SALMONELLA TYPHIMURIUM BASED GENOME-WIDE CRISPR/CAS9 KNOCKOUT SCREEN TO STUDY p16 DEGRADATION

by

Sebastian Dawo

B.Sc., The University of Applied Sciences Weihenstephan-Triesdorf, 2018

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

in

THE FACULTY OF GRADUATE AND POSTDOCTORAL STUDIES

(Medical Genetics)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

March 2022

The following individuals certify that they have read, and recommend to the Faculty of Graduate and Postdoctoral Studies for acceptance, a thesis entitled:

SALMONELLA TYPHIMURIUM BASED GENOME-WIDE CRISPR/CAS9 KNOCKOUT SCREEN TO STUDY p16 DEGRADATION

submitted by Sebastian Dawo in partial fulfillment of the requirements for the degree of Master of Science

in Medical Genetics

Examining Committee:

Prof. Josef Penninger, Medical Genetics, UBC
Supervisor

Prof. Philip Hieter, Medical Genetics, UBC
Supervisory Committee Member

Prof. Peter Stirling, Medical Genetics, UBC
Additional Examiner

Additional Supervisory Committee Members:

Prof. Christopher Overall, Centre for Blood Research, UBC
Supervisory Committee Member

Abstract

The state of a cell and cell function is ultimately characterized by the expression of specific proteins. The equilibrium of a "healthy" amount of protein is maintained by the balance between newly generated proteins via translation and the degradation. Proteins half-life in eukaryotic cells can span from minutes to several days and are tightly controlled by the selective degradation of proteins via the ubiquitin-proteasome system. Finding specific genes that are involved in the degradation of a protein of interest harbours many applications for translational medicine, but also enables a better understanding of basic cellular biology. To enable such screens, we have established a FACS-based genome-wide CRISPR/Cas9 knockout screening platform, based on protein delivery via the type three secretion system (T3SS) of an avirulent strain of *Salmonella* typhimurium. We discovered that the delivered proteins are dynamically degraded following transfer, therefore this protein delivery system offers a suitable and unique approach to study protein degradation. As a proof of concept, we studied the poorly understood degradation of the key tumor suppressor protein p16, which additionally is lysine-free, contradicting a longstanding concept that canonical ubiquitin is conjugated to a lysine residue. The screen yielded several promising candidate genes including genes of the Torsin family, the Sec62/63 complex, and

COPS6. Validation of the candidate genes uncovered *THAP1* as a potential general regulator of protein degradation. We anticipate this method will offer novel mechanistic insights into the degradation of intracellularly delivered proteins and might uncover new pathways of protein homeostasis.

Lay Summary

Proteins need to be recycled in every cell to maintain a healthy cellular environment and maintain homeostatic amounts of every protein. Most proteins are degraded through a very specific process, requiring it to be recognized, decorated with the marker molecule ubiquitin, and then broken apart by the proteasome. In my project, we developed a platform that allows us to find genes that are involved in the degradation of any protein we are interested in. As a proof of this principle, we studied the degradation of the tumor suppressor p16 and found the gene T*HAP1* as a potential general regulator of protein degradation. We anticipate this method will offer novel mechanistic insights into the degradation of intracellularly delivered proteins and might uncover new pathways involved in protein degradation.

Preface

Project conceptualization was carried out by Prof. Josef Penninger, Antoine Chabloz (PhD), Melanie de Almeida (MSc) and me. All experiments were carried out by me. Additional help for the genome-wide CRISPR/Cas9 screen was provided by Melanie de Almeida (MSc), Michaela Fellner, and the BioOptics core service at the Vienna Bio Center (VBC). Analysis of raw sequencing data was performed by the bioinformatics core service at the VBC and Melanie de Almeida (MSc). All other data analysis was performed by me. Additional scientific input and mentorship were provided by Prof. Josef Penninger, Prof. Phil Hieter, and Prof. Christopher Overall. All figures in the thesis were generated by me, if not attributed differently.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

AAA+	ATPases associated with a variety of cellular activities
AAVS1	adeno-associated virus integration site 1
ATP	adenosine triphosphate
ATR	ataxia telangiectasia and Rad3-related protein
BFP	blue fluorescent protein
BSA	bovine serum albumin
BZB	bortezomib
Cas9	CRISPR associated protein 9
CDK	cyclin dependent kinase
CM	chloramphenicol
CMV	*cytomegalovirus*
CRISPR	clustered regularly interspaced short palindrome repeats
DAPI	4′,6-diamidino-2-phenylindole
DMEM	Dulbecco's modified Eagle's medium
DNA	deoxyribonucleic acid
dox	doxycycline
E1	ubiquitin-activating enzyme
E2	ubiquitin-conjugating enzyme
E3	ubiquitin ligase
ECL	enhanced chemiluminescence
EDTA	Ethylenediaminetetraacetic acid
ERAD	Endoplasmic-reticulum-associated protein degradation
FACS	fluorescence-activated cell sorting
FAM111B	family with sequence similarity 111 member B
FBS	fetal bovine serum
FDR	false discovery rate
GAPDH	glyceraldehyde 3-phosphate dehydrogenase
gDNA	genomic DNA
GFP	green fluorescent protein
HCL	hydrochloride
HCV	hepatitis C virus
HECT	homologous to the E6AP carboxyl terminus
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
LB	lysogeny broth
LDH	lactate dehydrogenase
LFC	$\log_2$ fold change

MeOH	methanol
MOI	multiplicity of infection
NaCl	sodium chloride
NGS	next generation sequencing
OVA	ovalbumin
PBS	phosphate-buffered saline
PCR	polymerase chain reaction
PFA	paraformaldehyde
PGK	phosphoglycerate kinase
POLR3A	RNA polymerase III subunit A
PSM(B)	Proteasome subunit (beta type)
PVDF	polyvinylidene difluoride
Rb	retinoblastoma protein
RBR	RING between RING
REGγ	Proteasome Activator Subunit 3
RING	really interesting new gene
RNA	ribonucleic acid
RNAi	ribonucleic acid interference
RPMI	Roswell Park Memorial Institute
S. (Typhimurium)	*Salmonella* (Typhimurium)
sgRNA	single guide RNA
SKP2	S-phase kinase-associated protein 2
SPI	*Salmonella* pathogenicity island
SUMO	Small Ubiquitin-like Modifier
T3SS	type III secretion system
TBST	tris-buffered saline and Tween 20
THAP1	THAP domain-containing protein 1
Tris	tris(hydroxymethyl)aminomethane
U2SURP	U2 SnRNP associated SURP domain containing
Ub	ubiquitin
UBA1	ubiquitin-like modifier activating enzyme 1
UPS	ubiquitin-proteasome system

Acknowledgements

I would like to thank Josef Penninger for giving me the opportunity to start out in his new lab in Vancouver and for always supporting me in my project. I am grateful to have continued the project from Antoine Chabloz.

Additionally, I would like to thank all the members of the Penninger lab in Vancouver and Vienna that have made my Masters Thesis a very fun experience, albeit all the challenges we faced together. Specifically, I want to mention Ania Bogoslowski, Jun Wang, and Gustav Jonsson for your technical support and fruitful discussions. I am very grateful to Ivona Kozieradzki for always being there for any of our needs and offering great advice for life. Most importantly, I need to thank Melanie de Almeida for making this entire project possible and always taking the time to explain to me the ins and outs of CRISPR screening!

I am so happy that I came to Vancouver and all the friends I made. Thank you for making my time here a blast and something that I will always cherish and remember dearly: Sam, Veronika, Anna, Gabe, Caitlin, Julien, Jannis, Caroline, Leticia, Roula, Nayrouz, Omer, Laura, and Kevin.

Chapter 1: Introduction

1.1 Protein degradation by the ubiquitin-proteasome system

The state of a cell and cell function is ultimately characterized by the expression of specific proteins. To maintain protein homeostasis, the balance between newly generated proteins via translation and the degradation of proteins that are too abundant or have lost their function is precisely maintained[1]. Proteins half-life in eukaryotic cells can span from minutes to several days and are controlled by two major pathways: autophagy and the ubiquitin-proteasome system (UPS)[2,3]. Autophagy refers to several different pathways that ultimately converge to degrade proteins, mostly large unwanted structures such as organelles and protein aggregates, via lysosomes. While most autophagy processes are non-selective, the UPS is a very specific degradation pathway. More than 80% of proteins are turned over by the UPS and furthermore UPS degradation is essential to precisely control protein levels during development, differentiation, and many other cellular processes such as the cell cycle[4]. Therefore, it is not surprising that UPS dysregulating plays a key role in many diseases including cancer[5], neurodegenerative[6], or cardiac disease[7].

Proteins are degraded via the canonical UPS in a multistep process[4,8,9] (Figure 1):

(I) Adenosin triphosphate (ATP) dependent ubiquitin (Ub) activation by attaching to the ubiquitin-activating enzyme (E1).

(II) Transfer of Ub from E1 to a ubiquitin-conjugating enzyme (E2).

(III) Specific recruitment of a substrate and a ubiquitin bound E2 by a ubiquitin ligase (E3) and transfer of the ubiquitin to the substrate. This step may be repeated multiple times to form lysine 48-linked polyubiquitin chains.

(IV) Recognition of a (poly)ubiquitinated substrate by the 19S regulatory particle. Removal of ubiquitin from substrates by deubiquitylating enzymes.

(V) Entry of the unfolded substrate into the 20S catalytic core particle, catalyzed by the regulatory particle. Capping of a core particle by one or two regulatory particles leads to a complete 26S proteasome assembly and subsequent proteolysis of the substrate.

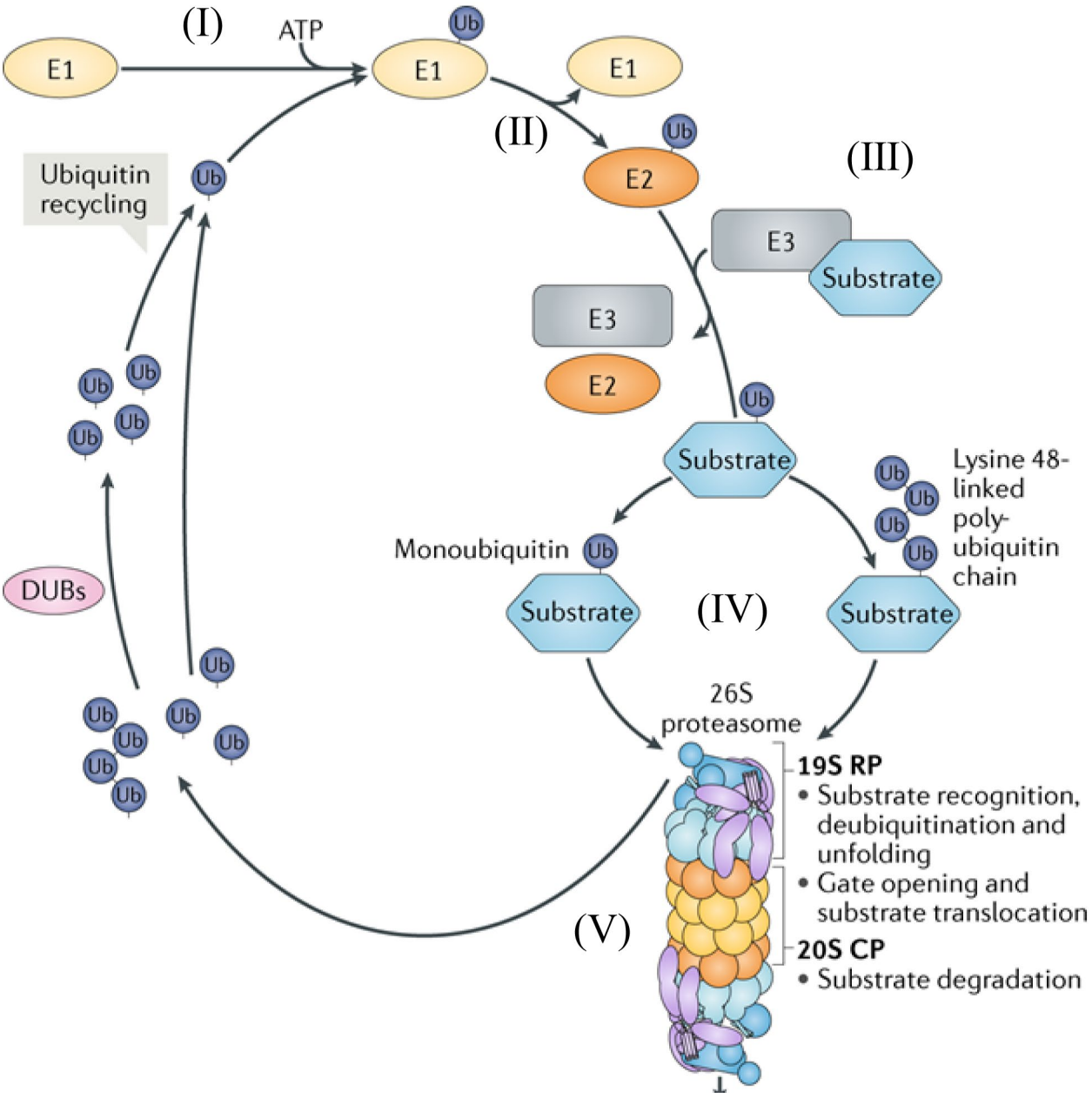

Figure 1. Schematic of the UPS. The five steps of protein degradation as described above are shown. Adapted from[8].

Currently, two ubiquitin E1s, 30-40 E2s, and more than 600 E3 ligases have been identified in the human genome[10–12] and thus E3 ligases confer the specificity of

the degradation system. This specificity makes E3 ligases an interesting target for cancer therapy[13], especially since almost 20% of all cancer driving genes affect UPS function[14]. E3 ligases are classified into three families: RING (really interesting new gene)-type E3 ligases, RBR (RING between RING)-type E3 ligases, and HECT (homologous to the E6AP carboxyl terminus) domain-containing E3 ligases[15,16]. The majority of E3 ligases belong to the RING family of E3s which mediate the direct transfer of Ub from the E2 to the substrate[15]. By contrast, HECT and RBR E3 ligases mediate ubiquitination by first transferring Ub to an active cystine site present in the E3 ligase and subsequent transfer of Ub to the substrate[16].

1.2 The tumor suppressor p16

The protein p16 (cyclin-dependent kinase (CDK) 4/6 inhibitor, *CDKN2A*) is an important tumor suppressor and loss-of-function of p16 has been observed in many types of cancer[17]. p16 exerts its tumor suppressor function through controlling the retinoblastoma protein (Rb) pathway (Figure 2), which primarily regulates the G1 to S transition in the cell cycle[18–21]. Briefly, p16's main role is to compete with D-cyclins to bind to CDK 4/6, thereby preventing the formation of the active cyclin D-CDK 4/6 complex. This in turn inhibits phosphorylation of Rb family members

and Rb remains associated with E2F transcription factors, consequently preventing transcription of E2F target genes that are required for G1 to S transition and thus leading to cell cycle arrest.

Figure 2. Schematic of the p16-CDK4/6-Rb-E2F pathway. p16 binds to CDK4/6, thereby blocking CDK4/6 activation by Cyclin D. Inactive CDK4/6 does not phosphorylate Rb, prohibiting the transcription of E2F target genes.

Interestingly, phosphorylation of Rb induces p16 expression through E2F1, forming a negative feedback loop[22,23]. However, upregulation of p16 has been found in a number of cancers and has been associated with increased patient survival in some tumors[24,25] and poor prognosis in others[26,27]. This dichotomy could be explained by non-canonical, cell cycle-independent, pathways that are regulated by p16, such as cell invasion, apoptosis, angiogenesis, tumor immune surveillance, and senescence[28,29] or simply an increase of p16 levels due to other downstream changes in the p16-Rb pathway, e.g. loss of Rb (positive feedback).

For example, p16 has been found to be overexpressed in cells at the invasive front of endometrial, basal cell, and colorectal carcinomas[30–32]. On the other hand,

increased cancer cell apoptosis and reduced angiogenesis have been reported in breast cancers treated with adenoviral transduction of p16 *in vivo*[33]. More recently, p16 has been found to play a role in tumor immune surveillance, with p16 increasing the anti-tumor activity of immune cells[29]. Additionally, p16 has a fundamental role in inducing senescence[28,34]. Specifically, p16 is important for oncogene induced premature senescence[28,34,35] and has been found to be overexpressed in benign and pre-malignant lesions that have not undergone malignant transformation[28]. Therefore, bypassing senescence via loss-of-function of the p16-Rb pathway could be an important step in the progression to a malignant tumor[28,29]. In summary, the regulation of p16 levels is important not only due to the loss of a major tumor suppressor pathway and therefore cancer initiation and neoplasia, but also the myriad of other tumor progression-related processes that are required for later malignant transformation.

