Univerzita Karlova

1. lékařská fakulta

Studijní program: Biomedicína

Studijní obor: Biochemie a patobiochemie

Mgr. Monika Sivá

Hledání biologické role rodiny proteinů podobných Ddi1

Deciphering the biological role of Ddi1-like protein family

Disertační práce

Vedoucí závěrečné práce/Školitel: Mgr. Klára Grantz Šašková, Ph.D.

Praha, 2019

Prohlášení:

Prohlašuji, že jsem závěrečnou práci zpracovala samostatně a že jsem řádně uvedla a citovala všechny použité prameny a literaturu. Současně prohlašuji, že práce nebyla využita k získání jiného nebo stejného titulu.

Souhlasím s trvalým uložením elektronické verze mé práce v databázi systému meziuniverzitního projektu Theses.cz za účelem soustavné kontroly podobnosti kvalifikačních prací.

V Praze, 1.3.2019

.....

Mgr. Monika Sivá

Identifikační záznam:

SIVÁ, Monika. Hledání biologické role rodiny proteinů podobných Ddi1. [Deciphering the biological role of Ddi1 like protein family]. Praha, 2018. 169 stran, 3 přílohy. Disertační práce. Univerzita Karlova, 1. lékařská fakulta, Přírodovědecká fakulta, Ústav organické chemie a biochemie AV ČR, v.v.i., Vedoucí závěrečné práce: Mgr. Klára Grantz Šašková, Ph.D.

Identification record:

SIVÁ, Monika. Deciphering the biological role of Ddi1 like protein family. [Hledání biologické role rodiny proteinů podobných Ddi1]. Prague, 2018. 169 pages, 3 appendices. PhD Thesis. Charles University, First Faculty of Medicine, Faculty of Science, Institute of Organic Chemistry and Biochemistry of the CAS. Supervisor: Mgr. Klára Grantz Šašková, Ph.D.

ABSTRACT

Ddi1-like protein family has been recently raised into the spotlight by the scientific community due to its important roles in cellular homeostasis maintenance. It represents a specific group among shuttling proteins of the ubiquitin-proteasome system. When compared to other shuttles, Ddi1-like protein family members harbor a unique retroviral-protease like domain besides the conventional ubiquitin-like (UBL) domain and domains interacting with ubiquitin. In addition, a helical domain of Ddi (HDD) has been recently found in most of the orthologs.

In this thesis, I focus on characterization of several members of Ddi1-like protein family, both on molecular level using NMR and in model mouse strains via a variety of biological methods.

Solution structure of the UBL domain of Ddi1p of *S. cerevisiae* was solved and its characteristics were compared to those of the UBL domain of its human ortholog. Furthermore, we show that human DDI2 specifically binds to ubiquitin with its terminal domains, both the UBL and the UIM; however, with very low affinity in contrast to binding properties of its yeast counterpart. Our study also show that hDDI2 does not form a head-to-tail homodimer. Based on our structural studies, we hypothesize that human DDI2 might have evolved a different function compared to its yeast ortholog. Next, we focused on Ddi1 protein analysis using model mouse strains. Our expression studies of the *Ddi1* homolog contributed to proposition of relevance of DDI1 in clinical research of Angelman syndrome. Moreover, our two mouse models exhibit embryonal lethality in mid-late gestation period, which suggest an essential role of the mammalian Ddi2 proteins.

Results acquired during this work shed light on possible functions of several Ddi1-like family members and will pave the way towards understanding of their compact biological roles.

Key words: ubiquitin-proteasome system, shuttling protein, Ddi1, aspartic protease, UBL, Nrf1 (NFE2L1), embryonic lethality, Angelman syndrome, mouse knockout model

ABSTRAKT

Představitelé rodiny proteinů podobných Ddi1 (z angl. DNA damage-inducible protein homolog 1) patří do skupiny tzv. „přenašečů“, které jsou v buňce zodpovědné za regulaci degradace proteinů v ubikvitin-proteasomálním systému. Proteiny podobné Ddi1 se od ostatních přenašečů liší specifickou doménou podobnou retrovirálním proteasam, která patří mezi aspartové proteasy. Nedávno byla v jejich struktuře taktéž objevena vysoce konservovaná helikální doména. V poslední době se objevují studie, které popisují nové funkce členů rodiny protein podobných Ddi1 a to v rámci udržování buněčné homeostázy v odpovědi na stresové podněty, např. v reakci na proteotoxické podmínky nebo v opravě poškozené DNA.

Tato disertační práce se zabývá charakterizací vybraných členů rodiny proteinů podobných Ddi1 a to jak na molekulární úrovni, tak v biologických studiích na myších modelech.

Byla vyřešena struktura domény podobné ubikvitinu (ubiquitin-like domain, UBL) v rámci kvasinkového proteinu Ddi1. Následně byly porovnány vazebné vlastnosti domén UBL kvasinkového proteinu Ddi1 a lidského proteinu DDI2. Doména UBL a motiv vázající ubikvitin (ubiquitin interacting motif, UIM) lidského homologu DDI2 sice specificky vážou ubikvitin, zároveň je však tato vazba velice slabá v porovnání s jinými ubikvitin vazebnými doménami. Naše studie taktéž vedly k objasnění celkové konformace homodimeru lidského DDI2. Na základě našich zjištění jsme formulovali hypotézu o odlišné funkci jednotlivých orthologů jinak vysoce konservovaných členů rodiny proteinů podobných Ddi1. Myší modely, ve kterých jsme inaktivovali expresi genu Ddi2, případně funkce proteasové domény, vykazují embryonální letalitu, což svědčí důležité biologické roli tohoto proteinu. Pomocí mapování exprese myšího genu Ddi1 jsme zároveň přispěly k objasnění relevance tohoto genu pro klinický výzkum Angelmanovho syndromu. Výsledky dosažené v této práci poukazují na některé vlastnosti a funkce členů rodiny proteinů podobných Ddi1, čímž představují pevný základ pro další studie, které by mohly vést k porozumění komplexní biologické role této proteinové rodiny.

Klíčová slova: retrovirální proteasa, Ddi1, UBL, Nrf1 (NFE2L1), embryonální letalita, Angelmanův syndrom, myší model

ACKNOWLEDGEMENTS

To my bosses and colleagues, without whom this would have been impossible...

I would like to express my gratitude to my supervisor Klara Grantz Šašková for giving me the opportunity to join her group few years ago, for introducing me into the Ddi2 world and for expressing a great deal of trust in my skills throughout my studies. I would like to thank for plenty of valuable advices and for proofreading of my thesis.

A great deal of gratefulness goes to Jan Konvalinka, for taking me in to his team 9 years ago and for his long-lasting support he has been giving me since, both professional and personal.

All colleagues from the Grantz Šašková, Konvalinka and Hlouchová labs created a pleasant atmosphere full of laughter, friendship, excellent science talks and great parties! ☺ I am blessed with the privilege to have worked with them.

Kudos to my guys from Quatro Happiness, for their support, advices, jokes, songs, experimental help and all the Gucci stuff! Jardič, Staníčkus and Miško, you were the best Don't Do It (Ddi) team I could have ever wished for. Special thanks goes to Miško for being my support in times of a PhD crisis, both our PhD studies were just easier when we went through it together.

Bimča and Karolína are special members of the Konvalinka lab that cannot be missing in the acknowledgement. They were here for me anytime I needed a hug, a kind word, a coffee, chocolate or a rum shot break and made the gloomy days look brighter. I would like to give thanks to Iva Flaisigová for all her help during my PhD studies. She always carried a smile and knew how to cheer me up, even though all of our experiment went wrong. Kudos to Jana Pokorná for the great coffee she took care of in our lab. Franta, thank you for all your advices, histology lectures, your empathy and for your saying that everything will be all right. Michael, thanks for helping me with formatting the damn Word document, when it kept going wrong! Additional thanks go to the newcomers in our lab throughout the years I had the chance to collaborate with: Jardič, Staníčkus, Zuzka, Kristie and Domi. By trying out to tutor them, I always got tutored on my own, which enabled me to get some healthy selfcriticism and grow both personaly and professionaly.

I would like to thank to our collaborators from the CCP and from the team of Dr. Radislav Sedláček, who contributed with enormous deal to this work on our mouse model projects. Special thanks goes to Honza and Míša Procházkoví, who have been my tutors, collaborators and advisors since they joined the Ddi2 mouse project. Kallayanee Chawengsaksophak, thank you for being such a patient tutor and for your advice on how to survive PhD studies.

Words of gratitude go to Dr. Vašek Veverka, who introduced me to the secrets of NMR during my undergraduate studies and has been a supportive advisor ever since.

This project was realized at following institutions:

FACULTY OF SCIENCE
Charles University

ÚOCHB AV ČR
IOCB PRAGUE

My family and friends, without whom this would have been impossible...

First of all, I would like to thank to my family for their love, support and patience they showed me throughout my studies. My beloved parents, my beloved brother and sister and their lovely growing families always managed to cheer me up (huge thanks go to my nieces and nephews!), recharge my batteries and they just made all the hard tasks look like a piece of cake.

A great deal of gratefulness goes to my dear Jakub, MPA, who coped with all the ups and downs I have been bringing home from work. During writing of the thesis, he was always coming at night to remind me my working hours are over and it is time to get some sleep. And! He has been a loving and lovely supporter and never-ending, never-complaining source of cheerful spirits, hugs, good food, chocolate, wine and great coffee!

Barunka Vorlová, thank you for being by my side for so many years, for being such a great friend, supporter and my source of inspiration. I would not have come so far without your friendship. Unfortunately, we became ex-colleagues and I will miss you and our coffees or just gossip breaks on daily basis. But, this also means that… wait for it… We finally made it!

I would like to thank to my dear friends, mainly Terezka, Pe3čka and Jeník, for their huge deal of support and to my beloved cousins and great friends Janka and Veronika, who have always been here for me. Prague just felt more like home with all of them, for which I am very grateful. I would like to especially depick Anežka for going through the finalization of our theses in such a good spirit.
Kudos to Tibor for his ability to make fun of everything. And for proofreading my acknowledgements. :-D

Big shiny THANKS go to all the friends that puzzled out that the killer question: "So when are you going to finish your PhD studies?" is not allowed to be said out aloud and on the contrary "No worries, everything will be ok, you can make it!" is very welcome. ;-)

Inspiring people, without whom this would have never happened...

I would like to acknowledge one of my professors from the Faculty of Science, who dragged me into the scientific world and kick-started my interest in science. Jan Černý, thank you for your own fascination by science and for your ability to “inoculate” your students with it.

I would especially like to acknowledge Buena Vista Social Club musicians, cast of Broadway musical The Lion King and Mercedes Sosa for their beautiful music that guided me throughout writing of this thesis.

Special thanks go to my dear grandma Terezka, who has always claimed her girls are strong and hardworking and they can achieve whatever they set their stubborn mind on. She has always longed for a doctor in her family (well, a real MD one…) and now she finally gets one!

My dear godmother Oľga, who not knowing, had inspired me the most. I think she is the upmost reason why I chose to become a scientist in the first place. To follow her steps.

Na záver by som chcela dodať, že laboratórne práce sa mi veľmi páčili.

- Jakub Nedoma, MPA

TABLE OF CONTENTS

ABBREVIATIONS

Abbreviaton	Meaning
AAA	ATPases Associated with diverse cellular Activities
AD	acidic domain
ADP	adenosine diphosphate
ADRM1	adhesion regulating molecule 1
AIDS	acquired immune deficiency syndrome
AMP	adenosine monophosphate
AP	alkaline phosphatase
AP-1	activator protein 1
APE1	apurinic/apyrimidinic endodeoxyribonuclease 1
APS	ammonium persulfate
ARE	antioxidant response element
ATF	activating transcription factor
ATM	ataxia telangiectasia mutated protein
ATP	adenosine triphosphate
ATR	ataxia telangiectasia and Rad3-related protein
BER	base excision repair
BHT	butylated hydroxytoluene
BiP	binding immunoglobulin protein
BRCA	breast cancer protein
BSA	bovine serum albumin
β-TrCP	β-transducin repeat-containing protein
bZIP	basic-leucine zipper
CAS	Czech Academy of Science
CCP	Czech Center for Phenogenomics
cDNA	complementary DNA
CENT2	centrin 2
CHOP	CCAAT-enhancer-binding protein homologous protein
CNC	cap-and-collar
CSP	chemical shift perturbation
CP	core particle
CREBP	cAMP response element binding protein
Cul1	cullin 1
DDR	DNA damage repair
DEPC	diethyl pyrocarbonate
Derlin	degradation in endoplasmic reticulum like protein
DEUBAD	deubiquitinase adaptor domain
DIG	digoxigenin
DMEM	Dulbecco's modified Eagle's medium
DMSO	dimethylsulfoxide
DNA	deoxyribonucleic acid
Doa10	degradation of alpha2-10
DPC	DNA-protein crosslink
DSB	double-strand break

DSF ..... differential scanning fluorimetry
DTT ..... dithiotreitol
DUB ..... deubiquitinase
DVC-1 ..... DNA Damage Protein Targeting VCP 1
E ..... embryonic day
EDEM1 ..... ER degradation enhancing alpha mannosidase-like protein 1
EDTA ..... ethylenediaminetetraacetic acid
EGTA ..... egtazic acid
eIF2α ..... eukaryotic translation initiation factor 2 subunit alpha
dNTP ..... deoxy-nucleotide triphosphate
ER ..... endoplasmic reticulum
ERAD ..... endoplasmic reticulum associated degradation
EUCOMM ..... European Conditional Mouse Mutagenesis
FBS ..... fetal bovine serum
Fbw7 ..... F-Box/WD repeat-containing protein 7
FDA ..... Food and Drug Administration
FPLC ..... fast protein liquid chromatography
GADD34 ..... growth arrest and DNA damage-inducible protein 34
GRP78 ..... 78-kDa glucose-regulated protein
GSH ..... glutathione
H2afz ..... H2A Histone Family Member Z
HDD ..... helical domain of Ddi
hDDI1 ..... human DNA damage-inducible protein 1 homolog 1
hDDI2 ..... human DNA damage-inducible protein 1 homolog 2
HDR ..... homology-directed repair
HECT ..... homologous to the E6AP carboxyl terminus
Herpud1 ..... homocysteine inducible ER protein with ubiquitin like domain 1
HIV-1 ..... human immunodeficiency virus 1
HOIP ..... HOIL-1-interacting protein
HPLC ..... high-performance liquid chromatography
hPLIC2 ..... human protein linking IAP with cytoskeleton
HR ..... homologous recombination
Hrd1 ..... β-hydroxy β-methylglutaryl-coenzyme A reductase degradation 1
HSQC ..... heteronuclear single quantum coherence
IKK ..... IκB kinase
IMDM ..... Iscove's modified Dulbecco's medium
IMG ..... Institute of Molecular Genetics
IMPC ..... International Mouse Phenotyping Consortium
IOCB ..... Institute of Organic Chemistry and Biochemistry
IpGTT ..... intra-peritoneal glucose tolerance test
IPTG ..... isopropyl β-D-1-thiogalactopyranoside
IRE1 ..... inostiol requiring enzyme 1
ISH ..... *in situ* hybridization
JAMM ..... Jab1/Mov34/Mpr1 Pad1 N-terminal+ domain-containing protease
Keap1 ..... kelch like, erythroid cell-derived protein with cnc homology associated protein 1
KO ..... knockout
LB ..... Luria-Bertani (lysogeny broth)

LCR-F1 ..... locus control region-factor 1
LIG ..... DNA ligase
Maf ..... avian musculoaponeurotic fibrosarcoma gene
MAG ..... 3-MethylAdenine DNA Glycosylase
MAPK ..... mitogen-activated protein kinase
MAT ..... mating type region
mDdi1 ..... murine DNA damage-inducible protein 1
mDdi2 ..... murine DNA damage-inducible protein 2
MEF ..... mouse embryonic fibroblast
MEM ..... minimum essential medium
MLH1 ..... MutL homolog 1
MMR ..... mismatch repair
MPN ..... Mpr1 Pad1 N-terminal+ domain
Mre11 ..... meiotic recombination protein 11
mRNA ..... messenger ribonucleic acid
mTORC1 ..... mammalian target of rapamycin complex 1
N/A ..... not available
NADPH ..... nicotinamide adenine dinucleotide phosphate (reduced)
Neh (domain) ..... Nrf2- erythroid, cell-derived protein with cnc homology, homology 2 domain
NER ..... nucleotide excision repair
NF-κB ..... nuclear factor kappa B
NFE2 ..... nuclear factor, erythroid 2
NFE2L ..... nuclear factor, erythroid-derived 2-related factor
NHB ..... N-terminal homology box
NHEJ ..... non-homologous DNA end joining
Ni-NTA ..... nickel - nitrilotriacetic acid
NMR ..... nuclear magnetic resonance
NOE ..... nuclear Overhauser effect
NOESY ..... nuclear Overhauser effect spectroscopy
Npl4 ..... nuclear protein localization homolog 4
Nrf1 ..... NF-E2-Related Factor 1
Nrf1 A ..... Nrf1 activated variant
Nrf1 FL ..... Nrf1 full-length variant
NST ..... Asp/Ser/Thr-rich region
NTD ..... N-terminal domain
OS9 ..... osteosarcoma amplified protein 9
PARP1 ..... poly(ADP-ribose) polymerase 1
PBS ..... phosphate buffered saline
PCNA ..... proliferating cell nuclear antigen
PCR ..... polymerase chain reaction
PD ..... protease defective
PDB ..... Protein Data Bank
PDS1 ..... precocious dissociation of sisters 1
PERK ..... protein kinase A-like ER kinase
PFA ..... paraformaldehyde
PNGase ..... peptide-N-glycanase
PICS ..... proteomic identification of protease cleavage sites

Psm ..... proteasomal subunit mammalian
PUB ..... PNGase/UBA domain
RAD..... RADiation sensitive
RBR ..... RING-betweenRING
REDAC ..... redundant dihedral angle constraints
RFC ..... replication factor C
RIN ..... RNA integrity number
RING ..... really interesting new gene
RMSD..... root mean square deviaiton
RNA..... ribonucleic acid
Rngo ..... rings lost
ROS ..... reactive oxygen species
RP ..... regulatory particle
RPA ..... replication protein A
Rpn ..... regulatory particle non-ATPase
Rpt ..... regulatory particle triphosphatase
RSC1A1..... regulatory solute carrier protein A1
RTF2..... replication termination factor 2
RUVBL ..... RuvB like AAA ATPase
RVP ..... retroviral protease (-like domain)
qPCR ..... quantitative polymerase chain reaction
SATB..... special AT-rich sequence binding protein
SAXS..... small-angle X-ray scattering
SCF ..... Skp1-Cul1-F-box protein complex
Sc UIM ..... scrambled ubiquitin interacting motif
SDS..... sodium dodecyl sulfate
SDSA..... synthesis-dependent strand annealing
SDS-PAGE..... sodium dodecyl sulfate – polyacrylamide gel electrophoresis
Sel1L ..... suppressor of lin-12-like protein 1
Sem1 ..... split hand/foot malformation (ectrodactyly) type 1
siRNA..... small interfering ribonucleic acid
Skn-1 ..... skinhead-1
Skp1..... S-phase kinase associated protein 1
Snc2..... synaptobrevin homolog 2
SNP..... single nucleotide polymorphism
SP1 and 2..... site-1 and site-2 proteases
SPRTN..... SprT-like N-terminal domain
SREBP1..... sterol regulatory element-binding protein 1
Sso1B..... supressor of Sec 1-binding
Sti1-like ..... stress-inducible phosphoprotein 1-like
TALEN..... Transcription Activator-Like Effector Nuclease
Tbp..... TATA-Box binding protein
TCF11..... transcription factor 11
TEMED ..... tetramethylethylendiamine
TEV ..... tobacco etch virus
TFIIH..... transcription initiation factor IIH
TLS..... translesion synthesis
TOCSY..... total correlated spectroscopy

TRAIP ..... tumor necrosis factor receptor associated factor Interacting protein
t-SNARE ..... target soluble N-ethylmaleimide-sensitive factor attachment protein receptor
UBA ..... ubiquitin-associated domain
Ube3a ..... ubiquitin protein ligase E3A
UBL..... ubiquitin-like domain
Ubp6..... ubiquitin-specific protease 6
UBQ ..... ubiquitin
UBX ..... ubiquitin regulatory X domain
Uch37 ..... ubiquitin C-terminal hydrolase 37
UCHL5..... ubiquitin C-terminal hydrolase L1
Ufd` ..... ubiquitin fusion degradation 1
UHMK1..... U2AF homology motif kinase 1
UIM..... ubiquitin-interacting motif
UPR..... unfolded protein response
UPS ..... ubiquitin-proteasome system
Usp14 ..... ubiquitin specific peptidase 14
UV..... ultraviolet
UVB ..... ultraviolet B
UV-DDB ..... UV radiation–DNA damage-binding protein
VBM..... VCP binding motif
VCP..... valosin containing protein
VIM..... VCP-interacting motif
VIMP..... VCP-interacting membrane protein
Vsm-1..... v-SNARE master protein 1
v-SNARE ..... vesicle soluble N-ethylmaleimide-sensitive factor attachment protein receptor
Wss1..... weak suppresor of *smt3*
WT..... wild-type
XBP1..... X-box binding protein 1
xCT..... XAP5 cysteine/glutamate transporter
XPA..... xeroderma pigmentosum group A protein
XPC..... xeroderma pigmentosum group C protein
XPD..... xeroderma pigmentosum group D protein
XTP3B..... HBV X-transactivated gene 3 protein-transactivated gene B protein
yDdi1..... yeast DNA damage-inducible protein homolog 1
YOD1 ..... yeast OTU deubiquinating enzyme 1

1 INTRODUCTION

1.1 THE CONCEPT OF HOMEOSTASIS – BRIEF HISTORY

One of the first theories on constancy maintenance and its importance in a biological system (human body), humorism (from Greek χυμός – chymós, translates to juice), dates back to ancient Egyptian or Mesopotamian medicine (Sertima I. V., 1992, Sudhoff K. and Garrison F. H., 1985). Humorism presents health of human body as a state, when the four bodily fluids: black bile, yellow bile, blood and phlegm, are in balance. Balance distortion of bodily liquids - dyscrasia - causes disease (mental or physical condition) (Jackson W. A., 2001). Holistic basis of this medical system interconnects mental and physical health and so the four body fluids additionally correspond to human temperaments: a choleric, a melancholic, a sanguinic and a phlegmatic(Jackson W. A., 2001).

The concept of internal balance of bodily fluids was not only used in ancient Egyptian or Mesopotamian medicine, however it most probably simultaneously set basis in Indian Ayurveda and traditional Chinese medical practices (Lutz P. L., 2002, Magner L. N., 2002, Sertima I. V., 1992, Sudhoff K. and Garrison F. H., 1985). Even though this theory had been used in ancient medicines previously, it was first described and systematized in ancient Greek collection of medical works called Hippocratic Corpus (from Latin: *Corpus Hippocraticum*) (Conrad L. I. N., Michael; Nutton, Vivian ; Porter, Roy; Wear, Andrew 1995). In spite of its name, this collection was created not by Hippocrates himself, but by many ancient Greek physicians and philosophers, most probably his students and followers. Hippocratic treatment was passive and moreover, as dissection of human body was permitted in ancient Grece, many deductions on diseases were based only on observations and incorrect. Next significant Greek physician, Galen of Pergamum (129 to 200 C. E.), not only formulated body disposition and human temper type interconnection with dominance of each of the bodily fluids, he described many anatomical observations based on necropsies he performed on animals (Schultz S. G., 2002). Since Hippocratic medicine and Galen´s observations spread widely to Roman, Persian and later most European cultures, they constituted the principia of medicine and deeply influenced scientific advancement from ancient Greece up to beginning of 19th century (Conrad L. I. N., Michael; Nutton, Vivian ; Porter, Roy; Wear, Andrew 1995).

First challenges to humorism occurred during Byzantine Empire with high influence of religion on medicine (Conrad L. I. N., Michael; Nutton, Vivian ; Porter, Roy; Wear, Andrew 1995). Galen´s study had become almost dogmatic and was not seriously challenged until 17th century. Then, Hippocratic medicine practices suffered more severe setback with arrival of renaissance and allowance of human body necropsies at universities for purpose of

medical and knowledge advancement. Anatomical discoveries of Andreas Vesalius published in 1543 in "De Humani Corporis Fabrica" (On the Fabric of the Human Body), where he identified errors in Galen´s treatises (Abbott A., 2015, Mesquita E. T. *et al.*, 2015), followed by William Harvey´s monograph "Exercitatio Anatomica de Motu Cordis et Sanguinis in Animalibus" (Anatomical Essay on the Motion of the Heart and Blood in Animals) published in 1628 (Schultz S. G., 2002), revolutionized physiology and medicine and contradicted Galenistic theories. Moreover, Harvey´s research set the basis of scientific methodology in systematic experimentation and computation opposed to previously performed observations. Complete discreditation of humorism occured in 1858 with publication "Die Cellularpathologie in ihrer Begründung auf physiologische und pathologische Gewebelehre" (Cellular Pathology) of German scientist Rudolf Ludwig Carl Virchow (Virchow R., 1858). Cell theory had already been established, however, Virchow´s significant contribution to the theory and promotion of microscopic insight into medicine and science designate him the founder of cellular pathology (Lin J. I., 1983). His observations completely opposed the theory of a balance of four bodily fluids determining human health. With sudden rapid advancement of cellular biology, microscopy and genetics since the middle of 19th century, humorism fell into shade and parted with modern science and conventional medicine (Conrad L. I. N., Michael; Nutton, Vivian ; Porter, Roy; Wear, Andrew 1995).

Even though humorism was left out of modern medical practices since, the perception of constancy in biological system, ironically basis of which was originally set in humorism, outlived even modern science. While French histologist Charles-Philippe Robin started to use phrase "milieu de l'intérieur" (the internal environment), which was comparable to the original Hippocratic balance between humors (Gross C. G., 2016), French physiologist Claude Bernard defined interstitial fluid in multicellular organisms as the "milieu intérieur". He described the ability of extra-cellular fluid to maintain healthy condition in tissues and organs of the whole body and defined it as disease protective element (Bernard C., 1865, Bernard C., 1949). He stated that internal bodily environment stability was basal condition for healthy life (Bernard C., 1974). This principle of internal constancy was later named "homeostasis" by American physiologist Walter Bradford Cannon in 1926 (Cannon W. B., 1926). The term comes from two Greek words: ὅμοιος (homoios), which means similar, and στάσις (stasis), which stands for standing still. Cannon defined homeostasis in 1929 as follows: "The coordinated physiological reactions which maintain most of the steady states in the body are so complex and are so peculiar to the living organism that it has been suggested

that a specific designation for these states be employed – homeostasis" (Cannon W. B., 1929). He built up his homeostasis regulation theory on Bernard´s scheme which emphasizes the need of constant values of material supply and environmental conditions. Cannon studied the flow of food and water intake, simultaneous excretion and maintenance of body temperature and blood pH and identified the importance of sympathetic pathway and hormonal secretion in regulation of all these processes (Cooper S. J., 2008). In 1956, Hungarian-Canadian endocrinologist Hans Selye introduced the concept of stress in physiology and biology (Goldstein D. S. and Kopin I. J., 2007, Selye H., 1956). He defined stress as a state "resulting in the nonspecific response of the body to any demand upon it" (Selye H., 1956). He actually expanded Cannon´s previously described theory of fight-or-flight response (Cannon, 1915). Modern concept of stress interprets it as a sensed threat to homeostasis with a specific response towards homeostasis restoration (McEwen B. S. and Stellar E., 1993).

Since beginning of 20th century, homeostasis maintenance has been considered the driving force of regulation of physiological and cellular processes towards ideal steady states to prevent pathological consequences of homeostasis disturbance.

1.1.1 From systemic to cellular homeostasis

Homeostasis has become one of the 8 core concepts of biology (Modell H. *et al.*, 2015). Its upkeep is key process at different levels in biological systems, from the complexity of an organism down to each organ, tissue compartment and individual cells. At all of these levels, the process of regulation of variables (internal environment conditions, e.g. temperature, pH, concentration of ions) towards stress response follows the same pattern (Buchman T. G., 2002, Goldstein D. S. and Kopin I. J., 2007). First, a sensor detects values of the variable and a homeostat compares it with the set point (a dynamic range of acceptable values) (Goldstein D. S., 1995). If the variable comes to be out of respectable range, control apparatus generates response and engages an effector. Effector changes the variable back towards steady state for homeostasis restoration (Goldstein D. S. and Kopin I. J., 2007). In general, generation of response can be performed in two ways: a positive feedback, which speeds up the stimuli process/change in condition or a negative feedback, which reverses the initiating impulse (Cooper S. J., 2008). To set an example, when the environment temperature and our core body temperature (variable) drops below acceptable range, thermosensors in skin and hypothalamus recognize this change. Skeletal muscles, as the effectors, act pro heat production causing shivering in a negative feedback loop (Modell H.*, et al.*, 2015).

Systemic homeostasis (at whole organism level) is ensured by autonomic nervous and endocrine system, which maintain variables such as core body temperature, osmolarity, blood pH, blood levels of ions, glucose, oxygen, etc. At tissue level, for example, the steady states of interstitial fluid volume, osmolarity and pH, cell positioning in tissue architecture, integrity of cell junctions are monitored (Chovatiya R. and Medzhitov R., 2014). In cell, variable changes in nutrient supply or protein folding and modifications (and many more processes – for few representative examples see Table 1) are detected by signaling proteins. If acceptable dynamic range of variables (nutrients, pH, all cellular processes) is violated, cell is stressed and engages stress response pathways towards cellular homeostasis restoration or in extreme cases towards apoptosis (Chovatiya R. and Medzhitov R., 2014). Apoptosis of cells might lead to protection of steady state in tissue environment (for example, regulation of cell number in tumorigenesis prevention) or homeostasis preservation at the level of whole organism (e.g. control of T-cell repertoire in autoimmunity prevention, apoptosis of infected host cells) (Giovannetti A. *et al.*, 2008, Hacker G., 2018).

Table 1: Representative sensor molecules in cellular homeostasis. Adapted from (Chovatiya R. and Medzhitov R., 2014)

Source of stress	Sensor molecule
ER stress	ATF6, IRE-1α, PERK
Genotoxic stress	p53
Heat shock	HSF-1
Hypoxia	HIF-1α
Oxidative stress	NFR2
Environmental stress	NF-κB, MAPK pathways
Amino acid deprivation	ATF4, mTOR

1.1.2 Cellular homeostasis

Since the definition of homeostasis by Walter Bradford Cannon in 1929, it was initially studied at organ or whole organism level (Cannon W. B., 1929). However, a revolution concerning cellular homeostasis started with description of a dynamic state of proteins, the never-ending synthesis followed by degradation over and over again, depending on internal and external environment (Simpson M. V., 1953). Practically, it is mostly proteins that control and regulate not only cell cycle and functions of the cell, but even the osmolarity of internal and external cellular fluids, ion concentration in cytoplasm, etc. Protein turnover involves a variety of cellular processes, that are crucial for proper function of cells, starting with DNA damage repair for preservation of genetic information, followed by regulation of gene expression, control of transcription and translation of proteins, their modifications, transportation in cell, activity control and eventual protein degradation (Calamini B. and

Morimoto R. I., 2012). Half-lives of proteins may differ in orders of magnitude even when present in one cell. Regulatory proteins mostly have lifetime of several minutes, however, there are also so-called long-lived proteins that evade degradation (Toyama B. H. and Hetzer M. W., 2013). Some proteins, such as myelin and nuclear pore complex (Rodrijguez de Lores A. *et al.*, 1971, Savas J. N. *et al.*, 2012), last for days. Some, for example collagen as an extracellular protein or crystalline in eye lens, live up to years or even decades (Masters P. M. *et al.*, 1977, Verzijl N. *et al.*, 2000).

Protein synthesis, from DNA via RNA to primary protein sequence, folding into a 3D structure and posttranslational modification of proteins are complex and variable processes, which can be impaired in a high number of ways. As native conformation of proteins is essential for sustaining of their biological function, protein homeostasis (proteostasis) is essential for health of the cell, or even whole organism from a broader perspective (Balch W. E. *et al.*, 2008, Calamini B. and Morimoto R. I., 2012). Defects in the synthesis pathway, such as mutations (or even polymorphism), error-prone protein synthesis, physical or chemical stress causing protein misfolding resulting in protein aggregation, may cause loss- or gain-of-function proteinopathies (e.g. cystic fibrosis or Alzheimer´s disease, respectively) (Cohen F. E. and Kelly J. W., 2003, Powers E. T. *et al.*, 2009). As our organism ages, the activity of stress response signaling pathways declines. Additionally, long-lived proteins that bypass degradation process and turnover are prone to accumulation of damage. That results in functional impairment and cellular aging (Toyama B. H. and Hetzer M. W., 2013). The overall gradual loss of proteostasis leads to susceptibility to chronic diseases, metabolic or neurological dysfunctions and cancer (Balch W. E*., et al.*, 2008, Kim Y. E. *et al.*, 2013).

1.2 PROTEIN HOMEOSTASIS MAINTENANCE

Proteostasis network consists of quality control mechanisms and stress response pathways that maintain stable and functional proteome. This is achieved by constant regulation and control of protein synthesis and degradation and via protein folding and trafficking (Balch W. E*., et al.*, 2008). At least around 1400 proteins are involved in proteostasis network in human cells, of which around 300 is a family of molecular chaperones and their regulators (Brehme M. *et al.*, 2014, Kim Y. E*., et al.*, 2013). The role of chaperones and folding enzymes in proteostasis network is to smooth energy barriers during acquisition of the native state of their substrates (Powers E. T*., et al.*, 2009).

There are several signaling pathways focusing on different stages and areas of proteostasis: DNA damage response, heat shock response, histone deacetylases system,

oxidative-stress response, inflammatory defense pathways, autophagic-lysosomal system, unfolded protein responses (UPR) in endoplasmic reticulum and mitochondria, Ca^{2+} cytoplasm-endoplasmic reticulum gradient regulatory pathway connected to endoplasmic reticulum associated degradation (ERAD) and the ubiquitin-proteasome system (UPS) (Galluzzi L. *et al.*, 2018, Powers E. T.*, et al.*, 2009, Walter P. and Ron D., 2011). Those relevant for our study are briefly introduced below.

1.2.1 DNA damage response (DDR) or DNA repair

DNA is constantly exposed to damaging agents or events of exogenous and endogenous origin. The exogenous causes can be, for example, xenobiotics or exposition to ultraviolet light, while chromatin remodeling, double-strand break (DSB) repair and redox homeostasis maintenance belong to main three endogenous damage-prone mechanisms (Turgeon M. O. *et al.*, 2018). Unsuccessful repair of genetic lesions may have severe effect on physiological tissue and systemic homeostasis (Galluzzi L.*, et al.*, 2018). DNA damage response is represented by a complex network of pathways operating differently, based on the origin of damage and mechanism of repair.

Base excision repair (BER) pathway focuses on DNA lesions that cause minor distortion to the double helix, such as oxidation, deamination or alkylation (Krokan H. E. and Bjoras M., 2013). The damaged base is removed by a DNA glycosylase (Lindahl T., 1974). Apurinic/apyrimidinic endodeoxyribonuclease 1 (APE1) recognizes the abasic site and recruits DNA polymerase β and DNA ligase 1 (LIG1) or LIG3 complexed with X-ray repair cross complementing protein 1 (XRCC1) to complete the repair process (Krokan H. E. and Bjoras M., 2013, Tell G. *et al.*, 2009). Another possibility is that poly(ADP-ribose) polymerase 1 (PARP1) denotes single-strand break site, which is recognized by APE1, and BER pathway is initiated (Durkacz B. W. *et al.*, 1980).

DNA adducts and bulky structures that cause double helix deformation are resolved by nucleotide excision repair (NER) pathway. The damage is sensed by a protein complex consisting of DNA damage recognition and repair factor (XPC), UV excision repair protein RAD23B and centrin 2 (CETN2), followed by recruitment of TFIIH (transcription initiation factor IIH), which bears a helicase subunit XPD for lesion verification (Volker M. *et al.*, 2001). In case of ultraviolet (UV) irradiation damage, first, UV–DDB (UV radiation–DNA damage-binding protein) complex is recruited to the site, followed by recognition of XPC complex (Wakasugi M. *et al.*, 2002). Once DNA is unwound, the undamaged strand is covered with replication protein A (RPA) for protection throughout the whole process (de

Laat W. L. *et al.*, 1998). XPA protein scans the damaged strand for nucleotides with altered chemical structures and engages structure-specific endonuclease complexes XPF–ERCC1 and XPG (encoded by ERCC5) into the repair process (Camenisch U. *et al.*, 2006, Fagbemi A. F. *et al.*, 2011). Finally, excised DNA gap is recognized by PCNA and replication factor C (RFC), which recruit synthesis and ligation enzymes such as DNA polymerase δ, ε or κ, and DNA ligase 1 or XRCC1– DNA ligase 3, depending on the state of the cell (Marteijn J. A. *et al.*, 2014, Ogi T. *et al.*, 2010).

Base mismatches and small deletions or insertions are restored via DNA mismatch repair (MMR) pathway. The site of damage is recognized by MutSα in case of single base damage (or MutSβ for larger insertions/deletions) that recruits endonuclease complex of MLH1 and PSM2 (Kunkel T. A. and Erie D. A., 2015, Schaetzlein S. *et al.*, 2013). PCNA sliding clamp then loads onto nascent DNA, activates MLH1, which incises the DNA in ATP-dependent manner (Modrich P., 2006). Mismatch is then removed by exonuclease 1 downstream of the recruitment protein complex and the cleaved DNA strand is resynthesized by polymerases and connected by ligases (Genschel J. *et al.*, 2002, Schaetzlein S.*, et al.*, 2013).

Double-strand breaks of genomic DNA are repaired via two mechanisms, non-homologous DNA end joining (NHEJ) and homology-directed repair (HDR). NHEJ is a more common pathway, which is able to restore chromosomal structure after DNA breakage, however at the same time causes deletion or insertion of a few nucleotides at one of the DNA ends, or both (Lieber M. R., 2010). Canonical NHEJ has a first and main actor, Ku protein that binds to broken DNA ends and recruits DNA-dependent protein kinase (Walker J. R. *et al.*, 2001). Other enzymes, such as nucleases (e.g. artemis), polymerases (e.g. polymerases μ and λ) and ligases (e.g. DNA ligase IV) of this pathway, are assembled to this DNA:protein complex, according to the requirements of the breakage repair mechanism (Lee J. W. *et al.*, 2004, Ma Y. *et al.*, 2002, Nick McElhinny S. A. and Ramsden D. A., 2003, Nick McElhinny S. A. *et al.*, 2000). Alternative NHEJ is a "backup" pathway that is able to mediate end joining depending on microhomology (less than approx. 10 bp) (Wang H. *et al.*, 2005). In contrast, HDR system is error-free. However, it requires a homology donor that is not present in diploid cells outside S and G2 phase of cell cycle (Lieber M. R., 2010, Vitale I. *et al.*, 2017). HDR has three forms of mechanism, single-strand annealing, breakage-induced replication and most frequent homologous recombination (HR), which requires the longest homology sequence (San Filippo J. *et al.*, 2008). Activation machinery of homologous recombination at site of DSB initiates PARP1 localization and DNA poly(ADP-ribosyl)ation at the damage

site (Van Meter M. *et al.*, 2016). Then, a protein complex consisting of Mre11, Rad50 and Nbs1 (called the MRN complex in humans) is recruited to the site of damage (Haince J. F. *et al.*, 2008). Resection of the 5′-end follows and RPA protein binds to the free 3′ overhang (szostak 1983). Rad51 recombinase recognizes the RPA coated DNA single strand and with help of other assisting proteins (e.g. BRCA1, BRCA2) searches for homology sequences at the sister chromatid (Prakash R. *et al.*, 2015, Vitale I.*, et al.*, 2017). Repair process can be performed in two ways, the double Holliday junction model or the synthesis-dependent strand annealing (SDSA) pathway (Vitale I.*, et al.*, 2017).

Translesion synthesis (TLS) is a process, which avoids DSB creation or replication fork collapse, when replication is stalled due to DNA damage (e.g. in case of thymine dimers) (Sale J. E., 2013). Proliferating cell nuclear antigen (PCNA) is ubiquitinated by a heterodimer of ubiquitin-conjugating enzyme E2A or B with E3 ubiquitin-protein ligase RAD18 (Bailly V. *et al.*, 1994). This modification attracts polymerases from Y family to the DNA damage site (Sale J. E., 2012). These polymerases are able to facilitate the required insertion at the damage site, which common replication polymerases are not able to perform (Waters L. S. *et al.*, 2009). These polymerases are error-prone, which is a high risk of mutagenesis, however, TLS is smoother process than other more challenging DNA damage responses, which could lead to cellular death.

Another DNA damage response, which deals with DNA-protein crosslinks (DPCs) that escaped NER pathway, has been recently described (Stingele J. *et al.*, 2015). DPCs, which stall replication fork and could lead to genome instability, are formed either when enzymes, e.g. topoisomerases, are covalently trapped in the otherwise transient DNA-protein intermediate or upon exposure to crosslinking agents, such as UV radiation or formaldehyde (Barker S. *et al.*, 2005, Pommier Y. *et al.*, 2006). In yeast models, metalloprotease Weak suppressor of *smt3* (Wss1), is activated by interaction with Cdc48 (VCP/p97 in higher eukaryotes) and with DNA and cleaves off the bulky protein body of DPC leaving only a DNA-bound peptide behind (Balakirev M. Y. *et al.*, 2015, Stingele J. *et al.*, 2014). This leads to ssDNA accumulation, attraction of PCNA and TLS polymerases that produce mutations, but promote replication restart (Stingele J.*, et al.*, 2014). In higher eukaryotes, DPCs are removed from nascent DNA strands by two mechanisms: proteolytic cleavage of the protein part by similar mechanism as in yeast by a homolog of Wss1, SPRTN metalloprotease, or by ubiquitination of the trapped protein by E3 ubiquitin ligase TRAIP and subsequent degradation in proteasome (Larsen N. B. *et al.*, 2018, Lopez-Mosqueda J. *et al.*, 2016, Morocz M. *et al.*, 2017, Stingele J. *et al.*, 2016).

There are two main actors common for DNA repair pathways adjacent to DSBs or ssDNA damage - ataxia telangiectasia mutated (ATM) serine/threonine kinase and ataxia telangiectasia and Rad3-related protein (ATR) (Marechal A. and Zou L., 2013). ATM and ATR phosphorylate H2A histone family member X, which leads to activation of check point kinases 1 and 2. This results either in cell cycle arrest until the repair process is finished or in apoptosis via regulation of p53 protein in case the damage cannot be resolved (Vitale I*., et al.*, 2017).

1.2.2 Oxidative stress response

Elevation of oxidation causing elements with toxic effect (oxidative stressors), such as reactive oxygen species (ROS), reactive nitrogen species, reactive sulfur species, reactive selenium species and reactive carbonyl species, leads to imbalance in redox homeostasis which might have pathogenic implication (Sies H. *et al.*, 2017). Most abundant group is ROS (covers superoxide, hydroxyl and peroxide radicals), which arise from disruption of mitochondrial electron transport chain or activity of several enzymes, e.g. NADPH oxidase (Vaquero E. C. *et al.*, 2004). Basal production of ROS is convenient as they act as second messengers in a variety of cellular signaling pathways. However, higher levels of ROS are detrimental as they cause damage to DNA, lipids and proteins (Martin K. R. and Barrett J. C., 2002, McCord J. M., 1995, Rhee S. G., 1999, Rhee S. G., 2006, Sauer H. *et al.*, 2001). Excessive generation of ROS has been reported to result in activation of transcription factors, such as NF-κB, mitogen-activated protein kinases (MAPK), activator protein 1 (AP-1) or factors activating apoptosis (Chen Q. *et al.*, 1995, Jacobson M. D., 1996, Pahl H. L. and Baeuerle P. A., 1994, Schreck R. *et al.*, 1991).

Essential regulatory genes that are activated upon oxidative stress are nuclear factor erythroid-derived 2-related factor 1 and 2, NFE2L1/Nrf1 and NFE2L2/Nrf2 (Chan J. Y. *et al.*, 1993b, Moi P. *et al.*, 1994, Venugopal R. and Jaiswal A. K., 1998). They belong to the cap-and-collar (CNC) subfamily of basic-leucine zipper (bZIP) transcription factors together with p45 NF-E2, Nrf3, Bach1 and Bach2 genes in mice and humans. Members of this group were originally described as regulators of beta-globin gene activation bearing a 43-residue long homology region prior DNA-binding domain (Andrews N. C. *et al.*, 1993, Biswas M. and Chan J. Y., 2010, Schultz M. A. *et al.*, 2010). Nrf1 and Nrf2 of vertebrate form functional heterodimers with either member of cAMP response element binding protein family (CREBP), ATF4 or small Maf proteins, e.g. MafG, MafF or MafK, which regulate expression of oxidative stress related genes via binding to the antioxidant response elements (AREs)

upstream of their sequence (Johnsen O. *et al.*, 1998, Johnsen O. *et al.*, 1996, Kaspar J. W. *et al.*, 2009, Motohashi H. *et al.*, 2004). Production of prooxidants is compensated by antioxidant gene expression (by Nrf1 or Nrf2) that are able to maintain redox homeostasis, such as *glutathione peroxidase 1*, *cytochromes P-450*, *peroxiredoxin-1*, *thioredoxin-1*, *superoxide dismutase*, *heme oxygenase-1*, *NADPH-quinone oxidoreductase* or enzymes involved in biosynthesis of the main antioxidant - glutathione (Biswas M. and Chan J. Y., 2010, Kim Y. J. *et al.*, 2007, Myhrstad M. C. *et al.*, 2001, Nioi P. *et al.*, 2003, Osburn W. O. and Kensler T. W., 2008, Wu G. *et al.*, 2004). ARE (also called electrophile response sequence) is a cis-active sequence 5′-TGACXXXGC-3′ in the target gene promotor region (Friling R. S. *et al.*, 1992, Rushmore T. H. *et al.*, 1991, Telakowski-Hopkins C. A. *et al.*, 1988).

Nrf2 is considered the main regulator of response to oxidative stress called the phase II response (Dinkova-Kostova A. T. *et al.*, 2005). ARE sequences of genes regulated by Nrf2 occur in promotors of around 1055 genes involved in not only oxidative stress, but as well in DNA repair, detoxification, cellular proliferation, signaling and immune response (Dodson M. *et al.*, 2019, Silva-Islas C. A. and Maldonado P. D., 2018). The 605 residues of Nrf2 form a seven domain structure and provide modification and interaction sites important for regulation of Nrf2 activity. Neh1 domain in the C-terminal part is the DNA binding and Maf interaction site (Itoh K. *et al.*, 1999, Li W. *et al.*, 2008a). Under basal conditions, Nrf2 is present in an inactive form bound to cytoplasmic chaperone Keap1 via its N-terminal Neh2 domain. Keap1 anchors Nrf2 to cytoskeleton and at the same time bridges it to Cul3-based E3-ubiquitin ligase for ubiquitination. Ubiquitinated Nrf2 is detached from Keap1 by p97-UBNX7-UFD1/NPL4 complex and subsequently degraded in proteasome – the half-life of Nrf2 is approx. 13 min (Cullinan S. B. *et al.*, 2004, Itoh K.*, et al.*, 1999, Stewart D. *et al.*, 2003, Tao S. *et al.*, 2017, Tong K. I. *et al.*, 2006). Under oxidative stress, Nrf2 can be activated by two mechanisms: canonical and non-canonical (Silva-Islas C. A. and Maldonado P. D., 2018). In the first mechanism, the lysines of Nrf2 are not available for ubiquitination and subsequent proteasomal degradation due to conformational change of the interacting Keap1, which arises from oxidation of Keap1 Cys residues by oxidizing compounds. In this case, Keap1 remains bound to Nrf2 and is unable to fish out newly synthesized Nrf2 molecules that escape proteasome and are subsequently activated (Baird L. *et al.*, 2013, Itoh K. *et al.*, 2003). In the non-canonical mechanism, Keap1 detaches Nrf2, which results in its translocation into nucleus and activation of antioxidant genes, instead of its proteasomal degradation (Holtzclaw W. D. *et al.*, 2004, Kobayashi A. *et al.*, 2004). A set of genes (e.g., p21, BRCA1) that are able

to bind either Keap1 or Nrf2, and thereby disrupt their interaction are responsible for this activation (Silva-Islas C. A. and Maldonado P. D., 2018). Nrf2 molecules that escape degradation are phosphorylated and acetylated prior to their translocation into nucleus via importin α5 and importin β1 (Huang H. C. *et al.*, 2002, Joo M. S. *et al.*, 2016, Kawai Y. *et al.*, 2011, Theodore M. *et al.*, 2008). A variety of additional modifications of Nrf2 are required for its binding to ARE sequences (Silva-Islas C. A. and Maldonado P. D., 2018). The pool of activated Nrf2 molecules in nucleus is regulated by E3 ubiquitin ligase Skp1-Cul1-F-box protein complex (SCF) together with β-transducin repeat-containing protein (β-TrCP). (β-TrCP) works as an adaptor, which binds to sequence DpSGX(1-4)pS phosphorylated on serine residues, which is situated in Neh6 domain (Chowdhry S. *et al.*, 2013, Rada P. *et al.*, 2011, Rada P. *et al.*, 2012, Skowyra D. *et al.*, 1997). This motif is conserved also in Nrf1 and Nrf3 proteins (Tsuchiya Y. *et al.*, 2011). This brings Nrf2 to proximity with a RING E3 ligase Skp1, which ubiquitinates Nrf2 and thereby targets it for degradation.

The role of Nrf2 in stress response leads to protection of cells from damage and death. However, Nrf2 may also be misused in cancer cells in reaction to radiotherapy or chemotherapeutics for activation of response pathways that result in resistance to apoptosis or proliferation (Wang X. J. *et al.*, 2006). Both these factors make Nrf2 an attractive therapeutic target from different points of view (Rojo de la Vega M. *et al.*, 2018). In spite of that, only one FDA approved compound that activates Nrf2, an electrophilic dimethyl fumarate, is used for treatment of relapsing multiple sclerosis (Satoh T. and Lipton S., 2017).

Structure of Nrf1 is similar to Nrf2 to a certain extent, as they contain conserved sequences in the DNA binding region, a β-TrCP binding motif and Neh2 domain. Nrf1 Neh2 is able to bind Keap1 in spite of the fact that Keap1 does not participate in the degradation nor activation cascade of Nrf1 (Kim H. M. *et al.*, 2016, Wang W. and Chan J. Y., 2006, Zhang Y. *et al.*, 2006). Life cycle of Nrf1 differs from that of Nrf2 even though it is based on similar principle (described in chapter 1.3.2.1.1 on page 54). Some target genes of Nrf1 and Nrf2 are regulated by both transcription factors; however, some are unique to the individuals (Kim H. M.*, et al.*, 2016). Ohtsuji and colleagues described distinct roles of Nrf1 and Nrf2 in regulation of ARE regulated genes on a number of mouse knockout models. Some antioxidant genes are responsive to both Nrf1 and Nrf2, however, some, e.g. metalothionein-1 and -2, are exclusive to activation by Nrf1 (Ohtsuji M. *et al.*, 2008). Nrf1 activity is essential for proper hepatocyte function as was shown in full and liver specific knockout mouse models (Chen L. *et al.*, 2003, Xu Z. *et al.*, 2005). Nrf2 knockout mice show failure in adequate antioxidant gene induction in response to oxidative stress (Chanas S. A. *et al.*, 2002, Itoh K. *et al.*, 1997). Strikingly, Nrf1

knockout mice die during embryonal development due to anemia before they reach embryonic day 13.5 (Chan J. Y. *et al.*, 1998). As Nrf1/Nrf2 double knockout mice die earlier, at embryonic day 11.5, Nrf2 seems to (at least partially) provide compensation for Nrf1 function (Leung L. *et al.*, 2003, Ohtsuji M.*, et al.*, 2008). Nrf2 differs from Nrf1 in several ways: Nrf2 is not glycosylated, it activates a different subset of stress response genes, and even though it does regulate the expression of proteasomal subunits, it is upon oxidative stress conditions and not upon proteasomal inhibition opposed to Nrf1 (Kwak M. K. *et al.*, 2003, Radhakrishnan S. K. *et al.*, 2010, Steffen J. *et al.*, 2010). It was therefore proposed that Nrf1 counteracts constitutive oxidative stress opposed to Nrf2 that shall respond to severe stress states (Ohtsuji M.*, et al.*, 2008). As the role of Nrf1 reaches beyond oxidative stress response and is closely related to this study, its degradation, activation and other functions are described in detail in chapter 1.3.2.1.1 on page 54.

1.2.3 Proteostasis maintenance pathways in ER

1.2.3.1 Unfolded protein response

Unfolded protein response pathway in ER, the center of protein secretion and surface display in cell, is main regulatory strategy, with which cells cope with the burden of misfolded protein accumulation in the ER lumen. Only in some specific cell types (e.g. immune system cells), the threshold for UPR activation is elevated in behalf of their function in systemic homeostasis (Frakes A. E. and Dillin A., 2017). UPR operates towards expansion of ER protein-folding capacity via three unique signal transduction mechanisms. These are each defined by ER-resident stress sensors: inositol requiring enzyme 1(IRE1), activating transcription factor 6 (ATF6) and protein kinase RNA (PKR)-like ER kinase (PERK) (Frakes A. E. and Dillin A., 2017, Walter P. and Ron D., 2011). All three UPR key components are transmembrane proteins that harbor luminal and cytosolic domains essential for their function. Luminal domains represent the sensor of folding homeostasis: they are bound to a chaperone binding immunoglobulin protein (BiP) and therefore inactive in stress-free environment. When the capacity of ER folding system is overwhelmed, misfolded proteins attract BiP, which frees the luminal domain of the sensor and activates it (Bertolotti A. *et al.*, 2000, Frakes A. E. and Dillin A., 2017, Shen J. *et al.*, 2005). Moreover, UPR sensors can be activated by direct interaction with misfolded proteins (Walter P. and Ron D., 2011). Cytosolic domains provide response to activation via regulation of transcription or translation pathways (Frakes A. E. and Dillin A., 2017).

The most conserved and best-studied branch of UPR starts with activation of the bifunctional kinase/endoribonuclease IRE1. Upon activation, IRE1 undergoes autophosphorylation and subsequent oligomerization, which activates its ribonuclease domain. IRE1 targets X-box binding proteins 1 (XBP1) and performs cleavage leading to formation of a unique XBP1 splice variant. Truncated XBP1 isoform regulates expression of a variety of proteins necessary for folding assistance and lipid synthesis for membrane expansion (Calfon M. *et al.*, 2002, Frakes A. E. and Dillin A., 2017, Lee A. H. *et al.*, 2003). Unspliced version of XBP1 is translated and provides a feedback loop regulation by binding to the spliced form of XBP1 hence targeting it to proteasome (Yoshida H. *et al.*, 2006).

Another branch is controlled by ATF6. Upon activation, ATF6 is translocated into Golgi apparatus in vesicles, where its luminal domain and transmembrane anchor are cleaved off by site-1 and site-2 proteases (SP1 and SP2) (Haze K. *et al.*, 1999, Schindler A. J. and Schekman R., 2009). The cytosolic N-terminal domain of ATF6 is translocated into the nucleus and induces expression of XBP1, chaperones (Hsp70 and Hsp90 family proteins) and genes involved in ERAD pathway (Hetz C. and Papa F. R., 2018, Lee A. H.*, et al.*, 2003).

PERK kinase activation is the third possibility of UPR. When the luminal domain of PERK kinase is unbound, the protein undergoes autophosphorylation and homomultimerization. Cytosolic kinase domain then phosphorylates the α subunit of eukaryotic translation initiation factor 2 (eIF2α), thereby inhibiting guanine nucleotide exchange factor (eIF2B) and mRNA translation *in toto* (Harding H. P. *et al.*, 1999). There are few important exceptions, for example the activating transcription factor 4 (ATF4), which is preferably translated under these conditions. ATF4 then targets genes such as pro-apoptotic transcription factor CHOP and GADD34 (growth arrest and DNA damage-inducible protein 34) (Tsaytler P. *et al.*, 2011). GADD34 encodes a protein phosphatase subunit PP1C, which dephosphorylates eIF2α and therefore autoregulates the activity of PERK kinase (marciniak).

1.2.3.2 ERAD pathway

ERAD is a quality control pathway that maintains ER homeostasis by retrotranslocation of terminally misfolded proteins into cytosol for subsequent degradation in cytosolic proteasomes (Qi L. *et al.*, 2017). During retrotranslocation, substrates are ubiquitinated by a membrane bound E3 ligase (for more information see chapter 1.2.4.3) and pulled out of ER in an ATP hydrolysis driven process (for ERAD pathway scheme see Figure 1) (Ruggiano A. *et al.*, 2014).

Substrates for ERAD are recognized and recruited to the retrotranslocation complex at luminal site of ER membrane either after specific deglycosylation (trimming of mannose) by mannosidases or by ER resident chaperones, such as binding immunoglobulin protein/78-kDa glucose-regulated protein (BiP/GRP78), osteosarcoma amplified protein 9 (OS9), XTP3-transactivayed gene B protein (XTP3B) or ER degradation enhancing alpha mannosidase-like protein 1 (EDEM1) (Bhamidipati A. *et al.*, 2005, Christianson J. C. *et al.*, 2008, Cormier J. H. *et al.*, 2009, Kim W. *et al.*, 2005, Molinari M. *et al.*, 2003, Oda Y. *et al.*, 2003, Plemper R. K. *et al.*, 1997, Thibault G. and Ng D. T., 2012). Retrotranslocation channel is built by ER transmembrane domains of membrane-embedded E3 ubiquitin ligase complexes. Yeast Hrd1p and Doa10p E3 ligases are the most characterized complexes (Bordallo J. *et al.*, 1998, Swanson R. *et al.*, 2001), which are specializing in degradation of misfolded proteins: Hrd1p recognized and designates proteins with lesions in luminal, membrane or translocon-associated domains and Doa10p ubiquitinates proteins with errors in cytosolic and membrane structures (Ruggiano A.*, et al.*, 2014, Thibault G. and Ng D. T., 2012). Mammals have several ERAD E3 ligases, Hrd1, Gp78, Rma1/Rnf5, Trc8, Rfp2, Rnf170 and Rnf185 (Claessen J. H. *et al.*, 2012). The most conserved retrotranslocation complex in mammals is represented by Hrd1 E3 ubiquitin ligase and its cofactor Sel1L (Qi L.*, et al.*, 2017). Hrd1 harbors six transmembrane domains that form the dislocation channel and a RING finger domain with a C-terminal proline-rich region projecting into the cytoplasm. RING domain ubiquitinates not only misfolded substrates, but in addition, the dislocation channel is shaded in a regulatory process when RING domain autoubiquitinates Hrd1 itself (Baldridge R. D. and Rapoport T. A., 2016, Carvalho P. *et al.*, 2010, Stein A. *et al.*, 2014). Second partner of the dislocation complex is Sel1L, an integral ER membrane protein that possesses an N-terminal ubiquitin associated domain (UBA) that binds ubiquitin and an evolutionary conserved C-terminal ubiquitin regulatory X domain (UBX) that facilitates the recruitment of ATPase into proximity with the dislocation complex and substrate (Buchberger A., 2002, Buchberger A. *et al.*, 2001, Hofmann K. and Bucher P., 1996, Schuberth C. and Buchberger A., 2005). There are other ER transmembrane proteins, such as Derlins and Herpud1, which bind to the Hrd1/Sel1L complex and assist the retrotranslocation process (Greenblatt E. J. *et al.*, 2011, Kokame K. *et al.*, 2000). Moreover, it has been shown recently that gp78 E3 ligase might second the Hrd1/Sel1L complex in solubilization of the substrates for further processing (Zhang T. *et al.*, 2015).

The act of substrate ubiquitination initiates enrollment of the valosin-containing protein (VCP)/p97 (Cdc48 in yeast), a member of type II AAA+ (ATPases associated with

diverse cellular activities) family (Meyer H. *et al.*, 2012). Apart from Hrd1/Sel1L complex interaction via a C-terminal VCP binding motif (VBM), other ER transmembrane proteins are also able to bind and recruit p97, e.g., the SHP domains in Derlin-1 and Derlin-2 and a VCP-interacting motif (VIM) in VIMP (Christensen L. C. *et al.*, 2012, Greenblatt E. J.*, et al.*, 2011, Lilley B. N. and Ploegh H. L., 2004, Neuber O. *et al.*, 2005, Ye Y. *et al.*, 2004). VCP uses ATP hydrolysis as a source of energy for pulling force by which it is able to extract modified substrates from organelle membranes (e.g., ER, mitochondria) or chromatin and target them to proteasome (Franz A. *et al.*, 2016, Heo J. M. *et al.*, 2010, Qi L.*, et al.*, 2017). The ATPase monomer consists of an N-terminal domain, which serves as an interaction site for ubiquitin and cofactor molecules, two conserved ATPase domains D1 and D2 that form a hexameric ring in a barrel-shape like structure around a central pore and a C-terminal 76 amino acid long tail available for cofactor interactions (Davies J. M. *et al.*, 2008, DeLaBarre B. and Brunger A. T., 2003, Meyer H. H. *et al.*, 2000, Peters J. M. *et al.*, 1992). ATP hydrolysis by D2 domain triggers a conformational change of the ring and thereby facilitate disassembly of molecular complexes (Pye V. E. *et al.*, 2006). Substrate release from p97 is thought to be driven by ATP hydrolysis in the D1 domain (Bodnar N. O. and Rapoport T. A., 2017).

Components of ERAD pathway possess a variety of p97 interaction domains, e.g., VIM, 8 residue interacting motif called SHP box or UBX (Buchberger A.*, et al.*, 2001, Hitt R. and Wolf D. H., 2004, Sato B. K. and Hampton R. Y., 2006). Thanks to these domains, more than 30 cofactors of p97 in ERAD pathway are able to recruit p97 to dislocation complex or coordinate correct orientation of p97 and the substrate (Ye Y. *et al.*, 2017). The N-terminal domain of p97 can bind modified substrate thanks to ubiquitin interaction sites held by p97 cofactors, for example nuclear protein localization homolog 4 (Npl4) and ubiquitin fusion degradation 1 (Ufd1) heterodimer (Meyer H. H.*, et al.*, 2000, Stein A.*, et al.*, 2014, Ye Y. *et al.*, 2003). The p97-Ufd1-Npl4 complex recognizes chains longer than a pentaubiquitin, preferably highly branched heterotypic chains (Blythe E. E. *et al.*, 2017, Bodnar N. O. and Rapoport T. A., 2017). In fact, it is the combinations of cofactors assembled at p97 protein that determine the pathway p97 will engage (van den Boom J. and Meyer H., 2018). The C-terminal tail of p97 is able to interact with other cofactors of ERAD pathway, like peptide-N-glycanase (PNGase) that harbors a PNGase/UBA domain (PUB), E3 ubiquitin ligase HOIP or substrate recruiters Ufd2 and Ufd3 (Bohm S. *et al.*, 2011, Schaeffer V. *et al.*, 2014, Zhao G. *et al.*, 2007). PNGase removes N-glycans from misfolded proteins that were retrotranslocated from ER into the cytosol by cleavage of the β-aspartyl-glucosamine bond

(Blom D. *et al.*, 2004, Tarentino A. L. and Plummer T. H., Jr., 1994). Phosphorylation of p97 prevents interaction with PNGase and causes accumulation of ubiquitinated substrates (Li G. *et al.*, 2008). p97 as well interacts with deubiquitinases such as YOD1 or Ataxin-3 that add another level of ERAD regulation into the pathway (Papadopoulos C. *et al.*, 2017, van den Boom J. and Meyer H., 2018). Adaptor proteins such as Rad23 and Dsk2 play an important role in substrate delivery from p97 and DUB assembly to the proteasome. Their function is further described in chapter 1.2.4.5 on page 46 (Doss-Pepe E. W. *et al.*, 2003, Li G. *et al.*, 2006).

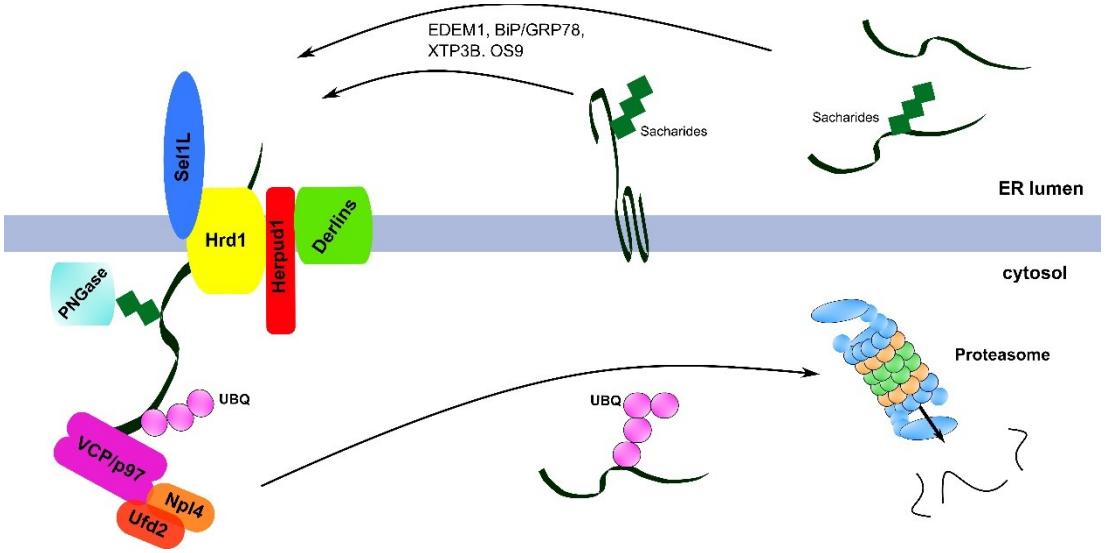


Figure 1: Scheme of endoplasmic reticulum associated pathway (ERAD) in mammals. Both glycosylated and non-glycosylated misfolded proteins are delivered to the ER transmembrane proteins that form a retrotranslocon complex – Hrd1/Sel1L, Derlins, Herpud1. Six transmembrane domains of Hrd1/Sel1L create a transmembrane channel for substrates that are ubiquitinated at the cytosolic site by the RING E3 ligase domain of Hrd1. Modified substrates are recognized by p97-Ufd1-Nlp1 complex, which provides further interaction with deubiquitinases. These modify the polyubiquitin chain for recognition by intrinsic or extrinsic proteasomal ubiquitin receptors or completely remove the ubiquitin moieties for substrate escape from the ERAD pathway.

1.2.4 Ubiquitin-proteasome system

Ubiquitin-proteasome system is the main machinery for degradation of short-lived cytosolic and nuclear proteins (Rubinsztein D. C., 2006). The core component of the pathway is a multisubunit protease complex called proteasome, which represents the main protease of the cell from the AAA+ family. Substrates destined for degradation are covalently modified with ubiquitin or polyubiquitin chains and therefore targeted to proteasome, where the ubiquitin tag is recognized, cut off, and the substrate is subsequently unfolded in ATP hydrolysis driven process and further cleaved into short peptides (Bard J. A. M. *et al.*, 2018, Ciechanover A. *et al.*, 1978, Ciechanover A. *et al.*, 1980, Etlinger J. D. and Goldberg A. L., 1977).

1.2.4.1 Ubiquitin

Ubiquitin belongs to one of the main regulatory tools in cellular homeostasis. It is a highly conserved 8.5 kDa protein consisting of 76 residues. Ubiquitin folds into a globule that consists of five antiparallel β-sheets, a long α-helix a short 3_{10}-helix (see Figure 2A) (Vijay-Kumar S. *et al.*, 1987). Main binding site on ubiquitin, which is important for its interaction with proteasome, is the Ile44 hydrophobic patch, represented by Leu8, Ile44, His68 and Val70 (Beal R. *et al.*, 1996). Another hydrophobic surface, which is important for interubiquitin interaction along polyubiquitin chains or for interaction with deubiquitinases, is situated around Ile36 and includes residues Leu71 and Leu73 from C-terminal tail (Hu M. *et al.*, 2002). Third patch that has been described by now, interacting with specific deubiquitinases, is localized around Phe4 and involves Gln2 and Thr12 as well (Hu M.*, et al.*, 2002, Sloper-Mould K. E. *et al.*, 2001). All three hydrophobic patches are shown in Figure 2B.

C-terminal glycine residue of ubiquitin can form a covalent isopeptide bond with lysine of a target protein. This post-translational modification, ubiquitination or ubiquitinylation, can activate/silence the activity of a protein, in case of transcription factors it can regulate expression of key molecules in their downstream pathways, or for histones alter expression profile of the cell, or contrarily aim the target towards degradation in the proteasome (Ciechanover A.*, et al.*, 1980, Goldknopf I. L. *et al.*, 1977, Pickart C. M., 2000). Monoubiquitination has been described to have regulatory function in DNA repair processes and chromosome remodeling (Hoege C. *et al.*, 2002, Robzyk K. *et al.*, 2000, Ulrich H. D. and Walden H., 2010). Moreover, ubiquitin possesses seven lysines, which are essential for its function as they provide bonding sites for ubiquitin linkage between the C-terminal glycine of distal ubiquitin and the individual lysines (K8, K11, K27, K29, K33, K48 and K63) of the proximal ubiquitin (these residues are highlighted in Figure 2A). This possibility gives rise to different polyubiquitin chains that can harbor one or more types of linkages or even can be branched (Swatek K. N. and Komander D., 2016). Most common chain type is K48-linked polyubiquitin, which is the main protein degradation signal in the UPS (Swatek K. N. and Komander D., 2016). Another quite abundant type is K63-linkage with degradation free roles, for example in inflammatory signaling, where they function as a scaffold for recruitment of IKK complex and activation of NF-κB pathway (Chen Z. J. and Sun L. J., 2009, Kanayama A. *et al.*, 2004). The role of K63 chains has been described as well in intracellular trafficking and DNA damage response (Komander D. and Rape M., 2012, Swatek K. N. and Komander D., 2016). K11-linked chains function as proteasomal degradation signal specifically in cell

cycle regulatory pathway during mitosis when they are attached to the substrates by anaphase promoting complex/cyclosome (Bremm A. and Komander D., 2011, Wickliffe K. E. *et al.*, 2011). The K29-linked chains have been proposed to be regulators of protein degradation as they modify proteasomal ubiquitin receptor Rpn13 (see chapter 1.2.4.2, page 40) and autoimmune inflammatory response (Besche H. C. *et al.*, 2014, Jin J. *et al.*, 2016). K33-linked chains have role in protein trafficking and their recruitment to trans-Golgi network (Yuan W. C. *et al.*, 2014). K6-linked chains belong to the less understood, however their function has been proposed in response to genotoxic stress and mitophagy in several studies (Durcan T. M. *et al.*, 2014, Morris J. R. and Solomon E., 2004, Wu-Baer F. *et al.*, 2003). Chains attached to K27 are probably the least characterized type, however there have been several studies recently that suggest their role as a scaffold for recruitment of the proteins of DNA repair process after double strand breaks and additional regulatory role in protein secretion (Gatti M. *et al.*, 2015, Palicharla V. R. and Maddika S., 2015). Additional polyubiquitin chain type exists with linkage via interaction with the N-terminal methionine (M1-linkage), which is a positive regulator of NF-κB signaling and has important role in inflammatory and immune responses (Boisson B. *et al.*, 2012, Kirisako T. *et al.*, 2006, Rahighi S. *et al.*, 2009).

1.2.4.2 Proteasome

Eukaryotic 26S proteasome is a proteolytic complex with molecular mass of 2.5 MDa, which consists of a central 20S barrel-shaped core particle (CP) and a 19S regulatory particle (RP) capping one or both ends of the barrel (see Figure 2C). The regulatory particle is responsible for recognition, deubiquitination and translocation of substrates into the degradation chamber of 20S core (Baumeister W. *et al.*, 1988, Glickman M. H. *et al.*, 1998a, Seemuller E. *et al.*, 1995, Zwickl P. *et al.*, 1999).

The 670 kDa 20S core particle is composed of four heptameric rings with α and β subunits forming the two outer and two inner circles, respectively (see Figure 2C and D) (Groll M. *et al.*, 1997, Seemuller E.*, et al.*, 1995). Alpha rings interact with the base of 19S RP. Highly conserved N-terminal domains of α subunits create a narrow entrance (approx. 13 Å) into the CP even in open state to avoid degradation of properly folded proteins (Groll M. *et al.*, 2000). Proteolytic chamber is assembled from β subunits in their precursor form (Zwickl P. *et al.*, 1994). Three of the β subunits show catalytic activity depending on the type of residue they cut after: post-acidic or caspase-like (or post-glutamyl peptide hydrolase) β1, post-basic or trypsin-like β2 and post-hydrophobic or chymotrypsin-like β5 subunit (Groll M. *et al.*, 1999, Heinemeyer W. *et al.*, 1997). They are activated by proteolytic removal of their

N-terminal propeptides. This leads to exposure of a threonine residue at their N-terminus, which serves as a nucleophile during hydrolysis (Brannigan J. A. *et al.*, 1995, Chen P. and Hochstrasser M., 1996, Dahlmann B. *et al.*, 1992, Lowe J. *et al.*, 1995, Seemuller E. *et al.*, 1996).

The approx. 900 kDa 19S regulatory particle is formed by two subcomplexes, a base and a lid (see Figure 2D) (Glickman M. H. *et al.*, 1998b). The nomenclature of all the RP subunits will be adducted here with its name in *Saccharomyces cerevisiae* and with its human gene name and protein name in brackets. The base of RP consists of three non-ATPase subunits, Rpn1 (PSMD2, S2), Rpn2 (PSMD1, S1) and Rpn13 (ADRM1) (Glickman M. H.*, et al.*, 1998b). Rpn1 and Rpn2 are receptors for ubiquitin and ubiquitin-like protein domains (UBLs) thanks to their α-solenoid domains. These are composed of 11 proteasome/cyclosome repeats, which are 35-40 residues long helix-turn-helix hairpins forming toroid-like structures that provide interaction surface of both RP subunits (He J. *et al.*, 2012, Kajava A. V., 2002, Lupas A. *et al.*, 1997). Rpn1 preferentially binds K6- and K48-linked ubiquitin chains with its toroid 1 site as described by Shi and colleagues (Shi Y. *et al.*, 2016). This is the same site with which Rpn1 interacts with UBLs of UPS adaptor proteins such as Rad23 and Dsk2(for interaction site of Rpn1 with Rad23 UBL domain see Figure 2E) (Chen X. *et al.*, 2016, Elsasser S. *et al.*, 2002, Saeki Y. *et al.*, 2002a, Shi Y.*, et al.*, 2016). Rpn2 interacts with the core particle, Rpn1, Rpn13 and two ATPase subunits Rpt4 and Rpt6 (Rosenzweig R. *et al.*, 2008, Rosenzweig R. *et al.*, 2012). Rpn13 harbors a pleckstrin-like receptor for ubiquitin (Pru) domain at its N-terminus, which interacts with ubiquitin (preferentially K48-linked chains) and UBLs (very strong interaction with the UBL of hPLIC2) (Chen X.*, et al.*, 2016, Husnjak K. *et al.*, 2008, Schreiner P. *et al.*, 2008). Six ATPase subunits, Rpt1 (PSMC2, S7), Rpt2 (PSMC1, S1), Rpt3 (PSMC4, S6), Rpt4 (PSMC6, S10), Rpt5 (PSMC3, S6a) and Rpt6 (PSMC5, S8), form a heterohexameric ring in the center of the base (Glickman M. H.*, et al.*, 1998b). N-terminal alpha helices of Rpt subunits create coiled coil structures between the Rpt1/Rpt2, Rpt63/Rpt6 and Rpt4/Rpt5 dimers during assembly (Inobe T. and Genmei R., 2015, Tomko R. J., Jr. *et al.*, 2010). The oligonucleotide/oligosaccharide binding fold domains of Rpt subunits create an N-ring above the AAA+ domain ring positioned over the gate of the CP. The motor pulling force in the unfolding process is performed by highly conserved loops of the AAA+ domains, which protrude into the center of the hexameric circle (Erales J. *et al.*, 2012, Maillard R. A. *et al.*, 2011, Martin A. *et al.*, 2008). There is another non-ATPase ubiquitin receptor creating a bridge between the base and the lid of assembled

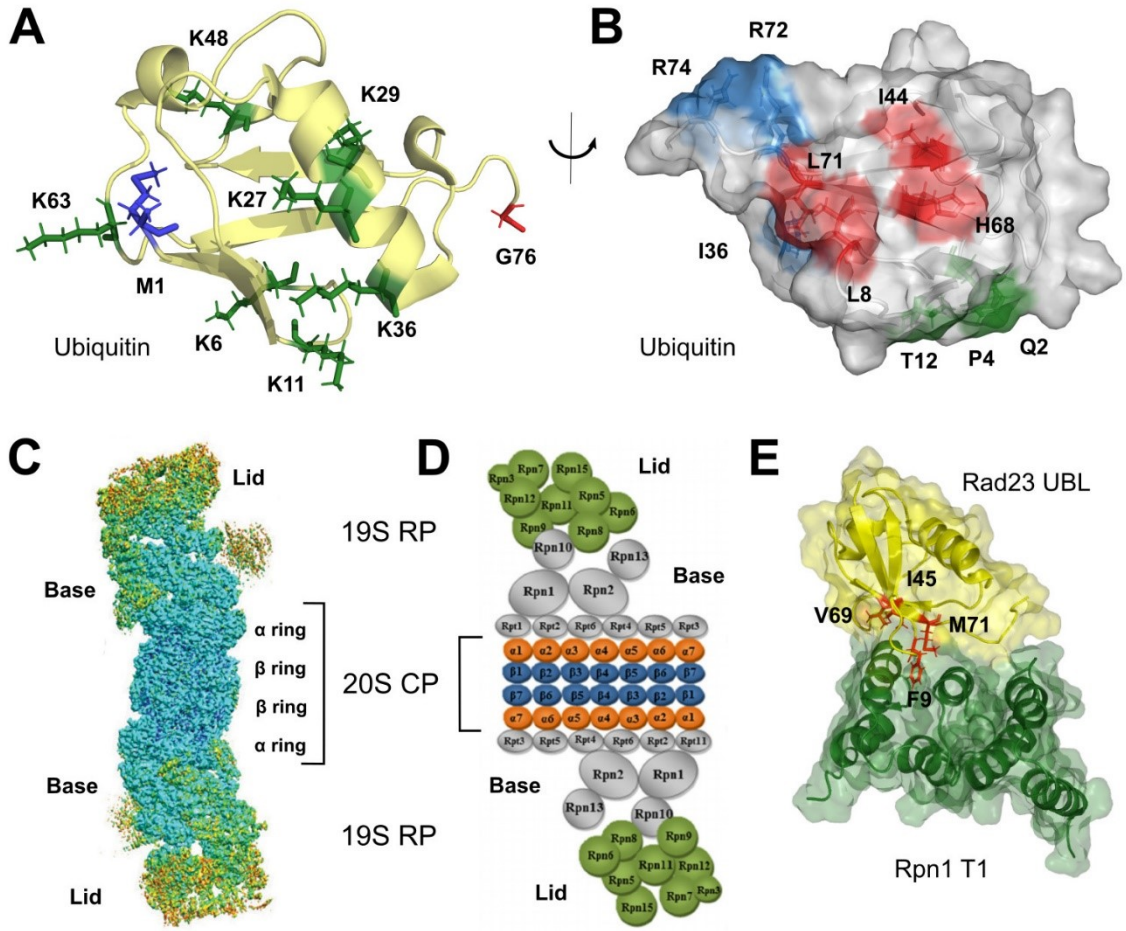


Figure 2: Structures of key components of ubiquitin-proteasome system. A) Structure of ubiquitin (cartoon representation): Ubiquitin harbors 7 lysine residues (in green) and an N-terminal methionine (blue) that can form specific linkages in polyubiquitin chains via bondage with C-terminal glycine (red). **B)** Surface representation of ubiquitin structure: Ubiquitin has a specific fold that provides three hydrophobic interaction patches: Phe4 (green), Ile36 (blue) and Ile44 (red). Side chains of significant interaction residues of the patches are shown. The figure was created in program PyMOL with PDB entry 1D3Z (Cornilescu G. *et al.*, 1998, Schrodinger, LLC, 2015). **C)** Cryo-EM map of human 26S proteasome at resolution of 3.9 Å (Huang X. et al., 2016). **D)** Scheme of 26S proteasome from S. cerevisiae and position of individual subunits of the whole complex. Adapted from (Diaz-Villanueva J. F. et al., 2015). **E)** Close-up on the interaction of the UBL domain of Rad23 (yellow) with the T1 domain of Rpn1 (green) from S. cerevisiae. Similarly to the Ile44 patch of ubiquitin, Rad23 UBL possesses a hydrophobic interaction patch with adequate interaction residues P9, I45, V69 and M71 (red) (PDB entry 2NBW) (Chen X., et al., 2016).

