798316002369

Bragg, W. H., & Bragg, W. L. (1913). The Reflection of X-rays by Crystals. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, *88*(605), 428–438.

Busam, K. J., Jungbluth, A. A., Rekthman, N., Coit, D., Pulitzer, M., Bini, J., Arora, R., Hanson, N. C., Tassello, J. A., Frosina, D., Moore, P., & Chang, Y. (2009). Merkel Cell Polyomavirus Expression in Merkel Cell Carcinomas and Its Absence in Combined Tumors and Pulmonary Neuroendocrine Carcinomas. *The American Journal of Surgical Pathology*, *33*(9), 1378–1385. https://doi.org/10.1097/PAS.0b013e3181aa30a5

Campello, C., Comar, M., D'Agaro, P., Minicozzi, A., Rodella, L., & Poli, A. (2011). A molecular case–control study of the Merkel cell polyomavirus in colon cancer. *Journal of Medical Virology*, *83*(4), 721–724. https://doi.org/10.1002/jmv.22004

Caragliano, E., Bonazza, S., Frascaroli, G., Tang, J., Soh, T. K., Grünewald, K., Bosse, J. B., & Brune, W. (2022). Human cytomegalovirus forms phase-separated compartments at viral genomes to facilitate viral replication. *Cell Reports*, *38*(10), 110469. https://doi.org/10.1016/j.celrep.2022.110469

Carter, J. J., Daugherty, M. D., Qi, X., Bheda-Malge, A., Wipf, G. C., Robinson, K., Roman, A., Malik, H. S., & Galloway, D. A. (2013). Identification of an overprinting gene in Merkel cell polyomavirus provides evolutionary insight into the birth of viral genes. *Proceedings of the National Academy of Sciences*, *110*(31), 12744–12749. https://doi.org/10.1073/pnas.1303526110

Carter, J. J., Paulson, K. G., Wipf, G. C., Miranda, D., Madeleine, M. M., Johnson, L. G., Lemos, B. D., Lee, S., Warcola, A. H., Iyer, J. G., Nghiem, P., & Galloway, D. A. (2009). Association of Merkel Cell Polyomavirus–Specific Antibodies With Merkel Cell Carcinoma. *JNCI Journal of the National Cancer Institute*, *101*(21), 1510–1522. https://doi.org/10.1093/jnci/djp332

Chang, Y., & Moore, P. S. (2012). Merkel cell carcinoma: A virus-induced human cancer. *Annual Review of Pathology*, *7*, 123–144. https://doi.org/10.1146/annurev-pathol-011110-130227

Chang, Y. P., Xu, M., Machado, A. C. D., Yu, X. J., Rohs, R., & Chen, X. S. (2013). Mechanism of Origin DNA Recognition and Assembly of an Initiator-Helicase Complex by SV40 Large Tumor Antigen. *Cell Reports*, *3*(4), 1117–1127. https://doi.org/10.1016/j.celrep.2013.03.002

Chapman, H. N., Fromme, P., Barty, A., White, T. A., Kirian, R. A., Aquila, A., Hunter, M. S., Schulz, J., DePonte, D. P., Weierstall, U., Doak, R. B., Maia, F. R. N. C., Martin, A. V., Schlichting, I., Lomb, L., Coppola, N., Shoeman, R. L., Epp, S. W., Hartmann, R., … Spence, J. C. H. (2011). Femtosecond X-ray protein nanocrystallography. *Nature*, *470*(7332), 73–77. https://doi.org/10.1038/nature09750

Chari, A., Haselbach, D., Kirves, J.-M., Ohmer, J., Paknia, E., Fischer, N., Ganichkin, O., Möller, V., Frye, J. J., Petzold, G., Jarvis, M., Tietzel, M., Grimm, C., Peters, J.-M., Schulman, B. A., Tittmann, K., Markl, J., Fischer, U., & Stark, H. (2015). ProteoPlex: Stability optimization of macromolecular complexes by sparse-matrix screening of chemical space. *Nature Methods*, *12*(9), 859–865. https://doi.org/10.1038/nmeth.3493

Chen, J., Noble, A. J., Kang, J. Y., & Darst, S. A. (2019). Eliminating effects of particle adsorption to the air/water interface in single-particle cryo-electron microscopy: Bacterial RNA polymerase and CHAPSO. *Journal of Structural Biology: X*, *1*, 100005. https://doi.org/10.1016/j.yjsbx.2019.100005

Chen, T., Hedman, L., Mattila, P. S., Jartti, T., Ruuskanen, O., Söderlund-Venermo, M., & Hedman, K. (2011). Serological evidence of Merkel cell polyomavirus primary infections in childhood. *Journal of Clinical Virology*, *50*(2), 125–129. https://doi.org/10.1016/j.jcv.2010.10.015

Cheng, J., Park, D. E., Berrios, C., White, E. A., Arora, R., Yoon, R., Branigan, T., Xiao, T., Westerling, T., Federation, A., Zeid, R., Strober, B., Swanson, S. K., Florens, L., Bradner, J. E., Brown, M., Howley, P. M., Padi, M., Washburn, M. P., & DeCaprio, J. A. (2017). Merkel cell polyomavirus recruits MYCL to the EP400 complex to promote oncogenesis. *PLoS Pathogens*, *13*(10), e1006668. https://doi.org/10.1371/journal.ppat.1006668

Cheng, J., Rozenblatt-Rosen, O., Paulson, K. G., Nghiem, P., & DeCaprio, J. A. (2013). Merkel Cell Polyomavirus Large T Antigen Has Growth-Promoting and Inhibitory Activities. *Journal of Virology*, *87*(11), 6118–6126. https://doi.org/10.1128/JVI.00385-13

Cho, U. S., Morrone, S., Sablina, A. A., Arroyo, J. D., Hahn, W. C., & Xu, W. (2007). Structural basis of PP2A inhibition by small t antigen. *PLoS Biology*, *5*(8), e202. https://doi.org/10.1371/journal.pbio.0050202

Chono, K., Katsumata, K., Kontani, T., Kobayashi, M., Sudo, K., Yokota, T., Konno, K., Shimizu, Y., & Suzuki, H. (2010). ASP2151, a novel helicase-primase inhibitor, possesses antiviral activity against varicella-zoster virus and herpes simplex virus types 1 and 2. *The Journal of Antimicrobial Chemotherapy*, *65*(8), 1733–1741. https://doi.org/10.1093/jac/dkq198

Coulibaly, F., Chiu, E., Ikeda, K., Gutmann, S., Haebel, P. W., Schulze-Briese, C., Mori, H., & Metcalf, P. (2007). The molecular organization of cypovirus polyhedra. *Nature*, *446*(7131), 97–101. https://doi.org/10.1038/nature05628

Cowey, C. L., Mahnke, L., Espirito, J., Helwig, C., Oksen, D., & Bharmal, M. (2017). Real-world treatment outcomes in patients with metastatic Merkel cell carcinoma treated with chemotherapy in the USA. *Future Oncology (London, England)*, *13*(19), 1699–1710. https://doi.org/10.2217/fon-2017-0187

Cowie, A., & Kamen, R. (1986). Guanine nucleotide contacts within viral DNA sequences bound by polyomavirus large T antigen. *Journal of Virology*, *57*(2), 505–514.

Cricca, M., Venturoli, S., Leo, E., Costa, S., Musiani, M., & Zerbini, M. (2009). Disruption of HPV 16 E1 and E2 genes in precancerous cervical lesions. *Journal of Virological Methods*, *158*(1), 180–183. https://doi.org/10.1016/j.jviromet.2009.01.005

Cuesta, I., Núñez-Ramírez, R., Scheres, S. H. W., Gai, D., Chen, X. S., Fanning, E., & Carazo, J. M. (2010). Conformational Rearrangements of SV40 Large T Antigen during Early Replication Events. *Journal of Molecular Biology*, *397*(5), 1276–1286. https://doi.org/10.1016/j.jmb.2010.02.042

Czech-Sioli, M., Günther, T., Therre, M., Spohn, M., Indenbirken, D., Theiss, J., Riethdorf, S., Qi, M., Alawi, M., Wülbeck, C., Fernandez-Cuesta, I., Esmek, F., Becker, J. C., Grundhoff, A., & Fischer, N. (2020). High-resolution analysis of Merkel Cell Polyomavirus in Merkel Cell Carcinoma reveals distinct integration patterns and suggests NHEJ and MMBIR as underlying mechanisms. *PLOS Pathogens*, *16*(8), e1008562. https://doi.org/10.1371/journal.ppat.1008562

Czech-Sioli, M., Siebels, S., Radau, S., Zahedi, R. P., Schmidt, C., Dobner, T., Grundhoff, A., & Fischer, N. (2019). The Ubiquitin-Specific Protease Usp7, a Novel Merkel Cell Polyomavirus Large T-Antigen Interaction Partner, Modulates Viral DNA Replication. *Journal of Virology*. https://doi.org/10.1128/JVI.01638-19

Daniels, R., Rusan, N. M., Wadsworth, P., & Hebert, D. N. (2006). SV40 VP2 and VP3 insertion into ER membranes is controlled by the capsid protein VP1: Implications for DNA translocation out of the ER. *Molecular Cell*, *24*(6), 955–966. https://doi.org/10.1016/j.molcel.2006.11.001

Dao, T. P., Kolaitis, R.-M., Kim, H. J., O'Donovan, K., Martyniak, B., Colicino, E., Hehnly, H., Taylor, J. P., & Castañeda, C. A. (2018). Ubiquitin Modulates Liquid-Liquid Phase Separation of UBQLN2 via Disruption of Multivalent Interactions. *Molecular Cell*, *69*(6), 965-978.e6. https://doi.org/10.1016/j.molcel.2018.02.004

Das, R. K., Ruff, K. M., & Pappu, R. V. (2015). Relating sequence encoded information to form and function of intrinsically disordered proteins. *Current Opinion in Structural Biology*, *32*, 102–112. https://doi.org/10.1016/j.sbi.2015.03.008

Davey, M. J., & O'Donnell, M. (2003). Replicative helicase loaders: Ring breakers and ring makers. *Current Biology: CB*, *13*(15), R594-596. https://doi.org/10.1016/s0960-9822(03)00523-2

Dean, F. B., Borowiec, J. A., Eki, T., & Hurwitz, J. (1992). The simian virus 40 T antigen double hexamer assembles around the DNA at the replication origin. *Journal of Biological Chemistry*, *267*(20), 14129–14137. https://doi.org/10.1016/S0021-9258(19)49688-9

Dean, F. B., Borowiec, J. A., Ishimi, Y., Deb, S., Tegtmeyer, P., & Hurwitz, J. (1987). Simian virus 40 large tumor antigen requires three core replication origin domains for DNA unwinding and replication in vitro. *Proceedings of the National Academy of Science*, *84*, 8267–8271. https://doi.org/10.1073/pnas.84.23.8267

Deb, S., DeLucia, A. L., Koff, A., Tsui, S., & Tegtmeyer, P. (1986). The adenine-thymine domain of the simian virus 40 core origin directs DNA bending and coordinately regulates DNA replication. *Molecular and Cellular Biology*, *6*(12), 4578–4584.

DeCaprio, J. A., & Garcea, R. L. (2013). A cornucopia of human polyomaviruses. *Nature Reviews. Microbiology*, *11*(4), 264–276. https://doi.org/10.1038/nrmicro2992

Dela Cruz, F. N., Giannitti, F., Li, L., Woods, L. W., Del Valle, L., Delwart, E., & Pesavento, P. A. (2013). Novel polyomavirus associated with brain tumors in free-ranging raccoons, western United States. *Emerging Infectious Diseases*, *19*(1), 77–84. https://doi.org/10.3201/eid1901.121078

Diaz, J., Wang, X., Tsang, S. H., Jiao, J., & You, J. (2014). Phosphorylation of Large T Antigen Regulates Merkel Cell Polyomavirus Replication. *Cancers*, *6*(3), Article 3. https://doi.org/10.3390/cancers6031464

Dilworth, S. M. (1990). Cell alterations induced by the large T-antigens of SV40 and polyoma virus. *Seminars in Cancer Biology*, *1*(6), 407–414.

D'Imprima, E., Floris, D., Joppe, M., Sánchez, R., Grininger, M., & Kühlbrandt, W. (2018). *The deadly touch: Protein denaturation at the water-air interface and how to prevent it* (p. 400432). bioRxiv. https://doi.org/10.1101/400432

D'Imprima, E., Floris, D., Joppe, M., Sánchez, R., Grininger, M., & Kühlbrandt, W. (2019). Protein denaturation at the air-water interface and how to prevent it. *ELife*, *8*, e42747. https://doi.org/10.7554/eLife.42747

D'Imprima, E., & Kühlbrandt, W. (2021). Current limitations to high-resolution structure determination by single-particle cryoEM. *Quarterly Reviews of Biophysics*, *54*, e4. https://doi.org/10.1017/S0033583521000020

Dixon, R. A., & Nathans, D. (1985). Purification of simian virus 40 large T antigen by immunoaffinity chromatography. *Journal of Virology*, *53*(3), 1001–1004.

Dobson, S. J., Anene, A., Boyne, J. R., Mankouri, J., Macdonald, A., & Whitehouse, A. (2020). Merkel cell polyomavirus small tumour antigen activates the p38 MAPK pathway to enhance cellular motility. *Biochemical Journal*, *477*(14), 2721–2733. https://doi.org/10.1042/BCJ20200399

Drozdetskiy, A., Cole, C., Procter, J., & Barton, G. J. (2015). JPred4: A protein secondary structure prediction server. *Nucleic Acids Research*, *43*(W1), W389–W394. https://doi.org/10.1093/nar/gkv332

Du, M., & Chen, Z. J. (2018). DNA-induced liquid phase condensation of cGAS activates innate immune signaling. *Science*, *361*(6403), 704–709. https://doi.org/10.1126/science.aat1022

Dubois, M., Abi Rached, H., Escande, A., Dezoteux, F., Darloy, F., Jouin, A., Kyheng, M., Labreuche, J., Dziwniel, V., Mirabel, X., & Mortier, L. (2021). Outcome of early stage Merkel carcinoma treated by exclusive radiation: A study of 53 patients. *Radiation Oncology (London, England)*, *16*(1), 90. https://doi.org/10.1186/s13014-021-01815-4

Ea, E., Ha, K., Nm, N., Jf, W., H, H., F, G., Ps, R., & M, F. (2003). Cancer incidence in Denmark following exposure to poliovirus vaccine contaminated with simian virus 40. *Journal of the National Cancer Institute*, *95*(7). https://doi.org/10.1093/jnci/95.7.532

Eddy, B. E., Borman, G. S., Grubbs, G. E., & Young, R. D. (1962). Identification of the oncogenic substance in rhesus monkey kidney cell cultures as simian virus 40. *Virology*, *17*(1), 65–75. https://doi.org/10.1016/0042-6822(62)90082-X

Eisemann, N., Jansen, L., Castro, F. A., Chen, T., Eberle, A., Nennecke, A., Zeissig, S. R., Brenner, H., Katalinic, A., & for the GEKID Cancer Survival Working Group. (2016). Survival with nonmelanoma skin cancer in Germany. *British Journal of Dermatology*, *174*(4), 778–785. https://doi.org/10.1111/bjd.14352

Enemark, E. J., & Joshua-Tor, L. (2006). Mechanism of DNA translocation in a replicative hexameric helicase. *Nature*, *442*(7100), Article 7100. https://doi.org/10.1038/nature04943

Enemark, E. J., Stenlund, A., & Joshua-Tor, L. (2002). Crystal structures of two intermediates in the assembly of the papillomavirus replication initiation complex. *The EMBO Journal*, *21*(6), 1487–1496. https://doi.org/10.1093/emboj/21.6.1487

Erdős, G., Pajkos, M., & Dosztányi, Z. (2021). IUPred3: Prediction of protein disorder enhanced with unambiguous experimental annotation and visualization of evolutionary conservation. *Nucleic Acids Research*, *49*(W1), W297–W303. https://doi.org/10.1093/nar/gkab408

Erstad, D. J., & Cusack, J. C. (2014). Mutational Analysis of Merkel Cell Carcinoma. *Cancers*, *6*(4), Article 4. https://doi.org/10.3390/cancers6042116

Evrin, C., Clarke, P., Zech, J., Lurz, R., Sun, J., Uhle, S., Li, H., Stillman, B., & Speck, C. (2009). A double-hexameric MCM2-7 complex is loaded onto origin DNA during licensing of eukaryotic DNA replication. *Proceedings of the National Academy of Sciences*, *106*(48), 20240–20245. https://doi.org/10.1073/pnas.0911500106

Fazio, N., Maisonneuve, P., Spada, F., Gervaso, L., Cella, C. A., Pozzari, M., Zerini, D., Pisa, E., Fumagalli, C., Barberis, M., Laffi, A., Grana C, C. M., Orsolini, G., Prestianni, P., Bonomo, G., Funicelli, L., Bertani, E., Queirolo, P., Ravizza, D., … Pennacchioli, E. (2022). Nodal Merkel Cell Carcinoma with Unknown Primary Site and No Distant Metastasis: A Single-Center Series. *Cancers*, *14*(19), 4777. https://doi.org/10.3390/cancers14194777

Feng, H., Kwun, H. J., Liu, X., Gjoerup, O., Stolz, D. B., Chang, Y., & Moore, P. S. (2011). Cellular and Viral Factors Regulating Merkel Cell Polyomavirus Replication. *PLOS ONE*, *6*(7), e22468. https://doi.org/10.1371/journal.pone.0022468

Feng, H., Shuda, M., Chang, Y., & Moore, P. S. (2008). Clonal Integration of a Polyomavirus in Human Merkel Cell Carcinoma. *Science*, *319*(5866), 1096–1100. https://doi.org/10.1126/science.1152586

Feng, X., Fu, Z., Kaledhonkar, S., Jia, Y., Shah, B., Jin, A., Liu, Z., Sun, M., Chen, B., Grassucci, R. A., Ren, Y., Jiang, H., Frank, J., & Lin, Q. (2017). A Fast and Effective Microfluidic Spraying-Plunging Method for High-Resolution Single-Particle Cryo-EM. *Structure (London, England: 1993)*, *25*(4), 663-670.e3. https://doi.org/10.1016/j.str.2017.02.005

Field, N., Low, W., Daniels, M., Howell, S., Daviet, L., Boshoff, C., & Collins, M. (2003). KSHV vFLIP binds to IKK-γ to activate IKK. *Journal of Cell Science*, *116*(18), 3721–3728. https://doi.org/10.1242/jcs.00691

Fischer, N., Brandner, J., Fuchs, F., Moll, I., & Grundhoff, A. (2010). Detection of Merkel cell polyomavirus (MCPyV) in Merkel cell carcinoma cell lines: Cell morphology and growth phenotype do not reflect presence of the virus. *International Journal of Cancer*, *126*(9), 2133–2142. https://doi.org/10.1002/ijc.24877

Fiser, A. (2010). Template-Based Protein Structure Modeling. *Methods in Molecular Biology (Clifton, N.J.)*, *673*, 73–94. https://doi.org/10.1007/978-1-60761-842-3_6

Foulongne, V., Courgnaud, V., Champeau, W., & Segondy, M. (2011). Detection of Merkel cell polyomavirus on environmental surfaces. *Journal of Medical Virology*, *83*(8), 1435–1439. https://doi.org/10.1002/jmv.22110

Fowler, N. J., & Williamson, M. P. (2022). The accuracy of protein structures in solution determined by AlphaFold and NMR. *Structure*, *30*(7), 925-933.e2. https://doi.org/10.1016/j.str.2022.04.005

Frappier, L. (2012). The Epstein-Barr Virus EBNA1 Protein. *Scientifica*, *2012*, 438204. https://doi.org/10.6064/2012/438204