Despite the discovery of p16 as a major tumor suppressor almost 30 years ago[36], and substantial research on this protein, many of the non-cell cycle related functions of p16 are still poorly understood. Another unanswered question is raised by the subcellular localization of p16. p16's most studied function is cell cycle control; this occurs in the nucleus. Interestingly however, p16 is also observed in

the cytoplasm in some cancers[37,38]. Here, p16 has been shown to form complexes with CDK4/6[37,39] and to interact with other proteins (e.g. actin, tubulin)[39], but no studies have addressed its actual function in the cytoplasm so far.

Interestingly, despite the prevalence of loss of p16 in many cancers[28], relatively little is known about its degradation and if such degradation events are differently regulated in the cytoplasm versus the nucleus.

1.2.1 p16 degradation – what is known

p16 has a relatively short half-life of a few hours[40–44]. Interestingly, p16 is lysine-free and conjugation of ubiquitin to an internal lysine residue is the first step in the degradation of most proteins through the canonical UPS (see chapter 1.1). Although N-terminal ubiquitination and ubiquitination on serine, threonine, and cysteine residues has been reported for some proteins in recent years, these non-canonical ubiquitin-dependent proteasomal degradation processes are still poorly understood[45]. A few studies have reported, at times contradicting, results on p16 degradation since the first published study in 2004[46]. The authors first reported that p16 is ubiquitinated and degraded by the proteasome in sparse HeLa cells but more stable in confluent cells[46]. The first study to suggest that p16 degradation follows a

ubiquitination independent pathway was published by Chen et al. in 2007[42]. Mass spectrometry results showed that endogenous cellular p16 is entirely acetylated at its N-terminus, therefore preventing it from being a suitable substrate for N-terminal ubiquitination and degradation[42]. Additionally, the authors demonstrated that p16 is degraded by 20S proteasomes but not 26S proteasomes *in vitro*, suggesting that ubiquitination and the 19S subunit are not required for p16 degradation. Instead, the REGγ (Proteasome Activator Subunit 3; also known as 11Sγ, PA28γ, *PSME3*) proteasomal activator interacts with p16 and knockdown of REGγ leads to increased steady-state levels and half-life of p16[41,42]. In contrast to the 19S subunit of the canonical UPS (see chapter 1.1), the REGγ regulatory subunit mediates degradation of substrates in a ubiquitin- and ATP-independent manner[47–49]. Furthermore, the protein p14 (ARF), expressed as an overlapping transcript at the same genetic locus as p16, regulates the proteasomal degradation of p16 by controlling SUMOylation (Small Ubiquitin-like Modifier) and nuclear export of REGγ[41]. Interestingly, REGγ is mostly localized to the nucleus but the SUMOylated form is translocated into the cytoplasm[50], potentially contributing to the degradation of cytoplasmic p16.

Of note, the hepatitis C virus (HCV) core protein activates REGγ expression, directly leading to the ubiquitin-independent proteasomal degradation of p16[43]. Additionally, the HCV core protein has also been shown to downregulate p16 expression by inducing promoter hypermethylation[51]. This adds another level of control over p16 by HCV and highlights the critical role of p16 downregulation in the development of hepatocellular carcinoma. In human gastric tumors and mouse xenografts, p16 is downregulated via COP9 (Constitutive photomorphogenesis 9) Signalosome Subunit 6 (COPS6)-mediated ubiquitin-independent degradation by the 20S core particle[52]. Furthermore, COPS6 and p16 interact with REGγ and knockdown of REGγ rescues COPS6-mediated downregulation of p16[52]. The COP9 signalosome itself is a paralogue of the 19S regulatory particle and functions as a regulator of protein degradation by directly promoting deneddylation, thereby controlling the activity of cullin-RING ligases[53].

Another study suggested that the poorly characterized protein FAM111B (family with sequence similarity 111 member B) promotes degradation of p16 in A549 lung cancer cells[54], but the underlying pathway was not further investigated. One study has found that the phosphorylation of p16 is important for its degradation: Al-Khalaf and colleagues observed UV-dependent polyubiquitination of p16 but

not of phosphorylated p16 in cells deficient of the protein kinase ATR (ataxia telangiectasia and Rad3-related protein)[44]. Furthermore, phosphorylated p16 was more stable than the non-phosphorylated form and p16 polyubiquitination was mediated by the E3 ligase SKP2 (S-phase kinase-associated protein 2), which is inhibited by ATR[44]. Lastly, a recent publication investigated and implicated p62-mediated autophagy as a pathway of p16 degradation upon cellular stress[55].

In summary, p16 degradation has been studied in different settings over the last two decades. Surprisingly however, our understanding of the pathways that the cell uses to regulate p16 levels is still incomplete and remains controversial, implicating different pathways, genes, and modulators. Furthermore, studying p16 degradation can give insight into "non-canonical" protein degradation pathways (e.g., N-terminal, serine, threonine, and cysteine ubiquitination) and thereby might lead to a better understanding of these fundamental cellular processes that have only been described recently. To study p16 degradation in a genome-wide context we established a genome-wide CRISPR/Cas9 knockout screen based on protein delivery via *Salmonella* serovar Typhimurium (*S.* Typhimurium), which will be introduced below.

1.3 *Salmonella*

Salmonella is a genus of the family *Enterobacteriaceae.* It is comprised of two species, *Salmonella bongori* and *Salmonella enterica* (*S. enterica*), which is additionally divided into six subspecies[56]. *Salmonella* are characterized as gram-negative, rod-shaped, generally motile bacteria with peritrichous flagella, ranging from 1-5 μm in size. *Salmonella* strains, specifically the serotypes Enteritidis and Typhimurium, are frequently responsible for bacteria related foodborne disease outbreaks in humans[57–59], instigating extensive research into this pathogen since its first description in 1884[60,61]. The non-typhoidal serotypes generally cause a self-limiting gastrointestinal illness, while other serotypes that are strictly adapted to humans, such as *S.* Typhi, can cause typhoid fever that requires immediate treatment[62]. *Salmonella* are facultative intracellular pathogens and can infect and invade a range of cells, but primarily proliferate and persist in macrophages[59]. An important tool that *Salmonella* uses for infecting and invading hosts is the type III secretion system (T3SS).

1.3.1 Type III Secretion System

The T3SS allows many gram-negative bacteria, like *Salmonella*, to deliver unfolded bacterial effector proteins directly into the cytoplasm of eukaryotic

cells[63]. *S.* Typhimurium has two discrete T3SSs encoded in so called *Salmonella* pathogenicity islands (SPI-1 and SPI-2)[64]. The T3SS from SPI-1 is active during initial contact with the host cell membrane and translocates effector proteins into the cytoplasm, an important step in host cell invasion[64,65]. After invading the target cell, the SPI-2 T3SS is responsible for ensuring bacterial survival and replication within the so called *Salmonella*-containing vacuoles, transferring proteins into the host endomembrane system and cytoplasm[64,66].

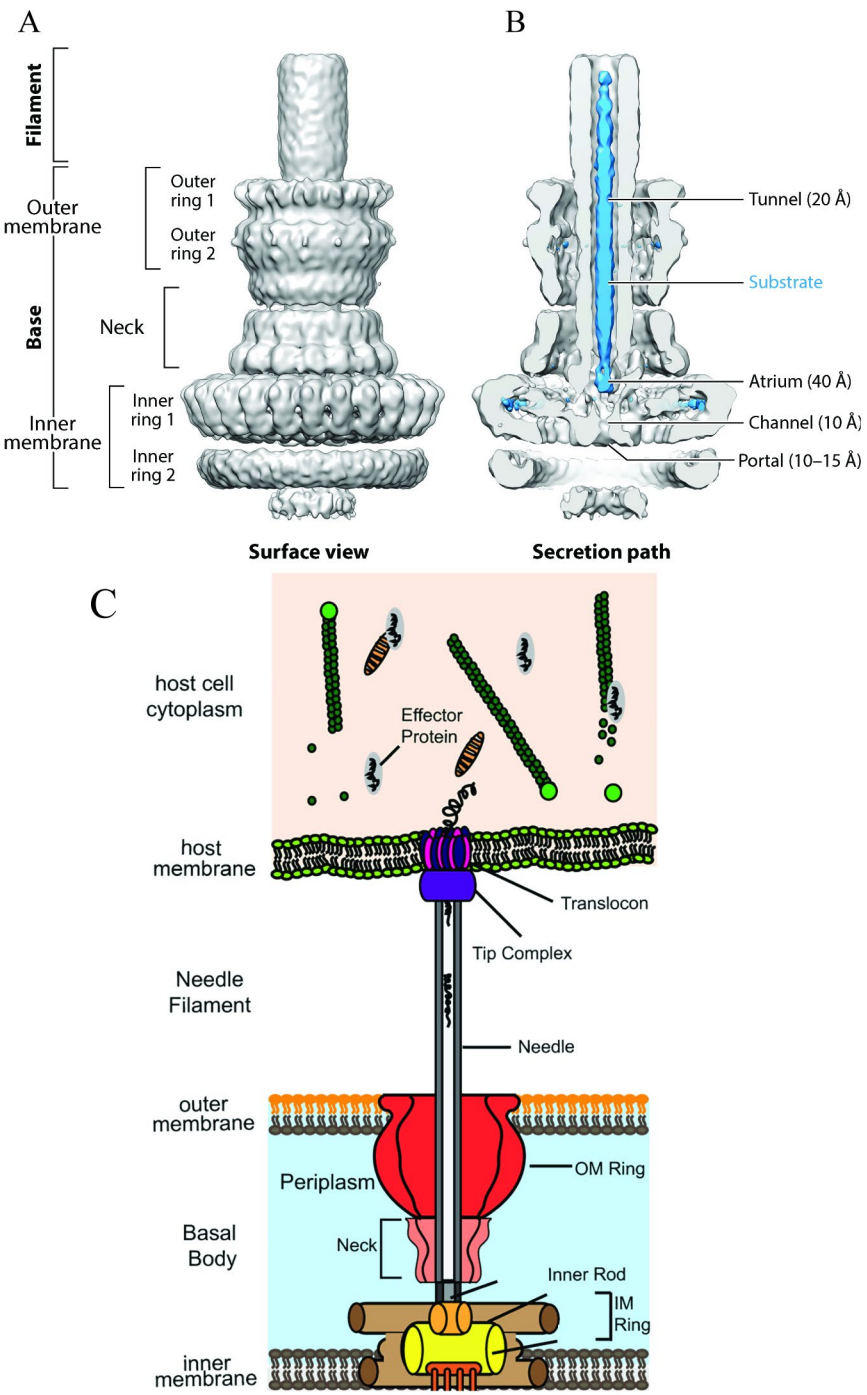

Figure 3. Structure of the needle complex from *S.* Typhimurium. 3-D reconstruction of a cryo-electron microscopy map of the T3SS of *S.* Typhimurium surface view (A) and half-sectioned needle complex with substrate (blue) trapped in the tunnel (B). Adapted from[67]. (C) Schematic of needle complex. Relevant structures are indicated. Adapted from[68].

The T3SS has two main components: (1) a highly conserved multi-protein nanomachine present on the cell membranes of bacteria, also called “injectisomes” or needle complex and (2) the effector proteins and transcriptional regulators that are translocated to manipulate host cell function[69,70]. The needle complex is composed of multiple structural subunits (Figure 3). An inner and an outer ring form the base that is rooted in the bacterial envelope and allows to translocate the effector proteins across the inner and outer bacterial membrane. Mounted on the circular base is a ~50 nm[71] needle or syringe-shaped filament that connects the bacteria to the external environment or the host cell membrane. The tip of the needle is capped and serves as a platform for the translocon that is secreted and directly creates a pore in the host cell membrane[68,72]. The selection and sequential delivery of effector proteins, recognized by encoded secretion signals[73], to the membrane base of the needle complex is orchestrated by a cytoplasmic multi-protein complex called the sorting platform[74]. Additionally, effector proteins are associated to chaperones that are necessary for targeting the proteins to the sorting platform[74] and to keep them in a partially unfolded state[75]. Since the diameter of the needle complex is only 10 Å at the narrowest point[76] (Figure 3), all substrates must be unfolded by a ATPase prior to translocation[73]. Interestingly, the unfolding

activity of the T3SS ATPase is relatively weak and therefore the mechanical stability of a protein in part predicts whether a protein can be translocated[77].

1.3.2 Non-virulent *S.* Typhimurium strain ASB2519 and pCASP-HilA expression plasmid

To harness the power of T3SS mediated cytoplasmatic protein delivery, we have recently engineered an improved version of the non-virulent *S.* Typhimurium strain SB2519[78]. Two additional genes (SopA, a E3 ligase required for invasion and SsaK, an indispensable SPI-2 effector) were knocked out in this strain to create a strain named ASB2519, which was used in this thesis[63]. This strain is currently approved for handling under BSL-1 conditions in Canada and Germany.

The pCASP-HilA plasmid is employed to express proteins of interest and boost their delivery[63] (Appendix A). Here, the protein of interest (such as p16), N-terminally fused to the first 120 amino acids of the secretion signal SptP (SptP120), is under the control of the SicA promoter, a *S.* Typhimurium SPI-1 effector protein, and is thus co-regulated with SPI-1 gene expression[63]. Furthermore, to boost the expression of SPI-1 genes in the ASB2519 strain, the pCASP-HilA plasmid contains the main regulator of SPI-1 genes, HilA, under the

control of an arabinose inducible pBAD promoter[63]. The inducible HilA and SptP120 enhance expression and secretion of the proteins of interest, respectively[63,79]. Therefore, ASB2519 in combination with the pCASP-HilA plasmid is a versatile and highly efficient protein delivery system, allowing the rapid cytosolic delivery of a wide range of proteins with high amounts of protein in all target cells[63,79]. Bacteria mediated protein delivery emerges as a powerful tool that can be used in many different applications to study basic biology (e.g., protein degradation), but also may be used in future translational efforts.

1.4 Pooled genome-wide CRISPR/Cas9 knockout screening

Less than a decade after researchers first described genetic perturbations using CRISPR (clustered regularly interspaced short palindrome repeats) in combination with the nuclease Cas9[80,81] it has evolved to a crucial tool in research and clinical sciences, highlighted by the Nobel Prize in Chemistry 2020 for Emmanuelle Charpentier and Jennifer A. Doudna.

Genome-scale loss-of-function screening is an important tool of functional genomics. Historically, this has been mostly done with RNA interference (RNAi), but the utility of RNAi has been hindered by incomplete protein depletion

(knockdown instead of knockout) and off-target effects[82–85]. Off-targets effects are also observed in Cas9 induced gene knockouts, but using an inducible Cas9 can mitigate the potential effects of long-term constitutional Cas9 expression[86,87]. CRISPR/Cas9 screens uncovered 3-4 times more essential genes compared to previous RNAi screening[88] and thus CRISPR/Cas9 screening has mostly replaced RNAi screening in recent years.

In a pooled CRISPR/Cas9 knockout screen a library of single guide RNAs (sgRNAs) is designed to target every gene of interest with multiple (typically 4-6) sgRNAs, as well as negative controls. These libraries are commercially available and therefore do not need to be designed, synthesized, and cloned by every individual researcher. Typically, lentiviral vectors are produced from the pooled sgRNA constructs and target cells transduced at a low multiplicity of infection (MOI) of > 0.3, ensuring that every cell only receives a single sgRNA. After selection of sgRNA positive cells, the cells are expanded and can be frozen for future screens. Cas9 can either be delivered simultaneously with the sgRNA library or a stable Cas9 cell line can be engineered prior to library transduction. Most pooled library screens use cell growth as a phenotype and readout in positive or negative selection[84], but fluorescence-activated cell sorting (FACS)-based screens

have, albeit more challenging, also shown good results in recent years[89]. The stable integration of the sgRNA expression cassette enables screen readout using next generation sequencing (NGS). Genomic DNA (gDNA) from target cell populations and control cells is extracted, a NGS library of the sgRNA-encoding region generated and subsequently sequenced. The raw sequencing data is commonly analyzed using bioinformatic analysis platforms such as MAGeCK-VISPR, PinAPL-Py, CRISPRCloud2, and CRISPRAnalyzeR[90]. These end-to-end packages generally allow a researcher with minimal knowledge of Pyhton or R to analyze screening data and further include a number of quality control tools[90].

1.5 Rationale, hypothesis, and objectives

Protein degradation is a major cellular pathway and a way for a cell to selectively clear excessive amounts of protein levels, protein aggregates, or abnormal proteins. It is fundamental to maintain the proper balance of proteins and therefore required for cellular function and cell health. Furthermore, knowing the mechanisms by which a specific protein is degraded not only expands our basic understanding of that protein's degradation, but also has direct implications for the treatment of many pathologies.

We hypothesized that cytoplasmic protein delivery via *S.* Typhimurium is a viable platform to rapidly study the degradation of any protein of interest using genome-wide CRISPR/Cas9 knockout screening. Thus, the objectives of my thesis were:

1. Find genes involved in the degradation of a sample protein of interest (p16) using *S.* Typhimurium as a protein delivery tool using genome-wide CRISPR/Cas9 screening.
2. Validate genes that are discovered in the screen to gain insight into (a) the degradation of p16 and (b) the degradation of lysine-free proteins.