RP, Rpn10 (PSMD4, S5a) (Deveraux Q. *et al.*, 1994). Rpn10 harbors an N-terminal von Willebrand factor type A domain that is responsible for interaction with RP subunits (Rpn1 and Rpn2) and C-terminal ubiquitin interaction motifs (UIMs), that provide binding surface for ubiquitin and UBLs (Erales J.*, et al.*, 2012, Sakata E. *et al.*, 2012, Verma R. *et al.*, 2004). UIM is an amphipatic helix consisting of 20 amino acids in a defined sequence, which docks into the hydrophobic patch on ubiquitin with affinity ranging from 0.1 to 2 mM (Fisher R. D. *et al.*, 2003, Hofmann K. and Falquet L., 2001, Young P. *et al.*, 1998). The lid of RP consists of six proteasome-CSN-initiation factor 3 domain containing subunits Rpn3 (PSMD3, S3), Rpn5 (PSMD12), Rpn6 (PSMD11, S9), Rpn7 (PSMD6, S10), Rpn9 (PSMD13, S11) and

Rpn12 (PSMD8, S14) (Lander G. C. *et al.*, 2012). Additionally, there are two Mpr1-Pad1 N-terminal domain containing subunits Rpn8 (PSMD7, S12) and Rpn11 (PSMD14, Poh1, Pad1) (Lander G. C.*, et al.*, 2012, Rinaldi T. *et al.*, 1998). There is another subunit, Sem1 (PSMD9, Dss1, Rpn15), which functions as a ubiquitin receptor for K63- and K48-linked chains thanks to its intrinsically disordered region (Paraskevopoulos K. *et al.*, 2014).

1.2.4.2.1 Proteasome inhibitors

Proteasome as a key component of regulatory protein degradation became a target for drug development, however most of the compounds are limited to laboratory use because of their non-specificity or poor metabolic stability (Adams J., 2003). The structure and specificity of inhibitors varies quite significantly. There are synthetic peptide aldehydes available, such as tripeptide aldehyde carbobenzoxy-Leu-Leu-leucinal called MG132 that targets several types of proteases such as serine proteases or calpain, in addition to reversible proteasome inhibition (Saito Y. *et al.*, 1990, Tsubuki S. *et al.*, 1996, Tsubuki S. *et al.*, 1993). Lactacystin is a metabolite of *Streptomyces*, which undergoes hydrolysis in cells and its intermediate product β-lactone is responsible for covalent modification of the catalytic threonine residues in mammalian 20S core particles (Dick L. R. *et al.*, 1996, Fenteany G. *et al.*, 1995). The α^1,β^1-epoxy-ketone tetrapeptide epoxomicin isolated from *Actinomycetes* irreversibly inhibits all three types of proteolytic activity of the β subunits (Hanada M. *et al.*, 1992, Meng L. *et al.*, 1999). The most successful proteasome inhibitor that made it into clinical use is bortezomib, a water-soluble reversible proteasome selective inhibitor of the chymotrypsin-like activity with K_i = 0.6 nM that belongs to a family of dipeptidyl boronic acids (Adams J. *et al.*, 1999). It was introduced by Millenium Pharmaceuticals, Inc. under the name Velcade and used in therapeutic approaches in treatment of patients suffering from multiple myeloma and mantel cell lymphoma (Field-Smith A. *et al.*, 2006). In addition, an analog of epoxomicin invented in Proteolix, Inc., carfilzomib, is another FDA approved anti-cancer drug used in relapsed or refractory multiple myeloma after previous treatment with bortezomib (Vij R. *et al.*, 2012).

1.2.4.3 Ubiquitination process

Substrates of the UPS pathway shall possess two main recognition elements: a covalently attached polyubiquitin chain and an unstructured initiation region either at their terminus (20-30 residues) or represented by an internal flexible loop (Prakash S. *et al.*, 2004, Takeuchi J. *et al.*, 2007).

In ubiquitination, the ubiquitin moiety must be first activated to be recognized and further utilized by individual enzymes of the cascade. This first step is performed by one predominant E1 ubiquitin-activating enzyme, which binds ATP and Mg^{2+} and catalyzes adenylation of the C-terminal glycine of ubiquitin. Adenylated ubiquitin is then attacked by the catalytic Cys of E1, which results in formation of macroergic thioester bond between Cys sulfhydryl and Gly carboxyl in the activated E1-ubiquitin complex (see Figure 3A). This is directly followed by adenylation of another ubiquitin molecule at the adenylation site. E1 reactions are driven by dephosphorylation process of ATP into inorganic phosphate and AMP and are reversible (Ciechanover A. *et al.*, 1981, Ciechanover A. *et al.*, 1982, Haas A. L. and Rose I. A., 1982, Haas A. L. *et al.*, 1983, Haas A. L. *et al.*, 1982, Pickart C. M. and Rose I. A., 1985). The ubiquitin from Cys on E1 is transferred onto active Cys on recruited E2 ubiquitin-conjugating enzyme (see Figure 3B) (Pickart C. M. and Rose I. A., 1985). Another enzyme, an E3 ubiquitin ligase, then associates with the E2-ubiquitin complex. There are several families of E3s that based on their domain composition operate in different ways. In general, there are two possible ways: E3 either binds the ubiquitin from E2 with its cysteine residue prior to transfer onto the substrate (e.g. HECT E3s, RBR E3s; see Figure 3D) or it works as a scaffold for bringing the E2-ubiquitin complex and substrate together into proximity and right conformation to facilitate the transfer of ubiquitin from E2 directly onto the substrate (e.g. RING E3s; see Figure 3C) (Huang L. *et al.*, 1999, Wenzel D. M. *et al.*, 2011, Yokouchi M. *et al.*, 1999). The ε-amino group of a lysine residue of the substrate attacks the thioester of the associated charged E2 or E3 and creates an isopeptide bond between ubiquitin and substrate (Scheffner M. *et al.*, 1995). After bond formation, E2 is discharged and leaves E3 enzyme. Dissociation of E2 can be followed by association of the E3 with another charged E2 in second round of ubiquitination performed either on another lysine residue of the substrate or on the ubiquitin towards chain formation (Schulman B. A. and Harper J. W., 2009). As there are altogether two E1s, around 38 E2s and over 800 E3s encoded in human genome, it seems that it is the E3s that decide selection of the substrate (Jin J. *et al.*, 2007, Li W. *et al.*, 2008b, Ye Y. and Rape M., 2009, Zheng N. and Shabek N., 2017).

The requirements for the initiation of unstructured region are strict, as the polypeptide must be long enough to reach the pore in the AAA+ ring, which is 30-40 Å away from the entrance in the RP of proteasome (Aufderheide A. *et al.*, 2015, Bard J. A. M.*, et al.*, 2018). Substrates that do not have the intrinsic initiation sequence are unfolded by a Cdc48 (VCP, p97) AAA+ unfoldase complex prior to proteasomal degradation, which is considered as

pre-processing of the substrates prior to proteasomal degradation (Bodnar N. O. and Rapoport T. A., 2017, Ye Y., *et al.*, 2017).

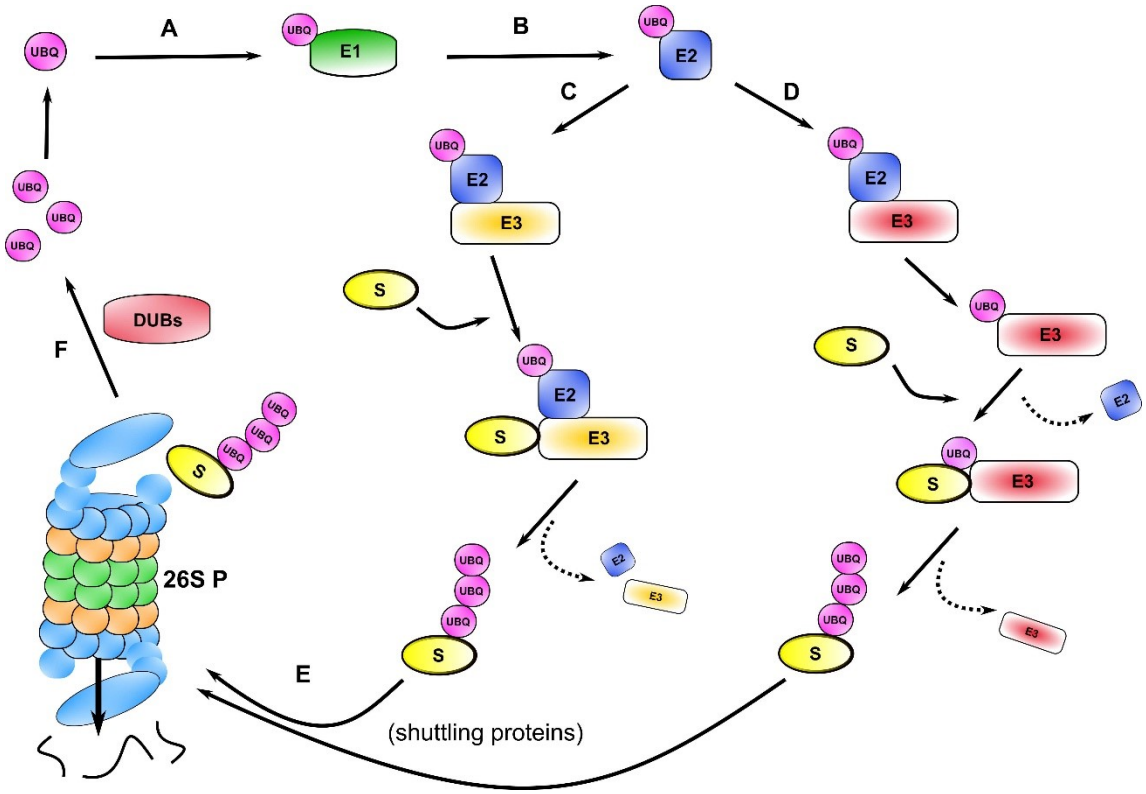


Figure 3: Scheme of substrate modification by ubiquitin. Collaboration of E1 (**A**), E2 (**B**) and E3 (**C** and **D**) enzymes of the UPS cascade results in ubiquitination of the substrate, which targets substrate protein towards degradation (**E**). After recognition of polyubiquitin chains by intrinsic or extrinsic ubiquitin receptors and removal of the ubiquitin moieties by DUBs (**F**), the substrate is ready to enter degradation chamber in the core particle of 26S proteasome where it is cleaved into short peptides. Adapted from (Weissman A. M. et al., 2011).

1.2.4.4 Proteasome associated deubiquitinases

There are two and three described deubiquitinases (DUBs) associated with yeast and mammalian proteasome, respectively. These DUBs cleave off the ubiquitin chain of the substrate prior its entrance into the AAA+ ring and unfolding (see Figure 3F). Rpn11 is Zn^{2+}-dependent metalloprotease that belongs to the family of JAMM/MPN deubiquitinase family. According to electron microscopy studies, it is situated directly above the N ring in the RP base, which is suitable for its function as it hydrolyzes the bond between the substrate and the first ubiquitin of the chain (Aufderheide A., *et al.*, 2015, Lander G. C., *et al.*, 2012, Verma R. *et al.*, 2002, Yao T. and Cohen R. E., 2002). Rpn11 forms a heterodimer with Rpn8, which changes its conformation and moves its Insert-1 loop region from an autoinhibitory closed state into active β-hairpin conformation, which allows its DUB activity (Worden E. J. *et al.*, 2014). Yeast and mammalian proteasomes both have DUB Ubp6 (Usp14), which interacts with Rpn1 subunit and cleaves supernumerary ubiquitin chains en bloc (Aufderheide

A., *et al.*, 2015, Hanna J. *et al.*, 2006, Lee B. H. *et al.*, 2016). Ubp6 cuts K48-linked long chains in case the substrate is tagged with more than one polyubiquitin chain (Lee B. H., *et al.*, 2016, Mansour W. *et al.*, 2015). Activation of Ubp6 is dependent on the interaction of its UBL domain with toroid 2 interaction site of Rpn1 and interaction of its catalytic domain with the N-ring and the AAA+ ring, which together performs conformation change that removes inhibitory loops out of the active site (Aufderheide A., *et al.*, 2015, Leggett D. S. *et al.*, 2002, Shi Y., *et al.*, 2016). Mammalian proteasomes additionally interact via their Rpn13 subunit with a ubiquitin C-terminal hydrolase Uch37 (alternatively named UCHL5) which removes distal K48-, K6- and K11-linked ubiquitin chains (Hamazaki J. *et al.*, 2006, Lam Y. A. *et al.*, 1997, Qiu X. B. *et al.*, 2006, Yao T. *et al.*, 2006). Uch37 interacts with the deubiquitinase adaptor (DEUBAD) domain of Rpn13 bound to proteasome, which moves a loop that covers the active cysteine residue and amplifies the affinity for ubiquitin up to fivefold (Vander Linden R. T. *et al.*, 2015).

1.2.4.5 Extrinsic ubiquitin receptors

In addition to the intrinsic ubiquitin receptors, there is a family of UBL/UBA-containing proteins called shuttling proteins or adaptor proteins of UPS, represented by Rad23 (hHR23A and hHR23B) and Dsk2 (hPLIC1, UBIQUILIN-1) (Bertolaet B. L. *et al.*, 2001). Rad23 interacts with Rad4 or XPC, which stabilizes them and enables them to recognize of DNA lesions in nucleotide excision repair pathway (Dantuma N. P. *et al.*, 2009, Guzder S. N. *et al.*, 1998, Jansen L. E. *et al.*, 1998, Masutani C. *et al.*, 1994). Rad23 also functions as a shuttling protein for proteasomal substrates in ERAD, downstream of VCP/p97 complex (Richly H. *et al.*, 2005). In addition it was reported to be involved in cell cycle regulation, spindle body formation and phosphate metabolism in budding yeast (Auesukaree C. *et al.*, 2008, Biggins S. *et al.*, 1996, Clarke D. J. *et al.*, 2001). hPLIC1 has significant role in clearance of aggregated proteins by their targeting to autophagosomes and aggresome formation, and in neuroprotection on whole organism level in conditions such as Alzheimer´s disease or ALS (Deng H. X. *et al.*, 2011, Lu K. *et al.*, 2017, Stieren E. S. *et al.*, 2011).

These proteins harbor N-terminal UBL domain that provides interaction with proteasomal receptors and a C-terminal ubiquitin associated domain (UBA) that binds polyubiquitin chains of substrates targeted for degradation (Chen L. *et al.*, 2001, Elsasser S., *et al.*, 2002, Funakoshi M. *et al.*, 2002, Schauber C. *et al.*, 1998, Verma R., *et al.*, 2004, Wilkinson C. R. *et al.*, 2001). Their UBL domains are built of a long α-helix, a 3_{10}-helix and

five antiparallel β-sheets, which form a hydrophobic patch responsible for their interaction with proteasomal subunits (Walters K. J. *et al.*, 2002). UBA domains consist of three helices that form a bundle, its helices 1 and 3 interact with the ubiquitin Ile44 hydrophobic patch or UBL domains (Mueller T. D. and Feigon J., 2002, Mueller T. D. *et al.*, 2004, Raasi S. *et al.*, 2004). Highly helical domain of Sti1-like family (or Rad4/XPC-binding domain in Rad homologs), which is responsible for interaction with DNA binding proteins during DNA damage response, was found in both above-mentioned representatives (Kaye F. J. *et al.*, 2000, Masutani C. *et al.*, 1997). For scheme of the domain architecture see Figure 4.

Another member of this family is DNA damage-inducible protein 1 (Ddi1), which will be further described in detail in following chapters (Bertolaet B. L.*, et al.*, 2001, Clarke D. J.*, et al.*, 2001).

1.3 DDI1-LIKE PROTEIN FAMILY

Ddi1-like protein family members possess special domain architecture, which dedicates them to a variety of biological roles. Because of their conserved N-terminal UBL and C-terminal UBA domains and interaction with polyubiquitinated proteins and proteasomal subunits, they were first suggested to perform the role of shuttling factors of UPS (Kottemann M. C. *et al.*, 2018, Morawe T. *et al.*, 2011, Saeki Y.*, et al.*, 2002a). When compared to other shuttling proteins, Ddi1-like family harbors additional retroviral protease-like domain (RVP) in the center of the protein. The fold of Ddi1-like protein RVPs is very similar to HIV-1 protease, it represents an aspartic protease with catalytic triad DT/SG, facilitates dimerization of full-length proteins and is highly conserved in all eukaryotes (Gabriely G. *et al.*, 2008, Krylov D. M. and Koonin E. V., 2001, Kumar S. and Suguna K., 2018, Sirkis R. *et al.*, 2006, Siva M. *et al.*, 2016, Trempe J. F. *et al.*, 2016). Recently, a novel helical domain of Ddi (HDD) was structurally characterized in the linker region between UBL and RVP domains in Ddi1p of *S. cerevisiae* and human DDI2 protein (Siva M.*, et al.*, 2016, Trempe J. F.*, et al.*, 2016). For domain architecture see Figure 4.

Loss of UBA domain occurred early during evolution in vertebrates, as the RSC1A1 gene was inserted into the DDI1 locus, which gave rise to a UBA-possessing RSC1A1 protein. Interestingly, mammals embed two DDI1-like genes in their genome, DDI1 and DDI2. DDI2 is quite similar to non-mammalian Ddi1-like genes and therefore is considered the original version of the gene. Mammalian DDI1 was most presumably duplicated in a retrotransposition event from the ancestral DDI2 (Siva M.*, et al.*, 2016).

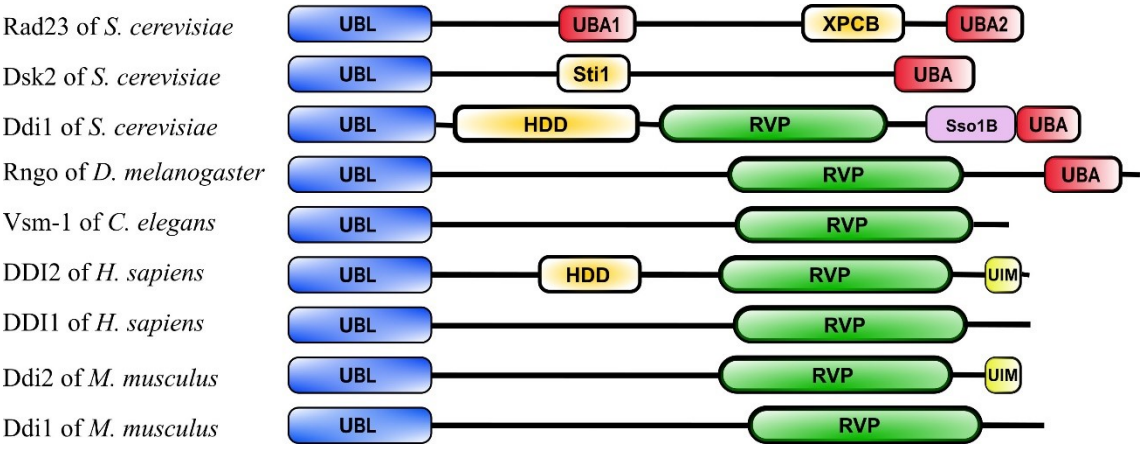


Figure 4: Domain architecture of shuttling proteins of UPS and Ddi1-like protein family. N-terminal UBL domains that interact with proteasomal receptors are in blue. Ubiquitin binding regions at C-terminal are highlighted in color: UBAs in red and UIMs in yellow. Highly helical domains of the shuttling proteins represented by XPC-binding domain in Rad23, Sti1-like domain in Dsk2 and HDDs for yDdi1p and hDDI2 are colored orange. The unique RVP domains in Ddi1-like protein family are colored green. Sso1 binding region of yDdi1p is also highlighten (in lilac).

Ddi1-like family members are involved in a variety of pathways: regulation of cell cycle progression, check point control, genome integrity or proteostasis maintenance and regulation of late exocytotic processes. Related to the role in UPS, interaction with Rpn10 proteasomal subunit with Ddi1 of a fruit fly was first observed in a pull-down experiment, however NMR titration data from several studies did not confirm the original data (Morawe T.*, et al.*, 2011, Nowicka U. *et al.*, 2015, Trempe J. F.*, et al.*, 2016, Zhang D. *et al.*, 2009). First direct evidence of proteolytic activity of the RVP in Ddi1-like protein family was published in a study of Ddi1 from Leishmania major, which claimed that Ddi1 was able to cleave BSA together with several HIV-1 and cathepsin D substrates. However, this was not observed for human DDI2 (Perteguer M. J. *et al.*, 2013, Siva M.*, et al.*, 2016). Effect of aspartic protease inhibitors on Ddi1-like proteins was identified by both *in vivo* and *in vitro* experiments, by impairment of deletion-rescue effect in DDI1 knockout in yeast strains and in enzymatic reactions with parasitic Ddi1 (Perteguer M. J.*, et al.*, 2013, White R. E. *et al.*, 2011a). The parasitic family of the trypanosomatid is mainly responsible for opportunistic infections. Several studies have shown that antiviral therapy in AIDS patients suffering from parasitic infection results in decrease of development of the opportunistic disease (Savoia D. *et al.*, 2005, Skinner-Adams T. S. *et al.*, 2004). Therefore, HIV-1 protease inhibitors are used in a variety of studies targeting the function of Ddi-1 protein family.

Ddi1 in higher eukaryotes is essential for embryonal development. Rngo (Ddi1 ortholog in Drosophila melanogaster) is highly expressed and ubiquitinated during embryonal neural development (Franco M. *et al.*, 2011). Interestingly, Rngo was found to be specifically

ubiquitinated by E3 ligase Ube3a in fruit fly neurons, which is not a signal for degradation in the UPS (Ramirez J. *et al.*, 2018). Moreover, fruit flies with depleted od Rngo exhibit defects in oogenesis and die at pupal stages, while the phenotype cannot be rescued by an inactive protein mutant (D257A), which highlights the function of the RVP domain (Morawe T.*, et al.*, 2011). In Caenorhabditis elegans, Vsm-1 (v-SNARE master protein 1) inhibits synaptogenesis in nematodes, as Vsm-1-null roundworms exhibit profound synaptic density (Guthmueller K. L. *et al.*, 2011). Recent study from Ruvkun laboratory revealed function of VSM-1 in activation of transcription factor SKN-1, which regulates expression of many proteasomal subunits (Lehrbach N. J. and Ruvkun G., 2016).

Ddi1 from *S. cerevisiae* and both human DDI1 and DDI2 are described in more detail in following chapters. Studying roles of mouse Ddi1 and Ddi2 genes is one of the aims of this thesis.

1.3.1 Ddi1 in *Saccharomyces cerevisiae*

DDI1 gene of *S. cerevisiae* is encoded on chromosome V, position 456,319-457,605 on the forward strand (YER143W, NCBI ID: 856886, SGD: S000000945). The protein consists of 428 amino acids (P40087) and harbors four domains, UBL, HDD, RVP and UBA (Geer L. Y. *et al.*, 2010, The UniProt C., 2017, Zerbino D. R. *et al.*, 2018).

Ddi1p, as a shuttling protein of UPS, interacts with polyubiquitin chains and proteasomal subunit Rpn1, however the later interaction is very weak (Gomez T. A. *et al.*, 2011, Rosenzweig R.*, et al.*, 2012, Saeki Y. *et al.*, 2002b, Wilkinson C. R.*, et al.*, 2001). Interestingly, Ddi1 interaction with ubiquitin was identified not only for the UBA as expected, but as well for the UBL domain. The interaction site was mapped onto the Ile44 patch of ubiquitin and analogous patch on UBL formed mainly by residues Ile13, Leu70 and Leu72 (Nowicka U.*, et al.*, 2015). Our study showed that a dimer of full-length Ddi1 protein binds two molecules of K48-linked diubiquitin (Trempe J. F.*, et al.*, 2016). Another domain was discovered to be docked between the UBL and RVP domains, a helical domain of Ddi (HDD). Ddi1 HDD of S. cerevisiae consists of two alpha-helical structured regions connected with a 10-residue linker: a four helical bundle with conserved hydrophobic core at the N-terminus and a helix-turn-helix at the C-terminus. The N-terminal bundle of HDD exhibits structural similarity to several DNA binding domains of transcription factors, yet this possible function has not been verified (Trempe J. F.*, et al.*, 2016). This domain proposes a multifunctional potential for the function of Ddi1p as was confirmed for other shuttling proteins, as they

indeed harbor Sti1-like domains that connect them to DNA repair processes (Kaye F. J.*, et al.*, 2000, Masutani C.*, et al.*, 1997, Schauber C.*, et al.*, 1998).

Expression of Ddi1p is induced upon DNA damage, as its transcription is regulated by a DNA damage-inducible promoter. This promoter is bidirectional and it controls expression of either DDI1 or MAG1 (a 3-methyladenine DNA glycosylase with role in base excision repair) alternatively, in response to divergent DNA damage (Fu Y. *et al.*, 2008, Liu Y. and Xiao W., 1997, Liu Y. *et al.*, 1997, Zhu Y. and Xiao W., 1998). Upon MEC1-dependent DDR pathway activation, a mating type locus MAT allele switching enzyme Ho endonuclease is rapidly degraded. As Ho endonuclease accumulates in Ddi1p-deficient cells and moreover, it interacts with Ufo1, factor that binds and brings phosphorylated Ho endonuclease into proximity with E3 ubiquitin ligase prior proteasomal degradation. Ddi1p is therefore considered a regulatory component of the Ho endonuclease degradation (Ivantsiv Y. *et al.*, 2006, Kaplun L. *et al.*, 2000, Kaplun L. *et al.*, 2003, Kaplun L. *et al.*, 2005).

Ddi1p was found to be involved in regulation of mitotic checkpoint control protein Pds1 degradation, which is required for partition of sister chromatids for G/M phase and anaphase onset (Clarke D. J.*, et al.*, 2001, Diaz-Martinez L. A. *et al.*, 2006). It was identified as a suppressor of temperature sensitivity in *PDS1* mutants (Clarke D. J.*, et al.*, 2001). Deletion of *DDI1* gene in yeast cells results in augmentation of protein secretion into media, which determines Ddi1p as a negative regulator of exocytosis (White R. E. *et al.*, 2011b). Ddi1p interacts with three SNARE proteins, an endocytic v-SNARE protein Snc2 and exocytic v-SNARE and t-SNARE proteins Snc1 and Sso1, respectively (Lustgarten V. and Gerst J. E., 1999, Marash M. and Gerst J. E., 2003). Ddi1p binds to Sso1 via a linker between RVP and UBA domains. Phosphorylation of autoinhibitory domain of Sso1 and of T348 in the interacting linker sequence of Ddi1p is as well a regulatory factor of exocytosis (Gabriely G.*, et al.*, 2008, Marash M. and Gerst J. E., 2003). Ddi1p was found to be one of regulatory genes required for the G protein α subunit exocytosis, which is essential for efficient mating (Dixit G. *et al.*, 2014).

Recently, a high throughput synthetic lethality screen in yeast exhibited genetic connection between *DDI1* and *WSS1*, which encodes the metalloprotease responsible for DNA-protein crosslink removal (the process is described in chapter 1.2.1) (Costanzo M. *et al.*, 2016). Double mutant Δddi1, Δwss1 strain of budding yeast is highly sensitive to DNA damage caused by hydroxyurea, as opposed to the single knockout mutant strains, which do not show any or mild sensitivity for Δddi1 and Δwss1 single mutants, respectively (O'Neill

B. M. *et al.*, 2004) (laboratory of Dr. Grantz Šašková – unpublished data). It is the proteolytic activity and the four helical bundle of the HDD domain that are essential and sufficient for DNA repair response and, surprisingly, not the shuttling protein role of Ddi1p, as UBA and UBL domains are redundant for phenotype rescue (laboratory of Dr. Grantz Šašková – unpublished data).

Recent publication revealed yeast Ddi1p as a natural substrate for metacaspases that cleave off its UBA domain under highly specific conditions. This processing that most probably modulates the function of Ddi1p, was observed as well for Ddi1-like protein in trypanosomes (Bouvier L. A. *et al.*, 2018).

1.3.2 DDI2 in *homo sapiens*

Human *DDI2* gene (ENSG00000197312, NCBI ID: 84301) is located on the first chromosome (15,617,500-15,669,044) on forward strand and consists of 10 exons, 9 of which comprise a protein coding transcript (ENST00000480945.5 - Ensemble release 95, NM_032341.5). This transcript encodes a DDI2 protein formed by 399 amino acids (Q5TDH0, NP_115717.3). No other transcription variants for *DDI2* have been discovered hitherto (Geer L. Y.*, et al.*, 2010, The UniProt C., 2017, Zerbino D. R.*, et al.*, 2018). *DDI2* is expressed quite ubiquitously in a variety of tissues in adult humans according to online databases. Its expression is significantly increased in several cell lines derived from carcinoma, e.g., prostate, breast or skin cancer (Uhlen M. *et al.*, 2017).

The structure of DDI2 was characterized in our laboratory in collaboration with J.-F. Trempe (Siva M.*, et al.*, 2016). Solution structure of N-terminal UBL domain reveals the conserved ubiquitin β-grasp fold typical for UBLs. Beta-sheet interaction site of DDI2 UBL is however moderately charged in comparison to negatively charged interaction patch of yDdi1 UBL, which is able to bind positively charged site on ubiquitin (Nowicka U.*, et al.*, 2015, Siva M.*, et al.*, 2016, Trempe J. F.*, et al.*, 2016). This is surprising, as UBLs had not been previously reported to bind ubiquitin itself (Nowicka U.*, et al.*, 2015). A novel domain, helical domain of Ddi (HDD), was identified in the sequence between already described UBL and RVP domains. Residues 125 to 212 form four helices that pack into a compact α-helical bundle with hydrophobic core and an arginine rich region that has been reported to figure in DNA binding domains, such as XPC-binding domain of Rad23 (Kim B. *et al.*, 2005, Lee J. H. *et al.*, 2005, Siva M.*, et al.*, 2016). The RVP forms a constitutive dimer of the DDI2 protein. DDI2 RVP structure adopts a typical aspartic-protease fold, however the substrate chamber is larger as flaps only distantly lay over the active site. The catalytic site of DDI2

consists of DSGA motif (Siva M., *et al.*, 2016). Additionally, a 24-residue region was identified at the C-terminus of DDI2 as potential ubiquitin binding sequence based on its similarity with conserved UIMs (Siva M., *et al.*, 2016). This four domain architecture provides multifunctional potential: N-terminal UBL and C-terminal UIM could set the basis for the role of a shuttling protein in UPS, the activity of DDI2 RVP has already been recently identified (described below) and the DNA-binding potential of HDD unlocks door to a set of new possible roles. Recently, several studies revealed distinct functions of human DDI2 as described below.

Koizumi and colleagues identified DDI2 as an activator of transcription factor NFE2L1 (NRF1) in siRNA screen and complementation experiments on human knockout and knock-in cell lines (Koizumi S. *et al.*, 2016). Transcription factor NRF1 is a constitutively expressed protein that regulates basal and stress-induced expression of a broad spectrum of genes (Kim H. M., *et al.*, 2016). Function of NRF1 (and its homolog NRF2) in oxidative stress response has been already briefly described in this work in chapter 1.2.2 on page 31. Additionally to antioxidant gene regulation, activated NRF1 targets ARE sequences of several proteasomal subunit genes and thereby is able to induce their expression (Biswas M. and Chan J. Y., 2010, Radhakrishnan S. K. *et al.*, 2014). Under normal conditions, NRF1 is degraded in ERAD pathway. Koizumi and colleagues showed, that upon proteotoxic stress, DDI2 is activated and cleaves deglycosylated NRF1 that escapes degradation and is translocated into nucleus where it induces expression of proteasomal subunit genes in a "bounce back" response (for scheme of the mechanism see Figure 6) (Biswas M. and Chan J. Y., 2010, Koizumi S., *et al.*, 2016, Radhakrishnan S. K., *et al.*, 2014, Radhakrishnan S. K., *et al.*, 2010, Steffen J., *et al.*, 2010). DDI2 cleaves NRF1 p120 form between the P1:Trp103 and P1′:Leu104 onto the active p110 form, which is a possible cleavage motif of RVP domain (Koizumi S., *et al.*, 2016, Radhakrishnan S. K., *et al.*, 2014). In addition, proteolytic activity of Ddi1 in *C. elegans* and its activation potential for a SKN-1 protein, an ortholog of NRF1 and NRF2, was recently described (Lehrbach N. J. and Ruvkun G., 2016).

Another recently discovered substrate of DDI2, transcription factor NFE2L3 (NRF3), is post-translationally housed in the ER lumen and constitutively sequestrated by ERAD pathway, similarly to NRF1 (Chowdhury A. *et al.*, 2017, Zhang Y. *et al.*, 2009). DDI2 cleaves NRF3 between Trp111 and Leu112, in sequence of NHB2 domain homologous to NRF1 NHB2 domain (Chowdhury A., *et al.*, 2017). Detailed characterization of NRF1 and NRf3, plus their up-to-date pathway regulation coverage is provided in following chapter 1.3.2.1.

The findings of NRF1 and NRF3 activation by DDI2 arose from biological experiments; however, its proteolytic activity has not been yet characterized on molecular level or enzymatically. Despite the effort, no proteolytic activity has been detected for DDI2 under normal conditions using a variety of biochemical methods, such as HPLC enzymatic analysis or PICS on mammalian-cell-peptide derived libraries (Siva M*., et al.*, 2016).

It was recently proposed that DDI2 and DDI1 may function as proteasomal shuttling factors for hitherto understudied replication termination factor 2 (RTF2) (Kottemann M. C*., et al.*, 2018). Its only identified homolog, Rtf2 in *S. pombe,* is responsible for maintenance of fork stalling for preservation of unidirectional replication of the mating type locus (Inagawa T. *et al.*, 2009). Human RTF2 associates with nascent DNA and most probably causes uncoupling of helicase from replicative polymerases (Dungrawala H. *et al.*, 2015, Kottemann M. C*., et al.*, 2018). The persistence of RTF2 on DNA in DDI1/DDI2-depleted cells under replication stress (hydroxyurea treatment) resulted in failure of full replication and cell cycle progression leading to cellular death. DDI1 and DDI2 were identified as key regulators of RTF2-dependent stalled replication fork recovery and RTF2-dependent maintenance of genome integrity (Kottemann M. C*., et al.*, 2018). Moreover, immunoprecipitation and crosslinking experiments revealed interaction of DDI2 with several proteasomal 19S RP subunits and replication factors, such as PCNA, polymerase δ, members of MCM helicase or RPA (Kottemann M. C*., et al.*, 2018).

1.3.2.1 Substrates of hDDI2

Recently discovered substrates of human DDI2 protease, NFE2L1 (NRF1) and NFE2L3 (NRF3), belong to the CNC/bZIP family of transcription factors (Chan J. Y*., et al.*, 1993b). Once translocated into nucleus, they form heterodimers with small Maf proteins or CREB protein, which bind to ARE sequences in promoters of their target genes and thereby induce their expression.

Both proteins quite resemble the structure of their homolog NFE2l2 (NRF2) (see Figure 5). In addition to very well conserved Neh1 domain (CNC and bZIP regions), they harbor Neh3-, Neh6- and Neh5-like domains (Andrews N. C*., et al.*, 1993, Zhang Y*., et al.*, 2009, Zhang Y. *et al.*, 2014a). In addition, they possess an additional N-terminal domain (NTD) that includes so-called N-terminal homology box 1 (NHB1: residues 7-24 and 12-31 in NRF1 and NRF3, respectively), which is rich in hydrophobic residues and mediates their anchoring to the ER membrane (Wang W. and Chan J. Y., 2006, Zhang Y*., et al.*, 2006, Zhang Y*., et al.*, 2009). NTD encodes a NHB2 domain, which comprises a cleavage site for DDI2.

A savoir, NRF1 has almost identical domain organization as NRF3; however, they differ in sequence (Zhang Y., *et al.*, 2009). In NRF1, there are two acidic domains (AD1 and AD2) near the N-terminus separated by Asp/Ser/Thr-rich region called NST, representing its glycosylation site (Zhang Y. and Hayes J. D., 2010, Zhang Y., *et al.*, 2009, Zhang Y., *et al.*, 2014a). Out of these, NRF3 lacks the acidic region 1 (Zhang Y., *et al.*, 2009).

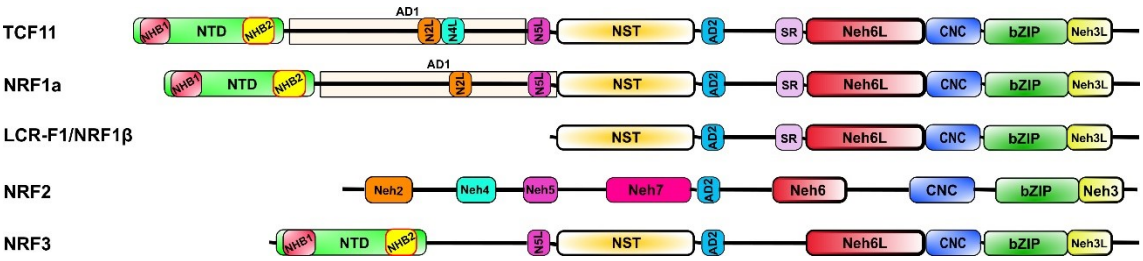


Figure 5: Domain organization of homologs NRF1, NRF2 and NRF3. The full-length NRF2 and NRF3 proteins with three transcription variants of NRF1 (TCF11, NRF1a and LCR-F1) are shown. Several domains, such as CNC, bZIP, Neh6, Neh3, Neh5 and AD2 are very well conserved among all the enlisted variants and homologs. TCF11, NRF1a and NRF3 harbor the N-terminal NTD, which serves as anchor in the ER membrane and comprises cleavage site for DDI2 protease in the NHB2 region. Adapted from (Zhang Y., *et al.*, 2014a).

NRF1 protein harbors a serine-rich domain at its C-terminus close to the Neh1 domain (Chan J. Y. *et al.*, 1993a, Zhang Y., *et al.*, 2014a). In humans, there are two transcription variants harboring the DDI2 cleavage site in NHB2 domain for *NFE2L1*, a 772 residue long TCF11 and 742 residue long NRF1a (see Figure 5). Another variant, LCR-F1/NRF1β, functions as transcription regulator as well. The individual splice variants are discussed in following chapter.

As the expression profiles, roles and depletion phenotypes differ for NRF1 and NRF3, they are described individually below.

1.3.2.1.1 NRF1

NFE2L1 (NRF1) gene is located on chromosome 17, in the region 48,048,329-48,061,487 on forward strand (ENSG00000082641 - Ensemble release 95, NCBI ID: 4779). It is a 6 exon gene exhibiting a number of splice variants (Geer L. Y., *et al.*, 2010, The UniProt C., 2017, Zerbino D. R., *et al.*, 2018). The longest transcript consists of 772 amino acids and is called TCF11. A splice variant without exon 4, which is 742 residues long is termed NRF1α (Luna L. *et al.*, 1995). Both these isoforms generate protein products that can be detected as around 120 kDa. A shorter transcript of 572 amino acids called LCRF1 (or NRF1β, p65NRF1) with molecular weight approx. 65kDa might be product of alternative translation initiation derived from internal ATG codons (Chan J. Y., *et al.*, 1993b). Domain organization of the three above mentioned splice variants is shown in Figure 5. LCRF1 is as well able to

heterodimerize with Maf proteins and bind ARE sequences (with weaker capacity), nevertheless it functions as a repressor of NRF2 activity (Caterina J. J. *et al.*, 1994, Wang W. *et al.*, 2007, Zhang Y. *et al.*, 2014b). Transcript variant NRF1b, created via alternative promotor using an alternative exon 1, is 583 acids long with molecular weight around 100 kDa (Kwong E. K. *et al.*, 2012). The NRF1γ and NRF1δ, presumably proteasome-processed isoforms of NRF1, are shorter proteins of molecular mass of 36 and 25 kDa, respectively, representing dominant-negative inhibitors of NRF2 and long forms of NRF1 (Zhang Y.*, et al.*, 2014b, Zhang Y. *et al.*, 2015).

NRF1 expression is ubiquitous in adult humans. Higher levels of expression were detected in heart, kidney, skeletal muscle, brain and fat (Chan J. Y.*, et al.*, 1993a, Kim H. M.*, et al.*, 2016). mRNA levels of *NRF1* are elevated in a number of cancer cell lines according to online databases. A regulatory SNP was reported by Hirotsu and colleagues, whose study indicates that rs3764400 correlates with increased *NRF1* expression (Hirotsu Y. *et al.*, 2014).

In cells, NRF1 is expressed constitutively. Levels of expression are regulated by a variety of factors, from nutrients to stress stimuli. Aside from stress response, NRF1 activation of proteasomal subunits can be modulated by mTORC1-mediated SREBP1 transcription factor (Zhang Y. and Manning B. D., 2015, Zhang Y. *et al.*, 2014c). Under normal conditions, NRF1 is degraded by ERAD pathway (for detailed description see chapter 1.2.3.2 on page 35), which means that after translation and glycosylation in ER, it is subsequently retrotranslocated into cytosol via HRD1/SEL1L complex, deglycosylated by PNGase, ubiquitinated by HRD1, partially unfolded by VCP/p97 and degraded in the proteasome (Sha Z. and Goldberg A. L., 2014, Steffen J.*, et al.*, 2010). The half-life on NRF1 protein in this ERAD cascade is approx. 12min (Steffen J.*, et al.*, 2010). Upon proteasomal inhibition, NRF1 is cleaved by DDI2 instead of its transport to proteasome and its C-terminal p110 (of approx. 110 kDa) active form is translocated into nucleus – in case of TCF11 variant of NRF1 cleavage (Biswas M. and Chan J. Y., 2010, Radhakrishnan S. K.*, et al.*, 2014). In humans, both TCF11 and NRF1a are processed by DDI2. The processed form of NRF1 is able to binds to ARE and induces expression of NRF1 downstream genes. This occurs not only under stressed but as well under normal conditions, as was described for some specific cell types, such as hepatocytes or neuronal tissue (Kim H. M.*, et al.*, 2016, Lee C. S. *et al.*, 2013, Lee C. S. *et al.*, 2011). This basal function of NRF1 arises as well from knock-in HCT116 cell line experiments, where RVP domain inactive DDI2 D252N mutant knock-in cells showed decrease in proteasomal activity when compared to WT and DDI2 WT knock-in cells (Koizumi S.*, et al.*, 2016). Transcriptional activity of nuclear p110 NRF1 form is

regulated by ubiquitination and subsequent degradation in proteasome facilitated by SKP1-β-TRCP complex that binds to the DSGLS recognition motif for β-TrCP, phosphorylated at serine residues (Tsuchiya Y., *et al.*, 2011). Similarly, NRF1 can be phosphorylated by glycogen synthase kinase 3 in its Cdc4 phosphodegron sequence (residues 350–354). This modification facilitates its interaction with Fbw7, another SCF (Skp1-Cul1-Fbox protein-Rbx1)-type ubiquitin ligase, which ubiquitinates NRF1 and targets it for proteasomal degradation (Biswas M. *et al.*, 2011, Biswas M. *et al.*, 2013). Additional level of NRF1 regulation is phosphorylation of S497 by casein kinase 2, which leads to decrease of proteasome gene expression (Tsuchiya Y. *et al.*, 2013). Recent finding pointed out another control in transcription activity of NRF1 by RUVBL1/RUVBL2 heterohexamer in transcription complex TIP60 (Vangala J. R. and Radhakrishnan S. K., 2018).

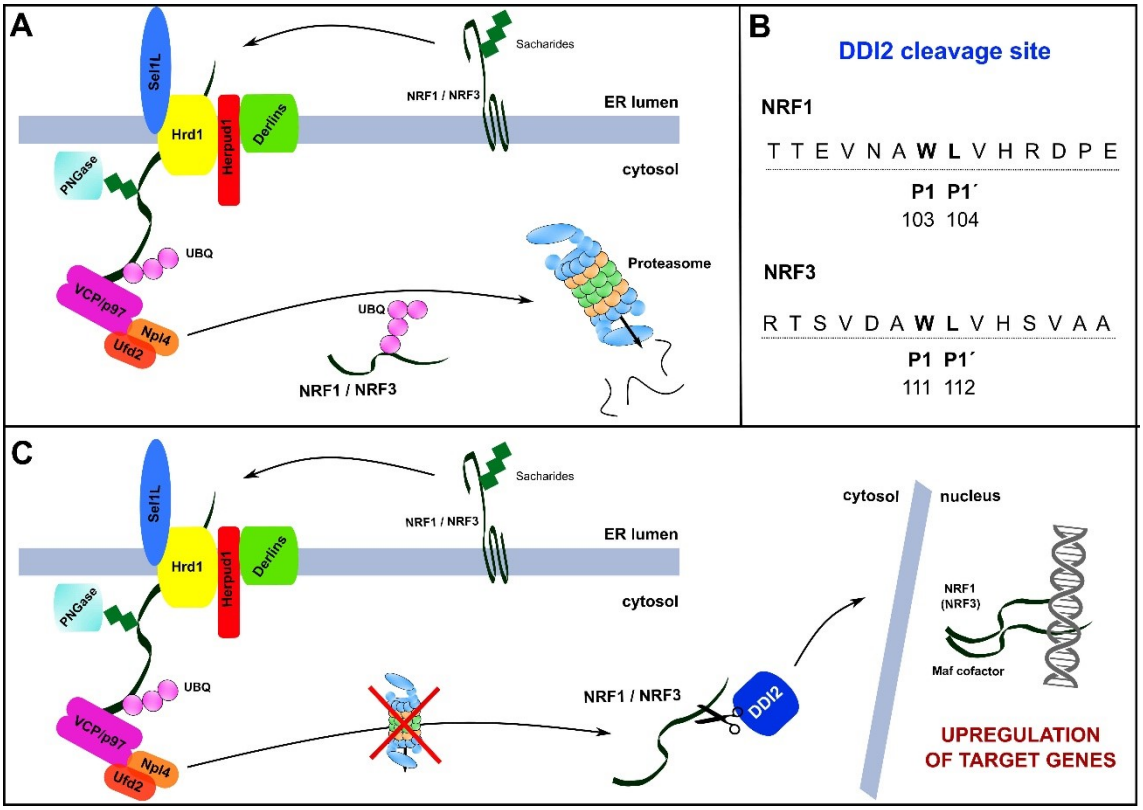


Figure 6: Activation of NRF1 and NRF3 transcription factors by DDI2 protein. A) Cascade of NRF1 processing under normal conditions. NRF1 is constitutively expressed, translocated into ER, retrotranslocated into the cytoplasm by HRD1/SEL1 and VCP/p97 complexes and subsequently degraded in the proteasome. **B)** Human DDI2 targeted sequences of NRF1 and NRF3 proteins. **C)** Upon proteotoxic stress, NRF1 (NRF3) is cleaved by DDI2 and the shorter form is translocated into nucleus where it binds to ARE and regulates its downstream gene expression.

As transcription factor recognizing quite abundant ARE sequences, role of NRF1 has been reported in a variety of cellular pathways. Glutathione (GSH) synthesis pathway genes (*glutamate-cysteine ligase, glutathione synthetase*), *glutathione S-transferase*,

metallothionein-1 and -2, *NADPH-quinone oxidoreductase 1*, *heme oxygenase 1* and *glutathione peroxidase 1* belong to genes controlled by NRF1 in response to oxidative stress (described in chapter 1.2.2 on page 31) (Chan J. Y.*, et al.*, 1998, Chen L.*, et al.*, 2003, Kwong M. *et al.*, 1999, Lu S. C., 2009, Myhrstad M. C.*, et al.*, 2001, Ohtsuji M.*, et al.*, 2008, Song M. O. *et al.*, 2014, Venugopal R. and Jaiswal A. K., 1998, Xu Z.*, et al.*, 2005). In addition to activation of GSH synthesis genes, NRF1 is important for cellular glutathione level maintenance as well by repression of *xCT* expression, which regulates cysteine uptake by xc- transporters (Tsujita T. *et al.*, 2014). Studies on mice and murine cell cultures have reported several other functions for mNrf1. Interestingly, GSH control indirectly regulates expression of proapoptotic *Bik* and *Xpc* functioning in NER in response to UVB DNA damage in keratinocytes (Han W. *et al.*, 2012). Another study suggested the role of NRF1 in genomic integrity and chromosomal stability maintenance, as they found that depletion of NRF1 leads to downregulation of kinetochore and mitotic checkpoint genes *Nuf2*, *Spc25*, *Sgo1* and *Ndc80* (Oh D. H. *et al.*, 2012). Mouse Nrf1 also belongs to control genes of cellular differentiation during formation of mineralized tissue: it activates *Dspp* (in heterodimer with C/EBP-β) and *Osx* genes in odontoblast and osteoblast differentiation, respectively, and *Dmp1* in both processes (Jacob A. *et al.*, 2014, Narayanan K. *et al.*, 2004, Xing W. *et al.*, 2007). Involvement in muscle regeneration was proposed for mNrf1, as its mRNA levels were augmented in macrophages, neutrophils and myofibroblasts during early inflammation in response to muscle injury (Zhang S. T. *et al.*, 2013). mNrf1 was found to feature in cellular immune response: it is able to reverse downregulation of nitric oxide synthase caused by a cytokine TGF-β (Berg D. T. *et al.*, 2007), and moreover, it directly modulates expression levels of *Tnf-α* via its promoter (Novotny V. *et al.*, 1998, Prieschl E. E. *et al.*, 1998). mNrf1 has important regulatory position in metabolism homeostasis. Concerning regulation of lipid metabolism, mNrf1 directly modulates expression of *Lipin1*, *Pgc-1β*, *Apoer2* and *Vldlr* receptor genes, Fads3 desaturase gene and Alox5ap, which results in elevation of triacyl glycerides and alteration of fatty acid composition leading to non-alcoholic steatosis in mNrf1-depleted livers of transgenic mice (Hirotsu Y. *et al.*, 2012, Tsujita T.*, et al.*, 2014). In glucose metabolism, mNrf1 is linked to glucose-stimulated insulin release and fasting hyperinsulinemia, as mNrf1 was found to regulate the expression of *Gck, Glut2, Ldh1, Hk1, Aldob, Pgk1, Pklr, Fbp1, Pck1*, and *Gapdh* genes (Hirotsu Y.*, et al.*, 2014, Zheng H. *et al.*, 2015).

Regulation of proteasome subunit expression is probably the best studied function of NRF1. This function is more profoundly exhibited in so-called “bounce-back” response,

when a set of proteasomal subunits is induced by NRF1 upon proteotoxic stress (Radhakrishnan S. K.*, et al.*, 2010, Steffen J.*, et al.*, 2010). Partial inhibition of proteasome leads to upregulation of proteasomal subunits, however once proteasomal function is abolished, suppression of NRF1 processing occurs (Sha Z. and Goldberg A. L., 2014). The target genes of NRF1 include subunit genes of all the proteasome components, in the regulatory particle (*PsmC1, PsmC4, PsmD1, PsmD12, PsmD14*), and in the core (*PsmA3, PsmA7, PsmB3, PsmB4, PsmB6, PsmB7*) (Radhakrishnan S. K.*, et al.*, 2010, Sha Z. and Goldberg A. L., 2014, Steffen J.*, et al.*, 2010). In addition, NRF1 regulates expression of ERAD pathway components Herpud1 and VCP/p97 (Ho D. V. and Chan J. Y., 2015, Sha Z. and Goldberg A. L., 2014).

A number of knockout mouse model has been prepared for understanding the role of mNrf1 *in vivo*. Nrf1-null model exhibits late gestational embryonic lethality from abnormal fetal liver erythropoiesis (Chan J. Y.*, et al.*, 1998). Another study on mouse chimeras addressed function of mNrf1 to maintenance of hepatocytes (Chen L.*, et al.*, 2003). Hepatocyte specific Nrf1 knockout models show high levels of apoptosis leading to steatosis, inflammation and tumorigenesis (Lee C. S.*, et al.*, 2013, Ohtsuji M.*, et al.*, 2008, Xu Z.*, et al.*, 2005). *Nrf1* deletion in neurons and glial cells leads to neurodegeneration, motor ataxia and forebrain atrophy developing with age (Kobayashi A. *et al.*, 2011, Lee C. S.*, et al.*, 2011). mNrf1 depletion in osteoblasts resulted in reduction of bone mass, size and mechanical strength (Kim J. *et al.*, 2010). On the contrary, a study of Hirotsu and colleagues with over-expression of mNrf1 in mice revealed insulin resistance of the transgenic mice. Insulin signaling is suppressed in liver and skeletal muscle by mNrf1 indirectly - via protein kinase B activation due to enhanced fatty acid oxidation (Hirotsu Y.*, et al.*, 2014).

1.3.2.1.2 NRF3

NFE2L3 (NRF3) is localized on chromosome 7, 26,152,240-26,187,125 on forward strand. It is a 4 exon gene (ENSG00000050344 - Ensemble release 95, NCBI ID: 9603) with one known splice variant (ENST00000056233.3, NM_004289.7) for a 694-residue long protein (Q9Y4A8, NP_004280.5) (Geer L. Y.*, et al.*, 2010, The UniProt C., 2017, Zerbino D. R.*, et al.*, 2018). For structure organization see chapter 1.3.2.1. and Figure 6. *NRF3* is expressed in a variety of human tissues according to up-to-date online databases; however, it exhibits high expression levels specifically in placenta (Chenais B. *et al.*, 2005). In addition, transcript levels are increased in lymphoma, breast and testicular carcinoma and colorectal adenocarninoma cell lines (Chevillard G. and Blank V., 2011, Chowdhury A.*, et al.*, 2017).

The degradation of NRF3 is performed by HDR1/p97 connected ERAD pathway in cytosol and by SKP1-β-TRCP complex in nucleus such as was described for its homologs NRF1 and NRF2 (for degradation mechanism details see chapter 1.2.2 on page 31) (Chowdhury A.*, et al.*, 2017).

NRF3 was found to be a negative regulator of NRF2 target antioxidant genes *NAD(P)H:Quinone Oxidoreductase1* and *peroxiredoxin 6* (Chowdhury I. *et al.*, 2009, Sankaranarayanan K. and Jaiswal A. K., 2004). Cell cycle regulator *U2AF homology motif kinase 1* (*UHMK1*) is another gene regulated by NRF3 in colon carcinoma cells, via which NRF3 could promote cell proliferation (Chowdhury A.*, et al.*, 2017). Pro-apoptotic role of NRF3 was identified in NRF3-dependent regulation of cell adhesion-related proteins function human keratinocytes upon exposure to UV radiation (Siegenthaler B. *et al.*, 2018). Study of Wang and colleagues recently discovered correlation between NRF3 expression levels and metastasis in pancreatic cancer tissues. NRF3 expression levels were remarkably increased in cancerous tissue in comparison to adjacent non-cancerous tissue suggesting that it most probably acts as an oncogene in pancreatic cancer (Wang H. *et al.*, 2018). Another role for NRF3 was identified in smooth muscle cell differentiation from stem cells. mNrf3 induces expression of smooth muscle cell markers, *phospholipase* A2, group 7 and pro-oxidant genes that produce ROS, which are essential for advancement of the differentiation process (Pepe A. E. *et al.*, 2010, Xiao Q. *et al.*, 2012).

Nrf3 knockout mice do not show any developmental defects or growth abnormalities under non-challenging conditions, however they are sensitive to a variety of stressing stimuli (Derjuga A. *et al.*, 2004). Nrf3 null mice showed profound weight loss after treatment with antioxidant BHT (butylated hydroxytoulene) in comparison with their WT littermates and developed acute lung and adipose tissue damage. Levels of Nrf1 and Nrf3 expression in WT mice decreased upon exposure to BHT as opposed to Nrf2 mRNA levels, which suggests that they represent different regulatory responses to these conditions (Chevillard G. *et al.*, 2010). Another study showed that Nrf3 null mice are prone to lymphomagenesis of T-cell origin in response to chemical carcinogen treatment (Chevillard G. *et al.*, 2011). As well, regulation of response to inflammation was suggested as one of the functions for mNrf3 transcription factor in connection to its compensatory role in Nrf2 deficient mice (Braun S. *et al.*, 2002). Interestingly, double knockout *Nrf3*$^{-/-}$*/Nrf2*$^{-/-}$ or *Nrf3*$^{-/-}$*/Nfe2*$^{-/-}$ mice did not show any obvious phenotype as well as their single knockout parental strains (Derjuga A.*, et al.*, 2004).

1.3.3 DDI1 in *homo sapiens*

Human *DDI1* is encoded by one exon gene (ENSG00000170967, NCBI ID: 414301) and therefore has one transcript (ENST00000302259.4 - Ensemble release 95, NM_001001711.2) that corresponds to the 396 residue long protein (Q8WTU0, NP_001001711.1) (Geer L. Y.*, et al.*, 2010, The UniProt C., 2017, Zerbino D. R.*, et al.*, 2018). Secondary structure predictions and sequence alignments of human DDI1 and DDI2 proteins, which share 72% sequence identity, suggest that DDI1 protein shall harbor three domains previously characterized for yDdi1 and DDI2: UBL, HDD and RVP (for sequence alignment see Figure 7 on page 100).

Human DDI1 has been scarcely studied, which is reflected also in amount of data available in online expression databases. Its expression is restricted exclusively to testes in healthy adult humans, although its production is slightly elevated in melanoma and lymphoma cell lines. Interestingly, mutations in *DDI1* were recently identified in patients suffering from familial neurodegenerative disorder, a variant of Alzheimer´s disease (Alexander J. *et al.*, 2016).

DDI1 is specifically ubiquitinated by an E3 ubiquitin ligase UBE3A in human neuroblastoma cells. Surprisingly, this modification does not target DDI1 for proteasomal degradation (Ramirez J.*, et al.*, 2018). A complex neurodevelopmental disease called Angelman syndrome is caused by mutations in a single gene encoding the UBE3A ligase, which suggest that its substrates might affect pathways essential for proper brain development (Buiting K. *et al.*, 2016, Sadikovic B. *et al.*, 2014). Ramirez and colleagues showed that expression of murine Ddi1 increases rapidly on embryonic day 16.5, which might mean a specific function for DDI1 at this stage of embryonal development (Ramirez J.*, et al.*, 2018). Moreover, it had been previously suggested that Ddi1 of *C. elegans* may regulate synaptogenesis (Guthmueller K. L.*, et al.*, 2011). Altogether, these findings point out possible relevance of DDI1 for development of Angelman syndrome (Ramirez J.*, et al.*, 2018).

Another function has been recently described for DDI1 together with DDI2 (see chapter 1.3.2 on page 51), both of which participate in DNA damage response, specifically during recovery from replication stress at repair of stalled replication forks (Kottemann M. C.*, et al.*, 2018).

2 AIMS AND OBJECTIVES

Several studies identified members of the evolutionary highly conserved family of Ddi1-like proteins as important regulators of homeostasis maintenance in a variety of pathways, such as ubiquitin-proteasome system, DNA damage response or proteotoxic stress response. Despite the emerging evidence of their crucial role in these biological processes, Ddi1-like proteins have been poorly studied on the molecular level or in suitable biological systems. Therefore the work presented in the thesis will enrich up-to-date studies towards better understanding of the role and mechanism of function of several Ddi1-like protein family members.

These objectives were set towards fulfillment of specific aims:

1. Structure solution and characterization of the UBL domain of Ddi1p from *S. cerevisiae* using NMR spectroscopy.

2. Characterization of the binding properties of UBL domain and ubiquitin-interacting motif of human DDI2 protein, in context of its possible function in ubiquitin-proteasome system using NMR spectroscopy.

3. Expression profiling of murine homolog *Ddi1* in developing mouse brain using *in situ* hybridization method.

4. Generation and characterization of Ddi2 full knockout mouse strain as the most suitable animal model for studying the physiological function of human DDI2 protein.

5. Generation and characterization of Ddi2 protease defective mouse model strain for deciphering the function of retroviral protease-like domain of Ddi2.

6. Comparison and detailed analysis of the phenotypes of both gene-edited mouse model strains.

7. Identification of possible pathways involving Ddi2 that are connected to phenotype development (if any) of the mouse strains.

3 MATERIALS AND METHODS

3.1 MATERIALS

3.1.1 Chemicals and solutions

acetic acid (Penta, Prague, Czechia)
acrylamide (Sigma-Aldrich, St. Louis, USA)
agarose (Serva, Heidelberg, Germany)
Albumin standard (2 mg/ml) (Thermo Scientific, Waltham, USA)
All Blue pre-stained protein standard (Bio-Rad, Hercules, USA)
ammonium chloride (^{15}N-labeled) (Cambridge Isotope Laboratories, Andover, USA)
ammonium persulfate (Serva, Heidelberg, Germany)
ampicillin (Sigma-Aldrich, St. Louis, USA)
Aquatex (Millipore, Burlington, USA)
BM purple (Roche, Basel, Switzerland)
boric acid (Sigma-Aldrich, St. Louis, USA)
bortezomib (UBPBio, Aurora, USA)
bromphenol blue (Sigma-Aldrich, St. Louis, USA)
calcium chloride (Penta, Prague, Czechia)
cobalt dichloride (Sigma-Aldrich, St. Louis, USA)
cOmplete™ Mini, EDTA-free, Protease Inhibitor Cocktail (Roche, Basel, Switzerland)
cupric chloride (Sigma-Aldrich, St. Louis, USA)
D-biotin (Sigma-Aldrich, St. Louis, USA)
deuterium dioxide (Merck, Billerica, USA)
D-glucose (Sigma-Aldrich, St. Louis, USA)
D-glucose (^{13}N-labeled) (Cambridge Isotope Laboratories, Andover, USA)
dimethyl sulfoxide (Sigma-Aldrich, St. Louis, USA)
DirectPCR Lysis Reagent (Viagen Biotech, Inc, Los Angeles, USA)
dithiotreitol (Sigma-Aldrich, St. Louis, USA)
DMEM High Glucose w/o L-Glutamine (Biosera, Nuaille, France)
DNA ladder 100 bp (Qiagen, Hilden, Germany)
dNTP mix (Serva, Heidelberg, Germany)
dry milk - Blotting-Grade Blocker (Bio-Rad, Hercules, USA)
ethylenediaminetetraacetic acid (EDTA) (Sigma-Aldrich, St. Louis, USA)
egtazic acid (EGTA) (Sigma-Aldrich, St. Louis, USA)
ferrous chloride (Sigma-Aldrich, St. Louis, USA)
fetal bovine serum (FBS) (Sigma-Aldrich, St. Louis, USA)
formaldehyde (Penta, Prague, Czechia)
gelatine (Penta, Prague, Czechia)
GelRed Nucleic Acid Gel Stain (Biotinum, Fremont, USA)
glycerol (Penta, Prague, Czechia)
glycine (Duchefa, Haarlem, Netherlands)
hydrochloric acid (Penta, Prague, Czechia)
Igepal CA-630 (Sigma-Aldrich, St. Louis, USA)
IMDM medium (Thermo Fisher Scientific, Waltham, USA)
imidazole (Sigma-Aldrich, St. Louis, USA)
isopropanol (Penta, Prague, Czechia)
isopropyl β-D-1-thiogalactopyranoside (IPTG) (Biosynth AG, Staad, Switzerland)
kanamycin (Sigma-Aldrich, St. Louis, USA)
LB agar (Sigma-Aldrich, St. Louis, USA)
LB medium (Sigma-Aldrich, St. Louis, USA)
L-glutamine (Sigma-Aldrich, St. Louis, USA)
magnesium chloride (Sigma-Aldrich, St. Louis, USA)

magnesium sulfate (Sigma-Aldrich, St. Louis, USA)
MEM Eagle Vitamin Mixture 100X (Lonza Biotech, Kouřim, Czechia)
MEM non-essential amino acid solution (Sigma-Aldrich, St. Louis, USA)
2-mercaptoethanol (Sigma-Aldrich, St. Louis, USA)
methanol (Penta, Prague, Czechia)
N,N'-methylen-bis(acrylamide) (USB, Cleveland, USA)
Nonidet P-40 (Sigma-Aldrich, St. Louis, USA)
Nuclear Fast Red (Sigma-Aldrich, St. Louis, USA)
opti-MEM medium (Thermo Fisher Scientific, Waltham, USA)
Paraformaldehyde (Sigma-Aldrich, St. Louis, USA)
Penicillin-Streptomycin 100X (Sigma-Aldrich, St. Louis, USA)
Protein Assay Dye Reagent Concentrate (Bio-Rad, Hercules, USA)
polyethylenimine (Sigma-Aldrich, St. Louis, USA)
potassium chloride (Sigma-Aldrich, St. Louis, USA)
potassium dihydrogen phosphate (Sigma-Aldrich, St. Louis, USA)
potassium ferricyanide (Sigma-Aldrich, St. Louis, USA)
potassium ferrocyanide (Sigma-Aldrich, St. Louis, USA)
RNAlater® (Sigma-Aldrich, St. Louis, USA)
silver nitrate (Lachema, Brno, Czechia)
sodium azide (Sigma-Aldrich, St. Louis, USA)
sodium carbonate (Penta, Prague, Czechia)
sodium chloride (Lachema, Brno, Czechia)
sodium deoxycholate (Sigma-Aldrich, St. Louis, USA)
sodium dihydrogenphosphate (Lach-Ner, Neratovice, CZ)
sodium dodecyl sulfate (Sigma-Aldrich, St. Louis, USA)
sodium hydroxide (Penta, Prague, Czechia)
sodium molybdate (Sigma-Aldrich, St. Louis, USA)
sodium phosphate dibasic (Sigma-Aldrich, St. Louis, USA)
sodium pyruvate 100mM 100X (Sigma-Aldrich, St. Louis, USA)
sodium sulfate (Penta, Prague, Czechia)
sodium thiosulfate pentahydrate (Penta, Prague, Czechia)
sucrose (Sigma-Aldrich, St. Louis, USA)
SYPRO® Orange Protein Gel Stain (Sigma-Aldrich, St. Louis, USA)
TEMED (tetramethylethylendiamine) (Fluka, Buchs, Switzerland)
thiamine hydrochloride (Sigma-Aldrich, St. Louis, USA)
tris (tris(hydroxymethyl)aminomethane) (USB, Cleveland, USA)
trypsin-EDTA solution (Sigma-Aldrich, St. Louis, USA)
Tween 20 (USB, Cleveland, USA)
X-gal (Thermo Scientific, Waltham, USA)
zinc chloride (Sigma-Aldrich, St. Louis, USA)

3.1.2 Antibodies

Anti-Digoxigenin-AP, Fab fragments (Roche, Basel, Switzerland)
Mouse monoclonal Anti-α-Tubulin antibody T6199 (Sigma-Aldrich, St. Louis, USA)
Mouse monoclonal Anti-β-Actin antibody, clone AC-15 (Sigma-Aldrich, St. Louis, USA)
Rabbit monoclonal anti- TCF11/NRF1 D5B10 #8052 (Cell Signaling)
Rabbit polyclonal anti-Ddi2 antibody A304-630A (Bethyl)
IRDye® 680 RD Goat anti-Mouse IgG (LI-COR Biosciences, Lincoln, USA)
IRDye® 800 CW Goat anti-Rabbit IgG (LI-COR Biosciences, Lincoln, USA)

3.1.3 Cell cultures

E. coli BL21(DE3)RIL (Novagen – Merck KGaA, Darmstadt, Germany)
E. coli Top10 cells (Invitrogen, Carlsbad, USA)
E. coli DH5α (Novagen – Merck KGaA, Darmstadt, Germany)
HEK293offA2 cells (original strain ATTC, Manassas, USA)

3.1.4 Commercial kits

Agilent RNA 6000 Nano Kit (Agilent technologies, Santa Clara, USA)
cell culture dishes and plates (Biotech, Prague, Czechia)
DIG RNA Labeling Kit (SP6/T7) (Roche, Basel, Switzerland)
mMESSAGE mMACHINE T7 Kit (Thermo Fisher Scientific, Waltham, USA)
Mouse Direct PCR Kit (Bimake, Houston, USA)
Poly(A) Tailing Kit (Thermo Fisher Scientific, Waltham, USA)
RNase-free DNAse Set (QIAGEN, Hilden, Germany)
RNeasy Mini Kit (QIAGEN, Hilden, Germany)
RNeasy plus Micro Kit (QIAGEN, Hilden, Germany)
QIAGEN Plasmid Maxi Kit (QIAGEN, Hilden, Germany)
QIAGEN Plasmid Midi Kit (QIAGEN, Hilden, Germany)
QIAprep Spin Miniprep Kit (QIAGEN, Hilden, Germany)
QIAquick Gel Extraction Kit (QIAGEN, Hilden, Germany)
TATAA GrandScript cDNA Synthesis Kit (TATAA Biocenter, Göteborg, Sweden)
TATAA SYBR® GrandMaster® Mix (TATAA Biocenter, Göteborg, Sweden)
Tissue-Tek® O.C.T. Compound (SAKURA Finetek USA Inc, Torrance, USA)
Whatman filter paper, Grade 470 (GE-Healthcare, Chicago, USA)
Zero Blunt® Cloning® Kit (Invitrogen, Carlsbad, USA)
Zyppy™ Plasmid Minipirep Kit (Zymo Research, Irvine, USA)

3.1.5 Enzymes

Antarctic Phosphatase (New England BioLabs, Ipswich, USA)
Benzonase® Nuclease (Novagen – Merck KGaA, Darmstadt, Germany)
Pfu DNA Polymerase (Promega, Madison, USA)
Phusion® HF Polymerase (New England BioLabs, Ipswich, USA)
proteinase K (New England BioLabs, Ipswich,USA)
BamHI, DpnI, EcoRI, KpnI, NdeI, NheI, SacII, SalI and XbaI (New England BioLabs, Ipswich, USA)
T4 DNA Ligase (New England BioLabs, Ipswich, USA)

3.1.6 Primers

Table 2: List of primers used in this study.

Primer name	Sequence	Use
Ddi1Sc_16b_F	5′-ATCAACATATGGATTTAACAATTTC-3′	cloning
Ddi1Sc_U_16b_R	5′- ATCAAGGATCCCTATCAGGAATTGGAAATCTTACCCC-3′	cloning
Ddi1_F_AS1	5′-GTATTGTGTGCGTAGGGACC-3’	ISH probe cloning
Ddi1_R_AS1	5’-TGAGTCTGTGAGCCGGTAGT-3’	ISH probe cloning
Ddi1_F_AS2	5’-TGAGCTTGAGTCTGGTGTGC-3’	ISH probe cloning
Ddi1_R_AS2	5’-GTGCCTCCTGAGCATATCAAG-3’	ISH probe cloning
Ddi2F	5′-GTCTGGTCCTTGTCCGTGTT-3′	genotyping
Ddi2R	5′-AGTCTGTCATCCCGAGTTGG-3′	genotyping
Ddi2tm1b WT F	5′-GCATGGGCTTACAGTGGTTACTC-3′	genotyping
Ddi2tm1b RV	5′-CTTACTAGTTGCACAGCTGATGACATC-3′	genotyping
Ddi2 IN R	5′-GACTGTAAAACATAAGCCAC-3′	genotyping
Ddi2 long R	5′-CCTGGCAACCTGAAATCAAG-3′	genotyping
Ddi2 nested F	5′-GTGAGACCCTGACTCGGCAA-3′	genotyping
DDI2_HDD_F	5′- ATCAACATATGCAGCAGTCCCACTCA-3′	cloning

DDI2_KpnI_F	5′- CTTCCAGGTACCAAAGATGCTGCTCACC-3′	cloning
Ddi2_OT_1_F	5′-TCCCTTTCATGAGGCCATTC -3′	off-target screen
Ddi2_OT_1_R	5′-AGCGCAGAGAATGAAAAAGC-3′	off-target screen
Ddi2_OT_2_F	5′-TGCTGAATTAGTGCTTTCATGTGG-3′	off-target screen
Ddi2_OT_2_R	5′-TACCATGCACACGCATCTCA-3′	off-target screen
Ddi2_OT_3_F	5′-TTCTTTCCAACTAACCCACA-3′	off-target screen
Ddi2_OT_3_R	5′-CTGGGATGAGAAGTTTTGAG-3′	off-target screen
Ddi2_OT_4_F	5′-TGACCAATGTAGTGGATAG-3′	off-target screen
Ddi2_OT_4_R	5′-TGGTGGATGTCAAGGATTAT-3′	off-target screen
Ddi2_OT_5_F	5′-GAACCTGAGTCTTCTGCAA-3′	off-target screen
Ddi2_OT_5_R	5′-AAGCACTCTACTGCTTTCC-3′	off-target screen
Ddi2_OT_6_F	5′-GAGGAACCACCTAGGGCTGA-3′	off-target screen
Ddi2_OT_6_R	5′-CAGGTCAGAGATGGGTCTGC-3′	off-target screen
Ddi2_OT_7_F	5′-CTGGACACTGGCTCTTC-3′	off-target screen
Ddi2_OT_7_R	5′-CGCAGTAGAAACATTGCAA-3′	off-target screen
Ddi2_OT_8_F	5′-AGCATGGGTACCAATTCCAGA-3′	off-target screen
Ddi2_OT_8_R	5′-TGCACCATGTAGACATTGACG-3′	off-target screen
Ddi2_OT_9_F	5′-TCTCCCTTGCCCCTTTAG-3′	off-target screen
Ddi2_OT_9_R	5′-TAATGGGGGAGTAGGACAGT-3′	off-target screen
Ddi2_OT_10_F	5′-TATAAGCCTGGCCTTTCTTGT-3′	off-target screen
Ddi2_OT_10_R	5′-TGTGCTCTCACACCCAC-3′	off-target screen
Ddi2_OT_11_F	5′-GTTGCAGCTCACCTTGAACG-3′	off-target screen
Ddi2_OT_11_R	5′-TTTGCCAGTCTCAGGTTGCT-3′	off-target screen
Ddi2_OT_12_F	5′-TCTGCTGCATTGTTTTATTGC-3′	off-target screen
Ddi2_OT_12_R	5′-CACAGGAACTTCTGGTGACTT-3′	off-target screen
DDI2_RVP_R	5′- ATCAAGGATCCCTATCACTCTGGTAGCTC-3′	cloning
DDI2_XbaI_R	5′- GATGCGGCCGTCTAGACTATCATGGCTTCTG-3′	cloning
FseqpTRETight	5′-AGGCGTATCACGAGGCCCTTTCGT-3′	sequencing
LacZ_R	5′-ACGGTTTCCATATGGGGATT-3′	genotyping
M13 F	5′-GTAAAACGACGGCCAG-3′	sequencing
M13 R	5′-CAGGAAACAGCTATGAC-3′	sequencing
mDdi2 F cDNA	5′-AAATGCTGCTCACCGTGTAC-3′	cloning
mDdi2 R cDNA	5′-AATCATGGCTTCTGACGCTC-3′	cloning
mDdi2_F	5′-CACACAGAAGATTATTGGAAGG-3′	RT-PCR
mDdi2_p905_F	5′-CAAGCTAGCATGCTGCTCACCG-3′	cloning
mDdi2_p905_R	5′-AGTAAGAATTCCTATCATGGCTTCTGACGCTC-3′	cloning
mDdi2_R	5′-CGTTTCAGCATGTCCAGACC-3′	RT-PCR
mH2afz_F	5′-TAGGACAACCAGCCACGGA-3′	RT-PCR
mH2afz_R	5′-GACGAGGGGTGATACGCTTT-3′	RT-PCR
mTbp_F	5′-TATCTACCGTGAATCTTGGCTG-3′	RT-PCR
mTbp_R	5′-TTGTCCGTGGCTCTCTTATTCT-3′	RT-PCR
MutDDI2_XbaI_F	5′- GACCTTGAGAAATTTTCCAGAGTCCTGGTGGAGCAG-3′	mutagenesis
MutDDI2_XbaI_R	5′- CTGCTCCACCAGGACTCTGGAAAATTTCTCAAGGTC-3′	mutagenesis
RseqpTRETight	5′-TATTACCGCCTTTGAGTGAGCTGA-3′	sequencing
T7 F	5′-TAATACGACTCACTATAGGG-3′	sequencing
T7 R	5′-GCTAGTTATTGCTCAGCGG-3′	sequencing

3.1.7 Vectors

p905 (gift from Dr. Řezáčova laboratory at IOCB CAS, Prague, Czech Republic)
pCR™-Blunt (Thermo Fisher Scientific, Waltham, USA)
pGEM-T® easy plasmid (Promega, Madison, USA)
pET16b (Novagen – Merck KGaA, Darmstadt, Germany)
pTreTight (gift from Dr. Konvalinka laboratory at IOCB CAS, Prague, Czech Republic)

3.1.8 Consumables

96-well transparent plate, F-bottom (P-lab, Prague, Czechia)
Amicon® Ultra 0.5 mL (Millipore, Burlington, USA)
Amicon® Ultra 15 mL (Millipore, Burlington, USA)
dialysis membrane Spectrapor (Spectrum Laboratories – Repligen, Waltham, USA)
Dumont micro forceps (Fine Science Tools, North Vancouver, Canada)

LightCycler® 480 Multiwell Plate 384 (Roche, Basel, Switzerland)
LightCycler® 480 Multiwell Plate 96 (Roche, Basel, Switzerland)
LightCycler® 480 Sealing Foil (Roche, Basel, Switzerland)
Ni-NTA Superflow resin (QIAGEN, Hilden, Germany)
nitrocellulose membrane (Bio-Rad, Hercules, USA)
Superdex™ 75pg 16/60 FPLC Column (GE Healthcare, Chicago, USA)
Superdex™ 200pg 16/60 FPLC Column (GE Healthcare, Chicago, USA)
Sterivex™ GP 0.22 µm filter unit (Millipore, Burlington, USA)

3.2 INSTRUMENTS

autoclave: MLS-3020U, Sanyo (Osaka, Japan)
centrifuges: Beckman Allegra X-15R, Beckman Coulter (Brea, USA)
Beckman Avanti J-30I, Beckman Coulter (Brea, USA)
Centrifuge 5415R, Eppendorf (Hamburg, Germany)
Fresco Heraus 21, IEC CL10, Thermo Scientific (Waltham, USA)
Megafuge 2.0R, Heraeus Instruments (Hanau, Germany)
Sorvall Evolution RC, Thermo Scientific (Waltham, USA)
chromatography: ÄKTAExplorer FPLC, Amersham Pharmacia Biotech - GE Healthcare (Chicago, USA)
electrophoresis: Agilent 2100 Bioanalyzer, Agilent Technologies (Santa Clara, USA)
electrophoresis power supply EPS 301, GE Healthcare (Chicago, USA)
horizontal electrophoresis apparatus, Thermo Scientific (Waltham, USA)
Mini-PROTEAN® Tetra Vertical Electrophoresis Cell, Bio-Rad (Hercules, USA)
Mini Trans-Blot® Cell, Bio-Rad (Hercules, USA)
homogenizers: TissueLyser II, Qiagen (Hilden, Germany)
EmulsiFlex-C3 homogenizer, AVESTIN (Ottawa, Canada)
imaging systems: Odyssey® CLx Infrared Imaging System, LI-COR Biosciences (Lincoln, USA)
UV lamp UVT-20 SML, Herolab (Wiesloch, Germany)
Monochrome scientific grade camera Quantum ST4, Vilber Lourmat (Collegién, France)
incubators: Thermocell Mixing Block MB102, BIOER Technology (China)
CO2 incubator MCO-19AIC, Sanyo (Osaka, Japan)
Innova44, New Brunswick Scientific (Enfield, USA)
scales: EK-400H, A&D Company, (Tokyo, Japan)
PLS 4000-2, KERN & Sohn GmpH (Postfach, Germany)
XA 116/X, Radwag (Sumperk, Czechia)
microscopes: Stemi 305 EDU Microscope, Zeiss (Oberkochen, Germany)
Axio Imager Z2, Zeiss (Oberkochen, Germany)
AxioScan Z1, Zeiss (Oberkochen, Germany)
AxioZoom, Apotome module macroscope, Zeiss (Oberkochen, Germany)
pH-meter: pH 50, XS instruments (Carpi, Italy)
sonicator: sonication bath S 30 Elmasonic, Elma (Singen, Germany)
spectrophotometers:
spectrophotometer UNICAM UV 500, Thermo Scientific (Waltham, USA)
NanoDrop ND-1000, Thermo Scientific (Waltham, USA)
Infinite® microplate reader M1000 PRO, Tecan (Männedorf, Switzerland)
600 MHz Bruker Avance spectrometer, Bruker (Billerica, USA)
850 MHz Bruker Avance spectrometer, Bruker (Billerica, USA)
thermocyclers: TRIO 48, Biometra (Göttingen, Germany)
LightCycler®480 II, Roche (Basel, Switzerland)
vortexes: MX-S,Dragonlab (Beijing, China)

3.3 SOFTWARE

Gimp 2.10.4 (The GIMP Development Team)
GraphPad Prism 7.00 (GraphPad Software, La Jolla, California, USA)
Image Studio Lite Software (LI-COR Biosciences, Lincoln, USA)
Inkscape (The Inkscape Development Team)
Microsoft Office (Microsoft Corporation ,Redmond, USA)
Sparky (University of California, San Francisco, USA)
Vector NTI 11 (Invitrogen, Carlsbad, USA)
Topspin 3.2 (Bruker, Billerica, USA)

3.4 METHODS

3.4.1 DNA cloning and analysis

Several protein constructs were used in this work for characterization of proteins recombinantly expressed in bacterial or human cell cultures. hDDI2 full-length protein and hDDI2 UBL were cloned into pET16b vector (Novagen) by Monika Sivá for studies intended for her diploma thesis. Protein constructs of hDDI2 ΔUIM and hDDI2 RVP-full C were cloned also into pET16b vector (Novagen) expression vector by Michal Svoboda and Klára Grantz Šašková. Plasmids encoding human ubiquitin without (in pET24a) and with N-terminal His-tag (in pHISTEV30a) was kindly provided as a gift from the laboratory of Dr. Ron T. Hay at University of Dundee. Monika Sivá cloned constructs of yDdi1 UBL, hDDI2 HDD-RVP and both mouse $Ddi2^{WT}$ and $Ddi2^{PD}$ (Δ254-296). Sequences of all constructs were designed and later analyzed using Vector NTI software (Invitrogen).