Frigola, J., He, J., Kinkelin, K., Pye, V. E., Renault, L., Douglas, M. E., Remus, D., Cherepanov, P., Costa, A., & Diffley, J. F. X. (2017). Cdt1 stabilizes an open MCM ring for helicase loading. *Nature Communications*, *8*(1), Article 1. https://doi.org/10.1038/ncomms15720

Fu, Y. V., Yardimci, H., Long, D. T., Ho, T. V., Guainazzi, A., Bermudez, V. P., Hurwitz, J., van Oijen, A., Schärer, O. D., & Walter, J. C. (2011). Selective bypass of a lagging strand roadblock by the eukaryotic replicative DNA helicase. *Cell*, *146*(6), 931–941. https://doi.org/10.1016/j.cell.2011.07.045

Gai, D., Chang, Y. P., & Chen, X. S. (2010). Origin DNA melting and unwinding in DNA replication. *Current Opinion in Structural Biology*, *20*(6), 756–762. https://doi.org/10.1016/j.sbi.2010.08.009

Gai, D., Li, D., Finkielstein, C. V., Ott, R. D., Taneja, P., Fanning, E., & Chen, X. S. (2004). Insights into the Oligomeric States, Conformational Changes, and Helicase Activities of SV40 Large Tumor Antigen *. *Journal of Biological Chemistry*, *279*(37), 38952–38959. https://doi.org/10.1074/jbc.M406160200

Gai, D., Wang, D., Li, S.-X., & Chen, X. S. (2016). The structure of SV40 large T hexameric helicase in complex with AT-rich origin DNA. *ELife*, *5*, e18129. https://doi.org/10.7554/eLife.18129

Gai, D., Zhao, R., Li, D., Finkielstein, C., & Chen, X. (2004). Mechanisms of Conformational Change for a Replicative Hexameric Helicase of SV40 Large Tumor Antigen. *Cell*, *119*, 47–60. https://doi.org/10.1016/j.cell.2004.09.017

Gambichler, T., Brüggestrat, L. G., Skrygan, M., Scheel, C. H., Susok, L., & Becker, J. C. (2023). The Antineoplastic Effect of Dimethyl Fumarate on Virus-Negative Merkel Cell Carcinoma Cell Lines: Preliminary Results. *Cancers*, *15*(2), 547. https://doi.org/10.3390/cancers15020547

Garbutcheon-Singh, K. B., Curchin, D. J., McCormack, C. J., & Smith, S. D. (2020). Trends in the incidence of Merkel cell carcinoma in Victoria, Australia, between 1986 and 2016. *Australasian Journal of Dermatology*, *61*(1), e34–e38. https://doi.org/10.1111/ajd.13131

Gardner, S. D., Field, A. M., Coleman, D. V., & Hulme, B. (1971). New human papovavirus (B.K.) isolated from urine after renal transplantation. *Lancet (London, England)*, *1*(7712), 1253–1257. https://doi.org/10.1016/s0140-6736(71)91776-4

Garneski, K. M., Warcola, A. H., Feng, Q., Kiviat, N. B., Leonard, J. H., & Nghiem, P. (2009). Merkel cell polyomavirus is more frequently present in North American than Australian Merkel cell carcinoma tumors. *The Journal of Investigative Dermatology*, *129*(1), 246–248. https://doi.org/10.1038/jid.2008.229

Gasteiger, E., Gattiker, A., Hoogland, C., Ivanyi, I., Appel, R. D., & Bairoch, A. (2003). ExPASy: The proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Research*, *31*(13), 3784–3788. https://doi.org/10.1093/nar/gkg563

Gati, C., Bourenkov, G., Klinge, M., Rehders, D., Stellato, F., Oberthür, D., Yefanov, O., Sommer, B. P., Mogk, S., Duszenko, M., Betzel, C., Schneider, T. R., Chapman, H. N., & Redecke, L. (2014). Serial crystallography on in vivo grown microcrystals using synchrotron radiation. *IUCrJ*, *1*(2), Article 2. https://doi.org/10.1107/S2052252513033939

Gibson, B. A., Doolittle, L. K., Schneider, M. W. G., Jensen, L. E., Gamarra, N., Henry, L., Gerlich, D. W., Redding, S., & Rosen, M. K. (2019). Organization of Chromatin by Intrinsic and Regulated Phase Separation. *Cell*, *179*(2), 470-484.e21. https://doi.org/10.1016/j.cell.2019.08.037

Gilardini Montani, M. S., Santarelli, R., Falcinelli, L., Gonnella, R., Granato, M., Di Renzo, L., Cuomo, L., Vitillo, M., Faggioni, A., & Cirone, M. (2018). EBV up-regulates PD-L1 on the surface of primary monocytes by increasing ROS and activating TLR signaling and STAT3. *Journal of Leukocyte Biology*, *104*(4), 821–832. https://doi.org/10.1002/JLB.2A0118-029RR

Goh, G., Walradt, T., Markarov, V., Blom, A., Riaz, N., Doumani, R., Stafstrom, K., Moshiri, A., Yelistratova, L., Levinsohn, J., Chan, T. A., Nghiem, P., Lifton, R. P., & Choi, J. (2015). Mutational landscape of MCPyV-positive and MCPyV-negative Merkel cell carcinomas with implications for immunotherapy. *Oncotarget*, *7*(3), 3403–3415. https://doi.org/10.18632/oncotarget.6494

Gorbalenya, A. E., Koonin, E. V., & Wolf, Y. I. (1990). A new superfamily of putative NTP-binding domains encoded by genomes of small DNA and RNA viruses. *FEBS Letters*, *262*(1), 145–148. https://doi.org/10.1016/0014-5793(90)80175-I

Gosert, R., Kardas, P., Major, E. O., & Hirsch, H. H. (2010). Rearranged JC Virus Noncoding Control Regions Found in Progressive Multifocal Leukoencephalopathy Patient Samples Increase Virus Early Gene Expression and Replication Rate. *Journal of Virology*, *84*(20), 10448–10456. https://doi.org/10.1128/JVI.00614-10

Gosert, R., Rinaldo, C. H., Funk, G. A., Egli, A., Ramos, E., Drachenberg, C. B., & Hirsch, H. H. (2008). Polyomavirus BK with rearranged noncoding control region emerge in vivo in renal transplant patients and increase viral replication and cytopathology. *Journal of Experimental Medicine*, *205*(4), 841–852. https://doi.org/10.1084/jem.20072097

Götz, C., Koenig, M. G., Issinger, O. G., & Montenarh, M. (1995). A casein-kinase-2-related protein kinase is tightly associated with the large T antigen of simian virus 40. *European Journal of Biochemistry*, *233*(1), 327–334. https://doi.org/10.1111/j.1432-1033.1995.327_1.x

Goudsmit, J., Wertheim-van Dillen, P., van Strien, A., & van der Noordaa, J. (1982). The role of BK virus in acute respiratory tract disease and the presence of BKV DNA in tonsils. *Journal of Medical Virology*, *10*(2), 91–99. https://doi.org/10.1002/jmv.1890100203

Goujon, M., McWilliam, H., Li, W., Valentin, F., Squizzato, S., Paern, J., & Lopez, R. (2010). A new bioinformatics analysis tools framework at EMBL–EBI. *Nucleic Acids Research*, *38*(suppl_2), W695–W699. https://doi.org/10.1093/nar/gkq313

Graffi, A., Bender, E., Schramm, T., Kuhn, W., & Schneiders, F. (1969). Induction of transmissible lymphomas in syrian hamsters by application of dna from viral hamster papovavirus-induced tumors and by cell-free filtrates from human tumors. *Proceedings of the National Academy of Sciences*, *64*(4), 1172–1175. https://doi.org/10.1073/pnas.64.4.1172

Gravemeyer, J., Spassova, I., Verhaegen, M. E., Dlugosz, A. A., Hoffmann, D., Lange, A., & Becker, J. C. (2021). DNA-methylation patterns imply a common cellular origin of virus- and UV-associated Merkel cell carcinoma. *Oncogene*. https://doi.org/10.1038/s41388-021-02064-1

Green, D. W., Ingram, V. M., & Perutz, M. F. (1954). The Structure of Haemoglobin. IV. Sign Determination by the Isomorphous Replacement Method. *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences*, *225*(1162), 287–307.

Greenleaf, W. B., Shen, J., Gai, D., & Chen, X. S. (2008). Systematic Study of the Functions for the Residues around the Nucleotide Pocket in Simian Virus 40 AAA+ Hexameric Helicase. *Journal of Virology*, *82*(12), 6017–6023. https://doi.org/10.1128/JVI.00387-08

Griffiths, D. A., Abdul-Sada, H., Knight, L. M., Jackson, B. R., Richards, K., Prescott, E. L., Peach, A. H. S., Blair, G. E., Macdonald, A., & Whitehouse, A. (2013). Merkel cell polyomavirus small T antigen targets the NEMO adaptor protein to disrupt inflammatory signaling. *Journal of Virology*, *87*(24), 13853–13867. https://doi.org/10.1128/JVI.02159-13

Gross, L. (1953). A Filterable Agent, Recovered from Ak Leukemic Extracts, Causing Salivary Gland Carcinomas in C3H Mice. *Proceedings of the Society for Experimental Biology and Medicine*, *83*(2), 414–421. https://doi.org/10.3181/00379727-83-20376

Grundhoff, A., & Fischer, N. (2015). Merkel cell polyomavirus, a highly prevalent virus with tumorigenic potential. *Current Opinion in Virology*, *14*, 129–137. https://doi.org/10.1016/j.coviro.2015.08.010

Guastafierro, A., Feng, H., Thant, M., Kirkwood, J. M., Chang, Y., Moore, P. S., & Shuda, M. (2013). Characterization of an early passage Merkel cell polyomavirus-positive Merkel cell carcinoma cell line, MS-1, and its growth in NOD scid gamma mice. *Journal of Virological Methods*, *187*(1), 6–14. https://doi.org/10.1016/j.jviromet.2012.10.001

Guergnon, J., Godet, A. N., Galioot, A., Falanga, P. B., Colle, J.-H., Cayla, X., & Garcia, A. (2011). PP2A targeting by viral proteins: A widespread biological strategy from DNA/RNA tumor viruses to HIV-1. *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease*, *1812*(11), 1498–1507. https://doi.org/10.1016/j.bbadis.2011.07.001

Guo, P., Driver, D., Zhao, Z., Zheng, Z., Chan, C., & Cheng, X. (2019). Controlling the Revolving and Rotating Motion Direction of Asymmetric Hexameric Nanomotor by Arginine Finger and Channel Chirality. *ACS Nano*, *13*(6), 6207–6223. https://doi.org/10.1021/acsnano.8b08849

Halata, Z., Grim, M., & Bauman, K. I. (2003). Friedrich Sigmund Merkel and his "Merkel cell", morphology, development, and physiology: Review and new results. *The Anatomical Record Part A: Discoveries in Molecular, Cellular, and Evolutionary Biology*, *271A*(1), 225–239. https://doi.org/10.1002/ar.a.10029

Handala, L., Blanchard, E., Raynal, P.-I., Roingeard, P., Morel, V., Descamps, V., Castelain, S., Francois, C., Duverlie, G., Brochot, E., & Helle, F. (2020). BK Polyomavirus Hijacks Extracellular Vesicles for En Bloc Transmission. *Journal of Virology*, *94*(6), e01834-19. https://doi.org/10.1128/JVI.01834-19

Harms, P. W., Harms, K. L., Moore, P. S., DeCaprio, J. A., Nghiem, P., Wong, M. K. K., & Brownell, I. (2018). The biology and treatment of Merkel cell carcinoma: Current understanding and research priorities. *Nature Reviews Clinical Oncology*, *15*(12), Article 12. https://doi.org/10.1038/s41571-018-0103-2

Harms, P. W., Vats, P., Verhaegen, M. E., Robinson, D. R., Wu, Y.-M., Dhanasekaran, S. M., Palanisamy, N., Siddiqui, J., Cao, X., Su, F., Wang, R., Xiao, H., Kunju, L. P., Mehra, R., Tomlins, S. A., Fullen, D. R., Bichakjian, C. K., Johnson, T. M., Dlugosz, A. A., & Chinnaiyan, A. M. (2015). The Distinctive Mutational Spectra of Polyomavirus-Negative Merkel Cell Carcinoma. *Cancer Research*, *75*(18), 3720–3727. https://doi.org/10.1158/0008-5472.CAN-15-0702

Harrison, C., Jiang, T., Banerjee, P., Meinke, G., D'Abramo, C. M., Schaffhausen, B., & Bohm, A. (2013). Polyomavirus Large T Antigen Binds Symmetrical Repeats at the Viral Origin in an Asymmetrical Manner. *Journal of Virology*, *87*(24), 13751–13759. https://doi.org/10.1128/JVI.01740-13

Harrison, C., Meinke, G., Kwun, H. J., Rogalin, H., Phelan, P. J., Bullock, P. A., Chang, Y., Moore, P. S., & Bohm, A. (2011). Asymmetric assembly of Merkel cell polyomavirus large T-antigen origin binding domains at the viral origin. *Journal of Molecular Biology*, *409*(4), 529–542. https://doi.org/10.1016/j.jmb.2011.03.051

Harrison, R. W. (1993). Phase problem in crystallography. *JOSA A*, *10*(5), 1046–1055. https://doi.org/10.1364/JOSAA.10.001046

Hartig, T. (1855). *Botan Z*, *13*, 881.

Hasegawa, H. (2019). Simultaneous induction of distinct protein phase separation events in multiple subcellular compartments of a single cell. *Experimental Cell Research*, *379*(1), 92–109. https://doi.org/10.1016/j.yexcr.2019.03.010

Hasegawa, H., Forte, C., Barber, I., Turnbaugh, S., Stoops, J., Shen, M., & Lim, A. C. (2014). Modulation of in vivo IgG crystallization in the secretory pathway by heavy chain isotype class switching and N-linked glycosylation. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research*, *1843*(7), 1325–1338. https://doi.org/10.1016/j.bbamcr.2014.03.024

Hasegawa, H., Wendling, J., He, F., Trilisky, E., Stevenson, R., Franey, H., Kinderman, F., Li, G., Piedmonte, D. M., Osslund, T., Shen, M., & Ketchem, R. R. (2011). In Vivo Crystallization of Human IgG in the Endoplasmic Reticulum of Engineered Chinese Hamster Ovary (CHO) Cells. *Journal of Biological Chemistry*, *286*(22), 19917–19931. https://doi.org/10.1074/jbc.M110.204362

Hashizume, Y., Inaka, K., Furubayashi, N., Kamo, M., Takahashi, S., & Tanaka, H. (2020). Methods for Obtaining Better Diffractive Protein Crystals: From Sample Evaluation to Space Crystallization. *Crystals*, *10*(2), Article 2. https://doi.org/10.3390/cryst10020078

Hatakeyama, M. (2019). Malignant Helicobacter pylori-Associated Diseases: Gastric Cancer and MALT Lymphoma. *Advances in Experimental Medicine and Biology*, *1149*, 135–149. https://doi.org/10.1007/5584_2019_363

Heath, M., Jaimes, N., Lemos, B., Mostaghimi, A., Wang, L. C., Peñas, P., & Nghiem, P. (2008). Clinical characteristics of Merkel cell carcinoma at diagnosis in 195 patients: The "AEIOU" features. *Journal of the American Academy of Dermatology*, *58*(3), 375–381. https://doi.org/10.1016/j.jaad.2007.11.020

Hein, J., Boichuk, S., Wu, J., Cheng, Y., Freire, R., Jat, P. S., Roberts, T. M., & Gjoerup, O. V. (2009). Simian Virus 40 Large T Antigen Disrupts Genome Integrity and Activates a DNA Damage Response via Bub1 Binding. *Journal of Virology*, *83*(1), 117–127. https://doi.org/10.1128/JVI.01515-08

Hnisz, D., Shrinivas, K., Young, R. A., Chakraborty, A. K., & Sharp, P. A. (2017). A Phase Separation Model for Transcriptional Control. *Cell*, *169*(1), 13–23. https://doi.org/10.1016/j.cell.2017.02.007

Hofmeister, F. (1888). Zur Lehre von der Wirkung der Salze. *Archiv für experimentelle Pathologie und Pharmakologie*, *24*(4), 247–260. https://doi.org/10.1007/BF01918191

Holehouse, A. S., Das, R. K., Ahad, J. N., Richardson, M. O. G., & Pappu, R. V. (2017). CIDER: Resources to Analyze Sequence-Ensemble Relationships of Intrinsically Disordered Proteins. *Biophysical Journal*, *112*(1), 16–21. https://doi.org/10.1016/j.bpj.2016.11.3200

Holowaty, M. N., Zeghouf, M., Wu, H., Tellam, J., Athanasopoulos, V., Greenblatt, J., & Frappier, L. (2003). Protein Profiling with Epstein-Barr Nuclear Antigen-1 Reveals an Interaction with the Herpesvirus-associated Ubiquitin-specific Protease HAUSP/USP7 *. *Journal of Biological Chemistry*, *278*(32), 29987–29994. https://doi.org/10.1074/jbc.M303977200

Hong, A. M., Ferguson, P., Dodds, T., Jones, D., Li, M., Yang, J., & Scolyer, R. A. (2019). Significant association of PD-L1 expression with human papillomavirus positivity and its prognostic impact in oropharyngeal cancer. *Oral Oncology*, *92*, 33–39. https://doi.org/10.1016/j.oraloncology.2019.03.012

Hopman, A. H. N., Smedts, F., Dignef, W., Ummelen, M., Sonke, G., Mravunac, M., Vooijs, G. P., Speel, E.-J. M., & Ramaekers, F. C. S. (2004). Transition of high-grade cervical intraepithelial neoplasia to micro-invasive carcinoma is characterized by integration of HPV 16/18 and numerical

chromosome abnormalities. *The Journal of Pathology*, *202*(1), 23–33. https://doi.org/10.1002/path.1490

Host, K. M., Jacobs, S. R., West, J. A., Zhang, Z., Costantini, L. M., Stopford, C. M., Dittmer, D. P., & Damania, B. (2017). Kaposi's Sarcoma-Associated Herpesvirus Increases PD-L1 and Proinflammatory Cytokine Expression in Human Monocytes. *MBio*, *8*(5), e00917-17. https://doi.org/10.1128/mBio.00917-17

Houben, R., Adam, C., Baeurle, A., Hesbacher, S., Grimm, J., Angermeyer, S., Henzel, K., Hauser, S., Elling, R., Bröcker, E.-B., Gaubatz, S., Becker, J. C., & Schrama, D. (2012). An intact retinoblastoma protein-binding site in Merkel cell polyomavirus large T antigen is required for promoting growth of Merkel cell carcinoma cells. *International Journal of Cancer*, *130*(4), 847–856. https://doi.org/10.1002/ijc.26076

Houben, R., Angermeyer, S., Haferkamp, S., Aue, A., Goebeler, M., Schrama, D., & Hesbacher, S. (2015). Characterization of functional domains in the Merkel cell polyoma virus Large T antigen. *International Journal of Cancer*, *136*(5), E290-300. https://doi.org/10.1002/ijc.29200

Houben, R., Shuda, M., Weinkam, R., Schrama, D., Feng, H., Chang, Y., Moore, P. S., & Becker, J. C. (2010). Merkel Cell Polyomavirus-Infected Merkel Cell Carcinoma Cells Require Expression of Viral T Antigens. *Journal of Virology*, *84*(14), 7064–7072. https://doi.org/10.1128/JVI.02400-09

Howley, P. M., & Livingston, D. M. (2009). Small DNA Tumor Viruses: Large Contributors to Biomedical Sciences. *Virology*, *384*(2), 256. https://doi.org/10.1016/j.virol.2008.12.006

Huang, S. G., Weisshart, K., & Fanning, E. (1998). Characterization of the nucleotide binding properties of SV40 T antigen using fluorescent 3'(2')-O-(2,4,6-trinitrophenyl)adenine nucleotide analogues. *Biochemistry*, *37*(44), 15336–15344. https://doi.org/10.1021/bi981094g

Iakoucheva, L. M., Brown, C. J., Lawson, J. D., Obradović, Z., & Dunker, A. K. (2002). Intrinsic Disorder in Cell-signaling and Cancer-associated Proteins. *Journal of Molecular Biology*, *323*(3), 573–584. https://doi.org/10.1016/S0022-2836(02)00969-5

Inokuma, Y., Kawano, M., & Fujita, M. (2011). Crystalline molecular flasks. *Nature Chemistry*, *3*(5), 349–358. https://doi.org/10.1038/nchem.1031

Inoue, T., & Tsai, B. (2011). A Large and Intact Viral Particle Penetrates the Endoplasmic Reticulum Membrane to Reach the Cytosol. *PLOS Pathogens*, *7*(5), e1002037. https://doi.org/10.1371/journal.ppat.1002037

Israel, M. A., Vanderryn, D. F., Meltzer, M. L., & Martin, M. A. (1980). Characterization of polyoma viral DNA sequences in polyoma-induced hamster tumor cell lines. *The Journal of Biological Chemistry*, *255*(8), 3798–3805.