Chapter 2: Methods

2.1 Cell culture

Human colon carcinoma (RKO; ATCC Cat# CRL-2577) cells were cultivated in Roswell Park Memorial Institute (RPMI) 1640 (Gibco). Human pancreas carcinoma (MIA PaCa-2; ATCC Cat# CRL-1420), human fibrosarcoma (HT-1080; ATCC Cat# CCL-121), and human embryonic kidney (Lenti-X 293T; Takara Bio Cat# 632180) cell lines were cultivated in Dulbecco's modified Eagle's medium (DMEM; Gibco). All growth media were supplemented with 10% fetal bovine serum (FBS; Gibco), 4 mM L-Glutamine (Sigma-Aldrich), and 25 mM HEPES (Gibco). Cells were routinely maintained at 37°C and 5% CO_2.

2.2 *Salmonella* Typhimurium infection assay

5 ml cultures were inoculated with the ASB2519 *Salmonella* strains and grown to saturation (250 rpm, 37°C) in LB medium (Sigma-Aldrich) supplemented with 25 μg/ml chloramphenicol (CM). Saturated bacterial cultures were diluted 1/10 in LB, supplemented with 25 μg/ml CM, 0.3 M NaCl, and 0.012% arabinose and grown for 2 h (250 rpm, 37°C). The target cell culture medium was changed 30 min prior to infection and supplemented with 25 μg/mL CM and 0.012% arabinose. Cells were then infected at a multiplicity of infection (MOI) of 100 for

1 h, after which cell culture medium (containing ASB2519) was removed and the eukaryotic cells washed once with 1x PBS. Cells were then further incubated for up to 4 h with 200 µg/ml gentamicin. When indicated, cells were treated with 50 nM bortezomib (BZB) 30 min before infection and throughout the entire experiment.

2.3 Western blotting

For sample preparation, cells were harvested with accutase (Gibco), pelleted at 4°C (2000 g, 2 min) and snap frozen in liquid nitrogen for storage at -80°C. Cells were thawed on ice and resuspended in 1x PBS, supplemented with 0.002% digitonin, protease inhibitor cocktail (Sigma-Aldrich, cOmplete), and, if necessary, with phosphatase inhibitor cocktail (Sigma-Aldrich, PhosSTOP). Resuspended cells were incubated for 5 min on ice and centrifuged at 4°C (16000 g, 25 min). Digitonin does not lyse bacterial membranes[63] and therefore the supernatant containing only eukaryotic cell lysate was harvested and the total protein content was determined via the Bradford protein assay (Bio-rad, Cat# 5000205). Samples were separated on 4-12% Bolt Bis-Tris gels (ThermoFisher Cat# NW04120BOX) and blotted on polyvinylidene difluoride (PVDF) membranes (ThermoFisherm, Cat# 88520) by wet transfer (4°C, 100 V, 60 min). PVDF membranes were

blocked for 45mins at RT in 5% milk or 5% PhosphoBlocker (Cellbiolabs, Cat# AKR-103) in TBST and probed with rabbit anti-HA-Tag (1/1600, Cell Signaling, C29F4), mouse anti-GAPDH (1/500, Santa Cruz Biotechnology, 0411), mouse anti-Rb (1:2000, Cell Signaling, 4H1), rabbit anti-phospho-Rb (Ser807/811) (1:1000, Cell Signaling, D20B12), rabbit anti-phospho-Rb (Ser780) (1:1000, Cell Signaling, 9307), diluted in universal antibody dilution buffer (Sigma-Aldrich, Cat# U3635) overnight at 4°C. Blots were washed three times in 1x PBS and incubated with secondary anti-mouse IgG (1/4000, Promega, W4021) or anti-rabbit IgG (1/4000, Sigma-Aldrich, NA9340) conjugated to horseradish peroxidase in 5% milk or 5% PhosphoBlocker for 1 h at RT. Blots were developed with Clarity ECL substrate (Bio-Rad, Cat# 1705060) and imaged using a ChemiDoc MP system (Bio-Rad).

2.4 Lactate dehydrogenase (LDH) assay

LDH release was measured with the LDH-Glo Cytotoxicity Assay (Promega, Cat# J2380) according to the manufacturer's instructions and measured on a Spark microplate reader (Tecan).

2.5 Immunofluorescence staining and microscopy

Cells were grown on coverslips in 24 well plates and infected at a MOI of 100 for 1 h. Cells were then washed once in 1x PBS and fixed with 4% paraformaldehyde (PFA) (ThermoFisher, Cat# 50-980-487) for 10 min at RT. Subsequently, cells were permeabilized and blocked for 1 h at RT in blocking buffer (5% FBS, 2% bovine serum albumin (BSA) (Sigma-Aldrich, Cat# A7906), 1% glycine (Sigma-Aldrich, Cat# 410225), 0.2% triton X (Roche, Cat# 11332481001) in 1x PBS. The specimens were then stained with primary rabbit anti-HA-Tag (1/1000, Cell Signaling, C29F4) overnight at 4°C in a humid chamber, followed by three washes in washing buffer (1% glycine, 0.2% Triton X in 1x PBS) and staining with secondary Alexa 647-anti-mouse IgG (1/500, Abcam, ab150115) in blocking buffer for 1 h at RT. Cells were washed three times in washing buffer and stained with DAPI (Fisher Scientific, Cat# D3571). Cover slips were mounted onto microscope slides (Thermo Fisher, Cat# 4951WS44370) and images acquired on a Leica TCS SP8 laser scanning confocal system. Image analysis and assembly was performed using ImageJ (v1.53c), the ScientiFig (v3.2) plugin for ImageJ, and Adobe Illustrator.

2.6 Flow cytometry

At the end of an infection experiment, the cells were washed once in 1x PBS, harvested with trypsin (Gibco), and transferred to a 96 well analysis plate (Sarstedt, Cat# 82.1583). The cells were then stained with fixable viability dye eFluor780 (Thermo Fisher, Cat# 65-0865-18) for 15 min on ice and washed once with FACS buffer (5% FBS in 1x PBS). Subsequently, the cells were fixed with 4% PFA for 10 min at RT. The samples were then blocked and permeabilized in blocking buffer for 30 min at RT and incubated with primary rabbit anti-HA-Tag (1/1600, Cell Signaling, C29F4) antibody at 4°C overnight. The cells were then stained with secondary PE-anti-rabbit IgG (1/500, BioLegend, 406421) in blocking buffer for 30 min at RT. Cells were then resuspended in FACS buffer and analyzed on a CytoFLEX LX (Beckman Coulter), LSR2 (BD Biosciences), or FACSCanto (BD Biosciences).

2.6.1 Flow cytometry for validation studies

For validations, the cells were processed as described in 2.6, but after fixation the prechilled cells were resuspended in ice-cold methanol (MeOH) and incubated on ice. The samples were then blocked, stained, and washed in FACS buffer.

2.7 FACS-based Genome-wide CRISPR/Cas9 knockout screen

The pooled genome-wide CRISPR/Cas9 knockout screen was performed using the genome-wide Vienna sgRNA library[91]. The inducible RKO library was already prepared as described previously[92]. 1.5x10^8 cells were thawed and expanded, maintaining a coverage of at least 500x per sgRNA. Cas9 expression was induced with 0.2 µg/ml doxycycline (dox) (Sigma-Aldrich, Cat# D9891) for the two timepoints (2.5 and 5 days of induction). At these two timepoints the standard *S.* Typhimurium infection and protein delivery (see 2.2) was performed and the cells were incubated for 4 h to allow for degradation of p16 and OVA. The cells were then pooled and split to batches of 5x10^7 cells per 50 ml Falcon conical centrifuge tube, stained with Zombie Aqua viability dye (1:1000, BioLegend, Cat# 423101), and fixed for 30 min at 4°C in 5 ml fixation buffer from the Foxp3/Transcription Factor Staining Buffer Set (Thermo Fisher, Cat# 00-5523-00). Subsequently, the cells were blocked in 5ml permeabilization buffer supplemented with 5% FBS for 45 min at 4°C and stained with rabbit anti-HA-Tag (1/1600, Cell Signaling, C29F4) antibody in 500 µl permeabilization buffer supplemented with 5% FBS overnight at 4°C. The cells were then stained with secondary PE-anti-rabbit IgG (1/500, BioLegend, 406421), washed 3x in FACS buffer, strained through a 40 µM mesh and sorted in sort buffer (1% FBS, 1 mM EDTA, 25 mM HEPES in 1x PBS)

at 8-10,000 events/second on FACS Aria III (BD Biosciences) and FACS Aria IIu (BD Biosciences) cell sorters. After exclusion of aggregates, dead, and Cas9-(GFP-) cells, two fractions were collected: a mid-fraction (HA^{mid}) of 50-60% of the main population centered around the median of the HA-PE signal and the 1-2% of cells with the highest HA-PE signal (HA^{high}). Sorted cells and unsorted control samples (harvested just before dox induction and on the respective timepoints as uninfected controls) were pelleted, snap-frozen and stored at -80°C until further processing.

2.7.1 Generation of NGS library and sequencing

NGS libraries were prepared in technical replicates for all HA^{mid} and HA^{high} fractions as previously described[91,92]. Genomic DNA (gDNA) was extracted by cell lysis (10 mM Tris-HCl, 150 mM NaCl, 10 mM EDTA, 0.1% SDS), proteinase K treatment (New England Biolabs, Cat# P8107S), and DNAse-free RNAse digest (Thermo Fisher, Cat# 10977035), followed by two rounds of phenol extraction. Isolated gDNA was frozen and thawed ten times to increase fragmentation prior to nested PCR amplification of the sgRNA cassette.

Barcoded NGS libraries for each sample were created using a two-step PCR protocol. In a first PCR reaction with 28 cycles, the sgRNA cassette was amplified from the isolated gDNA using 0.2 µl AmpliTaq Gold (Invitrogen, Cat# 4311820) in 50 µl reactions each containing 1 µg of gDNA. All reactions of a sample were pooled, and the DNA purified with in-house magnetic beads. The purified DNA was used in a second PCR reaction (7 cycles) to add barcodes and Illumina sequencing adaptors using 10 ng DNA template per 50 µl reaction. The final Illumina libraries were pooled and deep sequenced on a HiSeq2500 (Illumina) platform. Primers used for library amplification are listed in Appendix B.

2.7.2 Analysis of pooled genome-wide CRISPR/Cas9 knockout screen

Analysis of the pooled genome-wide CRISPR/Cas9 knockout screens was performed as previously described using the crispr-process-nf Nextflow workflow (https://github.com/ZuberLab/crispr-process-nf) developed by the Johannes Zuber laboratory[92]. Raw sequencing reads were aligned to a previously created index using Bowtie 2 (v2.3.0)[92]. Prior to read count median-normalization, guide RNAs with fewer than 50 counts in the control and sorted samples were excluded. Then, enrichment or depletion of sgRNAs in HA^{high} samples was calculated against control samples and average $\log_2$ fold changes (LFCs), ρ-values, and false

discovery rates (FDRs) calculated using MAGeCK (0.5.9)[93]. This analysis was performed by trained bioinformaticians, Florian Andersch and Melanie De Almeida, at the Vienna Bio Center.

2.8 Lentiviral plasmid design for Validation

sgRNA containing lentiviral plasmids were generated based on pLentiCRISPRv2 (Addgene #52961), obtained from the laboratory of Johannes Zuber, IMP, Vienna[92]. sgRNA sequences targeting the genes of interest were cloned into pLentiV2-U6-sgRNA-mPGK-eBFP using the Esp3L restriction site and confirmed by Sanger sequencing. Oligonucleotides for the sgRNA inserts are listed in Appendix C.

2.9 Lentivirus production and infection

70-90% confluent Lenti-X cells were cultured in RPMI 1640 and co-transfected with pCMVR8.74 (Addgene Cat#22036), pCMV-Eco (Cell Biolabs), and sgRNA containing lentiviral plasmid (pLentiV2.U6-improvedTRACER-PGK-eBFP) using Lipofectamine 3000 (Invitrogen) reagent according to the manufacture's instructions. RKO cells containing a doxycycline-inducible Cas9 were infected at 30-70% infection efficiency with 4 μg/ml polybrene (Sigma-Aldrich) and infection

levels were determined by flow cytometry based on blue fluorescent protein (BFP) expression 7 days post transduction.

2.10 Competitive proliferation assays

Gene editing was activated by inducing Cas9 expression with 0.2 μg/ml dox (Sigma-Aldrich, Cat# D9891) and confirmed by green fluorescent protein (GFP) expression. The percentage of BFP^+ ($sgRNA^+$) cells was monitored by flow cytometry over 9 days in 24 h intervals to determine effects of the inducible knockouts on proliferation and viability.

2.11 Statistical analysis

Statistical information regarding the number of technical or biological replicates are indicated in the figure legends. To organize data Microsoft Excel (v 16.0.14527.20270) was used. Screen analysis was performed with the R (v4.0.2) programming environment using RStudio (v1.3.1073 and v1.1.456)[92]. Quantitative immunoblotting was evaluated using ImageJ (v1.53c) and statistical analysis was performed using GraphPad Prism (v9.2.0).

Chapter 3: Results

3.1 Degradation of p16, OVA, and secretion tag in 3 cell lines

3.1.1 SptP120-HA is rapidly degraded

To find genes that are specific for the degradation of p16 and do not stem from the secretion tag (SptP120), the lysine-free HA tag, or the protein being expressed in bacteria, we chose to perform a separate screen on a control protein. The most straight forward choice to control for confounding components from the SptP120 and HA tags would be to use the fusion peptide SptP120-HA. Initial experiments revealed that SptP120-HA is delivered in much lower quantities to RKO cells than SptP120-p16-HA (Figure 4). Furthermore, while p16 is completely degraded over 4-5 h and the degradation was inhibited by bortezomib (BZB), SptP120-HA is degraded more rapidly and BZB does not rescue its degradation (Figure 4). This data shows that the delivered SptP120-p16-HA is degraded at comparable rates to native p16[40–44], while the secretion tag, SptP120-HA, on its own is degraded more rapidly.

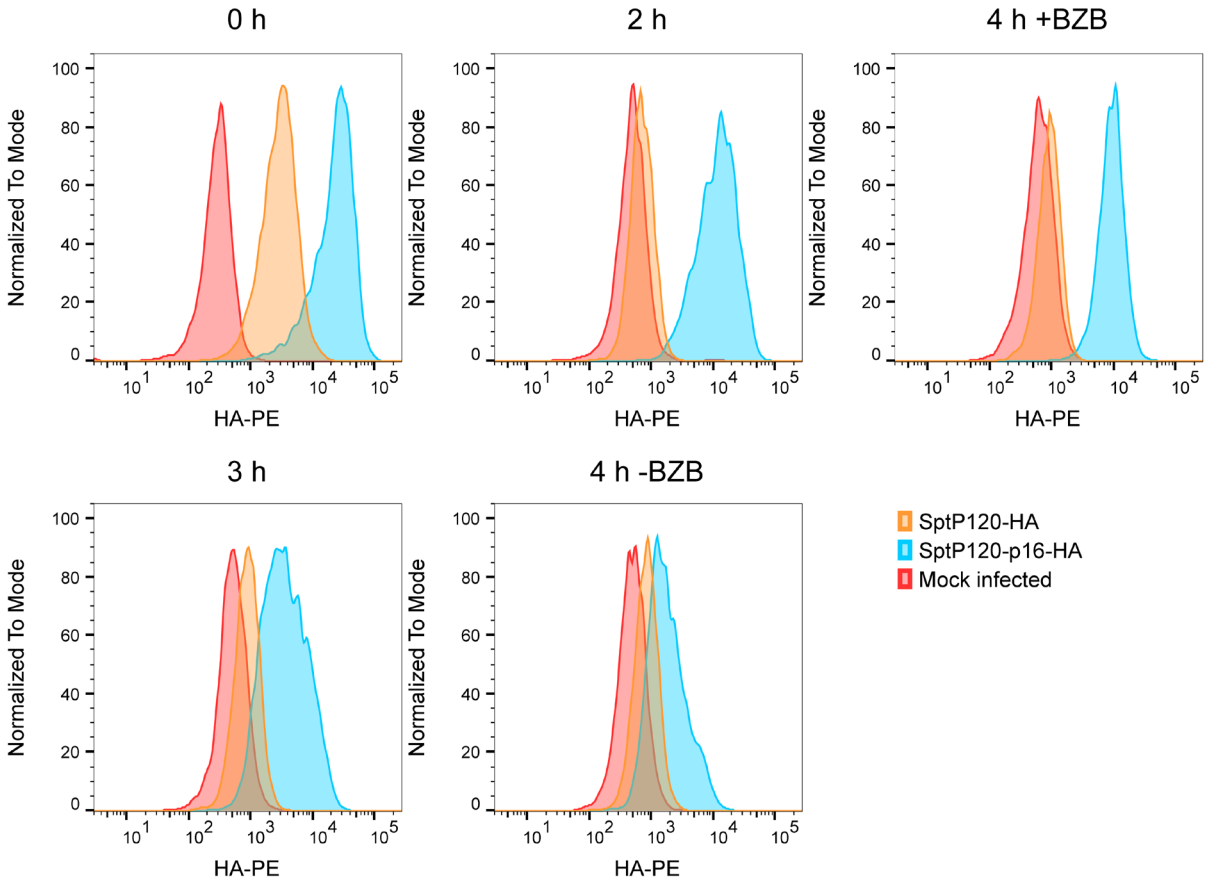

Figure 4. Fusion proteins delivered via S. Typhimurium are degraded by the proteasome in RKO cells. Flow cytometry analysis of SptP120-HA and SptP120-p16-HA transferred to RKO cells by 1 hour infection at a MOI of 100. Cells were incubated for the indicated times after infection in the presence of 50 nM of the proteasome inhibitor bortezomib (BZB) when indicated. *S.* Typhimurium ASB2519 with an empty pCASP-HilA vector served as a negative control (Mock infected). Data is representative of two independent experiments.