3.4.1.1 DNA construct cloning into bacterial expression vectors

First, common features of the cloning experiments are described for all the plasmid constructs. Primer sequences, DNA templates and DNA restriction with endonucleases is listed below for each construct individually.

Amplification of coding sequences of all constructs was performed with Phusion® HF Polymerase (New England BioLabs). The reactions were mixed as follows: 10 µl of 5X Phusion® HF Buffer, 200 µM dNTPs (Serva), 0.4 µM of relevant primers, 100 ng of DNA template, 0.2 µl of Phusion® High-Fidelity DNA Polymerase and nuclease-free water addition up to 50 µl of total volume. PCR amplification was performed PCR Thermocycler (Biometra) as follows: DNA template denaturation at 98°C for 5 minutes; 35 cycles of denaturation at 98°C for 20 seconds, annealing at temperature corresponding to individual primer pairs for 30 seconds and elongation at 72°C for 60 seconds; and final incubation at 72°C for 5 minutes. Amplified fragments were analyzed by agarose gel electrophoresis

(chapter 3.4.1.3) and subsequently digested with corresponding restriction endonucleases according to manufacturer´s manual for double digest. All digested products were again separated by horizontal agarose gel electrophoresis (control of vector linearization) (chapter 3.4.1.3) and extracted from gel (chapter 3.4.1.4). Dephosphorylation of the linearized vector with Antarctic Phosphatase (New England BioLabs) and following ligation of DNA fragments into the linearized vector with T4 DNA ligase (New England BioLabs) were performed according to manufacturer's instructions. Transformation of bacteria and preparation of plasmid DNA was performed as described in chapter 3.4.1.2. Purified pET16b, p905 and pTreTight plasmids encoding mDdi2 constructs were sequenced by GATC Biotech (Konstanz, Germany) using common sequencing primers T7 F, T7 R (pET16b, p905) or FseqpTRETight, RseqpTRETight (for primer sequences see Table 2 on page 69).

yDdi1 UBL and hDDI2 HDD-RVP coding sequences were amplified using primers Ddi1Sc_16b_F, Ddi1Sc_UBL_16b_R and DDI2_HDD_F, DDI2_RVP_R, respectively (for primer sequences see table 2 on page XX). Plasmids encoding full-length yDdi1p and hDDI2 full-length proteins from our laboratory were used as templates. Both PCR products and pET16b vector were cleaved with restriction endonucleases NdeI and BamHI (New England BioLabs).

Plasmids pCR™-Blunt containing correct $Ddi2^{WT}$ and $Ddi2^{PD}$ coding sequences (for cloning of DNA sequences from cDNA see 3.4.5.14) were used as templates for amplification of the fragments and cloning into a bacterial expression vectors p905 and pTreTight.

Primers mDdi2_p905_F and mDdi2_p905_R were used for amplification of the two variants of mDdi2 coding sequence. Restriction enzyme digestion of corresponding p905 plasmid and two Ddi2 variants was performed with NheI and EcoRI restriction enzymes (New England BioLabs).

Both pCR™-Blunt plasmids encoding mDdi2 variants were used as templates for site-directed mutagenesis producing silent mutation sequence non-cleavable by XbaI restriction endonuclease required for cloning of the DNA sequence into pTreTight vector. Reaction mixes contained 50 ng of the template plasmid, 1µM mutagenesis primers, 250 µM dNTPs (Serva), 2.5 U of Pfu DNA polymerase and 10X reaction buffer (Promega) and water addition up to 50 µl. The reactions were run on PCR Thermocycler (Biometra) as follows: 95 ºC (30 s), 18 cycles of 95 ºC (30 s), 55 ºC (60 s), 68 ºC (5 min) and final elongation of 10 min. The site-directed mutagenesis reaction mixtures were digested by 20 U of DpnI (New England BioLabs) for 1 hour at 37 ºC and analyzed by horizontal agarose gel electrophoresis (chapter 3.4.1.3). Uncleaved plasmids were extracted from gel (chapter 3.4.1.4) and

subsequently transformed into the *E. coli* Top10 competent cells (Invitrogen) as described in following chapter. Successful mutagenesis was identified by colony PCR, where 10 colonies from each agar plate after transformation were suspended in 10 µl of deionized water and 2 µl of the bacterial suspense were used as a template for PCR reaction described for cloning of all construct variants using Phusion® HF Polymerase (New England BioLabs) described at the beginning of this chapter. Primers DDI2_KpnI_F and DDI2_XbaI_R were used for mDdi2 variant DNA sequences amplification and subsequent sequencing by GATC Biotech (Konstanz, Germany). Correct sequences and pTreTight vector were digested by KnpI and XbaI.

3.4.1.2 Transformation of bacteria and amplification of plasmid DNA

The host strain *E. coli* DH5α (Novagen) was used for transformation and amplification of plasmids encoding Ddi1-like proteins used for NMR studies. In case of cloning mDdi2^{WT} and mDdi2^{PD} protein versions from cDNA into bacterial expression vectors, competent *E. coli* strain cells Top10 (Invitrogen) were used. The transformation was carried out as follows (Sambrook J., Fritschi, E.F. and Maniatis, T., 1989): 1 µl of ligation mixture was added to 30 µl of freshly unfrozen competent bacterial cells and left to incubate for 30 minutes on ice. Heat shock was performed at 42°C for 90 seconds, followed by cooling of the bacterial suspense down on ice for 5 minutes. Subsequently, the bacteria were incubated with approx. 400 µl of LB media (without antibiotic) at 37°C for 1 hour, then spread over the agar plates containing relevant antibiotic (100 mg/ml ampicillin or 40 mg/ml kanamycin) and incubated at 37°C overnight.

After overnight incubation, freshly grown colonies were individually picked and inoculated into 12.5 ml, 100 ml or 500 ml of sterile LB medium supplemented with antibiotic (100 mg/ml ampicillin or 40 mg/ml kanamycin) for DNA minipreparation, midipreparation and maxipreparation, respectively. Bacteria were grown in a rotatory incubator Innova 4300 (New Brunswick Scientific) at 37°C and 220 rpm overnight. Following day, the culture was centrifuged at 4000 g, 4°C for 10 minutes and cell pellets were further processed with Zyppy™ Plasmid Miniprep Kit (Zymo Research), QIAGEN Plasmid Midi or Maxi Kit (QIAGEN) according to the protocols provided by the manufacturers. Isolated DNA was eluted from columns with 40 µl of sterile water and its concentration and purity was measured using NanoDrop ND-1000 spectrophotometer (Thermo Scientific). Plasmids were sequenced by GATC Biotech (Konstanz, Germany).

3.4.1.3 Horizontal agarose gel electrophoresis

Horizontal agarose gel electrophoresis was used as analysis method for visualization of reaction products during cloning and genotyping procedures. 1% agarose (Serva) gel prepared in TAE buffer (40 mM Tris-acetate, pH 8.4, 1 mM EDTA) containing DNA stain GelRed (Biotinum) dissolved 20000X was used for separation of all products. The gel was run at 120 V for 20 minutes or 80V for 40 minutes and the DNA was subsequently visualized under UV lamp (Herolab) and photographed by monochrome scientific grade camera Quantum ST4 (Vilber Lourmat).

3.4.1.4 DNA isolation from agarose gel

DNA extractions from agarose gel after separation via electrophoresis were performed using QIAquick Gel Extraction Kit (QIAGEN). Briefly, approximately 200 miligrams of the gel was cut out and dissolved in 600 µl of buffer QG at 56°C for 10 min. After mixing with 200 µl of isopropanol, the DNA was further purified from the solution using microtubes and solution from the kit precisely according to the manufacturer´s protocol. At the end, the extracted DNA fragments were eluted from the microtube with 40 µl of sterile HPLC water via centrifugation at 13000 g for 1 min at 25°C.

3.4.2 Protein analysis methods

3.4.2.1 Sodium dodecyl sulfate – polyacrylamide gel electrophoresis

The process of recombinant protein expression and subsequent protein purification was monitored by discontinuous SDS-polyacrylamide (SDS-PAGE) gel electrophoresis followed by silver staining of the proteins separated in the polyacrylamide gel. Analysis of tissue lysate content was mainly performed by discontinuous SDS-PAGE followed by Western blotting. All the samples for SDS-PAGE were collected in different amounts according to expected protein concentrations, fraction or lysate volumes, which is individually specified at the end of each purification or lysation method description.

Samples collected for SDS-PAGE were mixed with SDS-PAGE loading buffer of 360 mM Tris, pH 6.8, 30% glycerol, 10% SDS, 4% 2-mercaptoethanol, 0.01% bromophenol blue in 5:1 ratio, vortexed and boiled for 10 minutes.

Polyacrylamide gels consisting of upper 5% stacking gel (250 mM Tris-HCl pH 6.8, 5% (v/v) acrylamide solution (acrylamide with N,N'-bisacrylamide in a ratio 35.7:1), 0.1% (w/v) SDS, 0.1% (w/v) ammonium persulfate (APS), 0.02% (v/v) TEMED) and lower 10% or 18% resolving gel (375 mM Tris-HCl pH 8.8, 10% or 18% (v/v) acrylamide solution

(acrylamide with N,N'-bisacrylamide in the ratio 35.7:1), 0.1% (w/v) SDS, 0.1% (w/v) ammonium persulfate (APS), 0.01% (v/v) TEMED) were used. Protein samples from purification processes were loaded onto an 18% polyacrylamide gel, while 10% resolving polyacrylamide gels were used for further Western blotting.

Protein separation was performed in polyacrylamide gel immersed in SDS-PAGE running buffer 25 mM Tris pH 8.8, 250 mM glycine, 0.1% SDS in a vertical electrophoresis apparatus (Bio-Rad) at a constant voltage of 140V. All Blue pre-stained protein standard (Bio-Rad) was used as a molecular weight marker. Time of separation was different for individual methods, as required: approximately 1.5 hour for analysis of purification process or until the bromophenol blue dye/25kDa band of All Blue pre-stained protein standard (Bio-Rad) reached the bottom of the gel in case of tissue lysate analysis.

3.4.2.2 Silver staining of proteins in polyacrylamide gel

After termination of protein separation in a polyacrylamide gel, the gels were fixed with 12% (v/v) acetic acid, 50% (v/v) methanol and 0.02% (v/v) formaldehyde for 30 minutes while constant gentle shaking. After washing 3 times with 50% (v/v) methanol for 5 minutes, the gels were exposed with 0.02% (w/v) sodium thiosulfate pentahydrate for 1 minute and rinsed 3 times with distilled water. The impregnation was performed with 0.2% (w/v) silver nitrate and 0.02% (v/v) formaldehyde for 20 minutes, followed by rinsing with distilled water 3 times. The gels were developed by incubation with 566 mM sodium carbonate, 16 μM sodium thiosulfate pentahydrate and 0.02% (v/v) formaldehyde while gentle shaking, until the protein bands became visible. After rinsing in distilled water, development was stopped by a 10-minute incubation with 12% (v/v) acetic acid and 50% (v/v) methanol. Resultant polyacrylamide gels were scanned for further analysis on a scanner (Canon).

3.4.2.3 Western blotting analysis

Protein expression in embryos and cell lines was monitored using Western blot analysis. Individual sample preparation, processing and sample loads onto polyacrylamide gels is described in chapters dedicated to individual above mentioned experiments.

In general, sample (tissue, embryo, cells) lysate after protein concentration measurement was mixed with lysis SDS-PAGE loading buffer (for content see chapter 3.4.2.1) in ratio 5:1, vortexed and boiled for 10 minutes. SDS-PAGE procedure was performed identically according to the protocol described in Chapter 3.3.2.1. Wet electroblotting was performed in the Mini Trans-Blot® Cell (Bio-Rad) in 12.5 mM Tris-glycine pH 8.3, 10% (v/v) methanol at 100 V for 1 hour. Proteins were transferred onto

nitrocellulose membrane (Bio-Rad). The membranes were blocked with 5% dry milk (Bio-RAD) in TBST′ (50 mM Tris-HCl, 150 mM NaCl, 0.1% Tween 20) with 0.01% NaN3 by overnight incubation at room temperature. All the following incubations with antibodies were performed while gentle shaking at 4°C. All the primary antibodies were diluted in the same blocking solution and incubated with the membrane for at least 4 hours. The membrane was then gently rinsed 3 times with TBST′ and incubated with secondary antibodies IRDye® 800 CW Goat anti-Rabbit IgG (LI-COR) or IRDye® 680 RD Goat anti-Mouse IgG (LI-COR) diluted 1:30000 in 5% dry milk (Bio-Rad) in TBST′ with 0.01% NaN3 for 2 hours. The membrane was then washed 3 times in TBST′ (approx. 5 minutes), dried and photographed on Odyssey® CLx Imaging System (LI-COR). Scanned picture was processed in Image Studio Lite Software (LI-COR).

List of primary antibodies used in this work: Rabbit polyclonal anti-Ddi2 antibody (Bethyl, A304-630A, dilution 1:1000), Mouse monoclonal Anti-α-Tubulin antibody (Sigma-Aldrich, T6199, dilution 1:2000), Rabbit monoclonal anti- TCF11/NRF1 (Cell Signaling, D5B10 #8052, dilution 1:1000), Mouse monoclonal Anti-β-Actin antibody, clone AC-15 (Sigma-Aldrich, dilution 1:5000).

3.4.2.4 Protein concentration determination using Bradford protein assay

Bradford protein assay (Bradford M. M., 1976) was used for determination of protein concentration in tissue or cell lysates and in samples acquired during individual protein purification steps. The measurement was performed in a 96-well plate format with total reaction volume of 200 µl per well. The calibration curve consisted of duplicates of Albumin standard (2 mg/ml, Thermo Scientific) dilutions at concentrations 0, 12.5, 25, 50, 100 and 200 mg/ml. To acquire most accurate results, measured samples were diluted so that the absorbance of each reached approximately the middle of the calibration curve. Each well contained 20 µls of Albumin standard or appropriately diluted sample and 180 µls of Protein Assay Dye (Bio-Rad) 4.5× diluted in deionized water. The 96-well transparent flat bottom plate (P-lab) was used for the measurement. After addition of all reagent, the plate was left at room temperature for 5 minutes and then the absorbance at 595 nm was measured using the Infinite microplate reader (Tecan). Final protein concentrations were calculated using Microsoft Excel software.

3.4.3 Recombinant preparation of proteins

All the protein constructs were recombinantly expressed in host bacterial strain *E. coli* BL21(DE3)RIL (Novagen). Proteins with DNA construct cloned into pET16b (Novagen) bacterial expression vector were further purified via nickel affinity chromatography and size-exclusion chromatography as described in chapters 3.4.3.3 and 3.4.3.4, respectively. Mouse proteins that were cloned into p905 bacterial expression vector (gift from Dr. Řezáčova laboratory at IOCB CAS, Prague) were additionally submitted for cleavage with TEV protease for removal of the His-tag during the purification process as more closely described in chapter 3.4.3.5. Non-tagged ubiquitin was recombinantly expressed and provided by Michal Svoboda, all the remaining proteins were expressed and purified by Monika Sivá with help of Iva Flaisigová.

All the collected fractions and samples were analyzed during each purification step by SDS-PAGE electrophoresis (see chapter 3.4.2.1.) followed by silver staining (chapter 3.3.2.2). Each individual method used during particular purification process is described in the following chapters.

3.4.3.1 Recombinant expression of proteins in *E. coli*

After transfection and overnight growth, freshly grown bacterial colonies were suspended in 10 µls of LB medium (Sigma-Aldrich) and then inoculated into 3 liters of the LB media supplemented with ampicillin (100 mg/ml). The cell culture was grown in a-rotatory incubator at 37°C and 220 rpm. Protein expression was induced at OD595 approximately 0.8 by final concentration of 0.75 mM IPTG. Culture was further grown at 20°C and 220 rpm overnight. Bacterial cells were harvested by centrifugation at 6000 g, 10 min, 10°C. The pellet was resuspended in 50 mM Tris pH 8.0, 50 mM NaCl, 1 mM EDTA buffer supplemented with cOmplete™ Mini, EDTA-free Protease Inhibitor Cocktail (Roche) and further homogenized at 1200 bar at 4°C using EmulziFlex-C3 homogenizer (AVESTIN, Canada). After centrifugation at 20 000 g, 20 min, 4°C, the supernatant was decanted and used for further purification process.

3.4.3.2 Recombinant expression of isotopically labeled proteins for NMR

Freshly grown bacterial colonies after transfection and overnight growth were suspended in 10 µl of minimal medium (M9) and then inoculated into 2 liters of minimal media supplemented with ampicillin (100 mg/ml). The minimal medium recipe has been previously described (Renshaw P. S. *et al.*, 2004, Veverka V. *et al.*, 2006). Depending on the

required isotope labeling, the medium contained 0.8 g/L [^{15}N] ammonium chloride and/or 2 g/L d-[^{13}C] glucose. Cell culture was grown in a rotatory incubator at 37°C and 220 rpm. Protein expression induction, bacterial cell growth, harvest and homogenization was performed precisely as described in chapter 3.4.3.1. After centrifugation at 20 000 g, 20 min, 4°C, the supernatant was used for further protein purification.

3.4.3.3 Nickel affinity chromatography (proteins expressed from pET16b vector)

Proteins were further purified from the supernatant by nickel affinity chromatography. Supernatants were incubated with 0.5 -1 ml of equilibrated Ni-NTA resin (QIAGEN) on a rocker at 4°C overnight. The suspense was centrifuged at 3000 g, 4°C for 5 minutes; the supernatant was decanted and stored for another round of nickel affinity chromatography. The resin was washed 3 times by addition of 4 mls of wash buffer (50 mM Tris pH 8.0, 50 mM imidazole) and subsequently centrifuged (3000 g, 4°C, 5 min). His-tagged proteins were then eluted with 4 ml of elution buffer of higher content of imidazole (50 mM Tris pH 8.0, 250 mM imidazole) rocking 1 hour at 4°C. The supernatant after final centrifugation at the same conditions as previously was further used in the purification process.

3.4.3.4 Size-exclusion chromatography

Eluted fractions from nickel affinity chromatography were dialyzed against 50 mM sodium phosphate pH 7.4 supplemented with 0.5 % (v/v) glycerol. Any precipitates were removed by centrifugation (4000 g, 4°C, 10 min) and the supernatant was concentrated with Amicon Ultra centrifugal filters (Millipore) up to 10 – 20 mg/ml and filtered with a 0.22 μm filter unit (Millipore) prior application onto chromatography column. The concentrates were purified by size-exclusion chromatography on an FPLC (ÄKTA explorer, Amersham Pharmacia Biotech) using either the Superdex^{TM} 75pg 16/60 or 200pg 16/60 FPLC columns (GE Healthcare) depending on molecular mass of individual protein constructs. Protein concentration of all the collected relevant fractions was determined by Bradford protein assay (see chapter 3.4.2.4).

3.4.3.5 Purification process of proteins expressed from p905 vector

Mouse Ddi2^{WT} and Ddi2^{PD} proteins that were cloned into p905 bacterial expression vector (gift from Dr. Řezáčova laboratory at IOCB CAS, Prague) were purified with a different purification protocol. The recombinant expression and first round of nickel affinity

chromatography were performed identically as described in chapters 3.4.3.1 and 3.3.3.3. Histidine-tag was cleaved off the mDdi2 proteins by TEV protease (1 mg/ml) during an overnight incubation at 4°C in 35:1 ratio of protein:TEV protease (TEV protease was recombinantly prepared in our laboratory by Jaroslav Kurfürst in a construct with N-terminal histidine tag). After overnight incubation, second round of nickel affinity chromatography was performed. Due to cleaved-off N-terminal histidine tag and His-tagged TEV protease bound to the Ni-NTA resin, Ni-NTA purification provided flow through fraction with mDdi2 protein construct only. The flow-through was further used for dialysis for buffer exchange and size-exclusion chromatography as described above.

3.4.4 NMR experiments and biophysical characterization of proteins

All the spectra were acquired from 350 µl protein samples at 25°C on a 600 MHz or 850 MHz Bruker Avance spectrometer (Bruker BioSpin GmbH) equipped with a triple resonance (^{15}N, ^{13}C, ^{1}H) cryoprobe. 1D protein spectra were measured in 50 mM phosphate buffer at pH 7.4 with 0.5% glycerol. The 2D and 3D NMR spectra were acquired from protein samples in identical phosphate buffer with deuterium/hydrogen content of 5% D_2O/95% H_2O if not stated otherwise. All the spectra were processed using Topspin 3.2 (Bruker).

3.4.4.1 One dimensional NMR spectroscopy

The one dimensional 1H HSQC spectra of all the protein constructs used in NMR experiments and in characterization of murine mDdi2^{WT} and mDdi2^{PD} proteins were acquired at concentrations from 50 to 100 mM, which varied according to the concentration used later in other experiments.

3.4.4.2 NMR spectra acquisitions and spectra assignment for structure determination

The spectra for yeast Ddi1 ubiquitin-like domain structure determination were acquired from samples of 0.5 mM ^{13}C/^{15}N-labeled protein as described previously (Renshaw P. S.*, et al.*, 2004, Veverka V.*, et al.*, 2006). ^{15}N/^{1}H HSQC, HNCO, HNCACB, CACB(CO)NH spectra were collected for sequence-specific backbone assignment. Aliphatic side-chain carbon resonances and corresponding protons were assigned from HCCH-TOCSY, ^{15}N-edited NOESY, ^{13}C-edited NOESY and ^{15}N-edited TOCSY spectra. Aromatic ring proton resonances were assigned from 2D-TOCSY and 2D-NOESY spectra. 3D ^{15}N/^{1}H NOESY-HSQC and ^{13}C/^{1}H HSQC-NOESY spectra were used for ^{1}H-^{1}H distance constraints calculations. All the 2D and 3D NOESY spectra were acquired with NOE mixing

time of 120 ms and the TOCSY spectra were acquired with mixing time of 60 ms. Raw data were processed for further calculations by Dr. Václav Veverka at the IOCB CAS. All the resonance assignments were carried out manually in program Sparky (Goddard T. D. a. K., D. G. , 2008).

3.4.4.3 Protein structure calculations

Preliminary structures were produced from NOE-derived restraints from 3D ^{15}N- and ^{13}C-edited NOESY spectra. First of all, the family of converged structures for yeast Ddi1 UBL domain was calculated using Cyana 2.1 program (Guntert P. *et al.*, 1997, Herrmann T. *et al.*, 2002). Secondly, program TALOS+ was used for generation of backbone torsion angle constraints from assigned chemical shifts (Shen Y. *et al.*, 2009). The calculations additionally included hydrogen bond constraints. These involved residues with slowly exchanging amide protons. Simulated annealing combined with redundant dihedral angle constraints (REDAC) was performed in five cycles (Guntert P. and Wuthrich K., 1991), which produced a set of 43 converged structures with no eminent restraint violations (distance constraint violations and van der Waals violations below 0.2Å, dihedral angle constraint violations below 5°) and the lowest Cyana target function. This structure set underwent further refinement in explicit solvent with YASARA forcefield. Final analysis of the family of structures was carried out using the programs Molmol, iCING and PyMol (Doreleijers J. F. *et al.*, 2012, Koradi R. *et al.*, 1996, Schrodinger, LLC, 2015).

3.4.4.4 Characterization of protein-protein interaction using NMR

First, a series of double and triple resonance spectra were collected for sequence-specific backbone assignment of hDDI2 UBL, hDDI2 RVP full-C protein construct and UBQ. All spectra were processed using the program Sparky (Goddard T. D. a. K., D. G. , 2008). Changes induced in the positions of backbone signals of ^{15}N-labeled proteins in ^{15}N/^{1}H HSQC spectra were used for monitoring the interaction site. Most significant shifts of the backbone amide groups of individual residues were used for binding site mapping. The formula $\Delta\delta = \sqrt{(\Delta\delta_H)^2 + (\Delta\delta_N \times 0.2)^2}$ was used for definition of weigthed-average chemical shift perturbations. A non-linear one site specific binding model was used for titration curve fitting in program GraphPad Prism. In case of single addition of the binding partner, the minimal shift approach was used for assessment of the changes in signals (Veverka V. *et al.*, 2008).

The UBQ/hDDI2-UIM peptide titration experiment was performed by acquisition of 2D HSQC spectra of 0.1 mM UBQ in 50 mM sodium phosphate buffer, pH 6.0, containing

3.9% DMSO without or with addition of hDDI2-UIM peptide to final concentrations of 0.69, 1.4, 2, 2.75 and 3.45 mM in individual samples. Cut-off for evaluation of the CSPs used for this experiment was set to 0.12. 6 best-fitting curves of shift changes for individual residues were used for calculation of the K_d. Control binding experiment with DDI2-scrambled UIM peptide was performed under the same conditions with 1.9 mM peptide (final concentration). Both peptides were synthesized, purified by reverse-phase HPLC and subsequently lyophilized in the core facility of IOCB CAS, Prague. They were dissolved in DMSO prior titration experiments.

The protein-protein interactions were characterized from changes in $^{15}N/^{1}H$ HSQC spectra of ^{15}N-labeled 0.1 mM UBQ in 50 mM sodium phosphate buffer, pH 7.4, without and with 1, 2, and 5-fold molar addition of non-labeled hDDI2 RVP full-C and *vice versa*.

The interaction of hDDI2 UBL with UBQ, 2D HSQC spectra of 0.042 mM ^{15}N-labeled DDI2 UBL were acquired without and with 1, 2, 4, 6, 8, and 10-fold molar addition of non-labeled UBQ in 20 mM phosphate buffer, 0.5 mM DTT. Numerical cut-off for evaluation of the CSPs was set to 0.075. The 10 best-fitting curves from shifts of individual residues throughout titration were used for calculation of the K_d. Reverse experiment was performed with 0.05 mM UBQ with single 6-fold molar addition of hDDI2 ΔUIM protein construct in 50 mM phosphate buffer, 0.5 mM DTT. Spectra acquisition of 0.05 mM UBQ without and with 6-fold addition of DDI2 HDD-RVP construct was used as a control experiment.

To identify intramolecular interaction of hDDI2 UBL domain and its flexibility in full-length protein structure, two measurements were performed: 2D HSQC spectra of 0.05 mM full-length hDDI2 protein and hDDI2 UBL domain were acquired individually and superimposed. Intramolecular DDI2 UBL/DDI2 UIM interaction was verified by acquisition and superimposition of 2D HSQC spectra of 0.22 mM ^{15}N-labeled hDdi2 full-length protein and 2D HSQC spectra of 0.093 mM hDDI2 ΔUIM in 50 mM sodium phosphate buffer, pH 7.4, with 0.1 mM DTT. Interaction of hDDI2 UBL with hDDI2-UIM peptide was studied by acquisition of 2D HSQC spectra of 0.05 mM protein in 50 mM sodium phosphate, pH 7.4, with addition of hDDI2-UIM peptide to a final concentration of 1.9 mM. A control experiment with single addition of hDDI2-scrambled UIM peptide reaching 1.2 mM final concentration was performed.

3.4.4.5 Thermofluor assay - differential scanning fluorimetry (DSF)

Murine Ddi2^{WT} and Ddi2^{PD} proteins were examined for their proper folding and stability using thermofluor assay on a LightCycler® 480 II (Roche). The measurement was performed with 20 μM proteins on a 96-well plate in 50 mM phosphate buffer at pH 7.4 with 0.5% glycerol with 5000x diluted SYPRO® Orange protein gel stain (Sigma-Aldrich) in total reaction volume 25 μl. The protocol was set as follows: after the pre-cooling temperature was held at 20°C for 10 min, the fluorescence itself was measured during continuous temperature increase from initial 20°C up to final 95°C with 82 acquisitions per 1°C accrue. The temperature increment was set at rate 0.01°C per second. The experiment was completed at 95°C for 10 s followed by 20°C for 10 s. LightCycler® 480 Software (Roche) was used for final calculation of the melting temperatures. Protein stability analysis was performed in collaboration with Michal Svoboda.

3.4.5 Methods linked to studies in mice and cell cultures

3.4.5.1 Generation of *Ddi2^{tm1b}* and *Ddi2$^{protease\ defective}$* mouse strains

Ddi2^{tm1b} (full name: C57BL/6NCrl-Ddi2$^{tm1b(EUCOMM)Hmgu/Ph}$) mouse strain was generated on the C57BL/6NCrl background at the IMG CAS and registered for phenotyping at the International Mouse Phenotyping Consortium (IMPC). The origin of the line is from ES clone HEPD0660_5_E02, which belongs under The European Conditional Mouse Mutagenesis Program (EUCOMM). This ES cell clone bears a *Ddi2$^{tm1a(EUCOMM)Hmgu}$* cassette with promoter-driven neomycin selection. We started to work with this mouse line in collaboration with Czech Centre for Phenogenomics hosted by the IMG CAS. Colony of *Ddi2^{tm1b}* strain was established by scientific staff of the IMG CAS.

Ddi2$^{protease\ defective}$ (full name: C57BL/6NCrl-Ddi2$^{em1/Rase}$) strain was generated by Transcription Activator-Like Effector Nuclease (TALEN)-mediated genome editing performed by Dr. Petr Kašpárek at the IMG CAS. TALENs specifically recognizing the intron 5 and intron 6 of Ddi2 gene were designed using TAL Effector Nucleotide Targeter 2.0 (Cermak T. *et al.*, 2011, Doyle E. L. *et al.*, 2012). TALENs were assembled using the Golden Gate Cloning system (Cermak T.*, et al.*, 2011) and cloned into the ELD-KKR backbone plasmid as described elsewhere (Flemr M. *et al.*, 2013). TALENs recognizing 5' site and 3´ site of target sequence within intron 5 contained the following repeats: HD NG NG HD NI HD NG NN NN NN NN HD NI NN HD NN NG and HD HD NI HD HD NI NI HD NI NN NI NI NI NI NG, respectively. TALENs recognizing target sequence within intron

6 contained the following repeats: for 5' site NN NG NN NG HD HD NG NG NN NG NN NG NI HD NN NN NN and for recognizing 3' site HD HD HD HD NI NN NG NN HD NG NN HD HD HD NG HD NG NN. Each plasmid was linearized with NotI and transcribed using the mMESSAGE mMACHINE T7 kit (Thermo Fisher Scientific). Polyadenylation of resulting mRNAs was performed using the Poly(A) Tailing kit (Thermo Fisher Scientific) and mRNA was purified with the RNeasy mini kit (QIAGEN). TALEN mRNAs, 10 ng/µl for each TALEN were microinjected into male nucleoli of zygotes isolated from C57BL/6N mice (Kasparek P. *et al.*, 2014). The mice were further maintained on C57BL/6N background.

All work with mice was approved by the Animal Care Committee of the IMG CAS according to institutional and national guidelines of Czech Central Commission for Animal Welfare and in accordance with European directive 2010/63/EU.

3.4.5.2 Establishment of *Ddi2*$^{protease\ defective}$ mouse colony, colony management and timed crossings

Four founder mice (F0) bearing successful exon 6 deletion were identified and the deletion region was sequenced by GATC Biotech (Konstanz, Germany). Genomic DNA (gDNA) extraction and subsequent deletion region amplification, agarose gel separation and gel extraction were performed as described in chapters 3.4.5.4. All four F0 mice were bred with C57Bl/6NCrl wild-type (WT) mice for production of F1 generation. After analysis of F1 generation by sequencing of deletion region of gDNA, only offspring of founder ID 38 were chosen for colony establishment. Mouse colony was established at F1 generation by Dr. Kašpárek and it was onward maintained on C57BL/6NCrl background by Monika Sivá. Mice were bred to F4 and F5 generation according to a breeding scheme where heterozygous mice were backcrossed to C57Bl/6NCrl wild-type mice. Heterozygous mice from F4 and F5 generation were bred with heterozygous mice to obtain offspring for experiments (embryo or adult mice). C57Bl/6NCrl WT mice were provided by the Animal Facility of IMG CAS.

Breedings for timed embryo sample collection were precisely planned. After breeding setup, females were checked for plug every morning. The noon of the day of vaginal plug detection was estimated as embryonal day 0.5 (E0.5). Additionally, the females were weighted on the day of breeding onset, on the day of plug and after 6 and 8 days from plug date by specially trained staff at the Animal Facility of IMG CAS. Plugged females were considered pregnant individually according to weight gain.

3.4.5.3 Off-target screen of TALEN-mediated gene alterations

A cohort of F1 generation mice used for further colony establishment were screen for off-target sites of TALENs used for *Ddi2*$^{protease\ defective}$ strain generation. Twelve sites within chromosome 4 with higher off-target score for all possible combinations of ELD and KKR heterodimers of TALENs were predicted in TAL Effector Nucleotide Targeter 2.0 (Doyle E. L.*, et al.*, 2012). These sites were analyzed for protein coding sequences using Ensembl online database. They were amplified by PCR from genomic DNA extracts (chapter 3.4.5.4) of F1 generation mice and sequenced by GATC Biotech (Konstanz, Germany). All primers were designed and examined using online tools Ensembl, UCSC in silico PCR and NCBI PrimerBlast and BLAST (Geer L. Y.*, et al.*, 2010, Kent W. J. *et al.*, 2002, Zerbino D. R.*, et al.*, 2018). Their sequences are listed in Table 2. Monitored DNA fragments were amplified in PCR reaction using Phusion® High-Fidelity DNA Polymerase (New England BioLabs). The reaction mix contained 10 µl of 5X Phusion® HF Buffer, 200 µM dNTPs (Serva), 0.4 µM primers (see Table 2), 1 µl of genomic DNA extract, 0.2 µl of Phusion® High-Fidelity DNA Polymerase and nuclease-free water addition up to 50 µl of total volume. PCR Thermocycler (Biometra) was used for amplification with cycles programed as follows: DNA template denaturation at 98°C for 5 minutes; 33 cycles of denaturation at 98°C for 20 seconds, annealing at 58-65°C (depending on melting temperatures of primer pairs designed for each off-target site) for 20 seconds and elongation at 72°C for 30 seconds; and final incubation at 72°C for 5 minutes. Amplified fragments were separated by agarose gel electrophoresis (chapter 3.4.1.3), extracted from gel (chapter 3.4.1.4) and sequenced using the same primers that were used for amplification by GATC Biotech (Konstanz, Germany).

3.4.5.4 Genomic DNA extraction and genotyping of *Ddi2*$^{protease\ defective}$ and *Ddi2*^{tm1b} mouse strains

Tail or ear biopsies were taken from three-week-old pups at weaning at the Animal Facility of the IMG CAS. Yolk sac or embryonal body samples for genotyping were collected during embryo harvest and dissection. Genomic DNA was isolated using DirectPCR Lysis Reagent (Viagen). Each mouse tail and embryo sample was immersed with 50 µl or 30 µl of DirectPCR and 0.5 µl or 0.25 µl of 10 mg/ml Proteinase K (New England Biolabs), respectively. After overnight incubation at 55 °C, the activity of Protease K was stopped by 5x dilution with sterile HPLC water.

Genomic DNA extract was subsequently used as template for PCR using Mouse Direct PCR Kit (Bimake). Genotyping of *Ddi2*$^{protease\ defective}$ strain mouse and embryo samples

was performed with a pair of Ddi2F and Ddi2R primers amplifying 1572 bps for WT allele and 1072 bps for exon 6 deleted allele form. PCR reaction mix contained 10 µl of M-PCR OPTI mix (component of Mouse Direct PCR Kit containing Taq DNA polymerase, dNTPs and reaction buffer), final concentration of 0.25 µM Ddi2F and Ddi2R primers, 1 µl of template of genomic DNA and water addition up to total volume of 20 µl. PCR cycles were preset and run on a PCR Thermocycler (Biometra) as follows: DNA template denaturation at 94°C for 3 minutes; 35 cycles of denaturation at 94°C for 20 seconds, annealing at 55°C for 30 seconds and elongation at 72°C for 60 seconds; and final incubation at 72°C for 5 minutes.

Two forward primers LacZ F, Ddi2^{tm1b} WT F and one reverse primer Ddi2^{tm1b} RV were used for genotyping of *Ddi2^{tm1b}* samples. PCR reaction mix contained 10 µl of M-PCR OPTI mix (component of Mouse Direct PCR Kit containing Taq DNA polymerase, dNTPs and reaction buffer), final concentration of 0.25 µM LacZ F and Ddi2tm1b WT F primers, final concentration of 0.375 µM Ddi2tm1b RV primers, 1 µl of template of genomic DNA and water addition up to total volume of 20 µl. The PCR reaction was performed as follows: DNA template denaturation at 94°C for 5 minutes; 35 cycles of denaturation at 94°C for 30 seconds, annealing at 55°C for 40 seconds and elongation at 72°C for 75 seconds; and final incubation at 72°C for 5 minutes. The set of 3 primers amplified fragments of 1080 bps for *Ddi2* wild-type allele and 640 bps for *Ddi2^{tm1b}* allele.

Thanks to M-PCR OPTI mix content, 2 µl of the PCR mix could be used for analysis of the amplification by agarose gel electrophoresis (chapter 3.4.1.3) with no need for loading buffer. The gel was run at constant voltage of 80V for 40 minutes. DNA fragments were visualized and analyzed under UV lamp (Herolab). Gel was photographed by monochrome scientific grade camera Quantum ST4 (Vilber Lourmat).

3.4.5.5 Genotyping of early embryonal stages using nested PCR

In case of early developmental stage *Ddi2$^{protease\ defective}$* strain embryo genotyping (9.5), another round of nested PCR was introduced to avoid contamination of the tissue samples by maternal blood or tissue during embryo dissection and MEF culture isolation.

Three primers were designed for two nested reactions: a forward primer Ddi2 nested F, Ddi2 long R and Ddi2 IN R reverse primers. The pair of Ddi2 nested F and Ddi2 long R primers in NESTED 1 reaction amplified fragments of 650 bps for WT allele and 150 bps for exon 6 deleted allele, while the pair of Ddi2 nested F and Ddi2 IN R in NESTED 2 reaction amplified a 350 bp long fragment in case of WT allele, however no DNA amplification product in case of exon 6 deletion.

Both NESTED reaction mixtures were prepared separately as follows: 10 µl of M-PCR OPTI mix (Bimake), final concentration of 0.25 µM primers, 1 µl of 200x diluted *Ddi2*$^{protease\ defective}$ genotyping PCR reactions after amplification from chapter 3.4.5.4 as template and water addition up to total volume of 20 µl. The PCR reaction was performed similarly to genotyping in chapter 3.4.5.4, as follows: DNA template denaturation at 94°C for 5 minutes; 27 cycles of denaturation at 94°C for 20 seconds, annealing at 55°C for 30 seconds and elongation at 72°C for 40 seconds; and final incubation at 72°C for 5 minutes. Amplified fragments were analyzed by agarose gel electrophoresis as described in chapter 3.4.1.3.

3.4.5.6 Embryo harvest

Pregnant mouse was sacrificed by cervical dislocation. The uterus was collected, washed from blood in sterile preheated (37°C) PBS in a 50 ml falcon tube and placed onto 100 mm Petri dish with preheated sterile PBS. Individual embryos were withdrawn from uterus with a pair of Dumont micro forceps (Fine Science Tools) under stereomicroscope Stemi 305 EDU Microscope (Zeiss). Yolk sac and amnion were separated from the embryonal body and stored as sample for genotyping. Embryo proper was either put into a pre-tarred microtube and frozen at -80°C for protein expression analysis, or fixed in 4% PFA for further RNA *in situ* hybridization experiments, or put into 500 µl of RNAlater solution (Ambion) for gene expression studies using qPCR. The sample for mRNA isolation was left 2 hours on the table to soak and then put to fridge for further soaking and short time storage. Whole-mount imaging of embryos was performed by Kallayanee Chawengsaksophak.

3.4.5.7 Isolation and culturing of primary mouse embryonal fibroblasts

The pregnancy of mice was monitored by weight gain: females were weight on the day of start of the breeding, on the day of plug and after 6 and 8 days from plug date. In case the female gained over 2 grams, it was considered pregnant. Mouse embryonic fibroblasts (MEFs) were isolated from embryos on embryonal day 10.5. The MEF medium used for culturing of primary cell line consists of DMEM High Glucose w/o L-Glutamine (Biosera), 10% heat-inactivated fetal bovine serum (Sigma-Aldrich), 2 mM L-glutamine (Sigma-Aldrich), 100X Penicillin-streptomycin (Sigma-Aldrich) and freshly added 100X MEM non-essential amino acid solution (Sigma-Aldrich) and 100X 100 mM sodium pyruvate (Sigma-Aldrich).

Embryos at specific stage of embryonal development were harvested as decribed in chaper 3.4.5.6. Trunk of each embryonal body was washed from organs and placed into individual well with 100 µl of sterile preheated PBS in a 24-well plate. Yolk sac and head

were separated from the embryonal body, washed in PBS and collected into sterile microtube as samples for genotyping. Trunk of each embryo was lysed separately by pipetting at least 5 times up and down with a 200 µl tip, followed by a thinner 20 µl tip and eventually lysed in a 26G gauge with 1ml injection. The 100 µl lysate was then transferred into 1ml of freshly prepared MEF medium in a well in 12-well plate (pre-coated with 0.1 % gelatin). Media was completely changed 2 days after isolation.

Medium was further exchanged every second day in ratio of original:new medium 1:3. The culture was passaged at minimum of 90% confluence onto plates pre-coated with 0.1 % gelatin as follows: P0 – isolated cells, P1 – from a well of 12-well plate onto one 60 mm Petri dish, P2 – from one 60 mm Petri dish onto 2x 100 mm Petri dish, P3 – one 100 mm Petri dish onto four 100 mm Petri dishes. For passage, MEF medium was withdrawn, cells were washed with preheated PBS and trypsinized twice with relevant amount of trypsin-EDTA solution for 2 minutes. Trypsinization was stopped with addition of MEF medium, cells were centrifuged at 250 g for 3 minutes, resuspended in MEF medium and placed onto a Petri dish. Final passage 3 was used for experiments.

3.4.5.8 Messenger RNA isolation

All the samples were individually weighted in sterile RNase free tared tubes. Messenger RNA isolations were performed using RNeasy Plus Micro Kit (QIAGEN) or RNeasy Mini Kit (QIAGEN), depending on whether the weight of sample was below 5 mg or 30 mg, respectively. Tissue samples were homogenized in 80 µl (RNeasy Plus Micro Kit) or 350 µl (RNeasy Mini Kit) of RLT Buffer with freshly added 1% β-mercaptoethanol using TissueLyzer II (Retsch) at frequency 30 Hz (1800 oscillations per minute) for 3 minutes. An RNase free iron ball was placed into the tubes for proper homogenization. Following procedures were performed according to further manufacturer´s instructions in both kit cases. DNase digestion was performed on gDNA Eliminator Mini Spin Columns (QIAGEN) when RNeasy Plus Micro Kit was used. In case of RNeasy Mini Kit, the RNase-free DNAse set (QIAGEN) was used according to manufacturer´s guide for on-column DNA digestion during RNA isolation. Purified RNA was eluted from columns with RNase-free water (QIAGEN) and stored at -80°C for further processing.

3.4.5.9 Quality control of mRNA and reverse transcription

The quality of isolated mRNA was characterized using Bioanalyzer RNA 6000 Nano assay (Agilent). Individual mRNA samples were diluted to reach concentrations (5–500 ng/µl) valid for measurement of RNA integrity number (RIN) in Total NANO RNA assay.

Twelve mRNA samples were run on each chip. Cleaning and setting up of the instrument and the experiment itself were performed exactly as described in manufacturer´s instructions in Agilent RNA 6000 Nano Kit Guide. Only RNA samples with RIN above 7 were used further in gene expression studies (chapter 3.4.5.10) and cloning of Ddi2 gene variants (chapter 3.4.5.14).

Messenger RNA was transcribed to cDNA using a reverse transcription set TATAA GrandScript cDNA Synthesis Kit (TATAA Biocenter). Individual reaction were performed according to manufacturer´s instructions in total volume of 20 µl with addition of 0.25 µg of mRNA. A reverse transcription negative control was prepared for samples dedicated for gene expression studies with qPCR by replacement of enzyme with nuclease-free water in reaction mix.

3.4.5.10 Quantitative PCR

Primer design, validation and qPCR experiments were carried out by Eva Rohlová and Filip Franko from the Gene Core facility at the IBT CAS in the BIOCEV center. Sequences of primers used for housekeeping genes *Tbp* and *H2afz,* and for *Ddi2* are listed in Table 2 on page 69. Reactions were run in duplicates on 384 plates (including interpolate calibration) on LightCycler® 480 (Roche) using TATAA SYBR® GrandMaster® Mix (TATAA Biocenter). Templates were diluted 5X into the final reaction volume of 10 µl. Reaction was performed as follows: 95°C (1 min), 45 cycles of 95°C (5 s), 60°C (30 s) and 72°C (10 s), and a final temperature gradient from 60°C to 95°C for melting curve acquisition.

Analysis of the results was performed by Monika Sivá and Vendula Novosadová from CCP hosted by IMG CAS, in Excell (Microsoft Office). *Ddi2* expression was analyzed relatively to *H2afz* expression, which was identified as appropriate stable reference gene for this screen.

3.4.5.11 Preparation of tissue lysates

Embryos were lysed in 30 up to 80 µl of RIPA buffer (20 mM Tris pH 7.4, 150 mM NaCl, 1 mM EDTA, 1% deoxycholate) with addition of cOmplete™ Mini, EDTA-free Protease Inhibitor Cocktail (Roche), depending on the weight of individual embryos. Tissue was homogenized in 80 to 300 µl of RIPA (depending on embryo weight) with an iron ball using TissueLyzer II (Retsch) at frequency 30 Hz (1800 oscillations per minute) for 3 minutes. The homogenized suspense was diluted in RIPA buffer with addition of igepal to 1% and thoroughly mixed by pipetting up and down. Lysates were centrifuged at 16000 g for 15 minutes at 4°C. Supernatants were measured for protein concentration

using Bradford protein assay (chapter 3.4.2.4) and used for preparation of SDS-PAGE samples as described in chapter 3.4.2.1.

3.4.5.12 MEF treatment and harvest for gene expression studies

MEF cultures for protein expression analysis by Western blotting were harvested without treatment or with treatment of inhibitor bortezomib (UBPBio) for 16 hours with 10 μM, 2.5 μM Botezomib diluted in DMSO or DMSO itself for control. Inhibitor treatments were performed on MEF cultures at minimum of 90% confluency of passage 3 in duplicates. For analysis of gene expression by qPCR, incubations with 1 μM Botezomib, DMSO for 16 hours or no treatment were prepared in triplicates.

MEF cells were harvested as follows: MEF medium was withdrawn, cells were washed with preheated PBS and trypsinized twice with relevant amount of trypsin-EDTA solution for 2 minutes. Trypsinization was stopped with addition of MEF medium (for composition see chapter 3.4.5.7). Cell suspension was centrifuged at 1200 rpm for 3 minutes (in pre-tarred microtubes), the pellet was washed with 200 μl of PBS and centrifuged at 1200 rpm for 3 minutes twice, weight at analytical scales and frozen until lysation. Further processing for gene expression analysis by qPCR or Western blotting is described in chapters 3.4.5.8 – 3.4.5.10 and 3.4.2.3, respectively.

3.4.5.13 Preparation of cell lysates for Western blotting

Cell pellets of MEF or HEK293offA2 were lysed in SDS sample buffer without dye (60 mM Tris pH 6.8, 60 nM SDS, 0.3 mM β-mercaptoethanol) with cOmplete™ Mini, EDTA-free Protease Inhibitor Cocktail (Roche) and with 1 μl of (15x diluted) Benzonase® (Novagen) by pipetting up and down several times, left 40 minutes on ice, sonicated three times at maximal frequency for 20 sec in cold-water sonication bath (Elma) with 1 min pause on ice after each sonication. Pipetting up and down was used again for each sample, and then the lysates were centrifuged at 16000 g for 20 minutes at 4°C. Pellets were discarded and protein concentration was measured from supernatant by Bradford protein assay (chapter 3.4.2.4). SDS-PAGE loading samples were prepared by addition of SDS-PAGE loading buffer dye (see chapter 3.4.2.1) to lysate in ratio 5:1, vortexed and boiled for 10 minutes. For description of Western blotting procedure see chapter 3.4.2.3.

3.4.5.14 Analysis of *Ddi2*$^{exon6\ +/-}$ and *Ddi2*$^{exon6\ -/-}$ mRNA products

The sequence of truncated mRNA product after exon 6 deletion in F1 generation mice was verified by mRNA extraction from mouse ears biopsies collected at the mice facility.

Expression of protein versions in heterozygous and homozygous mice was verified on mRNA level in embryos of all three genotypes at stage E10.5. The collection of material, mRNA isolation, quality control and cDNA synthesis were performed as described in chapters 3.4.5.6, 3.4.5.8 to 3.4.5.9.

Protein coding sequences of mDdi2^{WT} and mDdi2^{PD} (*Ddi2*$^{exon6\ -/-}$ encoded product) were amplified by PCR using a pair of primers mDdi2 F cDNA and mDdi2 R cDNA, encoding 5´and 3´ends of wild-type Ddi2 gene. The reaction mix contained 10 µl of 5X Phusion® HF Buffer, 200 µM dNTPs (Serva), 0.4 µM primers, 1 µl of cDNA, 0.2 µl of Phusion® High-Fidelity DNA Polymerase and nuclease-free water addition up to 50 µl of total volume. PCR Thermocycler (Biometra) was used for amplification with cycles programed as described for Phusion® High-Fidelity DNA Polymerase in chapter 3.4.1.1 with annealing temperature at 63°C. Amplified fragments for all three genotypes were analyzed by agarose gel electrophoresis (chapter 3.4.1.3).

The mDdi2^{WT} and mDdi2^{PD} coding sequence fragments amplified from cDNA as described in chapter XX were extracted from agarose gel after analysis and cloned into pCR™-Blunt vector using Zero Blunt® Cloning® Kit (Invitrogen) according to manufacturer´s instructions. After ligation, the ligation reactions and negative control reaction were transformed into competent *E. coli* strain Top10 (Invitrogen) similarly as described in chapter 3.4.1.2. Next day, 5 colonies from each agar plate were picked, cultured with LB medium containing 40 µg/ml kanamycin overnight at 37 ºC and purified according to protocol described in chapter 3.4.1.2. Purified plasmids were sequenced by GATC Biotech (Konstanz, Germany) using primers M13 F and M13 R (for primer sequences see Table 2 on page 69).

3.4.5.15 Overexpression of mDdi2 variants in human HEK293 cells

Plasmids pTreTight encoding mDdi2 variants were amplified in DNA maxipreparation as described in chapter 3.4.1.2 and were transfected into HEK293offA2 cells containing Tet-Off expression system provided by laboratory of Dr. Jan Konvalinka. HEK293offA2 cells were cultured in 12 well plates with 2 ml of IMDM complete medium (IMDM (Thermo Fisher Scientific) supplemented by 10% FBS and 40 mM L-glutamine). They were transfected with 400 µl of transfecting mix consisting of opti-MEM medium (Thermo Fisher Scientific), 7.5 µg of plasmid DNA and 10% (v/v) polyethylenimine (Sigma-Aldrich) at confluence of approx. 70%. The cells were incubated at 37 ºC and 5% CO2 until harvested, which happened 4, 8, 16, 24, 32, 40 and 48 hours after transfection. Harvest was performed by resuspendation of the cells in the IMDM complete medium

followed by centrifugation for 3 minutes at 250 g and washing with sterile pre-heated PBS twice. Cells were frozen at –20°C until further use.

3.4.5.16 *In situ* hybridization studies

The adult mice were mated as described in chapter 3.4.5.2. Embryos at the embryonal stage E9.5, E10.5, E14.5 and E16.5 were collected from pregnant CD-1 female mice and fixed in 4% PFA as described in chapter 3.4.5.6. Embryos at the age of E9.5 and E10.5 were processed in both whole mount and paraffin sections forms. The samples for whole mount method were frozen in methanol at -20°C straight after PFA fixation. Whole embryos at the two earlier stages and heads of the latter two used for ISH on sections, were dehydrated, embedded in paraffin, cut to 7 µm sagittal sections and rehydrated as described in standard protocols (Wilkinson D. G. and Nieto M. A., 1993). Tissue hydration, proteinase K treatment, acetylation and the pre-hybridization procedures were performed with DEPC water as described previously (Wilkinson D. G. and Nieto M. A., 1993). The hybridization was performed with DIG labelled probes diluted in the hybridization buffer (1.25X saline-sodium citrate, pH 7.0, 50% formamide, 0.1% Tween-20, 100X Denhardt´s solution, heparin (50 µg/ml), tRNA (50 µg/ml), salmon sperm DNA (50 µg/ml)) overnight at 70°C. Samples were developed using anti-DIG antibody conjugated with alkaline phosphatase and BM purple alkaline phosphatase substrate precipitating solution (Roche). Prior imaging on Zeiss ApoTome microscope, all the samples were post-fixed with 4% PFA and the slides were in addition mounted in Aquatex. ISH experiments were performed by Monika Sivá and Michaela Procházková.

Probes for specific recognition of murine Ddi1 mRNA and control sense probes were designed from the coding sequence of mDdi1 (NM_027942.1), which was synthetized by GenScript (New Jersey, USA) and ligated into pCR™-Blunt vector using Zero Blunt® Cloning® Kit (Invitrogen) according to manufacturer´s instructions. Two sequences were amplified for preparation of two anti-sense and two corresponding sense probes with following primers: for Ddi1 probe set No.1 with length 1018 bps, Ddi1_F_AS1 and Ddi1_R_AS1 and for Ddi1 probe set No.2 with length 993 bps, Ddi1_F_AS2 and Ddi1_R_AS2. The amplification and PCR product purification were performed similarly as described in chapter 3.4.1.1. Both sequences were cloned individually into pGEM-T® easy plasmid (Promega) for blue/white colony selection and the white colonies were picked and cultivated as described in chapter 3.4.1.2 by our collaborator Michaela Procházková at the CCP at IMG CAS. Glycerol stocks were prepared from 225 µl of sterile 80% glycerol and 1

ml of LB broth with certain bacterial clone and stored at -80°C, the rest was purified with QIAprep Spin Miniprep Kit (QIAGEN) and sequenced by GATC Biotech (Konstanz, Germany).

The plasmid DNAs containing correct Ddi1 sequences were used for synthesis of probes. First, the DNA was amplified by bacteria cultivation in 100 ml of LB media with final concentration of ampicillin 100 µg/ml inoculated with bacterial glycerol stock and plasmid DNA purification as described in chapter 3.4.1.2. For preparation of anti-sense probes or control sense probes, 10 µg of both plasmid DNAs were digested with either SacII or SalI endonuclease, respectively. The restriction enzyme DNA cleavage was performed by incubation at 37°C for at least two hours in a mix relevant for the specific endonuclease as instructed by the manufacturer. Sequences of antisense probes are shown in Table 3 below.

Table 3: Sequences of anti-sense probes for ISH

NAME	PROBE SEQUENCE
Ddi1 AS1	5′-gtattgtgtgcgtagggacctcacagaggtaacctttccctccaggtcaaccctgactttgagctctcaaacttcagagtcctctgtgag cttgagtctggtgtgcctgctgaggaggcccagatcgtttacatggaacagctcctcacagatgaccactgctccctgggctcctatggact caaagatggtgacatggttgtacttcttcagaaggataatgttggacttcggactccaggaaggactccaaaccatcctcgagcagatttca ctgggtctgggtcggccgtgcctggaacatcaagttcccgacacccgcatcagcatcaacaccattatcaccatcatcaacgtataccatc aacacagcaagcccacggattagcctctggagagaatatgacctttgctcaggaactcgacagccctgccctgattcgaagcatgctgctt tccaacccccatgatctgtccctgttgaaggaacggaatcctgctttggctgaagctctgcttagtggaaaccttgagacattttctcaggtcc tgatggagcagcagagagaaaggaccttgagagagcaagagatgtttcgtctttactctactaacccatttgatcaggaaactcaggctag aatagaagaagaaatccgacaacagaatattgaagaaaatatgaacatagctatggaagaggctccggagagttttggacaagtcgctat gctctatattaactgcaaagtgaatgggcatcctttaaaggcttttgttgactctggtgcccagatgactatcatgagccaagcttgtgctgag agatgtaatattatgagactggtagaccgacggtggggtggggttgctaagggagtaggcacacagaggattatgggccgcgttcatctg gctcagattcaaattgaaggtgatttcttacagtgctctttctctatacttgaagagcagcccatggatatccttctagggcttgatatgctcagg aggcac-3′
Ddi1 AS2	5′-tgagcttgagtctggtgtgcctgctgaggaggcccagatcgtttacatggaacagctcctcacagatgaccactgctccctgggctcct atggactcaaagatggtgacatggttgtacttcttcagaaggataatgttggacttcggactccaggaaggactccaaaccatcctcgagca gatttcactgggtctgggtcggccgtgcctggaacatcaagttcccgacacccgcatcagcatcaacaccattatcaccatcatcaacgtat accatcaacacagcaagcccacggattagcctctggagagaatatgacctttgctcaggaactcgacagccctgccctgattcgaagcat gctgctttccaacccccatgatctgtccctgttgaaggaacggaatcctgctttggctgaagctctgcttagtggaaaccttgagacattttct caggtcctgatggagcagcagagagaaaggaccttgagagagcaagagatgtttcgtctttactctactaacccatttgatcaggaaactc aggctagaatagaagaagaaatccgacaacagaatattgaagaaaatatgaacatagctatggaagaggctccggagagttttggacaa gtcgctatgctctatattaactgcaaagtgaatgggcatcctttaaaggcttttgttgactctggtgcccagatgactatcatgagccaagcttg tgctgagagatgtaatattatgagactggtagaccgacggtggggtggggttgctaagggagtaggcacacagaggattatgggccgcg ttcatctggctcagattcaaattgaaggtgatttcttacagtgctctttctctatacttgaagagcagcccatggatatccttctagggcttgatat gctcaggaggcaccagtgttccatcgacctaaagaaaaatgtgctggtgattggcactaccggctcacagactca-3′

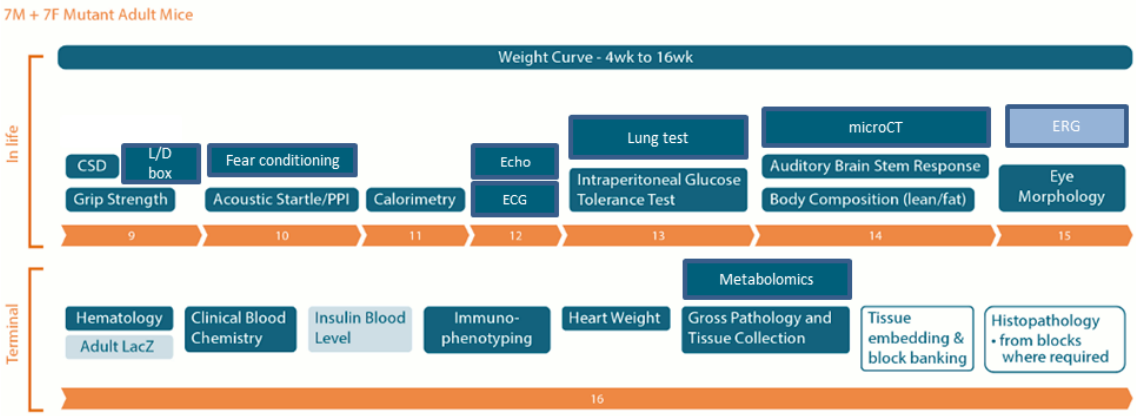
Each probe was prepared by *in vitro* transcription, which was performed in a mix containing 4 µl of 5x Transcription buffer, 2 µl 100 mM DTT, 2 µl of DIG RNA labeling mix, 1 µl RNase inhibitor, 2 µl relevant RNA polymerase, 1 µg of linearized DNA template and addition of sterile RNase-free water up to 20 µl (components of DIG RNA Labeling Kit (SP6/T7) from Roche). The mixture was incubated at 37°C for 4 hours and the probes were

purified using RNeasy Mini Kit (QIAGEN) according to the RNA cleanup protocol. Concentration of each probe was measured on NanoDrop ND-1000 spectrophotometer (Thermo Scientific).

Sp6 RNA polymerase was used for synthesis of anti-sense probes from the two DNA templates linearized by SacII endonuclease. T7 RNA polymerase was used for sense probes synthesis after linearization by SalI endonuclease. All four probes were used for ISH, however only data from probe Ddi1 AS1 are shown in Figure 13 in chapter 4.2.1.

3.4.5.17 Phenotyping of adult mice

Phenotyping of adult mice of both *Ddi2*^{tm1b} (C57BL/6NCrl-Ddi2$^{tm1b(EUCOMM)Hmgu}$/Ph) and *Ddi2*$^{protease\ defective}$ (C57Bl/6NCrl-Ddi2$^{em1/Rase}$) strains was performed at the CCP hosted by IMG CAS according to the international mouse phenotyping consortium (IMPC) pipeline workflow and standard operating procedures (for workflow see Scheme 1 below). Jan Procházka was in charge of the phenotyping team, Monika Sivá assisted during sample uptake of *Ddi2*$^{protease\ defective}$ strain mice.



Scheme 1: Pipeline workflow according to the international mouse phenotyping consortium. Adapted from https://www.mousephenotype.org/impress/procedures/44 (cited on February 20th 2019) (IMPC).

For phenotyping screen of *Ddi2*^{tm1b} strain, a cohort of 7 *Ddi2*$^{+/-}$ males and 8 *Ddi2*$^{+/-}$ females was used. The mice were studied for body composition and weight, behavioral tests, cardiovascular and lung function tests, glucose metabolism (IpGTT), hematology and biochemistry, gross pathology and histology after termination. The values and results were compared to results of phenotyping of a large (over 200 mice) C57Bl/6Ncrl cohort housed at the same facility.

Ddi2$^{protease\ defective}$ strain mice (8 of each sex and genotype available for adult mice) at the age of 16 weeks were subjected to screening of glucose metabolism (IpGTT),

biochemistry, hematology, gross pathology and histology. The results for *Ddi2*$^{exon6\ +/-}$ mice were compared with the values of *Ddi2*$^{exon6\ +/+}$ and C57Bl/6Ncrl in-house WT mice.

3.4.5.18 Mapping of *Ddi2* expression using LacZ staining

Embryos at the embryonal stage E9.5, E12.5, E14.5 and E17.5 were harvested (as described in chapter 3.4.5.6) from pregnant C57Bl/6Ncrl female mice mated with *Ddi2*$^{tm1b\ +/-}$ males or from *Ddi2*^{tm1b} heterozygous crosses.

Embryos for whole mount staining were first fixed in 4% PFA for 30 minutes and subsequently rinsed in 1 M phosphate buffer pH 7.5, 0.5 M EGTA, 0.01% sodium deoxycholate, 2 mM $MgCl_2$, and 0.02% Nonidet P-40 three times for 10 minutes. Whole embryos were immersed in X-gal staining solution (0.1 M phosphate buffer pH 7.5, 0.02% Nonidet P-40, 0.01% sodium deoxycholate, 5 mM potassium ferricyanide, 5 mM potassium ferrocyanide, 2 mM $MgCl_2$ and 1mg/ml X-Gal (Thermo Fisher Scientific)) and stained overnight at 37°C in dark. Embryos were then rinsed in PBS and post-fixed in 4% PFA prior imaging.

Embryos intended for cryo-sections were embedded in 30% sucrose/PBS overnight, frozen in OCT and cut to 10 µm sagittal sections according to CCP in-house standard operating protocols and frozen until staining. Sectioning was performed by embryology unit, CCP. Slides were washed in 1 M phosphate buffer pH 7.5, 0.5 M EGTA, 0.01% sodium deoxycholate, 2 mM $MgCl_2$ 0.02% Nonidet P-40 for 10 minutes each. Staining was performed overnight at 37°C in dark in identical solution as used for whole-mount staining. Slides were washed in PBS, post-fixed in 4% PFA for 10 minutes and washed twice in PBS for 10 minutes. Slides were then counter-stained with Nuclear Fast Red (Sigma-Aldrich), and mounted in Aquatex. Staining and imaging on Zeiss AxioImager Z2 (sections), Zeiss AxioScan Z1 (sections) and Zeiss AxioZoom with Apotome module macroscope (whole mount) was performed by Monika Sivá and embryology unit, CCP.

4 RESULTS

This dissertation represents part of a comprehensive study performed in the laboratory of Dr. Grantz Šašková, focusing on revealing the biological role/s of Ddi1-like protein family. The author, Monika Sivá, has contributed with her work to several ongoing projects: structural study of Ddi1p of *S. Cerevisiae* (chapter 4.1.2), interaction studies of human DDI2 protein using NMR spectroscopy (chapter 4.1.3), expression profiling of murine homolog Ddi1 in developing brain (chapter 4.2.1) and characterization of two diverse Ddi2-deficient mouse models (chapters 4.2.2). The thesis is therefore divided into two main chapters focusing on the study of Ddi1-like proteins on molecular level (Chapter 4.1.) and Ddi1-like proteins in biologically relevant models (Chapter 4.2). Combined together, the results from both perspectives will pave the way to the overall understanding of the biological roles played by members of Ddi1-like protein family.

4.1 CHARACTERIZATION OF DDI1-LIKE PROTEINS ON MOLECULAR LEVEL

Ddi1-like proteins throughout the eukaryotes exhibit high level of conservation in domain organization and structural features, as briefly discussed in the introduction chapter 1.3 on page 47. Sequence alignments of proteins of Ddi1-like family that are part of this study are shown in Figure 7A. All the hereby-studied proteins exhibit high sequence identities in the N-terminal UBL and central RVP domains, even though yDdi1p shows only around 35% sequence identity with each of the four mammalian orthologs individually. The high sequence identity applies as well to the HDD domain region preceding the RVP of the human and mouse orthologs, based on high residue identity with the hDDI2 HDD and on secondary structure predictions of the latter (data not shown). The overall identity of the full-length mammalian proteins is quite high, hDDI2 shares 72% of identical residues with hDDI1, and 71% with the murine Ddi1 homolog, which both seem to lack the C-terminal UIM motif (see Figure 7A). Murine Ddi2 harbors the UIM sequence and interestingly, it shares 96% sequence identity with hDDI2, thus supporting the use of *Ddi2* knockout mice as appropriate model for studying biological function of hDDI2. The overall high conservation of the domain architecture suggest more or less conserved function/s of the proteins, which are being elucidated further. Constructs of the recombinant proteins and synthesized peptides used in this work are depicted in Figure 7B.

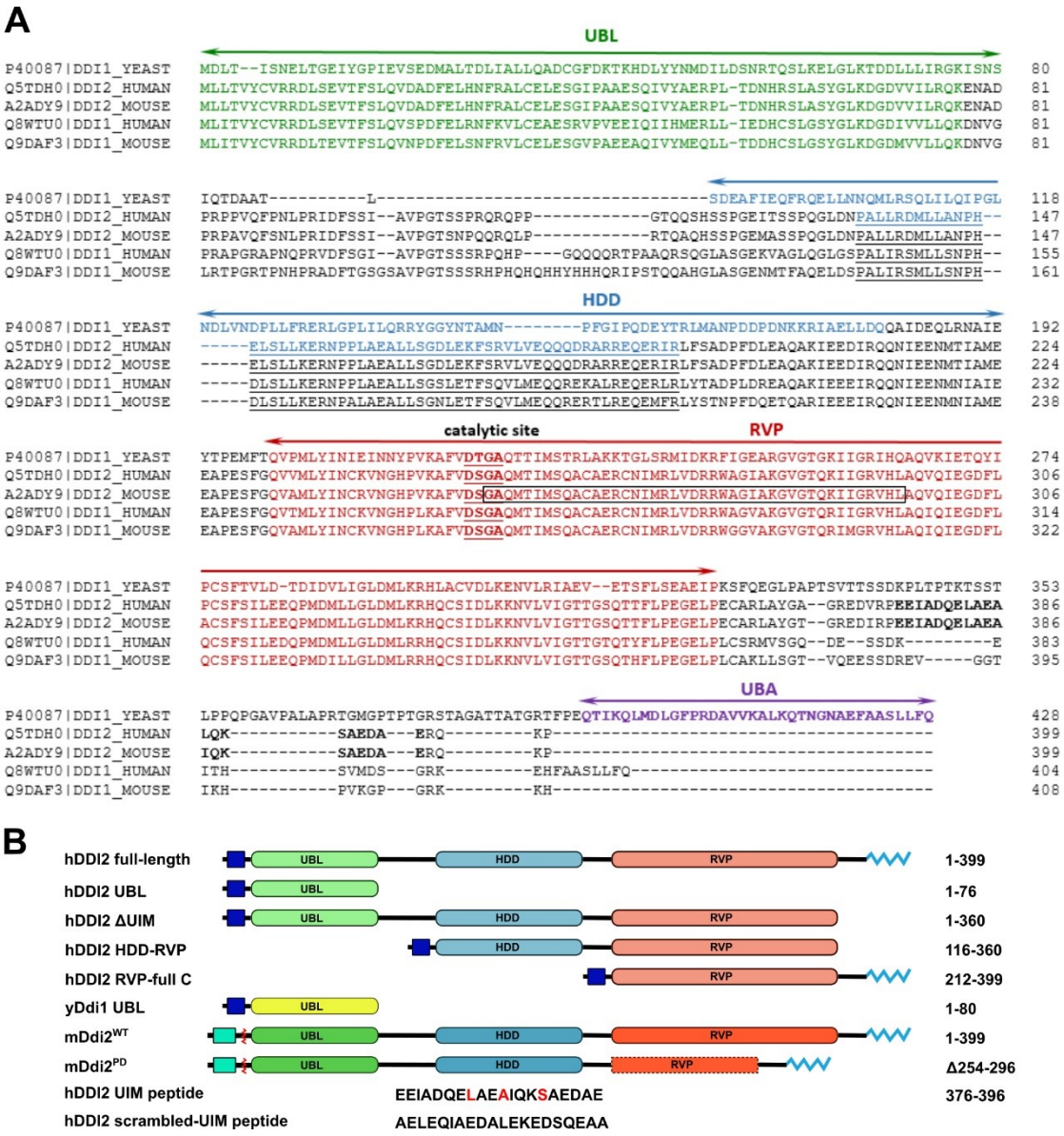


Figure 7: Ddi1-like protein family members studied in this thesis. A) Sequence alignment of Ddi1-like proteins from *S. cerevisiae*, *H. sapiens* and *M. musculus*. Highly conserved domains are distinguished by color: N-terminal ubiquitin-like domain (UBL) in green, helical domain of Ddi1-like proteins (HDD) in blue, retroviral protease-like domain (RVP) in red and C-terminal ubiquitin-associated domain (UBA) in violet. Very well conserved area of helical domain of Ddi1-like proteins is underlined. Catalytic active site residues inside the RVP domains are highlighted in bold. C-terminal ubiquitin-interacting motif (UIM) is in bold letters. Sequence encoded by exon 6 in mouse Ddi2 protein is shown in black rectangle. All sequence identity calculations were performed in Clustal Omega online tool (Chojnacki S. *et al.*, 2017). **B)** Schematic picture of protein constructs and peptide sequences used in this work. Individual domains are distinguished by color, UIM is represented as a blue zig-zag at the C-terminus. N-terminal histidine tag of proteins cloned into pET16b expression vector are colored in blue. mDdi2 WT and mDdi2^{PD} proteins were cloned into two vectors, p905 and pTreTight with His-tag and FLAG-tag at N-terminus showed in green.

4.1.1 Individual domains of Ddi1-like protein family members display high structural conservation

The conservation of Ddi1-like protein family members shown in sequence alignment in Figure 7A clearly designates the structure of individual members to be alike. Therefore, structural alignment of individual domains of all hereby-studied Ddi1-like proteins was performed using structures that were either solved in our laboratory (including yDdi1 UBL solved by Monika Sivá, see following chapter) or that are available in PDB. Superimposition of the UBL domains of yDdi1 and mDdi1 to hDDI2 shown in Figure 8A, revealed RMSD for backbone 1.22Å and 1.34Å, respectively. As apparent from both sequence and structure alignments (Figure 8B), the similar region of hDDI2 HDD and yDdi1 HDD is represented by the four helix bundle (in case of yeast ortholog at the N-terminal residues 89 – 141),

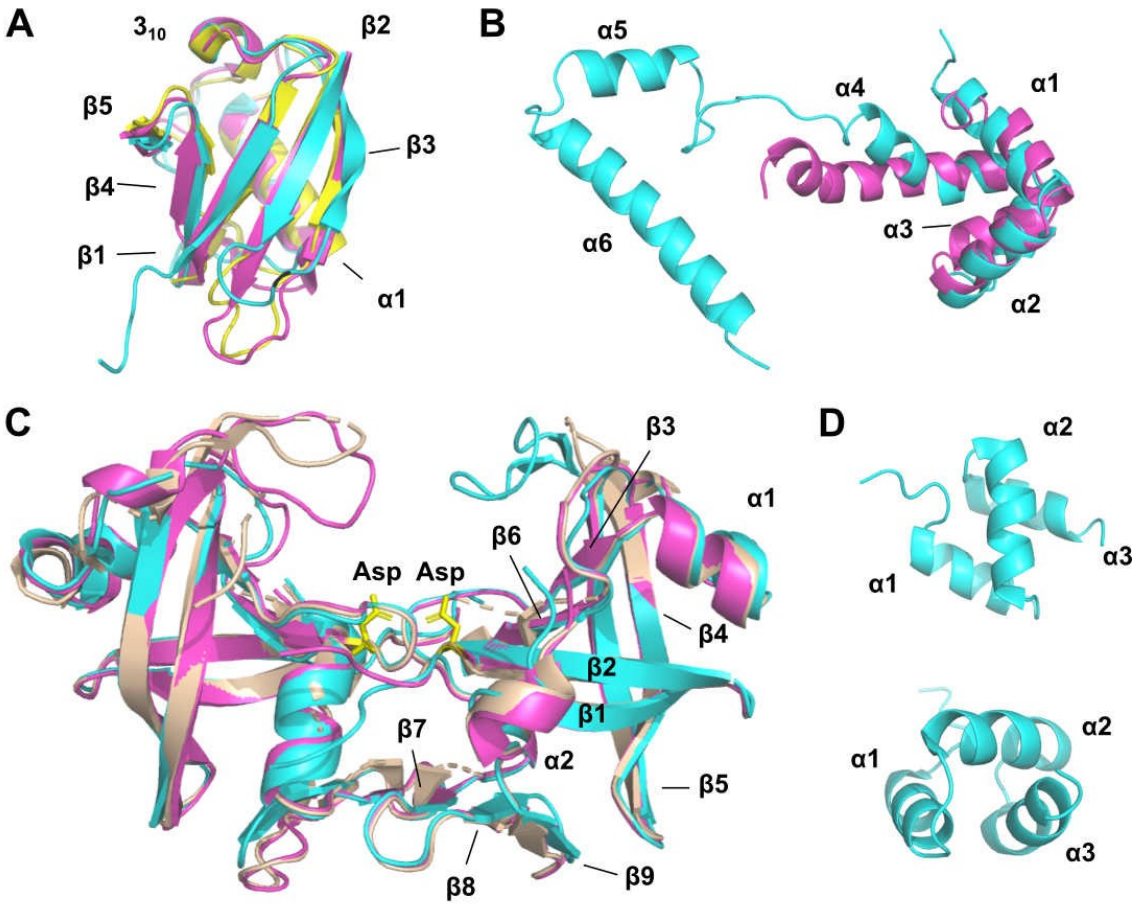


Figure 8: Ddi1-like proteins have highly conserved domain architecture among species. A) Ubiquitin fold is very well conserved among the UBLs of hDDI2 (magenta, 2N7D), yDdi1p (cyan, 2N7E) and mDdi1 (yellow, 1V5O, unpublished). **B)** Superimposition of hDDI2 HDD (magenta, 5K57) with HDD of yDdip (cyan, 5KES). **C)** Superimposition of RVP domain structures from hDDI2 (magenta, PDB entry 4RGH), Ddi1p (cyan, PDB entry 4Z2Z) and hDDI1 (wheat, PDB entry 3S8I, unpublished). **D)** Solution structure of UBA domain of yDdi1p (PDB entry 2MR9). Structures adapted from (Nowicka U., *et al.*, 2015, Siva M., *et al.*, 2016, Trempe J. F., *et al.*, 2016). Superimpositions and RMSD calculations were performed in program PyMOl (Schrodinger, LLC, 2015).

displaying RMSD of 2.96Å in superimposition (Siva M., *et al.*, 2016, Trempe J. F., *et al.*, 2016). The HDD of yDdi1p harbors additional two helices at its C-terminus compared to hDDI2 HDD. RVP domains exhibit remarkable conservation in the structured parts, the β-sheets and α-helices, the only difference is in the unstructured loops (flaps) covering the active site cavity, caused by their flexibility (see Figure 8C) (Siva M., *et al.*, 2016, Trempe J. F., *et al.*, 2016). The yDdi1 RVP and hDDI1 RVP superimposed to the hDDI2 RVP with RMSD of the backbone residues of 0.7Å and 0.35Å, respectively. Figure 8D shows the 3-helix bundle representing the UBA of yDdip, the primary ubiquitin interaction site, which is not present in the human and mouse orthologs (Nowicka U., *et al.*, 2015).