Iyer, L. M., Leipe, D. D., Koonin, E. V., & Aravind, L. (2004). Evolutionary history and higher order classification of AAA+ ATPases. *Journal of Structural Biology*, *146*(1), 11–31. https://doi.org/10.1016/j.jsb.2003.10.010

Jacobs, D., Huang, H., Olino, K., Weiss, S., Kluger, H., Judson, B. L., & Zhang, Y. (2021). Assessment of Age, Period, and Birth Cohort Effects and Trends in Merkel Cell Carcinoma Incidence in the United States. *JAMA Dermatology*, *157*(1), 59–65. https://doi.org/10.1001/jamadermatol.2020.4102

Jiang, D., Srinivasan, A., Lozano, G., & Robbins, P. D. (1993). SV40 T antigen abrogates p53-mediated transcriptional activity. *Oncogene*, *8*(10), 2805–2812.

Jiang, X., Klimovich, V., Arunkumar, A. I., Hysinger, E. B., Wang, Y., Ott, R. D., Guler, G. D., Weiner, B., Chazin, W. J., & Fanning, E. (2006). Structural mechanism of RPA loading on DNA during activation of a simple pre-replication complex. *The EMBO Journal*, *25*(23), 5516–5526. https://doi.org/10.1038/sj.emboj.7601432

Johnson, E. M. (2010). Structural evaluation of new human polyomaviruses provides clues to pathobiology. *Trends in Microbiology*, *18*(5), 215–223. https://doi.org/10.1016/j.tim.2010.01.001

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

papillomavirus. *Journal of Virology*, *83*(5), 2067–2074. https://doi.org/10.1128/JVI.02190-08

Joo, W. S., Kim, H. Y., Purviance, J. D., Sreekumar, K. R., & Bullock, P. A. (1998). Assembly of T-Antigen Double Hexamers on the Simian Virus 40 Core Origin Requires Only a Subset of the Available Binding Sites. *Molecular and Cellular Biology*, *18*(5), 2677–2687. https://doi.org/10.1128/MCB.18.5.2677

Joppe, M., D'Imprima, E., Salustros, N., Paithankar, K. S., Vonck, J., Grininger, M., & Kühlbrandt, W. (2020). The resolution revolution in cryoEM requires high-quality sample preparation: A rapid pipeline to a high-resolution map of yeast fatty acid synthase. *IUCrJ*, *7*(Pt 2), 220–227. https://doi.org/10.1107/S2052252519017366

Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., Bridgland, A., Meyer, C., Kohl, S. A. A., Ballard, A. J., Cowie, A., Romera-Paredes, B., Nikolov, S., Jain, R., Adler, J., … Hassabis, D. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature*, *596*(7873), Article 7873. https://doi.org/10.1038/s41586-021-03819-2

Kantola, K., Sadeghi, M., Lahtinen, A., Koskenvuo, M., Aaltonen, L.-M., Möttönen, M., Rahiala, J., Saarinen-Pihkala, U., Riikonen, P., Jartti, T., Ruuskanen, O., Söderlund-Venermo, M., & Hedman, K. (2009). Merkel cell polyomavirus DNA in tumor-free tonsillar tissues and upper respiratory tract samples: Implications for respiratory transmission and latency. *Journal of Clinical Virology*, *45*(4), 292–295. https://doi.org/10.1016/j.jcv.2009.04.008

Kaplan, D. L., Davey, M. J., & O'Donnell, M. (2003). Mcm4,6,7 Uses a "Pump in Ring" Mechanism to Unwind DNA by Steric Exclusion and Actively Translocate along a Duplex *. *Journal of Biological Chemistry*, *278*(49), 49171–49182. https://doi.org/10.1074/jbc.M308074200

Kassem, A., Schöpflin, A., Diaz, C., Weyers, W., Stickeler, E., Werner, M., & Zur Hausen, A. (2008). Frequent detection of Merkel cell polyomavirus in human Merkel cell carcinomas and identification of a unique deletion in the VP1 gene. *Cancer Research*, *68*(13), 5009–5013. https://doi.org/10.1158/0008-5472.CAN-08-0949

Kaufman, H. L., Russell, J., Hamid, O., Bhatia, S., Terheyden, P., D'Angelo, S. P., Shih, K. C., Lebbé, C., Linette, G. P., Milella, M., Brownell, I., Lewis, K. D., Lorch, J. H., Chin, K., Mahnke, L., von Heydebreck, A., Cuillerot, J.-M., & Nghiem, P. (2016). Avelumab in patients with chemotherapy-refractory metastatic Merkel cell carcinoma: A multicentre, single-group, open-label, phase 2 trial. *The Lancet Oncology*, *17*(10), 1374–1385. https://doi.org/10.1016/S1470-2045(16)30364-3

Kean, J. M., Rao, S., Wang, M., & Garcea, R. L. (2009). Seroepidemiology of human polyomaviruses. *PLoS Pathogens*, *5*(3), e1000363. https://doi.org/10.1371/journal.ppat.1000363

Kervarrec, T., Aljundi, M., Appenzeller, S., Samimi, M., Maubec, E., Cribier, B., Deschamps, L., Sarma, B., Sarosi, E.-M., Berthon, P., Levy, A., Bousquet, G., Tallet, A., Touzé, A., Guyétant, S., Schrama, D., & Houben, R. (2020). Polyomavirus-Positive Merkel Cell Carcinoma Derived from a Trichoblastoma Suggests an Epithelial Origin of this Merkel Cell Carcinoma. *Journal of Investigative Dermatology*, *140*(5), 976–985. https://doi.org/10.1016/j.jid.2019.09.026

Khatchaturian, E. M., & Zohrabian, N. (2018). A novel report of a primary Merkel cell carcinoma lesion in the nasal vestibule. *Clinical Case Reports*, *6*(11), 2033–2036. https://doi.org/10.1002/ccr3.1757

Khopde, S., & Simmons, D. T. (2008). Simian Virus 40 DNA Replication Is Dependent on an Interaction between Topoisomerase I and the C-Terminal End of T Antigen. *Journal of Virology*, *82*(3), 1136–1145. https://doi.org/10.1128/JVI.01314-07

Kim, D. E., Blum, B., Bradley, P., & Baker, D. (2009). Sampling bottlenecks in de novo protein structure prediction. *Journal of Molecular Biology*, *393*(1), 249–260. https://doi.org/10.1016/j.jmb.2009.07.063

Kim, H.-Y., Ahn, B.-Y., & Cho, Y. (2001). Structural basis for the inactivation of retinoblastoma tumor suppressor by SV40 large T antigen. *The EMBO Journal*, *20*(1–2), 295–304. https://doi.org/10.1093/emboj/20.1.295

Kitamura, T., Aso, Y., Kuniyoshi, N., Hara, K., & Yogo, Y. (1990). High incidence of urinary JC virus excretion in nonimmunosuppressed older patients. *The Journal of Infectious Diseases*, *161*(6), 1128–1133. https://doi.org/10.1093/infdis/161.6.1128

Klaes, R., Woerner, S. M., Ridder, R., Wentzensen, N., Duerst, M., Schneider, A., Lotz, B., Melsheimer, P., & von Knebel Doeberitz, M. (1999). Detection of high-risk cervical intraepithelial neoplasia and cervical cancer by amplification of transcripts derived from integrated papillomavirus oncogenes. *Cancer Research*, *59*(24), 6132–6136.

Knepper, T. C., Montesion, M., Russell, J. S., Sokol, E. S., Frampton, G. M., Miller, V. A., Albacker, L. A., McLeod, H. L., Eroglu, Z., Khushalani, N. I., Sondak, V. K., Messina, J. L., Schell, M. J., DeCaprio, J. A., Tsai, K. Y., & Brohl, A. S. (2019). The Genomic Landscape of Merkel Cell Carcinoma and Clinicogenomic Biomarkers of Response to Immune Checkpoint Inhibitor Therapy. *Clinical Cancer Research*, *25*(19), 5961–5971. https://doi.org/10.1158/1078-0432.CCR-18-4159

Knight, L. M., Stakaityte, G., Jennifer, J. W., Abdul-Sada, H., Griffiths, D. A., Howell, G. J., Wheat, R., Blair, G. E., Steven, N. M., Macdonald, A., Blackbourn, D. J., & Whitehouse, A. (2015). Merkel Cell Polyomavirus Small T Antigen Mediates Microtubule Destabilization To Promote Cell Motility and Migration. *Journal of Virology*, *89*(1), 35–47. https://doi.org/10.1128/JVI.02317-14

Knips, J., Czech-Sioli, M., Spohn, M., Heiland, M., Moll, I., Grundhoff, A., Schumacher, U., & Fischer, N. (2017). Spontaneous lung metastasis formation of human Merkel cell carcinoma cell lines transplanted into scid mice. *International Journal of Cancer*, *141*(1), 160–171. https://doi.org/10.1002/ijc.30723

Koiwai, K., Tsukimoto, J., Higashi, T., Mafuné, F., Miyajima, K., Nakane, T., Matsugaki, N., Kato, R., Sirigu, S., Jakobi, A., Wilmanns, M., Sugahara, M., Tanaka, T., Tono, K., Joti, Y., Yabashi, M., Nureki, O., Mizohata, E., Nakatsu, T., … Yumoto, F. (2019). Improvement of Production and Isolation of Human Neuraminidase-1 in Cellulo Crystals. *ACS Applied Bio Materials*. https://doi.org/10.1021/acsabm.9b00686

Koonin, E. V. (1993). A common set of conserved motifs in a vast variety of putative nucleic acid-dependent ATPases including MCM proteins involved in the initiation of eukaryotic DNA replication. *Nucleic Acids Research*, *21*(11), 2541–2547. https://doi.org/10.1093/nar/21.11.2541

Koopmann, R., Cupelli, K., Redecke, L., Nass, K., Deponte, D. P., White, T. A., Stellato, F., Rehders, D., Liang, M., Andreasson, J., Aquila, A., Bajt, S., Barthelmess, M., Barty, A., Bogan, M. J., Bostedt, C., Boutet, S., Bozek, J. D., Caleman, C., … Duszenko, M. (2012). In vivo protein crystallization opens new routes in structural biology. *Nature Methods*, *9*(3), 259–262. https://doi.org/10.1038/nmeth.1859

Krump, N. A., & You, J. (2021). From Merkel Cell Polyomavirus Infection to Merkel Cell Carcinoma Oncogenesis. *Frontiers in Microbiology*, *12*, 739695. https://doi.org/10.3389/fmicb.2021.739695

Kucukelbir, A., Sigworth, F. J., & Tagare, H. D. (2014). The Local Resolution of Cryo-EM Density Maps. *Nature Methods*, *11*(1), 63–65. https://doi.org/10.1038/nmeth.2727

Kwun, H. J., Chang, Y., & Moore, P. S. (2017). Protein-mediated viral latency is a novel mechanism for Merkel cell polyomavirus persistence. *Proceedings of the National Academy of Sciences*, *114*(20), E4040–E4047. https://doi.org/10.1073/pnas.1703879114

Kwun, H. J., Guastafierro, A., Shuda, M., Meinke, G., Bohm, A., Moore, P. S., & Chang, Y. (2009). The minimum replication origin of merkel cell polyomavirus has a unique large T-antigen loading architecture and requires small T-antigen expression for optimal replication. *Journal of Virology*, *83*(23), 12118–12128. https://doi.org/10.1128/JVI.01336-09

Kwun, H. J., Shuda, M., Camacho, C. J., Gamper, A. M., Thant, M., Chang, Y., & Moore, P. S. (2015). Restricted protein phosphatase 2A targeting by Merkel cell polyomavirus small T antigen. *Journal of Virology*, *89*(8), 4191–4200. https://doi.org/10.1128/JVI.00157-15

Kwun, H. J., Shuda, M., Feng, H., Camacho, C. J., Moore, P. S., & Chang, Y. (2013). Merkel cell polyomavirus small T antigen controls viral replication and oncoprotein expression by targeting the cellular ubiquitin ligase SCFFbw7. *Cell Host & Microbe*, *14*(2), 125–135. https://doi.org/10.1016/j.chom.2013.06.008

Langston, L. D., Yuan, Z., Georgescu, R., Li, H., & O'Donnell, M. E. (2022). SV40 T-antigen uses a DNA shearing mechanism to initiate origin unwinding. *Proceedings of the National Academy of Sciences of the United States of America*, *119*(49), e2216240119. https://doi.org/10.1073/pnas.2216240119

Lania, L., Hayday, A., & Fried, M. (1981). Loss of functional large T-antigen and free viral genomes from cells transformed in vitro by polyoma virus after passage in vivo as tumor cells. *Journal of Virology*, *39*(2), 422–431.

Lee, S., Paulson, K. G., Murchison, E. P., Afanasiev, O. K., Alkan, C., Leonard, J. H., Byrd, D. R., Hannon, G. J., & Nghiem, P. (2011). Identification and validation of a novel mature microRNA encoded by the Merkel cell polyomavirus in human Merkel cell carcinomas. *Journal of Clinical Virology : The Official Publication of the Pan American Society for Clinical Virology*, *52*(3), 272–275. https://doi.org/10.1016/j.jcv.2011.08.012

Leeman, J. E., Li, Y., Bell, A., Hussain, S. S., Majumdar, R., Rong-Mullins, X., Blecua, P., Damerla, R., Narang, H., Ravindran, P. T., Lee, N. Y., Riaz, N., Powell, S. N., & Higginson, D. S. (2019). Human papillomavirus 16 promotes microhomology-mediated end-joining. *Proceedings of the National Academy of Sciences of the United States of America*, *116*(43), 21573–21579. https://doi.org/10.1073/pnas.1906120116

Lemos, B. D., Storer, B. E., Iyer, J. G., Phillips, J. L., Bichakjian, C. K., Fang, L. C., Johnson, T. M., Liegeois-Kwon, N. J., Otley, C. C., Paulson, K. G., Ross, M. I., Yu, S. S., Zeitouni, N. C., Byrd, D. R., Sondak, V. K., Gershenwald, J. E., Sober, A. J., & Nghiem, P. (2010). Pathologic nodal evaluation improves prognostic accuracy in Merkel cell carcinoma: Analysis of 5823 cases as the basis of the first consensus staging system. *Journal of the American Academy of Dermatology*, *63*(5), 751–761. https://doi.org/10.1016/j.jaad.2010.02.056

Li, B., Zhu, D., Shi, H., & Zhang, X. (2021). Effect of charge on protein preferred orientation at the air–water interface in cryo-electron microscopy. *Journal of Structural Biology*, *213*(4), 107783. https://doi.org/10.1016/j.jsb.2021.107783

Li, D., Zhao, R., Lilyestrom, W., Gai, D., Zhang, R., DeCaprio, J. A., Fanning, E., Jochimiak, A., Szakonyi, G., & Chen, X. S. (2003). Structure of the replicative helicase of the oncoprotein SV40 large tumour antigen. *Nature*, *423*(6939), Article 6939. https://doi.org/10.1038/nature01691

Li, J., Diaz, J., Wang, X., Tsang, S. H., & You, J. (2015). Phosphorylation of Merkel Cell Polyomavirus Large Tumor Antigen at Serine 816 by ATM Kinase Induces Apoptosis in Host Cells*. *Journal of Biological Chemistry*, *290*(3), 1874–1884. https://doi.org/10.1074/jbc.M114.594895

Li, J., Wang, X., Diaz, J., Tsang, S. H., Buck, C. B., & You, J. (2013). Merkel Cell Polyomavirus Large T Antigen Disrupts Host Genomic Integrity and Inhibits Cellular Proliferation. *Journal of Virology*, *87*(16), 9173–9188. https://doi.org/10.1128/JVI.01216-13

Li, J., Zhang, M., Ma, W., Yang, B., Lu, H., Zhou, F., & Zhang, L. (2022). Post-translational modifications in liquid-liquid phase separation: A comprehensive review. *Molecular Biomedicine*, *3*(1), 13. https://doi.org/10.1186/s43556-022-00075-2

Li, N., Zhai, Y., Zhang, Y., Li, W., Yang, M., Lei, J., Tye, B.-K., & Gao, N. (2015). Structure of the eukaryotic MCM complex at 3.8 Å. *Nature*, *524*(7564), 186–191. https://doi.org/10.1038/nature14685

Li, P., Banjade, S., Cheng, H.-C., Kim, S., Chen, B., Guo, L., Llaguno, M., Hollingsworth, J. V., King, D. S., Banani, S. F., Russo, P. S., Jiang, Q.-X., Nixon, B. T., & Rosen, M. K. (2012). Phase transitions in the assembly of multivalent signalling proteins. *Nature*, *483*(7389), 336–340. https://doi.org/10.1038/nature10879

Li, T. L., Wang, Z., You, H., Ong, Q., Varanasi, V. J., Dong, M., Lu, B., Paşca, S. P., & Cui, B. (2019). Engineering a Genetically Encoded Magnetic Protein Crystal. *Nano Letters*, *19*(10), 6955–6963. https://doi.org/10.1021/acs.nanolett.9b02266

Li, X., Burton, E. M., & Bhaduri-McIntosh, S. (2017). Chloroquine triggers Epstein-Barr virus replication through phosphorylation of KAP1/TRIM28 in Burkitt lymphoma cells. *PLOS Pathogens*, *13*(3), e1006249. https://doi.org/10.1371/journal.ppat.1006249

Lilyestrom, W., Klein, M. G., Zhang, R., Joachimiak, A., & Chen, X. S. (2006). Crystal structure of SV40 large T-antigen bound to p53: Interplay between a viral oncoprotein and a cellular tumor suppressor. *Genes & Development*, *20*(17), 2373–2382. https://doi.org/10.1101/gad.1456306

Lin, H. J., Upson, R. H., & Simmons, D. T. (1992). Nonspecific DNA binding activity of simian virus 40 large T antigen: Evidence for the cooperation of two regions for full activity. *Journal of Virology*, *66*(9), 5443–5452. https://doi.org/10.1128/JVI.66.9.5443-5452.1992

Lin, Y., Protter, D. S. W., Rosen, M. K., & Parker, R. (2015). Formation and Maturation of Phase-Separated Liquid Droplets by RNA-Binding Proteins. *Molecular Cell*, *60*(2), 208–219. https://doi.org/10.1016/j.molcel.2015.08.018

Liu, W., Krump, N. A., MacDonald, M., & You, J. (2018). Merkel Cell Polyomavirus Infection of Animal Dermal Fibroblasts. *Journal of Virology*, *92*(4), e01610-17. https://doi.org/10.1128/JVI.01610-17

Liu, W., Yang, R., Payne, A. S., Schowalter, R. M., Spurgeon, M. E., Lambert, P. F., Xu, X., Buck, C. B., & You, J. (2016). Identifying the Target Cells and Mechanisms of Merkel Cell Polyomavirus Infection. *Cell Host & Microbe*, *19*(6), 775–787. https://doi.org/10.1016/j.chom.2016.04.024

Liu, X., Hein, J., Richardson, S. C. W., Basse, P. H., Toptan, T., Moore, P. S., Gjoerup, O. V., & Chang, Y. (2011). Merkel Cell Polyomavirus Large T Antigen Disrupts Lysosome Clustering by Translocating Human Vam6p from the Cytoplasm to the Nucleus *. *Journal of Biological Chemistry*, *286*(19), 17079–17090. https://doi.org/10.1074/jbc.M110.192856

Liu, X., & Marmorstein, R. (2007). Structure of the retinoblastoma protein bound to adenovirus E1A reveals the molecular basis for viral oncoprotein inactivation of a tumor suppressor. *Genes & Development*, *21*(21), 2711–2716. https://doi.org/10.1101/gad.1590607

Liu, X., Schuck, S., & Stenlund, A. (2007). Adjacent residues in the E1 initiator beta-hairpin define different roles of the beta-hairpin in Ori melting, helicase loading, and helicase activity. *Molecular Cell*, *25*(6), 825–837. https://doi.org/10.1016/j.molcel.2007.02.009

Loeber, G., Stenger, J. E., Ray, S., Parsons, R. E., Anderson, M. E., & Tegtmeyer, P. (1991). The zinc finger region of simian virus 40 large T antigen is needed for hexamer assembly and origin melting. *Journal of Virology*, *65*(6), 3167–3174. https://doi.org/10.1128/JVI.65.6.3167-3174.1991

Loyo, M., Guerrero-Preston, R., Brait, M., Hoque, M. O., Chuang, A., Kim, M. S., Sharma, R., Liégeois, N. J., Koch, W. M., Califano, J. A., Westra, W. H., & Sidransky, D. (2010). Quantitative detection of Merkel cell virus in human tissues and possible mode of transmission. *International Journal of Cancer*, *126*(12), 2991–2996. https://doi.org/10.1002/ijc.24737

Luo, X., Sanford, D. G., Bullock, P. A., & Bachovchin, W. W. (1996). Solution structure of the origin DNA-binding domain of SV40 T-antigen. *Nature Structural Biology*, *3*(12), Article 12. https://doi.org/10.1038/nsb1296-1034

Manos, M. M., & Gluzman, Y. (1984). Simian virus 40 large T-antigen point mutants that are defective in viral DNA replication but competent in oncogenic transformation. *Molecular and Cellular Biology*, *4*(6), 1125–1133.