3.1.2 Ovalbumin as a degradation control

Screening for a better suited control revealed that Ovalbumin (OVA) is delivered and degraded at a comparable rate to p16 in three different cancer cell lines (RKO, HT1080, and MiaPaCa2) that were available for screening (Figure 5A). Details

regarding these three cell lines can be found in chapter 2.1. Since RKO cells had the strongest shift from mock infected to p16 and OVA protein levels as measured by flow cytometry (the primary read out in the screen) and p16 delivery had low cytotoxic effects on these cells (Figure 5B), they were chosen for the genome-wide CRISPR/Cas9 knockout screens.

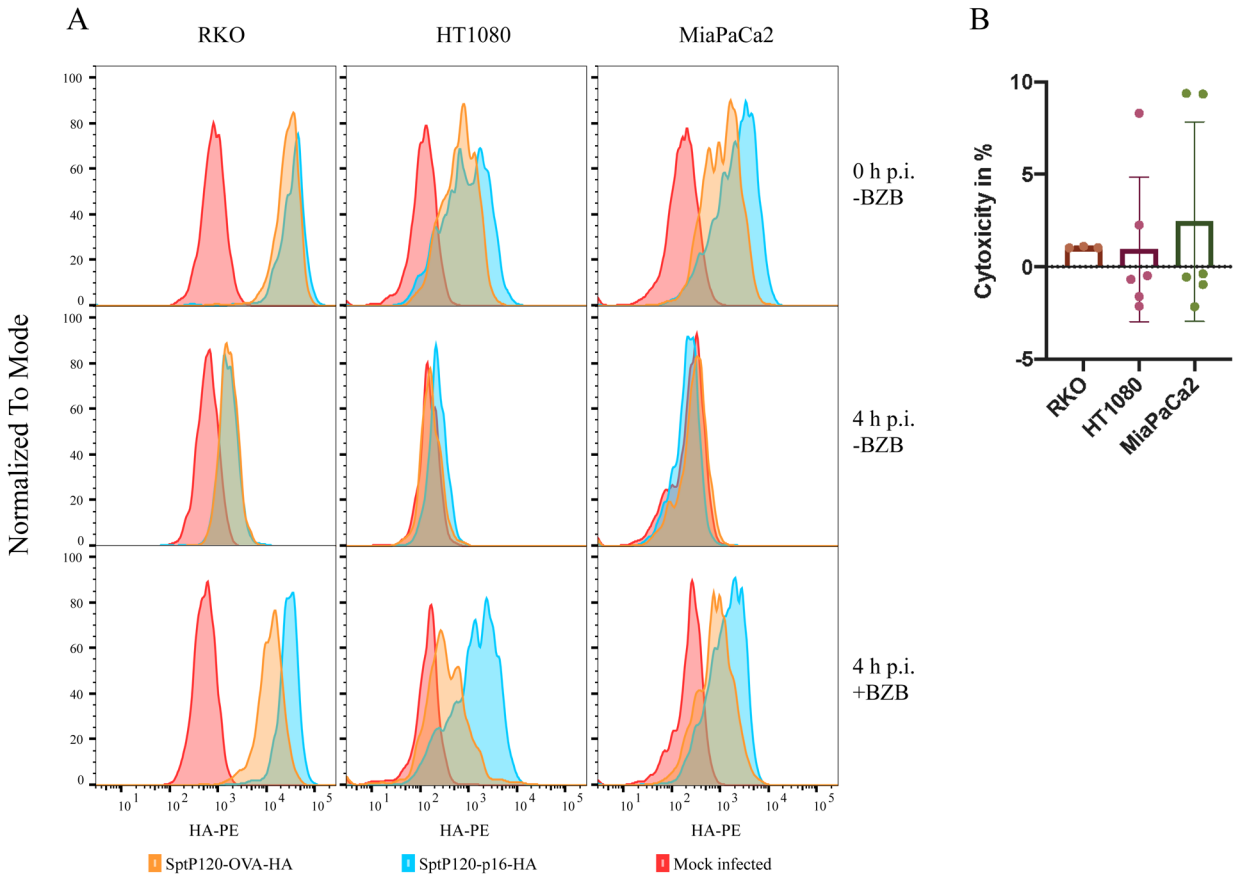

Figure 5. OVA degradation in RKO cells. (A) Flow cytometry analysis of SptP120-OVA-HA and SptP120-p16-HA transferred to RKO, HT1080, and MiaPaCa2 cells. Cells were infected at a MOI of 100 for 1 h and incubated for 0 and 4 h after infection with 50 nM of bortezomib (BZB) as indicated. *S.* Typhimurium ASB2519 with an

empty pCASP-HilA vector served as a negative control (Mock infected). Data is representative of four independent experiments. (B) Cytotoxicity of SptP120-p16-HA delivery via *S.* Typhimurium ASB2519 was assessed using lactate dehydrogenase release assays in the indicated cells infected at a MOI of 100 for 1 hour followed by incubation for 4 hours. Data are shown as mean ± s.d.. p.i., post infection.

3.2 Genome-wide CRISPR Cas-9 knockout screen

3.2.1 Layout of the CRISPR/Cas9 knockout screen

To find genes involved in the degradation of p16 we performed FACS-based genome-wide CRISPR/Cas9 knockout screens in RKO cells as laid out in Figure 6A. Briefly, cells were thawed and Cas9 expression induced for 2.5 and 5 days to capture essential genes and more stable proteins, respectively. At the two timepoints, SptP120-p16-HA and SptP120-OVA-HA were delivered to the knockout cell libraries for 1 h via *S.* Typhimurium and further incubated for 4 h after which the cells were fixed, the HA-tag immuno-stained, and cells sorted in two fractions: the Top 1% gate (HA^{high}) and the mid-fractions (HA^{mid}, of 50-60% of the main population centered around the HA-PE median) (Figure 6B). Genomic DNA from the sorted cells was extracted, the sgRNAs expression cassette deep sequenced in two technical replicates, and hits recovered using the MAGeCK platform.

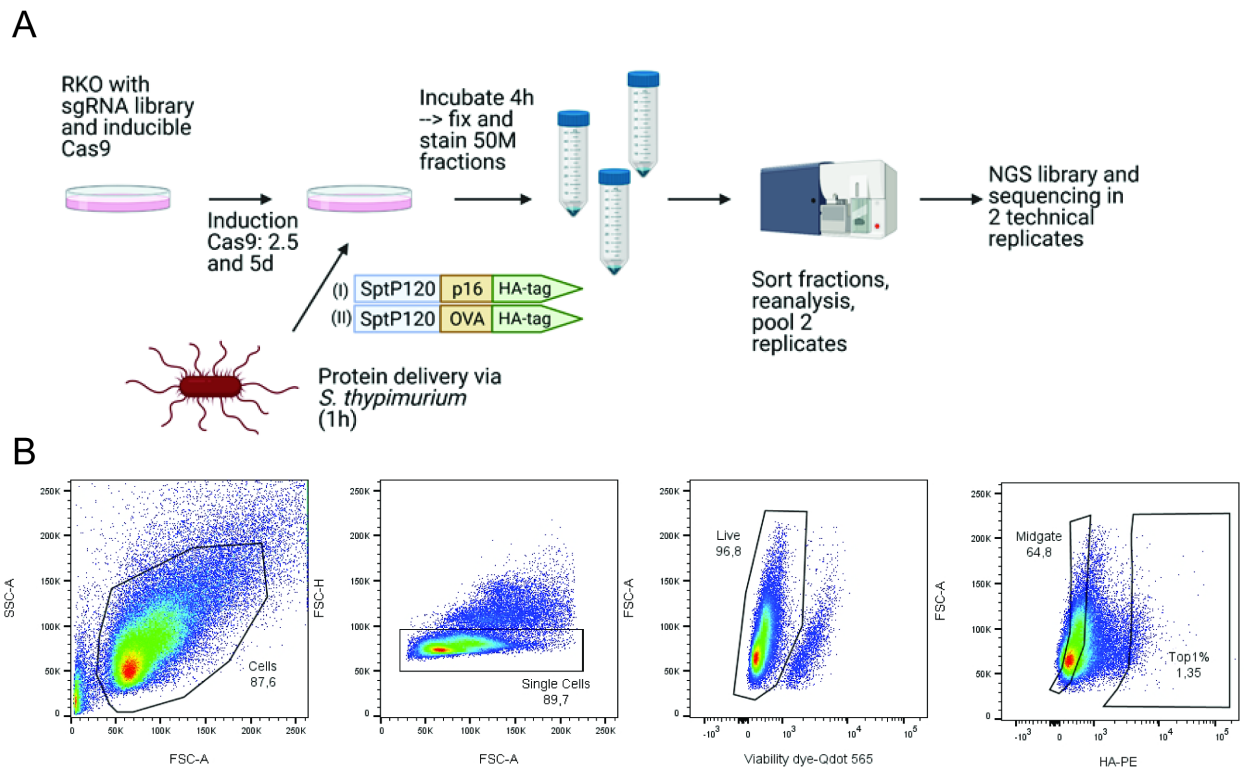

Figure 6. Layout of the genome-wide CRISPR/Cas9 knockout screen. (A) Schematic layout of the screen. (B) Sorting strategy used for the screens. First the general cell population is gated followed by gates excluding doublets and death cells. On this population HA^{high} and HA^{mid} are sorted.

3.2.2 Analysis of the genome-wide CRISPR/Cas9 knockout screen

The MAGeCK platform was used to normalize read counts based on the median and to calculate $\log_2$ fold changes (LFCs), ρ-values, and false discovery rates (FDRs)[93]. sgRNA enrichment was calculated between the sorted Top 1% gate, the mid-fractions, and the infected control samples harvested on the day of sorting. Since the representation of sgRNAs in many HA^{mid} samples was sub-par (many sgRNAs were not represented and variances between samples were different),

LFCs, ρ-values, and FDRs were more robust against unsorted control samples, and therefore HA^{mid} samples were not used for screen analysis. Furthermore, the technical replicates of HA^{high} samples were combined for statistical analysis. First, the obtained data was analyzed and general controls, such as genes of the 26S proteasome or other known essential genes that are involved in unrelated pathways, examined. Interestingly, more genes had stronger enrichment on day 5 than day 2.5 (Figure 7), indicating that most genes require more than 2.5 days to create a complete protein-level knockout.

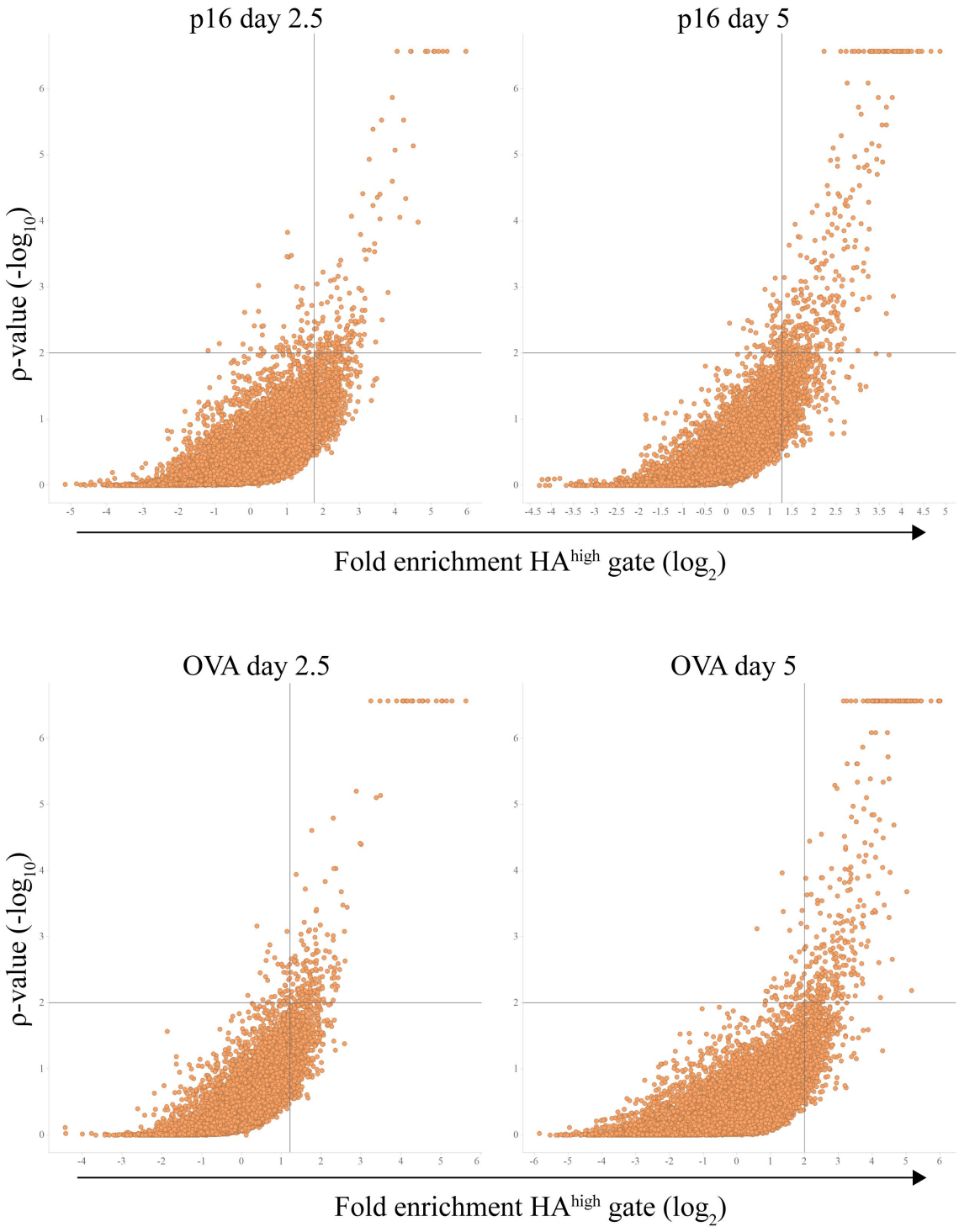

Figure 7. FACS-based CRISPR screens for genes involved in the degradation of SptP120-p16-HA and SptP120-OVA-HA. Gene-level enrichment of sgRNAs in HA^{high} cells as compared to unsorted control samples and one-sided MAGeCK ρ-values are shown. The lines indicate a ρ-value of 0.01 and the 95th percentile of enrichment.

All subunits of the 26S proteasome and other genes known to be involved in the UPS were analyzed in detail (Figure 8, Figure 9). This analysis revealed that the core particle is involved in degrading the delivered p16 and OVA, likely in a ubiquitin-dependent process as the E1 enzyme UBA1 and the 19S regulatory particle were also enriched in HA^{high} samples. Selected genes that have previously been described to be involved in p16 degradation and that have an established general role in protein degradation are also listed. For example, the alternative proteasome regulator REGγ (*PSME3*) is only moderately enriched in both p16 and OVA screens and therefore does not seem to be involved in their degradation in this screen. *COPS6* on the other hand, is only enriched in p16 and not OVA, while most other COP9 signalosome subunits are highly enriched in both, suggesting that *COPS6* has a specific role in p16 degradation.