Overall, the structural superimposition of individual domains revealed almost identical fold in case of RVP and UBL domains and highly similar fold for the four helical bundle of HDD. Based on these results, we decided to compare also functional properties of these domains on molecular level. Part of the analyses, performed specifically by Monika Sivá, is described in detail in following chapters.

4.1.2 Structural characterization of ubiquitin-like domain of Ddi1 from *S. cerevisiae*

The structure of Ddi1 UBL from baker´s yeast was solved using nuclear magnetic resonance. Protein construct encoding the yDdi1 UBL domain (residues 1-80), was cloned in-frame with N-terminal histidine tag (for the construct sequence see Figure 7 on page 100). The protein was expressed with yields of 1.67 mg and 1.12 mg per liter of medium for ^{15}N-labeled and ^{13}C/^{15}N-labeled protein, respectively.

Proper folding of the protein construct was verified with acquisition of 1D HSQC spectrum. Nearly complete ^{15}N, ^{13}C and ^{1}H-resonance assignments were acquired for yDdi1 UBL with N-terminal histidine tag and the structure was solved with high precision. A family of 43 converged structures was obtained with the RMSD to the mean structure at the ordered residue range (residues 1-78) for the backbone and heavy atoms 0.41Å and 0.81Å, respectively. The yeast Ddi1 UBL contains four β-sheets (β1:13I-19V, β2: 1M-7N, β3: 70L-75G, β4: 43H-46Y), one α-helix (25L-35D) and a 3_{10}-helix (60L-63L) as shown in Figure 8A. Despite its low sequence identity with ubiquitin (23.61% according to program Clustal Omega) (Chojnacki S., *et al.*, 2017), it adopts the ubiquitin fold (see Figure 9A, B and C). The β-sheet patch (Figure 9A) could serve as a potential interaction site as we further discuss in our publication Trempe *et al.,* 2016 (Trempe J. F., *et al.*, 2016). NMR constraints and structural statistics for the yDdi1 UBL domain are summarized in Table 4 below.

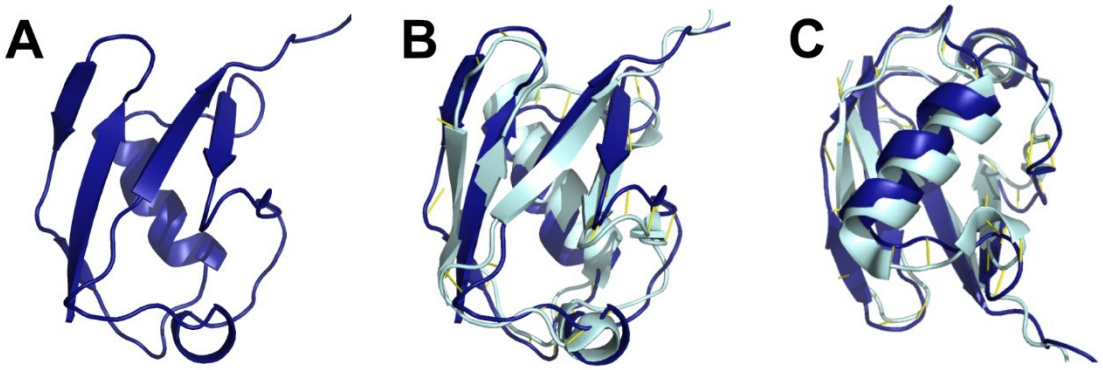


Figure 9: Solution structure of ubiquitin-like domain of Ddi1p from *S. cerevisiae*. Adapted from (Trempe J. F., et al., 2016). **A)** Solution structure of yDdi1 UBL (PDB entry 2N7E) bears ubiquitin fold. Resonance assignments were carried out manually in program Sparky (Goddard T. D. a. K., D. G. , 2008), calculations were carried out in programs Cyana 2.1 (Herrmann T., et al., 2002), TALOS+ (Shen Y., et al., 2009) and YASARA forcefield. **B)**, **C)** Superimposition of yDdi1-UBL with human ubiquitin (PDB entry 1D3Z) from two different views (Cornilescu G., et al., 1998). Structure drawing and alignment was performed in program PyMOL (Schrodinger, LLC, 2015).

Table 4: NMR constraints and structural statistics for yDdi1 UBL. Adapted from (Trempe J. F*., et al.*, 2016).

Non-redundant distance and angle constrains		
Total number of NOE constraints	1634	
Short-range NOEs (*i, i*+1)	880	
Medium-range NOEs (*i,i>1 i*≤4)	290	
Long-range NOEs (*i, i* ≥5)	464	
Tosion angles	128 (64 ϕ and 64 ψ)	
Total number of restricting constraints per restrained residue	20.7	
Maximum constraints violations and r.m.s		
Upper distance limits (Å)	0.07 ±0.02	0.0019 ±0.0005
Van der Waals contacts (Å)	0.14 ±0.02	
Torsion angle ranges (°)	3.71 ±0.22	0.475 ±0.036
Average CYANA target function (Å^2)	0.17 ±0.02	
Ramachandran plot		
Residues within the most favoured region	86%	
Residues within the additionally allowed region	14%	
Residues within generously allowed region	0%	
Residues within the disallowed region	0%	
r.m.s.d. to the mean structure	*ordered 100-177*	*all residues*
Backbone heavy atom (Å)	0.41 ± 0.06	1.71 ± 0.37
All heavy atom (Å)	0.81 ± 0.07	2.07 ± 0.36

4.1.3 Characterization of human DDI2 binding properties

4.1.3.1 DDI2 and its interaction with ubiquitin

To investigate the function of human DDI2 as an adaptor protein in the ubiquitin-proteasome system, in line with the published functional studies of the yeast Ddi1 ortholog, a series of hDDI2 - ubiquitin interaction experiments were performed using NMR. I focused on a detailed evaluation of two possible UBQ binding sites of human DDI2 protein: a UIM motif at the C-terminus identified by us (Siva M.*, et al.*, 2016)and a UBL at the N-terminus that was previously described for the yeast Ddi1 UBL domain as an alternative UBQ binding site (Nowicka U.*, et al.*, 2015).

First, titration experiments with ^{15}N-labeled ubiquitin and hDDI2-UIM peptide (residues 376-395 of human DDI2; see Figure R7A and B, page 100) were performed where a 35-fold molar excess of the peptide (final concentration of the hDDI2-UIM peptide was 3.45 mM) was reached as shown in Figure 10A. Chemical shift perturbation of individual amino acids were plotted at the endpoint of the titration (Figure 10D) and the residues with most significant shifts were zoomed in in the spectra overlay in Figure 10A. The K_d of 2.2-3.3 mM was calculated from 6 residues (K6, A46, G47, Q49, H68 and L71) with a 1:1 stoichiometry model for specific binding (Figure 10C) which were mapped onto the structure of ubiquitin (PDB entry 1D3Z, (Cornilescu G.*, et al.*, 1998))in Figure 10B. Based on these and other additional CSPs identified in this experiment (L8, R42, K48), we concluded that the interaction site differs from the common Isoleucine 44 interaction patch on ubiquitin (Bertolaet B. L.*, et al.*, 2001, Sloper-Mould K. E.*, et al.*, 2001).

To verify the specificity of the interaction of hDDI2-UIM peptide with UBQ, a similar experiment was performed with hDDI2-scrambled UIM peptide (for sequence see Figure 7, page 100). In the titration experiment, 1.9 mM final concentration of the hDDI2-scrambled UIM peptide was reached with no CSPs observed. The plot of CSPs for a 2.2 mM addition of hDDI2-UIM peptide and a 1.9 mM addition of the control hDDI2-scrambled UIM peptide (see Figure 10E) reveals the difference and confirms the specificity of the weak interaction between UBQ and the UIM sequence at the C-terminal end of hDDI2 protein.

Furthermore, we wanted to investigate this UBQ/hDDI2-UIM peptide interaction on a protein level, using more appropriate model. CSPs experiment with a ^{15}N-labeled ubiquitin with a 1-, 2- and 5-molar excess of a non-labeled hDDI2 RVP full-C protein construct (residues 212-399; see Figure 7) and a reverse experiment with ^{15}N-labelled hDDI2 RVP

full-C and non-labelled UBQ were performed. In the first titration, weak, yet specific changes for amino acids T7, R42, K48, Q49 and L71 (see Figure 10F, G and H) were observed,

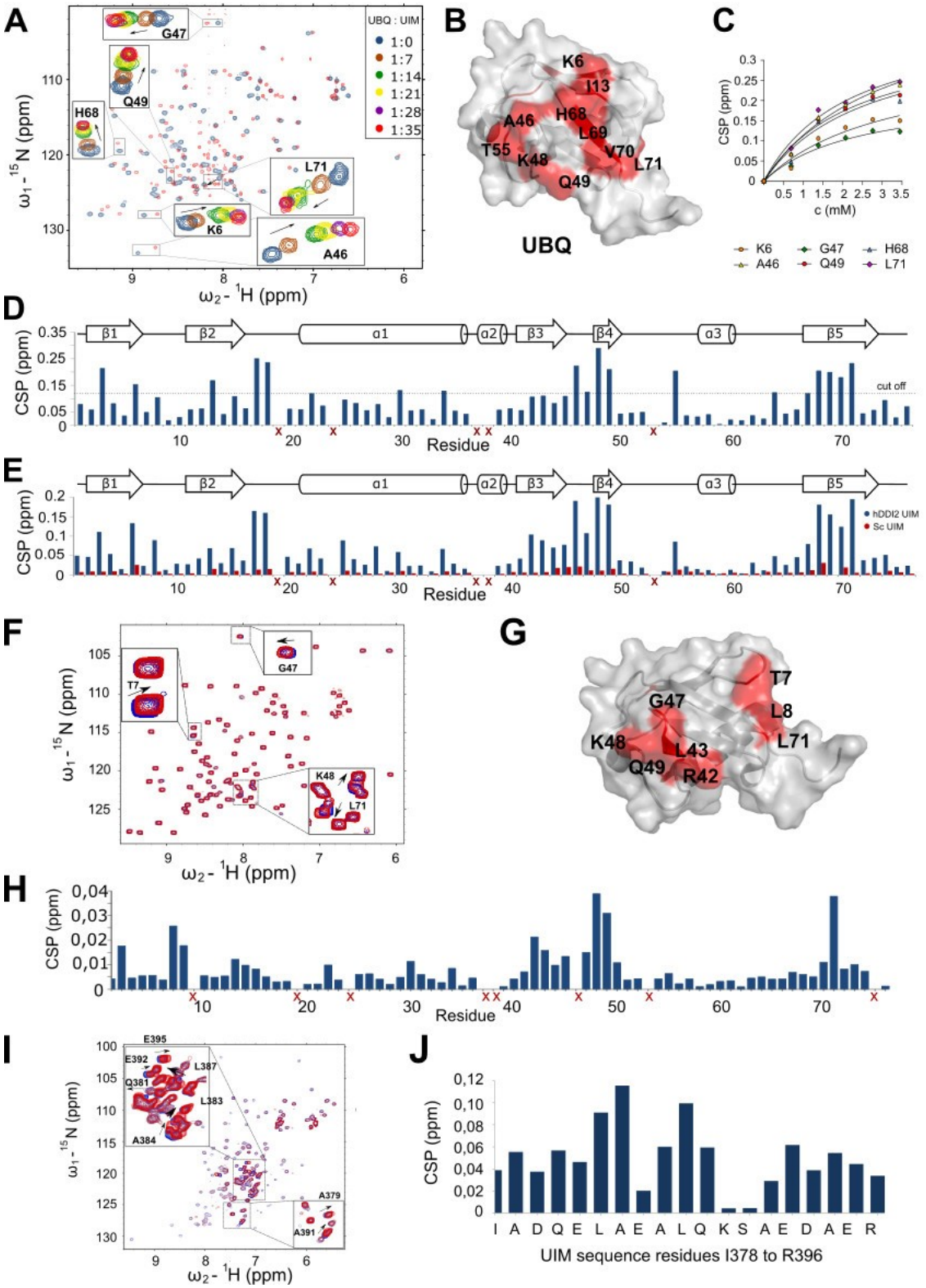


Figure 10: Characterization of interaction between ubiquitin and C-terminal UIM of human DDI2 protein using NMR spectroscopy. Adapted from (Siva M., *et al.*, 2016). **A)** Overlay of 2D HSQC spectra of ubiquitin acquired during titration with hDDI2-UIM peptide (individual additions of peptide are distinguished in color). **B)** Mapping of amino acids with most significant shifts in amide signals on UBQ

structure (Cornilescu G.*, et al.*, 1998). **C)** Titration curves of ubiquitin residues with largest CSPs upon interaction with hDDI2-UIM peptide. **D)** Plot of chemical shift perturbations of individual amino acids at 35-fold molar excess of hDDI2-UIM peptide. **E)** Plot of CSPs of ubiquitin residues at addition of hDDI2-UIM peptide up to final concentration of 2.2 mM (blue) and of hDDI2-scrambled UIM peptide up to 1.9 mM (red). **F)** 2D HSQC spectra of ^{15}N-labeled UBQ (blue) with 5-fold molar addition of non-labeled hDDI2 RVP full-C (red). **G)** Mapping of most significant shifts of UBQ residues (PDB entry 1D3Z) (Cornilescu G.*, et al.*, 1998) upon interaction with non-labeled hDDI2 RVP full-C. **H)** Very weak, however specific CSPs of UBQ backbone amide signals upon interaction with hDDI2 RVP full-C. **I)** Reverse mapping of the interaction on 2D HSQC spectra of ^{15}N-labeled hDDI2 RVP full C (blue) with 5-fold molar addition of non-labeled UBQ (red). **J)** Plot of CSPs of the hDDI2-UIM sequence locus upon UBQ binding. Red crosses in CSP plots mark residues that could not be used for evaluation. The figure was created in programs GraphPad Prism and Pymol (GraphPad S., Schrodinger, LLC, 2015).

which again suggests a different interaction site from the I44 patch on ubiquitin (Bertolaet B. L.*, et al.*, 2001, Sloper-Mould K. E.*, et al.*, 2001). In the reverse experiment, with 1-, 2- and 5- fold molar addition of UBQ, the CSPs in the backbone signals for residues of the hDDI2 RVP full-C were mapped onto the UIM sequence at the C-terminus of hDDI2 as shown in Figure 10I and 10J, which confirms the previous results of the UBQ/hDDI2-UIM peptide interaction.

Next, I focused on detailed evaluation of the second possible ubiquitin binding site. According to Nowicka and colleagues, yDdi1 UBL domain from *S. cerevisiae* binds ubiquitin (Nowicka U.*, et al.*, 2015). This surprising finding led our team to speculation about an alternative shuttle mechanism performed by yDdi1p and human DDI1 and DDI2 proteins. Therefore, NMR titration experiments with ^{15}N-labeled hDDI2 UBL and up to 10-fold molar addition of non-labeled UBQ (Figure 11A, p XX) were performed to decipher the possibility of UBQ/hDDI2 UBL interaction. CSPs were plotted for individual amino acids (Figure 11D) and the most significant perturbations were mapped onto the structure of hDDI2 UBL as shown in Figure 11B (PDB entry 2N7D, (Siva M.*, et al.*, 2016)). C7, V8, T16, F17, V21, F25, F30, Q46, D70 and I73 chemical shifts were used for calculation of the K_d, which resulted in the 0.42-1.1 mM range (Figure 11C). The reverse titration experiment with ^{15}N-labeled UBQ and 6-fold molar addition of non-labeled hDDI2 protein lacking UIM peptide at the C-terminus (hDDI2 ΔUIM; for sequence see Figure 7B, page 100) revealed the interaction site on UBQ (Figure 11E - G). The interaction was mapped onto the Isoleucine 44 patch as shown on the UBQ structure in Figure 11F (PDB entry 1D3Z, (Cornilescu G.*, et al.*, 1998)). To verify the localization of the interaction, additional control experiment with a 6-fold molar addition of hDDI2 protein lacking both N-terminal UBL domain and C-terminal UIM (hDDI2 HDD-RVP; for construct see Figure 7, page 100) to ubiquitin was performed. No significant chemical shift perturbations in the backbone amide signals of UBQ were

observed (Figure 11H and I), proposing very weak interaction of the hDDI2 UBL with UBQ as opposed to the yeast Ddi1 UBL/UBQ interaction (Nowicka U., *et al.*, 2015).

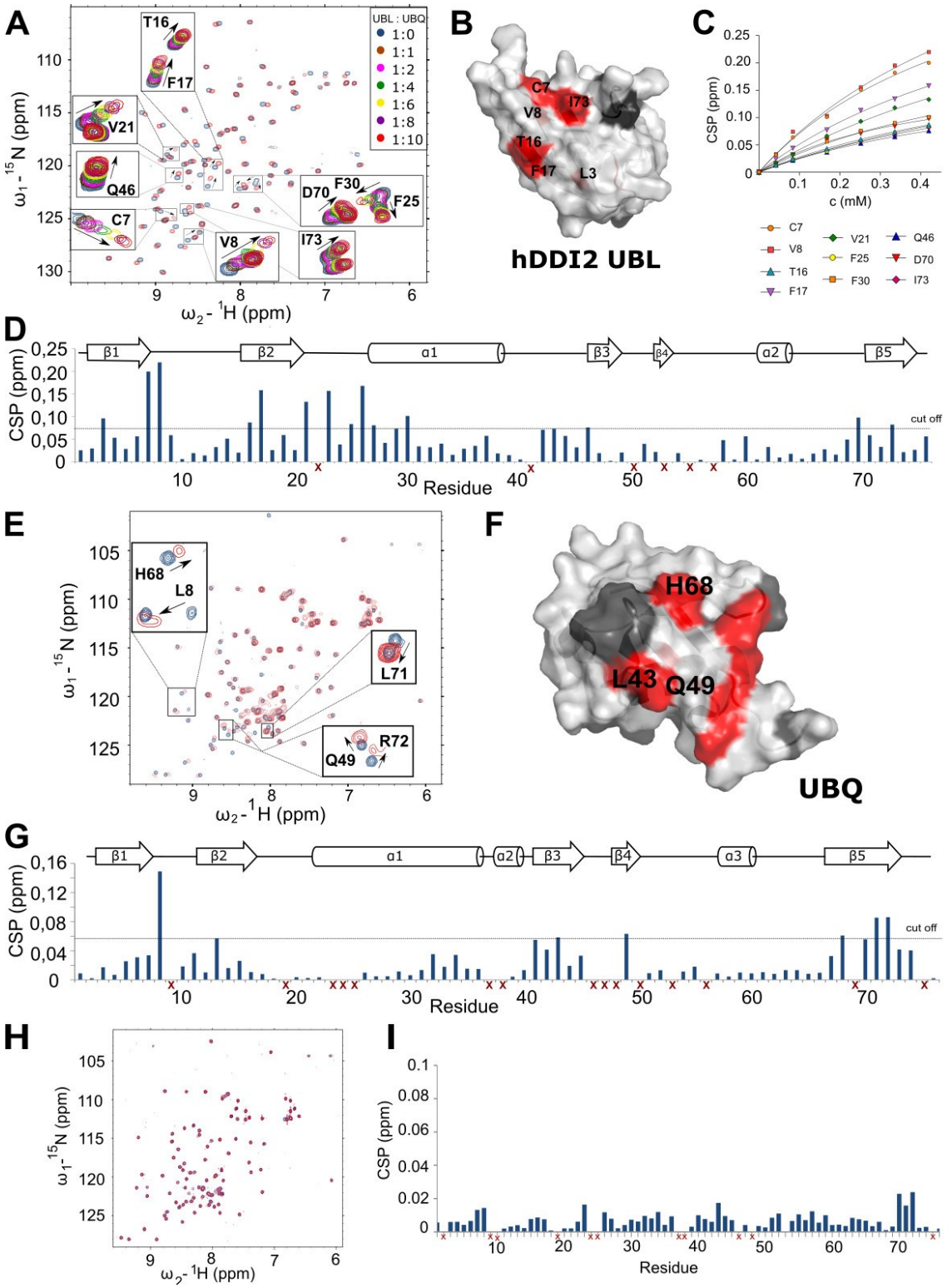


Figure 11: Characterization of UBQ – hDDI2 UBL interaction using NMR spectroscopy. Adapted from (Siva M., *et al.*, 2016). **A)** 2D HSQC titration spectra overlay of 15N-labeled hDDI2 UBL with up to 10-fold molar addition of non-labeled UBQ (individual additions of UBQ are distinguished in color). **B)** Mapping of the UBQ interaction site on hDDI2 UBL structure (PDB entry 2N7D) (Siva M., *et*

al., 2016) in red. **C)** Titration curves of shifts of hDDI2 UBL amino acids, that were used for K_d calculation. **D)** Plot of CSPs in hDDI2 UBL residues at the endpoint of titration with UBQ. Red crosses mark residues that could not be used for evaluation. **E)** 2D HSQC titration spectra of 15N-labeled UBQ prior (blue) and after addition of 6-fold molar excess of non-labeled hDDI2 ΔUIM protein construct (red). Signals of amides of most shifted amino acids are zoomed in and **F)** mapped onto the structure of UBQ (PDB entry 1D3Z) (Cornilescu G.*, et al.*, 1998). **G)** Plot of CSPs of UBQ residues upon interaction with hDDI2 ΔUIM. **H)** No significant chemical shift perturbations were observed in 2D HSQC spectra of 15N-labeled UBQ (blue) upon 6-fold molar addition of hDDI2 HDD-RVP (red) lacking the N-terminal UBL domain and C-terminal UIM motif. **I)** Plot of CSPs of backbone amides of UBQ upon addition of hDDI2 HDD-RVP. Red crosses in CSP plots mark residues that could not be used for evaluation. Amino acids that could not be used in mapping on protein structures are colored in black. The figure was created in programs GraphPad Prism and Pymol (GraphPad S., Schrodinger, LLC, 2015).

4.1.3.2 Investigation of possible intramolecular interactions of hDDI2

To compare the binding properties of UBL domain of human DDI2 with the yeast Ddi1 UBL, we acquired and superimposed 2D HSQC spectra of ^{15}N-labeled full-length hDDI2 protein and its UBL domain itself. We observed quite significant shifts in the backbone amide signals of the hDDI2 N-terminus (see Figure 12A), which clarifies the positioning of the hDDI2 UBL domain inside the full-length protein dimer. In contrast to the yDdi1 UBL domain localization as described by Nowicka and her colleagues (Nowicka U.*, et al.*, 2015), the UBL domain of hDDI2 does not extend out of the protein body.

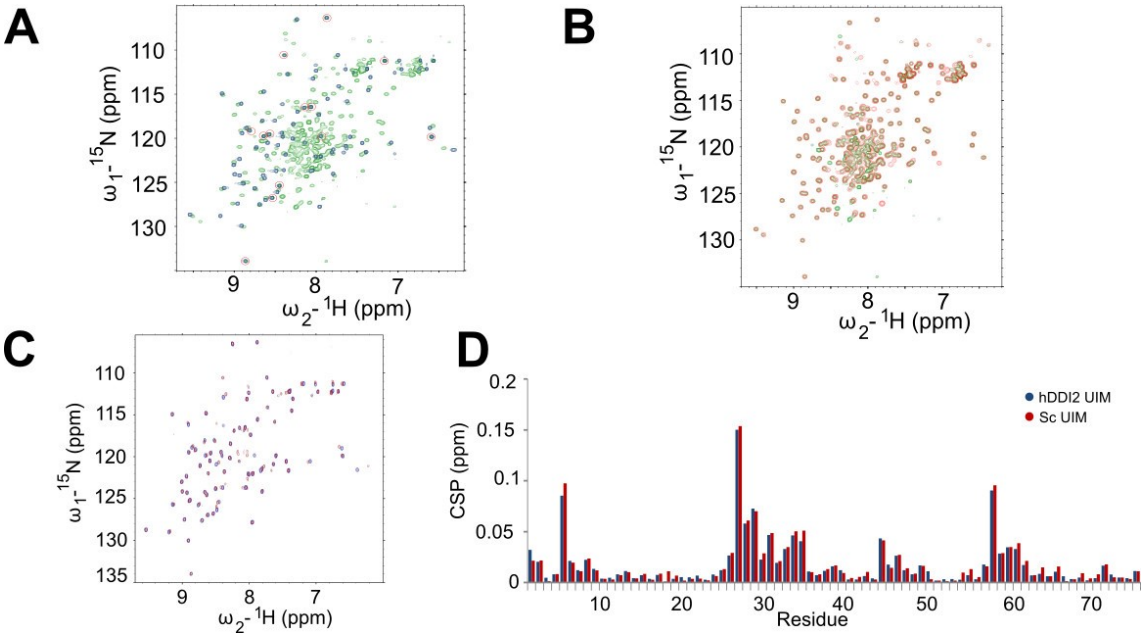


Figure 12: Examination of intramolecular interactions of hDDI2 protein using NMR spectrometry. Adapted from (Siva M.*, et al.*, 2016). **A)** hDDI2 UBL does not extend away from the body of the hDDI2 protein. 2D HSQC spectra of hDDI2 UBL (blue) superimposed with the 2D HSQC signals of the full-length protein (green) show differences in signals of the UBL, suggesting a rather compact structure for the full-length protein. The few amino acids lacking the difference are marked with red circles. **B)** The N-terminal UBL domain of hDDI2 does not bind C-terminal UIM peptide derived from hDDI2 protein. The overlay of 2D HSQC spectra of full-length hDDI2 protein (green) with 2D HSQC spectra of hDDI2 ΔUIM (red) does not reveal shifts in the signals of UBL domain. **C)** 2D HSQC spectra of hDDI2 UBL before

(blue) and after addition of the UIM peptide derived from the hDDI2 C-terminus (red) with final 1.9 mM concentration. **D)** Plot of CSPs after addition of hDDI2-UIM peptide (blue) and after addition of the hDDI2-scrambled UIM peptide (red) to a final concentration of 1.2 mM. The figure was created in programs GraphPad Prism and Pymol (GraphPad S., Schrodinger, LLC, 2015).

The possible intramolecular interaction of the N-terminal UBL domain with the C-terminal UIM motif of hDDI2 was further investigated. First, 2D HSQC spectra of ^{15}N-labeled full-length hDDI2 were acquired and superimposed with the 2D HSQC spectra of the truncated hDDI2 ΔUIM, where no difference was observed for the backbone amide signals for the UBL domain (see Figure 12B). NMR titration experiments with ^{15}N-labeled hDDI2 UBL domain with hDDI2-UIM peptide up to a final concentration of 1.9 mM were performed and revealed slight shifts in the backbone amide signals of UBL residues (see Figure 12C). However, a negative control experiment with addition of 1.9 mM hDDI2-scrambled UIM peptide resulted in the identical CSPs (Figure 12D). This suggests that hDDI2 UBL most likely does not bind to the UIM motif and so the full-length protein does not adopt a head-to-tail conformation.

In summary, our NMR titration studies revealed weak, yet specific ubiquitin interaction motif (UIM) at the C-terminus of human DDI2 protein and similarly very weak interaction of hDDI2 UBL with ubiquitin as opposed to yDdi1 UBL binding properties to ubiquitin. Based on our results, human DDI2 UBL does not extend from the "protein body" unlike its yeast counterpart and the full-length DDI2 protein never adopts a head-to-tail conformation, employing RVP domain as a central dimer interface and enclosing N-terminal UBL with C-terminal UIM in an autoinhibitory mode. This led us to conclusion that although yeast Ddi1 and human DDI2 proteins are structurally highly similar, at least some of the functional properties are not preserved and the proteins could differ in their biological roles.

4.2 Investigation of the role of Ddi1-like proteins using biological systems

In order to decipher biological role/s played by Ddi1-like proteins, we decided to use mouse models due to their experimental feasibility and based on their high sequence conservation when compared to human Ddi1-like proteins.

Mice possess two homologs of Ddi1-like genes. *Ddi2* (ENSMUSG00000078515, NCBI ID: 68817) is localized on the reverse strand of chromosome 4 (141,677,549-141,723,419), has one 10-exon transcript (ENSMUST00000102484.4, NM_001017966.2) and encodes a 399-residue long protein (A2ADY9, NP_001017966.1). *Ddi1* (ENSMUSG00000047619, NCBI ID: 71829) (Ensemble release 95) is localized on the

reverse strand of chromosome 9 (6,262,733-6,269,846) and encodes only one exon (ENSMUST00000051706.5, NM_027942.1) of a 408-residue protein (Q9DAF3, NP_082218.1) (Geer L. Y.*, et al.*, 2010, The UniProt C., 2017, Zerbino D. R.*, et al.*, 2018). These two homologs share 71% sequence identity (Chojnacki S.*, et al.*, 2017). As already mentioned in the introduction to Ddi1-gene family, *Ddi1* was generated from the original *Ddi2* gene as a copy via retrotransposition event during evolution (see chapter 1.3) (Siva M.*, et al.*, 2016).

Only scarce data is available on the expression of *Ddi1* and *Ddi2* in *M. musculus*. While *Ddi2* is quite ubiquitiously expressed, *Ddi1* was found only in testes of adult mice. In this work we show and comment specific expression of *Ddi1* in developing embryonic brain, which – in context with other observations described in our publication (Ramirez J.*, et al.*, 2018) - sheds light on its possible physiological function. Murine Ddi1 and Ddi2 proteins share 81% and 96% sequence identity with human DDI1 and DDI2 proteins, respectively (Chojnacki S.*, et al.*, 2017) (sequence alignment in Figure 7 on page 100). This further supported our choice to use mouse as a suitable model for studying the function of human *DDI2* gene on its closest homolog. In order to do so, two mouse models were established, a full knockout of Ddi2 and a protease domain defective model. Generation of these models, allele characterization and phenotyping screen together with expression studies of *Ddi1* are described in detail in following chapters.

4.2.1 Investigating the possible biological role of human Ddi1 protein

Our collaborators from the laboratory of Dr. Ugo Mayor have identified hDDI1 as a substrate of UBE3A in human neuroblastoma SH-SY5Y cells that is highly ubiquitinated without being targeted to the proteasome. Moreover, they found out that mDdi1 expression rises rapidly at a specific stage of mouse embryonal development, E16.5 (Ramirez J.*, et al.*, 2018). To better understand the function of mDdi1 as a highly specifically expressed gene (similarly to hDDI1), we performed *Ddi1* expression analysis by *in situ* hybridization in the developing brain of CD-1 mouse embryos. We focused on 4 different developmental stages – E9.5, E10.5, E14.5 and E16.5. As can be seen in Figure 13A-D, *Ddi1* is expressed in all parts of the developing brain (telencephalon, diencephalon, mesencephalon and rhombencephalon) at the stage E9.5 and E10.5 as shown by both whole mount ISH or on sagittal sections. Throughout later developmental stages, the expression localizes more into mesencephalic and telencephalic structures. At stage E14.5, the signal for *Ddi1* mRNA cumulates in the upper hill (colliculus tectum) and ventricular zone of pallium

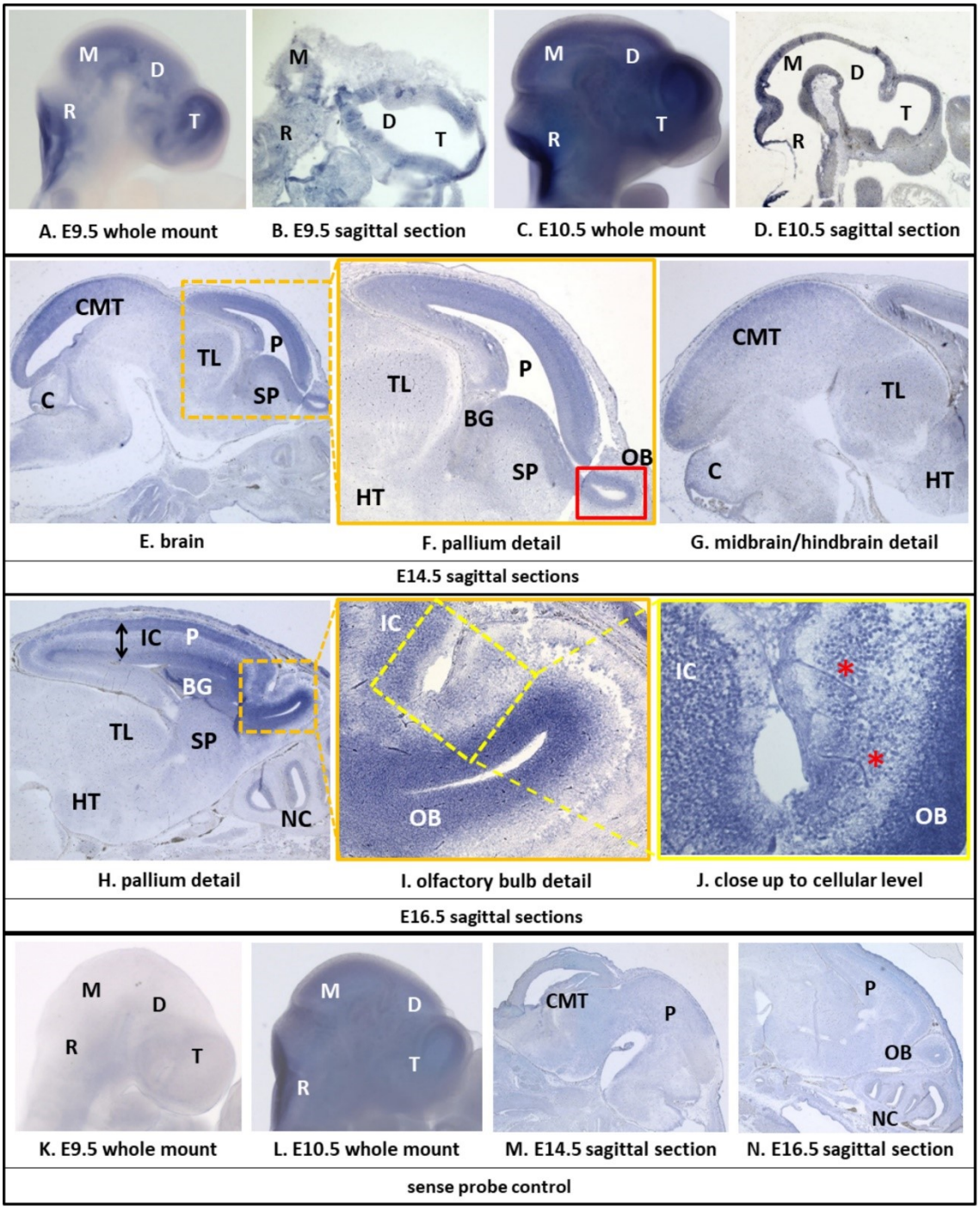


Figure 13: Spatial expression profiling of Ddi1 gene in mouse brain during embryonal development. Adapted from (Ramirez J., *et al.*, 2018). **A)** and **B)** Ddi1 is expressed in all parts of developing brain (telencephalon, diencephalon, mesencephalon and rhombencephalon) at embryonal stage E9.5 as shown on a whole mount and section ISH. The same pattern can be seen at stage E10.5 again for whole mount (**C**)) and for sagittal sections (**D)**). **E)** Sagittal section of embryonal brain at stage E14.5: Expression of Ddi1 is situated in pallial part of telencephalon (**F**) detail of pallium) and colliculus midbrain tectum (**G**) midbrain and hindbrain detail). **H)** and **I)** Specific expression of Ddi1 in isocortex and ventricular layer of olfactory bulb at stage E16.5. **J)** Detail of scan I: red stars highlight tubular structures, probably capillaries. **K)** Negative control whole mount ISH staining with a Ddi1 sense probe at embryonal stage E9.5. **L)** Non-specific background signal for Ddi1 sense probe is slightly increased at E10.5 for whole mount ISH. **M)** and **N)** Negative control ISH firmly confirms the specificity of Ddi1 expression studies performed on sagittal sections at stages E14.5 and E16.5, respectively. Abbreviations: T – telencephalon, D –

diencephalon, M – mesencephalon, R – rhombencephalon, C – cerebellum, CMT – colliculus midbrain tectum, TL – thalamus, P – pallium, SP – subpallium, HT – hypothalamus, BG – basal ganglia, OB – olfactory bulb, IC – isocortex, NC – nasal cavity.

(see Figure 16E-G). According to our collaborators, the expression of *Ddi1* in mouse embryonal brain reaches the highest level at stage E16.5 (Ramirez J.*, et al.*, 2018). The expression at this stage is localized in the ventricular layer and cortical plate of isocortex and the ventricular layer of olfactory bulb, while all these structures develop from telencephalon (see Figure 13H-J). As was suggested in (Ramirez J.*, et al.*, 2018), the expression could be situated in the neuroblast cells that undergo division and migration towards the external layer of isocortex.

We have observed increase in false positive signal of sense probe used for whole mount ISH at embryonal stage E10.5 as shown in Figure 13L. However, this embryonal stage was not crucial for our findings. It was rather used for presentation of expression localization throughout brain development until the essential stage at E16.5, where the expression was as quantified by qPCR experiments (Ramirez J.*, et al.*, 2018). Negative controls at other embryonal stages did not show any unusual elevation of background signal for whole mount ISH at E9.5 and for ISH on sections at E14.5 and E16.5 (Figure 13K, M and N).

Our expression profiling described the localization of the mDdi1 mRNA in the neuronal tissue of developing brain for the first time ever. In connection to the fact that hDDI1 is a unique, highly ubiquitinated substrate of UBE3A ligase without being targeted for proteasomal degradation, our study contributed to the formulation of hypothesis that hDDI1 specifically expressed in neuronal tissue (most probably also developing brain), might be of relevance in clinical research of Angelman syndrome.

4.2.2 Deciphering the biological role of DDI2 using mouse models

4.2.2.1 Introduction and nomenclature of our mouse models

Full knockout strain, C57BL/6NCrl-Ddi2$^{tm1b(EUCOMM)Hmgu}$/Ph mouse strain was created from original ES clone HEPD0660_5_E02, which belongs under The European Conditional Mouse Mutagenesis Program (EUCOMM) at the CCP hosted by IMG CAS. It records a loss of critical exon (exon2) of *Ddi2* gene resulting in frame-shift (see Figure 14B and C) (for more details see chapter 4.2.2.2). This strain will be further named as *Ddi2^{tm1b}*. Genotypes will be distinguished as *Ddi2$^{+/+}$*, *Ddi2$^{+/-}$* and *Ddi2$^{-/-}$*.

Ddi2 protease defective model C57Bl/6NCrl-Ddi2$^{em1/Rase}$ was generated at IMG CAS in the laboratory of Dr. Radislav Sedláček by TALEN-mediated excision of exon 6 of *Ddi2*

gene (em1/Rase) by Dr. Petr Kašpárek. This alteration of *Ddi2* gene resulted in alteration of the protein product and hence loss of its proteolytic activity and dimerization capability (see Figure 14D). Further information is stated in chapters 4.2.2.3. C57Bl/6NCrl-Ddi2$^{em1/Rase}$ strain will be further named *Ddi2*$^{protease\ defective}$ model. Genotypes will be labeled as *Ddi2*$^{exon6\ +/+}$, *Ddi2*$^{exon6\ +/-}$ and *Ddi2*$^{exon6\ -/-}$. The truncated protein construct is named mDdi2^{PD} (standing for protease defective).

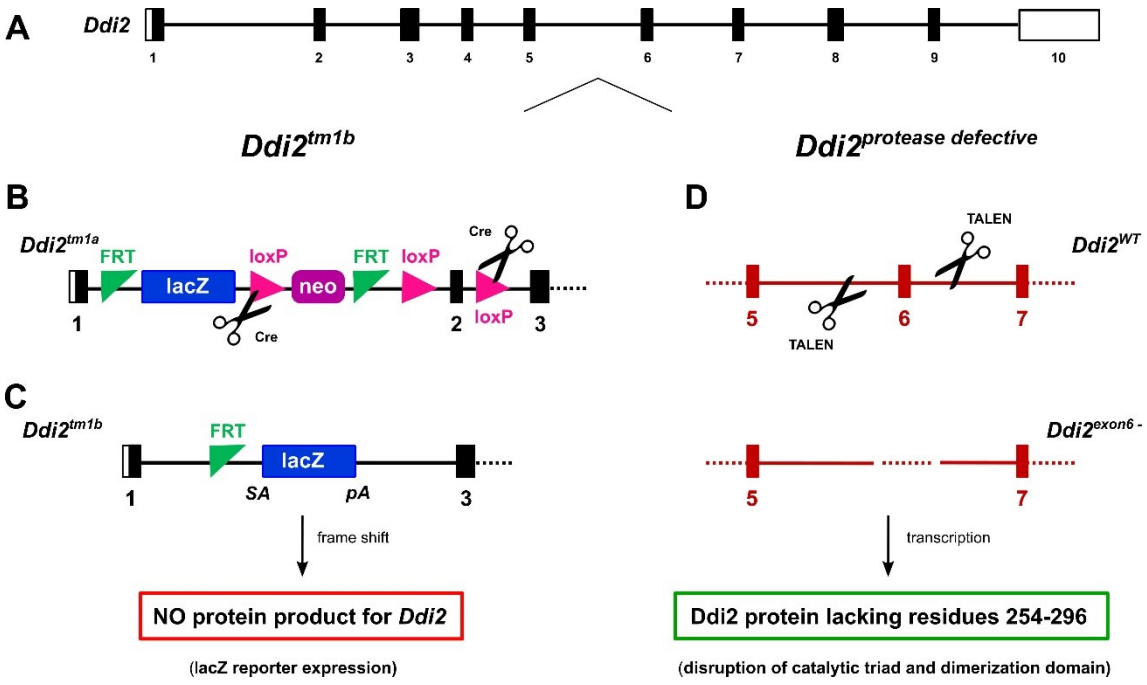


Figure 14: Schematic diagram of *Ddi2* gene alterations for generation of both mouse models, the *Ddi2*^{tm1b} and *Ddi2*$^{protease\ defective}$. A) Wild-type allele of murine *Ddi2* gene consists of 10 exons. **B)** Scheme of *Ddi2*^{tm1a} cassette inserted into ES cells for Ddi2 knockout mouse strain production. **C)** Scheme of *Ddi2*^{tm1b} cassette resulting in a frame-shift transcription. **D)** TALEN-induced excision of exon 6 in *Ddi2* gene resulting in protease domain alteration of the transcribed protein.

4.2.2.2 Generation of *Ddi2*^{tm1b} strain and genotyping

C57BL/6NCrl-Ddi2$^{tm1b(EUCOMM)Hmgu}$/Ph mouse strain was generated at the IMG CAS using Ddi2$^{tm1a(EUCOMM)Hmgu}$ (further mentioned as *Ddi2*^{tm1a}) embryonal stem cells which were created by ESCs manipulation. Heterozygous *Ddi2*^{tm1a} (see Figure 14B) adult mice were crossed with homozygous mice bearing ubiquitous expression of Cre-recombinase (Gt(ROSA)26Sor$^{tm1(ACTB-cre,-EGFP)Ics}$). The offspring heterozygous for *Ddi2* cassette and heterozygous for Cre-recombinase were further crossed with C57Bl/6NCrl to acquire F0 generation of C57BL/6NCrl-Ddi2$^{tm1b(EUCOMM)Hmgu}$/Ph.

As shown in Figure 15A, critical exon 2 of *Ddi2* gene was removed in this model, which creates a frame-shift and a null model. In contrast to *Ddi2*^{tm1a}, the promoter-driven

neomycin selection cassette was removed and *Ddi2* exon 2 was replaced with *LacZ* reporter for visualization of gene expression.

Genomic DNA for genotype estimation was extracted from tail biopsies collected at weaning or from yolk sac tissue collected at embryo harvest. Genotyping was designed to distinguish between wild-type *Ddi2* gene (1080 bps) and exchange of exon 2 with a *LacZ* reporter (640 bps) using a set of three primers: one common reverse primer and two different forward primers, the first one targeting sequence inside *Ddi2* that is cleaved out by Cre recombinase in *Ddi2*^{tm1b}, and the other one complementary with sequence inside *LacZ* (Figure 15A). Hence, it was possible to distinguish between wild-type, heterozygous and homozygous mice after one PCR reaction (Figure 15B).

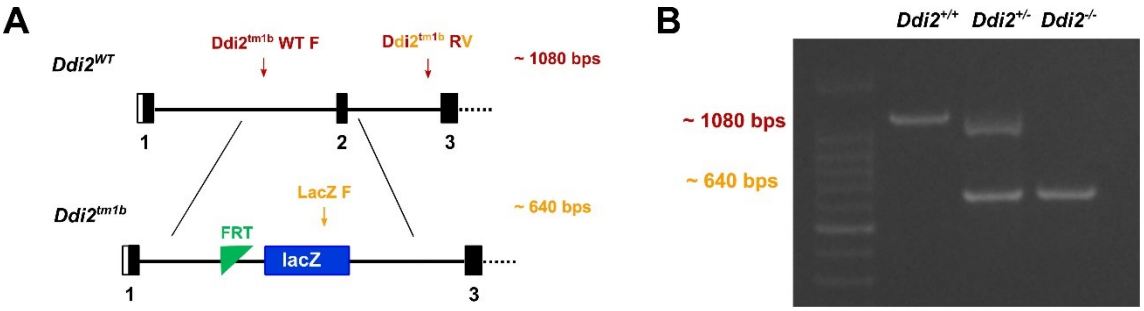


Figure 15: *Ddi2*^{tm1b} strain: design and evaluation of the results. A) Wild-type allele of murine *Ddi2* gene (close up on first three exons) and *Ddi2*^{tm1b} gene scheme. Target sequences for primers designed for genotyping of *Ddi2*^{tm1b} mouse strain are highlighted. **B)** Agarose gel showing genotyping results.

4.2.2.3 *Ddi2*$^{protease\ defective}$ mouse strain generation by TALEN-mediated *Ddi2* gene alteration and its genotyping strategy

Ddi2$^{protease\ defective}$ (C57Bl/6NCrl-Ddi2$^{em1/Rase}$) mouse strain was generated by Dr. Petr Kašpárek at IMG CAS using TALENs designed to target introns 5-6 and 6-7, which resulted in excision of exon 6 of *Ddi2* gene (design and experiment described in detail in chapter 3.4.5.1).

Two independent microinjections into mouse zygotes were performed. 51 mice in F0 generation were genotyped for nuclease-mediated alteration of *Ddi2* gene (Figure 14D). Four founder mice (IDs: 21, 30, 33 and 38) bearing exon 6 deletion were identified. The deletion resulted in shortage of the gene in approx. 550 bps (depending on the non-homologous end joining in the intron sequence). Targeted area was sequenced for each founder mouse (individual sequences are shown in Figure 16A), which were further crossed with wild-type partner to gain F1 generation. All of these mice were able to reproduce, however, we found out that one of the founder mice (ID: 33), was not able to produce offspring with desired exon 6 deletion, which meant that it did not bear *Ddi2* alteration in the germline cells. As all the

remaining founders produced offspring with no obvious phenotype and with identical truncation of mRNA as a result of gene alteration, only offspring of founder 38 was selected for *Ddi2*$^{protease\ defective}$ colony establishment.

Genomic DNA was acquired from tail biopsies collected at pup weaning or from yolk sac tissue collected at embryo harvest. Genotyping was performed as described in chapter 3.4.5.4. Primers were designed to anneal with sequences of *Ddi2* introns 5-6 and 6-7 outside the TALEN-targeted area. The PCR amplification product for wild-type allele was 1572 bp long and the allele with deletion of 550 base pairs (without exon6) resulted in 1072 bp length (see Figure 16B and C). Additional triplet of primers was used in nested PCR for determination of genotype of embryos under E9.5 or in cases of probable contamination by maternal sample as shown in Figure 16C.

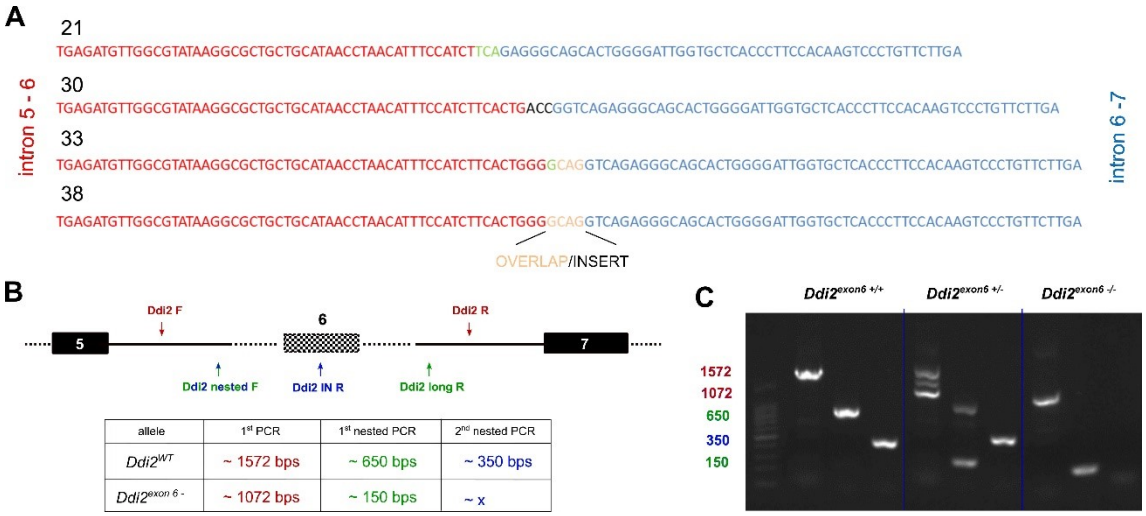


Figure 16: *Ddi2*$^{protease\ defective}$ strain: design of the TALEN-mediated *Ddi2* alteration and genotyping strategy. A) Sequencing results of the TALEN-altered area of *Ddi2* gene in four founder mice. **B)** Genotyping design of *Ddi2*$^{protease\ defective}$ strain. **C)** Agarose gel showing both first round PCR and two rounds of nested PCR required for genotyping of *Ddi2*$^{protease\ defective}$ mice.

4.2.2.4 *Ddi2* deficiency results in embryonic lethality

Heterozygous *Ddi2*^{tm1b} crossings failed to produce *Ddi2*$^{-/-}$ offspring, which reveals the embryonic lethality phenotype of *Ddi2*^{tm1b} mouse strain. *Ddi2*$^{-/-}$ embryos show development retardation from E12.5 and die prior E14.5 as observed from timed harvest of embryos (see Figure 17A). Only one heavily retarded embryo with beating heart was harvested at E14.5 (3 out of 4 *Ddi2*$^{-/-}$ embryos were already dead and undergoing resorption). No difference considering development timeline (normal number of somites at harvest on certain embryonal day) or retardation was observed for *Ddi2*$^{-/-}$ embryos prior stage E11.5 when compared to their *Ddi*$^{+/+}$ or *Ddi2*$^{+/-}$ littermates. Heterozygous littermates develop normally and shown no

obvious phenotype after birth or in adulthood under non-challenging conditions (see chapter 4.2.2.7).

Surprisingly, *Ddi2*$^{exon6\ -/-}$ embryos from heterozygous *Ddi2*^{exon6} crossings die in earlier stage of development, prior to E12.5 (see Figure 17B, for comparison of two mouse knockout models see Figure 17C). This might be the result of failure of Ddi2 protein production in *Ddi2*$^{-/-}$ (data not shown) as opposed to the expression of modified version of mDdi2 protein in *Ddi2*$^{exon6\ -/-}$ (see Figure 17D). The development of *Ddi2*$^{exon6\ -/-}$ up to E9.5 stage is normal compared to the WT littermates, however their yolk sac shows diminished vascularization (see Figure 17E). Moreover, embryos of harvests at E10.5 and E11.5 exhibit excessive growth retardation in both placenta and embryo proper, they lack mandible and heart development is delayed, even though they are still alive (see Figure 17F). Prenatal and adult heterozygous mice show no obvious phenotype, which is further discussed in chapter 4.2.2.7.

It is probable that the WT allele of *Ddi2* is capable of compensation for the *Ddi2*$^{-/-}$ allele and the *Ddi2*$^{exon6\ -/-}$ allele during embryonal development, as no anomalous phenotype was observed for heterozygous embryos.

A

Ddi2^{tm1b}						
Stage	Total # of embryos (# of litters)	*Ddi2* $^{+/+}$ (%)	*Ddi2* $^{+/-}$ (%)	*Ddi2* $^{-/-}$ (%)	*Ddi2* $^{-/-}$ resorbed (%)	N/A % (# of embryos)
E9.5	13 (2)	23	46	23	0	8 (1)
E10.5	23 (3)	13	43.5	30.5	0	13 (3)
E12.5	17 (2)	18	41	29	20	12 (2)
E14.5	35 (6)	23	66	11	75	0
E18.5	10 (2)	20	70	10	100	0

B

Ddi2$^{protease\ defective}$						
Stage	Total # of embryos (# of litters)	*Ddi2*$^{exon6\ +/+}$ (%)	*Ddi2*$^{exon6\ +/-}$ (%)	*Ddi2*$^{exon6\ -/-}$ (%)	*Ddi2*$^{exon6\ -/-}$ resorbed (%)	N/A % (# of embryos)
E9.5	181 (26)	21	52	25	2	2 (4)
E10.5	145 (17)	21.5	50.5	27	0	1 (2)
E11.5	56 (6)	20	50	23	23	7 (4)
E12.5	52 (5)	19	63.5	13.5	100	4 (2)

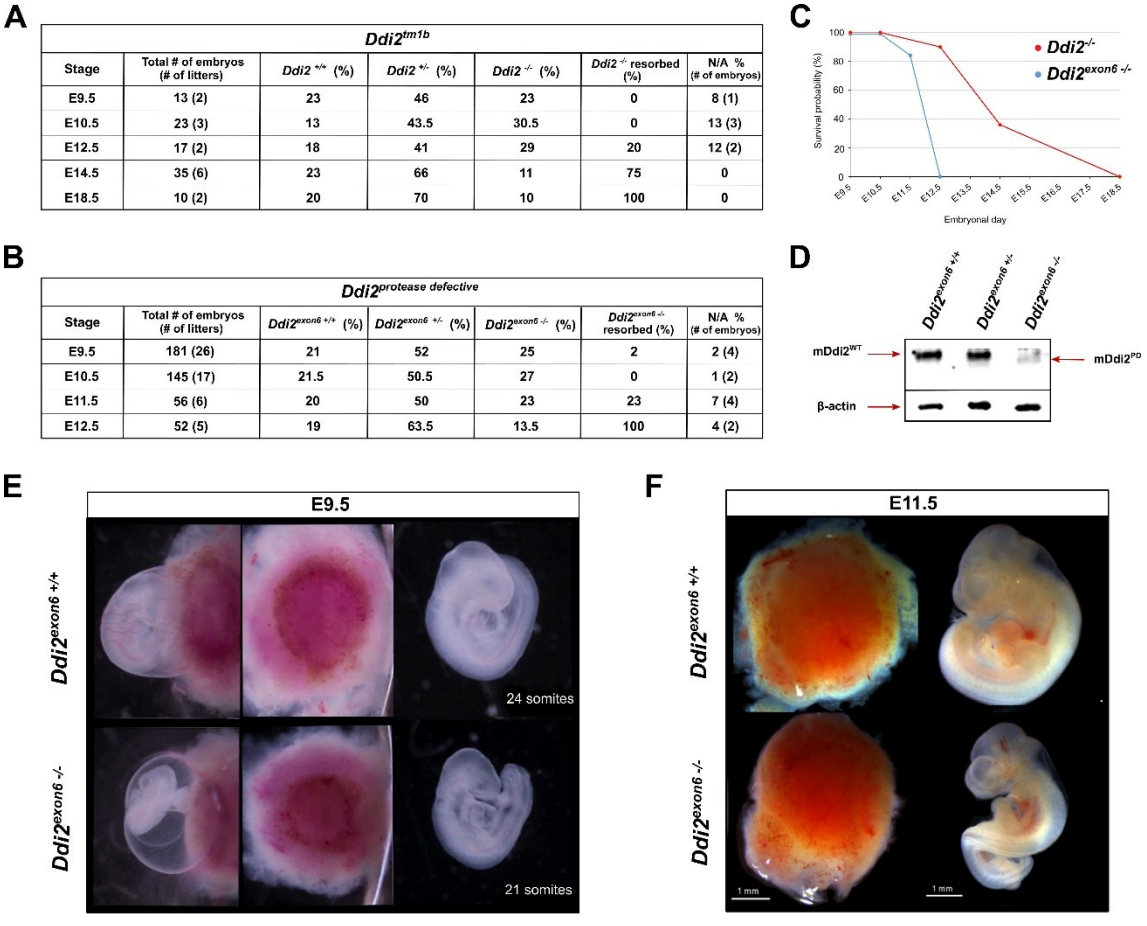


Figure 17: Ddi2 null and protease defective mice die during embryonal development. **A)** Table of embryo genotypes at different developmental stages of harvests from *Ddi2*$^{+/-}$ crossings. *Ddi2*$^{-/-}$

embryos die prior E14.5. **B)** Table of embryo genotypes of timed harvests from *Ddi2*$^{exon6\ +/-}$ crossings. *Ddi2*$^{exon6\ -/-}$ embryos die at E12.5. **C)** Kaplan-Meier plot of survival of both *Ddi2* mouse deficient models. **D)** *Ddi2*$^{exon6\ -/-}$ embryos produce a protease defective protein mDdi2^{PD} (Δ254-296), however it seems this truncation destines the protein to rapid degradation. **E)** Fresh preparation of *Ddi2*$^{exon6\ -/-}$ (down) embryo and placenta compared to the tissue of *Ddi2*$^{exon6\ +/+}$ (up) littermate at E9.5. **F)** Image of *Ddi2*$^{exon6\ -/-}$ (down) embryo and placenta compared to *Ddi2*$^{exon6\ +/+}$ (up) littermate at E11.5 reveals retardation in the *Ddi2*$^{exon6-/-}$ fetus. Images of fresh preparation were acquired by Kallayanee Chawengsaksophak.

4.2.2.5 Characterization of the mDdi2 protease defective protein

To characterize the mDdi2^{PD} (protease defective) protein construct expressed in *Ddi2*$^{exon6\ +/-}$ and *Ddi2*$^{exon6\ -/-}$ mice and to compare it with the mDdi2^{WT} protein, mRNA was isolated from *Ddi2*$^{exon6\ +/+}$ and *Ddi2*$^{exon6\ -/-}$ embryos at stage E9.5. Coding sequences of both proteins were amplified from cDNA and cloned into bacterial expression vectors p905 and pTreTight.

Recombinantly expressed and purified proteins were tested for proper folding using 1D NMR spectroscopy and differential scanning fluorimetry (DSF). As shown in Figure 18A, mDdi2^{WT} and mDdi2^{PD} protein 1D spectra correspond to proteins with acquired secondary structures. However, the DSF experiment revealed aggregation of the mDdi2^{PD} protein, as no melting temperature peak could be detected contrary to the mDdi2^{WT} (melting temperature was determined to 58.7 °C) (see Figure 18B).

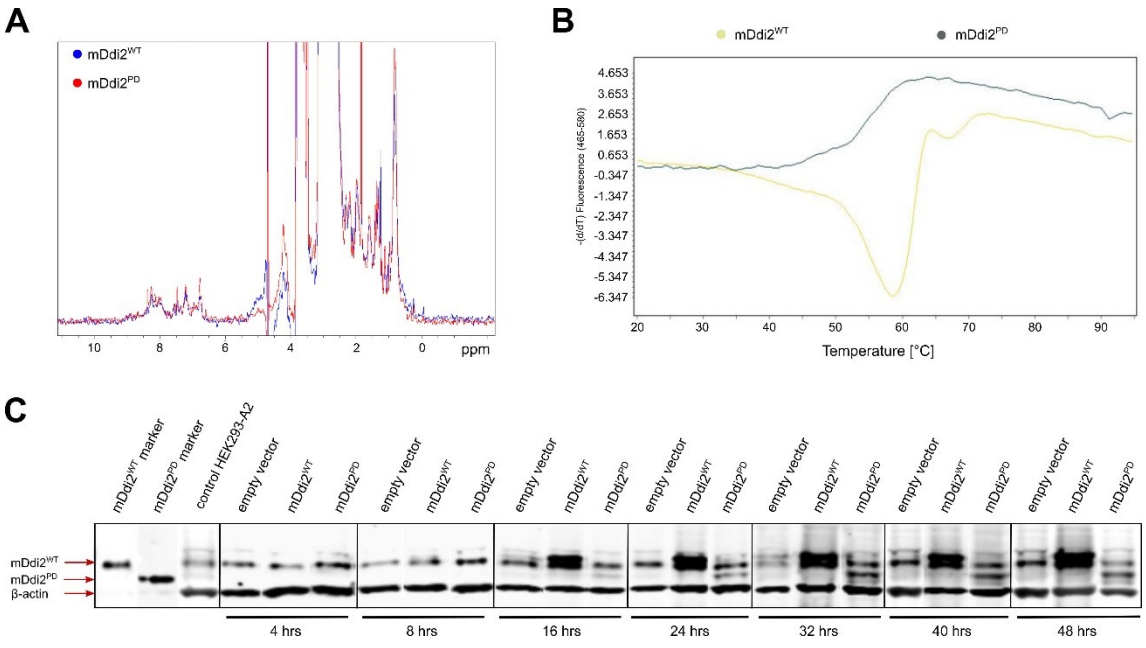


Figure 18: Characterization of mDdi2^{PD} protein expressed in *Ddi2*$^{exon6\ +/-}$ and *Ddi2*$^{exon6\ -/-}$ mice. **A)** Overlay of 1D NMR spectra of mDdi2^{WT} (blue) and mDdi2^{PD} protein (red). **B)** Melting peaks of mDdi2^{WT} (yellow) and mDdi2^{PD} (grey) proteins show impaired folding/aggregation of the truncated protein version. **C)** HEK293offA2 cells are not able to acquire high level of overexpression of mDdi2^{PD} protein compared to mDdi2^{WT}. Cells transfected with pTreTight vector encoding mDdi2^{PD}, mDdi2^{WT} or an empty vector were harvested after 4, 8, 16, 24, 32, 40 and 48 hours. β-actin was used as a loading control.

As shown in Figure 17D, we were able to detect mDdi2^{PD} protein with anti-DDI2 antibody recognizing the C-terminus of human DDI2 and its mouse ortholog, however only with very low signal just above background. We were not able to detect the mDdi2^{PD} specific peptide in Ddi2$^{exon6\ +/-}$ and Ddi2$^{exon6\ -/-}$ embryo lysates in MS experiments (data not shown). We therefore tested how normal cells (HEK293offA2) cope with overexpression of the mDdi2^{PD} protein construct. We transfected HEK293offA2 cells with mDdi2^{PD}, mDdi2^{WT} encoding or, empty vector and harvested cells in triplicates at different time points - 4, 8, 16, 24, 32, 40 and 48 hours after transfection. Interestingly, the increase in expression of the mDdi2^{PD} protein was smaller when compared to mDdi2^{WT} protein (see Figure 18C). Based on these studies, it seems mDdi2^{PD} protein is partially misfolded in the RVP domain region after translation, it might form aggregates and be quickly degraded.

4.2.2.6 Colony management for both *Ddi2^{tm1b}* and *Ddi2$^{protease\ defective}$* mouse strains

Colony of *Ddi2^{tm1b}* strain was established and common genotyping at weaning was performed by scientific staff of the IMG CAS; *Ddi2$^{protease\ defective}$* colony was managed and genotyped by Monika Sivá (author). As both models result in embryonic lethality, heterozygous *Ddi2$^{+/-}$* or *Ddi2$^{exon6\ +/-}$* mice were backcrossed with C57Bl/6NCrl wild-type mice for colony maintenance. Adult mice and embryos for phenotyping were collected from litter from heterozygote × heterozygote crossings.

4.2.2.7 Adult mice phenotyping and *Ddi2* expression studies

The phenotyping of adult mice of both *Ddi2*-altered strains was mainly performed by our colleagues from Czech center for Phenogenomics hosted by the IMG CAS under the supervision of Jan Procházka.

Ddi2^{tm1b} strain mice were subdued to standard screening which follows the international mouse phenotyping consortium pipeline (IMPC). These tests did not show any obvious difference or abnormality of the *Ddi2$^{+/-}$* mice. However, we observed problem in fertility in *Ddi2$^{+/-}$* mice crossings, where the pairs often failed to conceive despite positive vaginal plug.

Ddi2$^{protease\ defective}$ strain mice were subjected to smaller range screening of glucose metabolism (IpGTT), biochemistry, hematology, gross pathology and histology. No obvious phenotype features were observed in the adult *Ddi2$^{exon6\ +/-}$* mice when compared to *Ddi2$^{exon6\ +/+}$* and C57Bl/6Ncrl in-house WT mice.

4.2.2.8 Expression of Ddi2 during embryonal development

To better understand the lethal phenotype of *Ddi2*$^{-/-}$ and *Ddi2*$^{exon\ 6\ -/-}$ embryos during their development, we performed a qPCR experiment for estimation of *Ddi2* expression at the individual developmental stages and mapping of *Ddi2* expression via lacZ reporter gene in the *Ddi2*$^{+/-}$ embryos.

The *Ddi2* mRNA levels are increased 2-fold at embryonic stage E9.5 when compared to the basal expression levels at E10.5 and E11.5 (see Figure 19A).

Ddi2$^{+/-}$ embryos were used for *Ddi2* expression profiling via β-galactosidase (encoded by lacZ in the *Ddi2*^{tm1b}) activity screening with artificial substrate X-gal. Screening was performed on both whole-mount embryos and sagittal cryo-sections of embryos at different developmental stages. At stage E9.5, expression of *Ddi2* occurs in rapidly developing body parts, such as forelimb, hindlimb and tail buds, heart, and maxillary and mandibular arches (see Figure 19B). Later, at stage E12.5, *Ddi2* expression becomes ubiquitously spread all over the embryonal body, as shown in both whole-mount and paraffin section lacZ staining (see Figure 19C). Expression profile at later developmental stages is more specifically localized, e.g. sagittal section of head reveals expression in specific layer of cortex in forebrain and in trigeminal ganglion (E14.5, see Figure 19D, E17.5, see Figure 19F). Positive staining was as well observed in other ectodermal tissues, such as olfactory epithelium or skin (E14.5, Figure 19D) and in mesodermal tissues, such as smooth muscle of heart or cranium (E14.5, Figure 19D and E). In addition, positive *Ddi2* expression in fetal liver in both E14.5 and E17.5 stages is localized in cells that might correspond to fetal macrophages differentiating into Kupffer cells, and with very low signal in hepatocytes themselves (see Figure 19E and 19F, respectively). Negative controls acquired for each of the hereby studied developmental stages did not reveal any false positive signal in the lacZ staining data (some of the negative control data is shown in Figure 19B, D and E).

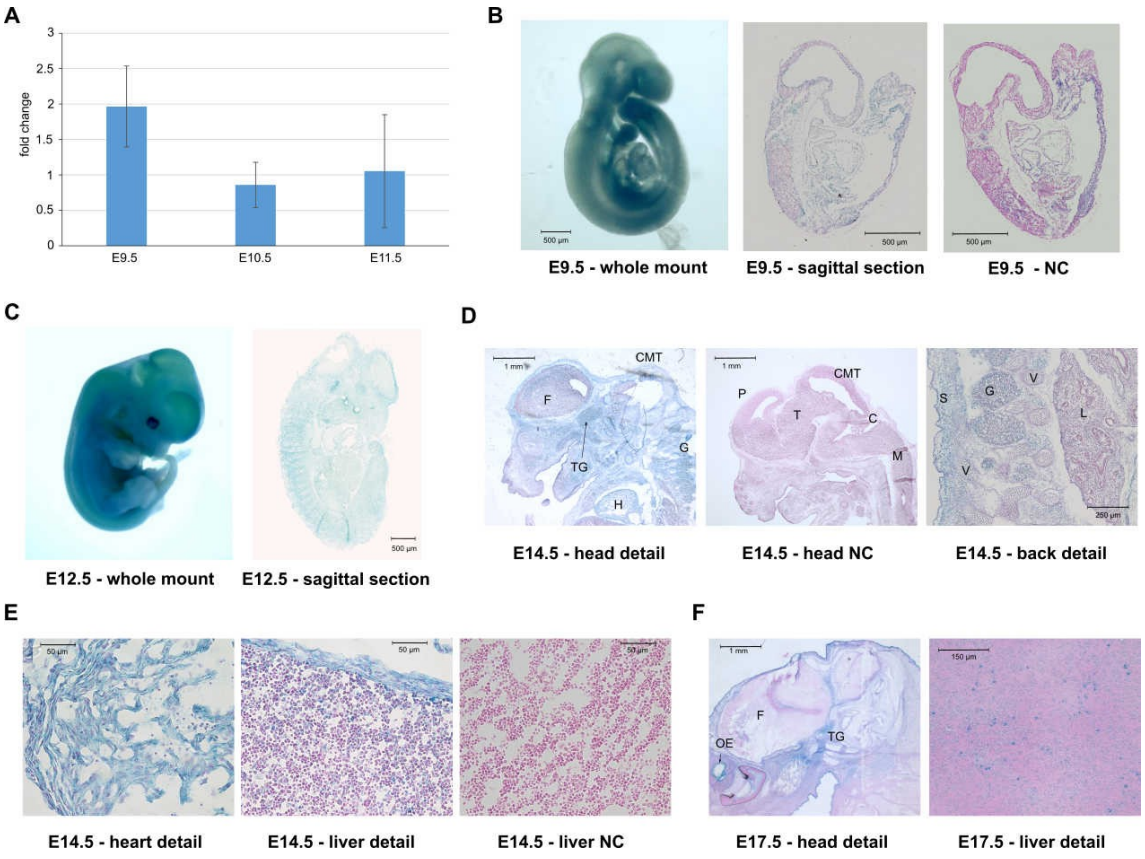


Figure 19: Expression profiling of *Ddi2* during embryonal development. **A)** Estimation of *Ddi2* expression at stages E9.5, E10.5 and E11.5 of embryonal development using qRT-PCR. Transcript levels of *H2afz* were used for normalization. Error bars denote SD (n=6). **B)** Mapping of *Ddi2* expression at stage E9.5 using lacZ reporter gene in *Ddi2*$^{+/-}$. **B)** *Ddi2* is expressed in developing parts of embryonal body at stage E9.5, such as limb buds, heart and mandibular and maxillary arches. **C**) *Ddi2* expression is ubiquitious at E12.5. **D)** Sagittal sections show specific localization of *Ddi2* expression at stages E14.5, in skin, brain (forebrain and trigeminal ganglion) and cranium. **E)** Smooth muscles of heart tissue and cells that most presumably correspond to Kupffer cells in liver at stage E14.5 also exhibit *Ddi2* expression. **F)** Expression of *Ddi2* at E17.5 in sagittal section of head (depicts localization in skin, trigeminal ganglion and olfactory epihelium) and liver. Abbreviations: C – cerebellum, F – forebrain CMT – colliculus midbrain tectum, G - ganglion, H – heart, L – lung, OE – olfactory epithelium, P – pallium, S – skin, M – medulla, T - thalamus, TG – trigeminal ganglion, V – vertebra

4.2.2.9 Activation of Nrf1 is diminished in both mouse model strains

To verify whether mDdi2-Nrf1 interplay could be the reason of embryonal lethality, experiments of Nrf1 activation via proteotoxic stress were performed in MEFs isolated from E10.5 embryos of each genotype and strain. Primary fibroblasts were cultured to passage 3, when the experiments were performed after the cells acquired more than 85% confluence. Both *Ddi2*$^{-/-}$ and *Ddi2*$^{exon6\ -/-}$ fibroblasts showed proliferation retardation in comparison to the cells isolated from wild-type and heterozygous embryos of both strains (data not shown). First, timeline setup and proteasome inhibitor screen was performed for refinement of experimental conditions (data not shown). The most representative results with the highest

response to proteasome inhibition in Nrf1 activation were acquired by treatment with 10 µM MG132 for 16 hours.

In wild-type and heterozygous cells of both strains, the treatment with MG132 resulted in overexpression (Nrf1 FL) and activation of Nrf1 (Nrf1 A) (see Figure 20A and B). In contrast, both *Ddi2*$^{-/-}$ and *Ddi2*$^{exon6\ -/-}$ fibroblast cultures failed to produce an adequate response to proteasomal inhibition, by only accumulation of the full-length Nrf1 with missing activation of Nrf1 by mDdi2 cleavage (Figure 20A and B).

A weak signal corresponding to mDdi2 full-length protein can be observed in the lysates of *Ddi2*$^{exon6\ -/-}$ fibroblasts (see Figure 20B). We have previously observed this phenomenon in human DDI2 KO cell lines (HTC116) prepared in our laboratory. Additionally, smears around this area also occur in *Ddi2*$^{-/-}$ fibroblast lysates. This could be result of high total protein load onto the SDS-PAGE gel (see β-actin loading control in Figure 20B) and therefore, we consider this to be a non-specificity of the antibody.

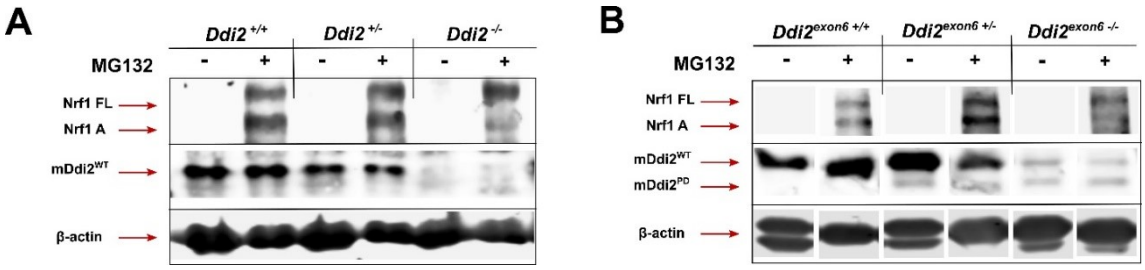


Figure 20: MEF culture derived from Ddi2-deficient embryos exhibit impairment in "bounce back" effect of Nrf1 protein activation. MEFs were treated with 10 µM MG132 for 16 hours for activation of Nrf1 protein by Ddi2 cleavage. Ddi2-deficient cells derived from both *Ddi2*^{tm1b} (A) and *Ddi2*$^{protease\ defective}$ (B) model strains fail to cleave Nrf1 sufficiently.

Impairment in Nrf1 pathway activation in Ddi2-deficient model cells is one of the first evidence supporting the hypothesis of Ddi2-Nrf1 interplay and its essential function during embryonal development. Whether failure of Nrf1 activation might be the main cause of the lethality shall be investigated in future studies. This hypothesis will be discussed further in the discussion.

4.2.2.10 My contribution to the project

Since this thesis represents a part of a comprehensive study of our laboratory, results presented here were generated by several members of our laboratory and our collaborators from the CCP unit hosted by IMG CAS. My specific contribution to the project is enlisted below.

1. Cloning, recombinant expression and purification of several protein constructs that were used for their characterization on molecular level.

2. Solving of structure of yDdi1 UBL using NMR spectroscopy.

3. Identification and characterization of binding properties of individual domains of human DDI2 protein using NMR spectrometry (titration experiments and binding site mapping).

4. Selection of founder mouse and colony management of *Ddi2*$^{protease\ defective}$ mouse strain.

5. Optimization of all protocols for protein purification, genotyping procedure, tissue lysis, cell lysis and Western blotting.

6. Majority of chromosomal DNA isolations and genotype identification of mice or embryos.

7. Harvest of majority of embryonal samples of *Ddi2*$^{protease\ defective}$, their subsequent processing for mRNA isolation or protein expression analysis using Western blotting.

8. Isolations and quality control of mRNA and reverse transcription.

9. Processing of raw qPCR data with Vendula Novosadová (CCP).

10. Harvest of all embryos for MEF isolation and realization of all related experiments.

11. Majority of Western blotting analysis.

12. *Ddi1* expression studies using ISH were performed in collaboration with Michaela Procházková as further specified in chapter 3.4.5.16.

13. Part of the lacZ staining experiment together with our collaborators from CCP. Complete *Ddi2* expression analysis based on lacZ staining data.

14. Design of most experimental procedures and data analysis related to both mouse strains.

15. Preparation of drafts of all manuscripts involving data obtained during my work.

5 DISCUSSION

Maintenance of homeostasis is the key to cell survival and to the fulfillment of its function in the concept of whole organism. Homeostasis is acquired by highly regulated processes, such as protein quality control mechanisms or responses to manifold stress inducers (Galluzzi L.*, et al.*, 2018, Walter P. and Ron D., 2011). In addition to the function as proteins targeting towards degradation in the ubiquitin-proteasome system, several members of the Ddi1-like family of shuttling proteins have been recently identified as important actors in DNA damage repair and regulation of protein expression by activation of a specific transcription pathway in response to proteotoxic stress (Kaplun L.*, et al.*, 2005, Koizumi S.*, et al.*, 2016, Kottemann M. C.*, et al.*, 2018, Lehrbach N. J. and Ruvkun G., 2016). Despite these few recent findings, that suggest an essential function of Ddi1-like protein family in homeostasis maintenance, all its members have been understudies in general. This dissertation presents a broad study of several members of the Ddi1-like protein family, focusing both on their characterization on molecular level and their role in relevant biological systems, specifically in mouse knockout models.

Ddi1-like protein family members have a unique domain architecture among other shuttling proteins of ubiquitin-proteasome system represented by Rad23 and Dsk2 (for domain architecture comparison see Figure 4 on page 48) (Bertolaet B. L.*, et al.*, 2001). Most of the non-mammalian Ddi1 orthologs harbor the conserved N-terminal UBL and the C-terminal UBA domains that in general facilitate the primary role of shuttling proteins (Bertolaet B. L.*, et al.*, 2001, Elsasser S.*, et al.*, 2002). Recently, we structurally characterized and structurally characterized a novel helical domain of Ddi (HDD), which exhibits similarity to DNA-binding domains, in both yDdi1 and hDDI2 protein (Siva M.*, et al.*, 2016, Trempe J. F.*, et al.*, 2016). Analogous helical domains of Sti1-like family had been previously described for Rad23 and Dsk2 (Kaye F. J.*, et al.*, 2000, Kim B.*, et al.*, 2005, Lee J. H.*, et al.*, 2005, Masutani C.*, et al.*, 1997). However, in addition to these regions, Ddi1-like proteins possess a central RVP domain, which gives them their uniqueness (Krylov D. M. and Koonin E. V., 2001, Sirkis R.*, et al.*, 2006).