Martinez-Zapien, D., Ruiz, F. X., Poirson, J., Mitschler, A., Ramirez, J., Forster, A., Cousido-Siah, A., Masson, M., Vande Pol, S., Podjarny, A., Travé, G., & Zanier, K. (2016). Structure of the E6/E6AP/p53 complex required for HPV-mediated degradation of p53. *Nature*, *529*(7587), 541–545. https://doi.org/10.1038/nature16481

Mastrangelo, I. A., Hough, P. V. C., Wall, J. S., Dodson, M., Dean, F. B., & Hurwrtz, J. (1989). ATP-dependent assembly of double hexamers of SV40 T antigen at the viral origin of DNA replication. *Nature*, *338*(6217), Article 6217. https://doi.org/10.1038/338658a0

McBride, A. A. (2013). The Papillomavirus E2 proteins. *Virology*, *445*(1), 57–79. https://doi.org/10.1016/j.virol.2013.06.006

McCafferty, C. L., Pennington, E. L., Papoulas, O., Taylor, D. W., & Marcotte, E. M. (2023). Does AlphaFold2 model proteins' intracellular conformations? An experimental test using cross-linking mass spectrometry of endogenous ciliary proteins. *Communications Biology*, *6*(1), Article 1. https://doi.org/10.1038/s42003-023-04773-7

McVey, D., Brizuela, L., Mohr, I., Marshak, D. R., Gluzman, Y., & Beach, D. (1989). Phosphorylation of large tumour antigen by cdc2 stimulates SV40 DNA replication. *Nature*, *341*(6242), 503–507. https://doi.org/10.1038/341503a0

Medina, L., Grove, K., & Haltiwanger, R. S. (1998). SV40 large T antigen is modified with O-linked N-acetylglucosamine but not with other forms of glycosylation. *Glycobiology*, *8*(4), 383–391. https://doi.org/10.1093/glycob/8.4.383

Meinke, G., Bullock, P. A., & Bohm, A. (2006). Crystal Structure of the Simian Virus 40 Large T-Antigen Origin-Binding Domain. *Journal of Virology*, *80*(9), 4304–4312. https://doi.org/10.1128/JVI.80.9.4304-4312.2006

Meinke, G., Phelan, P. J., Fradet-Turcotte, A., Bohm, A., Archambault, J., & Bullock, P. A. (2011). Structure-Based Analysis of the Interaction between the Simian Virus 40 T-Antigen Origin Binding Domain and Single-Stranded DNA. *Journal of Virology*, *85*(2), 818–827. https://doi.org/10.1128/JVI.01738-10

Meinke, G., Phelan, P. J., Kalekar, R., Shin, J., Archambault, J., Bohm, A., & Bullock, P. A. (2014). Insights into the initiation of JC virus DNA replication derived from the crystal structure of the T-antigen origin binding domain. *PLoS Pathogens*, *10*(2), e1003966. https://doi.org/10.1371/journal.ppat.1003966

Meinke, G., Phelan, P. J., Shin, J., Gagnon, D., Archambault, J., Bohm, A., & Bullock, P. A. (2016). Structural Based Analyses of the JC Virus T-Antigen F258L Mutant Provides Evidence for DNA Dependent Conformational Changes in the C-Termini of Polyomavirus Origin Binding Domains. *PLOS Pathogens*, *12*(1), e1005362. https://doi.org/10.1371/journal.ppat.1005362

Meinke, G., Phelan, P., Moine, S., Bochkareva, E., Bochkarev, A., Bullock, P. A., & Bohm, A. (2007). The Crystal Structure of the SV40 T-Antigen Origin Binding Domain in Complex with DNA. *PLOS Biology*, *5*(2), e23. https://doi.org/10.1371/journal.pbio.0050023

Mesri, E. A., Feitelson, M., & Munger, K. (2014). HUMAN VIRAL ONCOGENESIS: A CANCER HALLMARKS ANALYSIS. *Cell Host & Microbe*, *15*(3), 266–282. https://doi.org/10.1016/j.chom.2014.02.011

Metrick, M. A., Temple, J. E., & MacDonald, G. (2013). The effects of buffers and pH on the thermal stability, unfolding and substrate binding of RecA. *Biophysical Chemistry*, *184*, 29–36. https://doi.org/10.1016/j.bpc.2013.08.001

Mitrea, D. M., & Kriwacki, R. W. (2016). Phase separation in biology; functional organization of a higher order. *Cell Communication and Signaling*, *14*(1), 1. https://doi.org/10.1186/s12964-015-0125-7

Mittal, A., & Jayaram, B. (2011). The Newest View on Protein Folding: Stoichiometric and Spatial Unity in Structural and Functional Diversity. *Journal of Biomolecular Structure and Dynamics*, *28*(4), 669–674. https://doi.org/10.1080/073911011010524984

Moens, U., Van Ghelue, M., Ludvigsen, M., Korup-Schulz, S., & Ehlers, B. (2015). Early and late promoters of BK polyomavirus, Merkel cell polyomavirus, Trichodysplasia spinulosa-associated polyomavirus and human polyomavirus 12 are among the strongest of all known human polyomaviruses in 10 different cell lines. *The Journal of General Virology*, *96*(8), 2293–2303. https://doi.org/10.1099/vir.0.000181

Monaco, M. C., Jensen, P. N., Hou, J., Durham, L. C., & Major, E. O. (1998). Detection of JC virus DNA in human tonsil tissue: Evidence for site of initial viral infection. *Journal of Virology*, *72*(12), 9918–9923. https://doi.org/10.1128/JVI.72.12.9918-9923.1998

Monahan, Z., Ryan, V. H., Janke, A. M., Burke, K. A., Rhoads, S. N., Zerze, G. H., O'Meally, R., Dignon, G. L., Conicella, A. E., Zheng, W., Best, R. B., Cole, R. N., Mittal, J., Shewmaker, F., & Fawzi, N. L. (2017). Phosphorylation of the FUS low-complexity domain disrupts phase separation, aggregation, and toxicity. *The EMBO Journal*, *36*(20), 2951–2967. https://doi.org/10.15252/embj.201696394

Morris-Love, J., Gee, G. V., O'Hara, B. A., Assetta, B., Atkinson, A. L., Dugan, A. S., Haley, S. A., & Atwood, W. J. (2019). JC Polyomavirus Uses Extracellular Vesicles To Infect Target Cells. *MBio*, *10*(2), e00379-19. https://doi.org/10.1128/mBio.00379-19

Mortimer, E. A. J., Lepow, M. L., Gold, E., Robbins, F. C., Burton, G. J., & Fraumeni, J. F. J. (1981). Long-term follow-up of persons inadvertently inoculated with SV40 as neonates. *New England Journal of Medicine*, *305*(25), 1517–1518.

Müller, D., Ugi, I., Ballas, K., Reiser, P., Henning, R., & Montenarh, M. (1987). The AT-rich sequence of the SV40 control region influences the binding of SV40 T antigen to binding sites II and III. *Virology*, *161*(1), 81–90. https://doi.org/10.1016/0042-6822(87)90173-5

Nakamura, T., Sato, Y., Watanabe, D., Ito, H., Shimonohara, N., Tsuji, T., Nakajima, N., Suzuki, Y., Matsuo, K., Nakagawa, H., Sata, T., & Katano, H. (2010). Nuclear localization of Merkel cell polyomavirus large T antigen in Merkel cell carcinoma. *Virology*, *398*(2), 273–279. https://doi.org/10.1016/j.virol.2009.12.024

Nakanishi, A., Itoh, N., Li, P. P., Handa, H., Liddington, R. C., & Kasamatsu, H. (2007). Minor Capsid Proteins of Simian Virus 40 Are Dispensable for Nucleocapsid Assembly and Cell Entry but Are Required for Nuclear Entry of the Viral Genome. *Journal of Virology*, *81*(8), 3778–3785. https://doi.org/10.1128/JVI.02664-06

Naydenova, K., & Russo, C. J. (2017). Measuring the effects of particle orientation to improve the efficiency of electron cryomicroscopy. *Nature Communications*, *8*(1), Article 1. https://doi.org/10.1038/s41467-017-00782-3

Neumann, F., Borchert, S., Schmidt, C., Reimer, R., Hohenberg, H., Fischer, N., & Grundhoff, A. (2011). Replication, Gene Expression and Particle Production by a Consensus Merkel Cell Polyomavirus (MCPyV) Genome. *PLOS ONE*, *6*(12), e29112. https://doi.org/10.1371/journal.pone.0029112

Neuwald, A. F., Aravind, L., Spouge, J. L., & Koonin, E. V. (1999). AAA+: A Class of Chaperone-Like ATPases Associated with the Assembly, Operation, and Disassembly of Protein Complexes. *Genome Research*, *9*(1), 27–43. https://doi.org/10.1101/gr.9.1.27

Nghiem, P., Kaufman, H. L., Bharmal, M., Mahnke, L., Phatak, H., & Becker, J. C. (2017). Systematic literature review of efficacy, safety and tolerability outcomes of chemotherapy regimens in patients with metastatic Merkel cell carcinoma. *Future Oncology (London, England)*, *13*(14), 1263–1279. https://doi.org/10.2217/fon-2017-0072

Nghiem, P. T., Bhatia, S., Lipson, E. J., Kudchadkar, R. R., Miller, N. J., Annamalai, L., Berry, S., Chartash, E. K., Daud, A., Fling, S. P., Friedlander, P. A., Kluger, H. M., Kohrt, H. E., Lundgren, L., Margolin, K., Mitchell, A., Olencki, T., Pardoll, D. M., Reddy, S. A., … Cheever, M. A. (2016). PD-1 Blockade with Pembrolizumab in Advanced Merkel-Cell Carcinoma. *The New England Journal of Medicine*, *374*(26), 2542–2552. https://doi.org/10.1056/NEJMoa1603702

Nicol, J. T. J., Robinot, R., Carpentier, A., Carandina, G., Mazzoni, E., Tognon, M., Touzé, A., & Coursaget, P. (2013). Age-specific seroprevalences of merkel cell polyomavirus, human polyomaviruses 6, 7, and 9, and trichodysplasia spinulosa-associated polyomavirus. *Clinical and Vaccine Immunology: CVI*, *20*(3), 363–368. https://doi.org/10.1128/CVI.00438-12

Noble, A. J., Dandey, V. P., Wei, H., Brasch, J., Chase, J., Acharya, P., Tan, Y. Z., Zhang, Z., Kim, L. Y., Scapin, G., Rapp, M., Eng, E. T., Rice, W. J., Cheng, A., Negro, C. J., Shapiro, L., Kwong, P. D., Jeruzalmi, D.,

des Georges, A., … Carragher, B. (2018). Routine single particle CryoEM sample and grid characterization by tomography. *ELife*, *7*, e34257. https://doi.org/10.7554/eLife.34257

Noble, A. J., Wei, H., Dandey, V. P., Zhang, Z., Tan, Y. Z., Potter, C. S., & Carragher, B. (2018). Reducing effects of particle adsorption to the air–water interface in cryo-EM. *Nature Methods*, *15*(10), Article 10. https://doi.org/10.1038/s41592-018-0139-3

Noda, T., Satake, M., Yamaguchi, Y., & Ito, Y. (1987). Cooperation of middle and small T antigens of polyomavirus in transformation of established fibroblast and epithelial-like cell lines. *Journal of Virology*, *61*(7), 2253–2263. https://doi.org/10.1128/jvi.61.7.2253-2263.1987

Norkin, L. C., Anderson, H. A., Wolfrom, S. A., & Oppenheim, A. (2002). Caveolar endocytosis of simian virus 40 is followed by brefeldin A-sensitive transport to the endoplasmic reticulum, where the virus disassembles. *Journal of Virology*, *76*(10), 5156–5166. https://doi.org/10.1128/jvi.76.10.5156-5166.2002

Nott, T. J., Petsalaki, E., Farber, P., Jervis, D., Fussner, E., Plochowietz, A., Craggs, T. D., Bazett-Jones, D. P., Pawson, T., Forman-Kay, J. D., & Baldwin, A. J. (2015). Phase Transition of a Disordered Nuage Protein Generates Environmentally Responsive Membraneless Organelles. *Molecular Cell*, *57*(5), 936–947. https://doi.org/10.1016/j.molcel.2015.01.013

Nwogu, N., Boyne, J. R., Dobson, S. J., Poterlowicz, K., Blair, G. E., Macdonald, A., Mankouri, J., & Whitehouse, A. (2018). Cellular sheddases are induced by Merkel cell polyomavirus small tumour antigen to mediate cell dissociation and invasiveness. *PLoS Pathogens*, *14*(9), e1007276. https://doi.org/10.1371/journal.ppat.1007276

Nwogu, N., Ortiz, L. E., & Kwun, H. J. (2020a). Merkel Cell Polyomavirus Large T Antigen Unique Domain Regulates Its Own Protein Stability and Cell Growth. *Viruses*, *12*(9), Article 9. https://doi.org/10.3390/v12091043

Nwogu, N., Ortiz, L. E., & Kwun, H. J. (2020b). Surface charge of Merkel cell polyomavirus small T antigen determines cell transformation through allosteric FBW7 WD40 domain targeting. *Oncogenesis*, *9*(5), 53. https://doi.org/10.1038/s41389-020-0235-y

O'Donnell, M., & Li, H. (2018). The ring shaped hexameric helicases that function at DNA replication forks. *Nature Structural & Molecular Biology*, *25*(2), 122–130. https://doi.org/10.1038/s41594-018-0024-x

O'Hara, B. A., Morris-Love, J., Gee, G. V., Haley, S. A., & Atwood, W. J. (2020). JC Virus infected choroid plexus epithelial cells produce extracellular vesicles that infect glial cells independently of the virus attachment receptor. *PLoS Pathogens*, *16*(3), e1008371. https://doi.org/10.1371/journal.ppat.1008371

Oldfield, C. J., Meng, J., Yang, J. Y., Yang, M. Q., Uversky, V. N., & Dunker, A. K. (2008). Flexible nets: Disorder and induced fit in the associations of p53 and 14-3-3 with their partners. *BMC Genomics*, *9 Suppl 1*(Suppl 1), S1. https://doi.org/10.1186/1471-2164-9-S1-S1

Ondov, B. D., Starrett, G. J., Sappington, A., Kostic, A., Koren, S., Buck, C. B., & Phillippy, A. M. (2019). Mash Screen: High-throughput sequence containment estimation for genome discovery. *Genome Biology*, *20*, 232. https://doi.org/10.1186/s13059-019-1841-x

Ortiz, L. E., Pham, A. M., & Kwun, H. J. (2021). Identification of the Merkel Cell Polyomavirus Large Tumor Antigen Ubiquitin Conjugation Residue. *International Journal of Molecular Sciences*, *22*(13), 7169. https://doi.org/10.3390/ijms22137169

Padgett, B. L., Walker, D. L., ZuRhein, G. M., Eckroade, R. J., & Dessel, B. H. (1971). Cultivation of papova-like virus from human brain with progressive multifocal leucoencephalopathy. *Lancet (London, England)*, *1*(7712), 1257–1260. https://doi.org/10.1016/s0140-6736(71)91777-6

Pande, A., Pande, J., Asherie, N., Lomakin, A., Ogun, O., King, J., & Benedek, G. B. (2001). Crystal cataracts: Human genetic cataract caused by protein crystallization. *Proceedings of the National Academy of Sciences*, *98*(11), 6116–6120. https://doi.org/10.1073/pnas.101124798

Papassideri, I. S. (2007). Crystalline yolk spheroids in Drosophila melanogaster oocyte: Freeze fracture and two-dimensional reconstruction analysis. *Journal of Insect Physiology*, *53*(4), 370–376. https://doi.org/10.1016/j.jinsphys.2006.12.011

Park, D. E., Cheng, J., Berrios, C., Montero, J., Cortés-Cros, M., Ferretti, S., Arora, R., Tillgren, M. L., Gokhale, P. C., & DeCaprio, J. A. (2019). Dual inhibition of MDM2 and MDM4 in virus-positive Merkel cell carcinoma enhances the p53 response. *Proceedings of the National Academy of Sciences of the United States of America*, *116*(3), 1027–1032. https://doi.org/10.1073/pnas.1818798116

Parsons, R., & Tegtmeyer, P. (1992). Spacing is crucial for coordination of domain functions within the simian virus 40 core origin of replication. *Journal of Virology*, *66*(4), 1933–1942. https://doi.org/10.1128/jvi.66.4.1933-1942.1992

Pastrana, D. V., Tolstov, Y. L., Becker, J. C., Moore, P. S., Chang, Y., & Buck, C. B. (2009). Quantitation of Human Seroresponsiveness to Merkel Cell Polyomavirus. *PLOS Pathogens*, *5*(9), e1000578. https://doi.org/10.1371/journal.ppat.1000578

Pastrana, D. V., Wieland, U., Silling, S., Buck, C. B., & Pfister, H. (2012). Positive correlation between Merkel cell polyomavirus viral load and capsid-specific antibody titer. *Medical Microbiology and Immunology*, *201*(1), 17–23. https://doi.org/10.1007/s00430-011-0200-7