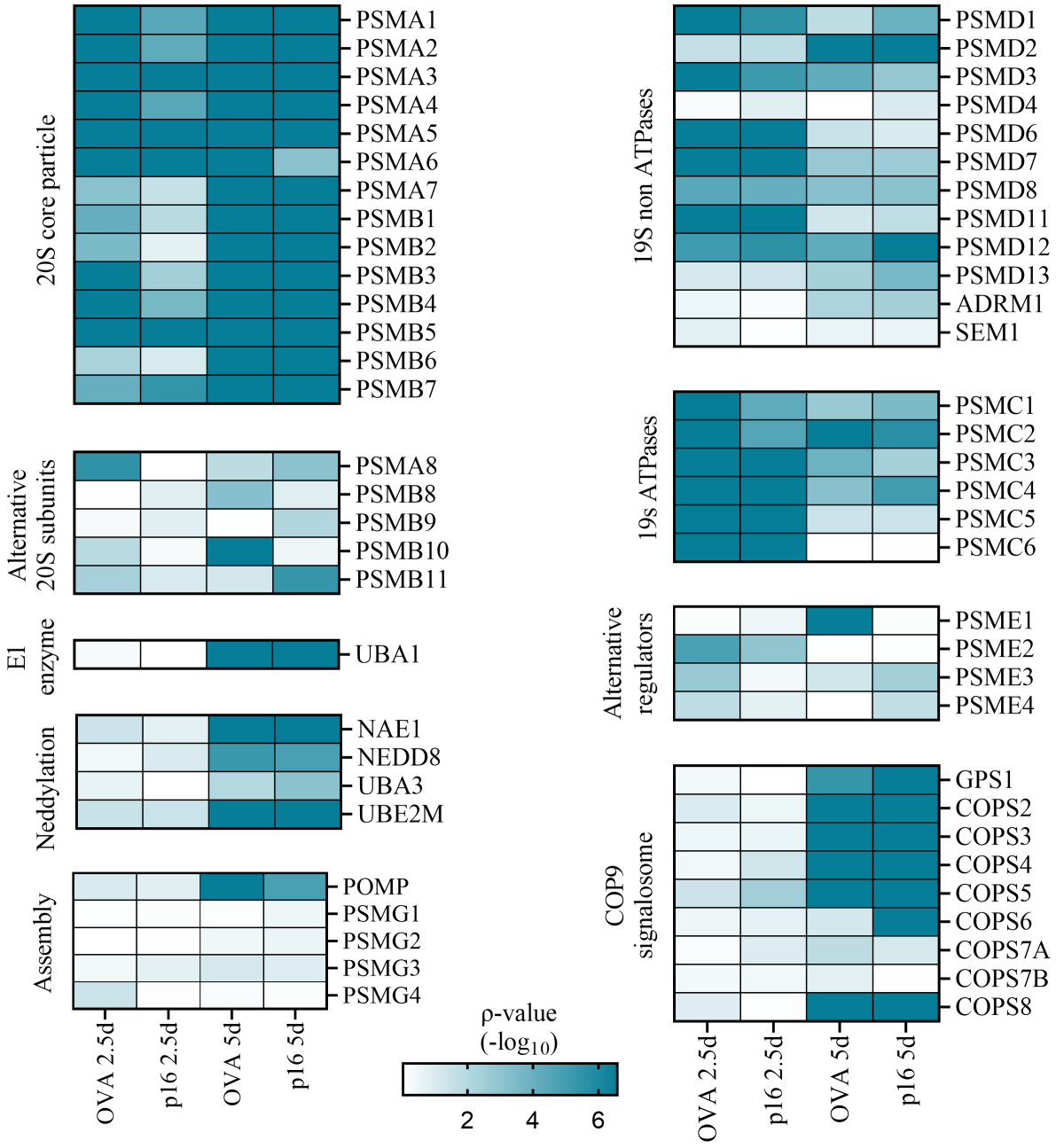

Figure 8. Analysis of selected UPS-related genes. Heat map of one-sided ρ-values calculated using MAGeCK for selected subunits of the proteasome and other genes involved in canonical protein degradation.

3.2.3 Selection of candidate genes

The candidates for validation studies were primarily selected based on their LFCs and ρ-value differences between p16 and OVA (Figure 9), and secondarily based on current literature. Criteria for candidate gene selection were further based on annotated functions and established interactions between them.

Three of the five known members of the Torsin family (*TOR1A*, *TOR1B*, and *TOR2A*) enriched in cells with delivered p16 but not in cells with delivered OVA and were therefore selected for validation. Torsins are atypical members of the AAA+ (ATPases associated with a variety of cellular activities) protein family and are the only known AAA+ ATPases to be located in the endoplasmic reticulum and nuclear envelope[94]. The mechanistic understanding of Torsins remains elusive due to their striking differences to their closest structural homologs, such as a non-canonical Walker A motif and the lack of an otherwise conserved arginine finger[94]. However, they have been found to play a role in ERAD (Endoplasmic-reticulum-associated protein degradation) of some proteins[95]. Therefore, the genes *SEC62* and *SEC63* were included, as they play a central role in translocating proteins into the ER and subsequent ERAD[96,97].

Remarkably, *COPS4*, a subunit of the COP9 signalosome, has been identified to interact with *TOR1A* and together they are required for the stability of snapin and synaptotagmin-specific endocytic adaptor stonin 2[98]. Further, COPS6, another component of the COP9 signalosome, has been shown to be directly involved in p16 degradation. This suggests that the COP9 signalosome and *TOR1A* might act in concert to regulate protein stability. Furthermore, RFWD2 (or COP1) is an E3 ligase that associates with COPS6 and together they regulated ubiquitin-dependent degradation of various proteins[99]. Additionally, a deletion in *TOR1A* has been identified as the cause of DYT1 dystonia[100]. Dystonias are a group of neurological hyperkinetic movement disorders, some of which are hereditary with known genetic causes. Interestingly, another clinically similar dystonia, DYT6, is caused by mutations in the transcription factor *THAP1*[101]. Interestingly, *TOR1A* expression can be repressed by wild type THAP1[101]. Furthermore, *THAP1* is involved in the regulation of pRB/E2F target genes[102], one of which is p16 itself. Of note, sgRNAs targeting the "safe harbor" locus AAVS1 and PSMB5 served as negative and positive controls, respectively. Based on these considerations, we focused on the validation of the genes shown in Figure 9 and listed in Table 1.

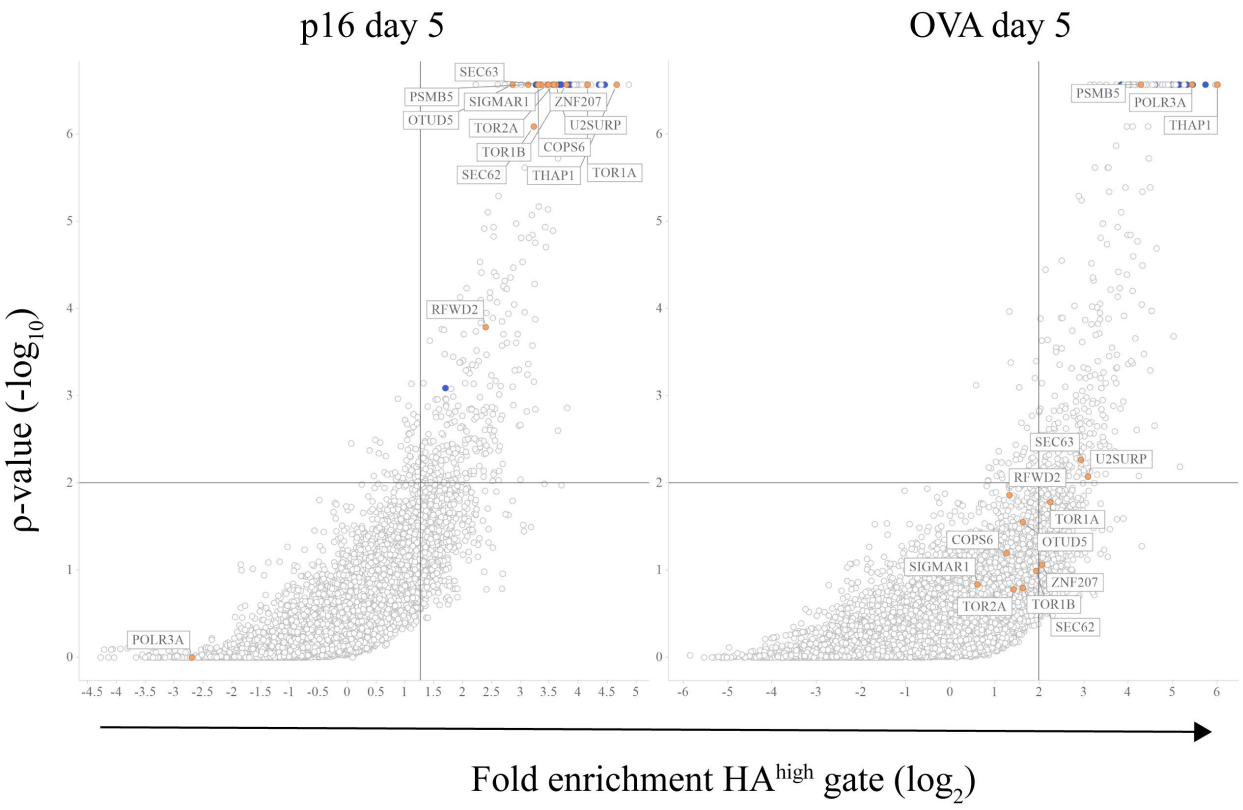

Figure 9. Selected hits for validation. Gene-level enrichment of sgRNAs in HA^{high} cells and one-sided MAGeCK ρ-values are shown. The lines indicate a ρ-value of 0.01 and the 95th percentile of enrichment. All genes that were used for validation studies are labeled and highlighted in orange, 20S proteasome core particle genes (as in **Figure 8**) are highlighted in blue.

Table 1: Selected candidate genes
Genes that were selected for validation are listed with respective gene-level enrichment of sgRNAs in HA^{high} cells and one-sided MAGeCK ρ-values. LFC, log_2 fold enrichment.

Gene	LFC p16 d2.5	ρ-value p16 d2.5	LFC p16 d5	ρ-value p16 d5	LFC OVA d2.5	ρ-value OVA d2.5	LFC OVA d5	ρ-value OVA d5
PSMB5	4.43	6.57	3.13	6.56	4.54	6.57	4.27	6.56
COPS6	2.16	0.71	3.30	6.56	0.78	0.45	1.25	1.19
TOR1A	3.57	4.03	4.16	6.56	1.84	2.99	2.24	1.78
TOR1B	3.27	4.93	3.79	6.56	2.29	2.95	1.62	0.80
TOR2A	2.82	2.24	3.57	6.56	2.31	2.45	1.41	0.78
SEC62	1.71	0.83	3.23	6.09	1.00	0.65	1.93	0.99
SEC63	2.50	2.10	3.35	6.56	1.14	0.68	2.93	2.27
RFWD2	2.76	2.48	2.40	3.79	0.33	0.42	1.33	1.86
THAP1	3.50	4.35	4.66	6.56	1.00	0.72	6.00	6.56
SIGMAR1	0.93	1.04	3.48	6.56	0.30	0.24	0.61	0.84
ZNF207	2.29	1.68	3.46	6.56	1.12	0.85	2.05	1.06
U2SURP	0.61	0.40	3.61	6.56	1.21	0.68	3.08	2.07
OTUD5	1.43	0.53	2.86	6.56	-0.75	0.04	1.63	1.55
POLR3A	-0.51	0.20	-2.69	0.00	-1.26	0.01	5.42	6.56

3.3 Validation of hits

3.3.1 Designing single gene inducible CRISPR/Cas9 knockouts

To validate and further characterize the selected genes, inducible single gene knockouts were created using the top ranking sgRNA as predicted by the Vienna Bioactivity CRISPR-score (https://www.vbc-score.org/)[91]. Additionally, two genes, *TOR1A* and *COPS6*, were targeted with the top scoring sgRNA from the CRISPR/Cas9 screen. The sgRNA sequences used can be found in Appendix C.

The RNA polymerase III subunit, *POLR3A,* was included as a control due to the unexpected strong enrichment in OVA. To this end, RKO cells containing a doxycycline (dox) inducible Cas9 were transduced with the respective sgRNA via pseudotyped lentivirus at 40-70% infection efficiency.

3.3.2 Competitive proliferation

To determine gene essentiality competitive proliferation assays were performed for every sgRNA (Figure 10). In addition to gene essentiality this assay provides crucial information about the required timing to obtain the strongest effect on overall function of the protein and at the same time mitigating effects that could stem from the cell dying. For the validation studies, Cas9 was then induced for the amount of time until knockout cells started to deplete in the proliferation assay, or for seven days when the gene was not essential for cell survival.

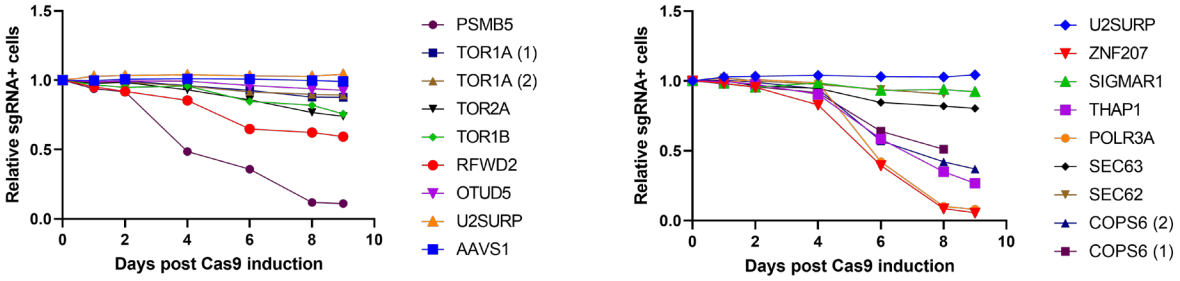

Figure 10. Competitive proliferation of candidate genes in RKO cells. Cell lines were grown for up to 9 days in the presence of dox and subsequently the frequency of sgRNA+ (BFP+) cells analyzed by flow cytometry. Data are normalized to day 0.

3.3.3 Validation of candidate genes

To validate candidate genes, the knockouts (Cas9) were induced for the appropriate amount of time as determined by the competition assay (Figure 10). Then, the standard *S.* Typhimurium-based protein delivery experiment was performed, followed by 4 hours incubation for protein degradation. The proteasomal subunit *PSMB5* served as a positive control and the "safe harbor" locus *AAVS1* as a negative control in all experiments.

As seen in Figure 11, most candidate genes did not validate since Cas9-induced cells contain the same amounts of protein as the uninduced controls. However, *THAP1* knockout results in a small fraction of cells with high p16 and OVA levels after the 4 hour incubation time. While *THAP1* knockout showed a measurable effect on p16 and OVA stability (as found in the genome-wide CRISPR/Cas9 screen), most other gene knockouts did not show a convincing phenotype. We hypothesized that this was because the analyzed cells were infected at low efficiencies and therefore cells that did not have the sgRNA were included. Therefore, the cell lines were sorted to enrich sgRNA+ cells, obtaining populations with at least 90% of cells BFP+/sgRNA+, now termed "enriched cells".

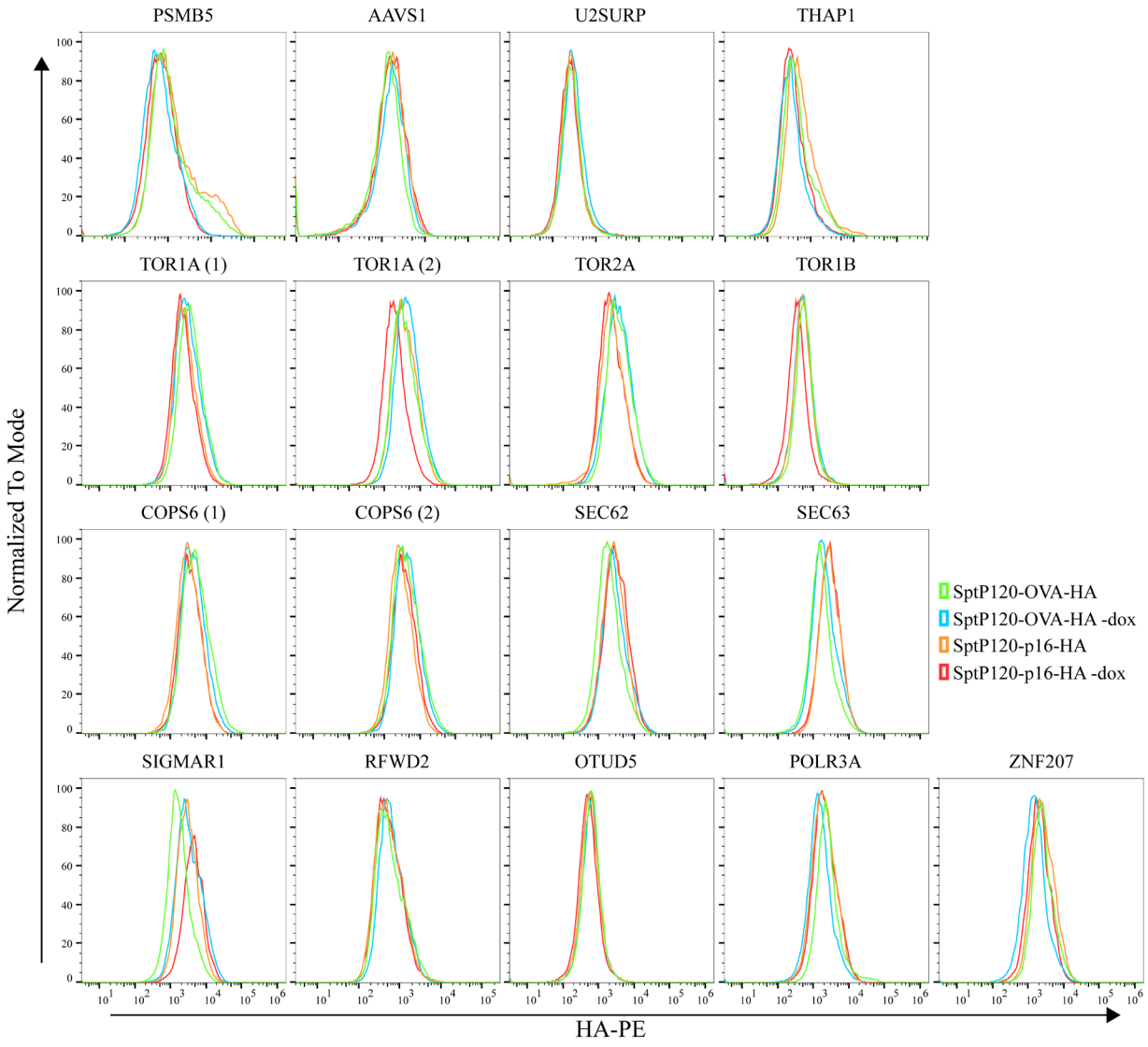

Figure 11. Validation of candidate genes in RKO cells. Flow cytometry analysis of SptP120-OVA-HA and SptP120-p16-HA transferred to RKO knockout cells (as indicated). Knockouts were induced with dox for 3 d (*PSMB5*), 4 d (*U2SURP*), 5 d (*COPS6* (1), *COPS6* (2), *TOR1A* (1)*, TOR1A* (2)*, TOR2A, THAP1, TOR1B*), 6 d (*U2SURP, SEC62, SIGMAR1, ZNF207, OTUD5)*, or 7 d (*AAVS1, POLR3A, SEC63*). Cells were infected at a MOI of 100 for 1 h and incubated for 4 h after infection. Non-induced cells served as negative control (-dox). Data is representative of three independent experiments.