Primary sequences and hence the structure of individual domains of Ddi1-like protein family members are very well conserved from yeast to mammals (see the analysis of sequence identity and structural superimpositions in Figure 7 and Figure 8, respectively). It is therefore presumed that they might adopt similar binding properties and perform akin functions.

The unique RVP domain of Ddi1-like proteins adopts a conserved structure, which highly resembles the structure of HIV-1 protease. It forms a homodimer of the full-length protein with catalytic aspartate positioned in the center of the substrate cavity protected with

flaps (Sirkis R.*, et al.*, 2006, Siva M.*, et al.*, 2016, Trempe J. F.*, et al.*, 2016). The catalytic function of the RVP domain was first indirectly reported in complementation studies of yDdi1p knockout strain, where the inactive mutant was not able to rescue the observed secretion phenotype (White R. E.*, et al.*, 2011b). The hypothesis of proteolytic activity was recently supported by studies of both human DDI2 and *C. elegans* Ddi1 activity under proteotoxic stress. Human DDI2 and *C. elegans* Vsm-1 specifically cleaves its substrate, a transcription factor NRF1 (Skn1), which is thereby activated and translocated into nucleus to fulfill its function (Koizumi S.*, et al.*, 2016, Lehrbach N. J. and Ruvkun G., 2016). NRF3 was also identified as another specific substrate of hDDI2 (Chowdhury A.*, et al.*, 2017). Although the mechanism of cleavage has not yet been understood, the fact that DDI2 activates these two transcription factors and that DDI2 clearly plays an important role in DNA repair, puts DDI2 (specifically its protease domain) among clinically relevant targets.

Based on sequence alignment of members of Ddi1-like family in Figure 7A and on our structure prediction analysis, the four helical bundle of HDD is present in all four mouse and human Ddi1-like proteins and in Ddi1 (Rngo) of *D. melanogaster* (this work, (Siva M.*, et al.*, 2016)). As mentioned previously, HDD is structurally similar to Sti1-like domains of the shuttling proteins that mediate protein-protein interactions (Kaye F. J.*, et al.*, 2000, Masutani C.*, et al.*, 1997, Trempe J. F.*, et al.*, 2016). Similar protein interface could thus be represented by HDD, or it could serve as interaction platform for substrate of the RVP domain and mediate transport of the substrate into the active site. HDD could also mediate interaction with DNA at site of DNA damage, based on resemblance of the N-terminal helical bundle of yDdi1 and DNA-binding domains of transcription factors, such as CUT domain of SATB or bacteriophage λ cII transcription activator (Trempe J. F.*, et al.*, 2016). Moreover, the two additional C-terminal helices of yeast HDD together with RVP are crucial for DNA damage response in yeast upon hydroxyurea treatment (for details see chapter 1.3.1) (laboratory of Dr. Grantz Šašková – unpublished data). We hypothesize that these two helices represent the substrate interaction site for the yDdi1p. The function of the four helical bundle of Ddi1-like proteins, which have lost the two adjacent C-terminal helices throughout evolution, remains to be clarified.

The dimerization of the RVP and the position of the UBL and UBA domains at opposite ends offer two configurations of the full-length protein, a head-to-head or a head-to-tail homodimers. The fact that the UBL domain of yDdi1p binds ubiquitin allowed to propose a novel alternative shuttling mechanism, where both UBL and UBA would bind the polyubiquitin chain in a head-to-tail configuration. The putative dual functionality of the

hDDI1 and hDDI2 UBLs (both lacking UBA) would hence allow to bind both the polyubiquitin chain and the proteasome (Nowicka U*., et al.*, 2015). This hypothesis was disapproved by results reported here, showing that despite the resembling fold, hDDI2 UBL and yDdi1 UBL differ in their surface properties: hDDI2 UBL does not possess the highly negatively charged β-sheet patch identified for yeast UBL, which interacts with the positively charged UBQ (Nowicka U*., et al.*, 2015, Siva M*., et al.*, 2016, Trempe J. F*., et al.*, 2016). The low affinity of the hDDI2 UBL to ubiquitin was further confirmed in NMR titration experiments also in this work (see Figure 11).

In connection to ubiquitin interactions, both yDdi1p and hDDI2 harbor a ubiquitin interacting region at their C-terminus, a UBA and a UIM, respectively (Nowicka U*., et al.*, 2015, Siva M*., et al.*, 2016). As sequence alignment in Figure 7A shows, the UIM sequence is conserved as well in mDdi2 protein, however it is absent in both human and murine Ddi1 homologs. We managed to characterize the very weak yet specific interaction between ubiquitin and the C-terminal UIM of hDDI2 in NMR titration experiments. In addition, we mapped the UIM interaction site onto the Ile44 patch of ubiquitin (see Figure 10). However, we were not able to pull down any of the di-ubiquitins chain types, which altogether defines the binding properties to be quite divergent from those previously described for the Ddi1p of *S. cerevisiae* (Siva M*., et al.*, 2016).

Regarding the overall structural properties of the full-length Ddi1 proteins, the human and yeast orthologs also exhibit different character. Our SAXS modeling data showed that yDdi1 with central protease dimer is flanked with rather flexible linkers to HDD and UBL domains. The data is supported by NMR experiment, where we did not observe any interaction between UBL and HDD-RVP constructs (Trempe J. F*., et al.*, 2016). This is in agreement with Nowicka and colleagues, who reported that the individual UBL, RVP and UBA subunits of yDdi1p do not interact with each other according to NMR experiments (Nowicka U*., et al.*, 2015). The SAXS modeling of the hDDI2 RVP-HDD dimer showed asymmetric distribution specific for elongated proteins. This is caused by the long linker between the RVP and HDD domains of hDDI2 (40 residues), which is not present in the yeast ortholog (this work, (Siva M*., et al.*, 2016, Trempe J. F*., et al.*, 2016)). On the contrary, based on the difference in the 2D HSQC spectra of sole UBL domain and the full-length hDDI2 superimposition performed in our study (Figure 12A), the UBL domain interacts with the body of the protein, which suggests that the hDDI2 ortholog forms rather a compact dimer than an extended protein (see chapter 4.1.3.2 and Figure 12). In addition, we observed that the UBL of hDDI2 does not interact with the C-terminal UIM motif and therefore it is unlikely

that it would constitute a head-to-tail dimer. Altogether, we reason that while the HDD-RVP protein part is extended in the center of the dimer, both UBL and UIM sweep back towards the center of the protein. This, together with the low affinity of hDDI2 UBL towards ubiquitin (this work, (Siva M*., et al.*, 2016)), is contradictory to the proposed model of the so-called "alternative shuttle" (Nowicka U*., et al.*, 2015).

Despite the study of Kottemann and colleagues, who observed association of hDDI2 with proteasomal subunits in crosslinking/mass spectrometry experiments (Kottemann M. C*., et al.*, 2018), direct interaction with any of the intrinsic ubiquitin receptors of the proteasome has not been shown either for human or mouse members of the Ddi1-like protein family. In fact, the multidomain architecture of Ddi1-like protein family members opens the discussion for their multifunctional potential. The diverse features and binding properties of their otherwise conserved domains suggest that the individual orthologs and mammalian homologs might have evolved different cellular roles, which remain to be further investigated.

In order to study the role of both human homologs in biological systems, we chose mouse as a suitable, and experimentally feasible biological model organism, based on sequential and structural similarities between human and murine Ddi1-like family members (see Figure 7A and Figure 8 on pages 100 and 101, respectively).

First, we focused on the understudied Ddi1 protein. Our colleagues from the laboratory of Dr. Ugo Mayor identified Ddi1 (Rngo) from *D. melanogaster* as a unique substrate of the E3 ubiquitin ligase Ube3a via unbiased ubiquitin proteomics approach (Franco M*., et al.*, 2011, Ramirez J*., et al.*, 2018). They also identified the human DDI1 protein to be ubiquitinated by UBE3A in human SH-SY5Y neuroblastoma cells. Interestingly, this specific modification does not target the substrate (hDDI1) towards degradation in the proteasome and clearly has other signaling function (Ramirez J*., et al.*, 2018). UBE3A dysregulation is tied with a complex neurodevelopmental disorder called Angelman syndrome. How mutations in *UBE3A* gene influence the development of nervous system is not fully understood. It has been suggested that substrates of UBE3A could play specific and/or essential roles in brain development (Buiting K*., et al.*, 2016, Sadikovic B*., et al.*, 2014). Our *Ddi1* expression profiling using ISH in developing brain of mouse embryos revealed specific expression localization of this gene throughout different stages of mouse brain development. The ventricular layer and cortical plate of isocortex and the ventricular layer of olfactory bulb exhibit specific expression of *Ddi1* at stage E16.5 (see Figure 13H-J), when the expression significantly increases multiple-fold in comparison to adjacent developmental stages (Ramirez J*., et al.*, 2018). Based on the similar expression profile of the

mouse and human Ddi1 proteins in adults, and the expression and specific ubiquitination of hDDI1 in human SH-SY5Y neuroblastoma cells, we suggest that hDDI1 could presumably be upregulated during development of human brain. In addition, several mutations of *DDI1* were recently identified and linked to a familial neurodegenerative disorder, which further supports possible function of *DDI1* in neuronal tissue (Alexander J.*, et al.*, 2016). Therefore, it might represent one of the substrates of UBE3A with role in fetal brain that could be of importance in connection to Angelman syndrome and possibly other neuronal disorders and their clinical research.

Next, we focused on deciphering the function of mammalian Ddi2 protein and generated two knockout mouse models. Human and mouse Ddi2 proteins share 96% sequence identity and highly conserved structural properties. As shown in Figure 7A on page 100, mouse Ddi2 harbors the HDD domain that precedes the RVP and the C-terminal UIM described in hDDI2 protein (Siva M.*, et al.*, 2016).

Two mouse model strains with alteration of *Ddi2* gene were generated, a full knockout and a Ddi2 protease defective strain. *Ddi2*^{tm1b} strain, which was generated by ESC manipulation and insertion of gene cassette encoding loxP sites intended for removal of critical exon 2 by Cre recombinase and a reporter-gene insert, represents the full knockout (Mansour S. L. *et al.*, 1990). Ddi2 protease defective strain was generated in order to abolish the functions of the RVP domain – catalytic activity and dimerization of the full-length protein. Therefore, this second mouse model strain (C57Bl6/NCrl-Ddi2$^{em1/Rase}$, here distinguished as *Ddi2*$^{protease\ defective}$) was generated by TALEN-mediated excision of exon 6, which resulted in alteration of the RVP domain. The mDdi2^{PD} protein is missing region of residues 254-296 that encodes catalytic aspartate and a part of the dimerization domain of the RVP. Expression of the altered mDdi2 protein variant was confirmed on mRNA level (sequencing of corresponding cDNA) and analysis of protein expression in embryo lysates using Western blotting. We were able to repeatedly observe small band with low signal at 44 kDa, where the mDdi2^{PD} variant is expected (see Figure 17D on page 116). Despite our effort, we failed to detect the protease defective variant in embryo or derived MEF lysates using mass spectrometry.

Ddi2 gene belongs to the one third of mammalian genes that are essential for life, as both mouse models show embryonic lethality at mid-late gestation period (Dickinson M. E. *et al.*, 2016). Strikingly, the homozygous individuals die at different developmental stages: while *Ddi2*$^{-/-}$ embryos die by E14.5 (mid-late gestation), *Ddi2*$^{exon6\ -/-}$ embryos die earlier, by E12.5 (mid-gestation) (for comparison of survival of embryos of both strains see

Kaplan-Meier curve in Figure 17C). In both models, the heterozygous embryos and adult mice do not exhibit any obvious phenotype, as revealed by our phenotyping study described in chapter 4.2.2.7, which suggests that the original allele might be able to entirely supplement the missing functional allele under non-challenging conditions. Despite no obvious phenotype, we have observed defects in conception after identification of vaginal plug in the heterozygous crossings of *Ddi2*^{tm1b} strain mice.

Ddi2$^{-/-}$ embryos show retardation at stage E12.5 when compared to their littermates, and die by E14.5. We have not yet entirely described the molecular and physiological reason of their death *in utero* at this specific stage. Due to fertility problems of the heterozygous adult mice, it is complicated to acquire sufficient number of embryonal samples. However, part of the harvested embryonal samples we have managed to acquire (see Figure 17A on page 116), is being processed for future studies, such as µCT scanning at the CCP embryology unit for detailed analysis of the retardation or qPCR analysis targeted on the influence of Ddi2 ablation on Nrf1-driven pathways. In contrast, as *Ddi2*$^{exon6\ +/-}$ adult mice exhibit normal rate in conception, we were able to harvest a significantly higher number of embryo samples and perform more experiments using this model. *Ddi2*$^{exon6\ -/-}$ embryo retardation starts after E9.5, as it is already distinguishable at stage E10.5 and quite visible at E11.5 (see Figure 17F, page 116). The embryos are smaller, they have lower number of somites at the same stage of harvest, lack mandibula and maxilla, the yolk sac is pale and exhibits less vascularization when compared to the littermates. The qPCR and lacZ staining experiments (see Figure 19 on page 120) revealed quantitative and qualitative data on *Ddi2* expression. The embryos at stage E9.5 exhibit two fold higher expression when compared to later developmental stages. This upregulation could be connected to the early onset of retardation in growth of the *Ddi2*$^{exon6\ -/-}$ embryos compared to the *Ddi2*$^{-/-}$ embryos, possibly explained by a dominant negative effect of the mDdi2^{PD} protein (will be further discussed below). They also shown defects in development of body parts with specific expression of *Ddi2*, based on our lacZ expression studies (see Figure 19B). The expression profile of *Ddi2* in both stages E9.5 and E12.5 acquired by lacZ staining were consistent with our previously obtained data from *Ddi2* ISH studies on whole mount embryos (data not shown). Nevertheless, the reason of difference in stage of embryonic lethality of our two mouse model strains has to be further investigated.

Next, we focused on the reason of embryonic death of our model mDdi2^{WT} deficient mice. In general, high percentage of embryonically lethal knockout mouse strains exhibit defects in development of extraembryonic structures (yolk sac and placenta) in addition to the retardation of the embryo proper (Perez-Garcia V. *et al.*, 2018). Yolk sac is essential for

nutrition supply during early embryo gestation prior to chorioallantoic attachment at E8.5 that thereafter provides nutrition for the rapidly growing embryo from the mother (Brett K. E. *et al.*, 2014, Cross J. C. *et al.*, 2003, Munro H. N. *et al.*, 1983, Rossant J. and Cross J. C., 2001). Vasculogenesis first occurs in the yolk sac prior to vascular system development in the embryo proper (Boucher D. M. and Pedersen R. A., 1996). Furthermore, hematopoiesis during embryonal development starts in a primitive form already in the yolk sac. It is initiated with primitive erythroid progenitors in the first wave between E7.25 – E9.0, and in the second wave with definitive erythroid progenitors between E8.25 – E10.0 (Dieterlen-Lievre F., 1978, Lux C. T. *et al.*, 2008, Palis J. and Yoder M. C., 2001, Wong P. M. *et al.*, 1986, Yamane T., 2018). At stages E10.0 – E11.0, the hematopoietic cells are transferred into fetal liver (Houssaint E., 1981, Zovein A. C. *et al.*, 2010). Defects in placenta, representing the main nutrient supplier from E9.5 onward, are also closely linked to the embryo proper retardation. The primary placental phenotype shall be studied in the models that exhibit embryonic lethality, so that malnutrition as the primary cause of embryo proper retardation could be excluded. At the stage of formation of placenta-embryo connection at E8.5, the embryos are challenged by a major change in metabolism. Due to connection with maternal blood between maternal and fetal capillaries in placental labyrinth zone, the metabolism changes from glycolytic to oxidative (Bulusu V. *et al.*, 2017, Shepard T. H. *et al.*, 1997, Watson E. D. and Cross J. C., 2005). Here, at the mid-gestation stage, the primitive erythroid cells are essential for oxygen supply to peripheral tissue of the embryo proper (Yamane T., 2018). Interestingly, hypoxia is a very important modulator of vascularization in both extraembryonic and embryonic tissues (Dunwoodie S. L., 2009). All these above-mentioned developmental processes could be connected to the function of Ddi2 protein and are therefore further discussed below in comparison with our *Ddi2*-deficient mouse model strains.

As shown in experiments on MEF cultures derived from our model *Ddi2*$^{-/-}$ and *Ddi2*$^{exon6\ -/-}$ embryos, the activation of the transcription factor Nrf1 is diminished under proteotoxic stress (Figure 20 on page 121). We used proteasomal inhibition by MG132, as this experimental method has been previously established in several studies and the effect of Ddi2 functional deficiency can be easily visualized by Western blotting (studies from our laboratory, (Radhakrishnan S. K.*, et al.*, 2010, Sha Z. and Goldberg A. L., 2014, Xiang Y. *et al.*, 2018). When we compared our mouse models with the model knockout mice of the *Ddi2* substrate *Nrf1*, the *Ddi2*$^{-/-}$ embryos exhibited lethality at similar developmental stage. Functional *Nrf1* knockout embryos (bearing disruption of the CNC bZIP domain) die due to hematopoiesis failure in liver (Chan J. Y.*, et al.*, 1998) and quite interestingly, they do not

show any obvious phenotype prior death except for growth retardation and anemia. As *Ddi2*$^{-/-}$ embryos exhibit growth retardation, but not any specific developmental defects in individual body parts with specific *Ddi2* expression (similarly to the *Nrf1*$^{-/-}$ embryos), the reason for their death in mid-late gestational stage could be as well caused by impaired processes in the yolk sac, for example the dismantled hematopoiesis due to failure of Nrf1 activation by Ddi2. This hypothesis is as well supported by our preliminary data from placental rescue of *Ddi2*^{tm1b} model strain (Sox-2 Cre driver, data not shown), in which *Ddi2*$^{-/-}$ embryos die at stage E18.5 – P0. This means, the lethality could be most likely induced by impairment of processes in placenta, in addition to the processes in the yolk sac and embryo proper. Similarly, the onset of retardation in the fetus and in the extraembryonic tissues (vascular network in the yolk sac and growth retardation of placenta) of *Ddi2*$^{exon6\ -/-}$ embryo at stage E9.5 could be connected as well to erythropoiesis, but in earlier phase of the hematopoiesis pathway. Moreover, this profound lethal phenotype might be based on the attachment of embryo to the placenta after E8.5, when it is challenged with offset of oxidative processes (Bulusu V.*, et al.*, 2017, Shepard T. H.*, et al.*, 1997, Watson E. D. and Cross J. C., 2005). Note, that one of the main roles of Nrf1 has been described in response pathways to oxidative stress (see chapters 1.2.2 and 1.3.2.1.1) (Chan J. Y.*, et al.*, 1998, Venugopal R. and Jaiswal A. K., 1998). The diminished vascularization observed in yolk sac of the *Ddi2*$^{exon6\ -/-}$ embryos might be the result of inability of the Ddi2 protease to activate Nrf1 that leads to excessive production of ROS. The role of nitric oxide and elevation in ROS production was reported in several studies to modulate vasculogenesis in the yolk sac (Nath A. K. *et al.*, 2004, Wang G. *et al.*, 2016). In conclusion, both Ddi2 deficient and protease defective model strain mice show phenotype that might be linked with its function in Nrf1 activation, while the *Ddi2*$^{exon6\ -/-}$ embryos exhibit profound sensitivity to developmental challenges as they die at earlier gestation stages. Furthermore, the loss of activation of the transcription factor Nrf3, another Ddi2 substrate that exhibits high expression levels in placenta, may second the loss of function in Nrf1 activation pathway and therefore contribute to the observed phenotype. This needs to be further verified (Chenais B.*, et al.*, 2005, Chowdhury A.*, et al.*, 2017). In addition, the embryonic lethality of our model strains could be linked to the involvement of hDDI2 in DNA repair processes. Previous studies of genes functionally related to mDdi2, such as *DVC-1* (murine ortholog of metalloprotease Wss1) and Rad23B, exhibit pre-implantation lethality and offset of lethality during embryonal development at E13.5 with only 10% survival of *Rad23B*$^{-/-}$ mice up to adulthood (Maskey R. S. *et al.*, 2014, Ng J. M. *et al.*, 2002).

We considered the observation, that heterozygous adult mice of both strains do not show any abnormal phenotype, very interesting and wanted to reveal, whether these animals would respond to stressing conditions. Based on previous study of Lee and colleagues on Nrf1$^{+/-}$ mice, the author Monika Sivá in collaboration with the CCP hosted by IMG CAS performed a full phenotyping screening of *Ddi2*$^{exon6\ +/-}$ aged males that were challenged with administration of proteasomal inhibitor bortezomib (data not shown). These animals did not show any obvious phenotype and in contrast to *Nrf1*$^{+/-}$ mice, which exhibited enhanced sensitivity to ER stress and steatosis upon inhibition of proteasome (Lee C. S.*, et al.*, 2013). *Ddi2*$^{exon6\ +/-}$ mice did not evolve any pathologic condition in the liver that could be detected in biochemistry test from murine blood or in histopathology of the livers (laboratory of Dr. Grantz Šašková – unpublished data). These findings additionally support the full complementation of *Ddi2* function in the heterozygous mice by the *Ddi2*^{WT} allele. However, based on the same study of Lee and colleagues and the lethal phenotype of our mouse model strains, we generated a liver-specific *Ddi2* knockout mouse strain. Based on our preliminary data (data not shown), the animals seem to be prone to elevated triglyceride content under non-challenged conditions (laboratory of Dr. Grantz Šašková – unpublished data). These findings of similarity between the knockout model strain mice link abolishment of *Ddi2* function with the function of *Nrf1* and suggest the need of further studying and understanding of the interplay of these two essential genes.

In order to decipher the differential lethality between the two model strains and validate proper folding of the altered protein, we performed characterization of the mDdi2^{PD} protein. To do so, we first isolated the mRNA of both alleles of *Ddi2*$^{protease\ defective}$ strain, the WT and the Δ254-296, prepared corresponding cDNA and cloned the protein coding sequences into bacterial expression vectors. Both mDdi2^{WT} and mDdi2^{PD} protein variants were recombinantly expressed and subjected to several studies. 1D NMR spectra showed proper folding in the secondary structures of both protein variants (Figure 18A, page 117), however, we were not able to estimate the melting temperature of mDdi2^{PD} using differential scanning fluorimetry, as the protein did not undergo any denaturing processes that would allow binding of the Sypro® Orange fluorescent dye into revealed hydrophobic regions upon denaturation (see Figure 18B, page 117). Both protein variants were also characterized using a semi-analytic chromatography and dynamic light scattering (data not shown), which revealed increased molecular mass of the recombinantly expressed mDdi2^{PD} protein when compared to mDdi2^{WT}. It seems, mDdi2^{PD} forms low mass aggregates right after translation in bacterial cells during its recombinant expression. We wanted to verify, whether human

cells could cope with expression of the altered mDdi2^{PD} protein variant and therefore we performed transient expression studies of both mDdi2^{WT} and mDdi2^{PD} with analysis of protein expression in harvested cells at several time points after transfection (see Figure 18C, page 117). This experiment revealed higher expression of the wild-type protein variant and only weak signal for the protease defective variant, even at 48 hours after transfection. Our observations led us to assume that aggregation of the protein via unstructured part of RVP domain in regions that originally surround the exon 6 encoded sequence results in rapid mDdi2^{PD} protein degradation by one of the response mechanisms to misfolding stress right after translation, therefore the low signal for this protein variant in Figure 18C.

When comparing the lethality stages of the two mouse model strains, we hypothesize that one of the explanations could be partial complementation of the function of mDdi2 by its homolog mDdi1. It has been previously reported in two different studies in human cell lines that hDDI1 might act complementary to its homolog hDDI2. Kottemann and colleagues identified both hDDI1 and hDDI2 as shuttling proteins of UPS responsible for stalled fork restart via removal of RTF2 and another study showed impairment of NRF1 activation in hDDI1 knockout cells (Kottemann M. C.*, et al.*, 2018, Xiang Y.*, et al.*, 2018). While *Ddi2*$^{-/-}$ mice could benefit from the complementation of the activity by mDdi1, the existence of mDdi2^{PD} protein in the *Ddi2*$^{exon6\ -/-}$ individuals could abolish the shuttling factor function of mDdi1 by occupying the interaction site (such as the one on proteasome receptors), which could provide explanation for the earlier lethality observed in this model. The interaction of mDdi2^{PD} protein with proteasome could be facilitated with its extended N-terminal and C-terminal domains, in spite of formation of aggregates by the misfolded core region of the protein. In fact, preventing of interaction of any molecules with proteasomal receptors would most probably lead to disruption of homeostasis, cellular apoptosis and result in organism death, in a dominant negative effect.

In addition to the above-mentioned full knockout and liver-specific knockout mouse models, there is a number of options for studying the role of Ddi2 in a variety of tissues in adult mice, as was described for its substrates, for example neuronal tissue or osteoblasts (Kim J.*, et al.*, 2010, Kobayashi A.*, et al.*, 2011, Lee C. S.*, et al.*, 2011). Moreover, the discovery of Crispr/Cas9 system opened a pandora box with plentiful possibilities in studying the function of genes in diverse cell culture lines (Jinek M. *et al.*, 2012). Here, we focused on human *DDI2* and its up-to-date identified roles. It is clearly one of mammalian essential genes, involved in DNA repair response and most presumably in a variety of important cellular processes via its substrates Nrf1 and Nrf3. Indeed, the linkage of hDDI2 protein to

other pathways besides the recently described response to proteotoxic stress shall be further studied.

The members of Ddi1-like protein family are proteins with remarkable evolutionary conservation and based on findings of our own and others, a presumably broad spectrum of functions. Even though a few biological functions have been described for the hereby-characterized members of the Ddi1-like protein family, their full potential and importance in the context of homeostasis maintenance of individual cells as well as on the level of the whole organism remains to be elucidated.

6 CONCLUSIONS

1. Ddi1-like proteins relevant for NMR studies were cloned, recombinantly expressed and purified in sufficient yields and purity for their further biophysical characterization.

2. Solution structure of the UBL domain of Ddi1p from *S. cerevisiae* was acquired using NMR spectroscopy and adopts a conserved ubiquitin fold.

3. Based on NMR titration studies, we characterized the binding properties of the UBL domain and the UIM of hDDI2, which differ from those of domains of yDdi1p. Both the UBL domain and UIM of human DDI2 specifically bind ubiquitin, however with very weak affinity as opposed to the yeast ortholog. The hDDI2 UBL does not bind the C-terminal UIM region, hence the protein does not adopt a head-to-tail conformation in the homodimer.

4. The *in situ* hybridization studies revealed specific expression of *Ddi1* in mouse developing brain, which together with other findings lead to hypothesis of possible relevance of *DDI1* to neurodevelopmental diseases.

5. Two mouse model strains were produced in order to study the biological role of Ddi2. The full knockout model was generated by ESC manipulation and insertion of a lacZ reporter gene instead of the critical exon of *Ddi2*. The second model with ablated function of the protease domain of Ddi2 was prepared by TALEN-mediated excision of exon 6.

6. Both mouse strains exhibit embryonal lethality of the homozygous mice in mid-late gestation period, while adult heterozygous animals did not reveal any obvious abnormality in phenotyping studies.

7. Defect in the RVP domain of Ddi2 results in dominant negative effect and homozygous embryos exhibit earlier lethality (by E12.5) then the full knockout embryos (by E14.5).

8. The protease defective Ddi2 variant ($mDdi2^{PD}$) was reverse transcribed from isolated embryonal mRNA and cloned into bacterial expression vectors. Characterization of both recombinantly expressed $mDdi2^{WT}$ and $mDdi2^{PD}$ proteins revealed aggregate formation in the case of the $mDdi2^{PD}$.

9. Expression profiling of *Ddi2* that was performed using qPCR and lacZ staining, revealed difference among critical embryonal stage relevant for our model strains and localization of the expression in ectodermal and mesodermal tissues.

10. The morphologic phenotype was described for both strains as well as their inability in activation of Nrf1 transcription factor.

11. Application of proteotoxic stress onto primary MEF cultures isolated from embryos of our model strains reveals evidence of the ablation of the Nrf1 activation pathway, which links this failure to the phenotype of the retarded embryos.

7 LIST OF PUBLICATIONS

Publications related to dissertation project

a) Research articles

Siva, M., Svoboda, M., Veverka, V., Trempe, J. F., Hofmann, K., Kozisek, M., Hexnerova, R., Sedlak, F., Belza, J., Brynda, J., Sacha, P., Hubalek, M., Starkova, J., Flaisigova, I., Konvalinka, J. and Saskova, K. G. (2016). "Human DNA-Damage-Inducible 2 Protein Is Structurally and Functionally Distinct from Its Yeast Ortholog". *Sci Rep* **6** 30443. **IF (2016) = 4.259**

Trempe, J. F., Saskova, K. G., Siva, M., Ratcliffe, C. D., Veverka, V., Hoegl, A., Menade, M., Feng, X., Shenker, S., Svoboda, M., Kozisek, M., Konvalinka, J. and Gehring, K. (2016). "Structural studies of the yeast DNA damage-inducible protein Ddi1 reveal domain architecture of this eukaryotic protein family". *Sci Rep* **6** 33671. **IF (2016) = 4.259**

Ramirez, J., Lectez, B., Osinalde, N., Siva, M., Elu, N., Aloria, K., Prochazkova, M., Perez, C., Martinez-Hernandez, J., Barrio, R., Saskova, K. G., Arizmendi, J. M. and Mayor, U. (2018). "Quantitative proteomics reveals neuronal ubiquitination of Rngo/Ddi1 and several proteasomal subunits by Ube3a, accounting for the complexity of Angelman syndrome". *Hum Mol Genet* **27** (11): 1955-1971. **IF (2017)= 4.902**

b) Meeting abstracts

Siva, M., Chawengsaksophak, K., Svoboda, M., Kasparek, P., Prochazka, J., Sedlacek, R., Konvalinka, J. and Saskova, K. G. (2016). "Specific domain knockout of murine DNA damage-inducible protein homolog 2 (Ddi2)". *Transgenic Res* **25** (2): 262-263. **IF (2016) = 2.993**

8 REFERENCES

Abbott, A. (2015). "Neurophysiology: The man who bared the brain". *Nature* **521** 160.

Adams, J. (2003). "The proteasome: structure, function, and role in the cell". *Cancer Treat Rev* **29 Suppl 1** 3-9.

Adams, J., Palombella, V. J., Sausville, E. A., Johnson, J., Destree, A., Lazarus, D. D., Maas, J., Pien, C. S., Prakash, S. and Elliott, P. J. (1999). "Proteasome inhibitors: a novel class of potent and effective antitumor agents". *Cancer Res* **59** (11): 2615-2622.

Alexander, J., Kalev, O., Mehrabian, S., Traykov, L., Raycheva, M., Kanakis, D., Drineas, P., Lutz, M. I., Strobel, T., Penz, T., Schuster, M., Bock, C., Ferrer, I., Paschou, P. and Kovacs, G. G. (2016). "Familial early-onset dementia with complex neuropathologic phenotype and genomic background". *Neurobiol Aging* **42** 199-204.

Andrews, N. C., Erdjument-Bromage, H., Davidson, M. B., Tempst, P. and Orkin, S. H. (1993). "Erythroid transcription factor NF-E2 is a haematopoietic-specific basic-leucine zipper protein". *Nature* **362** (6422): 722-728.

Auesukaree, C., Fuchigami, I., Homma, T., Kaneko, Y. and Harashima, S. (2008). "Ddi1p and Rad23p play a cooperative role as negative regulators in the PHO pathway in Saccharomyces cerevisiae". *Biochem Biophys Res Commun* **365** (4): 821-825.

Aufderheide, A., Beck, F., Stengel, F., Hartwig, M., Schweitzer, A., Pfeifer, G., Goldberg, A. L., Sakata, E., Baumeister, W. and Forster, F. (2015). "Structural characterization of the interaction of Ubp6 with the 26S proteasome". *Proc Natl Acad Sci U S A* **112** (28): 8626-8631.

Bailly, V., Lamb, J., Sung, P., Prakash, S. and Prakash, L. (1994). "Specific complex formation between yeast RAD6 and RAD18 proteins: a potential mechanism for targeting RAD6 ubiquitin-conjugating activity to DNA damage sites". *Genes Dev* **8** (7): 811-820.

Baird, L., Lleres, D., Swift, S. and Dinkova-Kostova, A. T. (2013). "Regulatory flexibility in the Nrf2-mediated stress response is conferred by conformational cycling of the Keap1-Nrf2 protein complex". *Proc Natl Acad Sci U S A* **110** (38): 15259-15264.

Balakirev, M. Y., Mullally, J. E., Favier, A., Assard, N., Sulpice, E., Lindsey, D. F., Rulina, A. V., Gidrol, X. and Wilkinson, K. D. (2015). "Wss1 metalloprotease partners with Cdc48/Doa1 in processing genotoxic SUMO conjugates". *Elife* **4**

Balch, W. E., Morimoto, R. I., Dillin, A. and Kelly, J. W. (2008). "Adapting proteostasis for disease intervention". *Science* **319** (5865): 916-919.

Baldridge, R. D. and Rapoport, T. A. (2016). "Autoubiquitination of the Hrd1 Ligase Triggers Protein Retrotranslocation in ERAD". *Cell* **166** (2): 394-407.

Bard, J. A. M., Goodall, E. A., Greene, E. R., Jonsson, E., Dong, K. C. and Martin, A. (2018). "Structure and Function of the 26S Proteasome". *Annu Rev Biochem* **87** 697-724.

Barker, S., Weinfeld, M. and Murray, D. (2005). "DNA-protein crosslinks: their induction, repair, and biological consequences". *Mutat Res* **589** (2): 111-135.

Baumeister, W., Dahlmann, B., Hegerl, R., Kopp, F., Kuehn, L. and Pfeifer, G. (1988). "Electron microscopy and image analysis of the multicatalytic proteinase". *FEBS Lett* **241** (1-2): 239-245.

Beal, R., Deveraux, Q., Xia, G., Rechsteiner, M. and Pickart, C. (1996). "Surface hydrophobic residues of multiubiquitin chains essential for proteolytic targeting". *Proc Natl Acad Sci U S A* **93** (2): 861-866.

Berg, D. T., Gupta, A., Richardson, M. A., O'Brien, L. A., Calnek, D. and Grinnell, B. W. (2007). "Negative regulation of inducible nitric-oxide synthase expression mediated through transforming growth factor-beta-dependent modulation of transcription factor TCF11". *J Biol Chem* **282** (51): 36837-36844.

Bertolaet, B. L., Clarke, D. J., Wolff, M., Watson, M. H., Henze, M., Divita, G. and Reed, S. I. (2001). "UBA domains of DNA damage-inducible proteins interact with ubiquitin". *Nat Struct Biol* **8** (5): 417-422.

Bertolotti, A., Zhang, Y., Hendershot, L. M., Harding, H. P. and Ron, D. (2000). "Dynamic interaction of BiP and ER stress transducers in the unfolded-protein response". *Nat Cell Biol* **2** (6): 326-332.

Besche, H. C., Sha, Z., Kukushkin, N. V., Peth, A., Hock, E. M., Kim, W., Gygi, S., Gutierrez, J. A., Liao, H., Dick, L. and Goldberg, A. L. (2014). "Autoubiquitination of the 26S proteasome on Rpn13 regulates breakdown of ubiquitin conjugates". *EMBO J* **33** (10): 1159-1176.

Bhamidipati, A., Denic, V., Quan, E. M. and Weissman, J. S. (2005). "Exploration of the topological requirements of ERAD identifies Yos9p as a lectin sensor of misfolded glycoproteins in the ER lumen". *Mol Cell* **19** (6): 741-751.

Biggins, S., Ivanovska, I. and Rose, M. D. (1996). "Yeast ubiquitin-like genes are involved in duplication of the microtubule organizing center". *J Cell Biol* **133** (6): 1331-1346.

Biswas, M. and Chan, J. Y. (2010). "Role of Nrf1 in antioxidant response element-mediated gene expression and beyond". *Toxicol Appl Pharmacol* **244** (1): 16-20.

Biswas, M., Phan, D., Watanabe, M. and Chan, J. Y. (2011). "The Fbw7 tumor suppressor regulates nuclear factor E2-related factor 1 transcription factor turnover through proteasome-mediated proteolysis". *J Biol Chem* **286** (45): 39282-39289.

Biswas, M., Kwong, E. K., Park, E., Nagra, P. and Chan, J. Y. (2013). "Glycogen synthase kinase 3 regulates expression of nuclear factor-erythroid-2 related transcription factor-1 (Nrf1) and inhibits pro-survival function of Nrf1". *Exp Cell Res* **319** (13): 1922-1931.

Blom, D., Hirsch, C., Stern, P., Tortorella, D. and Ploegh, H. L. (2004). "A glycosylated type I membrane protein becomes cytosolic when peptide: N-glycanase is compromised". *EMBO J* **23** (3): 650-658.

Blythe, E. E., Olson, K. C., Chau, V. and Deshaies, R. J. (2017). "Ubiquitin- and ATP-dependent unfoldase activity of P97/VCP*NPLOC4*UFD1L is enhanced by a mutation that causes multisystem proteinopathy". *Proc Natl Acad Sci U S A* **114** (22): E4380-E4388.

Bodnar, N. O. and Rapoport, T. A. (2017). "Molecular Mechanism of Substrate Processing by the Cdc48 ATPase Complex". *Cell* **169** (4): 722-735 e729.

Bohm, S., Lamberti, G., Fernandez-Saiz, V., Stapf, C. and Buchberger, A. (2011). "Cellular functions of Ufd2 and Ufd3 in proteasomal protein degradation depend on Cdc48 binding". *Mol Cell Biol* **31** (7): 1528-1539.

Boisson, B., Laplantine, E., Prando, C., Giliani, S., Israelsson, E., Xu, Z., Abhyankar, A., Israel, L., Trevejo-Nunez, G., Bogunovic, D., Cepika, A. M., MacDuff, D., Chrabieh, M., Hubeau, M., Bajolle, F., Debre, M., Mazzolari, E., Vairo, D., Agou, F., Virgin, H. W., Bossuyt, X., Rambaud, C., Facchetti, F., Bonnet, D., Quartier, P., Fournet, J. C., Pascual, V., Chaussabel, D., Notarangelo, L. D., Puel, A., Israel, A., Casanova, J. L. and Picard, C. (2012). "Immunodeficiency, autoinflammation and amylopectinosis in humans with inherited HOIL-1 and LUBAC deficiency". *Nat Immunol* **13** (12): 1178-1186.

Bordallo, J., Plemper, R. K., Finger, A. and Wolf, D. H. (1998). "Der3p/Hrd1p is required for endoplasmic reticulum-associated degradation of misfolded lumenal and integral membrane proteins". *Mol Biol Cell* **9** (1): 209-222.

Boucher, D. M. and Pedersen, R. A. (1996). "Induction and differentiation of extra-embryonic mesoderm in the mouse". *Reprod Fertil Dev* **8** (4): 765-777.

Bouvier, L. A., Niemirowicz, G. T., Salas-Sarduy, E., Cazzulo, J. J. and Alvarez, V. E. (2018). "DNA-damage inducible protein 1 is a conserved metacaspase substrate that is cleaved and further destabilized in yeast under specific metabolic conditions". *FEBS J* **285** (6): 1097-1110.

Bradford, M. M. (1976). "A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding". *Anal Biochem* **72** 248-254.

Brannigan, J. A., Dodson, G., Duggleby, H. J., Moody, P. C., Smith, J. L., Tomchick, D. R. and Murzin, A. G. (1995). "A protein catalytic framework with an N-terminal nucleophile is capable of self-activation". *Nature* **378** (6555): 416-419.

Braun, S., Hanselmann, C., Gassmann, M. G., auf dem Keller, U., Born-Berclaz, C., Chan, K., Kan, Y. W. and Werner, S. (2002). "Nrf2 transcription factor, a novel target of keratinocyte growth factor action which regulates gene expression and inflammation in the healing skin wound". *Mol Cell Biol* **22** (15): 5492-5505.

Brehme, M., Voisine, C., Rolland, T., Wachi, S., Soper, J. H., Zhu, Y., Orton, K., Villella, A., Garza, D., Vidal, M., Ge, H. and Morimoto, R. I. (2014). "A chaperome subnetwork safeguards proteostasis in aging and neurodegenerative disease". *Cell Rep* **9** (3): 1135-1150.

Bremm, A. and Komander, D. (2011). "Emerging roles for Lys11-linked polyubiquitin in cellular regulation". *Trends Biochem Sci* **36** (7): 355-363.

Brett, K. E., Ferraro, Z. M., Yockell-Lelievre, J., Gruslin, A. and Adamo, K. B. (2014). "Maternal-fetal nutrient transport in pregnancy pathologies: the role of the placenta". *Int J Mol Sci* **15** (9): 16153-16185.

Buchberger, A. (2002). "From UBA to UBX: new words in the ubiquitin vocabulary". *Trends Cell Biol* **12** (5): 216-221.

Buchberger, A., Howard, M. J., Proctor, M. and Bycroft, M. (2001). "The UBX domain: a widespread ubiquitin-like module". *J Mol Biol* **307** (1): 17-24.

Buchman, T. G. (2002). "The community of the self". *Nature* **420** (6912): 246-251.

Buiting, K., Williams, C. and Horsthemke, B. (2016). "Angelman syndrome - insights into a rare neurogenetic disorder". *Nat Rev Neurol* **12** (10): 584-593.

Bulusu, V., Prior, N., Snaebjornsson, M. T., Kuehne, A., Sonnen, K. F., Kress, J., Stein, F., Schultz, C., Sauer, U. and Aulehla, A. (2017). "Spatiotemporal Analysis of a Glycolytic Activity Gradient Linked to Mouse Embryo Mesoderm Development". *Dev Cell* **40** (4): 331-341 e334.

Calamini, B. and Morimoto, R. I. (2012). "Protein homeostasis as a therapeutic target for diseases of protein conformation". *Curr Top Med Chem* **12** (22): 2623-2640.

Calfon, M., Zeng, H., Urano, F., Till, J. H., Hubbard, S. R., Harding, H. P., Clark, S. G. and Ron, D. (2002). "IRE1 couples endoplasmic reticulum load to secretory capacity by processing the XBP-1 mRNA". *Nature* **415** (6867): 92-96.

Camenisch, U., Dip, R., Schumacher, S. B., Schuler, B. and Naegeli, H. (2006). "Recognition of helical kinks by xeroderma pigmentosum group A protein triggers DNA excision repair". *Nat Struct Mol Biol* **13** (3): 278-284.

Cannon, W. B. (1926). "Physiological regulation of normal states: some tentative postulates concerning biological homeostatics". *Les Éditions Médicales* **Jubilee volume** 91-93.

Cannon, W. B. (1929). "Organization for physiological homeostasis". *Physiological Reviews* **9** (3): 33.

Carvalho, P., Stanley, A. M. and Rapoport, T. A. (2010). "Retrotranslocation of a misfolded luminal ER protein by the ubiquitin-ligase Hrd1p". *Cell* **143** (4): 579-591.

Caterina, J. J., Donze, D., Sun, C. W., Ciavatta, D. J. and Townes, T. M. (1994). "Cloning and functional characterization of LCR-F1: a bZIP transcription factor that activates erythroid-specific, human globin gene expression". *Nucleic Acids Res* **22** (12): 2383-2391.

Cermak, T., Doyle, E. L., Christian, M., Wang, L., Zhang, Y., Schmidt, C., Baller, J. A., Somia, N. V., Bogdanove, A. J. and Voytas, D. F. (2011). "Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting". *Nucleic Acids Res* **39** (12): e82.

Chan, J. Y., Han, X. L. and Kan, Y. W. (1993a). "Cloning of Nrf1, an NF-E2-related transcription factor, by genetic selection in yeast". *Proc Natl Acad Sci U S A* **90** (23): 11371-11375.

Chan, J. Y., Han, X. L. and Kan, Y. W. (1993b). "Isolation of cDNA encoding the human NF-E2 protein". *Proc Natl Acad Sci U S A* **90** (23): 11366-11370.

Chan, J. Y., Kwong, M., Lu, R., Chang, J., Wang, B., Yen, T. S. and Kan, Y. W. (1998). "Targeted disruption of the ubiquitous CNC-bZIP transcription factor, Nrf-1, results in anemia and embryonic lethality in mice". *EMBO J* **17** (6): 1779-1787.

Chanas, S. A., Jiang, Q., McMahon, M., McWalter, G. K., McLellan, L. I., Elcombe, C. R., Henderson, C. J., Wolf, C. R., Moffat, G. J., Itoh, K., Yamamoto, M. and Hayes, J. D. (2002). "Loss of the Nrf2 transcription factor causes a marked reduction in constitutive and inducible expression of the glutathione S-transferase Gsta1, Gsta2, Gstm1, Gstm2, Gstm3 and Gstm4 genes in the livers of male and female mice". *Biochem J* **365** (Pt 2): 405-416.

Chen, L., Shinde, U., Ortolan, T. G. and Madura, K. (2001). "Ubiquitin-associated (UBA) domains in Rad23 bind ubiquitin and promote inhibition of multi-ubiquitin chain assembly". *EMBO Rep* **2** (10): 933-938.

Chen, L., Kwong, M., Lu, R., Ginzinger, D., Lee, C., Leung, L. and Chan, J. Y. (2003). "Nrf1 is critical for redox balance and survival of liver cells during development". *Mol Cell Biol* **23** (13): 4673-4686.

Chen, P. and Hochstrasser, M. (1996). "Autocatalytic subunit processing couples active site formation in the 20S proteasome to completion of assembly". *Cell* **86** (6): 961-972.

Chen, Q., Olashaw, N. and Wu, J. (1995). "Participation of reactive oxygen species in the lysophosphatidic acid-stimulated mitogen-activated protein kinase kinase activation pathway". *J Biol Chem* **270** (48): 28499-28502.

Chen, X., Randles, L., Shi, K., Tarasov, S. G., Aihara, H. and Walters, K. J. (2016). "Structures of Rpn1 T1:Rad23 and hRpn13:hPLIC2 Reveal Distinct Binding Mechanisms between Substrate Receptors and Shuttle Factors of the Proteasome". *Structure* **24** (8): 1257-1270.

Chen, Z. J. and Sun, L. J. (2009). "Nonproteolytic functions of ubiquitin in cell signaling". *Mol Cell* **33** (3): 275-286.

Chenais, B., Derjuga, A., Massrieh, W., Red-Horse, K., Bellingard, V., Fisher, S. J. and Blank, V. (2005). "Functional and placental expression analysis of the human NRF3 transcription factor". *Mol Endocrinol* **19** (1): 125-137.

Chevillard, G. and Blank, V. (2011). "NFE2L3 (NRF3): the Cinderella of the Cap'n'Collar transcription factors". *Cell Mol Life Sci* **68** (20): 3337-3348.

Chevillard, G., Paquet, M. and Blank, V. (2011). "Nfe2l3 (Nrf3) deficiency predisposes mice to T-cell lymphoblastic lymphoma". *Blood* **117** (6): 2005-2008.

Chevillard, G., Nouhi, Z., Anna, D., Paquet, M. and Blank, V. (2010). "Nrf3-deficient mice are not protected against acute lung and adipose tissue damages induced by butylated hydroxytoluene". *FEBS Lett* **584** (5): 923-928.

Chojnacki, S., Cowley, A., Lee, J., Foix, A. and Lopez, R. (2017). "Programmatic access to bioinformatics tools from EMBL-EBI update: 2017". *Nucleic Acids Res* **45** (W1): W550-W553.

Chovatiya, R. and Medzhitov, R. (2014). "Stress, inflammation, and defense of homeostasis". *Mol Cell* **54** (2): 281-288.

Chowdhry, S., Zhang, Y., McMahon, M., Sutherland, C., Cuadrado, A. and Hayes, J. D. (2013). "Nrf2 is controlled by two distinct beta-TrCP recognition motifs in its Neh6 domain, one of which can be modulated by GSK-3 activity". *Oncogene* **32** (32): 3765-3781.

Chowdhury, A., Katoh, H., Hatanaka, A., Iwanari, H., Nakamura, N., Hamakubo, T., Natsume, T., Waku, T. and Kobayashi, A. (2017). "Multiple regulatory mechanisms of the biological function of NRF3 (NFE2L3) control cancer cell proliferation". *Sci Rep* **7** (1): 12494.

Chowdhury, I., Mo, Y., Gao, L., Kazi, A., Fisher, A. B. and Feinstein, S. I. (2009). "Oxidant stress stimulates expression of the human peroxiredoxin 6 gene by a transcriptional mechanism involving an antioxidant response element". *Free Radic Biol Med* **46** (2): 146-153.

Christensen, L. C., Jensen, N. W., Vala, A., Kamarauskaite, J., Johansson, L., Winther, J. R., Hofmann, K., Teilum, K. and Ellgaard, L. (2012). "The human selenoprotein VCP-interacting membrane protein (VIMP) is non-globular and harbors a reductase function in an intrinsically disordered region". *J Biol Chem* **287** (31): 26388-26399.

Christianson, J. C., Shaler, T. A., Tyler, R. E. and Kopito, R. R. (2008). "OS-9 and GRP94 deliver mutant alpha1-antitrypsin to the Hrd1-SEL1L ubiquitin ligase complex for ERAD". *Nat Cell Biol* **10** (3): 272-282.

Ciechanover, A., Hod, Y. and Hershko, A. (1978). "A heat-stable polypeptide component of an ATP-dependent proteolytic system from reticulocytes.". *Biochem Biophys Res Commun* **81** 1100-1105.

Ciechanover, A., Heller, H., Katz-Etzion, R. and Hershko, A. (1981). "Activation of the heat-stable polypeptide of the ATP-dependent proteolytic system". *Proc Natl Acad Sci U S A* **78** (2): 761-765.

Ciechanover, A., Elias, S., Heller, H. and Hershko, A. (1982). ""Covalent affinity" purification of ubiquitin-activating enzyme". *J Biol Chem* **257** (5): 2537-2542.

Ciechanover, A., Elias, S., Heller, H., Ferber, S. and Hershko, A. (1980). "Characterization of the heat-stable polypeptide of the ATP-dependent proteolytic system from reticulocytes". *J Biol Chem* **255** (16): 7525-7528.

Claessen, J. H., Kundrat, L. and Ploegh, H. L. (2012). "Protein quality control in the ER: balancing the ubiquitin checkbook". *Trends Cell Biol* **22** (1): 22-32.

Clarke, D. J., Mondesert, G., Segal, M., Bertolaet, B. L., Jensen, S., Wolff, M., Henze, M. and Reed, S. I. (2001). "Dosage suppressors of pds1 implicate ubiquitin-associated domains in checkpoint control". *Mol Cell Biol* **21** (6): 1997-2007.

Cohen, F. E. and Kelly, J. W. (2003). "Therapeutic approaches to protein-misfolding diseases". *Nature* **426** (6968): 905-909.

Cooper, S. J. (2008). "From Claude Bernard to Walter Cannon. Emergence of the concept of homeostasis". *Appetite* **51** (3): 419-427.

Cormier, J. H., Tamura, T., Sunryd, J. C. and Hebert, D. N. (2009). "EDEM1 recognition and delivery of misfolded proteins to the SEL1L-containing ERAD complex". *Mol Cell* **34** (5): 627-633.

Cornilescu, G., Marquardt, J. L., Ottinger, M. and Bax, A. (1998). "Validation of Protein Structure from Anisotropic Carbonyl Chemical Shifts in a Dilute Liquid Crystalline Phase". *Journal of the american chemical society* **120** (27): 2.

Costanzo, M., VanderSluis, B., Koch, E. N., Baryshnikova, A., Pons, C., Tan, G., Wang, W., Usaj, M., Hanchard, J., Lee, S. D., Pelechano, V., Styles, E. B., Billmann, M., van Leeuwen, J., van Dyk, N., Lin, Z. Y., Kuzmin, E., Nelson, J., Piotrowski, J. S., Srikumar, T., Bahr, S., Chen, Y., Deshpande, R., Kurat, C. F., Li, S. C., Li, Z., Usaj, M. M., Okada, H., Pascoe, N., San Luis, B. J., Sharifpoor, S., Shuteriqi, E., Simpkins, S. W., Snider, J., Suresh, H. G., Tan, Y., Zhu, H., Malod-Dognin, N., Janjic, V., Przulj, N., Troyanskaya, O. G., Stagljar, I., Xia, T., Ohya, Y., Gingras, A. C., Raught, B., Boutros, M., Steinmetz, L. M., Moore, C. L., Rosebrock, A. P., Caudy, A. A., Myers, C. L., Andrews, B. and Boone, C. (2016). "A global genetic interaction network maps a wiring diagram of cellular function". *Science* **353** (6306):

Cross, J. C., Simmons, D. G. and Watson, E. D. (2003). "Chorioallantoic morphogenesis and formation of the placental villous tree". *Ann N Y Acad Sci* **995** 84-93.

Cullinan, S. B., Gordan, J. D., Jin, J., Harper, J. W. and Diehl, J. A. (2004). "The Keap1-BTB protein is an adaptor that bridges Nrf2 to a Cul3-based E3 ligase: oxidative stress sensing by a Cul3-Keap1 ligase". *Mol Cell Biol* **24** (19): 8477-8486.

Dahlmann, B., Kuehn, L., Grziwa, A., Zwickl, P. and Baumeister, W. (1992). "Biochemical properties of the proteasome from Thermoplasma acidophilum". *Eur J Biochem* **208** (3): 789-797.

Dantuma, N. P., Heinen, C. and Hoogstraten, D. (2009). "The ubiquitin receptor Rad23: at the crossroads of nucleotide excision repair and proteasomal degradation". *DNA Repair (Amst)* **8** (4): 449-460.

Davies, J. M., Brunger, A. T. and Weis, W. I. (2008). "Improved structures of full-length p97, an AAA ATPase: implications for mechanisms of nucleotide-dependent conformational change". *Structure* **16** (5): 715-726.

de Laat, W. L., Appeldoorn, E., Sugasawa, K., Weterings, E., Jaspers, N. G. and Hoeijmakers, J. H. (1998). "DNA-binding polarity of human replication protein A positions nucleases in nucleotide excision repair". *Genes Dev* **12** (16): 2598-2609.

DeLaBarre, B. and Brunger, A. T. (2003). "Complete structure of p97/valosin-containing protein reveals communication between nucleotide domains". *Nat Struct Biol* **10** (10): 856-863.

Deng, H. X., Chen, W., Hong, S. T., Boycott, K. M., Gorrie, G. H., Siddique, N., Yang, Y., Fecto, F., Shi, Y., Zhai, H., Jiang, H., Hirano, M., Rampersaud, E., Jansen, G. H., Donkervoort, S., Bigio, E. H., Brooks, B. R., Ajroud, K., Sufit, R. L., Haines, J. L., Mugnaini, E., Pericak-Vance, M. A. and Siddique, T. (2011). "Mutations in UBQLN2 cause dominant X-linked juvenile and adult-onset ALS and ALS/dementia". *Nature* **477** (7363): 211-215.

Derjuga, A., Gourley, T. S., Holm, T. M., Heng, H. H., Shivdasani, R. A., Ahmed, R., Andrews, N. C. and Blank, V. (2004). "Complexity of CNC transcription factors as revealed by gene targeting of the Nrf3 locus". *Mol Cell Biol* **24** (8): 3286-3294.

Deveraux, Q., Ustrell, V., Pickart, C. and Rechsteiner, M. (1994). "A 26 S protease subunit that binds ubiquitin conjugates". *J Biol Chem* **269** (10): 7059-7061.

Diaz-Martinez, L. A., Kang, Y., Walters, K. J. and Clarke, D. J. (2006). "Yeast UBL-UBA proteins have partially redundant functions in cell cycle control". *Cell Div* **1** 28.

Diaz-Villanueva, J. F., Diaz-Molina, R. and Garcia-Gonzalez, V. (2015). "Protein Folding and Mechanisms of Proteostasis". *Int J Mol Sci* **16** (8): 17193-17230.

Dick, L. R., Cruikshank, A. A., Grenier, L., Melandri, F. D., Nunes, S. L. and Stein, R. L. (1996). "Mechanistic studies on the inactivation of the proteasome by lactacystin: a central role for clasto-lactacystin beta-lactone". *J Biol Chem* **271** (13): 7273-7276.

Dickinson, M. E., Flenniken, A. M., Ji, X., Teboul, L., Wong, M. D., White, J. K., Meehan, T. F., Weninger, W. J., Westerberg, H., Adissu, H., Baker, C. N., Bower, L., Brown, J. M., Caddle, L. B., Chiani, F., Clary, D., Cleak, J., Daly, M. J., Denegre, J. M., Doe, B., Dolan, M. E., Edie, S. M., Fuchs, H., Gailus-Durner, V., Galli, A., Gambadoro, A., Gallegos, J., Guo, S., Horner, N. R., Hsu, C. W., Johnson, S. J., Kalaga, S., Keith, L. C., Lanoue, L., Lawson, T. N., Lek, M., Mark, M., Marschall, S., Mason, J., McElwee, M. L., Newbigging, S., Nutter, L. M., Peterson, K. A., Ramirez-Solis, R., Rowland, D. J., Ryder, E., Samocha, K. E., Seavitt, J. R., Selloum, M., Szoke-Kovacs, Z., Tamura, M., Trainor, A. G., Tudose, I., Wakana, S., Warren, J., Wendling, O., West, D. B., Wong, L., Yoshiki, A., International Mouse Phenotyping, C., Jackson, L., Infrastructure Nationale Phenomin, I. C. d. l. S., Charles River, L., Harwell, M. R. C., Toronto Centre for, P., Wellcome Trust Sanger, I., Center, R. B., MacArthur, D. G., Tocchini-Valentini, G. P., Gao, X., Flicek, P., Bradley, A., Skarnes, W. C., Justice, M. J., Parkinson, H. E., Moore, M., Wells, S., Braun, R. E., Svenson, K. L., de Angelis, M. H., Herault, Y., Mohun, T., Mallon, A. M., Henkelman, R. M., Brown, S. D., Adams, D. J., Lloyd, K. C., McKerlie, C., Beaudet, A. L., Bucan, M. and Murray, S. A. (2016). "High-throughput discovery of novel developmental phenotypes". *Nature* **537** (7621): 508-514.

Dieterlen-Lievre, F. (1978). "Yolk sac erythropoiesis". *Experientia* **34** (3): 284-289.

Dinkova-Kostova, A. T., Holtzclaw, W. D. and Kensler, T. W. (2005). "The role of Keap1 in cellular protective responses". *Chem Res Toxicol* **18** (12): 1779-1791.

Dixit, G., Baker, R., Sacks, C. M., Torres, M. P. and Dohlman, H. G. (2014). "Guanine nucleotide-binding protein (Galpha) endocytosis by a cascade of ubiquitin binding domain proteins is required for sustained morphogenesis and proper mating in yeast". *J Biol Chem* **289** (21): 15052-15063.

Dodson, M., de la Vega, M. R., Cholanians, A. B., Schmidlin, C. J., Chapman, E. and Zhang, D. D. (2019). "Modulating NRF2 in Disease: Timing Is Everything". *Annu Rev Pharmacol Toxicol* **59** 555-575.

Doreleijers, J. F., Sousa da Silva, A. W., Krieger, E., Nabuurs, S. B., Spronk, C. A., Stevens, T. J., Vranken, W. F., Vriend, G. and Vuister, G. W. (2012). "CING: an integrated residue-based structure validation program suite". *J Biomol NMR* **54** (3): 267-283.

Doss-Pepe, E. W., Stenroos, E. S., Johnson, W. G. and Madura, K. (2003). "Ataxin-3 interactions with rad23 and valosin-containing protein and its associations with ubiquitin chains and the proteasome are consistent with a role in ubiquitin-mediated proteolysis". *Mol Cell Biol* **23** (18): 6469-6483.

Doyle, E. L., Booher, N. J., Standage, D. S., Voytas, D. F., Brendel, V. P., Vandyk, J. K. and Bogdanove, A. J. (2012). "TAL Effector-Nucleotide Targeter (TALE-NT) 2.0: tools for TAL effector design and target prediction". *Nucleic Acids Res* **40** (Web Server issue): W117-122.

Dungrawala, H., Rose, K. L., Bhat, K. P., Mohni, K. N., Glick, G. G., Couch, F. B. and Cortez, D. (2015). "The Replication Checkpoint Prevents Two Types of Fork Collapse without Regulating Replisome Stability". *Mol Cell* **59** (6): 998-1010.

Dunwoodie, S. L. (2009). "The role of hypoxia in development of the Mammalian embryo". *Dev Cell* **17** (6): 755-773.

Durcan, T. M., Tang, M. Y., Perusse, J. R., Dashti, E. A., Aguileta, M. A., McLelland, G. L., Gros, P., Shaler, T. A., Faubert, D., Coulombe, B. and Fon, E. A. (2014). "USP8 regulates mitophagy by removing K6-linked ubiquitin conjugates from parkin". *EMBO J* **33** (21): 2473-2491.

Durkacz, B. W., Omidiji, O., Gray, D. A. and Shall, S. (1980). "(ADP-ribose)n participates in DNA excision repair". *Nature* **283** (5747): 593-596.

Elsasser, S., Gali, R. R., Schwickart, M., Larsen, C. N., Leggett, D. S., Muller, B., Feng, M. T., Tubing, F., Dittmar, G. A. and Finley, D. (2002). "Proteasome subunit Rpn1 binds ubiquitin-like protein domains". *Nat Cell Biol* **4** (9): 725-730.

Erales, J., Hoyt, M. A., Troll, F. and Coffino, P. (2012). "Functional asymmetries of proteasome translocase pore". *J Biol Chem* **287** (22): 18535-18543.

Etlinger, J. D. and Goldberg, A. L. (1977). "A soluble ATP-dependent proteolytic system responsible for the degradation of abnormal proteins in reticulocytes". *Proc Natl Acad Sci U S A* **74** (1): 54-58.

Fagbemi, A. F., Orelli, B. and Scharer, O. D. (2011). "Regulation of endonuclease activity in human nucleotide excision repair". *DNA Repair (Amst)* **10** (7): 722-729.

Fenteany, G., Standaert, R. F., Lane, W. S., Choi, S., Corey, E. J. and Schreiber, S. L. (1995). "Inhibition of proteasome activities and subunit-specific amino-terminal threonine modification by lactacystin". *Science* **268** (5211): 726-731.

Field-Smith, A., Morgan, G. J. and Davies, F. E. (2006). "Bortezomib (Velcadetrade mark) in the Treatment of Multiple Myeloma". *Ther Clin Risk Manag* **2** (3): 271-279.

Fisher, R. D., Wang, B., Alam, S. L., Higginson, D. S., Robinson, H., Sundquist, W. I. and Hill, C. P. (2003). "Structure and ubiquitin binding of the ubiquitin-interacting motif". *J Biol Chem* **278** (31): 28976-28984.

Flemr, M., Malik, R., Franke, V., Nejepinska, J., Sedlacek, R., Vlahovicek, K. and Svoboda, P. (2013). "A retrotransposon-driven dicer isoform directs endogenous small interfering RNA production in mouse oocytes". *Cell* **155** (4): 807-816.

Frakes, A. E. and Dillin, A. (2017). "The UPR(ER): Sensor and Coordinator of Organismal Homeostasis". *Mol Cell* **66** (6): 761-771.

Franco, M., Seyfried, N. T., Brand, A. H., Peng, J. and Mayor, U. (2011). "A novel strategy to isolate ubiquitin conjugates reveals wide role for ubiquitination during neural development". *Mol Cell Proteomics* **10** (5): M110 002188.

Franz, A., Ackermann, L. and Hoppe, T. (2016). "Ring of Change: CDC48/p97 Drives Protein Dynamics at Chromatin". *Front Genet* **7** 73.

Friling, R. S., Bergelson, S. and Daniel, V. (1992). "Two adjacent AP-1-like binding sites form the electrophile-responsive element of the murine glutathione S-transferase Ya subunit gene". *Proc Natl Acad Sci U S A* **89** (2): 668-672.

Fu, Y., Zhu, Y., Zhang, K., Yeung, M., Durocher, D. and Xiao, W. (2008). "Rad6-Rad18 mediates a eukaryotic SOS response by ubiquitinating the 9-1-1 checkpoint clamp". *Cell* **133** (4): 601-611.

Funakoshi, M., Sasaki, T., Nishimoto, T. and Kobayashi, H. (2002). "Budding yeast Dsk2p is a polyubiquitin-binding protein that can interact with the proteasome". *Proc Natl Acad Sci U S A* **99** (2): 745-750.

Gabriely, G., Kama, R., Gelin-Licht, R. and Gerst, J. E. (2008). "Different domains of the UBL-UBA ubiquitin receptor, Ddi1/Vsm1, are involved in its multiple cellular roles". *Mol Biol Cell* **19** (9): 3625-3637.

Galluzzi, L., Yamazaki, T. and Kroemer, G. (2018). "Linking cellular stress responses to systemic homeostasis". *Nat Rev Mol Cell Biol* **19** (11): 731-745.

Gatti, M., Pinato, S., Maiolica, A., Rocchio, F., Prato, M. G., Aebersold, R. and Penengo, L. (2015). "RNF168 promotes noncanonical K27 ubiquitination to signal DNA damage". *Cell Rep* **10** (2): 226-238.

Geer, L. Y., Marchler-Bauer, A., Geer, R. C., Han, L., He, J., He, S., Liu, C., Shi, W. and Bryant, S. H. (2010). "The NCBI BioSystems database". *Nucleic Acids Res* **38** (Database issue): D492-496.

Genschel, J., Bazemore, L. R. and Modrich, P. (2002). "Human exonuclease I is required for 5' and 3' mismatch repair". *J Biol Chem* **277** (15): 13302-13311.

Giovannetti, A., Pierdominici, M., Di Iorio, A., Cianci, R., Murdaca, G., Puppo, F., Pandolfi, F. and Paganelli, R. (2008). "Apoptosis in the homeostasis of the immune system and in human immune mediated diseases". *Curr Pharm Des* **14** (3): 253-268.

Glickman, M. H., Rubin, D. M., Fried, V. A. and Finley, D. (1998a). "The regulatory particle of the Saccharomyces cerevisiae proteasome". *Mol Cell Biol* **18** (6): 3149-3162.

Glickman, M. H., Rubin, D. M., Coux, O., Wefes, I., Pfeifer, G., Cjeka, Z., Baumeister, W., Fried, V. A. and Finley, D. (1998b). "A subcomplex of the proteasome regulatory particle required for ubiquitin-conjugate degradation and related to the COP9-signalosome and eIF3". *Cell* **94** (5): 615-623.

Goldknopf, I. L., French, M. F., Musso, R. and Busch, H. (1977). "Presence of protein A24 in rat liver nucleosomes". *Proc Natl Acad Sci U S A* **74** (12): 5492-5495.

Goldstein, D. S. (1995). "Stress as a scientific idea: a homeostatic theory of stress and distress". *Homeostasis* **4** 177–215.

Goldstein, D. S. and Kopin, I. J. (2007). "Evolution of concepts of stress". *Stress* **10** (2): 109-120.

Gomez, T. A., Kolawa, N., Gee, M., Sweredoski, M. J. and Deshaies, R. J. (2011). "Identification of a functional docking site in the Rpn1 LRR domain for the UBA-UBL domain protein Ddi1". *BMC Biol* **9** 33.

Greenblatt, E. J., Olzmann, J. A. and Kopito, R. R. (2011). "Derlin-1 is a rhomboid pseudoprotease required for the dislocation of mutant alpha-1 antitrypsin from the endoplasmic reticulum". *Nat Struct Mol Biol* **18** (10): 1147-1152.

Groll, M., Ditzel, L., Lowe, J., Stock, D., Bochtler, M., Bartunik, H. D. and Huber, R. (1997). "Structure of 20S proteasome from yeast at 2.4 A resolution". *Nature* **386** (6624): 463-471.

Groll, M., Heinemeyer, W., Jager, S., Ullrich, T., Bochtler, M., Wolf, D. H. and Huber, R. (1999). "The catalytic sites of 20S proteasomes and their role in subunit maturation: a mutational and crystallographic study". *Proc Natl Acad Sci U S A* **96** (20): 10976-10983.

Groll, M., Bajorek, M., Kohler, A., Moroder, L., Rubin, D. M., Huber, R., Glickman, M. H. and Finley, D. (2000). "A gated channel into the proteasome core particle". *Nat Struct Biol* **7** (11): 1062-1067.

Gross, C. G. (2016). "Claude Bernard and the Constancy of the Internal Environment". *The Neuroscientist* **4** (5): 380-385.

Guntert, P. and Wuthrich, K. (1991). "Improved efficiency of protein structure calculations from NMR data using the program DIANA with redundant dihedral angle constraints". *J Biomol NMR* **1** (4): 447-456.

Guntert, P., Mumenthaler, C. and Wuthrich, K. (1997). "Torsion angle dynamics for NMR structure calculation with the new program DYANA". *J Mol Biol* **273** (1): 283-298.

Guthmueller, K. L., Yoder, M. L. and Holgado, A. M. (2011). "Determining genetic expression profiles in C. elegans using microarray and real-time PCR". *J Vis Exp* (53):

Guzder, S. N., Sung, P., Prakash, L. and Prakash, S. (1998). "Affinity of yeast nucleotide excision repair factor 2, consisting of the Rad4 and Rad23 proteins, for ultraviolet damaged DNA". *J Biol Chem* **273** (47): 31541-31546.

Haas, A. L. and Rose, I. A. (1982). "The mechanism of ubiquitin activating enzyme. A kinetic and equilibrium analysis". *J Biol Chem* **257** (17): 10329-10337.

Haas, A. L., Warms, J. V. and Rose, I. A. (1983). "Ubiquitin adenylate: structure and role in ubiquitin activation". *Biochemistry* **22** (19): 4388-4394.

Haas, A. L., Warms, J. V., Hershko, A. and Rose, I. A. (1982). "Ubiquitin-activating enzyme. Mechanism and role in protein-ubiquitin conjugation". *J Biol Chem* **257** (5): 2543-2548.

Hacker, G. (2018). "Apoptosis in infection". *Microbes Infect* **20** (9-10): 552-559.

Haince, J. F., McDonald, D., Rodrigue, A., Dery, U., Masson, J. Y., Hendzel, M. J. and Poirier, G. G. (2008). "PARP1-dependent kinetics of recruitment of MRE11 and NBS1 proteins to multiple DNA damage sites". *J Biol Chem* **283** (2): 1197-1208.

Hamazaki, J., Iemura, S., Natsume, T., Yashiroda, H., Tanaka, K. and Murata, S. (2006). "A novel proteasome interacting protein recruits the deubiquitinating enzyme UCH37 to 26S proteasomes". *EMBO J* **25** (19): 4524-4536.

Han, W., Ming, M., Zhao, R., Pi, J., Wu, C. and He, Y. Y. (2012). "Nrf1 CNC-bZIP protein promotes cell survival and nucleotide excision repair through maintaining glutathione homeostasis". *J Biol Chem* **287** (22): 18788-18795.

Hanada, M., Sugawara, K., Kaneta, K., Toda, S., Nishiyama, Y., Tomita, K., Yamamoto, H., Konishi, M. and Oki, T. (1992). "Epoxomicin, a new antitumor agent of microbial origin". *J Antibiot (Tokyo)* **45** (11): 1746-1752.

Hanna, J., Hathaway, N. A., Tone, Y., Crosas, B., Elsasser, S., Kirkpatrick, D. S., Leggett, D. S., Gygi, S. P., King, R. W. and Finley, D. (2006). "Deubiquitinating enzyme Ubp6 functions noncatalytically to delay proteasomal degradation". *Cell* **127** (1): 99-111.

Harding, H. P., Zhang, Y. and Ron, D. (1999). "Protein translation and folding are coupled by an endoplasmic-reticulum-resident kinase". *Nature* **397** (6716): 271-274.

Haze, K., Yoshida, H., Yanagi, H., Yura, T. and Mori, K. (1999). "Mammalian transcription factor ATF6 is synthesized as a transmembrane protein and activated by proteolysis in response to endoplasmic reticulum stress". *Mol Biol Cell* **10** (11): 3787-3799.

He, J., Kulkarni, K., da Fonseca, P. C., Krutauz, D., Glickman, M. H., Barford, D. and Morris, E. P. (2012). "The structure of the 26S proteasome subunit Rpn2 reveals its PC repeat domain as a closed toroid of two concentric alpha-helical rings". *Structure* **20** (3): 513-521.

Heinemeyer, W., Fischer, M., Krimmer, T., Stachon, U. and Wolf, D. H. (1997). "The active sites of the eukaryotic 20 S proteasome and their involvement in subunit precursor processing". *J Biol Chem* **272** (40): 25200-25209.

Heo, J. M., Livnat-Levanon, N., Taylor, E. B., Jones, K. T., Dephoure, N., Ring, J., Xie, J., Brodsky, J. L., Madeo, F., Gygi, S. P., Ashrafi, K., Glickman, M. H. and Rutter, J. (2010). "A stress-responsive system for mitochondrial protein degradation". *Mol Cell* **40** (3): 465-480.

Herrmann, T., Guntert, P. and Wuthrich, K. (2002). "Protein NMR structure determination with automated NOE assignment using the new software CANDID and the torsion angle dynamics algorithm DYANA". *J Mol Biol* **319** (1): 209-227.

Hetz, C. and Papa, F. R. (2018). "The Unfolded Protein Response and Cell Fate Control". *Mol Cell* **69** (2): 169-181.

Hirotsu, Y., Hataya, N., Katsuoka, F. and Yamamoto, M. (2012). "NF-E2-related factor 1 (Nrf1) serves as a novel regulator of hepatic lipid metabolism through regulation of the Lipin1 and PGC-1beta genes". *Mol Cell Biol* **32** (14): 2760-2770.

Hirotsu, Y., Higashi, C., Fukutomi, T., Katsuoka, F., Tsujita, T., Yagishita, Y., Matsuyama, Y., Motohashi, H., Uruno, A. and Yamamoto, M. (2014). "Transcription factor NF-E2-related factor 1 impairs glucose metabolism in mice". *Genes Cells* **19** (8): 650-665.

Hitt, R. and Wolf, D. H. (2004). "Der1p, a protein required for degradation of malfolded soluble proteins of the endoplasmic reticulum: topology and Der1-like proteins". *FEMS Yeast Res* **4** (7): 721-729.

Ho, D. V. and Chan, J. Y. (2015). "Induction of Herpud1 expression by ER stress is regulated by Nrf1". *FEBS Lett* **589** (5): 615-620.

Hoege, C., Pfander, B., Moldovan, G. L., Pyrowolakis, G. and Jentsch, S. (2002). "RAD6-dependent DNA repair is linked to modification of PCNA by ubiquitin and SUMO". *Nature* **419** (6903): 135-141.

Hofmann, K. and Bucher, P. (1996). "The UBA domain: a sequence motif present in multiple enzyme classes of the ubiquitination pathway". *Trends Biochem Sci* **21** (5): 172-173.

Hofmann, K. and Falquet, L. (2001). "A ubiquitin-interacting motif conserved in components of the proteasomal and lysosomal protein degradation systems". *Trends Biochem Sci* **26** (6): 347-350.

Holtzclaw, W. D., Dinkova-Kostova, A. T. and Talalay, P. (2004). "Protection against electrophile and oxidative stress by induction of phase 2 genes: the quest for the elusive sensor that responds to inducers". *Adv Enzyme Regul* **44** 335-367.

Houssaint, E. (1981). "Differentiation of the mouse hepatic primordium. II. Extrinsic origin of the haemopoietic cell line". *Cell Differ* **10** (5): 243-252.

Hu, M., Li, P., Li, M., Li, W., Yao, T., Wu, J. W., Gu, W., Cohen, R. E. and Shi, Y. (2002). "Crystal structure of a UBP-family deubiquitinating enzyme in isolation and in complex with ubiquitin aldehyde". *Cell* **111** (7): 1041-1054.

Huang, H. C., Nguyen, T. and Pickett, C. B. (2002). "Phosphorylation of Nrf2 at Ser-40 by protein kinase C regulates antioxidant response element-mediated transcription". *J Biol Chem* **277** (45): 42769-42774.

Huang, L., Kinnucan, E., Wang, G., Beaudenon, S., Howley, P. M., Huibregtse, J. M. and Pavletich, N. P. (1999). "Structure of an E6AP-UbcH7 complex: insights into ubiquitination by the E2-E3 enzyme cascade". *Science* **286** (5443): 1321-1326.

Huang, X., Luan, B., Wu, J. and Shi, Y. (2016). "An atomic structure of the human 26S proteasome". *Nat Struct Mol Biol* **23** (9): 778-785.

Husnjak, K., Elsasser, S., Zhang, N., Chen, X., Randles, L., Shi, Y., Hofmann, K., Walters, K. J., Finley, D. and Dikic, I. (2008). "Proteasome subunit Rpn13 is a novel ubiquitin receptor". *Nature* **453** (7194): 481-488.

Inagawa, T., Yamada-Inagawa, T., Eydmann, T., Mian, I. S., Wang, T. S. and Dalgaard, J. Z. (2009). "Schizosaccharomyces pombe Rtf2 mediates site-specific replication termination by inhibiting replication restart". *Proc Natl Acad Sci U S A* **106** (19): 7927-7932.

Inobe, T. and Genmei, R. (2015). "N-Terminal Coiled-Coil Structure of ATPase Subunits of 26S Proteasome Is Crucial for Proteasome Function". *PLoS One* **10** (7): e0134056.

Itoh, K., Wakabayashi, N., Katoh, Y., Ishii, T., O'Connor, T. and Yamamoto, M. (2003). "Keap1 regulates both cytoplasmic-nuclear shuttling and degradation of Nrf2 in response to electrophiles". *Genes Cells* **8** (4): 379-391.

Itoh, K., Wakabayashi, N., Katoh, Y., Ishii, T., Igarashi, K., Engel, J. D. and Yamamoto, M. (1999). "Keap1 represses nuclear activation of antioxidant responsive elements by Nrf2 through binding to the amino-terminal Neh2 domain". *Genes Dev* **13** (1): 76-86.

Itoh, K., Chiba, T., Takahashi, S., Ishii, T., Igarashi, K., Katoh, Y., Oyake, T., Hayashi, N., Satoh, K., Hatayama, I., Yamamoto, M. and Nabeshima, Y. (1997). "An Nrf2/small Maf heterodimer mediates the induction of phase II detoxifying enzyme genes through antioxidant response elements". *Biochem Biophys Res Commun* **236** (2): 313-322.

Ivantsiv, Y., Kaplun, L., Tzirkin-Goldin, R., Shabek, N. and Raveh, D. (2006). "Unique role for the UbL-UbA protein Ddi1 in turnover of SCFUfo1 complexes". *Mol Cell Biol* **26** (5): 1579-1588.

Jackson, W. A. (2001). "A short guide to humoral medicine". *Trends Pharmacol Sci* **22** (9): 487-489.

Jacob, A., Zhang, Y. and George, A. (2014). "Transcriptional regulation of dentin matrix protein 1 (DMP1) in odontoblasts and osteoblasts". *Connect Tissue Res* **55 Suppl 1** 107-112.