Patel, S. S., & Picha, K. M. (2000). Structure and function of hexameric helicases. *Annual Review of Biochemistry*, *69*, 651–697. https://doi.org/10.1146/annurev.biochem.69.1.651

Paulson, K. G., Carter, J. J., Johnson, L. G., Cahill, K. W., Iyer, J. G., Schrama, D., Becker, J. C., Madeleine, M. M., Nghiem, P., & Galloway, D. A. (2010). Antibodies to Merkel Cell Polyomavirus T Antigen Oncoproteins Reflect Tumor Burden in Merkel Cell Carcinoma Patients. *Cancer Research*, *70*(21), 8388–8397. https://doi.org/10.1158/0008-5472.CAN-10-2128

Paulson, K. G., Iyer, J. G., Tegeder, A. R., Thibodeau, R., Schelter, J., Koba, S., Schrama, D., Simonson, W. T., Lemos, B. D., Byrd, D. R., Koelle, D. M., Galloway, D. A., Leonard, J. H., Madeleine, M. M., Argenyi, Z. B., Disis, M. L., Becker, J. C., Cleary, M. A., & Nghiem, P. (2011). Transcriptome-Wide Studies of Merkel Cell Carcinoma and Validation of Intratumoral CD8+ Lymphocyte Invasion As an Independent Predictor of Survival. *Journal of Clinical Oncology*, *29*(12), 1539–1546. https://doi.org/10.1200/JCO.2010.30.6308

Paulson, K. G., Park, S. Y., Vandeven, N. A., Lachance, K., Thomas, H., Chapuis, A. G., Harms, K. L., Thompson, J. A., Bhatia, S., Stang, A., & Nghiem, P. (2018). Merkel cell carcinoma: Current US incidence and projected increases based on changing demographics. *Journal of the American Academy of Dermatology*, *78*(3), 457-463.e2. https://doi.org/10.1016/j.jaad.2017.10.028

Paulson, K. G., Tegeder, A., Willmes, C., Iyer, J. G., Afanasiev, O. K., Schrama, D., Koba, S., Thibodeau, R., Nagase, K., Simonson, W. T., Seo, A., Koelle, D. M., Madeleine, M., Bhatia, S., Nakajima, H., Sano, S., Hardwick, J. S., Disis, M. L., Cleary, M. A., … Nghiem, P. (2014). Downregulation of MHC-I Expression Is Prevalent but Reversible in Merkel Cell Carcinoma. *Cancer Immunology Research*, *2*(11), 1071–1079. https://doi.org/10.1158/2326-6066.CIR-14-0005

Peng, Q., Wang, L., Qin, Z., Wang, J., Zheng, X., Wei, L., Zhang, X., Zhang, X., Liu, C., Li, Z., Wu, Y., Li, G., Yan, Q., & Ma, J. (2020). Phase Separation of Epstein-Barr Virus EBNA2 and Its Coactivator EBNALP Controls Gene Expression. *Journal of Virology*, *94*(7), e01771-19. https://doi.org/10.1128/JVI.01771-19

Peng, Q., Wang, L., Wang, J., Liu, C., Zheng, X., Zhang, X., Wei, L., Li, Z., Wu, Y., Wen, Y., Cao, P., Liao, Q., Yan, Q., & Ma, J. (2021). Epstein-Barr virus EBNA2 phase separation regulates cancer-associated alternative RNA splicing patterns. *Clinical and Translational Medicine*, *11*(8), e504. https://doi.org/10.1002/ctm2.504

Pettersen, E. F., Goddard, T. D., Huang, C. C., Meng, E. C., Couch, G. S., Croll, T. I., Morris, J. H., & Ferrin, T. E. (2021). UCSF ChimeraX: Structure visualization for researchers, educators, and developers. *Protein Science: A Publication of the Protein Society*, *30*(1), 70–82. https://doi.org/10.1002/pro.3943

Pipas, J. M., & Levine, A. J. (2001). Role of T antigen interactions with p53 in tumorigenesis. *Seminars in Cancer Biology*, *11*(1), 23–30. https://doi.org/10.1006/scbi.2000.0343

Ponomarenko, E. A., Poverennaya, E. V., Ilgisonis, E. V., Pyatnitskiy, M. A., Kopylov, A. T., Zgoda, V. G., Lisitsa, A. V., & Archakov, A. I. (2016). The Size of the Human Proteome: The Width and Depth.

International Journal of Analytical Chemistry, *2016*, 7436849. https://doi.org/10.1155/2016/7436849

Prives, C., Covey, L., Scheller, A., & Gluzman, Y. (1983). DNA-binding properties of simian virus 40 T-antigen mutants defective in viral DNA replication. *Molecular and Cellular Biology*, *3*(11), 1958–1966.

Pyeon, D., Pearce, S. M., Lank, S. M., Ahlquist, P., & Lambert, P. F. (2009). Establishment of Human Papillomavirus Infection Requires Cell Cycle Progression. *PLOS Pathogens*, *5*(2), e1000318. https://doi.org/10.1371/journal.ppat.1000318

Qin, J. Y., Zhang, L., Clift, K. L., Hulur, I., Xiang, A. P., Ren, B.-Z., & Lahn, B. T. (2010). Systematic Comparison of Constitutive Promoters and the Doxycycline-Inducible Promoter. *PLOS ONE*, *5*(5), e10611. https://doi.org/10.1371/journal.pone.0010611

Rana, S., Maples, P. B., Senzer, N., & Nemunaitis, J. (2008). Stathmin 1: A novel therapeutic target for anticancer activity. *Expert Review of Anticancer Therapy*, *8*(9), 1461–1470. https://doi.org/10.1586/14737140.8.9.1461

Randhawa, P., Zeng, G., Bueno, M., Salgarkar, A., Lesniak, A., Isse, K., Seyb, K., Perry, A., Charles, I., Hustus, C., Huang, M., Smith, M., & Glicksman, M. A. (2014). Inhibition of large T antigen ATPase activity as a potential strategy to develop anti-polyomavirus JC drugs. *Antiviral Research*, *112*, 113–119. https://doi.org/10.1016/j.antiviral.2014.10.004

Razinkov, I., Dandey, V., Wei, H., Zhang, Z., Melnekoff, D., Rice, W. J., Wigge, C., Potter, C. S., & Carragher, B. (2016). A new method for vitrifying samples for cryoEM. *Journal of Structural Biology*, *195*(2), 190–198. https://doi.org/10.1016/j.jsb.2016.06.001

Redecke, L., Nass, K., DePonte, D. P., White, T. A., Rehders, D., Barty, A., Stellato, F., Liang, M., Barends, T. R. M., Boutet, S., Williams, G. J., Messerschmidt, M., Seibert, M. M., Aquila, A., Arnlund, D., Bajt, S., Barth, T., Bogan, M. J., Caleman, C., … Chapman, H. N. (2013). Natively Inhibited Trypanosoma brucei Cathepsin B Structure Determined by Using an X-ray Laser. *Science*, *339*(6116), 227–230. https://doi.org/10.1126/science.1229663

Remus, D., Beuron, F., Tolun, G., Griffith, J. D., Morris, E. P., & Diffley, J. F. X. (2009). Concerted loading of Mcm2-7 double hexamers around DNA during DNA replication origin licensing. *Cell*, *139*(4), 719–730. https://doi.org/10.1016/j.cell.2009.10.015

Reynisdóttir, I., Lorimer, H. E., Friedman, P. N., Wang, E. H., & Prives, C. (1993). Phosphorylation and active ATP hydrolysis are not required for SV40 T antigen hexamer formation. *Journal of Biological Chemistry*, *268*(33), 24647–24654. https://doi.org/10.1016/S0021-9258(19)74515-3

Richards, K. F., Guastafierro, A., Shuda, M., Toptan, T., Moore, P. S., & Chang, Y. (2015). Merkel cell polyomavirus T antigens promote cell proliferation and inflammatory cytokine gene expression. *The Journal of General Virology*, *96*(12), 3532–3544. https://doi.org/10.1099/jgv.0.000287

Rohl, C. A., Strauss, C. E. M., Chivian, D., & Baker, D. (2004). Modeling structurally variable regions in homologous proteins with rosetta. *Proteins*, *55*(3), 656–677. https://doi.org/10.1002/prot.10629

Rohou, A., & Grigorieff, N. (2015). CTFFIND4: Fast and accurate defocus estimation from electron micrographs. *Journal of Structural Biology*, *192*(2), 216–221. https://doi.org/10.1016/j.jsb.2015.08.008

Rosen, S. T., Gould, V. E., Salwen, H. R., Herst, C. V., Le Beau, M. M., Lee, I., Bauer, K., Marder, R. J., Andersen, R., & Kies, M. S. (1987). Establishment and characterization of a neuroendocrine skin carcinoma cell line. *Laboratory Investigation; a Journal of Technical Methods and Pathology*, *56*(3), 302–312.

Rossmann, M. G., & Blow, D. M. (1962). The detection of sub-units within the crystallographic asymmetric unit. *Acta Crystallographica*, *15*(1), 24–31. https://doi.org/10.1107/S0365110X62000067

Ruff, K. M., & Pappu, R. V. (2021). AlphaFold and Implications for Intrinsically Disordered Proteins. *Journal of Molecular Biology*, *433*(20), 167208. https://doi.org/10.1016/j.jmb.2021.167208

Rundell, K., Major, E. O., & Lampert, M. (1981). Association of cellular 56,000- and 32,000-molecular-weight protein with BK virus and polyoma virus t-antigens. *Journal of Virology*, *37*(3), 1090–1093.

Sabari, B. R., Dall'Agnese, A., Boija, A., Klein, I. A., Coffey, E. L., Shrinivas, K., Abraham, B. J., Hannett, N. M., Zamudio, A. V., Manteiga, J. C., Li, C. H., Guo, Y. E., Day, D. S., Schuijers, J., Vasile, E., Malik, S., Hnisz, D., Lee, T. I., Cisse, I. I., … Young, R. A. (2018). Coactivator condensation at super-enhancers links phase separation and gene control. *Science*, *361*(6400), eaar3958. https://doi.org/10.1126/science.aar3958

Sahli, R., Freund, R., Dubensky, T., Garcea, R., Bronson, R., & Benjamin, T. (1993). Defect in entry and altered pathogenicity of a polyoma virus mutant blocked in VP2 myristylation. *Virology*, *192*(1), 142–153. https://doi.org/10.1006/viro.1993.1016

Samel, S. A., Fernández-Cid, A., Sun, J., Riera, A., Tognetti, S., Herrera, M. C., Li, H., & Speck, C. (2014). A unique DNA entry gate serves for regulated loading of the eukaryotic replicative helicase MCM2–7 onto DNA. *Genes & Development*, *28*(15), 1653–1666. https://doi.org/10.1101/gad.242404.114

Samimi, M. (2019). Immune Checkpoint Inhibitors and Beyond: An Overview of Immune-Based Therapies in Merkel Cell Carcinoma. *American Journal of Clinical Dermatology*, *20*(3), 391–407. https://doi.org/10.1007/s40257-019-00427-9

Sanders, C. M., Kovalevskiy, O. V., Sizov, D., Lebedev, A. A., Isupov, M. N., & Antson, A. A. (2007). Papillomavirus E1 helicase assembly maintains an asymmetric state in the absence of DNA and nucleotide cofactors. *Nucleic Acids Research*, *35*(19), 6451–6457. https://doi.org/10.1093/nar/gkm705

Saraste, M., Sibbald, P. R., & Wittinghofer, A. (1990). The P-loop—A common motif in ATP- and GTP-binding proteins. *Trends in Biochemical Sciences*, *15*(11), 430–434. https://doi.org/10.1016/0968-0004(90)90281-f

Sarkari, F., Sanchez-Alcaraz, T., Wang, S., Holowaty, M. N., Sheng, Y., & Frappier, L. (2009). EBNA1-Mediated Recruitment of a Histone H2B Deubiquitylating Complex to the Epstein-Barr Virus Latent Origin of DNA Replication. *PLOS Pathogens*, *5*(10), e1000624. https://doi.org/10.1371/journal.ppat.1000624

Sato, Y., Kamura, T., Shirata, N., Murata, T., Kudoh, A., Iwahori, S., Nakayama, S., Isomura, H., Nishiyama, Y., & Tsurumi, T. (2009). Degradation of phosphorylated p53 by viral protein-ECS E3 ligase complex. *PLoS Pathogens*, *5*(7), e1000530. https://doi.org/10.1371/journal.ppat.1000530

Sawaya, M. R., Cascio, D., Gingery, M., Rodriguez, J., Goldschmidt, L., Colletier, J.-P., Messerschmidt, M. M., Boutet, S., Koglin, J. E., Williams, G. J., Brewster, A. S., Nass, K., Hattne, J., Botha, S., Doak, R. B., Shoeman, R. L., DePonte, D. P., Park, H.-W., Federici, B. A., … Eisenberg, D. S. (2014). Protein crystal structure obtained at 2.9 Å resolution from injecting bacterial cells into an X-ray free-electron laser beam. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(35), 12769–12774. https://doi.org/10.1073/pnas.1413456111

Scampa, M., Merat, R., Tzika, E., Kalbermatten, D. F., & Oranges, C. M. (2022). Survival outcomes and epidemiology of Merkel cell carcinoma of the lower limb and hip: A Surveillance, Epidemiology, and End Results analysis 2000-2018. *JAAD International*, *7*, 13–21. https://doi.org/10.1016/j.jdin.2021.12.010

Scheffner, M., Wessel, R., & Stahl, H. (1989). Sequence independent duplex DNA opening reaction catalysed by SV40 large tumor antigen. *Nucleic Acids Research*, *17*(1), 93–106. https://doi.org/10.1093/nar/17.1.93

Scheres, S. H. W. (2012). RELION: Implementation of a Bayesian approach to cryo-EM structure determination. *Journal of Structural Biology*, *180*(3), 519–530. https://doi.org/10.1016/j.jsb.2012.09.006

Schlemeyer, T., Ohnezeit, D., Virdi, S., Körner, C., Weißelberg, S., Starzonek, S., Schumacher, U., Grundhoff, A., Indenbirken, D., Albertini, S., & Fischer, N. (2022). Merkel Cell Carcinoma and Immune Evasion: Merkel Cell Polyomavirus Small T-Antigen–Induced Surface Changes Can Be Reverted by Therapeutic Intervention. *The Journal of Investigative Dermatology*, *142*(11), 3071-3081.e13. https://doi.org/10.1016/j.jid.2022.04.029

Schmit, J. D., & Dill, K. (2012). Growth Rates of Protein Crystals. *Journal of the American Chemical Society*, *134*(9), 3934–3937. https://doi.org/10.1021/ja207336r

Schneider, C. A., Rasband, W. S., & Eliceiri, K. W. (2012). NIH Image to ImageJ: 25 years of image analysis. *Nature Methods*, *9*(7), Article 7. https://doi.org/10.1038/nmeth.2089

Schönherr, R., Klinge, M., Rudolph, J. M., Fita, K., Rehders, D., Lübber, F., Schneegans, S., Majoul, I. V., Duszenko, M., Betzel, C., Brandariz-Nuñez, A., Martinez-Costas, J., Duden, R., & Redecke, L. (2015). Real-time investigation of dynamic protein crystallization in living cells. *Structural Dynamics*, *2*(4). https://doi.org/10.1063/1.4921591

Schönherr, R., Rudolph, J. M., & Redecke, L. (2018). Protein crystallization in living cells. *Biological Chemistry*, *399*(7), 751–772. https://doi.org/10.1515/hsz-2018-0158

Schowalter, R. M., & Buck, C. B. (2013). The Merkel Cell Polyomavirus Minor Capsid Protein. *PLOS Pathogens*, *9*(8), e1003558. https://doi.org/10.1371/journal.ppat.1003558

Schowalter, R. M., Pastrana, D. V., & Buck, C. B. (2011). Glycosaminoglycans and Sialylated Glycans Sequentially Facilitate Merkel Cell Polyomavirus Infectious Entry. *PLOS Pathogens*, *7*(7), e1002161. https://doi.org/10.1371/journal.ppat.1002161

Schowalter, R. M., Pastrana, D. V., Pumphrey, K. A., Moyer, A. L., & Buck, C. B. (2010). Merkel cell polyomavirus and two previously unknown polyomaviruses are chronically shed from human skin. *Cell Host & Microbe*, *7*(6), 509–515. https://doi.org/10.1016/j.chom.2010.05.006

Schrama, D., Ugurel, S., & Becker, J. C. (2012). Merkel cell carcinoma: Recent insights and new treatment options. *Current Opinion in Oncology*, *24*(2), 141–149. https://doi.org/10.1097/CCO.0b013e32834fc9fe

Schuck, S., & Stenlund, A. (2011). Mechanistic Analysis of Local Ori Melting and Helicase Assembly by the Papillomavirus E1 Protein. *Molecular Cell*, *43*(5), 776–787. https://doi.org/10.1016/j.molcel.2011.06.026

Seguin, S. P., Ireland, A. W., Gupta, T., Wright, C. M., Miyata, Y., Wipf, P., Pipas, J. M., Gestwicki, J. E., & Brodsky, J. L. (2012). A Screen for Modulators of Large T Antigen's ATPase Activity Uncovers Novel Inhibitors of Simian Virus 40 and BK Virus Replication. *Antiviral Research*, *96*(1), 70–81. https://doi.org/10.1016/j.antiviral.2012.07.012

SenGupta, D. J., & Borowiec, J. A. (1992). Strand-specific recognition of a synthetic DNA replication fork by the SV40 large tumor antigen. *Science (New York, N.Y.)*, *256*(5064), 1656–1661. https://doi.org/10.1126/science.256.5064.1656

Seo, G. J., Chen, C. J., & Sullivan, C. S. (2009). Merkel cell polyomavirus encodes a microRNA with the ability to autoregulate viral gene expression. *Virology*, *383*(2), 183–187.

Seo, G. J., Fink, L. H. L., O'Hara, B., Atwood, W. J., & Sullivan, C. S. (2008). Evolutionarily conserved function of a viral microRNA. *Journal of Virology*, *82*(20), 9823–9828. https://doi.org/10.1128/jvi.01144-08

Seyffert, M., Georgi, F., Tobler, K., Bourqui, L., Anfossi, M., Michaelsen, K., Vogt, B., Greber, U. F., & Fraefel, C. (2021). The HSV-1 Transcription Factor ICP4 Confers Liquid-Like Properties to Viral Replication Compartments. *International Journal of Molecular Sciences*, *22*(9), Article 9. https://doi.org/10.3390/ijms22094447

Shen, C., Tung, C., Chao, C., Jou, Y., Huang, S., Meng, M., Chang, D., & Chen, P. (2021). The differential presence of human polyomaviruses, JCPyV and BKPyV, in prostate cancer and benign prostate hypertrophy tissues. *BMC Cancer*, *21*(1), 1141. https://doi.org/10.1186/s12885-021-08862-w

Shen, J., Gai, D., Patrick, A., Greenleaf, W. B., & Chen, X. S. (2005). The roles of the residues on the channel β-hairpin and loop structures of simian virus 40 hexameric helicase. *Proceedings of*

the National Academy of Sciences, *102*(32), 11248–11253. https://doi.org/10.1073/pnas.0409646102

Shi, X., & Jarvis, D. L. (2007). Protein N-Glycosylation in the Baculovirus-Insect Cell System. *Current Drug Targets*, *8*(10), 1116–1125.