Figure 12 shows the validation experiments performed on the enriched cells. Interestingly, the phenotype for *PSMB5* and *THAP1* was stronger in the enriched cells than in the unsorted population. However, knockout of the other candidate genes still did not show an effect on p16 or OVA stability. In summary, while some candidate genes did not validate strongly, this data supports *THAP1* as a hit that influences p16 and OVA stability.

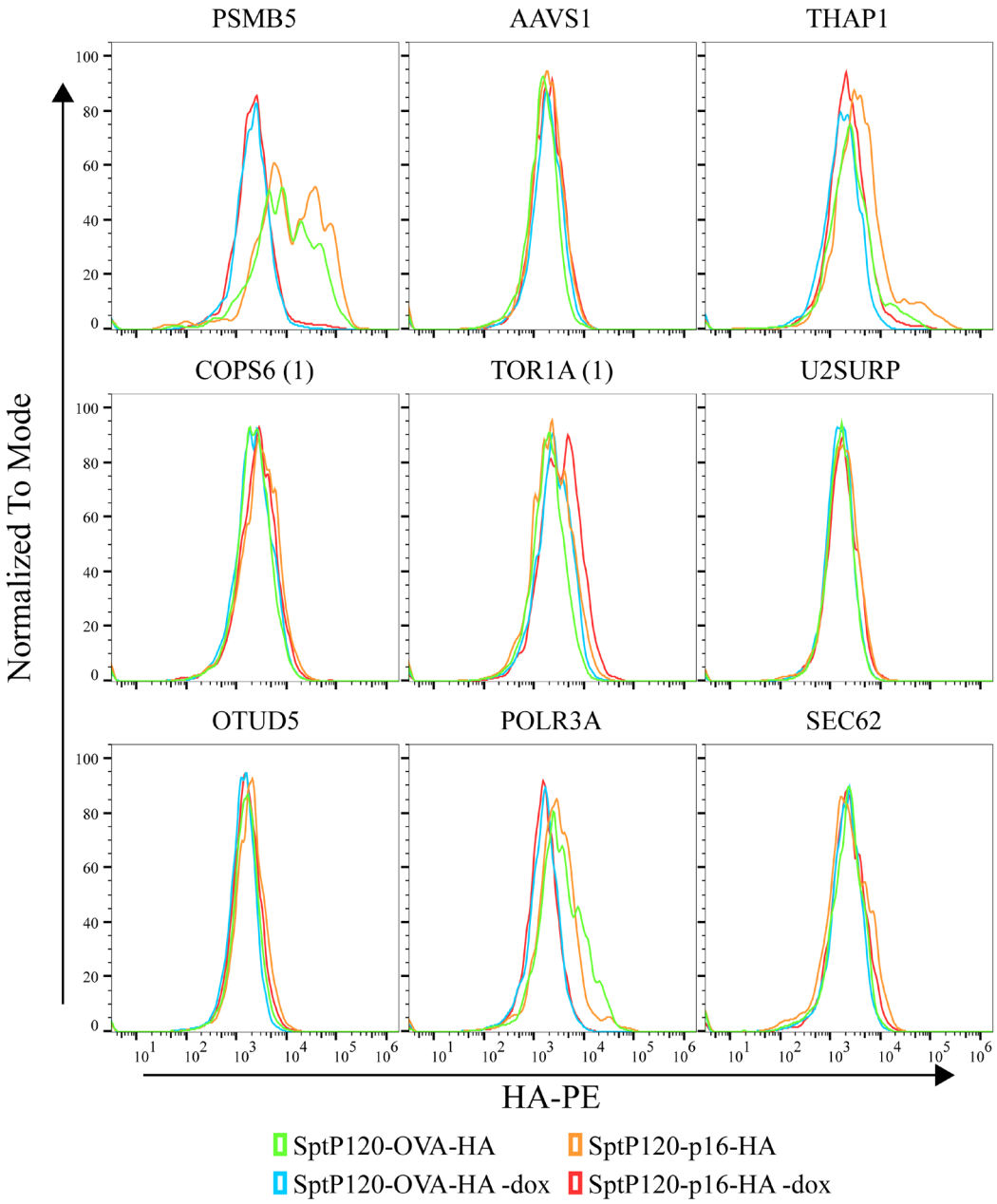

Figure 12. Validation of candidate genes in "enriched" RKO cells. Cells from **Figure 11** were sorted for sgRNA+ cells and re-cultivated. Flow cytometry analysis of SptP120-OVA-HA and SptP120-p16-HA transferred to RKO knockout cells (as indicated). Knockouts were induced with dox for 4 d (*PSMB5*), 5 d (*COPS6* (1), *TOR1A* (1), *THAP1*), 6 d (*U2SURP, POLR3A*), or 7 d (*AAVS1, OTUD5, SEC62*). Cells were infected at a MOI of 100 for 1 h and incubated for 4 h after infection. Non-induced cells served as negative control (-dox). Data is representative of three independent experiments.

3.4 Delivered p16 does not aggregate and is functional

Since *TOR1A* has been found to be involved in the degradation of protein aggregates[103,104] and some of the Torsins were highly enriched in the genome-wide CRISPR/Cas9 knockout screen, we investigated if the *S.* Typhimurium delivered SptP120-p16-HA forms aggregates within RKO cells. Immunofluorescence analysis of delivered p16 and OVA did not reveal any obvious aggregates (Figure 13A). To effectively demonstrate that the delivered p16 does not aggregate we further investigated if the delivered p16 is functional. Since p16 inhibits Rb phosphorylation through CDK4/6, we hypothesized that Rb phosphorylation levels should be reduced in cells that receive a high dose of p16. Indeed, quantitative western blotting with two antibodies against the Rb phosphorylation sites (S780 and S807/811) revealed that Rb phosphorylation is reduced (Figure 13B). These experiments show that the delivered SptP120-p16-HA does not aggregate and appears to be functional at the timepoint of analysis.

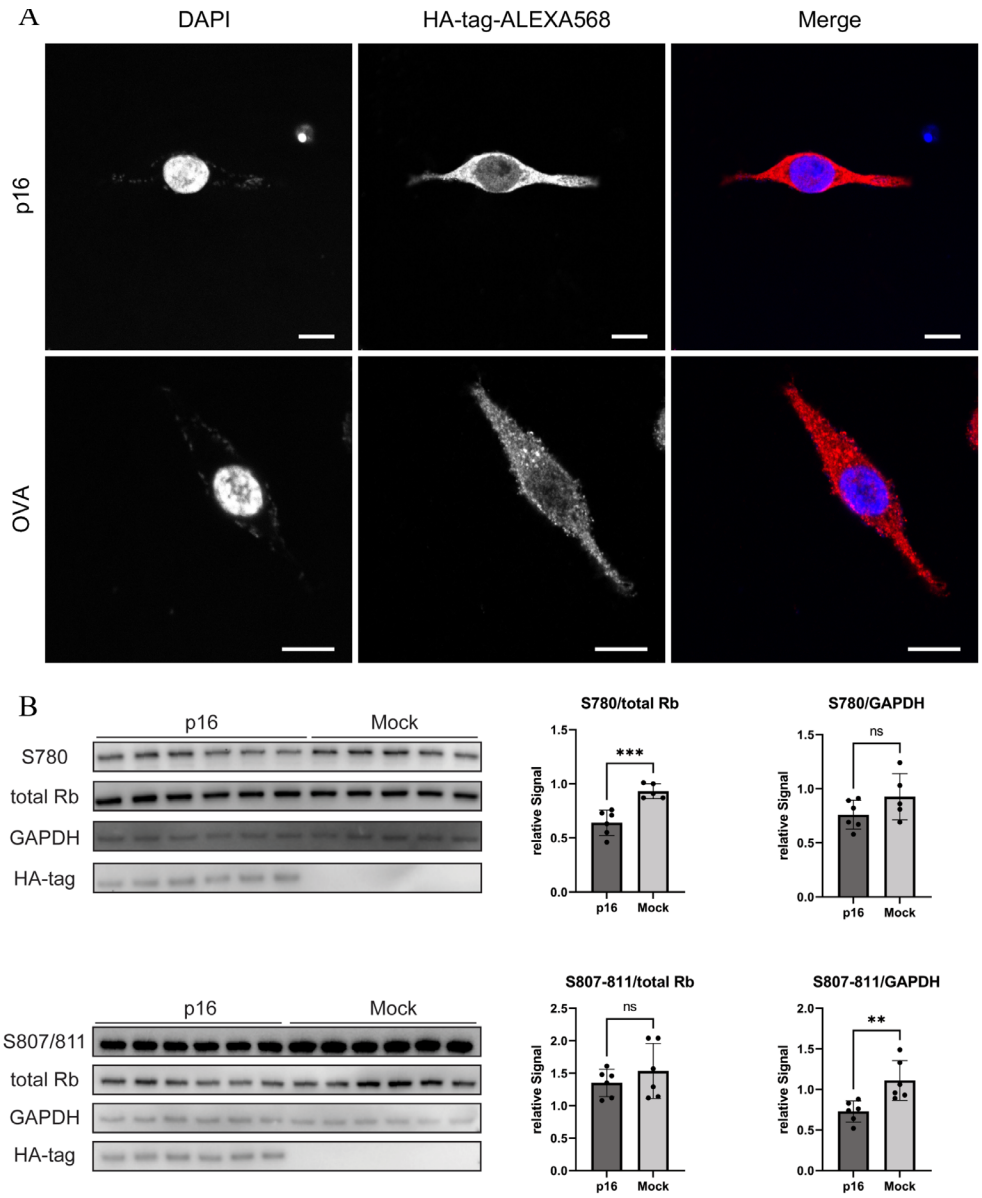

Figure 13. SptP120-p16-HA does not form aggregates and is functional. (A) Representative anti-HA immunostaining confocal images of SptP120-p16-HA and SptP120-OVA-HA delivered to RKO cells via *S.* Typhimurium ASB2519 for 1 h at a MOI of 100. In the merged image DAPI and anti-HA are pseudo colored as blue and red, respectively. 50 different cells were analyzed per condition. Scale bars 10 µm. (B) Quantitative immunoblot of RKO cells. SptP120-p16-HA was delivered to RKO cells via *S.* Typhimurium ASB2519 for 1 h and further incubated in the absence of ASB2519 for 1 h. *S.* Typhimurium ASB2519 with an empty pCASP-HilA vector served as negative control (Mock infected). HA-tags are displayed to show effective delivery of the protein complex

SptP120-p16-HA. Quantification of S780 and S807/811 phosphorylation was performed against total Rb protein and GAPDH. Statistical analysis was performed with two-sided t-test (n=6 biological replicates). Data is shown as mean ± s.d.

Chapter 4: Discussion

Protein degradation is a fundamental process in the daily life of every cell and the degradation of most proteins is mediated by specific enzymes[4]. Since the abundance of a protein often determines its cellular function and degradation is one of the two major means of a cell to regulate protein homeostasis, it is important to understand how a specific protein of interest is degraded. We hypothesized that cytoplasmic protein delivery via *S.* Typhimurium is a viable platform to rapidly study the degradation of proteins of interest using genome-wide CRISPR/Cas9 knockout screening. Thus, in this thesis we performed an exemplary screen to study the degradation of the tumor suppressor p16 and validated the candidate genes by performing single gene knockout experiments. Additionally, we were able to show that the delivered SptP120-p16-HA does not aggregate and is functional in cells.

By using an inducible Cas9 we were able to screen the entire genome, instead of just non-essential genes which has been a major limitation in many other genetic screens. By screening at two timepoints, 2.5 and 5 days after Cas9 induction, we can also detect temporally resolved effects of Cas9 mediated gene-knockouts. Coupled with the highly efficient protein delivery via the avirluent *S.*

Typhimurium strain ASB2519, inducible genome-wide CRISPR/Cas9 knockout screening becomes a powerful tool to study the degradation of proteins of interest with minimal setup time. This is facilitated by the ease and speed of cloning the pCASP-HilA plasmid for intracellular protein delivery via *S.* Typhimurium.

While the analysis of the raw sequencing data with MAGeCK revealed high quality data, the validation of candidate genes posed more challenging. Many of the candidate genes in both (days 2.5 and 5) HA^{high} fractions, despite having good LFCs, FDRs and ρ-values (almost equal to the positive control *PSMB5*), did not validate in single gene knockout cells as true positives. Given that the entire proteasome (and other genes that have a more specific link to p16 degradation) enrich in the HA^{high} fractions, we hypothesize that this was an issue of the delivery assay not being sensitive enough for the validations. More specificially, the problem is rooted in mean fluorescence intensities varying too much from well to well to confidently detect small changes, and not of detecting false positives in the genome-wide CRISPR/Cas9 knockout screen. Furthermore, in the screen every sgRNA was interogated against each other, therefore serving as internal controls and eliminating well to well variation which allows for the detection of smaller changes in phenotypes. To solve this validation problem, the sgRNA coupled BFP

reporter could be used to directly compare cells that have the gene knockout to cells without an integrated sgRNA within a well, however that did not work reproducibly in our hands. The BFP signal is relatively weak (compared to for example the GFP signal that is coupled to Cas9 expression) and even further quenched by fixation. Therefore, it was not possible for us to gate BFP+ and BFP- populations in the flow cytometry analysis.

An issue with FACS-based screens is that the effect sizes in the screen do not represent biological effect sizes and hits are therefore not directly comparable. Furthermore, they are limited to measure one parameter and therefore only produce a single enrichment score that makes it impossible to directly draw any information (e.g. pathway or dynamic and complex phenotypes) from these hits. Recently, optical pooled screening setups with microscopy-based read outs of multiple phenotypes[105] or temporally resolved information[106] coupled to *in situ* sequencing have shown a way to overcome these issues.

Interestingly, based on the defined criteria, the REGγ proteasomal activator (*PSME3)* did not qualify as a hit for p16 degration, despite multiple studies suggesting its involvement in the degradation of p16[41,42]. This might be due to

subcellular sequestration: REGγ is preferentially localized to the nucleus and the delivered SptP120-p16-HA localizes primarily to the cytoplasm (Figure 13A). However, SUMOylation of REGγ results in cytosolic translocation of the protein[50] which is mediated by p14[41]. In RKO cells, p14 expression is relatively low according to DepMap[107] and therefore sufficient amounts of REGγ might not translocate into the cytoplasm. Regardless of REGγ, p16 is located to both the nucleus and the cytoplasm and therefore even if nuclear p16 is degraded by REGγ, a cytoplasmic p16 degradation pathway must still be in place. The functionality of the delivered p16 (Figure 13B), despite being expressed in bacteria and the addition of the secretion signal SptP120 and HA-tag, provides additional proof that the delivered p16 is degraded via its canonical cytoplasmic pathway. Furthermore, *COPS6*, also previously found to be involved in the degradation of p16[52], was a top hit for p16 but not OVA degradation, suggesting that the obtained screen results are, at least in part, reliable.