Jacobson, M. D. (1996). "Reactive oxygen species and programmed cell death". *Trends Biochem Sci* **21** (3): 83-86.

Jansen, L. E., Verhage, R. A. and Brouwer, J. (1998). "Preferential binding of yeast Rad4.Rad23 complex to damaged DNA". *J Biol Chem* **273** (50): 33111-33114.

Jin, J., Li, X., Gygi, S. P. and Harper, J. W. (2007). "Dual E1 activation systems for ubiquitin differentially regulate E2 enzyme charging". *Nature* **447** (7148): 1135-1138.

Jin, J., Xie, X., Xiao, Y., Hu, H., Zou, Q., Cheng, X. and Sun, S. C. (2016). "Epigenetic regulation of the expression of Il12 and Il23 and autoimmune inflammation by the deubiquitinase Trabid". *Nat Immunol* **17** (3): 259-268.

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A. and Charpentier, E. (2012). "A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity". *Science* **337** (6096): 816-821.

Johnsen, O., Murphy, P., Prydz, H. and Kolsto, A. B. (1998). "Interaction of the CNC-bZIP factor TCF11/LCR-F1/Nrf1 with MafG: binding-site selection and regulation of transcription". *Nucleic Acids Res* **26** (2): 512-520.

Johnsen, O., Skammelsrud, N., Luna, L., Nishizawa, M., Prydz, H. and Kolsto, A. B. (1996). "Small Maf proteins interact with the human transcription factor TCF11/Nrf1/LCR-F1". *Nucleic Acids Res* **24** (21): 4289-4297.

Joo, M. S., Kim, W. D., Lee, K. Y., Kim, J. H., Koo, J. H. and Kim, S. G. (2016). "AMPK Facilitates Nuclear Accumulation of Nrf2 by Phosphorylating at Serine 550". *Mol Cell Biol* **36** (14): 1931-1942.

Kajava, A. V. (2002). "What curves alpha-solenoids? Evidence for an alpha-helical toroid structure of Rpn1 and Rpn2 proteins of the 26 S proteasome". *J Biol Chem* **277** (51): 49791-49798.

Kanayama, A., Seth, R. B., Sun, L., Ea, C. K., Hong, M., Shaito, A., Chiu, Y. H., Deng, L. and Chen, Z. J. (2004). "TAB2 and TAB3 activate the NF-kappaB pathway through binding to polyubiquitin chains". *Mol Cell* **15** (4): 535-548.

Kaplun, L., Ivantsiv, Y., Kornitzer, D. and Raveh, D. (2000). "Functions of the DNA damage response pathway target Ho endonuclease of yeast for degradation via the ubiquitin-26S proteasome system". *Proc Natl Acad Sci U S A* **97** (18): 10077-10082.

Kaplun, L., Ivantsiv, Y., Bakhrat, A. and Raveh, D. (2003). "DNA damage response-mediated degradation of Ho endonuclease via the ubiquitin system involves its nuclear export". *J Biol Chem* **278** (49): 48727-48734.

Kaplun, L., Tzirkin, R., Bakhrat, A., Shabek, N., Ivantsiv, Y. and Raveh, D. (2005). "The DNA damage-inducible UbL-UbA protein Ddi1 participates in Mec1-mediated degradation of Ho endonuclease". *Mol Cell Biol* **25** (13): 5355-5362.

Kaspar, J. W., Niture, S. K. and Jaiswal, A. K. (2009). "Nrf2:INrf2 (Keap1) signaling in oxidative stress". *Free Radic Biol Med* **47** (9): 1304-1309.

Kasparek, P., Krausova, M., Haneckova, R., Kriz, V., Zbodakova, O., Korinek, V. and Sedlacek, R. (2014). "Efficient gene targeting of the Rosa26 locus in mouse zygotes using TALE nucleases". *FEBS Lett* **588** (21): 3982-3988.

Kawai, Y., Garduno, L., Theodore, M., Yang, J. and Arinze, I. J. (2011). "Acetylation-deacetylation of the transcription factor Nrf2 (nuclear factor erythroid 2-related factor 2) regulates its transcriptional activity and nucleocytoplasmic localization". *J Biol Chem* **286** (9): 7629-7640.

Kaye, F. J., Modi, S., Ivanovska, I., Koonin, E. V., Thress, K., Kubo, A., Kornbluth, S. and Rose, M. D. (2000). "A family of ubiquitin-like proteins binds the ATPase domain of Hsp70-like Stch". *FEBS Lett* **467** (2-3): 348-355.

Kent, W. J., Sugnet, C. W., Furey, T. S., Roskin, K. M., Pringle, T. H., Zahler, A. M. and Haussler, D. (2002). "The human genome browser at UCSC". *Genome Res* **12** (6): 996-1006.

Kim, B., Ryu, K. S., Kim, H. J., Cho, S. J. and Choi, B. S. (2005). "Solution structure and backbone dynamics of the XPC-binding domain of the human DNA repair protein hHR23B". *FEBS J* **272** (10): 2467-2476.

Kim, H. M., Han, J. W. and Chan, J. Y. (2016). "Nuclear Factor Erythroid-2 Like 1 (NFE2L1): Structure, function and regulation". *Gene* **584** (1): 17-25.

Kim, J., Xing, W., Wergedal, J., Chan, J. Y. and Mohan, S. (2010). "Targeted disruption of nuclear factor erythroid-derived 2-like 1 in osteoblasts reduces bone size and bone formation in mice". *Physiol Genomics* **40** (2): 100-110.

Kim, W., Spear, E. D. and Ng, D. T. (2005). "Yos9p detects and targets misfolded glycoproteins for ER-associated degradation". *Mol Cell* **19** (6): 753-764.

Kim, Y. E., Hipp, M. S., Bracher, A., Hayer-Hartl, M. and Hartl, F. U. (2013). "Molecular chaperone functions in protein folding and proteostasis". *Annu Rev Biochem* **82** 323-355.

Kim, Y. J., Ahn, J. Y., Liang, P., Ip, C., Zhang, Y. and Park, Y. M. (2007). "Human prx1 gene is a target of Nrf2 and is up-regulated by hypoxia/reoxygenation: implication to tumor biology". *Cancer Res* **67** (2): 546-554.

Kirisako, T., Kamei, K., Murata, S., Kato, M., Fukumoto, H., Kanie, M., Sano, S., Tokunaga, F., Tanaka, K. and Iwai, K. (2006). "A ubiquitin ligase complex assembles linear polyubiquitin chains". *EMBO J* **25** (20): 4877-4887.

Kobayashi, A., Kang, M. I., Okawa, H., Ohtsuji, M., Zenke, Y., Chiba, T., Igarashi, K. and Yamamoto, M. (2004). "Oxidative stress sensor Keap1 functions as an adaptor for Cul3-based E3 ligase to regulate proteasomal degradation of Nrf2". *Mol Cell Biol* **24** (16): 7130-7139.

Kobayashi, A., Tsukide, T., Miyasaka, T., Morita, T., Mizoroki, T., Saito, Y., Ihara, Y., Takashima, A., Noguchi, N., Fukamizu, A., Hirotsu, Y., Ohtsuji, M., Katsuoka, F. and Yamamoto, M. (2011). "Central nervous system-specific deletion of transcription factor Nrf1 causes progressive motor neuronal dysfunction". *Genes Cells* **16** (6): 692-703.

Koizumi, S., Irie, T., Hirayama, S., Sakurai, Y., Yashiroda, H., Naguro, I., Ichijo, H., Hamazaki, J. and Murata, S. (2016). "The aspartyl protease DDI2 activates Nrf1 to compensate for proteasome dysfunction". *Elife* **5**

Kokame, K., Agarwala, K. L., Kato, H. and Miyata, T. (2000). "Herp, a new ubiquitin-like membrane protein induced by endoplasmic reticulum stress". *J Biol Chem* **275** (42): 32846-32853.

Komander, D. and Rape, M. (2012). "The ubiquitin code". *Annu Rev Biochem* **81** 203-229.

Koradi, R., Billeter, M. and Wuthrich, K. (1996). "MOLMOL: a program for display and analysis of macromolecular structures". *J Mol Graph* **14** (1): 51-55, 29-32.

Kottemann, M. C., Conti, B. A., Lach, F. P. and Smogorzewska, A. (2018). "Removal of RTF2 from Stalled Replisomes Promotes Maintenance of Genome Integrity". *Mol Cell* **69** (1): 24-35 e25.

Krokan, H. E. and Bjoras, M. (2013). "Base excision repair". *Cold Spring Harb Perspect Biol* **5** (4): a012583.

Krylov, D. M. and Koonin, E. V. (2001). "A novel family of predicted retroviral-like aspartyl proteases with a possible key role in eukaryotic cell cycle control". *Curr Biol* **11** (15): R584-587.

Kumar, S. and Suguna, K. (2018). "Crystal structure of the retroviral protease-like domain of a protozoal DNA damage-inducible 1 protein". *FEBS Open Bio* **8** (9): 1379-1394.

Kunkel, T. A. and Erie, D. A. (2015). "Eukaryotic Mismatch Repair in Relation to DNA Replication". *Annu Rev Genet* **49** 291-313.

Kwak, M. K., Wakabayashi, N., Greenlaw, J. L., Yamamoto, M. and Kensler, T. W. (2003). "Antioxidants enhance mammalian proteasome expression through the Keap1-Nrf2 signaling pathway". *Mol Cell Biol* **23** (23): 8786-8794.

Kwong, E. K., Kim, K. M., Penalosa, P. J. and Chan, J. Y. (2012). "Characterization of Nrf1b, a novel isoform of the nuclear factor-erythroid-2 related transcription factor-1 that activates antioxidant response element-regulated genes". *PLoS One* **7** (10): e48404.

Kwong, M., Kan, Y. W. and Chan, J. Y. (1999). "The CNC basic leucine zipper factor, Nrf1, is essential for cell survival in response to oxidative stress-inducing agents. Role for Nrf1 in gamma-gcs(l) and gss expression in mouse fibroblasts". *J Biol Chem* **274** (52): 37491-37498.

Lam, Y. A., Xu, W., DeMartino, G. N. and Cohen, R. E. (1997). "Editing of ubiquitin conjugates by an isopeptidase in the 26S proteasome". *Nature* **385** (6618): 737-740.

Lander, G. C., Estrin, E., Matyskiela, M. E., Bashore, C., Nogales, E. and Martin, A. (2012). "Complete subunit architecture of the proteasome regulatory particle". *Nature* **482** (7384): 186-191.

Larsen, N. B., Gao, A. O., Sparks, J. L., Gallina, I., Wu, R. A., Mann, M., Raschle, M., Walter, J. C. and Duxin, J. P. (2018). "Replication-Coupled DNA-Protein Crosslink Repair by SPRTN and the Proteasome in Xenopus Egg Extracts". *Mol Cell*

Lee, A. H., Iwakoshi, N. N. and Glimcher, L. H. (2003). "XBP-1 regulates a subset of endoplasmic reticulum resident chaperone genes in the unfolded protein response". *Mol Cell Biol* **23** (21): 7448-7459.

Lee, B. H., Lu, Y., Prado, M. A., Shi, Y., Tian, G., Sun, S., Elsasser, S., Gygi, S. P., King, R. W. and Finley, D. (2016). "USP14 deubiquitinates proteasome-bound substrates that are ubiquitinated at multiple sites". *Nature* **532** (7599): 398-401.

Lee, C. S., Ho, D. V. and Chan, J. Y. (2013). "Nuclear factor-erythroid 2-related factor 1 regulates expression of proteasome genes in hepatocytes and protects against endoplasmic reticulum stress and steatosis in mice". *FEBS J* **280** (15): 3609-3620.

Lee, C. S., Lee, C., Hu, T., Nguyen, J. M., Zhang, J., Martin, M. V., Vawter, M. P., Huang, E. J. and Chan, J. Y. (2011). "Loss of nuclear factor E2-related factor 1 in the brain leads to dysregulation of proteasome gene expression and neurodegeneration". *Proc Natl Acad Sci U S A* **108** (20): 8408-8413.

Lee, J. H., Choi, J. M., Lee, C., Yi, K. J. and Cho, Y. (2005). "Structure of a peptide:N-glycanase-Rad23 complex: insight into the deglycosylation for denatured glycoproteins". *Proc Natl Acad Sci U S A* **102** (26): 9144-9149.

Lee, J. W., Blanco, L., Zhou, T., Garcia-Diaz, M., Bebenek, K., Kunkel, T. A., Wang, Z. and Povirk, L. F. (2004). "Implication of DNA polymerase lambda in alignment-based gap filling for nonhomologous DNA end joining in human nuclear extracts". *J Biol Chem* **279** (1): 805-811.

Leggett, D. S., Hanna, J., Borodovsky, A., Crosas, B., Schmidt, M., Baker, R. T., Walz, T., Ploegh, H. and Finley, D. (2002). "Multiple associated proteins regulate proteasome structure and function". *Mol Cell* **10** (3): 495-507.

Lehrbach, N. J. and Ruvkun, G. (2016). "Proteasome dysfunction triggers activation of SKN-1A/Nrf1 by the aspartic protease DDI-1". *Elife* **5**

Leung, L., Kwong, M., Hou, S., Lee, C. and Chan, J. Y. (2003). "Deficiency of the Nrf1 and Nrf2 transcription factors results in early embryonic lethality and severe oxidative stress". *J Biol Chem* **278** (48): 48021-48029.

Li, G., Zhao, G., Schindelin, H. and Lennarz, W. J. (2008). "Tyrosine phosphorylation of ATPase p97 regulates its activity during ERAD". *Biochem Biophys Res Commun* **375** (2): 247-251.

Li, G., Zhao, G., Zhou, X., Schindelin, H. and Lennarz, W. J. (2006). "The AAA ATPase p97 links peptide N-glycanase to the endoplasmic reticulum-associated E3 ligase autocrine motility factor receptor". *Proc Natl Acad Sci U S A* **103** (22): 8348-8353.

Li, W., Yu, S., Liu, T., Kim, J. H., Blank, V., Li, H. and Kong, A. N. (2008a). "Heterodimerization with small Maf proteins enhances nuclear retention of Nrf2 via masking the NESzip motif". *Biochim Biophys Acta* **1783** (10): 1847-1856.

Li, W., Bengtson, M. H., Ulbrich, A., Matsuda, A., Reddy, V. A., Orth, A., Chanda, S. K., Batalov, S. and Joazeiro, C. A. (2008b). "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** (1): e1487.

Lieber, M. R. (2010). "The mechanism of double-strand DNA break repair by the nonhomologous DNA end-joining pathway". *Annu Rev Biochem* **79** 181-211.

Lilley, B. N. and Ploegh, H. L. (2004). "A membrane protein required for dislocation of misfolded proteins from the ER". *Nature* **429** (6994): 834-840.

Lin, J. I. (1983). "Rudolf Virchow: Creator of Cellular Pathology ". *Laboratory Medicine* **14** (12): 791–794.

Lindahl, T. (1974). "An N-glycosidase from Escherichia coli that releases free uracil from DNA containing deaminated cytosine residues". *Proc Natl Acad Sci U S A* **71** (9): 3649-3653.

Liu, Y. and Xiao, W. (1997). "Bidirectional regulation of two DNA-damage-inducible genes, MAG1 and DDI1, from Saccharomyces cerevisiae". *Mol Microbiol* **23** (4): 777-789.

Liu, Y., Dai, H. and Xiao, W. (1997). "UAS(MAG1), a yeast cis-acting element that regulates the expression of MAG1, is located within the protein coding region of DDI1". *Mol Gen Genet* **255** (5): 533-542.

Lopez-Mosqueda, J., Maddi, K., Prgomet, S., Kalayil, S., Marinovic-Terzic, I., Terzic, J. and Dikic, I. (2016). "SPRTN is a mammalian DNA-binding metalloprotease that resolves DNA-protein crosslinks". *Elife* **5**

Lowe, J., Stock, D., Jap, B., Zwickl, P., Baumeister, W. and Huber, R. (1995). "Crystal structure of the 20S proteasome from the archaeon T. acidophilum at 3.4 A resolution". *Science* **268** (5210): 533-539.

Lu, K., den Brave, F. and Jentsch, S. (2017). "Pathway choice between proteasomal and autophagic degradation". *Autophagy* **13** (10): 1799-1800.

Lu, S. C. (2009). "Regulation of glutathione synthesis". *Mol Aspects Med* **30** (1-2): 42-59.

Luna, L., Skammelsrud, N., Johnsen, O., Abel, K. J., Weber, B. L., Prydz, H. and Kolsto, A. B. (1995). "Structural organization and mapping of the human TCF11 gene". *Genomics* **27** (2): 237-244.

Lupas, A., Baumeister, W. and Hofmann, K. (1997). "A repetitive sequence in subunits of the 26S proteasome and 20S cyclosome (anaphase-promoting complex)". *Trends Biochem Sci* **22** (6): 195-196.

Lustgarten, V. and Gerst, J. E. (1999). "Yeast VSM1 encodes a v-SNARE binding protein that may act as a negative regulator of constitutive exocytosis". *Mol Cell Biol* **19** (6): 4480-4494.

Lux, C. T., Yoshimoto, M., McGrath, K., Conway, S. J., Palis, J. and Yoder, M. C. (2008). "All primitive and definitive hematopoietic progenitor cells emerging before E10 in the mouse embryo are products of the yolk sac". *Blood* **111** (7): 3435-3438.

Ma, Y., Pannicke, U., Schwarz, K. and Lieber, M. R. (2002). "Hairpin opening and overhang processing by an Artemis/DNA-dependent protein kinase complex in nonhomologous end joining and V(D)J recombination". *Cell* **108** (6): 781-794.

Maillard, R. A., Chistol, G., Sen, M., Righini, M., Tan, J., Kaiser, C. M., Hodges, C., Martin, A. and Bustamante, C. (2011). "ClpX(P) generates mechanical force to unfold and translocate its protein substrates". *Cell* **145** (3): 459-469.

Mansour, S. L., Thomas, K. R., Deng, C. X. and Capecchi, M. R. (1990). "Introduction of a lacZ reporter gene into the mouse int-2 locus by homologous recombination". *Proc Natl Acad Sci U S A* **87** (19): 7688-7692.

Mansour, W., Nakasone, M. A., von Delbruck, M., Yu, Z., Krutauz, D., Reis, N., Kleifeld, O., Sommer, T., Fushman, D. and Glickman, M. H. (2015). "Disassembly of Lys11 and mixed linkage polyubiquitin conjugates provides insights into function of proteasomal deubiquitinases Rpn11 and Ubp6". *J Biol Chem* **290** (8): 4688-4704.

Marash, M. and Gerst, J. E. (2003). "Phosphorylation of the autoinhibitory domain of the Sso t-SNAREs promotes binding of the Vsm1 SNARE regulator in yeast". *Mol Biol Cell* **14** (8): 3114-3125.

Marechal, A. and Zou, L. (2013). "DNA damage sensing by the ATM and ATR kinases". *Cold Spring Harb Perspect Biol* **5** (9):

Marteijn, J. A., Lans, H., Vermeulen, W. and Hoeijmakers, J. H. (2014). "Understanding nucleotide excision repair and its roles in cancer and ageing". *Nat Rev Mol Cell Biol* **15** (7): 465-481.

Martin, A., Baker, T. A. and Sauer, R. T. (2008). "Pore loops of the AAA+ ClpX machine grip substrates to drive translocation and unfolding". *Nat Struct Mol Biol* **15** (11): 1147-1151.

Martin, K. R. and Barrett, J. C. (2002). "Reactive oxygen species as double-edged swords in cellular processes: low-dose cell signaling versus high-dose toxicity". *Hum Exp Toxicol* **21** (2): 71-75.

Maskey, R. S., Kim, M. S., Baker, D. J., Childs, B., Malureanu, L. A., Jeganathan, K. B., Machida, Y., van Deursen, J. M. and Machida, Y. J. (2014). "Spartan deficiency causes genomic instability and progeroid phenotypes". *Nat Commun* **5** 5744.

Masters, P. M., Bada, J. L. and Zigler, J. S., Jr. (1977). "Aspartic acid racemisation in the human lens during ageing and in cataract formation". *Nature* **268** (5615): 71-73.

Masutani, C., Araki, M., Sugasawa, K., van der Spek, P. J., Yamada, A., Uchida, A., Maekawa, T., Bootsma, D., Hoeijmakers, J. H. and Hanaoka, F. (1997). "Identification and characterization of XPC-binding domain of hHR23B". *Mol Cell Biol* **17** (12): 6915-6923.

Masutani, C., Sugasawa, K., Yanagisawa, J., Sonoyama, T., Ui, M., Enomoto, T., Takio, K., Tanaka, K., van der Spek, P. J., Bootsma, D. and et al. (1994). "Purification and cloning of a nucleotide excision repair complex involving the xeroderma pigmentosum group C protein and a human homologue of yeast RAD23". *EMBO J* **13** (8): 1831-1843.

McCord, J. M. (1995). "Superoxide radical: controversies, contradictions, and paradoxes". *Proc Soc Exp Biol Med* **209** (2): 112-117.

McEwen, B. S. and Stellar, E. (1993). "Stress and the individual. Mechanisms leading to disease". *Arch Intern Med* **153** (18): 2093-2101.

Meng, L., Mohan, R., Kwok, B. H., Elofsson, M., Sin, N. and Crews, C. M. (1999). "Epoxomicin, a potent and selective proteasome inhibitor, exhibits in vivo antiinflammatory activity". *Proc Natl Acad Sci U S A* **96** (18): 10403-10408.

Mesquita, E. T., Souza Junior, C. V. and Ferreira, T. R. (2015). "Andreas Vesalius 500 years--A Renaissance that revolutionized cardiovascular knowledge". *Rev Bras Cir Cardiovasc* **30** (2): 260-265.

Meyer, H., Bug, M. and Bremer, S. (2012). "Emerging functions of the VCP/p97 AAA-ATPase in the ubiquitin system". *Nat Cell Biol* **14** (2): 117-123.

Meyer, H. H., Shorter, J. G., Seemann, J., Pappin, D. and Warren, G. (2000). "A complex of mammalian ufd1 and npl4 links the AAA-ATPase, p97, to ubiquitin and nuclear transport pathways". *EMBO J* **19** (10): 2181-2192.

Modell, H., Cliff, W., Michael, J., McFarland, J., Wenderoth, M. P. and Wright, A. (2015). "A physiologist's view of homeostasis". *Adv Physiol Educ* **39** (4): 259-266.

Modrich, P. (2006). "Mechanisms in eukaryotic mismatch repair". *J Biol Chem* **281** (41): 30305-30309.

Moi, P., Chan, K., Asunis, I., Cao, A. and Kan, Y. W. (1994). "Isolation of NF-E2-related factor 2 (Nrf2), a NF-E2-like basic leucine zipper transcriptional activator that binds to the tandem NF-E2/AP1 repeat of the beta-globin locus control region". *Proc Natl Acad Sci U S A* **91** (21): 9926-9930.

Molinari, M., Calanca, V., Galli, C., Lucca, P. and Paganetti, P. (2003). "Role of EDEM in the release of misfolded glycoproteins from the calnexin cycle". *Science* **299** (5611): 1397-1400.

Morawe, T., Honemann-Capito, M., von Stein, W. and Wodarz, A. (2011). "Loss of the extraproteasomal ubiquitin receptor Rings lost impairs ring canal growth in Drosophila oogenesis". *J Cell Biol* **193** (1): 71-80.

Morocz, M., Zsigmond, E., Toth, R., Enyedi, M. Z., Pinter, L. and Haracska, L. (2017). "DNA-dependent protease activity of human Spartan facilitates replication of DNA-protein crosslink-containing DNA". *Nucleic Acids Res* **45** (6): 3172-3188.

Morris, J. R. and Solomon, E. (2004). "BRCA1 : BARD1 induces the formation of conjugated ubiquitin structures, dependent on K6 of ubiquitin, in cells during DNA replication and repair". *Hum Mol Genet* **13** (8): 807-817.

Motohashi, H., Katsuoka, F., Engel, J. D. and Yamamoto, M. (2004). "Small Maf proteins serve as transcriptional cofactors for keratinocyte differentiation in the Keap1-Nrf2 regulatory pathway". *Proc Natl Acad Sci U S A* **101** (17): 6379-6384.

Mueller, T. D. and Feigon, J. (2002). "Solution structures of UBA domains reveal a conserved hydrophobic surface for protein-protein interactions". *J Mol Biol* **319** (5): 1243-1255.

Mueller, T. D., Kamionka, M. and Feigon, J. (2004). "Specificity of the interaction between ubiquitin-associated domains and ubiquitin". *J Biol Chem* **279** (12): 11926-11936.

Munro, H. N., Pilistine, S. J. and Fant, M. E. (1983). "The placenta in nutrition". *Annu Rev Nutr* **3** 97-124.

Myhrstad, M. C., Husberg, C., Murphy, P., Nordstrom, O., Blomhoff, R., Moskaug, J. O. and Kolsto, A. B. (2001). "TCF11/Nrf1 overexpression increases the intracellular glutathione level and can transactivate the gamma-glutamylcysteine synthetase (GCS) heavy subunit promoter". *Biochim Biophys Acta* **1517** (2): 212-219.

Narayanan, K., Ramachandran, A., Peterson, M. C., Hao, J., Kolsto, A. B., Friedman, A. D. and George, A. (2004). "The CCAAT enhancer-binding protein (C/EBP)beta and Nrf1 interact to regulate dentin sialophosphoprotein (DSPP) gene expression during odontoblast differentiation". *J Biol Chem* **279** (44): 45423-45432.

Nath, A. K., Enciso, J., Kuniyasu, M., Hao, X. Y., Madri, J. A. and Pinter, E. (2004). "Nitric oxide modulates murine yolk sac vasculogenesis and rescues glucose induced vasculopathy". *Development* **131** (10): 2485-2496.

Neuber, O., Jarosch, E., Volkwein, C., Walter, J. and Sommer, T. (2005). "Ubx2 links the Cdc48 complex to ER-associated protein degradation". *Nat Cell Biol* **7** (10): 993-998.

Ng, J. M., Vrieling, H., Sugasawa, K., Ooms, M. P., Grootegoed, J. A., Vreeburg, J. T., Visser, P., Beems, R. B., Gorgels, T. G., Hanaoka, F., Hoeijmakers, J. H. and van der Horst, G. T. (2002). "Developmental defects and male sterility in mice lacking the ubiquitin-like DNA repair gene mHR23B". *Mol Cell Biol* **22** (4): 1233-1245.

Nick McElhinny, S. A. and Ramsden, D. A. (2003). "Polymerase mu is a DNA-directed DNA/RNA polymerase". *Mol Cell Biol* **23** (7): 2309-2315.

Nick McElhinny, S. A., Snowden, C. M., McCarville, J. and Ramsden, D. A. (2000). "Ku recruits the XRCC4-ligase IV complex to DNA ends". *Mol Cell Biol* **20** (9): 2996-3003.

Nioi, P., McMahon, M., Itoh, K., Yamamoto, M. and Hayes, J. D. (2003). "Identification of a novel Nrf2-regulated antioxidant response element (ARE) in the mouse NAD(P)H:quinone oxidoreductase 1 gene: reassessment of the ARE consensus sequence". *Biochem J* **374** (Pt 2): 337-348.

Novotny, V., Prieschl, E. E., Csonga, R., Fabjani, G. and Baumruker, T. (1998). "Nrf1 in a complex with fosB, c-jun, junD and ATF2 forms the AP1 component at the TNF alpha promoter in stimulated mast cells". *Nucleic Acids Res* **26** (23): 5480-5485.

Nowicka, U., Zhang, D., Walker, O., Krutauz, D., Castaneda, C. A., Chaturvedi, A., Chen, T. Y., Reis, N., Glickman, M. H. and Fushman, D. (2015). "DNA-damage-inducible 1 protein (Ddi1) contains an uncharacteristic ubiquitin-like domain that binds ubiquitin". *Structure* **23** (3): 542-557.

O'Neill, B. M., Hanway, D., Winzeler, E. A. and Romesberg, F. E. (2004). "Coordinated functions of WSS1, PSY2 and TOF1 in the DNA damage response". *Nucleic Acids Res* **32** (22): 6519-6530.

Oda, Y., Hosokawa, N., Wada, I. and Nagata, K. (2003). "EDEM as an acceptor of terminally misfolded glycoproteins released from calnexin". *Science* **299** (5611): 1394-1397.

Ogi, T., Limsirichaikul, S., Overmeer, R. M., Volker, M., Takenaka, K., Cloney, R., Nakazawa, Y., Niimi, A., Miki, Y., Jaspers, N. G., Mullenders, L. H., Yamashita, S., Fousteri, M. I. and Lehmann, A. R. (2010). "Three DNA polymerases, recruited by different mechanisms, carry out NER repair synthesis in human cells". *Mol Cell* **37** (5): 714-727.

Oh, D. H., Rigas, D., Cho, A. and Chan, J. Y. (2012). "Deficiency in the nuclear-related factor erythroid 2 transcription factor (Nrf1) leads to genetic instability". *FEBS J* **279** (22): 4121-4130.

Ohtsuji, M., Katsuoka, F., Kobayashi, A., Aburatani, H., Hayes, J. D. and Yamamoto, M. (2008). "Nrf1 and Nrf2 play distinct roles in activation of antioxidant response element-dependent genes". *J Biol Chem* **283** (48): 33554-33562.

Osburn, W. O. and Kensler, T. W. (2008). "Nrf2 signaling: an adaptive response pathway for protection against environmental toxic insults". *Mutat Res* **659** (1-2): 31-39.

Pahl, H. L. and Baeuerle, P. A. (1994). "Oxygen and the control of gene expression". *Bioessays* **16** (7): 497-502.

Palicharla, V. R. and Maddika, S. (2015). "HACE1 mediated K27 ubiquitin linkage leads to YB-1 protein secretion". *Cell Signal* **27** (12): 2355-2362.

Palis, J. and Yoder, M. C. (2001). "Yolk-sac hematopoiesis: the first blood cells of mouse and man". *Exp Hematol* **29** (8): 927-936.

Papadopoulos, C., Kirchner, P., Bug, M., Grum, D., Koerver, L., Schulze, N., Poehler, R., Dressler, A., Fengler, S., Arhzaouy, K., Lux, V., Ehrmann, M., Weihl, C. C. and Meyer, H. (2017). "VCP/p97 cooperates with YOD1, UBXD1 and PLAA to drive clearance of ruptured lysosomes by autophagy". *EMBO J* **36** (2): 135-150.

Paraskevopoulos, K., Kriegenburg, F., Tatham, M. H., Rosner, H. I., Medina, B., Larsen, I. B., Brandstrup, R., Hardwick, K. G., Hay, R. T., Kragelund, B. B., Hartmann-Petersen, R. and Gordon, C. (2014). "Dss1 is a 26S proteasome ubiquitin receptor". *Mol Cell* **56** (3): 453-461.

Pepe, A. E., Xiao, Q., Zampetaki, A., Zhang, Z., Kobayashi, A., Hu, Y. and Xu, Q. (2010). "Crucial role of nrf3 in smooth muscle cell differentiation from stem cells". *Circ Res* **106** (5): 870-879.

Perez-Garcia, V., Fineberg, E., Wilson, R., Murray, A., Mazzeo, C. I., Tudor, C., Sienerth, A., White, J. K., Tuck, E., Ryder, E. J., Gleeson, D., Siragher, E., Wardle-Jones, H., Staudt, N., Wali, N., Collins, J., Geyer, S., Busch-Nentwich, E. M., Galli, A., Smith, J. C., Robertson, E., Adams, D. J., Weninger, W. J., Mohun, T. and Hemberger, M. (2018). "Placentation defects are highly prevalent in embryonic lethal mouse mutants". *Nature* **555** (7697): 463-468.

Perteguer, M. J., Gomez-Puertas, P., Canavate, C., Dagger, F., Garate, T. and Valdivieso, E. (2013). "Ddi1-like protein from Leishmania major is an active aspartyl proteinase". *Cell Stress Chaperones* **18** (2): 171-181.

Peters, J. M., Harris, J. R., Lustig, A., Muller, S., Engel, A., Volker, S. and Franke, W. W. (1992). "Ubiquitous soluble Mg(2+)-ATPase complex. A structural study". *J Mol Biol* **223** (2): 557-571.

Pickart, C. M. (2000). "Ubiquitin in chains". *Trends Biochem Sci* **25** (11): 544-548.

Pickart, C. M. and Rose, I. A. (1985). "Functional heterogeneity of ubiquitin carrier proteins". *J Biol Chem* **260** (3): 1573-1581.

Plemper, R. K., Bohmler, S., Bordallo, J., Sommer, T. and Wolf, D. H. (1997). "Mutant analysis links the translocon and BiP to retrograde protein transport for ER degradation". *Nature* **388** (6645): 891-895.

Pommier, Y., Barcelo, J. M., Rao, V. A., Sordet, O., Jobson, A. G., Thibaut, L., Miao, Z. H., Seiler, J. A., Zhang, H., Marchand, C., Agama, K., Nitiss, J. L. and Redon, C. (2006). "Repair of topoisomerase I-mediated DNA damage". *Prog Nucleic Acid Res Mol Biol* **81** 179-229.

Powers, E. T., Morimoto, R. I., Dillin, A., Kelly, J. W. and Balch, W. E. (2009). "Biological and chemical approaches to diseases of proteostasis deficiency". *Annu Rev Biochem* **78** 959-991.

Prakash, R., Zhang, Y., Feng, W. and Jasin, M. (2015). "Homologous recombination and human health: the roles of BRCA1, BRCA2, and associated proteins". *Cold Spring Harb Perspect Biol* **7** (4): a016600.

Prakash, S., Tian, L., Ratliff, K. S., Lehotzky, R. E. and Matouschek, A. (2004). "An unstructured initiation site is required for efficient proteasome-mediated degradation". *Nat Struct Mol Biol* **11** (9): 830-837.

Prieschl, E. E., Novotny, V., Csonga, R., Jaksche, D., Elbe-Burger, A., Thumb, W., Auer, M., Stingl, G. and Baumruker, T. (1998). "A novel splice variant of the transcription factor Nrf1 interacts with the TNFalpha promoter and stimulates transcription". *Nucleic Acids Res* **26** (10): 2291-2297.

Pye, V. E., Dreveny, I., Briggs, L. C., Sands, C., Beuron, F., Zhang, X. and Freemont, P. S. (2006). "Going through the motions: the ATPase cycle of p97". *J Struct Biol* **156** (1): 12-28.

Qi, L., Tsai, B. and Arvan, P. (2017). "New Insights into the Physiological Role of Endoplasmic Reticulum-Associated Degradation". *Trends Cell Biol* **27** (6): 430-440.

Qiu, X. B., Ouyang, S. Y., Li, C. J., Miao, S., Wang, L. and Goldberg, A. L. (2006). "hRpn13/ADRM1/GP110 is a novel proteasome subunit that binds the deubiquitinating enzyme, UCH37". *EMBO J* **25** (24): 5742-5753.

Raasi, S., Orlov, I., Fleming, K. G. and Pickart, C. M. (2004). "Binding of polyubiquitin chains to ubiquitin-associated (UBA) domains of HHR23A". *J Mol Biol* **341** (5): 1367-1379.

Rada, P., Rojo, A. I., Chowdhry, S., McMahon, M., Hayes, J. D. and Cuadrado, A. (2011). "SCF/{beta}-TrCP promotes glycogen synthase kinase 3-dependent degradation of the Nrf2 transcription factor in a Keap1-independent manner". *Mol Cell Biol* **31** (6): 1121-1133.

Rada, P., Rojo, A. I., Evrard-Todeschi, N., Innamorato, N. G., Cotte, A., Jaworski, T., Tobon-Velasco, J. C., Devijver, H., Garcia-Mayoral, M. F., Van Leuven, F., Hayes, J. D., Bertho, G. and Cuadrado, A. (2012). "Structural and functional characterization of Nrf2 degradation by the glycogen synthase kinase 3/beta-TrCP axis". *Mol Cell Biol* **32** (17): 3486-3499.

Radhakrishnan, S. K., den Besten, W. and Deshaies, R. J. (2014). "p97-dependent retrotranslocation and proteolytic processing govern formation of active Nrf1 upon proteasome inhibition". *Elife* **3** e01856.

Radhakrishnan, S. K., Lee, C. S., Young, P., Beskow, A., Chan, J. Y. and Deshaies, R. J. (2010). "Transcription factor Nrf1 mediates the proteasome recovery pathway after proteasome inhibition in mammalian cells". *Mol Cell* **38** (1): 17-28.

Rahighi, S., Ikeda, F., Kawasaki, M., Akutsu, M., Suzuki, N., Kato, R., Kensche, T., Uejima, T., Bloor, S., Komander, D., Randow, F., Wakatsuki, S. and Dikic, I. (2009). "Specific recognition of linear ubiquitin chains by NEMO is important for NF-kappaB activation". *Cell* **136** (6): 1098-1109.

Ramirez, J., Lectez, B., Osinalde, N., Siva, M., Elu, N., Aloria, K., Prochazkova, M., Perez, C., Martinez-Hernandez, J., Barrio, R., Saskova, K. G., Arizmendi, J. M. and Mayor, U. (2018). "Quantitative proteomics reveals neuronal ubiquitination of Rngo/Ddi1 and several proteasomal subunits by Ube3a, accounting for the complexity of Angelman syndrome". *Hum Mol Genet* **27** (11): 1955-1971.

Renshaw, P. S., Veverka, V., Kelly, G., Frenkiel, T. A., Williamson, R. A., Gordon, S. V., Hewinson, R. G. and Carr, M. D. (2004). "Sequence-specific assignment and secondary structure determination of the 195-residue complex formed by the Mycobacterium tuberculosis proteins CFP-10 and ESAT-6". *J Biomol NMR* **30** (2): 225-226.

Rhee, S. G. (1999). "Redox signaling: hydrogen peroxide as intracellular messenger". *Exp Mol Med* **31** (2): 53-59.

Rhee, S. G. (2006). "Cell signaling. H2O2, a necessary evil for cell signaling". *Science* **312** (5782): 1882-1883.

Richly, H., Rape, M., Braun, S., Rumpf, S., Hoege, C. and Jentsch, S. (2005). "A series of ubiquitin binding factors connects CDC48/p97 to substrate multiubiquitylation and proteasomal targeting". *Cell* **120** (1): 73-84.

Rinaldi, T., Ricci, C., Porro, D., Bolotin-Fukuhara, M. and Frontali, L. (1998). "A mutation in a novel yeast proteasomal gene, RPN11/MPR1, produces a cell cycle arrest, overreplication of nuclear and mitochondrial DNA, and an altered mitochondrial morphology". *Mol Biol Cell* **9** (10): 2917-2931.

Robzyk, K., Recht, J. and Osley, M. A. (2000). "Rad6-dependent ubiquitination of histone H2B in yeast". *Science* **287** (5452): 501-504.

Rodrijguez de Lores, A., Alberici de Canal, M. and De Robertis, E. (1971). "Turnover of proteins in subcellular fractions of rat cerebral cortex". *Brain Res* **31** (1): 179-184.

Rojo de la Vega, M., Chapman, E. and Zhang, D. D. (2018). "NRF2 and the Hallmarks of Cancer". *Cancer Cell* **34** (1): 21-43.

Rosenzweig, R., Osmulski, P. A., Gaczynska, M. and Glickman, M. H. (2008). "The central unit within the 19S regulatory particle of the proteasome". *Nat Struct Mol Biol* **15** (6): 573-580.

Rosenzweig, R., Bronner, V., Zhang, D., Fushman, D. and Glickman, M. H. (2012). "Rpn1 and Rpn2 coordinate ubiquitin processing factors at proteasome". *J Biol Chem* **287** (18): 14659-14671.

Rossant, J. and Cross, J. C. (2001). "Placental development: lessons from mouse mutants". *Nat Rev Genet* **2** (7): 538-548.

Rubinsztein, D. C. (2006). "The roles of intracellular protein-degradation pathways in neurodegeneration". *Nature* **443** (7113): 780-786.

Ruggiano, A., Foresti, O. and Carvalho, P. (2014). "Quality control: ER-associated degradation: protein quality control and beyond". *J Cell Biol* **204** (6): 869-879.

Rushmore, T. H., Morton, M. R. and Pickett, C. B. (1991). "The antioxidant responsive element. Activation by oxidative stress and identification of the DNA consensus sequence required for functional activity". *J Biol Chem* **266** (18): 11632-11639.

Sadikovic, B., Fernandes, P., Zhang, V. W., Ward, P. A., Miloslavskaya, I., Rhead, W., Rosenbaum, R., Gin, R., Roa, B. and Fang, P. (2014). "Mutation Update for UBE3A variants in Angelman syndrome". *Hum Mutat* **35** (12): 1407-1417.

Saeki, Y., Sone, T., Toh-e, A. and Yokosawa, H. (2002a). "Identification of ubiquitin-like protein-binding subunits of the 26S proteasome". *Biochem Biophys Res Commun* **296** (4): 813-819.

Saeki, Y., Saitoh, A., Toh-e, A. and Yokosawa, H. (2002b). "Ubiquitin-like proteins and Rpn10 play cooperative roles in ubiquitin-dependent proteolysis". *Biochem Biophys Res Commun* **293** (3): 986-992.

Saito, Y., Tsubuki, S., Ito, H. and Kawashima, S. (1990). "The structure-function relationship between peptide aldehyde derivatives on initiation of neurite outgrowth in PC12h cells". *Neurosci Lett* **120** (1): 1-4.

Sakata, E., Bohn, S., Mihalache, O., Kiss, P., Beck, F., Nagy, I., Nickell, S., Tanaka, K., Saeki, Y., Forster, F. and Baumeister, W. (2012). "Localization of the proteasomal ubiquitin receptors Rpn10 and Rpn13 by electron cryomicroscopy". *Proc Natl Acad Sci U S A* **109** (5): 1479-1484.

Sale, J. E. (2012). "Competition, collaboration and coordination--determining how cells bypass DNA damage". *J Cell Sci* **125** (Pt 7): 1633-1643.

Sale, J. E. (2013). "Translesion DNA synthesis and mutagenesis in eukaryotes". *Cold Spring Harb Perspect Biol* **5** (3): a012708.

San Filippo, J., Sung, P. and Klein, H. (2008). "Mechanism of eukaryotic homologous recombination". *Annu Rev Biochem* **77** 229-257.

Sankaranarayanan, K. and Jaiswal, A. K. (2004). "Nrf3 negatively regulates antioxidant-response element-mediated expression and antioxidant induction of NAD(P)H:quinone oxidoreductase1 gene". *J Biol Chem* **279** (49): 50810-50817.

Sato, B. K. and Hampton, R. Y. (2006). "Yeast Derlin Dfm1 interacts with Cdc48 and functions in ER homeostasis". *Yeast* **23** (14-15): 1053-1064.

Satoh, T. and Lipton, S. (2017). "Recent advances in understanding NRF2 as a druggable target: development of pro-electrophilic and non-covalent NRF2 activators to overcome systemic side effects of electrophilic drugs like dimethyl fumarate". *F1000Res* **6** 2138.

Sauer, H., Wartenberg, M. and Hescheler, J. (2001). "Reactive oxygen species as intracellular messengers during cell growth and differentiation". *Cell Physiol Biochem* **11** (4): 173-186.

Savas, J. N., Toyama, B. H., Xu, T., Yates, J. R., 3rd and Hetzer, M. W. (2012). "Extremely long-lived nuclear pore proteins in the rat brain". *Science* **335** (6071): 942.

Savoia, D., Allice, T. and Tovo, P. A. (2005). "Antileishmanial activity of HIV protease inhibitors". *Int J Antimicrob Agents* **26** (1): 92-94.

Schaeffer, V., Akutsu, M., Olma, M. H., Gomes, L. C., Kawasaki, M. and Dikic, I. (2014). "Binding of OTULIN to the PUB domain of HOIP controls NF-kappaB signaling". *Mol Cell* **54** (3): 349-361.

Schaetzlein, S., Chahwan, R., Avdievich, E., Roa, S., Wei, K., Eoff, R. L., Sellers, R. S., Clark, A. B., Kunkel, T. A., Scharff, M. D. and Edelmann, W. (2013). "Mammalian Exo1 encodes both structural and catalytic functions that play distinct roles in essential biological processes". *Proc Natl Acad Sci U S A* **110** (27): E2470-2479.

Schauber, C., Chen, L., Tongaonkar, P., Vega, I., Lambertson, D., Potts, W. and Madura, K. (1998). "Rad23 links DNA repair to the ubiquitin/proteasome pathway". *Nature* **391** (6668): 715-718.

Scheffner, M., Nuber, U. and Huibregtse, J. M. (1995). "Protein ubiquitination involving an E1-E2-E3 enzyme ubiquitin thioester cascade". *Nature* **373** (6509): 81-83.

Schindler, A. J. and Schekman, R. (2009). "In vitro reconstitution of ER-stress induced ATF6 transport in COPII vesicles". *Proc Natl Acad Sci U S A* **106** (42): 17775-17780.

Schreck, R., Rieber, P. and Baeuerle, P. A. (1991). "Reactive oxygen intermediates as apparently widely used messengers in the activation of the NF-kappa B transcription factor and HIV-1". *EMBO J* **10** (8): 2247-2258.

Schreiner, P., Chen, X., Husnjak, K., Randles, L., Zhang, N., Elsasser, S., Finley, D., Dikic, I., Walters, K. J. and Groll, M. (2008). "Ubiquitin docking at the proteasome through a novel pleckstrin-homology domain interaction". *Nature* **453** (7194): 548-552.

Schuberth, C. and Buchberger, A. (2005). "Membrane-bound Ubx2 recruits Cdc48 to ubiquitin ligases and their substrates to ensure efficient ER-associated protein degradation". *Nat Cell Biol* **7** (10): 999-1006.

Schulman, B. A. and Harper, J. W. (2009). "Ubiquitin-like protein activation by E1 enzymes: the apex for downstream signalling pathways". *Nat Rev Mol Cell Biol* **10** (5): 319-331.

Schultz, M. A., Abdel-Mageed, A. B. and Mondal, D. (2010). "The nrf1 and nrf2 balance in oxidative stress regulation and androgen signaling in prostate cancer cells". *Cancers (Basel)* **2** (2): 1354-1378.

Schultz, S. G. (2002). "William Harvey and the circulation of the blood: the birth of a scientific revolution and modern physiology". *News Physiol Sci* **17** 175-180.

Seemuller, E., Lupas, A. and Baumeister, W. (1996). "Autocatalytic processing of the 20S proteasome". *Nature* **382** (6590): 468-471.

Seemuller, E., Lupas, A., Stock, D., Lowe, J., Huber, R. and Baumeister, W. (1995). "Proteasome from Thermoplasma acidophilum: a threonine protease". *Science* **268** (5210): 579-582.

Sha, Z. and Goldberg, A. L. (2014). "Proteasome-mediated processing of Nrf1 is essential for coordinate induction of all proteasome subunits and p97". *Curr Biol* **24** (14): 1573-1583.

Shen, J., Snapp, E. L., Lippincott-Schwartz, J. and Prywes, R. (2005). "Stable binding of ATF6 to BiP in the endoplasmic reticulum stress response". *Mol Cell Biol* **25** (3): 921-932.

Shen, Y., Delaglio, F., Cornilescu, G. and Bax, A. (2009). "TALOS+: a hybrid method for predicting protein backbone torsion angles from NMR chemical shifts". *J Biomol NMR* **44** (4): 213-223.

Shepard, T. H., Tanimura, T. and Park, H. W. (1997). "Glucose absorption and utilization by rat embryos". *Int J Dev Biol* **41** (2): 307-314.

Shi, Y., Chen, X., Elsasser, S., Stocks, B. B., Tian, G., Lee, B. H., Shi, Y., Zhang, N., de Poot, S. A., Tuebing, F., Sun, S., Vannoy, J., Tarasov, S. G., Engen, J. R., Finley, D. and Walters, K. J. (2016). "Rpn1 provides adjacent receptor sites for substrate binding and deubiquitination by the proteasome". *Science* **351** (6275):

Siegenthaler, B., Defila, C., Muzumdar, S., Beer, H. D., Meyer, M., Tanner, S., Bloch, W., Blank, V., Schafer, M. and Werner, S. (2018). "Nrf3 promotes UV-induced keratinocyte apoptosis through suppression of cell adhesion". *Cell Death Differ* **25** (10): 1749-1765.

Sies, H., Berndt, C. and Jones, D. P. (2017). "Oxidative Stress". *Annu Rev Biochem* **86** 715-748.

Silva-Islas, C. A. and Maldonado, P. D. (2018). "Canonical and non-canonical mechanisms of Nrf2 activation". *Pharmacol Res* **134** 92-99.

Simpson, M. V. (1953). "The release of labeled amino acids from the proteins of rat liver slices". *J Biol Chem* **201** (1): 143-154.

Sirkis, R., Gerst, J. E. and Fass, D. (2006). "Ddi1, a eukaryotic protein with the retroviral protease fold". *J Mol Biol* **364** (3): 376-387.

Siva, M., Svoboda, M., Veverka, V., Trempe, J. F., Hofmann, K., Kozisek, M., Hexnerova, R., Sedlak, F., Belza, J., Brynda, J., Sacha, P., Hubalek, M., Starkova, J., Flaisigova, I., Konvalinka, J. and Saskova, K. G. (2016). "Human DNA-Damage-Inducible 2 Protein Is Structurally and Functionally Distinct from Its Yeast Ortholog". *Sci Rep* **6** 30443.

Skinner-Adams, T. S., McCarthy, J. S., Gardiner, D. L., Hilton, P. M. and Andrews, K. T. (2004). "Antiretrovirals as antimalarial agents". *J Infect Dis* **190** (11): 1998-2000.

Skowyra, D., Craig, K. L., Tyers, M., Elledge, S. J. and Harper, J. W. (1997). "F-box proteins are receptors that recruit phosphorylated substrates to the SCF ubiquitin-ligase complex". *Cell* **91** (2): 209-219.

Sloper-Mould, K. E., Jemc, J. C., Pickart, C. M. and Hicke, L. (2001). "Distinct functional surface regions on ubiquitin". *J Biol Chem* **276** (32): 30483-30489.

Song, M. O., Mattie, M. D., Lee, C. H. and Freedman, J. H. (2014). "The role of Nrf1 and Nrf2 in the regulation of copper-responsive transcription". *Exp Cell Res* **322** (1): 39-50.

Steffen, J., Seeger, M., Koch, A. and Kruger, E. (2010). "Proteasomal degradation is transcriptionally controlled by TCF11 via an ERAD-dependent feedback loop". *Mol Cell* **40** (1): 147-158.

Stein, A., Ruggiano, A., Carvalho, P. and Rapoport, T. A. (2014). "Key steps in ERAD of luminal ER proteins reconstituted with purified components". *Cell* **158** (6): 1375-1388.

Stewart, D., Killeen, E., Naquin, R., Alam, S. and Alam, J. (2003). "Degradation of transcription factor Nrf2 via the ubiquitin-proteasome pathway and stabilization by cadmium". *J Biol Chem* **278** (4): 2396-2402.

Stieren, E. S., El Ayadi, A., Xiao, Y., Siller, E., Landsverk, M. L., Oberhauser, A. F., Barral, J. M. and Boehning, D. (2011). "Ubiquilin-1 is a molecular chaperone for the amyloid precursor protein". *J Biol Chem* **286** (41): 35689-35698.

Stingele, J., Habermann, B. and Jentsch, S. (2015). "DNA-protein crosslink repair: proteases as DNA repair enzymes". *Trends Biochem Sci* **40** (2): 67-71.

Stingele, J., Schwarz, M. S., Bloemeke, N., Wolf, P. G. and Jentsch, S. (2014). "A DNA-dependent protease involved in DNA-protein crosslink repair". *Cell* **158** (2): 327-338.

Stingele, J., Bellelli, R., Alte, F., Hewitt, G., Sarek, G., Maslen, S. L., Tsutakawa, S. E., Borg, A., Kjaer, S., Tainer, J. A., Skehel, J. M., Groll, M. and Boulton, S. J. (2016). "Mechanism and Regulation of DNA-Protein Crosslink Repair by the DNA-Dependent Metalloprotease SPRTN". *Mol Cell* **64** (4): 688-703.

Swanson, R., Locher, M. and Hochstrasser, M. (2001). "A conserved ubiquitin ligase of the nuclear envelope/endoplasmic reticulum that functions in both ER-associated and Matalpha2 repressor degradation". *Genes Dev* **15** (20): 2660-2674.

Swatek, K. N. and Komander, D. (2016). "Ubiquitin modifications". *Cell Res* **26** (4): 399-422.

Takeuchi, J., Chen, H. and Coffino, P. (2007). "Proteasome substrate degradation requires association plus extended peptide". *EMBO J* **26** (1): 123-131.

Tao, S., Liu, P., Luo, G., Rojo de la Vega, M., Chen, H., Wu, T., Tillotson, J., Chapman, E. and Zhang, D. D. (2017). "p97 Negatively Regulates NRF2 by Extracting Ubiquitylated NRF2 from the KEAP1-CUL3 E3 Complex". *Mol Cell Biol* **37** (8):

Tarentino, A. L. and Plummer, T. H., Jr. (1994). "Enzymatic deglycosylation of asparagine-linked glycans: purification, properties, and specificity of oligosaccharide-cleaving enzymes from Flavobacterium meningosepticum". *Methods Enzymol* **230** 44-57.

Telakowski-Hopkins, C. A., King, R. G. and Pickett, C. B. (1988). "Glutathione S-transferase Ya subunit gene: identification of regulatory elements required for basal level and inducible expression". *Proc Natl Acad Sci U S A* **85** (4): 1000-1004.

Tell, G., Quadrifoglio, F., Tiribelli, C. and Kelley, M. R. (2009). "The many functions of APE1/Ref-1: not only a DNA repair enzyme". *Antioxid Redox Signal* **11** (3): 601-620.

The UniProt, C. (2017). "UniProt: the universal protein knowledgebase". *Nucleic Acids Res* **45** (D1): D158-D169.

Theodore, M., Kawai, Y., Yang, J., Kleshchenko, Y., Reddy, S. P., Villalta, F. and Arinze, I. J. (2008). "Multiple nuclear localization signals function in the nuclear import of the transcription factor Nrf2". *J Biol Chem* **283** (14): 8984-8994.

Thibault, G. and Ng, D. T. (2012). "The endoplasmic reticulum-associated degradation pathways of budding yeast". *Cold Spring Harb Perspect Biol* **4** (12):

Tomko, R. J., Jr., Funakoshi, M., Schneider, K., Wang, J. and Hochstrasser, M. (2010). "Heterohexameric ring arrangement of the eukaryotic proteasomal ATPases: implications for proteasome structure and assembly". *Mol Cell* **38** (3): 393-403.

Tong, K. I., Katoh, Y., Kusunoki, H., Itoh, K., Tanaka, T. and Yamamoto, M. (2006). "Keap1 recruits Neh2 through binding to ETGE and DLG motifs: characterization of the two-site molecular recognition model". *Mol Cell Biol* **26** (8): 2887-2900.

Toyama, B. H. and Hetzer, M. W. (2013). "Protein homeostasis: live long, won't prosper". *Nat Rev Mol Cell Biol* **14** (1): 55-61.

Trempe, J. F., Saskova, K. G., Siva, M., Ratcliffe, C. D., Veverka, V., Hoegl, A., Menade, M., Feng, X., Shenker, S., Svoboda, M., Kozisek, M., Konvalinka, J. and Gehring, K. (2016). "Structural studies of the yeast DNA damage-inducible protein Ddi1 reveal domain architecture of this eukaryotic protein family". *Sci Rep* **6** 33671.

Tsaytler, P., Harding, H. P., Ron, D. and Bertolotti, A. (2011). "Selective inhibition of a regulatory subunit of protein phosphatase 1 restores proteostasis". *Science* **332** (6025): 91-94.

Tsubuki, S., Saito, Y., Tomioka, M., Ito, H. and Kawashima, S. (1996). "Differential inhibition of calpain and proteasome activities by peptidyl aldehydes of di-leucine and tri-leucine". *J Biochem* **119** (3): 572-576.

Tsubuki, S., Kawasaki, H., Saito, Y., Miyashita, N., Inomata, M. and Kawashima, S. (1993). "Purification and characterization of a Z-Leu-Leu-Leu-MCA degrading protease expected to regulate neurite formation: a novel catalytic activity in proteasome". *Biochem Biophys Res Commun* **196** (3): 1195-1201.

Tsuchiya, Y., Morita, T., Kim, M., Iemura, S., Natsume, T., Yamamoto, M. and Kobayashi, A. (2011). "Dual regulation of the transcriptional activity of Nrf1 by beta-TrCP- and Hrd1-dependent degradation mechanisms". *Mol Cell Biol* **31** (22): 4500-4512.

Tsuchiya, Y., Taniguchi, H., Ito, Y., Morita, T., Karim, M. R., Ohtake, N., Fukagai, K., Ito, T., Okamuro, S., Iemura, S., Natsume, T., Nishida, E. and Kobayashi, A. (2013). "The casein kinase 2-nrf1 axis controls the clearance of ubiquitinated proteins by regulating proteasome gene expression". *Mol Cell Biol* **33** (17): 3461-3472.

Tsujita, T., Peirce, V., Baird, L., Matsuyama, Y., Takaku, M., Walsh, S. V., Griffin, J. L., Uruno, A., Yamamoto, M. and Hayes, J. D. (2014). "Transcription factor Nrf1 negatively regulates the cystine/glutamate transporter and lipid-metabolizing enzymes". *Mol Cell Biol* **34** (20): 3800-3816.

Turgeon, M. O., Perry, N. J. S. and Poulogiannis, G. (2018). "DNA Damage, Repair, and Cancer Metabolism". *Front Oncol* **8** 15.

Uhlen, M., Zhang, C., Lee, S., Sjostedt, E., Fagerberg, L., Bidkhori, G., Benfeitas, R., Arif, M., Liu, Z., Edfors, F., Sanli, K., von Feilitzen, K., Oksvold, P., Lundberg, E., Hober, S., Nilsson, P., Mattsson, J., Schwenk, J. M., Brunnstrom, H., Glimelius, B., Sjoblom, T., Edqvist, P. H., Djureinovic, D., Micke, P., Lindskog, C., Mardinoglu, A. and Ponten, F. (2017). "A pathology atlas of the human cancer transcriptome". *Science* **357** (6352):

Ulrich, H. D. and Walden, H. (2010). "Ubiquitin signalling in DNA replication and repair". *Nat Rev Mol Cell Biol* **11** (7): 479-489.

van den Boom, J. and Meyer, H. (2018). "VCP/p97-Mediated Unfolding as a Principle in Protein Homeostasis and Signaling". *Mol Cell* **69** (2): 182-194.

Van Meter, M., Simon, M., Tombline, G., May, A., Morello, T. D., Hubbard, B. P., Bredbenner, K., Park, R., Sinclair, D. A., Bohr, V. A., Gorbunova, V. and Seluanov, A. (2016). "JNK Phosphorylates SIRT6 to Stimulate DNA Double-Strand Break Repair in Response to Oxidative Stress by Recruiting PARP1 to DNA Breaks". *Cell Rep* **16** (10): 2641-2650.

Vander Linden, R. T., Hemmis, C. W., Schmitt, B., Ndoja, A., Whitby, F. G., Robinson, H., Cohen, R. E., Yao, T. and Hill, C. P. (2015). "Structural basis for the activation and inhibition of the UCH37 deubiquitylase". *Mol Cell* **57** (5): 901-911.

Vangala, J. R. and Radhakrishnan, S. K. (2018). "Nrf1-mediated transcriptional regulation of the proteasome requires a functional TIP60 complex". *J Biol Chem*

Vaquero, E. C., Edderkaoui, M., Pandol, S. J., Gukovsky, I. and Gukovskaya, A. S. (2004). "Reactive oxygen species produced by NAD(P)H oxidase inhibit apoptosis in pancreatic cancer cells". *J Biol Chem* **279** (33): 34643-34654.

Venugopal, R. and Jaiswal, A. K. (1998). "Nrf2 and Nrf1 in association with Jun proteins regulate antioxidant response element-mediated expression and coordinated induction of genes encoding detoxifying enzymes". *Oncogene* **17** (24): 3145-3156.

Verma, R., Oania, R., Graumann, J. and Deshaies, R. J. (2004). "Multiubiquitin chain receptors define a layer of substrate selectivity in the ubiquitin-proteasome system". *Cell* **118** (1): 99-110.

Verma, R., Aravind, L., Oania, R., McDonald, W. H., Yates, J. R., 3rd, Koonin, E. V. and Deshaies, R. J. (2002). "Role of Rpn11 metalloprotease in deubiquitination and degradation by the 26S proteasome". *Science* **298** (5593): 611-615.

Verzijl, N., DeGroot, J., Thorpe, S. R., Bank, R. A., Shaw, J. N., Lyons, T. J., Bijlsma, J. W., Lafeber, F. P., Baynes, J. W. and TeKoppele, J. M. (2000). "Effect of collagen turnover on the accumulation of advanced glycation end products". *J Biol Chem* **275** (50): 39027-39031.

Veverka, V., Lennie, G., Crabbe, T., Bird, I., Taylor, R. J. and Carr, M. D. (2006). "NMR assignment of the mTOR domain responsible for rapamycin binding". *J Biomol NMR* **36 Suppl 1** 3.

Veverka, V., Crabbe, T., Bird, I., Lennie, G., Muskett, F. W., Taylor, R. J. and Carr, M. D. (2008). "Structural characterization of the interaction of mTOR with phosphatidic acid and a novel class of inhibitor: compelling evidence for a central role of the FRB domain in small molecule-mediated regulation of mTOR". *Oncogene* **27** (5): 585-595.

Vij, R., Siegel, D. S., Jagannath, S., Jakubowiak, A. J., Stewart, A. K., McDonagh, K., Bahlis, N., Belch, A., Kunkel, L. A., Wear, S., Wong, A. F. and Wang, M. (2012). "An open-label, single-arm, phase 2 study of single-agent carfilzomib in patients with relapsed and/or refractory multiple myeloma who have been previously treated with bortezomib". *Br J Haematol* **158** (6): 739-748.

Vijay-Kumar, S., Bugg, C. E. and Cook, W. J. (1987). "Structure of ubiquitin refined at 1.8 A resolution". *J Mol Biol* **194** (3): 531-544.

Vitale, I., Manic, G., De Maria, R., Kroemer, G. and Galluzzi, L. (2017). "DNA Damage in Stem Cells". *Mol Cell* **66** (3): 306-319.

Volker, M., Mone, M. J., Karmakar, P., van Hoffen, A., Schul, W., Vermeulen, W., Hoeijmakers, J. H., van Driel, R., van Zeeland, A. A. and Mullenders, L. H. (2001). "Sequential assembly of the nucleotide excision repair factors in vivo". *Mol Cell* **8** (1): 213-224.

Wakasugi, M., Kawashima, A., Morioka, H., Linn, S., Sancar, A., Mori, T., Nikaido, O. and Matsunaga, T. (2002). "DDB accumulates at DNA damage sites immediately after UV irradiation and directly stimulates nucleotide excision repair". *J Biol Chem* **277** (3): 1637-1640.

Walker, J. R., Corpina, R. A. and Goldberg, J. (2001). "Structure of the Ku heterodimer bound to DNA and its implications for double-strand break repair". *Nature* **412** (6847): 607-614.

Walter, P. and Ron, D. (2011). "The unfolded protein response: from stress pathway to homeostatic regulation". *Science* **334** (6059): 1081-1086.

Walters, K. J., Kleijnen, M. F., Goh, A. M., Wagner, G. and Howley, P. M. (2002). "Structural studies of the interaction between ubiquitin family proteins and proteasome subunit S5a". *Biochemistry* **41** (6): 1767-1777.

Wang, G., Zhong, S., Zhang, S. Y., Ma, Z. L., Chen, J. L., Lu, W. H., Cheng, X., Chuai, M., Lee, K. K., Lu, D. X. and Yang, X. (2016). "Angiogenesis is repressed by ethanol exposure during chick embryonic development". *J Appl Toxicol* **36** (5): 692-701.

Wang, H., Zhan, M., Yang, R., Shi, Y., Liu, Q. and Wang, J. (2018). "Elevated expression of NFE2L3 predicts the poor prognosis of pancreatic cancer patients". *Cell Cycle* **17** (17): 2164-2174.

Wang, H., Rosidi, B., Perrault, R., Wang, M., Zhang, L., Windhofer, F. and Iliakis, G. (2005). "DNA ligase III as a candidate component of backup pathways of nonhomologous end joining". *Cancer Res* **65** (10): 4020-4030.

Wang, W. and Chan, J. Y. (2006). "Nrf1 is targeted to the endoplasmic reticulum membrane by an N-terminal transmembrane domain. Inhibition of nuclear translocation and transacting function". *J Biol Chem* **281** (28): 19676-19687.

Wang, W., Kwok, A. M. and Chan, J. Y. (2007). "The p65 isoform of Nrf1 is a dominant negative inhibitor of ARE-mediated transcription". *J Biol Chem* **282** (34): 24670-24678.

Wang, X. J., Hayes, J. D. and Wolf, C. R. (2006). "Generation of a stable antioxidant response element-driven reporter gene cell line and its use to show redox-dependent activation of nrf2 by cancer chemotherapeutic agents". *Cancer Res* **66** (22): 10983-10994.

Waters, L. S., Minesinger, B. K., Wiltrout, M. E., D'Souza, S., Woodruff, R. V. and Walker, G. C. (2009). "Eukaryotic translesion polymerases and their roles and regulation in DNA damage tolerance". *Microbiol Mol Biol Rev* **73** (1): 134-154.

Watson, E. D. and Cross, J. C. (2005). "Development of structures and transport functions in the mouse placenta". *Physiology (Bethesda)* **20** 180-193.

Weissman, A. M., Shabek, N. and Ciechanover, A. (2011). "The predator becomes the prey: regulating the ubiquitin system by ubiquitylation and degradation". *Nat Rev Mol Cell Biol* **12** (9): 605-620.

Wenzel, D. M., Lissounov, A., Brzovic, P. S. and Klevit, R. E. (2011). "UBCH7 reactivity profile reveals parkin and HHARI to be RING/HECT hybrids". *Nature* **474** (7349): 105-108.

White, R. E., Powell, D. J. and Berry, C. (2011a). "HIV proteinase inhibitors target the Ddi1-like protein of Leishmania parasites". *FASEB J* **25** (5): 1729-1736.

White, R. E., Dickinson, J. R., Semple, C. A., Powell, D. J. and Berry, C. (2011b). "The retroviral proteinase active site and the N-terminus of Ddi1 are required for repression of protein secretion". *FEBS Lett* **585** (1): 139-142.

Wickliffe, K. E., Williamson, A., Meyer, H. J., Kelly, A. and Rape, M. (2011). "K11-linked ubiquitin chains as novel regulators of cell division". *Trends Cell Biol* **21** (11): 656-663.

Wilkinson, C. R., Seeger, M., Hartmann-Petersen, R., Stone, M., Wallace, M., Semple, C. and Gordon, C. (2001). "Proteins containing the UBA domain are able to bind to multi-ubiquitin chains". *Nat Cell Biol* **3** (10): 939-943.

Wilkinson, D. G. and Nieto, M. A. (1993). "Detection of messenger RNA by in situ hybridization to tissue sections and whole mounts". *Methods Enzymol* **225** 361-373.

Wong, P. M., Chung, S. W., Chui, D. H. and Eaves, C. J. (1986). "Properties of the earliest clonogenic hemopoietic precursors to appear in the developing murine yolk sac". *Proc Natl Acad Sci U S A* **83** (11): 3851-3854.

Worden, E. J., Padovani, C. and Martin, A. (2014). "Structure of the Rpn11-Rpn8 dimer reveals mechanisms of substrate deubiquitination during proteasomal degradation". *Nat Struct Mol Biol* **21** (3): 220-227.

Wu-Baer, F., Lagrazon, K., Yuan, W. and Baer, R. (2003). "The BRCA1/BARD1 heterodimer assembles polyubiquitin chains through an unconventional linkage involving lysine residue K6 of ubiquitin". *J Biol Chem* **278** (37): 34743-34746.

Wu, G., Fang, Y. Z., Yang, S., Lupton, J. R. and Turner, N. D. (2004). "Glutathione metabolism and its implications for health". *J Nutr* **134** (3): 489-492.

Xiang, Y., Wang, M., Hu, S., Qiu, L., Yang, F., Zhang, Z., Yu, S., Pi, J. and Zhang, Y. (2018). "Mechanisms controlling the multistage post-translational processing of endogenous Nrf1alpha/TCF11 proteins to yield distinct isoforms within the coupled positive and negative feedback circuits". *Toxicol Appl Pharmacol* **360** 212-235.

Xiao, Q., Pepe, A. E., Wang, G., Luo, Z., Zhang, L., Zeng, L., Zhang, Z., Hu, Y., Ye, S. and Xu, Q. (2012). "Nrf3-Pla2g7 interaction plays an essential role in smooth muscle differentiation from stem cells". *Arterioscler Thromb Vasc Biol* **32** (3): 730-744.

Xing, W., Singgih, A., Kapoor, A., Alarcon, C. M., Baylink, D. J. and Mohan, S. (2007). "Nuclear factor-E2-related factor-1 mediates ascorbic acid induction of osterix expression via interaction with antioxidant-responsive element in bone cells". *J Biol Chem* **282** (30): 22052-22061.

Xu, Z., Chen, L., Leung, L., Yen, T. S., Lee, C. and Chan, J. Y. (2005). "Liver-specific inactivation of the Nrf1 gene in adult mouse leads to nonalcoholic steatohepatitis and hepatic neoplasia". *Proc Natl Acad Sci U S A* **102** (11): 4120-4125.

Yamane, T. (2018). "Mouse Yolk Sac Hematopoiesis". *Front Cell Dev Biol* **6** 80.

Yao, T. and Cohen, R. E. (2002). "A cryptic protease couples deubiquitination and degradation by the proteasome". *Nature* **419** (6905): 403-407.

Yao, T., Song, L., Xu, W., DeMartino, G. N., Florens, L., Swanson, S. K., Washburn, M. P., Conaway, R. C., Conaway, J. W. and Cohen, R. E. (2006). "Proteasome recruitment and activation of the Uch37 deubiquitinating enzyme by Adrm1". *Nat Cell Biol* **8** (9): 994-1002.

Ye, Y. and Rape, M. (2009). "Building ubiquitin chains: E2 enzymes at work". *Nat Rev Mol Cell Biol* **10** (11): 755-764.

Ye, Y., Meyer, H. H. and Rapoport, T. A. (2003). "Function of the p97-Ufd1-Npl4 complex in retrotranslocation from the ER to the cytosol: dual recognition of nonubiquitinated polypeptide segments and polyubiquitin chains". *J Cell Biol* **162** (1): 71-84.

Ye, Y., Tang, W. K., Zhang, T. and Xia, D. (2017). "A Mighty "Protein Extractor" of the Cell: Structure and Function of the p97/CDC48 ATPase". *Front Mol Biosci* **4** 39.

Ye, Y., Shibata, Y., Yun, C., Ron, D. and Rapoport, T. A. (2004). "A membrane protein complex mediates retro-translocation from the ER lumen into the cytosol". *Nature* **429** (6994): 841-847.

Yokouchi, M., Kondo, T., Houghton, A., Bartkiewicz, M., Horne, W. C., Zhang, H., Yoshimura, A. and Baron, R. (1999). "Ligand-induced ubiquitination of the epidermal growth factor receptor involves the interaction of the c-Cbl RING finger and UbcH7". *J Biol Chem* **274** (44): 31707-31712.

Yoshida, H., Oku, M., Suzuki, M. and Mori, K. (2006). "pXBP1(U) encoded in XBP1 pre-mRNA negatively regulates unfolded protein response activator pXBP1(S) in mammalian ER stress response". *J Cell Biol* **172** (4): 565-575.

Young, P., Deveraux, Q., Beal, R. E., Pickart, C. M. and Rechsteiner, M. (1998). "Characterization of two polyubiquitin binding sites in the 26 S protease subunit 5a". *J Biol Chem* **273** (10): 5461-5467.

Yuan, W. C., Lee, Y. R., Lin, S. Y., Chang, L. Y., Tan, Y. P., Hung, C. C., Kuo, J. C., Liu, C. H., Lin, M. Y., Xu, M., Chen, Z. J. and Chen, R. H. (2014). "K33-Linked Polyubiquitination of Coronin 7 by Cul3-KLHL20 Ubiquitin E3 Ligase Regulates Protein Trafficking". *Mol Cell* **54** (4): 586-600.

Zerbino, D. R., Achuthan, P., Akanni, W., Amode, M. R., Barrell, D., Bhai, J., Billis, K., Cummins, C., Gall, A., Giron, C. G., Gil, L., Gordon, L., Haggerty, L., Haskell, E., Hourlier, T., Izuogu, O. G., Janacek, S. H., Juettemann, T., To, J. K., Laird, M. R., Lavidas, I., Liu, Z., Loveland, J. E., Maurel, T., McLaren, W., Moore, B., Mudge, J., Murphy, D. N., Newman, V., Nuhn, M., Ogeh, D., Ong, C. K., Parker, A., Patricio, M., Riat, H. S., Schuilenburg, H., Sheppard, D., Sparrow, H., Taylor, K., Thormann, A., Vullo, A., Walts, B., Zadissa, A., Frankish, A., Hunt, S. E., Kostadima, M., Langridge, N., Martin, F. J., Muffato, M., Perry, E., Ruffier, M., Staines, D. M., Trevanion, S. J., Aken, B. L., Cunningham, F., Yates, A. and Flicek, P. (2018). "Ensembl 2018". *Nucleic Acids Res* **46** (D1): D754-D761.

Zhang, D., Chen, T., Ziv, I., Rosenzweig, R., Matiuhin, Y., Bronner, V., Glickman, M. H. and Fushman, D. (2009). "Together, Rpn10 and Dsk2 can serve as a polyubiquitin chain-length sensor". *Mol Cell* **36** (6): 1018-1033.

Zhang, S. T., Zhao, R., Ma, W. X., Fan, Y. Y., Guan, W. Z., Wang, J., Ren, P., Zhong, K., Yu, T. S., Pi, J. B. and Guan, D. W. (2013). "Nrf1 is time-dependently expressed and distributed in the distinct cell types after trauma to skeletal muscles in rats". *Histol Histopathol* **28** (6): 725-735.

Zhang, T., Xu, Y., Liu, Y. and Ye, Y. (2015). "gp78 functions downstream of Hrd1 to promote degradation of misfolded proteins of the endoplasmic reticulum". *Mol Biol Cell* **26** (24): 4438-4450.

Zhang, Y. and Hayes, J. D. (2010). "Identification of topological determinants in the N-terminal domain of transcription factor Nrf1 that control its orientation in the endoplasmic reticulum membrane". *Biochem J* **430** (3): 497-510.

Zhang, Y. and Manning, B. D. (2015). "mTORC1 signaling activates NRF1 to increase cellular proteasome levels". *Cell Cycle* **14** (13): 2011-2017.

Zhang, Y., Crouch, D. H., Yamamoto, M. and Hayes, J. D. (2006). "Negative regulation of the Nrf1 transcription factor by its N-terminal domain is independent of Keap1: Nrf1, but not Nrf2, is targeted to the endoplasmic reticulum". *Biochem J* **399** (3): 373-385.

Zhang, Y., Kobayashi, A., Yamamoto, M. and Hayes, J. D. (2009). "The Nrf3 transcription factor is a membrane-bound glycoprotein targeted to the endoplasmic reticulum through its N-terminal homology box 1 sequence". *J Biol Chem* **284** (5): 3195-3210.

Zhang, Y., Ren, Y., Li, S. and Hayes, J. D. (2014a). "Transcription factor Nrf1 is topologically repartitioned across membranes to enable target gene transactivation through its acidic glucose-responsive domains". *PLoS One* **9** (4): e93458.

Zhang, Y., Qiu, L., Li, S., Xiang, Y., Chen, J. and Ren, Y. (2014b). "The C-terminal domain of Nrf1 negatively regulates the full-length CNC-bZIP factor and its shorter isoform LCR-F1/Nrf1beta; both are also inhibited by the small dominant-negative Nrf1gamma/delta isoforms that down-regulate ARE-battery gene expression". *PLoS One* **9** (10): e109159.

Zhang, Y., Li, S., Xiang, Y., Qiu, L., Zhao, H. and Hayes, J. D. (2015). "The selective post-translational processing of transcription factor Nrf1 yields distinct isoforms that dictate its ability to differentially regulate gene expression". *Sci Rep* **5** 12983.

Zhang, Y., Nicholatos, J., Dreier, J. R., Ricoult, S. J., Widenmaier, S. B., Hotamisligil, G. S., Kwiatkowski, D. J. and Manning, B. D. (2014c). "Coordinated regulation of protein synthesis and degradation by mTORC1". *Nature* **513** (7518): 440-443.

Zhao, G., Zhou, X., Wang, L., Li, G., Schindelin, H. and Lennarz, W. J. (2007). "Studies on peptide:N-glycanase-p97 interaction suggest that p97 phosphorylation modulates endoplasmic reticulum-associated degradation". *Proc Natl Acad Sci U S A* **104** (21): 8785-8790.

Zheng, H., Fu, J., Xue, P., Zhao, R., Dong, J., Liu, D., Yamamoto, M., Tong, Q., Teng, W., Qu, W., Zhang, Q., Andersen, M. E. and Pi, J. (2015). "CNC-bZIP protein Nrf1-dependent regulation of glucose-stimulated insulin secretion". *Antioxid Redox Signal* **22** (10): 819-831.

Zheng, N. and Shabek, N. (2017). "Ubiquitin Ligases: Structure, Function, and Regulation". *Annu Rev Biochem* **86** 129-157.

Zhu, Y. and Xiao, W. (1998). "Differential regulation of two closely clustered yeast genes, MAG1 and DDI1, by cell-cycle checkpoints". *Nucleic Acids Res* **26** (23): 5402-5408.

Zovein, A. C., Turlo, K. A., Ponec, R. M., Lynch, M. R., Chen, K. C., Hofmann, J. J., Cox, T. C., Gasson, J. C. and Iruela-Arispe, M. L. (2010). "Vascular remodeling of the vitelline artery initiates extravascular emergence of hematopoietic clusters". *Blood* **116** (18): 3435-3444.

Zwickl, P., Kleinz, J. and Baumeister, W. (1994). "Critical elements in proteasome assembly". *Nat Struct Biol* **1** (11): 765-770.

Zwickl, P., Voges, D. and Baumeister, W. (1999). "The proteasome: a macromolecular assembly designed for controlled proteolysis". *Philos Trans R Soc Lond B Biol Sci* **354** (1389): 1501-1511.