Shi, Y., Dodson, G. E., Shaikh, S., Rundell, K., & Tibbetts, R. S. (2005). Ataxia-telangiectasia-mutated (ATM) is a T-antigen kinase that controls SV40 viral replication in vivo. *The Journal of Biological Chemistry*, *280*(48), 40195–40200. https://doi.org/10.1074/jbc.C500400200

Shimazu, T., Komatsu, Y., Nakayama, K. I., Fukazawa, H., Horinouchi, S., & Yoshida, M. (2006). Regulation of SV40 large T-antigen stability by reversible acetylation. *Oncogene*, *25*(56), Article 56. https://doi.org/10.1038/sj.onc.1209731

Shoraka, H. R., Aboubakri, O., Naghibzadeh-Tahami, A., Mollaei, H. R., Bagherinezhad, Z., Malekpour Afshar, R., & Shahesmaeili, A. (2020). Prevalence of JC and BK viruses in Patients with Colorectal Cancer: A Systematic Review and Meta- Analysis. *Asian Pacific Journal of Cancer Prevention : APJCP*, *21*(6), 1499–1509. https://doi.org/10.31557/APJCP.2020.21.6.1499

Shorter, J. (2017). Liquidizing FUS via prion-like domain phosphorylation. *The EMBO Journal*, *36*(20), 2925–2927. https://doi.org/10.15252/embj.201798078

Shuda, M., Arora, R., Kwun, H. J., Feng, H., Sarid, R., Fernández-Figueras, M.-T., Tolstov, Y., Gjoerup, O., Mansukhani, M. M., Swerdlow, S. H., Chaudhary, P. M., Kirkwood, J. M., Nalesnik, M. A., Kant, J. A., Weiss, L. M., Moore, P. S., & Chang, Y. (2009). Human Merkel cell polyomavirus infection I. MCV T antigen expression in Merkel cell carcinoma, lymphoid tissues and lymphoid tumors. *International Journal of Cancer*, *125*(6), 1243–1249. https://doi.org/10.1002/ijc.24510

Shuda, M., Chang, Y., & Moore, P. S. (2014). Merkel cell polyomavirus-positive Merkel cell carcinoma requires viral small T-antigen for cell proliferation. *The Journal of Investigative Dermatology*, *134*(5), 1479–1481. https://doi.org/10.1038/jid.2013.483

Shuda, M., Feng, H., Kwun, H. J., Rosen, S. T., Gjoerup, O., Moore, P. S., & Chang, Y. (2008). T antigen mutations are a human tumor-specific signature for Merkel cell polyomavirus. *Proceedings of the National Academy of Sciences*, *105*(42), 16272–16277. https://doi.org/10.1073/pnas.0806526105

Shuda, M., Guastafierro, A., Geng, X., Shuda, Y., Ostrowski, S. M., Lukianov, S., Jenkins, F. J., Honda, K., Maricich, S. M., Moore, P. S., & Chang, Y. (2015). Merkel Cell Polyomavirus Small T Antigen Induces Cancer and Embryonic Merkel Cell Proliferation in a Transgenic Mouse Model. *PLOS ONE*, *10*(11), e0142329. https://doi.org/10.1371/journal.pone.0142329

Shuda, M., Kwun, H. J., Feng, H., Chang, Y., & Moore, P. S. (2011). Human Merkel cell polyomavirus small T antigen is an oncoprotein targeting the 4E-BP1 translation regulator. *The Journal of Clinical Investigation*, *121*(9). https://doi.org/10.1172/JCI46323

Siebels, S., Czech-Sioli, M., Spohn, M., Schmidt, C., Theiss, J., Indenbirken, D., Günther, T., Grundhoff, A., & Fischer, N. (2020). Merkel Cell Polyomavirus DNA Replication Induces Senescence in Human Dermal Fibroblasts in a Kap1/Trim28-Dependent Manner. *MBio*, *11*(2), e00142-20. https://doi.org/10.1128/mBio.00142-20

Simanis, V., & Lane, D. P. (1985). An immunoaffinity purification procedure for sv40 large t antigen. *Virology*, *144*(1), 88–100. https://doi.org/10.1016/0042-6822(85)90308-3

Singleton, M. R., Sawaya, M. R., Ellenberger, T., & Wigley, D. B. (2000). Crystal structure of T7 gene 4 ring helicase indicates a mechanism for sequential hydrolysis of nucleotides. *Cell*, *101*(6), 589–600. https://doi.org/10.1016/s0092-8674(00)80871-5

Small, M. B., Gluzman, Y., & Ozer, H. L. (1982). Enhanced transformation of human fibroblasts by origin-defective simian virus 40. *Nature*, *296*(5858), Article 5858. https://doi.org/10.1038/296671a0

Smelkova, N. V., & Borowiec, J. A. (1998). Synthetic DNA Replication Bubbles Bound and Unwound with Twofold Symmetry by a Simian Virus 40 T-Antigen Double Hexamer. *Journal of Virology*, *72*(11), 8676–8681.

Sontag, E., Sontag, J. M., & Garcia, A. (1997). Protein phosphatase 2A is a critical regulator of protein kinase C zeta signaling targeted by SV40 small t to promote cell growth and NF-kappaB activation. *The EMBO Journal*, *16*(18), 5662–5671. https://doi.org/10.1093/emboj/16.18.5662

Southworth-Davies, R. J., Medina, M. A., Carmichael, I., & Garman, E. F. (2007). Observation of Decreased Radiation Damage at Higher Dose Rates in Room Temperature Protein Crystallography. *Structure*, *15*(12), 1531–1541. https://doi.org/10.1016/j.str.2007.10.013

Sowd, G. A., & Fanning, E. (2012). A Wolf in Sheep's Clothing: SV40 Co-opts Host Genome Maintenance Proteins to Replicate Viral DNA. *PLOS Pathogens*, *8*(11), e1002994. https://doi.org/10.1371/journal.ppat.1002994

Spitkovsky, D., Hehner, S. P., Hofmann, T. G., Möller, A., & Schmitz, M. L. (2002). The Human Papillomavirus Oncoprotein E7 Attenuates NF-κB Activation by Targeting the IκB Kinase Complex *. *Journal of Biological Chemistry*, *277*(28), 25576–25582. https://doi.org/10.1074/jbc.M201884200

Spriggs, C. C., & Laimins, L. A. (2017). Human Papillomavirus and the DNA Damage Response: Exploiting Host Repair Pathways for Viral Replication. *Viruses*, *9*(8), 232. https://doi.org/10.3390/v9080232

Stakaitytė, G., Nwogu, N., Dobson, S. J., Knight, L. M., Wasson, C. W., Salguero, F. J., Blackbourn, D. J., Blair, G. E., Mankouri, J., Macdonald, A., & Whitehouse, A. (2018). Merkel Cell Polyomavirus Small T Antigen Drives Cell Motility via Rho-GTPase-Induced Filopodium Formation. *Journal of Virology*, *92*(2), e00940-17. https://doi.org/10.1128/JVI.00940-17

Starrett, G. J., Thakuria, M., Chen, T., Marcelus, C., Cheng, J., Nomburg, J., Thorner, A. R., Slevin, M. K., Powers, W., Burns, R. T., Perry, C., Piris, A., Kuo, F. C., Rabinowits, G., Giobbie-Hurder, A., MacConaill, L. E., & DeCaprio, J. A. (2020). Clinical and molecular characterization of virus-positive and virus-negative Merkel cell carcinoma. *Genome Medicine*, *12*(1), 30. https://doi.org/10.1186/s13073-020-00727-4

Stewart, S. E., Eddy, B. E., & Borgese, N. (1958). Neoplasms in Mice Inoculated with a Tumor Agent Carried in Tissue Culture2. *JNCI: Journal of the National Cancer Institute*, *20*(6), 1223–1243. https://doi.org/10.1093/jnci/20.6.1223

Stewart, S. E., Eddy, B. E., Gochenour, A. M., Borgese, N. G., & Grubbs, G. E. (1957). The induction of neoplasms with a substance released from mouse tumors by tissue culture. *Virology*, *3*(2), 380–400. https://doi.org/10.1016/0042-6822(57)90100-9

Stillman, B. (2008). DNA polymerases at the replication fork in eukaryotes. *Molecular Cell*, *30*(3), 259–260. https://doi.org/10.1016/j.molcel.2008.04.011

Stinson, B. M., Baytshtok, V., Schmitz, K. R., Baker, T. A., & Sauer, R. T. (2015). Subunit asymmetry and roles of conformational switching in the hexameric AAA+ ring of ClpX. *Nature Structural & Molecular Biology*, *22*(5), 411–416. https://doi.org/10.1038/nsmb.3012

Strom, T., Carr, M., Zager, J. S., Naghavi, A., Smith, F. O., Cruse, C. W., Messina, J. L., Russell, J., Rao, N. G., Fulp, W., Kim, S., Torres-Roca, J. F., Padhya, T. A., Sondak, V. K., Trotti, A. M., Harrison, L. B., & Caudell, J. J. (2016). Radiation Therapy is Associated with Improved Outcomes in Merkel Cell Carcinoma. *Annals of Surgical Oncology*, *23*(11), 3572–3578. https://doi.org/10.1245/s10434-016-5293-1

Stubdal, H., Zalvide, J., Campbell, K. S., Schweitzer, C., Roberts, T. M., & DeCaprio, J. A. (1997). Inactivation of pRB-related proteins p130 and p107 mediated by the J domain of simian virus 40 large T antigen. *Molecular and Cellular Biology*, *17*(9), 4979–4990. https://doi.org/10.1128/MCB.17.9.4979

Sullivan, C. S., Cantalupo, P., & Pipas, J. M. (2000). The Molecular Chaperone Activity of Simian Virus 40 Large T Antigen Is Required To Disrupt Rb-E2F Family Complexes by an ATP-Dependent Mechanism. *Molecular and Cellular Biology*, *20*(17), 6233–6243. https://doi.org/10.1128/MCB.20.17.6233-6243.2000

Sullivan, C. S., Sung, C. K., Pack, C. D., Grundhoff, A., Lukacher, A. E., Benjamin, T. L., & Ganem, D. (2009). Murine Polyomavirus encodes a microRNA that cleaves early RNA transcripts but is not essential for experimental infection. *Virology*, *387*(1), 157–167. https://doi.org/10.1016/j.virol.2009.02.017

Sun, R., Liang, D., Gao, Y., & Lan, K. (2014). Kaposi's Sarcoma-Associated Herpesvirus-Encoded LANA Interacts with Host KAP1 To Facilitate Establishment of Viral Latency. *Journal of Virology*, *88*(13), 7331–7344. https://doi.org/10.1128/JVI.00596-14

Sunshine, J. C., Jahchan, N. S., Sage, J., & Choi, J. (2018). Are there multiple cells of origin of Merkel cell carcinoma? *Oncogene*, *37*(11), Article 11. https://doi.org/10.1038/s41388-017-0073-3

Sweet, B. H., & Hilleman, M. R. (1960). The Vacuolating Virus, S.V.40. *Proceedings of the Society for Experimental Biology and Medicine*, *105*(2), 420–427. https://doi.org/10.3181/00379727-105-26128

Symonds, H., Krall, L., Remington, L., Saenz-Robles, M., Lowe, S., Jacks, T., & Van Dyke, T. (1994). P53-dependent apoptosis suppresses tumor growth and progression in vivo. *Cell*, *78*(4), 703–711. https://doi.org/10.1016/0092-8674(94)90534-7

Taherkhani, R., & Farshadpour, F. (2022). BK and JC polyomaviruses and risk of urothelial bladder carcinoma: A preliminary study in the northern shores of Persian Gulf, Iran. *Infectious Agents and Cancer*, *17*(1), 50. https://doi.org/10.1186/s13027-022-00463-x

Tai, P. T., Yu, E., Winquist, E., Hammond, A., Stitt, L., Tonita, J., & Gilchrist, J. (2000). Chemotherapy in neuroendocrine/Merkel cell carcinoma of the skin: Case series and review of 204 cases. *Journal of Clinical Oncology: Official Journal of the American Society of Clinical Oncology*, *18*(12), 2493–2499. https://doi.org/10.1200/JCO.2000.18.12.2493

Takeda, T., Tsubaki, M., Asano, R., Itoh, T., Imano, M., Satou, T., & Nishida, S. (2020). Dimethyl fumarate suppresses metastasis and growth of melanoma cells by inhibiting the nuclear translocation of NF-κB. *Journal of Dermatological Science*, *99*(3), 168–176. https://doi.org/10.1016/j.jdermsci.2020.07.004

Tan, Y. Z., Baldwin, P. R., Davis, J. H., Williamson, J. R., Potter, C. S., Carragher, B., & Lyumkis, D. (2017). Addressing preferred specimen orientation in single-particle cryo-EM through tilting. *Nature Methods*, *14*(8), Article 8. https://doi.org/10.1038/nmeth.4347

Tempera, I., & Lieberman, P. M. (2021). Oncogenic Viruses as Entropic Drivers of Cancer Evolution. *Frontiers in Virology*, *1*. https://www.frontiersin.org/articles/10.3389/fviro.2021.753366

Theiss, J. M., Günther, T., Alawi, M., Neumann, F., Tessmer, U., Fischer, N., & Grundhoff, A. (2015). A Comprehensive Analysis of Replicating Merkel Cell Polyomavirus Genomes Delineates the Viral Transcription Program and Suggests a Role for mcv-miR-M1 in Episomal Persistence. *PLOS Pathogens*, *11*(7), e1004974. https://doi.org/10.1371/journal.ppat.1004974

Thomsen, N. D., & Berger, J. M. (2009). Running in reverse: The structural basis for translocation polarity in hexameric helicases. *Cell*, *139*(3), 523–534. https://doi.org/10.1016/j.cell.2009.08.043

Toker, C. (1972). Trabecular carcinoma of the skin. *Archives of Dermatology*, *105*(1), 107–110.

Tolstov, Y. L., Pastrana, D. V., Feng, H., Becker, J. C., Jenkins, F. J., Moschos, S., Chang, Y., Buck, C. B., & Moore, P. S. (2009). Human Merkel cell polyomavirus infection II. MCV is a common human infection that can be detected by conformational capsid epitope immunoassays. *International Journal of Cancer*, *125*(6), 1250–1256. https://doi.org/10.1002/ijc.24509

Tomasini, R., Samir, A. A., Carrier, A., Isnardon, D., Cecchinelli, B., Soddu, S., Malissen, B., Dagorn, J.-C., Iovanna, J. L., & Dusetti, N. J. (2003). TP53INP1s and Homeodomain-interacting Protein Kinase-2 (HIPK2) Are Partners in Regulating p53 Activity *. *Journal of Biological Chemistry*, *278*(39), 37722–37729. https://doi.org/10.1074/jbc.M301979200

Topalian, S. L., Bhatia, S., Hollebecque, A., Awada, A., Boer, J. P. D., Kudchadkar, R. R., Goncalves, A., Delord, J.-P., Martens, U. M., Picazo, J. M. L., Oaknin, A., Spanos, W. C., Aljumaily, R., Sharfman, W. H., Rao, S., Soumaoro, I., Cao, A., Nghiem, P., & Schadendorf, D. (2017). Abstract CT074: Non-comparative, open-label, multiple cohort, phase 1/2 study to evaluate nivolumab (NIVO) in

patients with virus-associated tumors (CheckMate 358): Efficacy and safety in Merkel cell carcinoma (MCC). *Cancer Research*, *77*(13_Supplement), CT074. https://doi.org/10.1158/1538-7445.AM2017-CT074

Tourasse, N. J., & Li, W.-H. (2000). Selective Constraints, Amino Acid Composition, and the Rate of Protein Evolution. *Molecular Biology and Evolution*, *17*(4), 656–664. https://doi.org/10.1093/oxfordjournals.molbev.a026344

Tsang, S. H., Wang, R., Nakamaru-Ogiso, E., Knight, S. A. B., Buck, C. B., & You, J. (2016). The Oncogenic Small Tumor Antigen of Merkel Cell Polyomavirus Is an Iron-Sulfur Cluster Protein That Enhances Viral DNA Replication. *Journal of Virology*, *90*(3), 1544–1556. https://doi.org/10.1128/JVI.02121-15

Tsang, S. H., Wang, X., Li, J., Buck, C. B., & You, J. (2014). Host DNA Damage Response Factors Localize to Merkel Cell Polyomavirus DNA Replication Sites To Support Efficient Viral DNA Replication. *Journal of Virology*, *88*(6), 3285–3297. https://doi.org/10.1128/JVI.03656-13

Tsoi, P. S., Quan, M. D., Ferreon, J. C., & Ferreon, A. C. M. (2023). Aggregation of Disordered Proteins Associated with Neurodegeneration. *International Journal of Molecular Sciences*, *24*(4), Article 4. https://doi.org/10.3390/ijms24043380

Tsukada, H., Mochizuki, Y., & Fujiwara, S. (1966). THE NUCLEOIDS OF RAT LIVER CELL MICROBODIES: Fine Structure and Enzymes. *Journal of Cell Biology*, *28*(3), 449–460. https://doi.org/10.1083/jcb.28.3.449

Tsutsui, H., Jinno, Y., Shoda, K., Tomita, A., Matsuda, M., Yamashita, E., Katayama, H., Nakagawa, A., & Miyawaki, A. (2015). A diffraction-quality protein crystal processed as an autophagic cargo. *Molecular Cell*, *58*(1), 186–193. https://doi.org/10.1016/j.molcel.2015.02.007

Tunyasuvunakool, K., Adler, J., Wu, Z., Green, T., Zielinski, M., Žídek, A., Bridgland, A., Cowie, A., Meyer, C., Laydon, A., Velankar, S., Kleywegt, G. J., Bateman, A., Evans, R., Pritzel, A., Figurnov, M., Ronneberger, O., Bates, R., Kohl, S. A. A., … Hassabis, D. (2021). Highly accurate protein structure prediction for the human proteome. *Nature*, *596*(7873), Article 7873. https://doi.org/10.1038/s41586-021-03828-1

Tuteja, N., & Tuteja, R. (2004). Prokaryotic and eukaryotic DNA helicases. *European Journal of Biochemistry*, *271*(10), 1835–1848. https://doi.org/10.1111/j.1432-1033.2004.04093.x

Uversky, V. N. (2017). Intrinsically disordered proteins in overcrowded milieu: Membrane-less organelles, phase separation, and intrinsic disorder. *Current Opinion in Structural Biology*, *44*, 18–30. https://doi.org/10.1016/j.sbi.2016.10.015

Valle, M., Chen, X. S., Donate, L. E., Fanning, E., & Carazo, J. M. (2006). Structural Basis for the Cooperative Assembly of Large T Antigen on the Origin of Replication. *Journal of Molecular Biology*, *357*(4), 1295–1305. https://doi.org/10.1016/j.jmb.2006.01.021

Valle, M., Gruss, C., Halmer, L., Carazo, J. M., & Donate, L. E. (2000). Large T-Antigen Double Hexamers Imaged at the Simian Virus 40 Origin of Replication. *Molecular and Cellular Biology*, *20*(1), 34–41.

van der Meijden, E., Janssens, R. W. A., Lauber, C., Bouwes Bavinck, J. N., Gorbalenya, A. E., & Feltkamp, M. C. W. (2010). Discovery of a new human polyomavirus associated with trichodysplasia spinulosa in an immunocompromized patient. *PLoS Pathogens*, *6*(7), e1001024. https://doi.org/10.1371/journal.ppat.1001024

van Heel, M., & Schatz, M. (2005). Fourier shell correlation threshold criteria. *Journal of Structural Biology*, *151*(3), 250–262. https://doi.org/10.1016/j.jsb.2005.05.009

van Veenendaal, L. M., van Akkooi, A. C. J., Verhoef, C., Grünhagen, D. J., Klop, W. M. C., Valk, G. D., & Tesselaar, M. E. T. (2018). Merkel cell carcinoma: Clinical outcome and prognostic factors in 351 patients. *Journal of Surgical Oncology*, *117*(8), 1768–1775. https://doi.org/10.1002/jso.25090