The strongest candidate gene, *THAP1*, has, to our knowledge, not been shown to be involved in the degradation of any protein. Recently however, *THAP1* was found in a genome-wide CRISPR/Cas9 knockout screen to regulate MYC stability[92]. This implies that *THAP1* could be a general regulator of protein

degradation that has not been described previously. *THAP1* has been shown to be involved in the regulation of pRB/E2F target genes, and thereby regulate cell proliferation[102]. The minute control of protein levels (specifically of cyclins and CDK inhibitors such as p16) is important for cell cycle progression and thus cell proliferation[4]. As a transcription factors, *THAP1* regulates the expression of specific target genes. These genes could be components of the protein degradation machinery, such as E3 ligases, to tightly regulate the degradation of proteins that are involved in the cell cycle. *TOR1A*, for example, is directly involved in ERAD[95] and its expression can be repressed by wild type THAP1[101]. Additionally, it has been shown that *THAP1* directly regulates the expression of β-glucuronidase, a lysosomal enzyme involved in the catabolism of glycosaminoglycan in the extra cellular matrix thereby controlling the maturation of oligodendroglial progenitor cells into myelinating cells[108].

We used the T3SS delivery system of *S.* Typhimurium due to the high amounts of protein that can be delivered into the cytosol, the dynamic degradation of these proteins, its capability to infect many cell types, and the low cytotoxicity of the modified ASB2519 strain. None of the high ranking hits in the genome-wide CRISPR/Cas9 knockout screen were related to *S.* Typhimurium or the T3SS as the

mode of protein delivery. Therefore, the genome-wide CRISPR/Cas9 knockout screening via *S.* Typhimurium is a viable tool to obtain high quality data. However, it might be beneficial to employ a secondary screen of a smaller subset of target genes or using different cell lines. Here, the endogenous protein of interest could be stained directly or endogenously tagged with a fluorescent reporter to asses steady state levels, potentially also allowing to validate smaller phenotypic changes. Since this smaller follow-up screen would be focused to the significant targets found in the first *S.* Typhimurium based screen, other genes, e.g. required for transcription or translation, that would otherwise confound the screen analysis could be excluded. The capture of only genes that are directly involved in the degradation of the protein of interest is a major advantage of our system using intracellular protein delivery via *S.* Typhimurium.

This work showed a viable method to rapidly perform genome-wide CRISPR/Cas9 screening for genes that are involved in the degradation of p16. With an appropriate validation strategy, this method can be use to study the degradation of proteins of interest. We found and validated *THAP1* as a potential general regulator of protein degradation that warrants further research. Other genes that highly

enriched in the screen but did not validate could be validated using other methods or cell lines.

References

1. Clausen, L. *et al.* Protein stability and degradation in health and disease. *Adv. Protein Chem. Struct. Biol.* **114**, 61–83 (2019).
2. Mathieson, T. *et al.* Systematic analysis of protein turnover in primary cells. *Nat. Commun. 2018 91* **9**, 1–10 (2018).
3. Chen, W., Smeekens, J. M. & Wu, R. Systematic study of the dynamics and half-lives of newly synthesized proteins in human cells. *Chem. Sci.* **7**, 1393–1400 (2016).
4. Collins, G. A. & Goldberg, A. L. The Logic of the 26S Proteasome. *Cell* **169**, 792–806 (2017).
5. Hanahan, D. & Weinberg, R. A. Hallmarks of cancer: The next generation. *Cell* **144**, 646–674 (2011).
6. Zheng, Q. *et al.* Dysregulation of ubiquitin-proteasome system in neurodegenerative diseases. *Front. Aging Neurosci.* **8**, 303 (2016).
7. Willis, M. S. *et al.* The role of ubiquitin ligases in cardiac disease. *J. Mol. Cell. Cardiol.* **71**, 43–53 (2014).
8. Rousseau, A. & Bertolotti, A. Regulation of proteasome assembly and activity in health and disease. *Nat. Rev. Mol. Cell Biol. 2018 1911* **19**, 697–712 (2018).
9. Lecker, S. H., Goldberg, A. L. & Mitch, W. E. Protein Degradation by the Ubiquitin–Proteasome Pathway in Normal and Disease States. *J. Am. Soc. Nephrol.* **17**, 1807–1819 (2006).
10. Buetow, L. & Huang, D. T. Structural insights into the catalysis and regulation of E3 ubiquitin ligases. *Nat. Rev. Mol. Cell Biol. 2016 1710* **17**, 626–642 (2016).

11. Barghout, S. H. & Schimmer, A. D. E1 Enzymes as Therapeutic Targets in Cancer. *Pharmacol. Rev.* **73**, 1–58 (2021).

12. Li, W. *et al.* Genome-Wide and Functional Annotation of Human E3 Ubiquitin Ligases Identifies MULAN, a Mitochondrial E3 that Regulates the Organelle's Dynamics and Signaling. *PLoS One* **3**, (2008).

13. Cohen, P. & Tcherpakov, M. Will the Ubiquitin System Furnish as Many Drug Targets as Protein Kinases? *Cell* **143**, 686–693 (2010).

14. Tokheim, C. *et al.* Systematic characterization of mutations altering protein degradation in human cancers. *Mol. Cell* (2021) doi:10.1016/j.molcel.2021.01.020.

15. Metzger, M. B., Hristova, V. A. & Weissman, A. M. HECT and RING finger families of E3 ubiquitin ligases at a glance. *J. Cell Sci.* **125**, 531–537 (2012).

16. Berndsen, C. E. & Wolberger, C. New insights into ubiquitin E3 ligase mechanism. *Nat. Struct. Mol. Biol. 2014 214* **21**, 301–307 (2014).

17. Cairns, P. *et al.* Rates of p16 (MTS1) mutations in primary tumors with 9p loss. *Science* **265**, 415–417 (1994).

18. Kent, L. N. & Leone, G. The broken cycle: E2F dysfunction in cancer. *Nat. Rev. Cancer 2019 196* **19**, 326–338 (2019).

19. Harbour, J. W. & Dean, D. C. The Rb/E2F pathway: expanding roles and emerging paradigms. *Genes Dev.* **14**, 2393–2409 (2000).

20. Dick, F. A. & Rubin, S. M. Molecular mechanisms underlying RB protein function. *Nat. Rev. Mol. Cell Biol.* **14**, 297 (2013).

21. Giacinti, C. & Giordano, A. RB and cell cycle progression. *Oncogene 2006 2538* **25**,

5220–5227 (2006).

22. Khleif, S. N. *et al.* Inhibition of cyclin D-CDK4/CDK6 activity is associated with an E2F-mediated induction of cyclin kinase inhibitor activity. *Proc. Natl. Acad. Sci.* **93**, 4350–4354 (1996).

23. Li, Y., Nichols, M. A., Shay, J. W. & Xiong, Y. Transcriptional Repression of the D-Type Cyclin-dependent Kinase Inhibitor p16 by the Retinoblastoma Susceptibility Gene Product pRb. *Cancer Res.* **54**, (1994).

24. Lin, J., Albers, A. E., Qin, J. & Kaufmann, A. M. Prognostic Significance of Overexpressed p16INK4a in Patients with Cervical Cancer: A Meta-Analysis. *PLoS One* **9**, e106384 (2014).

25. Weinberger, P. M. *et al.* Prognostic Significance of p16 Protein Levels in Oropharyngeal Squamous Cell Cancer. *Clin. Cancer Res.* **10**, 5684–5691 (2004).

26. Lebok, P. *et al.* p16 overexpression and 9p21 deletion are linked to unfavorable tumor phenotype in breast cancer. *Oncotarget* **7**, 81322–81331 (2016).

27. Lee, C. T. *et al.* Overexpression of the Cyclin-dependent Kinase Inhibitor p16 Is Associated with Tumor Recurrence in Human Prostate Cancer. *Clin. Cancer Res.* **5**, (1999).

28. Romagosa, C. *et al.* p16Ink4a overexpression in cancer: a tumor suppressor gene associated with senescence and high-grade tumors. *Oncogene 2011 3018* **30**, 2087–2097 (2011).

29. Leon, K. E., Tangudu, N. K., Aird, K. M. & Buj, R. Loss of p16: A Bouncer of the Immunological Surveillance? *Life 2021, Vol. 11, Page 309* **11**, 309 (2021).

30. Jung, A. *et al.* The invasion front of human colorectal adenocarcinomas shows co-localization of nuclear beta-catenin, cyclin D1, and p16INK4A and is a region of low proliferation. *Am. J. Pathol.* **159**, 1613–1617 (2001).

31. Horrée, N., van Diest, P. J., Sie-Go, D. M. D. S. & Heintz, A. P. M. The invasive front in endometrial carcinoma: higher proliferation and associated derailment of cell cycle regulators. *Hum. Pathol.* **38**, 1232–1238 (2007).

32. Svensson, S., Nilsson, K., Ringberg, A. & Landberg, G. Invade or Proliferate? Two Contrasting Events in Malignant Behavior Governed by p16INK4a and an Intact Rb Pathway Illustrated by a Model System of Basal Cell Carcinoma. *Cancer Res.* **63**, (2003).

33. Lu, Y., Zhang, X. & Zhang, J. Inhibition of Breast Tumor Cell Growth by Ectopic Expression of p16/INK4A Via Combined Effects of Cell Cycle Arrest, Senescence and Apoptotic Induction, and Angiogenesis Inhibition. *J. Cancer* **3**, 333 (2012).

34. Rayess, H., Wang, M. B. & Srivatsan, E. S. Cellular senescence and tumor suppressor gene p16. *Int. J. Cancer* **130**, 1715 (2012).

35. Buj, R. *et al.* Suppression of p16 Induces mTORC1-Mediated Nucleotide Metabolic Reprogramming. *Cell Rep.* **28**, 1971-1980.e8 (2019).

36. Xiong, Y., Zhang, H. & Beach, D. Subunit rearrangement of the cyclin-dependent kinases is associated with cellular transformation. *Genes Dev.* **7**, 1572–1583 (1993).

37. Nilsson, K. & Landberg, G. Subcellular localization, modification and protein complex formation of the cdk-inhibitor p16 in Rb-functional and Rb-inactivated tumor cells. *Int. J. Cancer* **118**, 1120–1125 (2006).

38. Evangelou, K. *et al.* Electron microscopy evidence that cytoplasmic localization of the

p16(INK4A) ‘nuclear’ cyclin-dependent kinase inhibitor (CKI) in tumor cells is specific and not an artifact. A study in non-small cell lung carcinomas. *Biotech. Histochem.* **79**, 5–10 (2004).

39. Souza-Rodrígues, E. *et al.* Proteomic analysis of p16ink4a-binding proteins. *Proteomics* **7**, 4102–4111 (2007).

40. Gombart, A. F., Yang, R., Campbell, M. J., Berman, J. D. & Koeffler, H. P. Inhibition of growth of human leukemia cell lines by retrovirally expressed wild-type p16INK4A. *Leuk. 1997 1110* **11**, 1673–1680 (1997).

41. Kobayashi, T., Wang, J., Al-Ahmadie, H. & Abate-Shen, C. ARF regulates the stability of p16 protein via REGγ-dependent proteasome degradation. *Mol. Cancer Res.* **11**, 828 (2013).

42. Chen, X., Barton, L. F., Chi, Y., Clurman, B. E. & Roberts, J. M. Ubiquitin-independent degradation of cell cycle inhibitors by the REGγ proteasome. *Mol. Cell* **26**, 843 (2007).

43. Cha, S., Park, I. & Jang, K. L. Hepatitis C virus core protein activates proteasomal activator 28 gamma to downregulate p16 levels via ubiquitin-independent proteasomal degradation. *Heliyon* **7**, e06134 (2021).

44. Al-Khalaf, H. H., Hendrayani, S. F. & Aboussekhra, A. The Atr Protein Kinase Controls UV-Dependent Upregulation of p16INK4A Through Inhibition of Skp2-Related Polyubiquitination/Degradation. *Mol. Cancer Res.* **9**, 311–319 (2011).

45. McClellan, A. J., Laugesen, S. H. & Ellgaard, L. Cellular functions and molecular mechanisms of non-lysine ubiquitination. *Open Biol.* **9**, 190147 (2019).

46. Ben-Saadon, R. *et al.* The Tumor Suppressor Protein p16INK4a and the Human

Papillomavirus Oncoprotein-58 E7 Are Naturally Occurring Lysine-less Proteins That Are Degraded by the Ubiquitin System: DIRECT EVIDENCE FOR UBIQUITINATION AT THE N-TERMINAL RESIDUE *. *J. Biol. Chem.* **279**, 41414–41421 (2004).

47. Ben-Nissan, G. & Sharon, M. Regulating the 20S Proteasome Ubiquitin-Independent Degradation Pathway. *Biomol. 2014, Vol. 4, Pages 862-884* **4**, 862–884 (2014).

48. Shen, M. *et al.* Role of oncogenic REGγ in cancer. *Biomed. Pharmacother.* **130**, 110614 (2020).

49. Jariel-Encontre, I., Bossis, G. & Piechaczyk, M. Ubiquitin-independent degradation of proteins by the proteasome. *Biochim. Biophys. Acta - Rev. Cancer* **1786**, 153–177 (2008).

50. Wu, Y. *et al.* Regulation of REGγ cellular distribution and function by SUMO modification. *Cell Res.* **21**, 807 (2011).

51. Park, S. H., Lim, J. S., Lim, S. Y., Tiwari, I. & Jang, K. L. Hepatitis C virus Core protein stimulates cell growth by down-regulating p16 expression via DNA methylation. *Cancer Lett.* **310**, 61–68 (2011).

52. Du, W. *et al.* CSN6 promotes tumorigenesis of gastric cancer by ubiquitin-independent proteasomal degradation of p16INK4a. *Cancer Biol. Med.* **16**, 514 (2019).

53. Dubiel, W., Chaithongyot, S., Dubiel, D. & Naumann, M. The COP9 Signalosome: A Multi-DUB Complex. *Biomol. 2020, Vol. 10, Page 1082* **10**, 1082 (2020).

54. Kawasaki, K. *et al.* FAM111B enhances proliferation of KRAS-driven lung adenocarcinoma by degrading p16. *Cancer Sci.* **111**, 2635 (2020).

55. Coryell, P. R. *et al.* Autophagy regulates the localization and degradation of p16INK4a. *Aging Cell* **19**, e13171 (2020).

56. Tindall, B. J., Grimont, P. A. D., Garrity, G. M. & Euzéby, J. P. Nomenclature and taxonomy of the genus Salmonella. *Int. J. Syst. Evol. Microbiol.* **55**, 521–524 (2005).

57. Kozak, G. K., Macdonald, D., Landry, L. & Farber, J. M. Foodborne Outbreaks in Canada Linked to Produce: 2001 through 2009. *J. Food Prot.* **76**, 173–183 (2013).

58. Galanis, E. *et al.* Web-based Surveillance and Global Salmonella Distribution, 2000–2002. *Emerg. Infect. Dis.* **12**, 381 (2006).

59. Jantsch, J., Chikkaballi, D. & Hensel, M. Cellular aspects of immunity to intracellular Salmonella enterica. *Immunol. Rev.* **240**, 185–195 (2011).

60. Kauffmann, F. Das Fundament zur Geschichte und Bedeutung der Salmonella- und Escherichia-Forschung = The basis on the history and significance of Salmonella- and escherichia-research. 110 (1978).

61. Wray, C. & Wray, A. *Salmonella in domestic animals.* (CABI, 2000). doi:10.1079/9780851992617.0000.

62. House, D., Bishop, A., Parry, C., Dougan, G. & Wain, J. Typhoid fever: pathogenesis and disease. *Curr. Opin. Infect. Dis.* **14**, 573–578 (2001).

63. Chabloz, A. *et al.* Salmonella-based platform for efficient delivery of functional binding proteins to the cytosol. *Commun. Biol. 2020 31* **3**, 1–11 (2020).

64. dos Santos, A. M. P., Ferrari, R. G. & Conte-Junior, C. A. Virulence Factors in Salmonella Typhimurium: The Sagacity of a Bacterium. *Curr. Microbiol. 2018 766* **76**, 762–773 (2018).

65. Hansen-Wester, I. & Hensel, M. Salmonella pathogenicity islands encoding type III secretion systems. *Microbes Infect.* **3**, 549–559 (2001).

66. Figueira, R. & Holden, D. W. Functions of the Salmonella pathogenicity island 2 (SPI-2) type III secretion system effectors. *Microbiology* **158**, 1147–1161 (2012).

67. Galán, J. E., Lara-Tejero, M., Marlovits, T. C. & Wagner, S. Bacterial Type III Secretion Systems: Specialized Nanomachines for Protein Delivery into Target Cells. *http://dx.doi.org/10.1146/annurev-micro-092412-155725* **68**, 415–438 (2014).

68. Dey, S., Chakravarty, A., Guha Biswas, P. & De Guzman, R. N. The type III secretion system needle, tip, and translocon. *Protein Sci.* **28**, 1582 (2019).

69. Notti, R. Q. & Stebbins, C. E. The Structure and Function of Type III Secretion Systems. *Microbiol. Spectr.* **4**, (2016).