APPENDICES

APPENDIX 1

Siva, M., Svoboda, M., Veverka, V., Trempe, J. F., Hofmann, K., Kozisek, M., Hexnerova, R., Sedlak, F., Belza, J., Brynda, J., Sacha, P., Hubalek, M., Starkova, J., Flaisigova, I., Konvalinka, J. and Saskova, K. G. (2016). "Human DNA-Damage-Inducible 2 Protein Is Structurally and Functionally Distinct from Its Yeast Ortholog". *Sci Rep* **6** 30443. **IF (2016) = 4.259**

SCIENTIFIC REPORTS

OPEN

Human DNA-Damage-Inducible 2 Protein Is Structurally and Functionally Distinct from Its Yeast Ortholog

Received: 02 July 2015
Accepted: 04 July 2016
Published: 27 July 2016

Monika Sivá[1,2,3,*], Michal Svoboda[1,4,*], Václav Veverka[1], Jean-François Trempe[5], Kay Hofmann[6], Milan Kožíšek[1], Rozálie Hexnerová[1], František Sedlák[1,2,3], Jan Belza[1,3], Jiří Brynda[1], Pavel Šácha[1], Martin Hubálek[1], Jana Starková[1], Iva Flaisigová[1], Jan Konvalinka[1,3] & Klára Grantz Šašková[1,3]

Although Ddi1-like proteins are conserved among eukaryotes, their biological functions remain poorly characterized. Yeast Ddi1 has been implicated in cell cycle regulation, DNA-damage response, and exocytosis. By virtue of its ubiquitin-like (UBL) and ubiquitin-associated (UBA) domains, it has been proposed to serve as a proteasomal shuttle factor. All Ddi1-like family members also contain a highly conserved retroviral protease-like (RVP) domain with unknown substrate specificity. While the structure and biological function of yeast Ddi1 have been investigated, no such analysis is available for the human homologs. To address this, we solved the 3D structures of the human Ddi2 UBL and RVP domains and identified a new helical domain that extends on either side of the RVP dimer. While Ddi1-like proteins from all vertebrates lack a UBA domain, we identify a novel ubiquitin-interacting motif (UIM) located at the C-terminus of the protein. The UIM showed a weak yet specific affinity towards ubiquitin, as did the Ddi2 UBL domain. However, the full-length Ddi2 protein is unable to bind to di-ubiquitin chains. While proteomic analysis revealed no activity, implying that the protease requires other factors for activation, our structural characterization of all domains of human Ddi2 sets the stage for further characterization.

The ubiquitin-proteasome system (UPS) plays a crucial role in eukaryotic cell biology. Pathway components are involved in processes including protein degradation and trafficking, cell signaling, response to DNA damage, and cell cycle regulation. Ubiquitin (UBQ) is a central molecule in the pathway, and its ability to form various polymeric chains marks substrates for specific tasks[1,2]. Controlling mechanisms by which the chains are recognized are important for proper system function and cellular homeostasis. Imbalance in any step of the pathway can have significant impact on an organism, and thus, complete understanding of this central pathway is essential.

Polyubiquitination marks proteins for multiple fates, such as degradation or vesicle sorting. Polyubiquitinated proteins that undergo degradation are either recognized directly by proteasomal receptors (Rpn10, Rpn13) or "captured" by so-called shuttle (or adaptor) proteins (Rad23, Dsk2, and Ddi1 in budding yeast). The shuttles deliver their polyubiquitinated substrates to the regulatory part of the 26S proteasome[3–9]. Proteasomal shuttle proteins possess a typical domain architecture that includes an N-terminal ubiquitin-like domain (UBL) that binds the 26S proteasome and a C-terminal ubiquitin-associated domain (UBA) responsible for binding UBQ or poly-UBQ chains[10].

[1]Gilead Sciences and IOCB Research Center, Institute of Organic Chemistry and Biochemistry of the Academy of Sciences of the Czech Republic, Flemingovo n. 2, 166 10 Prague 6, Czech Republic. [2]First Faculty of Medicine, Charles University in Prague, Katerinska 32, 121 08, Prague 2, Czech Republic. [3]Department of Biochemistry, Faculty of Science, Charles University, Hlavova 8, 128 00 Prague 2, Czech Republic. [4]Department of Physical and Macromolecular Chemistry, Faculty of Science, Charles University, Hlavova 8, 128 00 Prague 2, Czech Republic. [5]Groupe de Recherche Axé sur la Structure des Protéines, Department of Pharmacology & Therapeutics, McGill University, Montreal, QC, H3G 1Y6, Canada. [6]Institute for Genetics, University of Cologne, Zülpicher Str. 47a, 50647 Cologne, Germany. *These authors contributed equally to this work. Correspondence and requests for materials should be addressed to K.G.Š. (email: saskova@uochb.cas.cz)

In line with this UBL-UBA domain architecture, DNA damage-inducible (Ddi1)-like proteins are thought to act as proteasomal shuttle proteins, although the evidence for this function is incomplete[9–12]. Recently, Nowicka and co-workers proposed an alternative mechanism for the yeast Ddi1 (yDdi1) shuttling process based on the surprising fact that yDdi1 UBL binds UBQ[13]. Yet another factor differentiates Ddi1-like proteins from classical proteasomal shuttles: Ddi1-like proteins contain an additional domain called the retroviral protease-like (RVP) domain, the 3D fold of which is strikingly reminiscent of HIV-1 protease. RVP is highly conserved in eukaryotes, and is present in human Ddi1-like orthologs. It contains the catalytic triad characteristic of aspartic proteases (D[T/S]G) and is responsible for dimerization of the protein (Fig. 1A)[11,14]. The physiological substrate of this putative aspartic protease, if any, remains unknown.

Ddi1 from *Saccharomyces cerevisiae* is by far the best-studied Ddi1-like ortholog. Its expression is DNA-damage inducible, and it is involved in cell cycle progression through the mitotic checkpoint protein Pds1[15,16]. Studies from the Raveh laboratory indicate that it plays a role in degradation of HO endonuclease, the enzyme responsible for switching alleles at the mating type locus *MAT*[9]. Furthermore, yDdi1 interacts with the exo- and endocytotic v-SNARE proteins Snc1 and Snc2 as well as exocytotic t-SNARE Sso1, playing a role as a negative regulator of exocytosis[11,17,18].

Overall, the current body of knowledge indicates that Ddi1-like proteins play a significant role in cell cycle control, growth control, and trafficking in yeast and may play a crucial role in embryogenesis in higher eukaryotes. Ddi1-like orthologs from higher eukaryotes have not been investigated in much detail. Notably, Ddi1-like protein from *Caenorhabditis elegans* (Vsm-1) may play a crucial role in synaptogenesis[19]. In *Drosophila melanogaster*, knock-out of the *Rngo* (fruit fly *DDI1* homolog) gene is lethal and forms ring canal defects in oogenesis[20]. Moreover, a high-throughput proteomics study identified Rngo protein as one of the most abundant ubiquitinated proteins during neural development in *Drosophila* embryogenesis[21].

The highly conserved RVP domain poses an interesting evolutionary puzzle. The 3D structure of yDdi1 RVP was solved by others (PDB code 2I1A)[22] at 2.3 Å resolution and very recently by us at 1.9 Å resolution. Our structure shows the conformation of the "flap" region in detail (HIV terminology), which was missing in the previous model (details are presented in our back-to-back publication, Trempe *et al.*, 2016)[22–24]. However, the structure of the RVP domain of human Ddi2 (hDdi2) has not been published to date. The putative active site of yDdi1 RVP is similar to that of HIV-1 protease, including a water molecule that could act as a nucleophile for peptide bond hydrolysis. The first direct evidence that Ddi1-like RVP can act as a protease was presented by Perteguer and coworkers, who showed that a *Leishmania major* Ddi1-like ortholog cleaves BSA at acidic pH[25]. In addition, they showed that it hydrolyzes one HIV peptide substrate and two cathepsin D substrates and that this activity can be inhibited by specific aspartic protease inhibitors. This evidence was supported by another finding showing that knock-out of yDdi1 leads to an increase in protein secretion into the media[17] and can be complemented by transfection of a plasmid encoding Ddi1. Complementation requires both the UBL and Asp220 of the RVP active site[26]. White and coworkers reported the similar finding that the yDdi1 knock-out phenotype can be rescued by a plasmid encoding human or leishmanial Ddi1. This rescue is inhibited by some HIV protease inhibitors[27]. Data obtained with Rngo, the Ddi1-like ortholog from *Drosophila*, also supports the hypothesis that Ddi1 is an active protease: the oogenesis-defect phenotype can be fully rescued by transgenes encoding full-length Rngo or Rngo lacking either the UBL or UBA domain. In contrast, the phenotype cannot be rescued by Rngo protein variant with a mutated catalytic aspartate in the RVP domain (D257A)[20]. Therefore, it is clear that Ddi1-like RVP is required for its biological function, although its physiological substrate remains elusive.

In the human genome, there are two genes (located on chromosome 11 and chromosome 1) encoding Ddi1-like proteins: the 396-amino-acid Ddi homolog 1 (hDdi1) and the 399-amino-acid Ddi homolog 2 (hDdi2). Based on its genomic organization, hDdi2 seems to be the "original" version of yDdi1 that later gave rise to hDdi1 through a retrotransposition event. To the best of our knowledge, neither protein has been specifically studied. They share 70% amino acid sequence identity and 81% similarity. Compared to the protein domain architecture of lower eukaryotes that of both human variants is conserved only to a certain extent. While the UBL and RVP domains are preserved, the UBA domain is missing. Therefore, the putative function of human Ddi1-like proteins as proteasomal shuttles is questionable, and their biological role remains elusive.

We present here the first structural and functional study of hDdi2. We first analyze the evolutionary pathway leading to the loss of the UBA domain. We identify a putative short UBQ-interacting motif (UIM) at the C-terminus, instead of UBA, and we show its specific but very weak binding to UBQ. Prompted by the recent results from Nowicka and coworkers, we solved the 3D structure of hDdi2 UBL and performed NMR titrations with UBQ. While the yDdi1 UBL binds to UBQ[13,23], we observe only a weak affinity of hDdi2 UBL for UBQ. We extended our investigations to UBQ conjugates and showed that hDdi2 does not bind any di-UBQ chains *in vitro*. We also present the first 3D structure of the hDdi2 RVP domain, together with its functional proteolytic analysis. Finally, we used NMR to elucidate the structure of the region preceding the RVP domain, which we named the Helical Domain of hDdi2 (HDD), and describe its characteristic features.

Results

Evolution of Ddi1-like proteins: loss of UBA and identification of a novel ubiquitin-interacting motif in human Ddi2. Ddi1-like proteins, which combine an N-terminal UBL domain with an intact RVP, arose early in eukaryotic evolution. Database searches with sequence profiles for UBL and RVP domains have detected widespread occurrence of these proteins in animals, plants, and fungi[28], as well as in protozoan lineages including apicomplexans, kinetoplastids, and oomycetes. The majority of UBL-RVP containing proteins also possess a C-terminal UBA domain, suggesting that they might act as proteasomal shuttling factors similar to yDdi1[29]. However, Ddi1-like proteins from all vertebrate families appear to have lost the UBA domain, although it is retained in other animal lineages. In the mammalian lineage, the UBA-deficient gene was duplicated, giving rise to two related UBL-RVP-containing genes, called *DDI1* and *DDI2* in humans. Despite their names, yDdi1

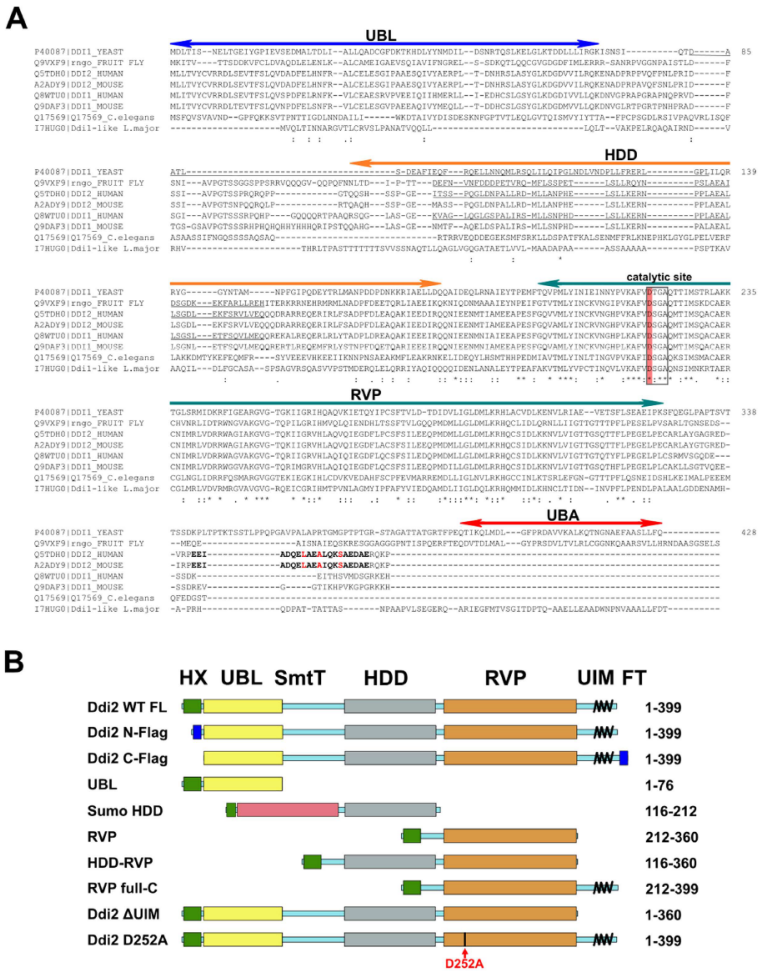


Figure 1. Sequence analysis of Ddi1 orthologs. (**A**) Sequence alignment of Ddi1-like proteins from various eukaryotic organisms. Domains are indicated with double-headed arrows. The highly conserved catalytic site of RVP is highlighted. The putative UIM motif is highlighted in bold, with residues important for ubiquitin binding in red. (**B**) Schematic diagram of full-length hDdi2 and the truncated constructs used in this study. Positions of the histidine tag including the factor Xa cleavage site (green), UBL domain (yellow), HDD (gray), RVP domain (orange), and C-terminal UIM (black helix) are indicated. Flexible regions are indicated with blue boxes. Mutation of the putative catalytic aspartate (D252A) is indicated with a red arrow.

and its non-mammalian homologs are more similar to hDdi2 than to hDdi1. Because the human *DDI2* gene also shares conserved synteny with the single *DDI1*-like gene of non-mammalian vertebrates, *DDI2* is assumed to be the "original" version that later gave rise to the intron-less mammalian *DDI1* through a retrotransposition event.

Closer inspection of the mammalian *DDI2* locus and corresponding loci in non-mammalian vertebrates shed light on the evolutionary fate of the C-terminal UBA domain. Early in vertebrate evolution, a novel vertebrate-specific gene called *RSC1A1* apparently became inserted into the ancestral *DDI2* locus, separating the N-terminal UBL-RVP portion from the C-terminal UBA-containing region. In extant vertebrates, the UBA

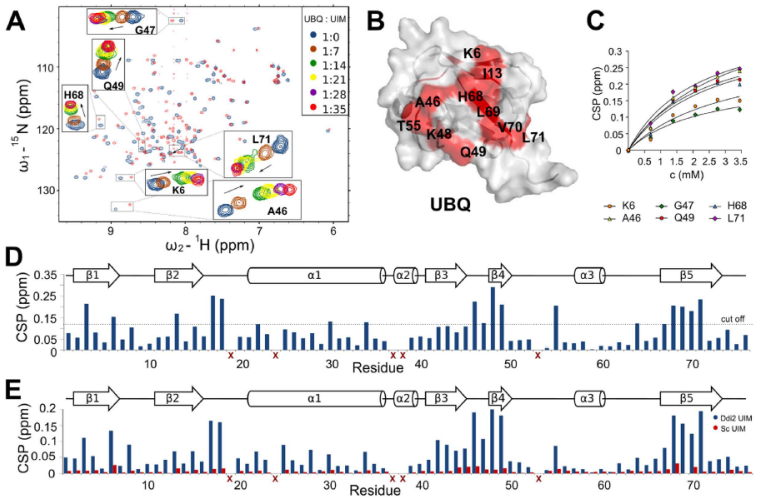


Figure 2. Mapping of the UBQ-hDdi2 interaction site. (**A**) $^{15}N/^{1}H$-HSQC titration spectra of UBQ with hDdi2-UIM peptide. (**B**) Identification of mapped residues shown on the UBQ structure (PDB entry 1D3Z)[32]. (**C**) Titration curves of selected amino acids on UBQ. (**D**) Plot of chemical shift perturbations of individual amino acids upon interaction at the end point of the titration (35-fold molar excess). Red crosses mark amino acids that were not reliably observed in the titration spectra. (**E**) Plots of chemical shift perturbations of UBQ residues upon interaction with 2.2 mM hDdi2-UIM peptide (blue) and upon addition of hDdi2-scrambled UIM peptide (red) to a final concentration of 1.9 mM.

domain has become part of the RSC1A1 polypeptide and might participate in this protein's function of regulating the trafficking of sugar transporters[30].

Considering the putative role of hDdi2 as a shuttle protein for the UPS, we performed a bioinformatics analysis of the newly evolved C-terminus to identify potential alternative UBQ-binding domains to the lost UBA domain. Alignment of Ddi1-like sequences from various organisms revealed a conserved region of 24 residues that is absent from yDdi1 and non-vertebrate Ddi1-like sequences. Comparison of this region to databases of annotated domains using the program HHPRED revealed significant similarity ($p < 0.0001$) to a family of ubiquitin-interacting motif (UIM) proteins[31]. As shown in Fig. 1, the pattern of UBQ-binding residues typical of UIM motifs is conserved in the Ddi2 family, suggesting that this newly identified UIM-like motif might replace the lost UBA domain as a UBQ receptor.

The C-terminal UIM motif of human Ddi2 binds weakly, yet selectively to mono-UBQ. To evaluate the putative ability of the C-terminal UIM of hDdi2 to bind UBQ, we performed NMR chemical shift perturbation (CSP) experiments with UBQ and either 1) hDdi2-UIM peptide (hDdi2 residues 376–396); 2) hDdi2-scrambled UIM peptide; 3) the full C-terminus of Ddi2 including the RVP domain (hDdi2 RVP-UIM full-C, residues 212–399). After assignment of both double and triple resonance spectra of ^{15}N and $^{15}N/^{13}C$-labeled protein constructs (RVP full-C and UBQ), we analyzed specific shifts in positions of backbone amide signals induced by the addition of non-labeled peptide or protein partner (Fig. 2).

First, we titrated UBQ with UIM peptide. We reached a UIM peptide concentration of 3.45 mM (35-fold molar excess over UBQ) and determined the K_d between 2.2–3.2 mM. The K_d was calculated from 6 residues (Lys6, Ala46, Gly47, Gln49, His68, and Leu71) by fitting the titration curves with a 1:1 stoichiometry model for specific binding (Fig. 2C). The CSPs are illustrated in the overlaid spectra, with and without final addition of the peptide, with a close-up on significantly shifted peaks (used for K_d calculation) that were mapped onto the UBQ structure (PDB 1D3Z) (Fig. 2A,B)[32]. Based on shifts in residues used for fitting the titration and in Leu8, Arg42, Lys48, Gln49, and Leu71, we concluded that the binding epitope is slightly different compared to the Ile44 hydrophobic patch (Fig. 2D). However, we observed different shifts in backbone amides of other amino acids (Ile3, Ile13, Val17, Glu18, Glu34, Thr55, Glu64, and Leu69). The control experiment with the hDdi2-scrambled UIM peptide revealed no significant CSPs in comparison to equimolar addition of the hDdi2-UIM peptide (Fig. 2E), suggesting that the weak interaction between the UIM and ubiquitin is nonetheless specific.

Guided by previous NMR data with isolated motifs, we next examined binding of ^{15}N-labeled UBQ with addition of a 1-, 2-, and 5-fold molar excess of non-labeled hDdi2 RVP full-C, which could provide a more refined map of the interaction (Figure S1A). Relatively small yet specific changes in positions of backbone signals

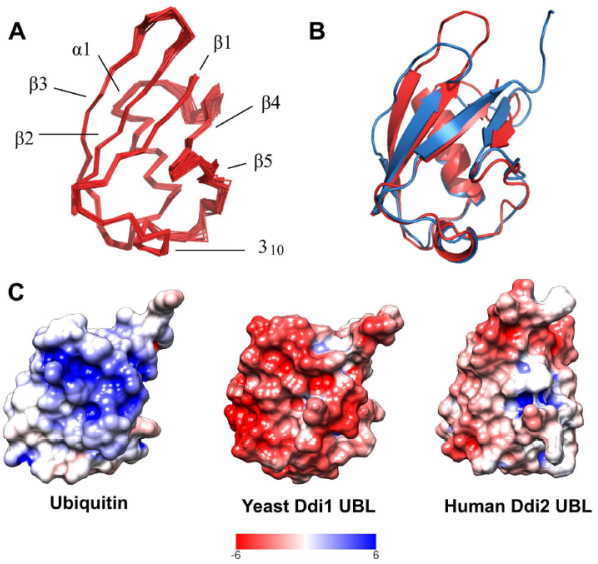


Figure 3. Solution structure of the hDdi2 UBL domain. (**A**) Superimposition of 40 converged structures of the UBL domain. (**B**) Structural alignment of solution structures of the yDdi1 UBL in blue (PDB code 2N7E) and hDdi2 UBL in orange (PDB code 2N7D). The structural alignment over 74 equivalent positions yields an RMSD of 1.66 Å[36]. (**C**) Comparison of the surface electrostatic potential of ubiquitin (PDB 1UBQ), yDdi1 UBL (accompanying paper by Trempe)[23], and hDdi2 UBL. For NMR structures, representative structures closest to the mean structure were used, but similar results were obtained with the first structures of the ensembles. All molecules are oriented based on secondary structure alignment, with the β-sheet area towards the reader. The surface is colored from red (negative values) to blue (positive values); the range is ±6 kT/e for all structures. Surface electrostatic potential maps were generated using the Adaptive Poisson Boltzmann Solver[57] package with structure preprocessing using the PDB2PQR tool[58] in the UCSF Chimera software package[55]. All calculations were performed using the SWANSON force field at pH 7.4; other settings were kept at default values. Chimera was also used for final surface visualization.

were observed for residues Thr7, Arg42, Lys48, Gln49, and Leu71, which were slightly different from those seen in the Ile44 patch known to interact with several UBAs and UIM[10,24,33] (Figure S1A). We also performed the reverse experiment with ^{15}N-labeled hDdi2 RVP full-C protein and addition of a 1-, 2-, and 5-fold molar excess of non-labeled UBQ. The alignment of HSQC spectra during the titration revealed shifts in individual residues located at the Ddi2-UIM peptide sequence (Figure S1B). Overall, the data suggest that UBQ binds to the C-terminal sequence harboring the putative UIM, but with very weak affinity.

Inspired by the work of Singh and co-workers showing specific interaction of yDdi1 and Rub1 (the closest relative of UBQ, Nedd8 in mammals)[34], we performed similar NMR CSP experiments to investigate the possibility of Nedd8 binding to hDdi2. In this case, we did not observe any significant perturbation with the C-terminal hDdi2 UIM peptide (Figure S2) nor with the N-terminal UBL domain of hDdi2 (Figure S3A). Therefore, we conclude that the C-terminal UIM of hDdi2 specifically binds UBQ.

The UBL domain from human Ddi2 binds more weakly to ubiquitin than the yeast Ddi1 UBL. To gain deeper structural information about hDdi2, we obtained nearly complete ^{15}N-, ^{13}C-, and ^{1}H-resonance assignments of its N-terminal UBL domain (residues 1–76, with N-terminal histidine tag) and determined the solution structure with high precision. The root mean-squared deviation (RMSD) to the mean structure for the backbone and heavy atoms for the final 40 converged structures was 0.4 Å overall and 1 Å at the ordered residue range (residues 1–76 of the protein sequence). The UBL of hDdi2 contains five β-sheets (β1: M1-V8, β2: V15-V21, β3: Q46-Y49, β4: R52-P53, β5: V71-R75), one α-helix (L27-S38), and a 3_{10}-helix (L61-Y64), which is consistent with the typical UBQ β-grasp fold (Fig. 3A). The distribution of NMR constraints and structural statistics for the hDdi2 UBL domain are summarized in Table S1.

To characterize the binding properties of hDdi2 UBL, we inspected its structure and performed a detailed comparison with the UBL structure of yDdi1 reported in our back-to-back publication[23]. The sequence similarity between the yeast and human UBL domains is 46%, and despite their low sequence identity (25%)[35], their

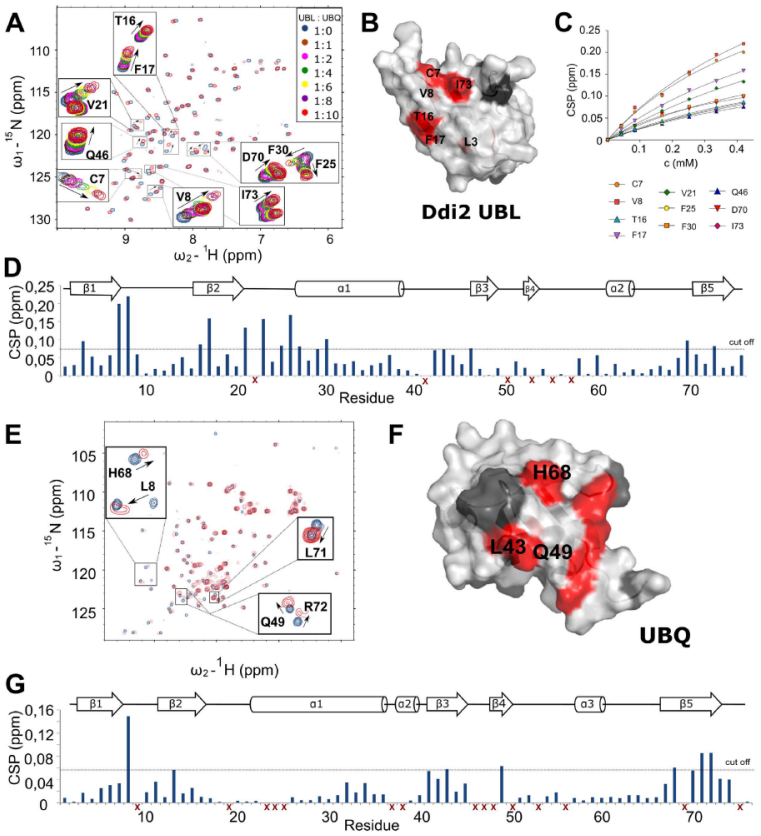


Figure 4. Characterization of the hDdi2 UBL interaction with UBQ. (**A**) $^{15}N/^{1}H$-HSQC titration spectra of Ddi2 UBL with addition of a 1-, 2-, 4-, 6-, 8-, or 10-fold molar excess of UBQ. Residues Cys7, Val8, Thr16, Phe17, Val21, Phe25, Phe30, Gln46, Asp70, and Ile73 were used for K_d calculation (0.42–1.1 mM). (**B**) The mapped interaction site shown on the UBL structure is most likely located in the β-sheet area, according to shifts in Leu3, Cys7, Val8, Thr16, Phe17, and Ile73 upon UBQ binding. Additional shifts in backbone amides observed in the spectra (Val21, Ala23, Phe25, Glu26, Phe30, and Asp70) at the other site of the domain could be the result of a structural change upon binding. Amino acids that could not be used for evaluation are marked black. (**C**) Titration curves of selected hDdi2 UBL amino acids used for K_d calculation according to the 1:1 stoichiometry model for specific binding. (**D**) CSP plot showing perturbation at the titration endpoint. Residues not considered in the evaluation are marked with red crosses. (**E**) $^{15}N/^{1}H$-HSQC titration spectra of UBQ with final 6-fold excess of hDdi2 ΔUIM with close-ups of the shifted signals of individual amino acids mapped (**F**) onto UBQ (PDB entry 1D3Z) (**G**) Plots of chemical shift perturbations of individual amino acids of UBQ.

secondary structure elements superimpose very well with a backbone RMSD of 1.66 Å[36] (Fig. 3B). We compared the surface properties of the interaction patches from both yDdi1 and hDdi2 UBLs and UBQ (Fig. 3C). As discussed by Nowicka and co-workers[13], the β-sheet interaction area of yDdi1 UBL is formed by positively charged side chains, which makes it complementary to the negatively charged UBQ patch. Interestingly, the surface electrostatic potential of hDdi2 UBL shows a small hydrophobic area that is moderately charged. We reasoned, that due to different charge distribution on the interaction patch of hDdi2 UBL and yDdi1 UBL, they might interact with different partners.

Prompted by the unexpected finding of Nowicka and co-workers that yDdi1 UBL binds UBQ with a K_d of $45 \pm 7\,\mu M$, we investigated whether hDdi2 UBL has any affinity for UBQ[13]. We performed NMR titration experiments on ^{15}N-labeled hDdi2 UBL with addition of UBQ up to a 10-fold molar excess (Fig. 4A). We mapped the

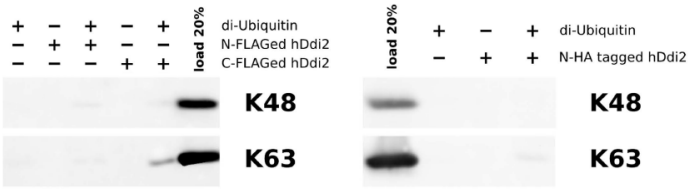


Figure 5. Human Ddi2 shows no strong interaction with di-ubiquitin chains. Western blot analysis of pull-down experiments with di-ubiquitin conjugates of Lys48 and Lys63 architecture. Human Ddi2 with a FLAG tag on either the N- or C-terminus or an HA tag on the N-terminus was immobilized on magnetic agarose beads. Beads were incubated with the di-ubiquitin conjugate of given linkage architecture, washed, and eluted by boiling in non-reducing SDS sample buffer. Samples were analyzed on 18% SDS-PAGE followed by immunoblotting with anti-UBQ antibody.

most relevant shifts onto the structure of hDdi2 UBL (Fig. 4B), which showed that this interaction is located in the β-sheet area, with a K_d in the 0.42–1.1 mM range, calculated from 10 residues (Fig. 4C). This interaction was supported by a reverse experiment with ^{15}N-labeled UBQ titrated with non-labeled hDdi2 ΔUIM (lacking UIM) to a 6-fold molar excess. We mapped the changes in HSQC spectra onto the site close to Ile44 patch (Fig. 4E–G). A negative control experiment with 6-fold molar addition of hDdi2 HDD-RVP (lacking both UIM and UBL) did not show any significant CSPs of the UBQ backbone amide signals (Figure S3B). On the basis of these data, we infer that unlike the yDdi1 UBL domain, the hDdi2 UBL domain interacts weakly with UBQ with a Kd in the low millimolar range.

We next examined whether the UBL of hDdi2 could bind the protein's C-terminal UIM motif. We performed NMR titration experiments with ^{15}N-labeled hDdi2 UBL with addition of hDdi2-UIM peptide to a final concentration of 1.9 mM (Figure S3C), as well as negative control experiment with the same molar addition of hDdi2-scrambled UIM peptide. Both resulted in the same low CSP response (Figure S3C). We next measured and superimposed HSQC spectra of ^{15}N-labeled full-length hDdi2 and the ΔUIM truncated form of hDdi2 to elucidate the potential intramolecular interaction (Figure S3D). No difference was observed in the chemical shifts corresponding to the hDdi2 UBL domain, suggesting that hDdi2 UBL cannot bind its own C-terminal UIM and most likely never adopts a "head-to-tail" auto-inhibited conformation. Interestingly, superimposition of the HSQC spectra of ^{15}N-labeled full-length protein with its UBL domain revealed shifts in almost all N-terminal amino acids of hDdi2 (Figure S3E). This demonstrates that the UBL domain binds and is not independent from the rest of the protein, in contrast to the yDdi1 UBL[13,23].

Polyubiquitin chain binding is not preserved in human Ddi2. Given that the interaction between hDdi2 and mono-UBQ is very weak and completely different from that of yDdi1 and UBQ, we wondered whether these weak interactions mediated by the UBL and UIM motifs could synergize to enable polyvalent binding to ubiquitin chains. Therefore, we tested the binding full-length hDdi2 to various UBQ chains (Fig. 5, Figure S4). N- and C-terminally FLAG-tagged hDdi2 and HA-tagged hDdi2 were immobilized on magnetic beads and mixed independently with all eight native linkage types of di-UBQ conjugates (Lys6-, Lys11-, Lys27-, Lys29, Lys33-, Lys48-, Lys63-linked, and linear). The same experiment was repeated also with in house synthetized Lys48- and Lys63-linked chains. The data clearly shows that hDdi2 does not pull down any of di-UBQ conjugates under physiological pH. This contrasts with yDdi1, which binds to polyubiquitin chains[10].

The structure of the helical domain of human Ddi2 reveals a conserved bundle fold. Given the weak interaction of hDdi2 with ubiquitin, we looked for other domains in the protein to gain further insight into the function of the protein. Bioinformatics sequence analysis revealed strong conservation in the region preceding the RVP domain of hDdi2 (positions 116–212; Fig. 1). Within this region, we detected similarity to the Sti1 domain (residues 125–178), an α-helical domain found in the proteasome shuttle proteins Rad23 and Dsk2 and their animal homologs (Figure S5). The remainder of the region shows helicity as well, but does not share detectable similarity with other protein families. We refer to the entire α-helical bundle spanning residues 125–212 as the helical domain of Ddi (HDD).

The NMR structure of the hDdi2 HDD domain confirmed our prediction that this region adopts an α-helical folded structure (Fig. 6A and Table S1). The hDdi2 HDD structure consists of a globular arrangement of 4 α-helices spanning the following residues (Fig. 6B): helix 1 (135–144), helix 2 (146–155), helix 3 (157–164), and helix 4 (168–190). The region is preceded by two turns of another α-helix that is not included in the numbering. All four major helices pack against each other, forming a compact bundle with a hydrophobic core made up mostly of leucine residues. The bundle is further supported by a salt bridge between helix 3 and the initial part of helix 4, including residues Ser165 and Lys170, with occasional contribution of Glu161 (Fig. 6C). Helix 4 spans 22 amino acids with an interesting accumulation of 6 arginine residues in proximity to Arg153 from helix 2. The end of helix 4 is flexible. Both the N- and C-terminal parts of HDD form unstructured linker regions, allowing flexibility between the individual structured domains of hDdi2.

We used the Dali server[37] to test whether HDD has structural homology with other known proteins, but surprisingly, we did not detect any significant structural homologs. We were also unable to manually superimpose

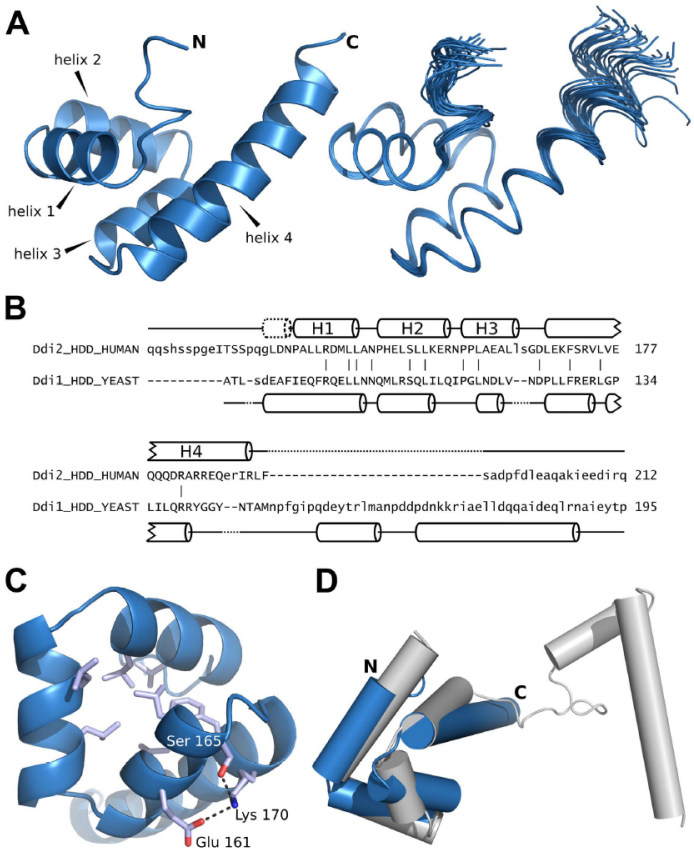


Figure 6. Solution structure of hDdi2 HDD. (**A**) Superimposition of 30 converged structures of HDD. (**B**) Structural alignment of hDdi2 HDD and yDdi1 HDD (PDB code 5KES) analyzed by Dali Pairwise comparison[37]. The Z score for these two structures is 4, and their RMSD is 5 Å. Secondary structures are shown; bars connect identical amino acids. (**C**) Hydrophobic core of the HDD bundle supported by a salt bridge between helix 3 and the initial part of helix 4, including residues Ser165 and Lys170, with occasional contribution from Glu161 (**D**) Superimposition of hDdi2 HDD (blue) with yDdi1 HDD (grey) represented by cylindrical helices. N-terminal parts of both HDDs superimpose with an RMSD of 0.95 Å.

the Sti1-like domain of Rad23 (PDB code 1 × 3W)[38] with our HDD structure, although they show broad similarities. Next, we examined the structural homology between yDdi1 HDD and hDdi2 HDD, which share 25% sequence identity[39]. As shown in Fig. 6D, yDdi1 HDD forms two independent subdomains connected by a flexible linker[23]. Superimposition of the N-terminal "bundle" region of both HDDs (hDdi2 HDD residues 116–178, yDdi1 HDD residues 86–134) yielded an RMSD of 0.95 Å (Fig. 6D), whereas the RMSD calculation for the full-length structures expectedly yielded a high number (3.55 Å). This led us to hypothesize that the two-domain architecture of yeast HDD is in human HDD compacted into a single bundle with an extremely long final helix. We conclude that the hDdi2 HDD possesses a novel α-helical architecture.

The human Ddi2 RVP domain adopts an aspartic protease-like structure. Next, we determined the crystal structure of the hDdi2 RVP domain (Ddi2 212–360) at 1.9 Å resolution (Fig. 7 and Table S2). The structure was solved by molecular replacement using PDB 2I1A as a starting model and refined to an R_{work}/R_{free} of 20.8/21.6%[22]. Comparison of the hDdi2 RVP structure with the previously reported yDdi1 RVP structure revealed conservation of the overall fold (Fig. 7A,B) and active site (Fig. 7E,F)[22]. Similar to yDdi1 RVP, hDdi2

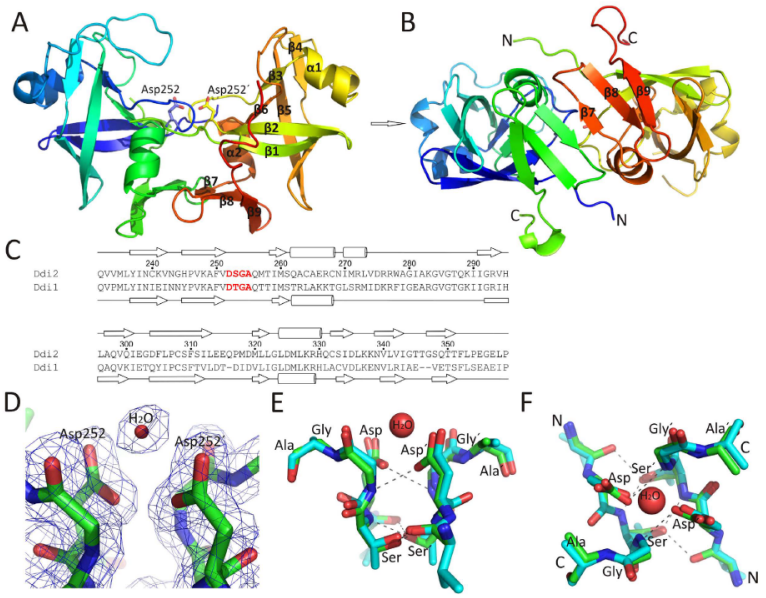


Figure 7. X-ray structure of the hDdi2 RVP domain. (**A**) A ribbon diagram of the structure of the hDdi2 RVP (residues 212–360) dimer (blue N-terminus to red C-terminus). The aspartate side chains that form the putative RVP active site are shown in stick representation. Secondary structure elements are labeled. (**B**) Second view of the RVP dimer related to A) by a 90° rotation about the horizontal axis. C- and N-termini, as well as secondary structure elements of the β-sheet platform, are highlighted. (**C**) Sequence alignment between the hDdi2 and yDdi1 RVP (PDB 2I1A)[22] domains spanning residues from Gln232 to Pro359 of Ddi2, which are visible in the structure. Secondary structure elements are indicated, with arrows representing β-strands and cylinders representing α-helices of the hDdi2 RVP structure (above the sequence) and yDdi1 RVP (below). The putative active site of both RVP domains is highlighted in red. (**D**) The putative active site of the hDdi2 RVP domain showing catalytic aspartates and a water molecule, with the calculated omit map contoured at 1.0 σ. (**E**) The same section of the hDdi2 RVP (in green) shown in (**D**) superposed with the yDdi1 RVP domain[23] (in blue). The hydrogen bonding pattern forming the "fireman's grip" is indicated with dotted gray lines. (**F**) The same section shown in (**E**) rotated by 90° about the horizontal axis. C- and N-termini are indicated.

RVP comprises a six-stranded β-barrel, three β-sheet dimerization platform, and two helices, with the latter quite atypical for retroviral proteases. The second helix precedes the loop that corresponds to the flap region characteristic of other retroviral proteases. The flap in our hDdi2 RVP structure covers the active site only to a certain extent and cannot form hydrogen bonds with the second flap loop, unlike, for example, the structure of HIV-1 protease. The substrate cavity is thus significantly larger than those of other retroviral proteases and potentially could even accommodate small proteins, as observed previously in the yeast Ddi1 RVP[22].

The putative catalytic cavity is formed by the typical amino acid signature of aspartic proteases (Asp-Ser-Gly-Ala). In yDdi1 RVP, Thr is present in place of Ser in the tetrapeptide. The RMSD for all atoms that form the Asp-Ser/Thr-Gly-Ala motif in the hDdi2 RVP and yDdi1 RVP structures is 0.353 Å. The RMSD calculated for the same monomer is 0.219 Å. Both values indicate perfect superposition of the active sites. Similar to other aspartic proteases, in hDdi2 RVP the putative catalytic Asp252 points to the area between the two β-barrel lobes. The residue following Asp252 is Ser, the side chain hydroxyl group of which participates in the "fireman's grip" by hydrogen bonding to the backbone amide group of Ser253´ across the dimer interface and to the backbone carbonyl group of Val251´ (Fig. 7E,F). In agreement with structures of other aspartic proteases, we found a catalytic water molecule within hydrogen bonding distance of the Asp dyad. In summary, the geometry of the hDdi2 RVP domain structure corresponds to that of other catalytically active aspartic proteases, although the catalytic cavity seems to be more open and could possibly accommodate larger substrates.

Small-angle X-ray scattering reveals that Ddi2 adopts an extended dimeric structure. To further inspect the overall shape of hDdi2, we used small-angle X-ray scattering (SAXS) to evaluate the molecular weight, radius of gyration, and low-resolution structure of the HDD-RVP domains of hDdi2. The SAXS invariant

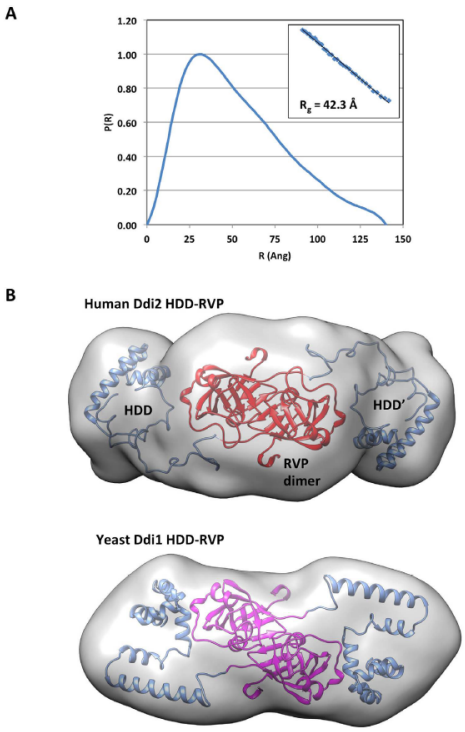


Figure 8. SAXS analysis of the HDD-RVP domains of hDdi2. (**A**) Pair-distance distribution from merged SAXS data, showing the asymmetric distribution characteristic of elongated structures. The inset shows the linearity of the Guinier plot for data collected at 5 mg/ml, indicating monodispersity. (**B**) Modeling of the HDD-RVP structure using the program BUNCH. Twenty models were superposed, averaged and converted to a map for surface visualization in Chimera (top). The structure of the HDD and RVP domains are displayed in blue and red, respectively for the two symmetry-related chains. The structure of the HDD-RVP module from yeast Ddi1 in showed at the bottom for comparison (back-to-back paper for details[23]).

V_c was used to calculate a molecular mass of 66 kDa, which corresponds to the expected dimer mass (monomer: 30 kDa). The large R_g value of 42 Å and the *P(r)* distribution suggest an elongated structure (Fig. 8A). Modeling of the dimeric structure using the crystal structure of the RVP domain and NMR structure of HDD revealed that the HDD extends on either side of the RVP, similar to the yDdi1 HDD-RVP model with a slightly larger D_{max} of 140 Å (Fig. 7B). The overall larger dimensions of the hDdi2 HDD-RVP module arise from the longer flexible linker between the HDD N-terminal bundle and the RVP (40 residues), which in yeast Ddi1 is a more rigid two-helix segment connected by only 9 residues to the RVP. In hDdi2, the longer linker allows for the HDD bundle to extend further and adopt greater range conformations, which increases D_{max} and R_g. Overall, the SAXS data confirmed the dimeric nature of hDdi2 in solution and the conserved structure of the HDD-RVP module between yeast and human Ddi1-like proteins.

Search for putative proteolytic activity and small-molecule binder of the RVP domain. To shed light on the putative proteolytic activity of RVP, we performed PICS with full-length hDdi2 expressed in bacterial and mammalian expression systems[40]. In both cases, the cleavage experiment was performed with a mammalian-cell-derived peptide library prepared using trypsin and GluC digestion. We analyzed the cleavage profile of full-length hDdi2 at pH 4.0, 5.0, and 7.0 with 300 mM NaCl. As negative controls, we used hDdi2 with a D252A mutation in the putative catalytic site and a mock reaction with buffer instead of enzyme. As a positive control, we tested the HIV-1 protease cleavage profile in 100 mM Na acetate, 300 mM NaCl, pH 4.7, using wild-type enzyme and the catalytically inactive D25N mutant with a 1:200 protease-to-library ratio. To our surprise, the data analysis showed no cleavage related to hDdi2 (Figure S6).

Driven by this finding, we subjected the hDdi2 RVP domain to a similar enzymatic analysis as previously reported by Perteguer and co-workers, who showed BSA and HIV-peptide-derived substrate cleavage by leishmanial Ddi1 in acidic conditions[25]. We therefore tested BSA, HSA, β-casein, insulin, and a complete set of HIV-polyprotein-derived peptide substrates for putative hydrolysis by hDdi2 RVP at pH 5.0 and pH 7.0 in various salt concentrations (150 to 500 mM NaCl) by HPLC assay. Again, we did not observe any cleavage (Figures S7–14). ITC further demonstrated that HIV protease inhibitors (saquinavir, ritonavir, indinavir, nelfinavir, amprenavir, darunavir, GS-8374, atazanavir, brecanavir, and acetyl-pepstatin) do not bind to the hDdi2 RVP domain (Figure S15). Thus, we hypothesized that hDdi2 is either catalytically inactive or requires some stimulus or protein partner for its activation.

Discussion

We report here the first structural and functional analysis of mammalian Ddi1-like protein, human Ddi2. The Ddi1-like protein family is intricately connected to the UBQ-proteasome pathway, as its UBL domain interacts both with the proteasome and UBQ and its UBA interacts with UBQ and UBQ-chains[11,13,24]. Based on sequence analysis and genomic organization, we suggest, that hDdi2 is the original version of yDdi1 and non-mammalian orthologs of Ddi1-like proteins. Strikingly, hDdi2 differs from yDdi1 on several levels. One obvious difference is the loss of the UBA domain at the hDdi2 C-terminus. Therefore, we inspected hDdi2 for another potential UBQ-interacting motif (-***L***-*X*-*X*-***A***-*X*-*X*-*X*-***S***-), which we subsequently identified at the C-terminus (-***L***-*A*-*E*-***A***-*L*-*Q*-*K*-***S***-). We applied NMR chemical shift perturbation analysis to reveal that UBQ binds to hDdi2 C-terminal UIM specifically, but with a K_d of 2.2–3.2 mM. It will be interesting to explore whether such binding has any physiological relevance.

Recent work by Nowicka and co-workers showed that the yDdi1 UBL domain can bind UBQ[13]. This surprising feature completely changed our view of the Ddi1-like protein acting as a classical shuttle, suggesting that it may have an alternative mechanism. Therefore, we inspected hDdi2 UBL for its structural and functional properties. Our NMR structure of hDdi2 UBL indicates that unlike the positively charged β-sheet interaction area of yDdi1 UBL, which is complementary to the UBQ patch, the hDdi2 UBL has a small hydrophobic area that is moderately charged. Due to dissimilar charge distribution on the interaction patch, the pattern of interaction partners might differ. This assumption supports the NMR CSP analysis of hDdi2 UBL and UBQ, which shows weak but specific interaction between these two proteins (K_d of 0.42–1.1 mM).

Prompted by the above findings, we subjected hDdi2 to pull-down experiments with all eight native di-ubiquitin conjugates. We assumed that, if the observed weak hDdi2-UBQ affinity has any significance within the cell, an increase in affinity towards some of the UBQ chains would be observed. Notably, neither FLAG-tagged hDdi2 nor HA-tagged hDdi2 were able to pull down any di-UBQs. These results indicate significant differences between hDdi2 and yDdi1.

Yet another interesting feature of all Ddi1-like proteins is the presence of a highly conserved RVP domain, the function of which is largely unresolved. We solved the X-ray structure of hDdi2 RVP and compared it with yDdi1 RVP. As expected, both RVPs are structurally almost identical and quite similar to HIV-1 protease. The structural conservation of the catalytic residues indicates that it could be proteolytically active, although the catalytic cavity is significantly larger than those of other retroviral proteases and might accommodate even small proteins. While some work indicates that leishmanial Ddi1 is catalytically active at acidic pH and cleaves HIV substrates and BSA, we could not confirm these findings with hDdi2 using an HPLC-based method (see Figure S11). Moreover, we did not detect any putative proteolytic activity of hDdi2 with peptide-derived HIV-1 substrates and other proteins. In addition, PICS with an HEK293-derived peptide library revealed no cleavage connected to hDdi2. We also found that no HIV protease inhibitors bind to the RVP domain, as monitored by ITC. From these data, we infer that the RVP domain of hDdi2 likely does not possess intrinsic proteolytic activity. On the other hand, recent data suggests a potential hydrolytic function of RVP that is important for *Drosophila* development and is dependent on intact RVP[20]. That led us to hypothesize that the hDdi2 RVP domain may become catalytically active in more complex arrangement with yet to be identified protein partner.

The identification of the hDdi2 HDD domain goes in line with our hypothesis. This helical arrangement precedes RVP in most Ddi1-like orthologs, suggesting its functional importance. We determined the solution structure of hDdi2 HDD. It consists of a globular arrangement of 4 α-helices and shares broad similarities with the Sti1-like domain of Rad23, which is not structurally similar to any other known protein. All helices pack against each other and form a compact bundle with a hydrophobic core. This bundle superimposes well with the N-terminal part of yDdi1 HDD (identification and structurally characterization of which are described in our back-to-back publication[23]), which may suggest a similar function. Whether HDD could act as an interaction platform for an RVP substrate remains to be determined.

Overall, we present the first detailed study of hDdi2. We determined the 3D structures of all individual protein domains, including the previously unknown helical domain of hDdi2 (HDD). We also identified a novel UBQ-interacting motif (UIM) at the C-terminus of hDdi2. Furthermore, we show that the *in vitro* binding of mono-UBQ to its cognate domains is very weak but specific. We did not observe any binding of any native di-ubiquitin conjugates, which makes hDdi2 unique and diverse from yDdi1. Moreover, we thoroughly studied the RVP domain of hDdi2, solved its 3D structure by protein crystallography, and showed that it is homologous to yDdi1 RVP and HIV-1 protease. It remains to be determined whether RVP processes any substrates in a cellular context, perhaps after activation by a yet-to-be-identified stimulus or protein partner, or whether it exerts a different structural or functional role not directly linked to peptide bond hydrolysis.

Methods

Protein expression and purification. All proteins, including full-length hDdi2 and its truncated forms (UBL, residues 1–76; HDD, residues 116–212; RVP, residues 212–360; HDD-RVP, residues 116–360; RVP full-C,

residues 212–399; and hDdi2 ΔUIM, residues 1–360), human ubiquitin, and Nedd8, were cloned into the vector pET16b (Novagen) in-frame with an N-terminal histidine tag (Fig. 1B). HDD was expressed in fusion with SUMO at the N-terminus. All constructs were expressed in *E. coli* BL21(DE3)RIL host cells; subsequently resuspended in buffer containing 50 mM Tris-HCl, pH 8.0, 50 mM NaCl, and 1 mM EDTA; and lysed by three passages through an EmulsiFlex-C3 high pressure homogenizer (Avestin, Canada) at 1200 bar. Proteins were purified using nickel affinity chromatography and eluted with 250 mM imidazole. Proteins were then dialyzed overnight into 50 mM HEPES, pH 7.4, 150 mM NaCl, and 10% glycerol and applied onto a Superdex 75 or 200 16/60 gel filtration column (GE Healthcare), depending on the protein mass. Individual fractions were analyzed by SDS-PAGE and/or Western blot.

For NMR experiments, hDdi2 UBL, hDdi2 HDD, hDdi2 RVP full-C, and human ubiquitin were expressed as ^{15}N- and ^{15}N/^{13}C-labeled proteins; Nedd8 was expressed as an ^{15}N-labeled protein. Cells were grown in minimal medium containing 0.8 g/l [^{15}N]ammonium chloride and 2 g/l d-[^{13}C]glucose, as required. Further procedures were the same as mentioned above, except the size-exclusion chromatography was carried out in buffers used for NMR titrations.

Mammalian-expressed protein immobilization for PICS assay. For PICS proteolytic activity experiments and pull-downs, DNA encoding both N- and C-terminally FLAG-tagged full-length hDdi2 were cloned into the pTRE-Tight vector, and the constructs were transfected into HEK293A2 cells grown on DMEM media, using lipofectamine to produce a stable transfected cell line. Clones with a high level of FLAG-hDdi2 expression were selected by Western blot. Cells from ten 100-mm cell culture dishes were harvested by washing into PBS followed by centrifugation (2 min, 225 g, RT) and washed 3x with PBS. Cells were resuspended in ice cold lysis buffer (50 mM HEPES. pH 7.8, 150 mM NaCl, 0.4% Igepal CA-630) and lysed on ice using 3 freeze/thaw cycles on dry ice, each followed by repeated aspiration of the cell suspension with a 30-gauge needle. The cell lysate was diluted 4x with lysis buffer without Igepal and cleared by centrifugation (15 min, 20,000 g, 4 °C). Supernatant was loaded on M2 anti-FLAG magnetic beads (Sigma-Aldrich) in batch format according to the manufacturer's recommendation. After a 1-h equilibration, beads were washed 4 times with PBS. The purification process and the final amount and purity of protein immobilized on magnetic beads were monitored by SDS-PAGE and Western blot. FLAG-tagged hDdi2 immobilized on magnetic beads was subsequently used for PICS experiments. As control samples, an identical amount of magnetic beads was incubated with an equal (in protein mass) amount of cell lysate from non-transfected cells and processed the same way.

Pull-down experiments. Beads with approximately 3 μg of hDdi2 immobilized *via* FLAG-tag on either the N- or C-terminus were equilibrated with UBQ-binding buffer (PBS, pH 7.4, 1% Triton X-100, 5 mM EDTA, 0.2 mg/ml BSA) and mixed with 1 μg of di-ubiquitin conjugate (UbiQ) of given linkage type in a total volume of 50 μl of the same buffer. The final mixture was incubated for 2 h at room temperature with mild agitation. Beads were washed twice with 150 μl and 100 μl of TBS, and bound proteins were eluted by heating to 95 °C for 3 min in 5 μl of 2x non-reducing SDS-PAGE sample buffer (125 mM Tris, pH 6.8, 4% SDS, 20% (v/v) glycerol, 0.004% bromphenol blue). The whole eluted fraction was separated by 18% Tris-glycine SDS-PAGE and blotted onto a PVDF membrane. The membrane was denatured (6 M guanidium chloride, 20 mM Tris, pH 7.5, 1 mM PMSF, 5 mM β-mercaptoethanol) and developed using anti-ubiquitin rabbit polyclonal antibody (Dako). Experiments were performed with di-ubiquitins of all eight native linkage types. In addition, potential binding was tested also with Lys48- and Lys63-linked chains synthetized in house according to Pickart and co-workers[41]. Negative controls with either no immobilized hDdi2 protein or without loaded di-ubiquitin were treated the same way.

X-ray crystallography. Crystals of hDdi2 RVP were grown by the hanging drop vapor diffusion technique at 19 °C with 0.2 M ammonium acetate, 0.1 M Bis-Tris, pH 5.5, and 25% PEG 3350 as precipitant. For cryoprotection, crystals were soaked in the reservoir solution supplemented with 25% (v/v) glycerol. Diffraction data were collected at 100 K at BESSY beamline 14.2 at the Hemholtz Zentrum Berlin, Germany[42]. Data were integrated using Mosflm v7.0.6 and later scaled with SCALA v3.3.20[43,44]. The crystal structure was solved by molecular replacement using the program Molrep and the structure of yDdi1 RVP (PDB code 2I1A) as a template[22,45]. Model refinement was carried out with REFMAC 5.6 from the CCP4 package[46,47], interspersed with manual adjustments using Coot[48]. Atomic coordinates and experimental structure factors have been deposited in the Protein Data Bank under the code 4RGH. Data collection and refinement statistics are given in Table S2.

Peptide synthesis. The UIM peptides (hDdi2 C-terminus-derived UIM of amino acid sequence EEIADQELAEALQKSAEDAE and its scrambled version AELEQIAEDALEKEDSQEAA) were synthetized on an ABI 433A solid phase synthesizer (Applied Biosystems, USA) at the peptide synthesis core facility of IOCB, Czech Republic. They were further purified in the form of C-terminal amides by reverse-phase high-performance liquid chromatography (HPLC) on a semipreparative C18 column (Labio a.s., Prague, Czech Republic). Purified fractions were frozen in liquid nitrogen, lyophilized, and dissolved in DMSO prior to further use.

Nuclear magnetic resonance spectroscopy. NMR spectra for interaction site identification were acquired from 350 μl samples of 0.1 mM (peptide binding) or 0.05 mM (protein – protein interaction) ^{15}N-labeled hDdi2 UBL and hDdi2 RVP full-C in 50 mM sodium phosphate buffer, pH 7.4, and from 0.1 or 0.05 mM UBQ in 50 mM sodium phosphate buffer, pH 6.0 and pH 7.4. All buffers contained 5% D_2O/95% H_2O. Spectra for structural determination and backbone assignments were acquired at 0.5 mM concentration of ^{13}C/^{15}N-labeled proteins. NMR data were collected at 25 °C on 600 and 850 MHz Bruker Avance spectrometers equipped with triple resonance (^{15}N/^{13}C/^{1}H) cryoprobes. Resonance assignments were obtained using a previously published approach[49,50]. Detailed experimental procedures for all the NMR measurements, structure calculations, and chemical shift mapping are described in the Supplementary Information.

Small-angle X-ray scattering. The His-tagged HDD-RVP construct of hDdi2 (residues 116–360) was purified and concentrated in SAXS buffer (25 mM Tris/HCl, 75 mM NaCl, 5% glycerol, 1 mM DTT, pH 7.4). A series of dilutions (10, 5, and 2.5 mg/ml) and buffer alone were frozen and shipped to the SIBYLS facility at the Advanced Light Source (ALS) for automated SAXS analysis as described[51]. SAXS data were acquired for 0.5, 1, 2, and 4 sec for each sample. Due to a slight concentration-dependent effect in the low-q region, the data at 10 mg/ml were discarded. The 5 and 2.5 mg/ml data were merged for data analysis using the ATSAS software suite[52]. The molecular weight was calculated using the *Qr* method as described[53]. BUNCH software was used for modeling, using the crystal structure of the RVP domain (residues 231–360) with a fixed P2 symmetry axis and the NMR structure of the HDD domain (residues 131–190). Twenty models were calculated with χ^2 fit to experimental data ranging between 1.64 and 2.39, and averaged using DAMAVER. The resulting bead model was converted into a volumetric map using the program SITUS and visualized in Chimera[54,55].

Isothermal titration calorimetry (ITC). The ability of hDdi2 RVP to bind HIV-1 protease inhibitors was analyzed at 25 °C using a high-throughput screening Auto-iTC$_{200}$ system (MicroCal, GE Healthcare Life Sciences). Aliquots (2 μl) of 120 μM protease inhibitors (saquinavir, ritonavir, indinavir, nelfinavir, amprenavir, darunavir, GS-8374, atazanavir, brecanavir, and acetyl-pepstatin) were injected stepwise into a sample cell containing 200 μl of 10 μM hDdi2 RVP (concentration calculated based on the molecular weight of the dimer; HPLC amino acid analysis was performed). The titrations were monitored by MicroCal software implemented in Origin 7.0 (MicroCal, GE Healthcare Life Sciences).

PICS assay. A HEK293-cell-derived peptide library for PICS experiments was prepared as described by Schilling *et al.*[56]. Isolated denatured proteins were cleaved into peptides using trypsin (Sigma Aldrich) and GluC as working proteases. After abolishing the working protease activity using PMSF, a second round of sulfhydryl reduction and alkylation was performed, and primary amines on peptide N-termini and lysine side chains were blocked using formaldehyde-cyanoborohydride reductive dimethylation. Excess modification reagents were removed by gel filtration, and the peptide library was purified and transferred to HPLC grade water using a Sep-Pak Plus C-18 solid phase extraction cartridge (Waters), following the manufacturer's protocol. The peptide concentration in the library was adjusted to 2 mg/ml. The integrity of the peptide library was confirmed by LC-MS/MS analysis. The final amine-protected mammalian proteome-derived peptide library was stored in aliquots at −80 °C until further use.

For the endopeptidase assay, peptide library (1 mg/ml) was incubated in 200 μl of 100 mM sodium acetate, 300 mM NaCl, pH 4.0, with 4 μg of recombinant full-length hDdi2. Reactions were incubated for 12 h at 37 °C, then heat-inactivated for 30 min at 70 °C and transferred to 200 mM HEPES, pH 8.0, using a Sep-Pak Light C-18 solid phase extraction cartridge (Waters), following the manufacturer's protocol.

Subsequently, newly formed peptide free N-termini (products of proteolytic cleavage) were biotinylated *in vitro* by incubation with 350 μM sulfo-NHS-SS-biotin (Thermo-Scientific) for 4 h at room temperature. Biotinylated products were then immobilized on streptavidin agarose (Solulink) by 2 h incubation with mild agitation at room temperature, followed by washing. Additional washing steps (2 M urea followed by 20% isopropanol, 5% DMSO, and finally 5% acetonitrile, all in washing buffer [50 mM HEPES, pH 7.5, 150 mM NaCl]) were added into the protocol, followed by ten washes with washing buffer alone. Immobilized peptides were eluted with 20 mM DTT and desalted using Pepclean C-18 reverse phase cartridges (Thermo Scientific), following the manufacturer's protocol, and analyzed by mass spectrometry.

As negative controls, we used D252A hDdi2 and a mock reaction with buffer added instead of enzyme. As a positive control, the HIV-1 protease cleavage profile in 100 mM Na acetate, 300 mM NaCl, pH 4.7, was tested using wild-type HIV-1 protease and a catalytically inactive mutant (D25N) in a 1:200 protease-to-library ratio. The proteolytic cleavage assay was carried out in 100 mM sodium acetate, 300 mM NaCl, pH 5.0, and 100 mM HEPES, 300 mM NaCl, pH 7.0, with processing and control reactions as described above.

Eukaryotic-expressed hDdi2 was also tested in the PICS assay. For those experiments, magnetic beads with immobilized FLAG-tagged hDdi2 in an amount corresponding to approximately 1 μg of immobilized protein (based on Western blot) were mixed with 200 μg of the peptide library. After a 12-h incubation at 37 °C, beads were magnetically removed, residual protein was heat-inactivated, and the sample was further processed as described above. This assay was carried out under three different buffer conditions (100 mM sodium acetate, 300 mM NaCl, pH 4.0; 100 mM sodium acetate, 300 mM NaCl, pH 5.0; and 100 mM HEPES, 300 mM NaCl, pH 7.0). As a control, an identical amount of magnetic beads incubated with the lysate of nontransfected cells was used.

Data analysis of the PICS assay. Data were analyzed by a series of predesigned queries in Microsoft Access database management software. First, lists of identified peptides from each MS run were loaded to the database and filtered for peptides containing products of N-terminal modification by biotinylation. Second, peptides with over 80% confidence were picked for the enzyme tested (hDdi2 or HIV-1 protease), while peptides with over 10% confidence were picked for control reactions (catalytically inactive mutants and mock reactions). To properly subtract the background signal, the list of peptides identified in the tested enzyme reaction was screened for peptides presented in the mock reaction as well as in the reaction with catalytically inactive enzyme (hDdi2 [D252A] or HIV-1 protease [D25N]), and those peptides were removed from processing. Finally, the tested enzyme reactions were screened for peptides identified in the original unprocessed peptide library. Such peptides were then removed from the analysis.

The final cleared list of identified peptides was then mapped on the FASTA database used for proteomics database searching. By alignment of identified peptides with the database, the N-terminal portions of cleaved peptides (preceding the cleavage site) were determined. If multiple computationally identified preceding sequences were

found for one identified peptide, they were removed from processing, while the identified peptide sequences were kept in the list for downstream analysis. The final list of substrate peptides containing sequences of four P-prime amino acids identified by MS and four P amino acids identified computationally was then created. The frequency of each amino acid in each particular position was calculated and plotted, yielding the substrate specificity matrix for the enzyme studied.

HPLC analysis. The hydrolysis of peptides corresponding to the HIV-1 Gag and Gag-Pol processing sites was performed in 50 mM sodium acetate, pH 5.0, containing 0.5 M NaCl and 2 mM EDTA, using 200 μM peptide and 75 nM hDdi2 RVP expressed in a prokaryotic system. Additionally, cleavage of 5 μM proteins (bovine serum albumin, human serum albumin, bovine casein, and bovine insulin) in 100 mM sodium acetate, pH 5.0, 1 M NaCl, and 4 mM EDTA using 200 nM hDdi2 RVP was monitored. The reaction mixture was incubated at 37 °C for 24 h. The reactions were stopped by addition of formic acid to a final concentration of 20% (v/v). Aliquots (5 μl) of the reaction mixtures were subsequently injected into a Zorbac SB-C18 reversed-phase chromatography column (Agilent), and peptides were resolved using a water-acetonitrile gradient on a high-performance liquid chromatograph (HPLC) (Agilent). The peptide cleavage was monitored at 220 nm.

References

1. Hershko, A. & Ciechanover, A. The ubiquitin system. *Annu Rev Biochem* **67**, 425–479 (1998).
2. Ciechanover, A. The ubiquitin-proteasome proteolytic pathway. *Cell* **79**, 13–21 (1994).
3. Clarke, D. J. *et al.* Dosage suppressors of pds1 implicate ubiquitin-associated domains in checkpoint control. *Mol Cell Biol* **21**, 1997–2007 (2001).
4. Elsasser, S. & Finley, D. Delivery of ubiquitinated substrates to protein-unfolding machines. *Nat Cell Biol* **7**, 742–749 (2005).
5. Elsasser, S. *et al.* Proteasome subunit Rpn1 binds ubiquitin-like protein domains. *Nat Cell Biol* **4**, 725–730 (2002).
6. Gomez, T. A., Kolawa, N., Gee, M., Sweredoski, M. J. & Deshaies, R. J. Identification of a functional docking site in the Rpn1 LRR domain for the UBA-UBL domain protein Ddi1. *Bmc Biol* **9** (2011).
7. Saeki, Y., Saitoh, A., Toh-e, A. & Yokosawa, H. Ubiquitin-like proteins and Rpn10 play cooperative roles in ubiquitin-dependent proteolysis. *Biochem Bioph Res Co* **293**, 986–992 (2002).
8. Husnjak, K. *et al.* Proteasome subunit Rpn13 is a novel ubiquitin receptor. *Nature* **453**, 481–488 (2008).
9. Kaplun, L. *et al.* The DNA damage-inducible UbL-UbA protein Ddi1 participates in Mec1-mediated degradation of Ho endonuclease. *Mol Cell Biol* **25**, 5355–5362 (2005).
10. Bertolaet, B. L. *et al.* UBA domains of DNA damage-inducible proteins interact with ubiquitin. *Nat Struct Biol* **8**, 417–422 (2001).
11. Gabriely, G., Kama, R., Gelin-Licht, R. & Gerst, J. E. Different domains of the UBL-UBA ubiquitin receptor, Ddi1/Vsm1, are involved in its multiple cellular roles. *Mol Biol Cell* **19**, 3625–3637 (2008).
12. Ivantsiv, Y., Kaplun, L., Tzirkin-Goldin, R., Shabek, N. & Raveh, D. Unique role for the UbL-UbA protein Ddi1 in turnover of SCFUfo complexes. *Mol Cell Biol* **26**, 1579–1588 (2006).
13. Nowicka, U. *et al.* DNA-Damage-Inducible 1 Protein (Ddi1) Contains an Uncharacteristic Ubiquitin-like Domain that Binds Ubiquitin. *Structure* **23**, 542–557 (2015).
14. Krylov, D. M. & Koonin, E. V. A novel family of predicted retroviral-like aspartyl proteases with a possible key role in eukaryotic cell cycle control. *Curr Biol* **11**, R584–R587 (2001).
15. Liu, Y. & Xiao, W. Bidirectional regulation of two DNA-damage-inducible genes, MAG1 and DDI1, from Saccharomyces cerevisiae. *Mol Microbiol* **23**, 777–789 (1997).
16. Diaz-Martinez, L. A., Kang, Y., Walters, K. J. & Clarke, D. J. Yeast UBL-UBA proteins have partially redundant functions in cell cycle control. *Cell Div* **1** (2006).
17. Lustgarten, V. & Gerst, J. E. Yeast VSM1 encodes a v-SNARE binding protein that may act as a negative regulator of constitutive exocytosis. *Mol Cell Biol* **19**, 4480–4494 (1999).
18. Marash, M. & Gerst, J. E. Phosphorylation of the autoinhibitory domain of the Sso t-SNAREs promotes binding of the Vsm1 SNARE regulator in yeast. *Mol Biol Cell* **14**, 3114–3125 (2003).
19. Guthmueller, K. L., Yoder, M. L. & Holgado, A. M. Determining genetic expression profiles in C. elegans using microarray and real-time PCR. *J Vis Exp* (2011).
20. Morawe, T., Honemann-Capito, M., von Stein, W. & Wodarz, A. Loss of the extraproteasomal ubiquitin receptor Rings lost impairs ring canal growth in Drosophila oogenesis. *J Cell Biol* **193**, 71–80 (2011).
21. Franco, M. I., Turin, L., Mershin, A. & Skoulakis, E. M. Molecular vibration-sensing component in Drosophila melanogaster olfaction. *Proc Natl Acad Sci USA* **108**, 3797–3802 (2011).
22. Sirkis, R., Gerst, J. E. & Fass, D. Ddi1, a eukaryotic protein with the retroviral protease fold. *J Mol Biol* **364**, 376–387 (2006).
23. Trempe, J. F. Structural studies of yeast DNA damage-inducible protein (Ddi1) reveal domain architecture of the Ddi eukaryotic protein family (2016).
24. Trempe, J. F. *et al.* Mechanism of Lys48-linked polyubiquitin chain recognition by the Mud1 UBA domain. *Embo J* **24**, 3178–3189 (2005).
25. Perteguer, M. J. *et al.* Ddi1-like protein from Leishmania major is an active aspartyl proteinase. *Cell Stress Chaperones* **18**, 171–181 (2013).
26. White, R. E., Dickinson, J. R., Semple, C. A., Powell, D. J. & Berry, C. The retroviral proteinase active site and the N-terminus of Ddi1 are required for repression of protein secretion. *FEBS Lett* **585**, 139–142 (2011).
27. White, R. E., Powell, D. J. & Berry, C. HIV proteinase inhibitors target the Ddi1-like protein of Leishmania parasites. *FASEB J* **25**, 1729–1736 (2011).
28. Bucher, P., Karplus, K., Moeri, N. & Hofmann, K. A flexible motif search technique based on generalized profiles. *Comput Chem* **20**, 3–23 (1996).
29. Voloshin, O., Bakhrat, A., Herrmann, S. & Raveh, D. Transfer of Ho endonuclease and Ufo1 to the proteasome by the UbL-UbA shuttle protein, Ddi1, analysed by complex formation *in vitro*. *Plos One* **7**, e39210 (2012).
30. Osswald, C. *et al.* Mice without the regulator gene Rsc1A1 exhibit increased Na+-D-glucose cotransport in small intestine and develop obesity. *Mol Cell Biol* **25**, 78–87 (2005).
31. Soding, J. Protein homology detection by HMM-HMM comparison. *Bioinformatics* **21**, 951–960 (2005).
32. Cornilescu, G., Marquardt, J. L., Ottiger, M. & Bax, A. Validation of protein structure from anisotropic carbonyl chemical shifts in a dilute liquid crystalline phase. *J Am Chem Soc* **120**, 6836–6837 (1998).
33. Sloper-Mould, K. E., Jemc, J. C., Pickart, C. M. & Hicke, L. Distinct functional surface regions on ubiquitin. *J Biol Chem* **276**, 30483–30489 (2001).
34. Singh, R. K. *et al.* Recognition and cleavage of related to ubiquitin 1 (Rub1) and Rub1-ubiquitin chains by components of the ubiquitin-proteasome system. *Mol Cell Proteomics* **11**, 1595–1611 (2012).
35. Rice, P., Longden, I. & Bleasby, A. EMBOSS: The European molecular biology open software suite. *Trends Genet* **16**, 276–277 (2000).

36. Ye, Y. Z. & Godzik, A. Flexible structure alignment by chaining aligned fragment pairs allowing twists. *Bioinformatics* **19**, Ii246–Ii255 (2003).
37. Hasegawa, H. & Holm, L. Advances and pitfalls of protein structural alignment. *Curr Opin Struc Biol* **19**, 341–348 (2009).
38. Lee, J. H., Choi, J. M., Lee, C. W., Yi, K. J. & Cho, Y. J. Structure of a peptide : N-glycanase-Rad23 complex: Insight into the deglycosylation for denatured glycoproteins. *P Natl Acad Sci USA* **102**, 9144–9149 (2005).
39. Sievers, F. *et al.* Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol* **7** (2011).
40. Schilling, O. & Overall, C. M. Proteome-derived, database-searchable peptide libraries for identifying protease cleavage sites. *Nat Biotechnol* **26**, 685–694 (2008).
41. Pickart, C. M. & Raasi, S. Controlled synthesis of polyubiquitin chains. *Ubiquitin and Protein Degradation, Pt B* **399**, 21–36 (2005).
42. Mueller, U. *et al.* Facilities for macromolecular crystallography at the Helmholtz-Zentrum Berlin. *J Synchrotron Radiat* **19**, 442–449 (2012).
43. Leslie, A. G. W. & Powell, H. R. Processing diffraction data with MOSFLM. *Nato Sci Ser Ii Math* **245**, 41–51 (2007).
44. Evans, P. Scaling and assessment of data quality. *Acta Crystallogr D* **62**, 72–82 (2006).
45. Vagin, A. & Teplyakov, A. An approach to multi-copy search in molecular replacement. *Acta Crystallogr D* **56**, 1622–1624 (2000).
46. Murshudov, G. N., Vagin, A. A. & Dodson, E. J. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D* **53**, 240–255 (1997).
47. Winn, M. D. *et al.* Overview of the CCP4 suite and current developments. *Acta Crystallogr D* **67**, 235–242 (2011).
48. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr D* **60**, 2126–2132 (2004).
49. Renshaw, P. S. *et al.* Sequence-specific assignment and secondary structure determination of the 195-residue complex formed by the Mycobacterium tuberculosis proteins CFP-10 and ESAT-6. *J Biomol Nmr* **30**, 225–226 (2004).
50. Veverka, V. *et al.* NMR assignment of the mTOR domain responsible for rapamycin binding. *J Biomol Nmr* **36**, 3–3 (2006).
51. Hura, G. L. *et al.* Robust, high-throughput solution structural analyses by small angle X-ray scattering (SAXS). *Nat Methods* **6**, 606–612 (2009).
52. Petoukhov, M. V. *et al.* New developments in the program package for small-angle scattering data analysis. *J Appl Crystallogr* **45**, 342–350 (2012).
53. Rambo, R. P. & Tainer, J. A. Accurate assessment of mass, models and resolution by small-angle scattering. *Nature* **496**, 477–481 (2013).
54. Wriggers, W. Using Situs for the integration of multi-resolution structures. *Biophys Rev* **2**, 21–27 (2010).
55. Pettersen, E. F. *et al.* UCSF Chimera–a visualization system for exploratory research and analysis. *J Comput Chem* **25**, 1605–1612 (2004).
56. Schilling, O., Huesgen, P. F., Barre, O., Keller, U. A. D. & Overall, C. M. Characterization of the prime and non-prime active site specificities of proteases by proteome-derived peptide libraries and tandem mass spectrometry. *Nat Protoc* **6**, 111–120 (2011).
57. Baker, N. A., Sept, D., Joseph, S., Holst, M. J. & McCammon, J. A. Electrostatics of nanosystems: Application to microtubules and the ribosome. *P Natl Acad Sci USA* **98**, 10037–10041 (2001).
58. Dolinsky, T. J., Nielsen, J. E., McCammon, J. A. & Baker, N. A. PDB2PQR: an automated pipeline for the setup of Poisson-Boltzmann electrostatics calculations. *Nucleic Acids Res* **32**, W665–W667 (2004).

Acknowledgements

This work was supported by the Ministry of Education, Youth and Sports of the Czech Republic within the National Sustainability Program II (Project BIOCEV-FAR) LQ1604 and by the project "BIOCEV" (CZ.1.05/1.1.00/02.0109). We thank the Advanced Light Source in Berkeley for access and data collection on the SIBYLS SAXS beamline and the Helmholtz Zentrum Berlin (HZB). This work was also supported by a Center of Excellence project [P208/12/G016] of the Grant Agency of the Czech Republic, as well as the Canada Research Chairs program and Canadian Institutes of Health Research (J.-F.T.). The authors thank Dr. Pavel Otáhal and Prof. Václav Hořejší for providing the initial DNA sample encoding hDdi2.

Author Contributions

M. Sivá, M. Svoboda, J.-F.T., R.H., M.K., F.S., J. Belza, J. Brynda, M.H. and K.G.Š. performed the experiments and analyzed the data. K.H. performed bioinformatic analysis. V.V., M.K., P.Š., J.K. and K.G.Š. analyzed and interpreted the data. J.S. and I.F. provided technical support. K.G.Š., M. Sivá, M. Svoboda, J.-F.T., K.H. and J.K. drafted the manuscript. All authors reviewed the manuscript.

Additional Information

Accession codes: Coordinates and structure factors for the RVP crystal structure were deposited in the PDB under accession code 4RGH. The structure and assigned chemical shifts for the UBL domain of hDdi2 were deposited in the PDB and BMRB databases under accession codes 2N7D and 25801, respectively. The structure and assigned chemicals shifts for the HDD of hDdi2 were deposited in the PDB and BMRB databases under accession codes 5K57 and 30097, respectively.