Vanchiere, J. A., Abudayyeh, S., Copeland, C. M., Lu, L. B., Graham, D. Y., & Butel, J. S. (2009). Polyomavirus Shedding in the Stool of Healthy Adults. *Journal of Clinical Microbiology*, *47*(8), 2388–2391. https://doi.org/10.1128/JCM.02472-08

Varadi, M., Anyango, S., Deshpande, M., Nair, S., Natassia, C., Yordanova, G., Yuan, D., Stroe, O., Wood, G., Laydon, A., Žídek, A., Green, T., Tunyasuvunakool, K., Petersen, S., Jumper, J., Clancy, E., Green, R., Vora, A., Lutfi, M., … Velankar, S. (2022). AlphaFold Protein Structure Database: Massively expanding the structural coverage of protein-sequence space with high-accuracy models. *Nucleic Acids Research*, *50*(D1), D439–D444. https://doi.org/10.1093/nar/gkab1061

Veenhuis, M., Kiel, J. A. K. W., & Klei, I. J. V. der. (2003). Peroxisome assembly in yeast. *Microscopy Research and Technique*, *61*(2), 139–150. https://doi.org/10.1002/jemt.10323

Verhaegen, M. E., Harms, P. W., Goor, J. J. V., Arche, J., Patrick, M. T., Wilbert, D., Zabawa, H., Grachtchouk, M., Liu, C.-J., Hu, K., Kelly, M. C., Chen, P., Saunders, T. L., Weidinger, S., Syu, L.-J., Runge, J. S., Gudjonsson, J. E., Wong, S. Y., Brownell, I., … Dlugosz, A. A. (2022). Direct cellular reprogramming enables development of viral T antigen–driven Merkel cell carcinoma in mice. *The Journal of Clinical Investigation*, *132*(7). https://doi.org/10.1172/JCI152069

Verhaegen, M. E., Mangelberger, D., Harms, P. W., Eberl, M., Wilbert, D. M., Meireles, J., Bichakjian, C. K., Saunders, T. L., Wong, S. Y., & Dlugosz, A. A. (2017). Merkel Cell Polyomavirus Small T Antigen Initiates Merkel Cell Carcinoma-like Tumor Development in Mice. *Cancer Research*, *77*(12), 3151–3157. https://doi.org/10.1158/0008-5472.CAN-17-0035

Viscidi, R. P., Rollison, D. E., Sondak, V. K., Silver, B., Messina, J. L., Giuliano, A. R., Fulp, W., Ajidahun, A., & Rivanera, D. (2011). Age-Specific Seroprevalence of Merkel Cell Polyomavirus, BK Virus, and JC Virus ▾. *Clinical and Vaccine Immunology : CVI*, *18*(10), 1737–1743. https://doi.org/10.1128/CVI.05175-11

Vladimirova, O., Leo, A. D., Deng, Z., Wiedmer, A., Hayden, J., & Lieberman, P. M. (2021). Phase separation and DAXX redistribution contribute to LANA nuclear body and KSHV genome dynamics during latency and reactivation. *PLOS Pathogens*, *17*(1), e1009231. https://doi.org/10.1371/journal.ppat.1009231

Vonck, J., & van Bruggen, E. F. (1992). Architecture of peroxisomal alcohol oxidase crystals from the methylotrophic yeast Hansenula polymorpha as deduced by electron microscopy. *Journal of Bacteriology*, *174*(16), 5391–5399. https://doi.org/10.1128/jb.174.16.5391-5399.1992

Voog, E., Biron, P., Martin, J. P., & Blay, J. Y. (1999). Chemotherapy for patients with locally advanced or metastatic Merkel cell carcinoma. *Cancer*, *85*(12), 2589–2595. https://doi.org/10.1002/(sici)1097-0142(19990615)85:12<2589::aid-cncr15>3.0.co;2-f

Walker, J. E., Saraste, M., Runswick, M. J., & Gay, N. J. (1982). Distantly related sequences in the alpha- and beta-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. *The EMBO Journal*, *1*(8), 945–951.

Wang, B., Zhang, L., Dai, T., Qin, Z., Lu, H., Zhang, L., & Zhou, F. (2021). Liquid–liquid phase separation in human health and diseases. *Signal Transduction and Targeted Therapy*, *6*(1), Article 1. https://doi.org/10.1038/s41392-021-00678-1

Wang, X., Li, J., Schowalter, R. M., Jiao, J., Buck, C. B., & You, J. (2012). Bromodomain Protein Brd4 Plays a Key Role in Merkel Cell Polyomavirus DNA Replication. *PLOS Pathogens*, *8*(11), e1003021. https://doi.org/10.1371/journal.ppat.1003021

Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., Heer, F. T., de Beer, T. A. P., Rempfer, C., Bordoli, L., Lepore, R., & Schwede, T. (2018). SWISS-MODEL: Homology modelling of protein structures and complexes. *Nucleic Acids Research*, *46*(W1), W296–W303. https://doi.org/10.1093/nar/gky427

Wei, H., Dandey, V. P., Zhang, Z., Raczkowski, A., Rice, W. J., Carragher, B., & Potter, C. S. (2018). Optimizing "self-wicking" nanowire grids. *Journal of Structural Biology*, *202*(2), 170–174. https://doi.org/10.1016/j.jsb.2018.01.001

Weisshart, K., Taneja, P., Jenne, A., Herbig, U., Simmons, D. T., & Fanning, E. (1999). Two regions of simian virus 40 T antigen determine cooperativity of double-hexamer assembly on the viral origin of DNA replication and promote hexamer interactions during bidirectional origin DNA unwinding. *Journal of Virology*, *73*(3), 2201–2211. https://doi.org/10.1128/JVI.73.3.2201-2211.1999

Wendler, P., Ciniawsky, S., Kock, M., & Kube, S. (2012). Structure and function of the AAA+ nucleotide binding pocket. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research*, *1823*(1), 2–14. https://doi.org/10.1016/j.bbamcr.2011.06.014

Wendzicki, J. A., Moore, P. S., & Chang, Y. (2015). Large T and small T antigens of Merkel cell polyomavirus. *Current Opinion in Virology*, *11*, 38–43. https://doi.org/10.1016/j.coviro.2015.01.009

Wessel, R., Ramsperger, U., Stahl, H., & Knippers, R. (1992). The interaction of SV40 large T antigen with unspecific double-stranded DNA: An electron microscopic study. *Virology*, *189*(1), 293–303. https://doi.org/10.1016/0042-6822(92)90705-T

Wright, C. M., Seguin, S. P., Fewell, S. W., Zhang, H., Ishwad, C., Vats, A., Lingwood, C. A., Wipf, P., Fanning, E., Pipas, J. M., & Brodsky, J. L. (2009). Inhibition of Simian Virus 40 replication by targeting the molecular chaperone function and ATPase activity of T antigen. *Virus Research*, *141*(1), 71–80. https://doi.org/10.1016/j.virusres.2008.12.018

Wright, P. E., & Dyson, H. J. (2009). Linking folding and binding. *Current Opinion in Structural Biology*, *19*(1), 31–38. https://doi.org/10.1016/j.sbi.2008.12.003

Wu, J. H., Narayanan, D., Limmer, A. L., Simonette, R. A., Rady, P. L., & Tyring, S. K. (2019). Merkel Cell Polyomavirus Small T Antigen Induces DNA Damage Response. *Intervirology*, *62*(2), 96–100. https://doi.org/10.1159/000501419

Wu, X., Cai, Q., Shen, Z., Chen, X., Zeng, M., Du, S., & Zhang, M. (2019). RIM and RIM-BP Form Presynaptic Active-Zone-like Condensates via Phase Separation. *Molecular Cell*, *73*(5), 971-984.e5. https://doi.org/10.1016/j.molcel.2018.12.007

Xie, Q., Cheng, J., Mei, W., Yang, D., Zhang, P., & Zeng, C. (2023). Phase separation in cancer at a glance. *Journal of Translational Medicine*, *21*(1), 237. https://doi.org/10.1186/s12967-023-04082-x

Yang, R., Lee, E. E., Kim, J., Choi, J. H., Kolitz, E., Chen, Y., Crewe, C., Salisbury, N. J. H., Scherer, P. E., Cockerell, C., Smith, T. R., Rosen, L., Verlinden, L., Galloway, D. A., Buck, C. B., Feltkamp, M. C., Sullivan, C. S., & Wang, R. C. (2021). Characterization of ALTO-encoding circular RNAs expressed by Merkel cell polyomavirus and trichodysplasia spinulosa polyomavirus. *PLOS Pathogens*, *17*(5), e1009582. https://doi.org/10.1371/journal.ppat.1009582

Yardimci, H., Wang, X., Loveland, A. B., Tappin, I., Rudner, D. Z., Hurwitz, J., van Oijen, A. M., & Walter, J. C. (2012). Bypass of a protein barrier by a replicative DNA helicase. *Nature*, *492*(7428), 205–209. https://doi.org/10.1038/nature11730

Yasuda, S., Tsuchiya, H., Kaiho, A., Guo, Q., Ikeuchi, K., Endo, A., Arai, N., Ohtake, F., Murata, S., Inada, T., Baumeister, W., Fernández-Busnadiego, R., Tanaka, K., & Saeki, Y. (2020). Stress- and ubiquitylation-dependent phase separation of the proteasome. *Nature*, *578*(7794), Article 7794. https://doi.org/10.1038/s41586-020-1982-9

Yoon-Robarts, M., Blouin, A. G., Bleker, S., Kleinschmidt, J. A., Aggarwal, A. K., Escalante, C. R., & Linden, R. M. (2004). Residues within the B' motif are critical for DNA binding by the superfamily 3 helicase Rep40 of adeno-associated virus type 2. *The Journal of Biological Chemistry*, *279*(48), 50472–50481. https://doi.org/10.1074/jbc.M403900200

You, Z., Falco, M. D., Kamada, K., Pisani, F. M., & Masai, H. (2013). The Mini-Chromosome Maintenance (Mcm) Complexes Interact with DNA Polymerase α-Primase and Stimulate Its Ability to Synthesize RNA Primers. *PLOS ONE*, *8*(8), e72408. https://doi.org/10.1371/journal.pone.0072408

Youlden, D. R., Soyer, H. P., Youl, P. H., Fritschi, L., & Baade, P. D. (2014). Incidence and survival for Merkel cell carcinoma in Queensland, Australia, 1993-2010. *JAMA Dermatology*, *150*(8), 864–872. https://doi.org/10.1001/jamadermatol.2014.124

Yuan, Z., Georgescu, R., Bai, L., Zhang, D., Li, H., & O'Donnell, M. E. (2020). DNA unwinding mechanism of a eukaryotic replicative CMG helicase. *Nature Communications*, *11*(1), 688. https://doi.org/10.1038/s41467-020-14577-6

Zapatka, M., Borozan, I., Brewer, D. S., Iskar, M., Grundhoff, A., Alawi, M., Desai, N., Sültmann, H., Moch, H., Alawi, M., Cooper, C. S., Eils, R., Ferretti, V., & Lichter, P. (2020). The landscape of viral

associations in human cancers. *Nature Genetics*, *52*(3), Article 3. https://doi.org/10.1038/s41588-019-0558-9

Zhang, L., Zhu, C., Guo, Y., Wei, F., Lu, J., Qin, J., Banerjee, S., Wang, J., Shang, H., Verma, S. C., Yuan, Z., Robertson, E. S., & Cai, Q. (2014). Inhibition of KAP1 Enhances Hypoxia-Induced Kaposi's Sarcoma-Associated Herpesvirus Reactivation through RBP-Jκ. *Journal of Virology*, *88*(12), 6873–6884. https://doi.org/10.1128/JVI.00283-14

Zhao, J., Jia, Y., Shen, S., Kim, J., Wang, X., Lee, E., Brownell, I., Cho-Vega, J. H., Lewis, C., Homsi, J., Sharma, R. R., & Wang, R. C. (2020). Merkel Cell Polyomavirus Small T Antigen Activates Noncanonical NF-κB Signaling to Promote Tumorigenesis. *Molecular Cancer Research: MCR*, *18*(11), 1623–1637. https://doi.org/10.1158/1541-7786.MCR-20-0587

Zhao, X., Madden-Fuentes, R. J., Lou, B. X., Pipas, J. M., Gerhardt, J., Rigell, C. J., & Fanning, E. (2008). Ataxia Telangiectasia-Mutated Damage-Signaling Kinase- and Proteasome-Dependent Destruction of Mre11-Rad50-Nbs1 Subunits in Simian Virus 40-Infected Primate Cells. *Journal of Virology*, *82*(11), 5316–5328. https://doi.org/10.1128/JVI.02677-07

Zheng, S. Q., Palovcak, E., Armache, J.-P., Verba, K. A., Cheng, Y., & Agard, D. A. (2017). MotionCor2: Anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nature Methods*, *14*(4), Article 4. https://doi.org/10.1038/nmeth.4193

Zhou, B., Arnett, D. R., Yu, X., Brewster, A., Sowd, G. A., Xie, C. L., Vila, S., Gai, D., Fanning, E., & Chen, X. S. (2012). Structural Basis for the Interaction of a Hexameric Replicative Helicase with the Regulatory Subunit of Human DNA Polymerase α-Primase *. *Journal of Biological Chemistry*, *287*(32), 26854–26866. https://doi.org/10.1074/jbc.M112.363655

Zlotkin, T., Kaufmann, G., Jiang, Y., Lee, M. Y., Uitto, L., Syväoja, J., Dornreiter, I., Fanning, E., & Nethanel, T. (1996). DNA polymerase epsilon may be dispensable for SV40- but not cellular-DNA replication. *The EMBO Journal*, *15*(9), 2298–2305.

zur Hausen, A., Rennspiess, D., Winnepenninckx, V., Speel, E.-J., & Kurz, A. K. (2013). Early B-Cell Differentiation in Merkel Cell Carcinomas: Clues to Cellular Ancestry. *Cancer Research*, *73*(16), 4982–4987. https://doi.org/10.1158/0008-5472.CAN-13-0616

9. List of figures

10. List of tables

11. Supplementary material

11.1. Hazardous substances according to GHS

Table S 1: Hazardous substances according to GHS

Substance	Hazard statements	Precautionary statements	Hazard pictogram
Acetic acid	H226, H314, H402	P210, P233, P240, P241, P242, P243, P260, P264, P273, P280, P301+330+331, P303+361+353	
Acrylamide	H301, H312, H315, H317, H319, H332, H340, H350, H361, H372	P201, P280, P301+310, P308+360+313	
Ampicillin sodium salt	H315, H317, H319, H334, H335	P261, P280, P305+351+338, P342+311	
APS	H272, H302, H315, H317, H319, H334, H335	P280, P302+352, P304+340, P305+351+338, P342+311	
BSA	H335	P261	
CaCl$_2$	H319	P264, P280, P305+351+338, P337+313	
cOmplete™, EDTA-free Protease Inhibitor Cocktail	H314, H318	P260, P280, P301+330+331, P301+361+353, P304+340+310, P305+351+338+310	
DAPI	H315, H317, H335	P264, P280	
DOC	H314, H318	P260, P280, P303+361+353,P304+340+310, P305+351+338, P363	
DTT	H302, H315, H319, H335	P261, P302+352, P305+351+338, P501	
EDTA	H302, H315, H319, H335	P261, P264, P270, P271, P280, P301+330+331, P302+352, P311, P305+351+338, P332+313, P337+313, P362, P403+233	
Epon 812 (Glycid ether 100)	H302, H315, H319, H341, H361f	P280, P302+352, P305+351+338, P308+313, P337+313, P501	
Ethane	H220, H280	P210, P410+403	
EtOH	H225, H319	P210, P233, P241, P243, P280, P337+P313, P403+P235	
Gentamycin sulfate	H317, H334	P260, P302+352, P342+311	

HCl	H290, H301, H314, H330, H334, H335, H370, H372	P260, P280, P301+310, P303+361+353, P304+340, P305+351+338, P308+311, P342+311	
HEPES	H315, H319, H335	P261, P264, P270, P271, P280, P301+312, P302+352, P304+312, P304+340, P305+351+338, P312, P321, P322, P330, P332+313, P337+313, P362, P363, P403+233, P405, P501	
Isopropanol	H225, H319, H336	P210, P240, P305+351+338, P403+233	
Kanamycin sulfate	H360	P201, P202, P280, P308, P313	
Liquid nitrogen	H281	P202, P271+403, P282,	
MeOH	H225, H301+311+331, H370	P210, P233, P240, P241, P242, P243, P260, P264, P270, P271, P280, P301+310, P303+361+353, P304+340, P330, P361+364, P370+378, P403+235, P404, P501	
MgCl2	H302, H312, H332	P102, P202, P261, P64, P280	
MnCl$_2$	H301, H318, H373, H411	P273, P280, P301+310+330, P305+351+338+310, P314	
Na$_2$CO$_3$	H319	P264, P280, P305+315+338, P337+313	
Na$_3$VO$_4$	H302+312+332, H315, H319	P261, P280, P301+312, P302+352+312, P304+340+312, P305+351+338	
NaCl	H314	P260, P264, P280, P301+330+331, P303+361+353, P305+351+338, P310, P363, P405, P501	
NaF	H301, H315, H319, H335	P302+352, P305+351+338, P308+310	
Nonidet P-40	H302, H315, H410	P273, P280, P301+312+330, P302, P352	
NP40	H318, H412	P280, p273,p305+351+338, P310, P501	
Paraformaldehyde	H228, H302+332, H315, H317, H318, H335, H341, H350	P210, P280, P301+312, P304+340+312, P305+351+338, P308+313	
PEI	H302, H317, H319, H411	P273, P280, P305+351+338	

145

Pen/Strep	H302, H317, H334, H361d	P260, P280, P302+352, P309+311	
PMSF	H301, H314	P260, P280, P303+361+353, P304+340+310, P305+351+338	
Potassium acetate	H302, H315, H319, H335	P261, P264, P270, P271, P280	
Propane	H220	P210	
SDS	H302, H311, H315, H319	P280, P305+351+338, P361, P405, P501	
TEMED	H225, H314, H302, H332	P210, P280, P305+351+338, P310	
Tetracycline hydrochloride	H302, H315, H361d, H362	P260, P280, P301+312, P302+352, P308+313	
TRIS-HCl	H315, H319, H332, H335	P261, P270, P280, P302+352, P304+340+312, P305+351+338, P321, P332+313, P337+313, P403+233, P501	
Triton X-100	H315, H319	P264, P280, P302+352, P305+351+338, P332+313, P338+313, P362	
Uranyl acetate	H300, H330, H373, H411	P260, P264, P273, P284, P301+310	
X-Gal	H302, H312, H332	P261, P264, P271, P280, P313, P501	
X-tremeGENE™ HP	H225, H319	P210, P233, P280, P303+361, P353, P337+313, P370+378	
β-Mercaptoethanol	H227, H301+331, H310, H315, H317, H318, H373, H410	P261, P273, P280, P302+350, P305+351+338, P310, P501	

11.2. Supplementary figures

		1	2	3	4	5	6	7	8
Mcm3	1		30,38	22,75	23,34	23,90	20,74	6,09	6,87
Mcm7	2	598		23,79	25,50	24,00	21,94	7,27	8,66
Mcm2	3	788	711		23,01	19,54	22,83	6,52	6,37
Mcm5	4	680	602	726		22,80	22,48	7,52	9,18
Mcm6	5	694	684	844	694		19,10	8,08	8,21
Mcm4	6	818	747	764	724	805		7,38	6,98
SV40 LT	7	925	854	990	824	876	929		34,31
MCPyV LT	8	935	844	1000	831	906	960	582	