70. Galán, J. E. Common themes in the design and function of bacterial effectors. *Cell Host Microbe* **5**, 571 (2009).

71. Kubori, T. *et al.* Supramolecular structure of the Salmonella typhimurium type III protein secretion system. *Science* **280**, 602–605 (1998).

72. Park, D. *et al.* Visualization of the type III secretion mediated Salmonella–host cell interface using cryo-electron tomography. *Elife* **7**, (2018).

73. Akeda, Y. & Galán, J. E. Chaperone release and unfolding of substrates in type III secretion. *Nat. 2005 4377060* **437**, 911–915 (2005).

74. Lara-Tejero, M., Kato, J., Wagner, S., Liu, X. & Galán, J. E. A Sorting Platform Determines the Order of Protein Secretion in Bacterial Type III Systems. *Science* **331**, 1188–1191 (2011).

75. Stebbins, C. E. & Galán, J. E. Maintenance of an unfolded polypeptide by a cognate chaperone in bacterial type III secretion. *Nat. 2001 4146859* **414**, 77–81 (2001).

76. Radics, J., Königsmaier, L. & Marlovits, T. C. Structure of a pathogenic type 3 secretion system in action. *Nat. Struct. Mol. Biol. 2013 211* **21**, 82–87 (2013).

77. LeBlanc, M. A., Fink, M. R., Perkins, T. T. & Sousa, M. C. Type III secretion system effector proteins are mechanically labile. *Proc. Natl. Acad. Sci.* **118**, (2021).

78. Spanò, S., Gao, X., Hannemann, S., Lara-Tejero, M. & Galán, J. E. A Bacterial Pathogen Targets a Host Rab-Family GTPase Defense Pathway with a GAP. *Cell Host Microbe* **19**, 216 (2016).

79. Chabloz, A. Engineered Salmonella as a platform for intracellular protein delivery. (Vienna University, 2018).

80. Jinek, M. *et al.* A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science (80-.).* **337**, 816–821 (2012).

81. Gasiunas, G., Barrangou, R., Horvath, P. & Siksnys, V. Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. *Proc. Natl. Acad. Sci. U. S. A.* **109**, E2579 (2012).

82. Jackson, A. L. *et al.* Widespread siRNA 'off-target' transcript silencing mediated by seed region sequence complementarity. *RNA* **12**, 1179–1187 (2006).

83. Buehler, E., Chen, Y. C. & Martin, S. C911: A Bench-Level Control for Sequence Specific siRNA Off-Target Effects. *PLoS One* **7**, e51942 (2012).

84. Shalem, O., Sanjana, N. E. & Zhang, F. High-throughput functional genomics using CRISPR–Cas9. *Nat. Rev. Genet. 2015 165* **16**, 299–311 (2015).

85. Echeverri, C. J. *et al.* Minimizing the risk of reporting false positives in large-scale RNAi screens. *Nat. Methods* **3**, 777–779 (2006).

86. Han, H. A., Pang, J. K. S. & Soh, B. S. Mitigating off-target effects in CRISPR/Cas9-mediated in vivo gene editing. *J. Mol. Med.* **98**, 615–632 (2020).

87. Pattanayak, V. *et al.* High-throughput profiling of off-target DNA cleavage reveals RNA-programmed Cas9 nuclease specificity. *Nat. Biotechnol. 2013 319* **31**, 839–843 (2013).

88. Hart, T., Brown, K. R., Sircoulomb, F., Rottapel, R. & Moffat, J. Measuring error rates in genomic perturbation screens: gold standards for human functional genomics. *Mol. Syst. Biol.* **10**, 733 (2014).

89. Yu, J. S. L. & Yusa, K. Genome-wide CRISPR-Cas9 screening in mammalian cells. *Methods* **164**–**165**, 29–35 (2019).

90. Hanna, R. E. & Doench, J. G. Design and analysis of CRISPR–Cas experiments. *Nat. Biotechnol. 2020 387* **38**, 813–823 (2020).

91. Michlits, G. *et al.* Multilayered VBC score predicts sgRNAs that efficiently generate loss-of-function alleles. *Nat. Methods 2020 177* **17**, 708–716 (2020).

92. de Almeida, M. *et al.* AKIRIN2 controls the nuclear import of proteasomes in vertebrates. *Nat. 2021* 1–6 (2021) doi:10.1038/s41586-021-04035-8.

93. Li, W. *et al.* MAGeCK enables robust identification of essential genes from genome-scale CRISPR/Cas9 knockout screens. *Genome Biol.* **15**, 554 (2014).

94. Rose, A. E., Brown, R. S. H. & Schlieker, C. Torsins: Not Your Typical AAA+ ATPases. *Crit. Rev. Biochem. Mol. Biol.* **50**, 532 (2015).

95. Nery, F. C. *et al.* TorsinA participates in endoplasmic reticulum-associated degradation. *Nat. Commun. 2011 21* **2**, 1–10 (2011).

96. Conti, B. J., Devaraneni, P. K., Yang, Z., David, L. L. & Skach, W. R. Cotranslational

Stabilization of Sec62/63 within the ER Sec61 Translocon is Controlled by Distinct Substrate-Driven Translocation Events. *Mol. Cell* **58**, 269 (2015).

97. Linxweiler, M., Schick, B. & Zimmermann, R. Let's talk about Secs: Sec61, Sec62 and Sec63 in signal transduction, oncology and personalized medicine. *Signal Transduct. Target. Ther. 2017 21* **2**, 1–10 (2017).

98. Granata, A., Koo, S. J., Haucke, V., Schiavo, G. & Warner, T. T. CSN complex controls the stability of selected synaptic proteins via a torsinA-dependent process. *EMBO J.* **30**, 181 (2011).

99. Choi, H. H. & Lee, M. H. CSN6-COP1 axis in cancer. *Aging (Albany NY)* **7**, 461 (2015).

100. Ozelius, L. J. *et al.* The early-onset torsion dystonia gene (DYT1) encodes an ATP-binding protein. *Nat. Genet. 1997 171* **17**, 40–48 (1997).

101. Kaiser, F. J. *et al.* The dystonia gene DYT1 is repressed by the transcription factor THAP1 (DYT6). *Ann. Neurol.* **68**, 554–559 (2010).

102. Cayrol, C. *et al.* The THAP–zinc finger protein THAP1 regulates endothelial cell proliferation through modulation of pRB/E2F cell-cycle target genes. *Blood* **109**, 584–594 (2007).

103. Caldwell, G. A. *et al.* Suppression of polyglutamine-induced protein aggregation in Caenorhabditis elegans by torsin proteins. *Hum. Mol. Genet.* **12**, 307–319 (2003).

104. McLean, P. J. *et al.* TorsinA and heat shock proteins act as molecular chaperones: suppression of α-synuclein aggregation. *J. Neurochem.* **83**, 846–854 (2002).

105. Funk, L. *et al.* The phenotypic landscape of essential human genes. *bioRxiv* 2021.11.28.470116 (2021) doi:10.1101/2021.11.28.470116.

106. Feldman, D. *et al.* Optical Pooled Screens in Human Cells. *Cell* **179**, 787-799.e17 (2019).

107. DepMap, Broad (2021): DepMap 21Q4 Public. figshare. Dataset. https://doi.org/10.6084/m9.figshare.16924132.v1 (2021).

108. Yellajoshyula, D. *et al.* THAP1 modulates oligodendrocyte maturation by regulating ECM degradation in lysosomes. *Proc. Natl. Acad. Sci. U. S. A.* **118**, (2021).

Appendices

Appendix A pCASP-HilA SptP120-p16-HA plasmid map

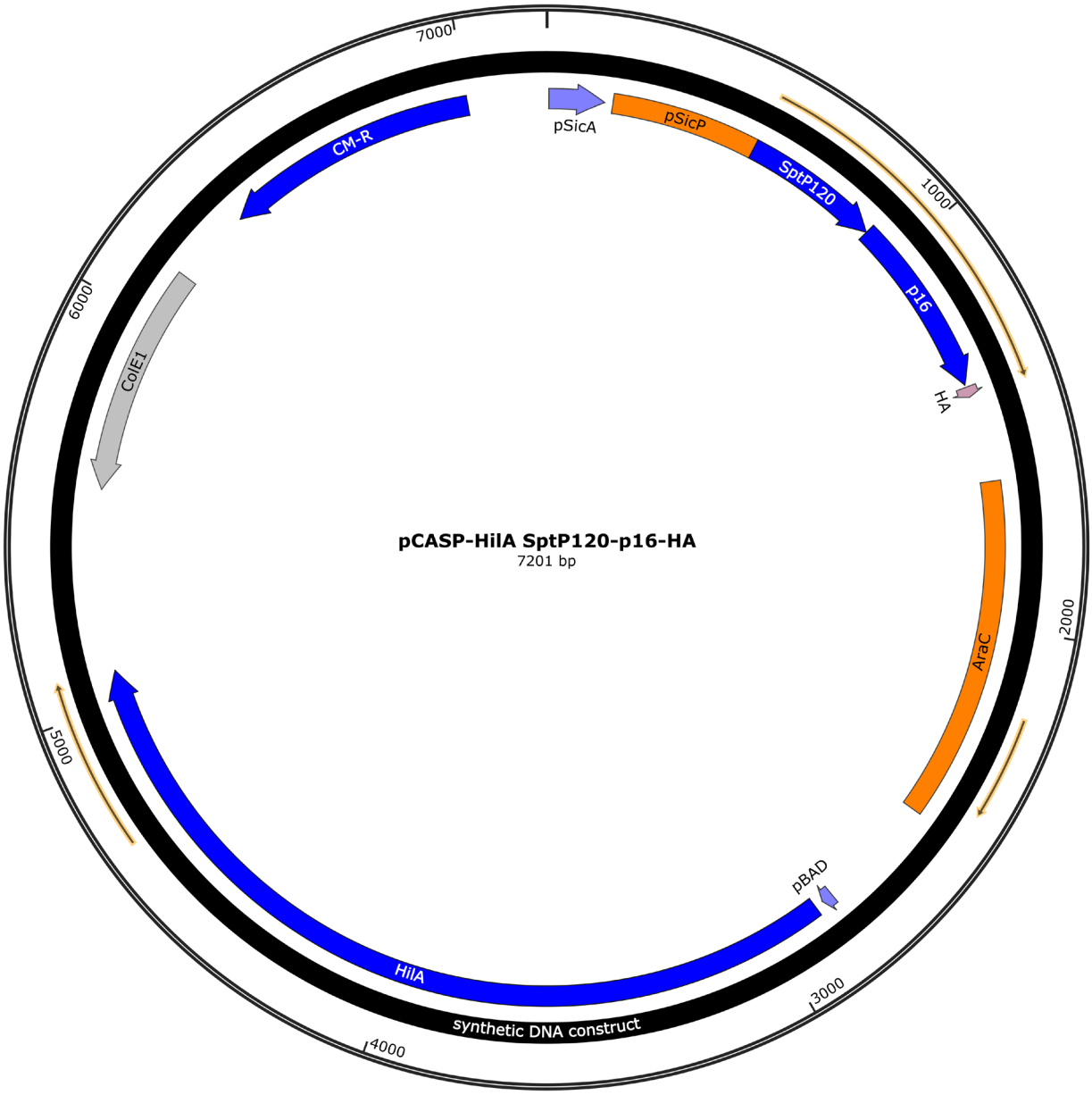

Figure 14. pCASP-HilA SptP120-p16-HA plasmid map. Purple boxes represent promotors with the arrow indicating the direction of read. (CM-R, chloramphenicol resistance; AraC, L-arabinose operon; pBAD, arabinose promotor; ColE1, origin)

Appendix B NGS library primer sequences

Table 2: Primers used for 2-step PCR to generate NGS libraries
NN denotes random nucleotides

Primer ID	Sequence (5' to 3')
F1	GCATACGAGATAGCTAGCCACC
R1_T0 unsorted	CTCTTTCCCTACACGACGCTCTTCCGATCTNNaCGGTTT CCAGCATAGCTCTTAAAC
R1_T2.5 unsorted	CTCTTTCCCTACACGACGCTCTTCCGATCTNNtaCGTGT TCCAGCATAGCTCTTAAAC
R1_T5 unsorted	CTCTTTCCCTACACGACGCTCTTCCGATCTNNctaGACG TTCCAGCATAGCTCTTAAAC
R1_T5 p16 unsorted	CTCTTTCCCTACACGACGCTCTTCCGATCTNNgctaACGT TTCCAGCATAGCTCTTAAAC
R1_T5 OVA unsorted	CTCTTTCCCTACACGACGCTCTTCCGATCTNNgTTCGTT CCAGCATAGCTCTTAAAC
R1_T2.5 p16 a mid	CTCTTTCCCTACACGACGCTCTTCCGATCTNNagCTGGT TCCAGCATAGCTCTTAAAC
R1_T2.5 p16 b mid	CTCTTTCCCTACACGACGCTCTTCCGATCTNNtagAATG TTCCAGCATAGCTCTTAAAC
R1_T2.5 OVA a mid	CTCTTTCCCTACACGACGCTCTTCCGATCTNNctagCAA GTTCCAGCATAGCTCTTAAAC
R1_T2.5 OVA b mid	CTCTTTCCCTACACGACGCTCTTCCGATCTNNcTTGTTT CCAGCATAGCTCTTAAAC
R1_T2.5 p16 a high	CTCTTTCCCTACACGACGCTCTTCCGATCTNNgcATAGT TCCAGCATAGCTCTTAAAC
R1_T2.5 p16 b high	CTCTTTCCCTACACGACGCTCTTCCGATCTNNagcACTC TTCCAGCATAGCTCTTAAAC
R1_T2.5 OVA a high	CTCTTTCCCTACACGACGCTCTTCCGATCTNNtagcGCA GTTCCAGCATAGCTCTTAAAC
R1_T2.5 OVA b high	CTCTTTCCCTACACGACGCTCTTCCGATCTNNtAGCTTT CCAGCATAGCTCTTAAAC

R1_T5 p16 a mid	CTCTTTCCCTACACGACGCTCTTCCGATCTNNctGGTCTTCCAGCATAGCTCTTAAAC
R1_T5 p16 b mid	CTCTTTCCCTACACGACGCTCTTCCGATCTNNgctACCGTTCCAGCATAGCTCTTAAAC
R1_T5 OVA a mid	CTCTTTCCCTACACGACGCTCTTCCGATCTNNagctGAGTTTCCAGCATAGCTCTTAAAC
R1_T5 OVA b mid	CTCTTTCCCTACACGACGCTCTTCCGATCTNNaTCTGTTCCAGCATAGCTCTTAAAC
R1_T5 p16 a high	CTCTTTCCCTACACGACGCTCTTCCGATCTNNtaTGAGTTCCAGCATAGCTCTTAAAC
R1_T5 p16 b high	CTCTTTCCCTACACGACGCTCTTCCGATCTNNctaCTTCTTCCAGCATAGCTCTTAAAC
R1_T5 OVA a high	CTCTTTCCCTACACGACGCTCTTCCGATCTNNgctaTAGGTTCCAGCATAGCTCTTAAAC
R1_T5 OVA b high	CTCTTTCCCTACACGACGCTCTTCCGATCTNNgGCTTTTCCAGCATAGCTCTTAAAC
R2	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCT
F2	CAAGCAGAAGACGGCATACGAGATAGCTAGCCACC

Appendix C Validation sgRNA sequences

Table 3: sgRNA sequences
sgRNAs were synthesiszed as single stranded oligonucleotides, annealed, cloned into pLentiV2-U6-sgRNA-mPGK-eBFP expression plasmid and used for gene knockout experiments. sgRNAs were picked based on the VBC score [91], or the sgRNA that enriched strongest in the screen as indicated.

sgRNA	Sequence
PSMB5	TTTGTACTGATACACCATGTTGG
COPS6	GGTGTGGGACAGCAGCTCAAAGG
TOR1A	AGCATGTGGAAAGTGCAATGTGG
TOR1B	GTCATTGTAGGACAGGTAGCCGG
TOR2A	GGTCCCGTATACCACCCAGGAGG
SEC62	GATTTGGCAAGAACCAAAAGTGG
SEC63	GATCACGATGAGCCCCACGAAGG
RFWD2	TATGACCCTTGAAGGAACGTAGG
THAP1	GCGGTTCTTGCAGCCGTAGGCGG
SIGMAR1	CGGACAGCGAGGCGTGCAGAAGG
ZNF207	GACAAAGGATACCAGCACCACGG
U2SURP	AGAAGAAATAGATCATCTGGTGG
OTUD5	AGAGATGTACAACCGTCCTGTGG
POLR3A	GTAGAGAGACACCTCATCGATGG
VCP	GATGAATTGCAGTTGTTCCGAGG
CKLF-CMTM1	CCAGCCGCAGCATCTTCACGTGG
AAVS1	GCTGTGCCCCGATGCACAC
COPS6_5 (2) (Screen)	GAGCTGCTGTCCCACACCG
TOR1A_10 (2) (Screen)	GCAATGTGGCCACAAACAGG