Figure S 1: Sequence alignment of human MCM 2-7 with SV40 and MCPyV LT. Sequence identity is depicted in % (blue). The number of differing alignment positions is shown in red. The alignment was carried out with CLC.
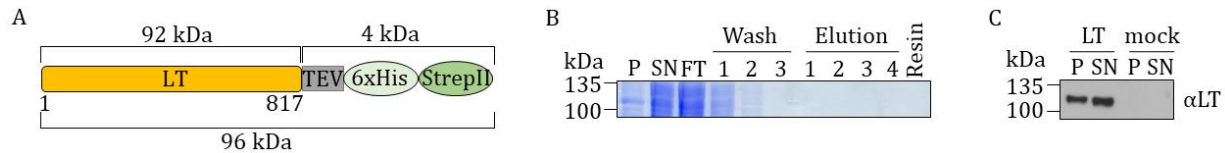

Figure S 2: Expression and purification of LT-His-StrepII. (A) LT construct which could not be purified. LT (yellow), C-terminally coupled to a TEV-cleavage site (gray), 6xHis (light green), and StrepII (green) tag, with amino acid numbers, depicted below the construct. The molecular weight of LT, with and without tags, is indicated in kDa. (B) SDS-PAGE of the LT-His-StrepII gravity flow StrepTactin affinity chromatography. 2.5×10^7 Sf9 insect cells were transduced with 0.5% (v/v) P2 baculovirus supernatant and harvested 4 d.p.t.. The cells were lysed with LB3 (Table 5), sonified, centrifuged, and the supernatant (SN) was incubated with 50 µl StrepTactin. LT-bound resin was washed 3x with 1 ml EB4 (Table 5) (Wash 1-3) and eluted with 4x 1 CV EB4, supplemented with 50 mM biotin (Elution 1-4). Similar volumes of the pellet (P), SN, flowthrough (FT) fractions, and resin were applied to an 8% gel and analyzed by Coomassie staining. (C) WB of LT-His-StrepII expression. 5×10^5 Sf9 cells, non-transfected (mock) or transfected with LT-His-StrepII bacmid (LT), were harvested 5 d.p.t. and lysed with 100 µl RIPA buffer (Table 5). The sample was centrifuged, and the pellet (P) was resuspended in 100 µl RIPA buffer. Similar volumes of P and the supernatant (SN) were analyzed by using the LT-antibody Cm2B4 (αLT).
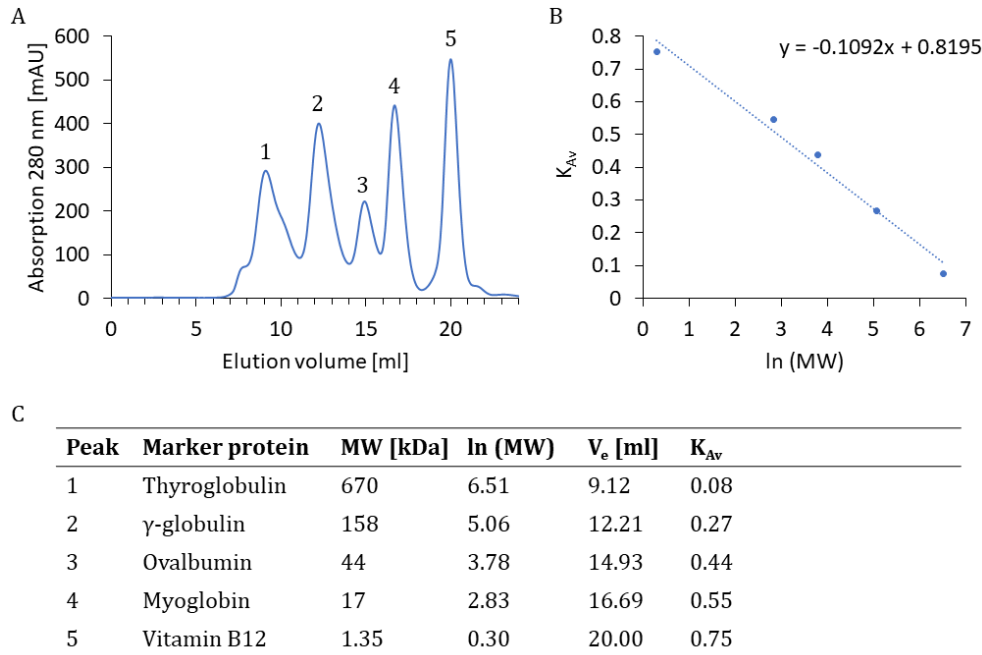
Figure S 3: Calibration of the size exclusion chromatography column HiLoad Superdex 200 10/300 GL. (A) Chromatography was carried out with EB4 (Table 5) at a flow rate of 0.25 ml/min with 60 µl gel filtration standard, supplemented with EB4 to a final volume of 500 µl. The absorption peaks are labeled with numbers corresponding to the elution volume of the proteins, outlined in (C). (B) Graphical illustration of (A), thereby plotting the natural logarithm of the molecular weight (ln (MW)) against the distribution coefficient (K_{av}), determined with the equation $(V_e – V_0)/(V_c\text{-}V_0)$. V_e: Elution volume; V_0: Void volume, which was determined at 7.89 ml; V_c: Colum volume, which is 24 ml for the respective column. (C) Used marker proteins with their associated MW, ln (MW), V_e, which was determined via (A) and K_{Av}, which was determined via (B).
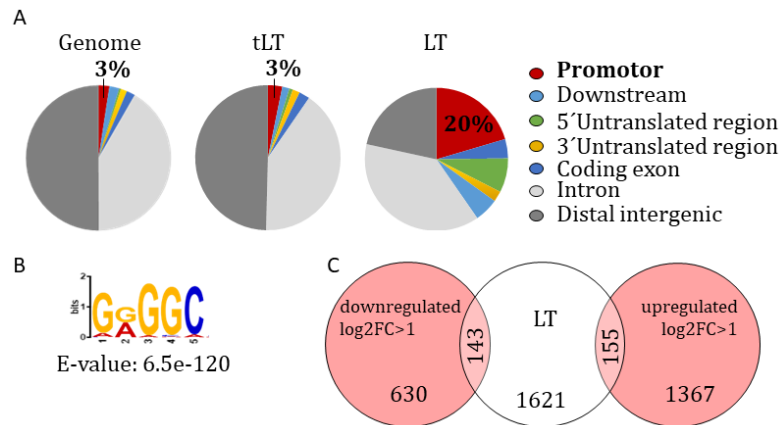

Figure S 4: LT-chromatin interaction. (A) LT-chromatin immunoprecipitation with subsequent DNA sequencing (ChIP-Seq). The MCPyV early region was transiently expressed in HEK293 cells, and LT-bound chromatin was enriched by the LT-antibody Cm2B4. Left pie chart shows percentage of genome annotations of hg19, given on the right. Center pie chart depicts tLT binding peaks (125) and right pie chart shows LT binding peaks (1581). (B) Motif calling of the in (A) identified LT binding sites. The letter size (y-axis) correlates with the statistical frequency at motif position 1-5 (x-axis). (C) Differential expression of LT-bound genes. Neonatal dermal fibroblasts were transduced with LT-expressing lentiviral supernatant and analyzed via RNA sequencing 8 d.p.t.. Venn diagram of all genes bound by LT (1621) and their down- (143) and upregulation (155) by log2FC>1. Additionally, down- (630) and upregulated (1367) genes versus vector control, independent of direct LT binding, are depicted.
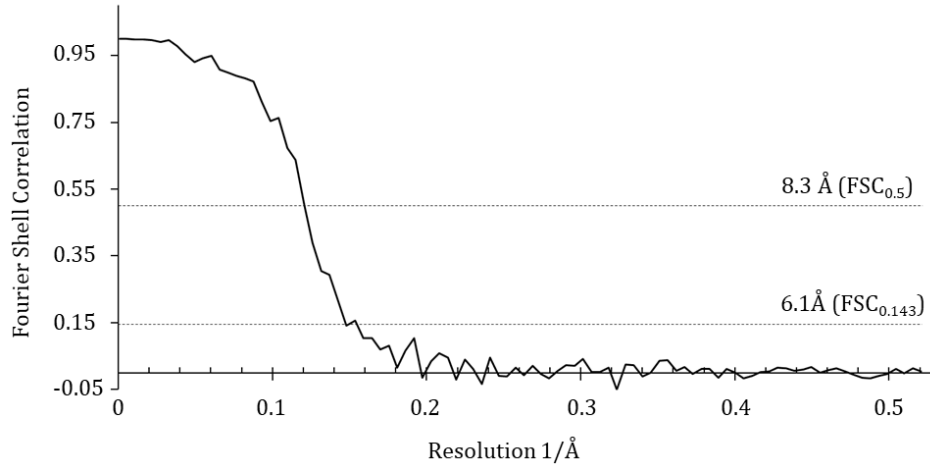
148

Figure S 5: FSC curve of the final cryoEM SPA LT model. The FSC curve was determined of two independently refined half maps and depicts the overall resolution as indicated by the cut-off at 0.5 and 0.143.

11.3. Supplementary tables

Table S 2: Overview of protein constructs screened for *in cellulo* crystallization. Expression was performed in (A) insect cells and (B) mammalian COS7 and HEK293 cells. Numbers represent the amino acid stop position (behind tLT/tsT) or starts (in front of tLT/tsT). Colon represents the co-expression of two proteins, a hyphen shows a protein fusion. Proteins with a peroxisomal target sequence (PTS) or an ER target sequence (ERTS) are sorted into the peroxisome or the ER, respectively. Crystal formation is coded with a green tick and crystal absence with a red cross.

A

Construct	Crystal formation	Molecular weight [kDa]
PAK4cat	✖	36
iBOX-PAK4cat	✔	40
sT-iBOX-Pak4cat	✖	60
tLT244-iBOX-PAK4cat	✖	67
GFP-iBox-PAK4cat	✖	67
GFP-Inka1:PAK4cat	✔	59
sT-Inka1:PAK4cat	✖	54
Inka1-DnaJ:PAK4cat	✖	41
PS:PAK4cat	✖	2
tLT244-PS:PAK4cat	✖	29
sT	✖	22
tLT265	✖	33
sT-PTS/ERTS	✖	22/25
122tsT-PTS/ERTS	✖	8/11
tLT328-PTS/ERTS	✖	36/39
tLT430-PTS/ERTS	✖	47/50
tLT544-PTS/ERTS	✖	61/64

B

Construct	Crystal formation	Molecular weight [kDa]
PAK4cat	✔	36
Inka1	✖	33
Inka1:PAK4cat	✔	
GFP-iBox-PAK4cat	✖	67
sT-Inka1:PAK4cat	✔	54
tLT244-Inka1:PAK4cat	✔	58
Inka1-122tsT:PAK4cat	✔	41
122tsT-Inka1:PAK4cat	✔	41
Inka1:PAK4cat-122tsT	✖	44
Inka1:122tsT-PAK4cat	✖	44
sT-PS:PAK4cat	✔	24

Table S 3: Buffers used for *in cellulo* crystal isolation.

Name	Composition
Hypo buffer	10 mM HEPES 10 mM KCl 5 mM MgCl$_2$ pH 7.5
Lysis buffer	50 mM TRIS-HCl 150 mM NaCl 5 mM EDTA 1% Triton X-100 1 g/l SDS 5 g/l Sodium deoxycholate pH 7.4

Table S4: Composition of the custom-made buffer screen used for the thermal stability assay. Column 12 was used for controls 1) 1M SPG Buffer contains: 0.125 M Succinic Acid, 0.5 M NaH2PO4, 0.375M Glycine adjusted to the appropriate pH with 10 M NaOH; 2) 1 M MMT Buffer contains: 0.2 M DL-Malic Acid, 0.4 M MES Monohydrate, 0.4 M Tris adjusted to the appropriate pH with either 10 M HCl or 10 M NaOH; 3) 1 M PCB Buffer contains: 0.4 M Sodium Propionate, 0.2 M Sodium Cacodylate trihydrate, 0.4 M Bis-tris Propane adjusted to the appropriate pH with 10 M HCL; 4) MES (2-(N-morpholino)ethanesulfonic acid) monohydrate 5) Bis(2-hydroxyethyl)amino-tris(hydroxymethyl)methane 6) N-(2-Acetamido)iminodiacetic acid 7) 4-(2-Hydroxyethyl)piperazine-1-ethanesulfonic acid 8) 2-Amino-2-hydroxymethyl-propane-1,3-diol.

	1	2	3	4	5	6	7	8	9	10	11	12
A	1 M SPG[1] Buffer, pH 5.6	1M MMT[2] Buffer, pH 5.6	1M PCB[3] Buffer, pH 5.6	1M MES[4] Buffer, pH 5.5	1M Na Citrate Buffer, pH 5.5	1M BisTris[5] Buffer, pH 5.8	1M Na/K Phosphate Buffer, pH 5.8	1M ADA[6] Buffer, pH 6.0	1M HEPES[7] Buffer, pH 6.8	1M Imidazole Buffer, pH 6.2	1 M Tris[8] Buffer, pH 7.5	
B	1 M SPG[1] Buffer, pH 6.0	1M MMT[2] Buffer, pH 6.0	1M PCB[3] Buffer, pH 5.0	1M MES[4] Buffer, pH 5.7	1M Na Citrate Buffer, pH 5.75	1M BisTris[5] Buffer, pH 6.0	1M Na/K Phosphate Buffer, pH 6.1	1M ADA[6] Buffer, pH 6.2	1M HEPES[7] Buffer, pH 7.0	1M Imidazole Buffer, pH 6.45	1 M Tris[8] Buffer, pH 7.7	
C	1 M SPG[1] Buffer, pH 6.4	1M MMT[2] Buffer, pH 6.4	1M PCB[3] Buffer, pH 6.4	1M MES[4] Buffer, pH 5.9	1M Na Citrate Buffer, pH 6.0	1M BisTris[5] Buffer, pH 6.2	1M Na/K Phosphate Buffer, pH 6.4	1M ADA[6] Buffer, pH 6.4	1M HEPES[7] Buffer, pH 7.2	1M Imidazole Buffer, pH 6.7	1 M Tris[8] Buffer, pH 7.9	
D	1 M SPG[1] Buffer, pH 6.8	1M MMT[2] Buffer, pH 6.8	1M PCB[3] Buffer, pH 6.8	1M MES[4] Buffer, pH 6.1	1M Na Citrate Buffer, pH 6.25	1M BisTris[5] Buffer, pH 6.4	1M Na/K Phosphate Buffer, pH 6.7	1M ADA[6] Buffer, pH 6.6	1M HEPES[7] Buffer, pH 7.4	1M Imidazole Buffer, pH 6.95	1 M Tris[8] Buffer, pH 8.1	
E	1 M SPG[1] Buffer, pH 7.2	1M MMT[2] Buffer, pH 7.2	1M PCB[3] Buffer, pH 7.2	1M MES[4] Buffer, pH 6.3	1M Na Citrate Buffer, pH 6.5	1M BisTris[5] Buffer, pH 6.6	1M Na/K Phosphate Buffer, pH 7.0	1M ADA[6] Buffer, pH 6.8	1M HEPES[7] Buffer, pH 7.6	1M Imidazole Buffer, pH 7.2	1 M Tris[8] Buffer, pH 8.3	
F	1 M SPG[1] Buffer, pH 7.6	1M MMT[2] Buffer, pH 7.6	1M PCB[3] Buffer, pH 7.6	1M MES[4] Buffer, pH 6.5	1M Na Citrate Buffer, pH 6.75	1M BisTris[5] Buffer, pH 6.8	1M Na/K Phosphate Buffer, pH 7.3	1M ADA[6] Buffer, pH 7.0	1M HEPES[7] Buffer, pH 7.8	1M Imidazole Buffer, pH 7.45	1 M Tris[8] Buffer, pH 8.5	
G	1 M SPG[1] Buffer, pH 8.0	1M MMT[2] Buffer, pH 8.0	1M PCB[3] Buffer, pH 8.0	1M MES[4] Buffer, pH 6.7	1M Na Citrate Buffer, pH 7.0	1M BisTris[5] Buffer, pH 7.0	1M Na/K Phosphate Buffer, pH 7.6	1M ADA[6] Buffer, pH 7.2	1M HEPES[7] Buffer, pH 8.0	1M Imidazole Buffer, pH 7.7	1 M Tris[8] Buffer, pH 8.8	
H	1 M SPG[1] Buffer, pH 8.4	1M MMT[2] Buffer, pH 8.4	1M PCB[3] Buffer, pH 8.4	1M MES[4] Buffer, pH 6.9	1M Na Citrate Buffer, pH 7.25	1M BisTris[5] Buffer, pH 7.2	1M Na/K Phosphate Buffer, pH 7.9	1M ADA[6] Buffer, pH 7.4	1M HEPES[7] Buffer, pH 8.2	1M Imidazole Buffer, pH 7.95	1 M Tris[8] Buffer, pH 9.0	

Table S 5: Known structures of PyV T antigens.

PyV	Protein/ Domain	Interaction partner	Resolution	Method	PDB	Source
MCPyV	VP1		3.5 Å	X-ray diffraction	6ZLZ	(Bayer et al., 2020)
SV40	sT	PP2A Aα	3.1 Å	X-rax diffraction	2PF4	(Cho et al., 2007)
MPyV	DnaJ			NMR	1FAF	(Berjanskii et al., 2000)
SV40	LT DnaJ-LxCxE	pRb 379-577 + 645-772	3.2 Å	X-rax diffraction	1GH6	(H.-Y. Kim et al., 2001)
SV40	LT NLS	Importin-α 79-507	1.8 Å	X-rax diffraction	4RXH	(Bernardes et al., 2015)
SV40	LT OBD			NMR	2TBD	(Luo et al., 1996)
SV40	LT OBD		1.5 Å	X-rax diffraction	2FUF	(Meinke et al., 2006)
SV40	LT OBD	Viral ori	1.7 Å	X-rax diffraction	2ITL	(Bochkareva et al., 2006)
SV40	LT OBD	DNA fork	1.7 Å	X-rax diffraction	5D9I	(Meinke et al., 2011)
SV40	LT OBD	Viral ori M1 and M3	2.4 Å	X-rax diffraction	2NTC	(Meinke et al., 2007)
JCPyV	LT OBD		2.6 Å	X-rax diffraction	4LIF	(Meinke et al., 2014)
JCPyV	LT OBD		2.7 Å	X-rax diffraction	5CYB	(Meinke et al., 2016)
MPyV	LT OBD	Viral ori	3.8 Å	X-rax diffraction	4FB3	(C. Harrison et al., 2013)
MCPyV	LT OBD	Viral ori	2.9 Å	X-rax diffraction	3QFQ	(C. Harrison et al., 2011)
SV40	LT OBD	RPA		NMR	1Z1D	(Arunkumar et al., 2005)
SV40	LT OBD-helicase	Viral ori M1-2 + EP	2.8 Å	X-rax diffraction	4GDF	(Y. P. Chang et al., 2013)
SV40	LT helicase		2.8 Å	X-rax diffraction	1N25	(D. Li et al., 2003)
SV40	LT helicase	DNA polymerase α-primase 1-78	5 Å	X-rax diffraction	4E2I	(Zhou et al., 2012)

SV40	LT helicase	p53 DNA-binding domain 92-292	3.2 Å	X-rax diffraction	2H1L	(Lilyestrom et al., 2006)
SV40	LT helicase	ATP	1.9 Å	X-rax diffraction	1SVM	(Gai, Zhao, et al., 2004)
JCPyV	LT helicase	ATP-competitive inhibitor	2.0 Å	X-rax diffraction	5J47	(Bonafoux et al., 2016)

12. Acknowledgement

13. Eidesstattliche Versicherung

Hiermit erkläre ich an Eides statt, dass ich die vorliegende Dissertationsschrift selbst verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe.

Hamburg, den 03.07.2023

_____

Veronika Brinschwitz