Supplementary information accompanies this paper at http://www.nature.com/srep

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: Sivá, M. *et al.* Human DNA-Damage-Inducible 2 Protein Is Structurally and Functionally Distinct from Its Yeast Ortholog. *Sci. Rep.* **6**, 30443; doi: 10.1038/srep30443 (2016).

APPENDIX 2

Trempe, J. F., Saskova, K. G., Siva, M., Ratcliffe, C. D., Veverka, V., Hoegl, A., Menade, M., Feng, X., Shenker, S., Svoboda, M., Kozisek, M., Konvalinka, J. and Gehring, K. (2016). "Structural studies of the yeast DNA damage-inducible protein Ddi1 reveal domain architecture of this eukaryotic protein family". *Sci Rep* **6** 33671. **IF (2016) = 4.259**

SCIENTIFIC REPORTS

OPEN

Structural studies of the yeast DNA damage-inducible protein Ddi1 reveal domain architecture of this eukaryotic protein family

Received: 28 June 2015
Accepted: 01 September 2016
Published: 20 September 2016

Jean-François Trempe[1,†], Klára Grantz Šašková[2,3], Monika Sivá[2,3,4], Colin D. H. Ratcliffe[1], Václav Veverka[2], Annabelle Hoegl[1], Marie Ménade[1], Xin Feng[1], Solomon Shenker[1], Michal Svoboda[2,5], Milan Kožíšek[2], Jan Konvalinka[2,3] & Kalle Gehring[1]

The eukaryotic Ddi1 family is defined by a conserved retroviral aspartyl protease-like (RVP) domain found in association with a ubiquitin-like (UBL) domain. Ddi1 from *Saccharomyces cerevisiae* additionally contains a ubiquitin-associated (UBA) domain. The substrate specificity and role of the protease domain in the biological functions of the Ddi family remain unclear. Yeast Ddi1 has been implicated in the regulation of cell cycle progression, DNA-damage repair, and exocytosis. Here, we investigated the multi-domain structure of yeast Ddi1 using X-ray crystallography, nuclear magnetic resonance, and small-angle X-ray scattering. The crystal structure of the RVP domain sheds light on a putative substrate recognition site involving a conserved loop. Isothermal titration calorimetry confirms that both UBL and UBA domains bind ubiquitin, and that Ddi1 binds K48-linked diubiquitin with enhanced affinity. The solution NMR structure of a helical domain that precedes the protease displays tertiary structure similarity to DNA-binding domains from transcription regulators. Our structural studies suggest that the helical domain could serve as a landing platform for substrates in conjunction with attached ubiquitin chains binding to the UBL and UBA domains.

The ubiquitin system is primarily a signaling pathway whereby substrates tagged with various types of ubiquitin chains or ubiquitin-like (UBL) modifiers undergo different fates in the cell[1]. Ubiquitinated substrates are recognized by receptor proteins that contain ubiquitin-binding domains such as ubiquitin-interacting motifs (UIM) and ubiquitin-associated (UBA) domains[2]. In *Saccharomyces cerevisiae*, three ubiquitin receptors (Ddi1, Rad23, and Dsk2) have C-terminal UBA domains that bind ubiquitin and Lys48-linked polyubiquitin[3–6]. These proteins also bear an N-terminal UBL domain that binds Rpn1 in the 19S proteasome subunit[7–10]. Ddi1 and Rad23 are DNA-damage inducible proteins, and both have been shown to suppress the temperature sensitivity of a *pds1* mutant[11]. The protein Pds1 (securin) is a mitotic checkpoint control protein, and its ubiquitination by the anaphase-promoting complex (APC) and subsequent degradation is required for the separation of sister chromatids. The triple-deletion mutant Δ*ddi1*Δ*rad23*Δ*dsk2* shows a synthetic effect and delays in the onset of G2/M phase and anaphase, suggesting redundant roles in cell cycle progression[12].

Over the last ten years, the biology of yeast Ddi1 has been investigated from different perspectives. The expression of the *DDI1* gene is controlled by a bidirectional DNA-damage inducible promoter that divergently transcribes *DDI1* and *MAG1*, a 3-methyladenine DNA glycosylase involved in a base-excision-repair pathway[13,14]. These two genes are differentially regulated in response to different DNA-damage checkpoint pathways[15–17].

[1]Groupe de Recherche Axé sur la Structure des Protéines, Department of Biochemistry, McGill University, 3649 Promenade Sir William Osler, Montreal, QC, H3G 0B1, Canada. [2]Gilead Sciences and IOCB Research Center, Institute of Organic Chemistry and Biochemistry of the Academy of Sciences of the Czech Republic, Flemingovo n. 2, 166 10 Prague 6, Czech Republic. [3]Department of Biochemistry, Faculty of Science, Charles University, Hlavova 8, 120 00 Prague 2, Czech Republic. [4]First Faculty of Medicine, Charles University in Prague, Katerinska 32, 121 08, Prague 2, Czech Republic. [5]Department of Physical and Macromolecular Chemistry, Faculty of Science, Charles University, Hlavova 8, 120 00 Prague 2, Czech Republic. [†]Present address: Department of Pharmacology & Therapeutics, McGill University, 3655 Promenade Sir William Osler, Montreal, QC, H3G 1Y6, Canada. Correspondence and requests for materials should be addressed to J.-F.T. (email: jeanfrancois.trempe@mcgill.ca)

Strong expression of *MAG1* and *DDI1* can thus be induced by the addition of methyl methane-sulfonate to yeast cells, which triggers the *CHK1*- and *MEC1*-dependent DNA-damage response (DDR) pathways. Recent studies also indicate a possible role for Ddi1 in degradation of the Ho endonuclease, the enzyme responsible for switching alleles at the mating type locus *MAT*[18]. Activation of the *MEC1*-dependent DDR pathway leads to the phosphorylation and rapid degradation of the Ho protein by the ubiquitin-proteasome system[19]. Phosphorylated nuclear Ho is exported to the cytoplasm *via* the Msn5 nuclear exportin and ubiquitinated by the SCF^{Ufo1} E3 ligase complex[20]. Interestingly, Ho accumulates in Δ*ddi1* cells, but not in Δ*rad23* or Δ*dsk2* cells[18]. This specificity was attributed to specific interactions between the UBL domain of Ddi1 and four tandem UIMs located at the C-terminus of Ufo1[21]. Ufo1 binds phosphorylated Ho through its F-box domain to mediate SCF-dependent ubiquitination[18].

Ddi1 (also known as Vsm1 from v-SNARE-master 1) was independently identified as a SNARE-interacting protein in a yeast two-hybrid screen using the endocytic Snc2 protein as a bait[22]. Ddi1 interacts with both exo- and endocytic v-SNARE proteins (Snc1 and Snc2). Overexpression of Ddi1 in yeast bearing a mutation in the *sec9* gene (t-SNARE) inhibits protein secretion, suggesting that Ddi1 is a negative regulator of exocytosis. It was later shown that Ddi1 binds to the exocytic t-SNARE Sso1, which precludes binding of Sso1 to its functional partner Sec9 and thus inhibits exocytosis[23]. Binding of Ddi1 to Sso1 is promoted by phosphorylation of the N-terminal autoinhibitory domain of Sso1. The interaction is mediated by a linker region of Ddi1 located between the protease and UBA domains. The linker includes a PEST motif and a phosphorylation site (T348) that also regulates exocytosis[24]. Consistent with these findings, Δ*ddi1* yeast cells show increased global protein secretion[23,25]. Ddi1 is also required for endocytosis of the guanine nucleotide-binding protein Gα[26]. Overall, these various studies point towards a role for Ddi1 in cell cycle and growth control, as well as protein trafficking.

Yeast Ddi1 has three structural domains. It has an N-terminal UBL domain that shares only 14% sequence identity with ubiquitin. Its central retroviral protease-like domain (RVP), which is common to all eukaryotic Ddi1 orthologs, is homologous to retroviral aspartic proteases[27]. The active site aspartate is required for repression of protein secretion in yeast[25], and this phenotype can be inhibited by HIV protease inhibitors[28], strongly suggesting that this function of Ddi1 is linked to its protease activity. The Ddi1-like protein from *Leishmania major* displays proteolytic activity at acidic pH[29], suggesting that the protein may be active only in acidic vesicular compartments. The three-dimensional structure of the isolated protease domain confirms that the domain adopts the typical aspartyl protease fold, with a two-fold dyad symmetry that allows Asp220 from two subunits to form hydrogen bonds with a catalytic water molecule[30]. Yeast Ddi1 also bears a C-terminal UBA domain that is found only in plants, fungi, and invertebrates, but not in vertebrate Ddi1 orthologs[3–5]. The UBA domain from the *Schizosaccharomyces pombe* ortholog Mud1 binds selectively to K48-linked diubiquitin (Ub_2) through two ubiquitin-binding sites[5]. In spite of all these studies, the overall function and substrate(s) of Ddi1 protease domain remain elusive.

Here, we report findings from structural and functional studies of full-length Ddi1 from *S. cerevisiae*. We determined a new crystal structure of the RVP domain of Ddi1 that provides insight into its putative substrate recognition mechanism. We determined the solution structure of the UBL domain by NMR spectroscopy and performed interaction studies with different proposed ligands. We found that UBA and UBL both bind ubiquitin and that Ddi1 binds K48-linked diubiquitin with enhanced affinity. We also determined the structure of a new α-helical domain (named HDD) that precedes the RVP domain and could play a role in substrate recognition. We used SAXS to investigate the structure and dynamics of the module formed by the UBL, HDD and RVP domains. Finally, we performed Proteomic Identification of protease Cleavage Sites (PICS) analysis with full-length Ddi1 at both acidic and neutral pH to explore substrate specificity[31].

Results

Crystal structure of the Ddi1 retroviral protease-like domain reveals a potential substrate-binding loop. The most conserved segment of the Ddi1 family is its RVP domain, which probably defines its biochemical function. While the structure of the RVP had previously been determined[30], the substrate-binding mode remained unknown. We obtained a new crystal structure of the RVP domain of Ddi1 (residues 185–325) at 1.8 Å resolution. The unit cell dimensions differ from those of the previously published structure. The structure was solved by molecular replacement and refined to R_{work}/R_{free} of 18.3/21.3% (Supplementary Table 1). The asymmetric unit consists of two chains, which form a dimer with non-crystallographic C2 symmetry. The structure is very similar overall to the previous one, with a backbone rmsd of 0.6 Å for residues that are common to both structures. In the new structure, electron density is visible for the N-terminal segment spanning residues 185–199. Intriguingly, this N-terminal segment binds into the active site (Asp220) of an adjacent protease dimer in a different asymmetric unit (Fig. 1a). The segment is positioned such that the active site Asp220 might cleave between amino acids 189 and 190 (Fig. 1b). However, the protein does not undergo auto-proteolysis in solution; mass spectrometry confirmed that the RVP domain remains intact even under acidic conditions (Supplementary Fig. S1). Nonetheless, while the observed arrangement is a likely an artefact of crystallization, the N-terminal segment acts as pseudo-substrate and reveals how Ddi1 might engage a substrate. The N-terminus adopts a β-strand conformation that interacts extensively with a loop formed by residues 245–258 (Fig. 1b). This loop is only visible in the chain that docks the N-terminus. There are multiple hydrogen bonds between the backbone atoms of the pseudo-substrate and the loop, suggesting that it might be involved in positioning the substrate for catalysis. The side chain of the conserved Gln224 (Fig. 1c) also makes a hydrogen bond with the backbone amide of the pseudo-substrate (Fig. 1b). The side chain of Phe246, which is conserved as an aromatic residue across Ddi1 orthologs, interacts with a methionine in the N-terminal fragment (Fig. 1c), suggesting this may be a specificity determinant for the protease activity of Ddi1. Arg251 is also conserved and could potentially interact with a substrate, but we did not observe any electron density for its side chain, implying that it is disordered. Finally, the side chain of Ile191 fits snugly into a hydrophobic groove formed by the loop and the rest of the RVP domain.

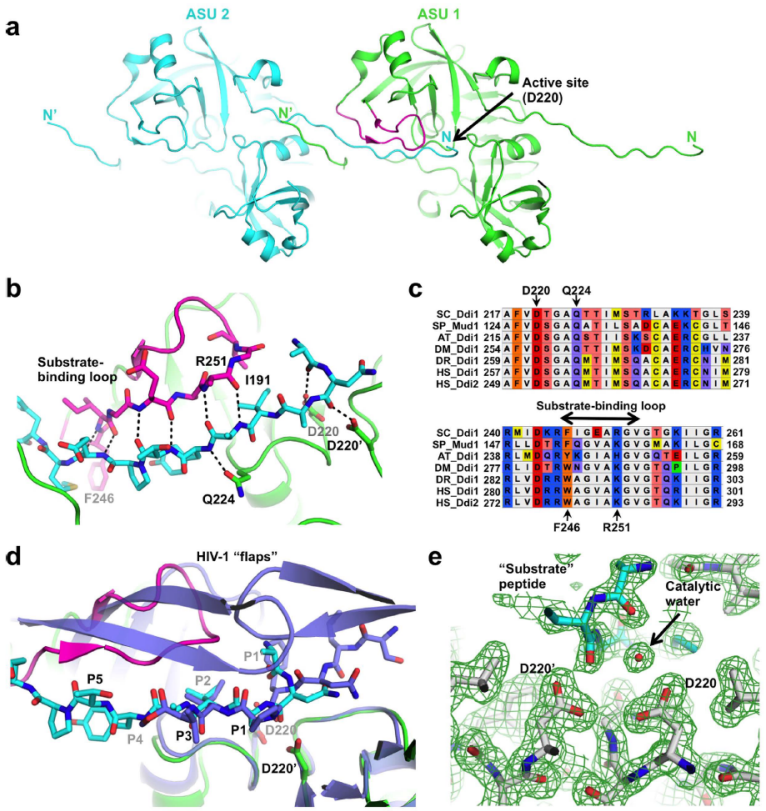


Figure 1. Crystal structure of the yeast Ddi1 protease domain reveals a potential substrate-binding mode. (**a**) Cartoon representation of two adjacent asymmetric units (ASU), showing the N-terminus of one molecule in ASU #2 (cyan) binding to the active site of a dimeric protease in ASU #1 (green). The loop that forms interactions with the N-terminal peptide is colored magenta. (**b**) Close-up view of the interaction between the N-terminal peptide and the active site. Hydrogen bonds are shown as dashed lines, and important residues are labeled. (**c**) Sequence alignment of Ddi1 orthologs from different species. SC, *Saccharomyces cerevisiae*; SP, *Schizosaccharomyces pombe*; AT, *Arabidopsis thaliana*; DM, *Drosophila melanogaster*; DR, *Danio rerio*; HS, *Homo sapiens*. (**d**) Superposition of the Ddi1 protease structure (green) with HIV-1 protease bound to a peptide substrate mimetic (violet, PDB 7HVP). The HIV substrate mimetic is shown in blue, and the Ddi1 pseudo-substrate N-terminal peptide is shown in cyan. (**e**) $2F_o$-F_c electron density maps of the active site, revealing the position of a water molecule that could act as a potential nucleophile in a proteolytic reaction.

Comparison with an HIV protease structure reveals substantial differences in substrate binding. In the structure of the HIV protease bound to a substrate-based hydroxyethylamine inhibitor[32], two flaps wrap around the substrate analog (Fig. 1d). Main-chain amides in the flaps of HIV-1 protease form hydrogen bonds with a water molecule that also binds the substrate analog. In Ddi1, the flap does not wrap around the pseudo-substrate, thus leaving it solvent-exposed. Remarkably, the pseudo-substrate and analogs in both Ddi1 and HIV-1 adopt very similar conformations, with main-chain atoms in the same configuration for amino acids in the P1 to P3 positions. Notably, both Ddi1 and HIV-1 protease substrates have an isoleucine in the P2 position. Overall, this suggests that the catalytic mechanism employed by both enzymes is similar, although they may engage their substrates differently.

The structure also reveals how Ddi1 could cleave a peptide bond: the catalytic residue Asp220 holds an ordered water molecule in place, which may act as the nucleophile for peptide bond hydrolysis (Fig. 1e). The symmetry-related Asp220′ forms a hydrogen bond with a carbonyl in the P1′ position, rendering it more

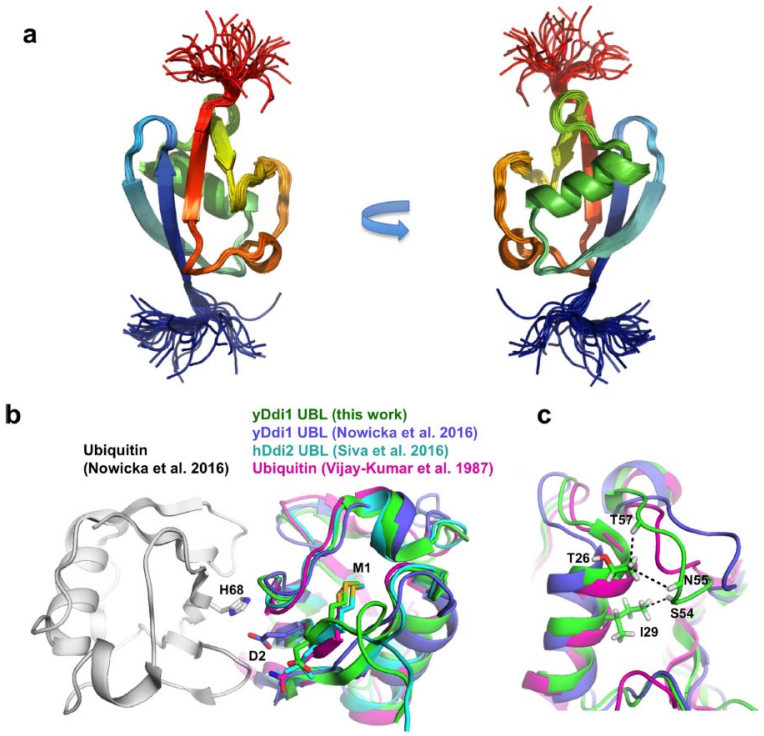


Figure 2. Solution structure of the Ddi1 UBL. (a) NMR solution structure of the UBL domain (a.a. 1–80; PDB 2N7E). An ensemble of 43 models is shown in cartoon representation, colored from blue to red from the N- to C-terminus. (**b**) Superposition of the yeast Ddi1 UBL NMR structure (green, pdb 2N7E) with the yeast Ddi1 UBL docked to ubiquitin (violet and white, pdb 2MWS), the human Ddi2 UBL NMR structure (cyan, pdb 2N7D), as well as the ubiquitin crystal structure (magenta, pdb 1UBQ). The side-chains of Met1 and Asp2 in the yDdi1 UBL (Gln2 and Leu2 in ubiquitin and hDdi2, respectively), and His68 in ubiquitin are shown as sticks. (**c**) Same as in (**b**), in a different orientation. The dashed lines indicate NOEs between Hα and methyl protons in the yeast Ddi1 UBL that confirms the proximity between the loop formed by a.a. 52–58 and the N-terminal segment of an a-helix (a.a. 26–29).

susceptible to nucleophilic attack. The water molecule is within hydrogen-bonding distance of two carbonyl oxygens in the N-terminal fragment, but it is not positioned in a way that would enable catalytic attack of the carbonyl carbon. Thus, the observed conformation would not lead to proteolysis. Thus, it remains unknown how the protease activity can be triggered and what the substrate(s) might be.

Proteomics screen for substrate(s) of Ddi1 protease. To characterize putative substrate(s) of Ddi1 RVP, we used a proteomic technique that employs a proteome-derived peptide library as a proteolytic substrate screen[31]. We used a peptide library derived from haploid yeast cells. We analyzed full-length Ddi1 as well as its "inactive" D220A variant at pH 4.0, 5.0, and 7.4. The data analysis revealed no Ddi1-dependent cleavage at all pH tested, whereas the HIV-1 positive control produced significant amount of proteolysis (Supplementary Fig. S2). This suggests that the protease domain of Ddi1 requires activation or may cleave intact proteins in their native conformations.

The UBL and UBA domains of Ddi1 bind to ubiquitin. UBL domains are known to be protein:protein interaction modules, and thus could potentially play a role in protease substrate recognition. We thus characterized the structure and interactions mediated by the Ddi1 UBL domain using NMR. The domain adopts the ubiquitin fold (Fig. 2a,b and Supplementary Table 2) in spite of its low sequence similarity to ubiquitin. Yeast Ddi1 UBL has a rather shallow hydrophobic patch that is located at the same sequential and structural location as in human Rad23A[33]. This β-sheet patch potentially may be a protein-protein interaction site. However, NMR

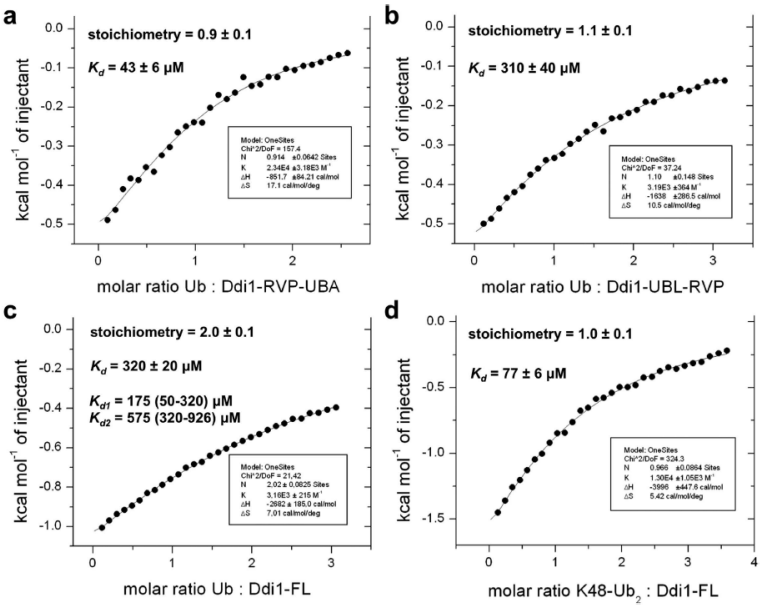


Figure 3. The UBL and UBA domains of Ddi1 bind to ubiquitin and diubiquitin. Ubiquitin or K48-linked diubiquitin (Ub_2) were added by syringe to Ddi1 constructs in the sample cell. The different titrations were (**a**) ubiquitin to RVP-UBA, (**b**) ubiquitin to UBL-RVP, (**c**) ubiquitin to full-length Ddi1, and (**d**) K48-Ub_2 to full-length Ddi1. For full-length Ddi1 binding to ubiquitin (**c**), the data was also fitted to a model with two independent binding sites (K_{d1} and K_{d2}), each with a stoichiometry of 1.

titrations of ^{15}N-Ddi1 UBL with potential UIM-containing ligands Ufo1 and Rpn10 showed no chemical shift perturbations, as reported by others (Supplementary Fig. S3)[34,35]. Moreover, no significant chemical shift perturbations were observed with the addition of Ddi1 86–325 (helical domains + protease) or Ddi1 388–428 (UBA), suggesting that the UBL would make no intramolecular contacts with other Ddi1 domains in the context of the full-length protein.

During preparation of this manuscript, the structure of Ddi1 UBL was published and its interaction with ubiquitin reported[34]. Comparison of the structures revealed small but significant differences. In our structure, the loop spanning a.a. 52–58 is in proximity to the N-terminus of the α-helix formed by a.a 24–34, with unambiguous NOEs between the two segments (Fig. 2c). This conformation, similar to the one found in ubiquitin, is different in the yeast Ddi1 UBL structure previously reported, where it is more distant from the α-helix. Moreover, our construct includes the N-terminal Met1 residue, which was absent from the construct used by Nowicka *et al.* The main-chain carbonyl of Met1 makes a hydrogen bond with Val19 and extends the first β strand, and its side-chain is oriented towards the core of the domain, as in ubiquitin or in the human Ddi2 UBL domain[36] (Fig. 2b). This confers a different orientation to Asp2, which side-chain points towards His68 in ubiquitin docked to yeast Ddi1 UBL[34].

To determine whether our Ddi1 constructs would bind to ubiquitin, we used isothermal titration calorimetry (ITC). We tested binding in the context of the full-length protein, with deletion of either the UBL or the UBA domain. Ubiquitin was found to bind the RVP-UBA construct (without the UBL) with a K_d of 43 µM, whereas it bound the UBL-RVP construct (without the UBA) with a K_d of 310 µM (Fig. 3a,b). Titration of full-length Ddi1 with ubiquitin yielded an average K_d of 320 µM, with a 2:1 stoichiometry (Fig. 3c). We also fitted the latter data to a model with two independent sites and found ranges of K_d values that are close to the values obtained for the deletion constructs. These results are thus consistent with both UBA and UBL interacting with ubiquitin. Finally, we found that K48-Ub_2 binds full-length Ddi1 with a K_d of 77 µM and a 1:1 stoichiometry, suggesting that each Ddi1 dimer binds two K48-Ub_2 molecules (Fig. 3d).

The structure of the helical domain of Ddi1 reveals similarities to DNA-binding domains. In an effort to elucidate the function of Ddi1, we extended our structural studies beyond its RVP and UBL domains. As previously observed[29], secondary structure prediction of Ddi1 shows that there is an α-helical region juxtaposed to the N-terminus of the protease domain (Fig. 4a). This domain is separated from the UBL domain by a linker of variable length in different organisms, but is always juxtaposed to the RVP domain. To confirm that this region

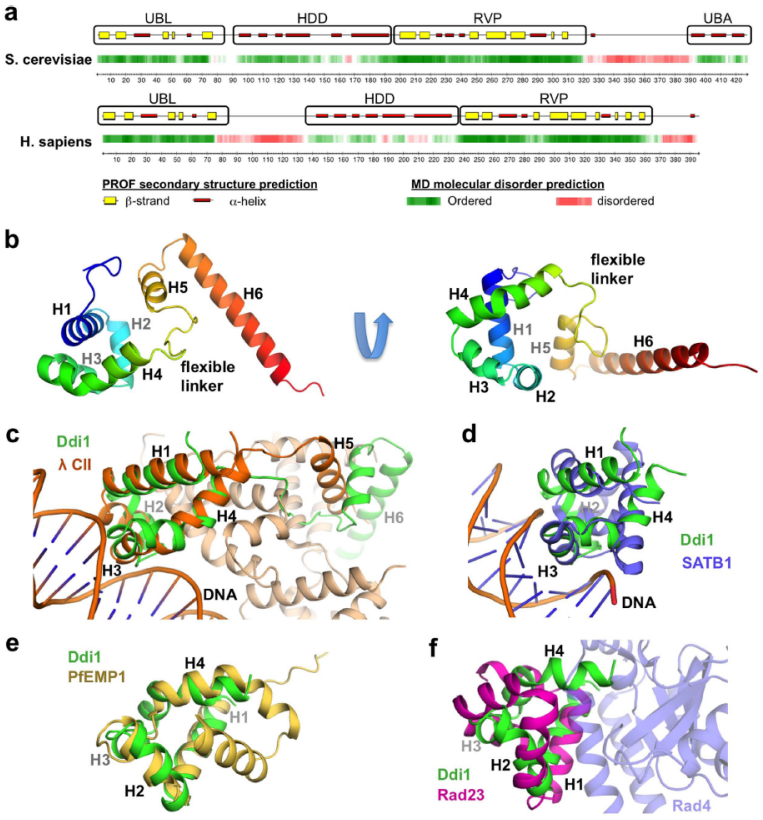


Figure 4. Ddi1 contains two helical domains preceding the protease domain. (a) Secondary structure and molecular disorder predictions of the yeast (top) and human (bottom) Ddi1 proteins. Predictions were performed on the PredictProtein server (http://www.predictprotein.org). (**b**) Cartoon of a representative model from the ensemble of HDD solution NMR structures, colored progressively from blue (N-terminus) to red (C-terminus). A flexible linker between helices 4–5 connects the two domains. (**c–f**) Structural superposition of the Ddi1 HDD N-terminal domain (green) with various homologous domains: (c) the bacteriophage λ CII transcription activator bound to a DNA duplex (orange, PDB 1ZS4). Other protein chains in the λ CII structure are colored in pale orange, showing the 5th helix mediating tetramerization. (**d**) SATB1 CUT domain bound to a DNA duplex (blue, PDB 2O4A). (**e**) Intracellular domain of *Plasmodium falciparum* erythrocyte membrane protein 1 (yellow, PDB 2LKL). (**f**) Rad23 XPCB domain (magenta, PDB 2F4M) bound to Rad4 (pale blue).

effectively forms one or many folded domains, we expressed ^{15}N,^{13}C-labeled Ddi1 (residues 86–196) and characterized its structure by NMR. Its ^{1}H,^{15}N HSQC spectrum showed good signal dispersion in the proton dimension, indicating that the construct is folded (Supplementary Fig. S4a). We confirmed that this region of Ddi1 effectively adopts a folded structure using ^{15}N-^{1}H heteronuclear NOE, which shows positive values except for the flexible N- and C-termini (Supplementary Fig. S4b). Because this helical region is folded and conserved across Ddi1 orthologs (Supplementary Fig. S4c), we named it the Helical Domain of Ddi1 (HDD).

The solution structure of HDD was determined using dihedral and NOE distance restraints, as well as residual dipolar couplings (RDCs) (Supplementary Table 3). The HDD actually consists of two alpha-helical domains (Fig. 4b): the N-terminal domain (residues 89–141) is a bundle of four helices with a hydrophobic core formed by some of the most conserved residues of the Ddi1 HDD. The C-terminal domain (residues 150–190) forms a hairpin with two helices, with a small hydrophobic core involving helix 1 and the first portion of the long helix 2. The structure calculation converged for each domain, with backbone average pairwise rmsd of 0.57 and 0.99 Å for the N- and

C-terminal domains, respectively. However, there is considerable variability in the relative position of each domain (Supplementary Fig. S5a). A 10-residue linker with lower heteronuclear NOE values (Supplementary Fig. S4b) as well as low sequence conservation amongst Ddi1 orthologs (Supplementary Fig. S4c) connects the two parts of HDD. No long-range ^{1}H-^{1}H NOE was detected between the N- and C-terminal domains, implying that they do not pack against each other. To determine the extent of the dynamic motion between the two domains, we analyzed small-angle X-ray scattering (SAXS) data using ensemble optimization[37,38]. The wide size and R_g distributions of the optimized ensembles of tethered domains structures hint to a dynamic regime with a slightly more compact configuration than expected from a random pool of structures (Supplementary Fig. S6). Finally, the alignment tensor rhombicities are significantly different, further confirming they tumble independently from each other (Supplementary Fig. S5b,c).

To gain insight into the potential function of the HDD region of Ddi1, we used the structure as a query to search for homologous domains in the Protein Data Bank (Supplementary Fig. S7a). Overall, we find that the HDD N-terminal domain (HDDnt) is similar to a wide and disparate set of alpha-helical bundle structures. However, the HDDnt displays striking similarity to DNA-binding domains from transcriptional regulators. Notably, the DNA-binding motif of the bacteriophage λ cII transcription activator has a Cα rmsd of 2.9 Å with HDDnt (Fig. 4c). The domain is followed by a flexible tether and a long helix that can take multiple positions[39], similar to the C-terminal domain of the HDD. This λ cII C-terminal helix mediates tetramer formation upon DNA-binding. The HDDnt domain is also similar to the POU-specific OCT-1 DNA-binding domain (rmsd 3.9 Å), as well as to the homologous CUT domain from human SATB1 (3.7 Å, Fig. 4d). These DNA-binding domains are all four-helical bundles that insert helix 3 into the major groove and recognize specific DNA sequences (Fig. 4c,d). The fold search also revealed similarity between HDDnt and the intracellular domain of the erythrocyte membrane protein 1 from *Plasmodium falciparum*, but the latter is a five-helical bundle, and only the first four helices are similar to the HDDnt, suggesting their function are likely unrelated (Fig. 4e). Finally, weak sequence homology prompted us to compare the structure of HDDnt to helical bundles of the Sti1-like family found in other UBL-UBA proteins such as Rad23 (Supplementary Fig. S7b). This domain binds to Rad4/XPC, a DNA-binding protein implicated in nucleotide excision repair. The first three helices of HDDnt are similar to XPC-binding motif of that Rad23 domain, but the 4th helix adopts a completely different orientation (Fig. 4f), suggesting it is unlikely that Ddi1 binds Rad4/XPC. Thus, the Ddi1 HDD is most similar to DNA-binding domains.

SAXS analysis suggests a dynamic structure for the Ddi1 dimer in solution. The Ddi1 protein consists of four domains with known structure, but how they are positioned to each other in the full-length protein is unclear. We therefore used SAXS to determine the relative position of the UBL and HDD domains with respect to the RVP dimer. Light scattering (dynamic and multiple-angle) was initially used to characterize and optimize solution scattering conditions for multiple Ddi1 constructs, except for the full-length protein, which formed aggregates that impeded analysis under all conditions tested. Constructs comprising the RVP domain (200–325) all formed dimers, as expected (Supplementary Fig. S8a). Ddi1 86–325 slightly aggregated at pH 7, but increasing the pH and adding glycerol reduced the aggregation.

SAXS data from the RVP domain fit well to the dimeric crystal structure reported here, with a χ^2 of 1.9, confirming that the Ddi1 RVP adopts the same conformation in solution as observed in the crystal structure (Fig. 5a). SAXS data were also acquired on Ddi1 86–325 (HDD-RVP) and 2-325 (UBL-HDD-RVP). The *P(r)* function reveals an asymmetrical pattern characteristic of elongated structures (Fig. 5b). To determine the relative positions of each domain, the SAXS data were fitted with the NMR structure of the UBL, the NMR structure of the HDD domain, and the crystal structure of the RVP domain as inputs. P2 symmetry was imposed, and the RVP domain position was kept constant. SAXS modeling of the Ddi1 HDD-RVP domains yielded two classes of structures, where the HDD domain extends on either side of the protease domain (Fig. 5c and Supplementary Fig. S8b). Calculations with the UBL-HDD-RVP data produced four classes of structures with similar overall shapes (Fig. 5c). The position of the UBL domain is highly variable and can be located on multiple sides of the HDD domain (Supplementary Fig. S8b). To determine the extent of the dynamics and reveal potential interactions among the domains, we also analyzed SAXS data using ensemble optimization[37,38]. In this approach, a pool of 10,000 models is generated from the structure of individual domains, connected by flexible linkers, to define the potential conformational space of the multi-domain protein. Then, a genetic algorithm is used to select a subset of conformers that best fit the experimental scattering data. In this case, P2 symmetry was maintained for the core RVP domain, but no symmetry was imposed on the UBL and HDD domains. Excellent fits were obtained for both data sets, and three independent calculations yielded similar R_g and D_{max} distributions for the best-fit ensemble (Fig. 5d and Supplementary Fig. S9). The ensembles of HDD-RVP structures have more compact structures with smaller R_g values than the pool (Fig. 5d). However analysis of D_{max} showed no clustering (Supplementary Fig. S9b), and the quantity R_{flex} is close to that of the pool, suggesting the protein is flexible. Similar results were obtained for the UBL-HDD-RVP construct, albeit with a class of relatively compact structures dominating the ensemble (Fig. 5d). However, an overlay of these compact structures revealed no favored arrangement, and R_{flex} is also high, reflecting flexibility. Overall, our analysis suggests that the Ddi1 protease dimer is flanked by flexible UBL and HDD domains that are in dynamic exchange in solution, albeit with a tendency towards more compact configurations.

Discussion

The focus of this study was to investigate the structure of yeast Ddi1 and the interactions mediated by its UBL domain. Surprisingly, we found that the UBL was unable to bind any of the UIMs that we tested, including all four UIMs found in the Ufo1 protein. Similar results were obtained by Nowicka *et al.*[34]. Based on the latter publication, we confirmed that the yeast UBL domain binds to ubiquitin. However, our ITC-based affinity measurements deviate from the reported affinities measured by NMR (150 and 45 μM for the UBA and UBL, respectively[34]),

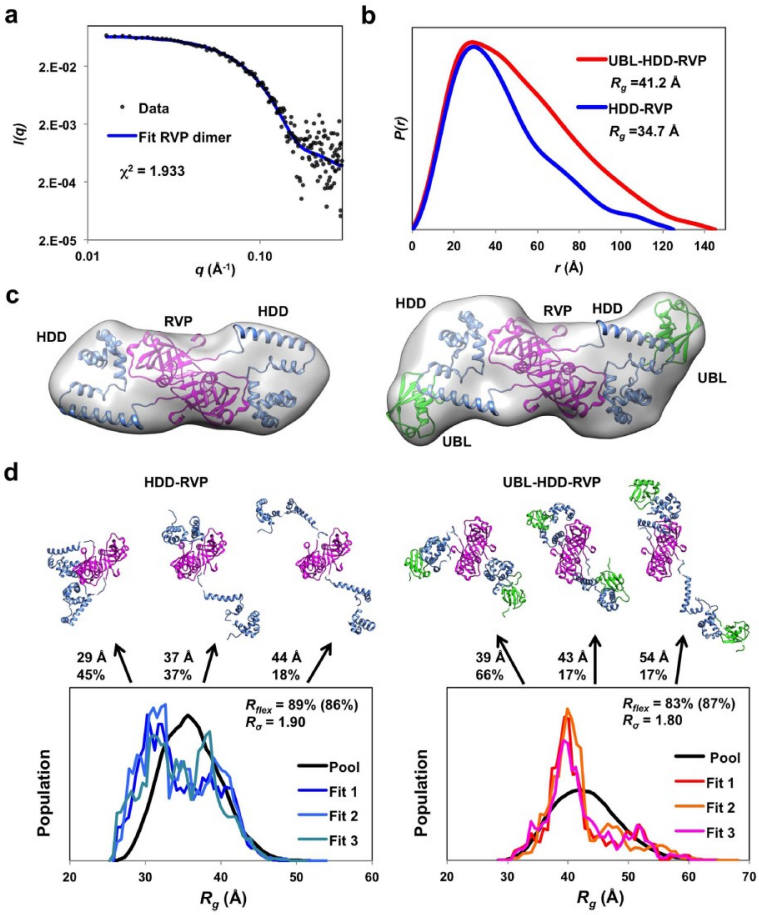


Figure 5. SAXS analysis of the Ddi1 dimer in solution. (a) SAXS data (circles) and calculated scattering curves (blue line) derived from the dimeric Ddi1 RVP domain crystal structure (200–325), displayed as a double-logarithmic plot. (**b**) SAXS pair-distance distribution functions for Ddi1 HDD-RVP (86–325) or UBL-HDD-RVP (2–325). (**c**) Modeling of SAXS data for HDD-RVP (left) and UBL-HDD-RVP (right), using the NMR structure of the UBL (green) and HDD (blue) domains, and the crystal structure of the dimeric protease domain (magenta). The surface represents the average of all models generated, contoured at 1.1× the volume of the particle. The cartoon displays a representative model from the ensemble. (**d**) Dynamic ensemble analysis of Ddi1 HDD-RVP (left) and UBL-HDD-RVP (right) using EOM. Each graph shows the distribution of R_g for a pool of 10,000 structures randomly generated, and three sets of 100 ensembles that best fit the data. The structures of the best ensemble is shown on top, with R_g and fraction indicated for each. These models suggest the behaviour of the flexible protein in solution, and does not represent the only solution.

which could be attributed to difference in buffers (sodium phosphate pH 6.8 versus HEPES pH 7.4 here) and/or the use of different constructs. Indeed, the NMR titrations were carried out with isolated domains, as opposed to the ITC titrations that were performed in the context of the full-length dimeric protein. The association rate constants of protein:protein interactions is dependent on translational and rotational diffusion rates and protein electrostatics, which vary with protein size and depends on the context in which the domain is positioned[40]. These factors could explain the 3-fold difference in K_d (about 0.7 kcal/mol in free energy) we measured for the Ddi1 UBA domain. Moreover, the construct used by Nowicka *et al.* lacked the initiator Met1, which alters the

position of Asp2, poised to interact with ubiquitin (Fig. 2b). This would reduce the affinity of our construct for ubiquitin, and we indeed observe a 6-fold difference. Whether the yeast Ddi1 protein retains its N-terminal Met1 is unknown, but proteins with a charged residue in the second position are typically not excised, and often acetylated[41]. Nevertheless, our data are overall consistent with yeast Ddi1 having two independent binding sites for ubiquitin. We also found that full-length yeast Ddi1 binds K48-Ub_2 with a 1:1 stoichiometry. These results could be explained by a mixture of binding modes, including binding to the UBL and UBA within the same Ddi1 molecule, or two UBL or two UBA domains in the dimer. The presence of two ubiquitin-binding domains in yeast Ddi1 may explain why some orthologs lack one or the other domain. For example, the *S. pombe* ortholog Mud1 does not have a UBL domain, but its UBA domain binds tightly to K48-Ub_2[5]. However, in an accompanying publication[36], we found that the UBL domain of human Ddi2, which lacks a UBA domain, does not bind to ubiquitin, implying that ubiquitin-binding is not a conserved attribute of the Ddi1-like eukaryotic family of proteins.

Our crystal structure of the RVP domain reveals that the protease domain binds the N-terminus of the protein construct, which adopts an extended β conformation through interaction with a conserved loop adjacent to the active site (Fig. 1). This loop normally forms a flap in aspartyl proteases of retroviruses such as HIV-1, but in our structure, the loop forms an extensive hydrogen-bonding network with the N-terminal segment pseudo-substrate. This conformation is likely a crystallization artefact for the following reasons: in the previous crystal structure of the yeast Ddi1 RVP domain[30], the segment 180–198 is disordered and not observed in the active site; the construct that was crystallized is not cleaved in solution (Supplementary Fig. S1); the segment 185–191 is actually part of the HDD domain, which adopts an alpha-helical conformation in solution (Fig. 4). Yet, the pseudo-substrate N-terminal segment, adopts a conformation similar to HIV-1 protease peptide substrates. This suggests that the yeast Ddi1 RVP domain indeed functions as a hydrolytic enzyme against polypeptides, and that the observed conformation likely represents a model of substrate binding.

Unlike retroviral aspartyl proteases, which are able to cleave peptide substrates *in vitro*, Ddi1 did not exhibit any protease activity in our proteomic screen against a library of peptides derived from the yeast proteome (Supplementary Fig. S2). While this could mean that this domain might simply not be a protease, it could also be that the protease is activated only in the context of its interactions with another protein. The newly identified HDD domain could serve as a substrate anchor in this context (Fig. 4). Our SAXS data show that the HDD extends on either side of the RVP dimer and could thus serve as a landing platform for a substrate (Fig. 5c). We report a similar configuration in human Ddi2[36] where the HDD is juxtaposed to the protease dimer, suggesting a conserved functional relationship between the HDD and RVP domains. However, ensemble modeling of the SAXS data shows that the UBL and HDD domains are likely dynamic and can adopt multiple configurations (Fig. 5d). Considering the underdetermined nature of SAXS data, we cannot conclude whether the UBL-HDD-RVP module adopts a single rigid or multiple dynamic conformations. However, we note that NMR titrations with the UBL and the HDD-RVP constructs revealed no interaction, in agreement with the dynamic nature of the UBL-HDD-RVP module.

The role of the HDD remains unclear, but it is likely mediating interactions with a potential substrate of the Ddi1 protease. Sti1-like domains homologous to the HDD have indeed been implicated in protein-protein interactions. In particular, the Rad23 Sti1-like domain forms a complex with the nucleotide excision repair protein Rad4/XPC[42,43], and the role of this interaction might be to protect Rad4/XPC from proteasomal degradation[44]. Dsk2/ubiquilin also contains a Sti1-like domain that has been proposed to bind the Hsp70-like protein Stch[45]. More strikingly, the structural homology with DNA-binding domains hints to the possibility that Ddi1 might be recognizing specific DNA motifs itself. The domain has some hallmarks of DNA-binding domains, such as a conserved basic residue (Arg131) that faces the DNA phosphate backbone in the structure alignment with the bacteriophage λ CII transcription activator bound to a DNA duplex. This is consistent with the observation that Ddi1 localizes to the nucleus[24]. As Ddi1 is implicated in the DNA-damage response and cell cycle checkpoints[11,15], it is possible that Ddi1 HDD binds to DNA damage sites whose regulation requires an ubiquitin-dependent proteolysis event. Future work should focus on identifying binding partners for Ddi1 HDD.

Methods

Protein expression and purification. The DDI1 gene was amplified by PCR from yeast genomic DNA and used as a template to generate fragments containing UBL (residues 2–80), HDD (86–196), RVP (185–325), HDD-RVP (86–325), and UBL-HDD-RVP (2–325). Yeast Ufo1 UIM1-4 (residues 512–668), UIM1 (512–539), UIM2 (542–569), and UIM3 (577–608) were similarly amplified by PCR. Rpn10 cloning was described previously[46]. These fragments were cloned into the pGEX-6p1 plasmid in-frame with an N-terminal GST tag *via* *Bam*H1 and *Xho*I sites. Constructs were expressed overnight with 0.5 mM IPTG at 20 °C in *E. coli* BL21(DE3) cells, resuspended in TBS buffer (50 mM Tris-HCl, pH 8.0, 100 mM NaCl, 5 mM β-mercaptoethanol) supplemented with 1 mM EDTA, and lysed by sonication. The fusion protein was purified using glutathione-sepharose affinity and eluted with 20 mM glutathione dissolved in TBS. The fusion protein was cleaved overnight at 4 °C with the 3C protease and applied onto size-exclusion Superdex 200 or Superdex 75 16/600 chromatography columns (GE Healthcare). Contaminant GST was removed using glutathione-sepharose resin. Gel filtration was performed in NMR buffer (10 mM HEPES-NaOH, pH 7.0, 50 mM NaCl, 5 mM β-mercaptoethanol) for Ddi1 UBL, HDD, HDD-RVP, and UBA and Ufo1 UIM1-4, or SAXS buffer (10 mM Tris-HCl, pH 8.0, 50 mM NaCl, 5% glycerol, 5 mM β-mercaptoethanol) for constructs 2–325, 86–325, 185–325, and 86–196. Ufo1 UIM4 (651–668) was synthesized chemically. Single UIMs were further purified by C18 reverse chromatography and lyophilized prior to resuspension in NMR buffer.

Full-length yeast Ddi1, UBL (1–80), UBL-RVP (1–325), and RVP-UBA (180–428) were also cloned into pET16b vector (Novagen) in-frame with an N-terminal histidine tag and used for ITC measurements. They were expressed in *E. coli* BL21(DE3)RIL host cells; subsequently resuspended in 50 mM Tris-HCl, pH 8.0, 50 mM NaCl, and 1 mM EDTA; and lysed by three passages through an EmulsiFlex-C3 high pressure homogenizer (Avestin,

Canada) at 1200 bar. Proteins were purified using nickel affinity chromatography and eluted with 250 mM imidazole. Afterwards, they were dialyzed overnight into 50 mM HEPES, pH 7.4, 150 mM NaCl, and 10% glycerol and further purified by size-exclusion chromatography on a Superdex 200 16/60 gel filtration column (GE Healthcare). Individual fractions were analyzed by SDS-PAGE and/or Western blot.

The Ddi1 UBL with an N-terminal His-tag (residues 1–80) was expressed as ^{15}N- and ^{15}N/^{13}C-labeled proteins in cells grown in minimal medium containing 0.8 g/L [^{15}N] ammonium chloride and 2 g/L d-[^{13}C] glucose, as required. The purification procedure was the same as described above.

K48-linked Ub_2 was synthesized using K48C and D77 ubiquitin mutants mixed with human E1 and yeast Cdc34, as previously described[46,47]. The products were purified by cation-exchange chromatography (mono S 5/50 GL, GE Healthcare) using 25 mM sodium acetate, pH 4.5, and 1 M NaCl for elution, and buffer-exchanged in ITC buffer.

X-ray crystallography. Ddi1 RVP (residues 185–325) was purified by gel filtration and concentrated to 8 mg/mL in 10 mM Tris-HCl, pH 7.4, 50 mM NaCl and 1 mM DTT. MALDI-TOF analysis revealed a single species with an average molecular weight of 16,320 Da (predicted 16,313 Da). The protein was crystallized by vapor diffusion using the sitting drop technique by mixing 1 μL of protein solution with 1 μL of crystallization solution (0.1 M phosphate-citrate, pH 4.2, 0.4 M NaCl, 20% PEG 8000). A crystal grew in 1–2 days. The crystal was cryo-protected by addition of 15% glycerol to the crystallization solution.

X-ray diffraction data at 100K were acquired at the CHESS beamline A1a (Supplementary Table 1). A total of 240 images with an oscillation angle of 0.5 were collected. Reflections were integrated using iMOSFLM and scaled with SCALA as implemented in the *CCP4* package[48]. The structure was determined by molecular replacement with the program Phaser[49], using chains A and B of the yeast Ddi1 RVP structure as a search model (PDB code 2I1A[30]). Model building was performed using the program COOT[50]. Restrained and TLS refinement were performed using Refmac5[48].

Nuclear magnetic resonance spectroscopy. NMR spectra were acquired from 350 μl samples of 0.2 mM ^{15}N-labeled Ddi1 UBL for binding site mapping or 0.5 mM ^{13}C/^{15}N-labeled Ddi1 UBL for structural determination in a 50 mM sodium phosphate buffer, pH 7.4, containing 0.5% glycerol and 5% D_2O/95% H_2O. All NMR data for the UBL were collected at 25 °C on a 600 MHz Bruker Avance spectrometer equipped with a triple resonance (^{15}N/^{13}C/^{1}H) cryoprobe. For determination of the sequence-specific resonance assignments for the UBL domain, a series of double and triple resonance spectra were collected as described previously[51,52]. ^{1}H-^{1}H distance constraints required to calculate the structure were derived from NOEs identified in 3D ^{15}N/^{1}H NOESY-HSQC, and ^{13}C/^{1}H HSQC-NOESY spectra, which were acquired with an NOE mixing time of 120 ms. Specific interaction of proteins and peptides with the Ddi1 UBL was monitored by changes induced in the positions of signals of ^{15}N-labeled Ddi1 UBL 2D ^{15}N/^{1}H HSQC spectra using a recently described combined minimal shift approach[53]. A two-fold molar excess (0.4 mM) of Rpn10, Ddi1 HDD-RVP, Ddi1 UBA, Ufo1 UIM1-4, UIM1, UIM2, and UIM4 was added in these experiments. All spectra were processed using Topspin 3.2 (Bruker) and analyzed using Sparky (www.cgl.ucsf.edu/home/sparky).

All data sets for the Ddi1 86–196 HDD domain were acquired in HEPES-based NMR buffer at 30 °C on a 600 or 850 MHz Bruker NMR spectrometer both equipped with a triple-resonance (^{1}H, ^{13}C, ^{15}N) cryoprobe. Heteronuclear ^{1}H-^{15}N NOE values were measured at 600 MHz, as described[54]. ^{15}N-^{1}H residual dipolar couplings were measured in 10 mg/mL Pf1 bacteriophage[55] at 600 MHz using a sensitivity-enhanced HSQC-IPAP experiment[56]. Backbone assignments were performed on ^{15}N,^{13}C-labeled protein samples (0.5 mM) using CBCACONH and HNCACB NMR experiments. ^{1}H-^{1}H distance constraints required to calculate the structure were derived from NOEs identified in 3D ^{15}N/^{1}H NOESY-HSQC, and ^{13}C/^{1}H HSQC-NOESY spectra, which were acquired at 850 MHz with an NOE mixing time of 120 ms.

The family of converged structures for Ddi1 UBL and HDD were initially calculated using Cyana 2.1[57,58]. NOE-derived restraints from 3D ^{15}N- and ^{13}C-edited NOESY spectra, which were assigned using combined automated NOE assignment and structure determination protocol, were used to produce preliminary structures. Backbone torsion angle constrains were generated from assigned chemical shifts using the program TALOS+[59]. For the UBL, hydrogen bond constraints involving residues with slowly exchanging amide protons were used in the calculations. Subsequently, five cycles of simulated annealing combined with redundant dihedral angle constrains were performed to produce a set of 43 converged structures with the lowest Cyana target function, no distance constraint violation and van der Waals violations greater than 0.2 Å, and no dihedral angle constraint violation greater than 5°. These were further refined in explicit solvent using the YASARA forcefield[60]. The structure of the HDD domain was further refined in XPLOR-NIH to incorporate residual dipolar couplings for residues displaying heteronuclear NOE values above 0.6, i.e. 90–141 (N-terminal domain) and 151–187 (C-terminal domain). Initial estimates and Monte Carlo calculations of the alignment tensor D_a and D_r were obtained using the software MODULE[61]. Satisfactory R_{dip} values of 30% and 35% were obtained for the N- and C-terminal domains, respectively, prior to the RDC refinement in XPLOR-NIH, indicating the accuracy of the NOE-derived structure. D_a and D_r were then optimized using a grid-search in XPLOR-NIH. Structures with the lowest total energy were selected.

Small-angle X-ray scattering (SAXS). Small-angle X-ray scattering data sets were collected on an in-house Anton Paar SAXSess camera equipped with a PANalytical PW3830 X-ray generator and a Roper/Princeton CCD detector. The beam length was set to 18 mm, and the beam profile was recorded using an image plate for subsequent desmearing. Scattering data were collected at 4 °C at protein concentrations of 4.0 and 8.0 mg/mL for 1 hour for Ddi1 185–325, 5.0 and 10.0 mg/mL for 2 hours for 2–325, 5.0 and 10.0 mg/mL for 2 hours for 86–325, and 5.0 and 10.0 mg/mL for 2 hours for 86–196. Background scattering from the SAXS buffer was

measured for 2 hours. Dark current correction, scaling, buffer subtraction, binning, desmearing, and merging were performed using SAXSquant 3.0 (Anton Paar). The merged scattering curves were then analyzed with different software included in the ATSAS package[62]. Scattering data were fitted to chains A & B of the Ddi1 RVP crystal structure using CRYSOL, pair-distance distributions and R_g values were calculated using GNOM. Molecular weights were estimated using the Q_r invariant as described[63]. Ensemble optimization of the Ddi1 86–196 structure against the SAXS data was performed using EOM 2.0 by generating 10,000 structures with the NMR structures of the N-terminal (a.a. 89–141) and C-terminal (a.a. 150–191) HDD domains, selected using a genetic algorithm. Modeling of data collected from Ddi1 2–325 and 86–325 was performed using CORAL with the NMR structure of the UBL (a.a. 2–75), the two HDD domains (a.a. 89–141 & 150–191) and the dimeric crystal structure of the RVP (a.a. 200–325). Twenty models with $\chi^2 < 1.6$ were generated and averaged using DAMAVER, with average χ^2 of 1.36 and 1.19 for HDD-RVP and UBL-HDD-RVP, respectively. The resulting coordinates were used to generate pseudo-densities using Situs-pdb2vol[64] and contoured 10% above the particle volume derived from the Porod invariant (109,000 and 148,000 Å^3 for HDD-RVP and UBL-HDD-RVP, respectively) using UCSF-Chimera[65]. EOM 2.0 was used to generate 10,000 structures using the same domains as used in CORAL, with P2 symmetry imposed only on the RVP domain, using the genetic algorithm for conformer selection. The genetic algorithm was performed 100 times thrice to estimate the variability in the distribution of D_{max} and R_g values.

Isothermal titration calorimetry (ITC). All calorimetric titrations of ubiquitin with full-length yeast Ddi1 and truncated variants were performed in 50 mM HEPES, pH 7.4, 150 mM NaCl at 25 °C using a VP-ITC system (MicroCal, GE Healthcare Life Sciences). For full-length Ddi1, 9 μL aliquots of 1.42 mM bovine ubiquitin (Sigma, cat. no. U6253) were injected stepwise into a sample cell containing 1.43 ml of 97 μM Ddi1 protein (concentration calculated to monomer). For UBL-RVP, 9 μl aliquots of 2 mM bovine ubiquitin were injected stepwise into a sample cell containing 1.43 ml of 133.1 μM Ddi1 UBL-RVP protein, and for RVP-UBA, 9 μl aliquots of 796 μM bovine ubiquitin were injected stepwise into a sample cell containing 1.43 ml of 64.8 μM Ddi1 RVP-UBA protein. The control dilution experiment, in which ubiquitin was injected into buffer alone, was also performed. All proteins used for titrations were properly dialyzed against buffer at 4 °C overnight, and their exact concentrations were determined by HPLC amino acid analysis. Titration of K48-Ub_2 with full-length yeast Ddi1 was performed in 50 mM HEPES, pH 7.4, 150 mM NaCl at 25 °C. Nine-microliter aliquots of 833.5 μM K48-Ub_2 were injected stepwise into a sample cell containing 1.43 ml of 48.5 μM Ddi1 protein. Data sets were analyzed using Origin, using a one-site model by varying N, K_d and ΔH. For the titration with Ddi1 FL and ubiquitin, the data were also fitted to a two-sites model, where N was fixed to 1.0 and K_d and ΔH were floating variables for both sites. The range of values was determined by allowing the χ^2 value to increase up to 37.3, observed at $K_{d1} = 50$ μM and $K_{d2} = 926$ μM, which still gives a satisfactory fit. The minimum χ^2 value of 21.4 was observed at $K_{d1} = 175$ μM and $K_{d2} = 575$ μM.

PICS assay and analysis. The PICS procedure was carried out as previously described[31], and further details are included in our back-to-back publication[34]. Briefly, the amine-protected yeast proteome-derived peptide library (1 mg/ml) was incubated in 200 μL buffer with 4 μg of full-length yeast Ddi1 WT. The reaction was incubated for 12 h at 37 °C. The proteolytic cleavage assays were carried out in of 100 mM sodium acetate, 300 mM NaCl, pH 4.0, 100 mM sodium acetate, 300 mM NaCl, pH 5.0, or 100 mM HEPES, 300 mM NaCl, pH 7.0. As negative controls, we used full-length Ddi1 with an inactivating mutation in its catalytic site (D220A), as well as a mock reaction with buffer. As a positive control, we tested the HIV-1 protease cleavage profile in 100 mM Na acetate, 300 mM NaCl, pH 4.7, using wild-type and the catalytically inactive D25N mutant in a 1:200 protease-to-library ratio.

Data were analyzed by a series of predesigned queries in Microsoft Access database management software. First, lists of identified peptides from each MS run were filtered for peptides containing products of N-terminal modification by biotinylation. Second, peptides with over 80% confidence were picked for the tested enzyme, while peptides with over 10% confidence were picked for control reactions. To properly subtract the background signal, the list of peptides identified in the tested enzyme reaction was screened for peptides present in the mock reaction as well as in the reaction with catalytically inactive, and those peptides were removed from processing. Finally, peptides identified in the original unprocessed peptide library were removed from the analysis.

The final cleared list of identified peptides was then aligned with a FASTA proteomics database used for proteomics database search to determine the N-terminal portions of cleaved peptides. If multiple computationally identified preceding sequences were found for one MS identified peptide, they were removed from processing, while the MS identified peptide sequences were kept in the list for downstream analysis. The final list of substrate peptides containing sequences of five P' amino acids identified in the MS experiment and five P amino acids identified computationally was then created. The frequency of each amino acid in each particular position was calculated and plotted, yielding the substrate specificity matrix.

References

1. Hershko, A. & Ciechanover, A. The ubiquitin system. *Annu Rev Biochem* **67**, 425–479 (1998).
2. Hicke, L. *et al.* Ubiquitin-binding domains. *Nat Rev Mol Cell Biol* **6**, 610–621 (2005).
3. Bertolaet, B. L. *et al.* UBA domains of DNA damage-inducible proteins interact with ubiquitin. *Nat Struct Biol* **8**, 417–422 (2001).
4. Wilkinson, C. R. *et al.* Proteins containing the UBA domain are able to bind to multi-ubiquitin chains. *Nat Cell Biol* **3**, 939–943 (2001).
5. Trempe, J. F. *et al.* Mechanism of Lys48-linked polyubiquitin chain recognition by the Mud1 UBA domain. *EMBO J* **24**, 3178–3189 (2005).
6. Raasi, S. *et al.* Diverse polyubiquitin interaction properties of ubiquitin-associated domains. *Nat Struct Mol Biol* **12**, 708–714 (2005).
7. Elsasser, S. *et al.* Proteasome subunit Rpn1 binds ubiquitin-like protein domains. *Nat Cell Biol* **4**, 725–730 (2002).
8. Saeki, Y. *et al.* Ubiquitin-like proteins and Rpn10 play cooperative roles in ubiquitin-dependent proteolysis. *Biochem Biophys Res Commun* **293**, 986–992 (2002).

9. Seeger, M. *et al.* Interaction of the anaphase-promoting complex/cyclosome and proteasome protein complexes with multiubiquitin chain-binding proteins. *J Biol Chem* **278**, 16791–16796 (2003).
10. Gomez, T. A. *et al.* Identification of a functional docking site in the Rpn1 LRR domain for the UBA-UBL domain protein Ddi1. *BMC Biol* **9**, 33 (2011).
11. Clarke, D. J. *et al.* Dosage suppressors of pds1 implicate ubiquitin-associated domains in checkpoint control. *Mol Cell Biol* **21**, 1997–2007 (2001).
12. Diaz-Martinez, L. A. *et al.* Yeast UBL-UBA proteins have partially redundant functions in cell cycle control. *Cell Div* **1**, 28 (2006).
13. Liu, Y. & Xiao, W. Bidirectional regulation of two DNA-damage-inducible genes, MAG1 and DDI1, from Saccharomyces cerevisiae. *Mol Microbiol* **23**, 777–789 (1997).
14. Liu, Y. *et al.* UAS(MAG1), a yeast cis-acting element that regulates the expression of MAG1, is located within the protein coding region of DDI1. *Mol Gen Genet* **255**, 533–542 (1997).
15. Zhu, Y. & Xiao, W. Differential regulation of two closely clustered yeast genes, MAG1 and DDI1, by cell-cycle checkpoints. *Nucleic Acids Res* **26**, 5402–5408 (1998).
16. Zhu, Y. & Xiao, W. Two alternative cell cycle checkpoint pathways differentially control DNA damage-dependent induction of MAG1 and DDI1 expression in yeast. *Mol Genet Genomics* **266**, 436–444 (2001).
17. Fu, Y. *et al.* Rad6-Rad18 mediates a eukaryotic SOS response by ubiquitinating the 9-1-1 checkpoint clamp. *Cell* **133**, 601–611 (2008).
18. Kaplun, L. *et al.* The DNA damage-inducible UbL-UbA protein Ddi1 participates in Mec1-mediated degradation of Ho endonuclease. *Mol Cell Biol* **25**, 5355–5362 (2005).
19. Kaplun, L. *et al.* Functions of the DNA damage response pathway target Ho endonuclease of yeast for degradation via the ubiquitin-26S proteasome system. *Proc Natl Acad Sci USA* **97**, 10077–10082 (2000).
20. Kaplun, L. *et al.* DNA damage response-mediated degradation of Ho endonuclease via the ubiquitin system involves its nuclear export. *J Biol Chem* **278**, 48727–48734 (2003).
21. Ivantsiv, Y. *et al.* Unique role for the UbL-UbA protein Ddi1 in turnover of SCFUfo1 complexes. *Mol Cell Biol* **26**, 1579–1588 (2006).
22. Lustgarten, V. & Gerst, J. E. Yeast VSM1 encodes a v-SNARE binding protein that may act as a negative regulator of constitutive exocytosis. *Mol Cell Biol* **19**, 4480–4494 (1999).
23. Marash, M. & Gerst, J. E. Phosphorylation of the autoinhibitory domain of the Sso t-SNAREs promotes binding of the Vsm1 SNARE regulator in yeast. *Mol Biol Cell* **14**, 3114–3125 (2003).
24. Gabriely, G. *et al.* Different domains of the UBL-UBA ubiquitin receptor, Ddi1/Vsm1, are involved in its multiple cellular roles. *Mol Biol Cell* **19**, 3625–3637 (2008).
25. White, R. E. *et al.* The retroviral proteinase active site and the N-terminus of Ddi1 are required for repression of protein secretion. *FEBS Lett* **585**, 139–142 (2011).
26. Dixit, G. *et al.* Guanine nucleotide-binding protein (Galpha) endocytosis by a cascade of ubiquitin binding domain proteins is required for sustained morphogenesis and proper mating in yeast. *J Biol Chem* **289**, 15052–15063 (2014).
27. Krylov, D. M. & Koonin, E. V. A novel family of predicted retroviral-like aspartyl proteases with a possible key role in eukaryotic cell cycle control. *Curr Biol* **11**, R584–R587 (2001).
28. White, R. E. *et al.* HIV proteinase inhibitors target the Ddi1-like protein of Leishmania parasites. *FASEB J* (2011).
29. Perteguer, M. J. *et al.* Ddi1-like protein from Leishmania major is an active aspartyl proteinase. *Cell Stress Chaperones* **18**, 171–181 (2013).
30. Sirkis, R. *et al.* Ddi1, a eukaryotic protein with the retroviral protease fold. *J Mol Biol* **364**, 376–387 (2006).
31. Schilling, O. *et al.* Characterization of the prime and non-prime active site specificities of proteases by proteome-derived peptide libraries and tandem mass spectrometry. *Nat Protoc* **6**, 111–120 (2011).
32. Swain, A. L. *et al.* X-ray crystallographic structure of a complex between a synthetic protease of human immunodeficiency virus 1 and a substrate-based hydroxyethylamine inhibitor. *Proc Natl Acad Sci USA* **87**, 8805–8809 (1990).
33. Mueller, T. D. & Feigon, J. Structural determinants for the binding of ubiquitin-like domains to the proteasome. *EMBO J* **22**, 4634–4645 (2003).
34. Nowicka, U. *et al.* DNA-Damage-Inducible 1 Protein (Ddi1) Contains an Uncharacteristic Ubiquitin-like Domain that Binds Ubiquitin. *Structure* **23**, 542–557 (2015).
35. Zhang, D. *et al.* Together, Rpn10 and Dsk2 can serve as a polyubiquitin chain-length sensor. *Mol Cell* **36**, 1018–1033 (2009).
36. Siva, M. *et al.* Human DNA-Damage-Inducible 2 Protein Is Structurally and Functionally Distinct from Its Yeast Ortholog. *Sci Rep* **6**, 30443 (2016).
37. Bernado, P. *et al.* Structural characterization of flexible proteins using small-angle X-ray scattering. *J Am Chem Soc* **129**, 5656–5664 (2007).
38. Tria, G. *et al.* Advanced ensemble modelling of flexible macromolecules using X-ray solution scattering. *IUCrJ* **2**, 207–217 (2015).
39. Jain, D. *et al.* Crystal structure of bacteriophage lambda cII and its DNA complex. *Mol Cell* **19**, 259–269 (2005).
40. Schreiber, G. *et al.* Fundamental aspects of protein-protein association kinetics. *Chem Rev* **109**, 839–860 (2009).
41. Bonissone, S. *et al.* N-terminal protein processing: a comparative proteogenomic analysis. *Mol Cell Proteomics* **12**, 14–28 (2013).
42. Min, J. H. & Pavletich, N. P. Recognition of DNA damage by the Rad4 nucleotide excision repair protein. *Nature* **449**, 570–575 (2007).
43. Masutani, C. *et al.* Identification and characterization of XPC-binding domain of hHR23B. *Mol Cell Biol* **17**, 6915–6923 (1997).
44. Ng, J. M. *et al.* A novel regulation mechanism of DNA repair by damage-induced and RAD23-dependent stabilization of xeroderma pigmentosum group C protein. *Genes Dev* **17**, 1630–1645 (2003).
45. Kaye, F. J. *et al.* A family of ubiquitin-like proteins binds the ATPase domain of Hsp70-like Sth. *FEBS Lett* **467**, 348–355 (2000).
46. Riedinger, C. *et al.* Structure of Rpn10 and its interactions with polyubiquitin chains and the proteasome subunit Rpn12. *J Biol Chem* **285**, 33992–34003 (2010).
47. Trempe, J. F. *et al.* A new crystal form of Lys48-linked diubiquitin. *Acta Crystallogr Sect F Struct Biol Cryst Commun* **66**, 994–998 (2010).
48. Winn, M. D. *et al.* Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol Crystallogr* **67**, 235–242 (2011).
49. McCoy, A. J. *et al.* Phaser crystallographic software. *J Appl Crystallogr* **40**, 658–674 (2007).
50. Emsley, P. *et al.* Features and Development of Coot. *Acta Crystallogr D Biol Crystallogr* **66**, 486–501 (2010).
51. Veverka, V. *et al.* NMR assignment of the mTOR domain responsible for rapamycin binding. *J Biomol NMR* **36** Suppl 1, 3 (2006).
52. Renshaw, P. S. *et al.* Sequence-specific assignment and secondary structure determination of the 195-residue complex formed by the Mycobacterium tuberculosis proteins CFP-10 and ESAT-6. *J Biomol NMR* **30**, 225–226 (2004).
53. Veverka, V. *et al.* Structural characterization of the interaction of mTOR with phosphatidic acid and a novel class of inhibitor: compelling evidence for a central role of the FRB domain in small molecule-mediated regulation of mTOR. *Oncogene* **27**, 585–595 (2008).
54. Kay, L. E. *et al.* Backbone dynamics of proteins as studied by 15N inverse detected heteronuclear NMR spectroscopy: application to staphylococcal nuclease. *Biochemistry* **28**, 8972–8979 (1989).
55. Hansen, M. R. *et al.* Tunable alignment of macromolecules by filamentous phage yields dipolar coupling interactions. *Nat Struct Biol* **5**, 1065–1074 (1998).

56. Cordier, F. *et al.* A doublet-separated sensitivity-enhanced HSQC for the determination of scalar and dipolar one-bond J-couplings. *J Biomol NMR* **13**, 175–180 (1999).
57. Guntert, P. *et al.* Torsion angle dynamics for NMR structure calculation with the new program DYANA. *J Mol Biol* **273**, 283–298 (1997).
58. Herrmann, T. *et al.* Protein NMR structure determination with automated NOE assignment using the new software CANDID and the torsion angle dynamics algorithm DYANA. *J Mol Biol* **319**, 209–227 (2002).
59. Shen, Y. *et al.* TALOS+: a hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. *J Biomol NMR* **44**, 213–223 (2009).
60. Harjes, E. *et al.* GTP-Ras disrupts the intramolecular complex of C1 and RA domains of Nore1. *Structure* **14**, 881–888 (2006).
61. Dosset, P. *et al.* A novel interactive tool for rigid-body modeling of multi-domain macromolecules using residual dipolar couplings. *J Biomol NMR* **20**, 223–231 (2001).
62. Petoukhov, M. V. *et al.* New developments in the ATSAS program package for small-angle scattering data analysis. *J Appl Cryst* **45**, 342–350 (2012).
63. Rambo, R. P. & Tainer, J. A. Accurate assessment of mass, models and resolution by small-angle scattering. *Nature* **496**, 477–481 (2013).
64. Wriggers, W. & Chacon, P. Using Situs for the Registration of Protein Structures with Low-Resolution Bead Models from X-ray Solution Scattering. *J Appl Cryst* **34**, 773–776 (2001).
65. Pettersen, E. F. *et al.* UCSF Chimera–a visualization system for exploratory research and analysis. *J Comput Chem* **25**, 1605–1612 (2004).

Acknowledgements

We thank the staff at MacCHESS for assistance with X-ray data collection. This work was supported by a Canadian Institutes of Health Research (CIHR) postdoctoral fellowship and startup funds from McGill University and Fonds de recherche du Québec - Santé (FRQS) to J.-F.T., a CIHR grant to K.G. (MOP-14219). The McGill SPR-MS Facility thanks the Canada Foundation for Innovation (CFI) for infrastructure support. This work was supported by the Ministry of Education, Youth and Sports of the Czech Republic within the National Sustainability Program II (Project BIOCEV-FAR) LQ1604 and by the project "BIOCEV" (CZ.1.05/1.1.00/02.0109) and the Ministry of Education of the Czech Republic (program "NAVRAT" LK11205 and program "NPU I" LO1304).

Author Contributions

J.-F.T. conceived experiments, collected and analyzed crystallography, mass spectrometry, NMR and SAXS data, synthesized diubiquitin, wrote the manuscript, and prepared the figures. M.S., M.S., M.K., V.V., J.K. and K.G.Š. performed and analyzed the UBL structure determination and ITC experiments. C.D.H.R. performed NMR assignments on the HDD and purified the SAXS samples. A.H. performed NMR titrations with the UBL. M.M. performed molecular cloning. X.F. purified the RVP for crystallization. S.S. performed NMR assignments on the UBL. J.-F.T., K.G., J.K. and K.G.Š. drafted the manuscript.

Additional Information

Accession Codes: Coordinates and structure factors for the RVP crystal structure were deposited in the PDB under accession code 4Z2Z. Coordinates and restraints for the Ddi1 UBL and HDD solution NMR structures were deposited in the PDB under accession codes 2N7E (BMRB code 25803) and 5KES (BMRB code 30102), respectively.

Supplementary information accompanies this paper at http://www.nature.com/srep

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: Trempe, J.-F. *et al.* Structural studies of the yeast DNA damage-inducible protein Ddi1 reveal domain architecture of this eukaryotic protein family. *Sci. Rep.* **6**, 33671; doi: 10.1038/srep33671 (2016).

APPENDIX 3

Ramirez, J., Lectez, B., Osinalde, N., Siva, M., Elu, N., Aloria, K., Prochazkova, M., Perez, C., Martinez-Hernandez, J., Barrio, R., Saskova, K. G., Arizmendi, J. M. and Mayor, U. (2018). "Quantitative proteomics reveals neuronal ubiquitination of Rngo/Ddi1 and several proteasomal subunits by Ube3a, accounting for the complexity of Angelman syndrome". *Hum Mol Genet* **27** (11): 1955-1971. **IF (2017)= 4.902**

The full version of the article may be found online on the below link, directly on the website of the publishing journal, which holds copyrights to the article. With respect to the said copyrights and bearing in mind that the thesis hereof will be published by Charles University in its online repository of theses, the article is not attached hereto directly. An author manuscript version is attached in compliance with the cypyright rules of the journal.

https://academic.oup.com/hmg/article/27/11/1955/4948678

This is a pre-copyedited, author-produced version of an article accepted for publication in Human Molecular Genetics following peer review.

The version of record (Ramirez, J., Lectez, B., Osinalde, N., Siva, M., Elu, N., Aloria, K., Prochazkova, M., Perez, C., Martinez-Hernandez, J., Barrio, R., Saskova, K. G., Arizmendi, J. M. and Mayor, U. (2018). "Quantitative proteomics reveals neuronal ubiquitination of Rngo/Ddi1 and several proteasomal subunits by Ube3a, accounting for the complexity of Angelman syndrome". Hum Mol Genet 27 (11): 1955-1971) is available online at: https://academic.oup.com/hmg/article/27/11/1955/4948678 .

DOI: 10.1093/hmg/ddy103.

Quantitative proteomics reveals neuronal ubiquitination of Rngo/Ddi1 and several proteasomal subunits by Ube3a, accounting for the complexity of Angelman syndrome

Juanma Ramirez [1], Benoit Lectez [1], Nerea Osinalde [2], Monika Sivá [3, 4, 5], Nagore Elu [1], Kerman Aloria [6], Michaela Procházková [7], Coralia Perez [8], Jose Martínez-Hernández [1, 9], Rosa Barrio [8], Klára Grantz Šašková [3, 4], Jesus M. Arizmendi [1] and Ugo Mayor [1, 9]

[1] Department of Biochemistry and Molecular Biology, Faculty of Science and Technology, University of the Basque Country (UPV/EHU), 48940 Leioa, Spain
[2] Department of Biochemistry and Molecular Biology, Faculty of Pharmacy (UPV/EHU), 01006 Vitoria-Gasteiz, Spain
[3] Department of Genetics and Microbiology, Charles University, 12843 Prague, Czech Republic
[4] Institute of Organic Chemistry and Biochemistry of the Czech Academy of Sciences, 16610 Prague, Czech Republic
[5] First Faculty of Medicine, Charles University, 12108 Prague, Czech Republic
[6] Proteomics Core Facility-SGIKER, University of the Basque Country (UPV/EHU), 48940 Leioa, Spain
[7] Institute of Molecular Genetics of the Czech Academy of Sciences, Czech Centre for Phenogenomics and Laboratory of Transgenic Models of Diseases, Division BIOCEV, Vestec, Czech Republic
[8] Functional Genomics Unit, CIC bioGUNE, 48160 Derio, Spain
[9] Ikerbasque, Basque Foundation for Science, 48013 Bilbao, Spain

Abstract

Angelman syndrome is a complex neurodevelopmental disorder caused by the lack of function in the brain of a single gene, *UBE3A*. The E3 ligase coded by this gene is known to build K48-linked ubiquitin chains, a modification historically considered to target substrates for degradation by the proteasome. However, a change in protein abundance is not proof that a candidate UBE3A substrate is indeed ubiquitinated by UBE3A. We have here used an unbiased ubiquitin proteomics approach, the ^{bio}Ub strategy, to identify 79 proteins that appear more ubiquitinated in the *Drosophila* photoreceptor cells when Ube3a is over-expressed. We found a significantly high number of those proteins to be proteasomal subunits or proteasome-interacting proteins, suggesting a wide proteasomal perturbation in the brain of Angelman patients. We focused on validating the ubiquitination by Ube3a of Rngo, a proteasomal component conserved from yeast (Ddi1) to humans (DDI1 and DDI2), but yet scarcely characterized. Ube3a-mediated Rngo ubiquitination in fly neurons was confirmed by immunoblotting. Using human neuroblastoma SH-SY5Y cells in culture, we also observed that human DDI1 is ubiquitinated by UBE3A, without being targeted for degradation. The novel observation that DDI1 is expressed in the developing mice brain, with a significant peak at E16.5, strongly suggests that DDI1 has biological functions not yet described that could be of relevance for Angelman syndrome clinical research.

Introduction

Angelman syndrome (AS) is a rare neurodevelopmental disorder (OMIM #105830), with an estimated incidence of 1/15 000 births, characterized by a severe developmental delay, language impairment, ataxic movements, epilepsy, sleep disturbances and episodes of frequent laughter (1). In contrast to other complex syndromes that are caused by large genetic duplications/deletions, the underlying cause for AS is the loss of maternal expression in neurons of the brain of one single enzyme, UBE3A (2,3), a HECT-type ubiquitin E3 ligase (4–6). Although the deficiency of this paternally imprinted gene (7) is commonly originated by maternally inherited deletions on the 15q11-q13 chromosomal region (8), the syndrome is also caused by mutations affecting exclusively the *UBE3A* gene (9). Some of these UBE3A mutants lack the ubiquitin-ligase activity when tested *in vitro* (10,11), indicating that AS is caused by the lack of the ubiquitin ligase activity of UBE3A in neurons. Ubiquitinated substrates of UBE3A are therefore likely to be the effecting pathways of the resulting brain connectivity and/or function alterations. Interestingly, excess ligase activity of UBE3A has also been associated with autism spectrum disorders (12–14).

Ubiquitin E3 ligase enzymes catalyze the covalent attachment of ubiquitin to the lysine residues on target proteins. According to *in vitro* studies, UBE3A catalyzes the preferential attachment of K48-linked poly-ubiquitin chains (15), presumably targeting its substrates for proteasomal

degradation. Based on this premise, several UBE3A putative substrates (AIB1, Bak, Blk, Mcm7, Pbl) whose levels changed in the presence/absence of this enzyme were reported (16–20). Nevertheless, ubiquitination of those proteins by UBE3A has not yet been described. Similarly, the ubiquitination of other proposed neuronal UBE3A substrates was only validated *in vitro* (Arc, Na^+/K^+ ATPase, p27, Ring1B, Adrm1, Rpt5) or using non-denaturing immunoprecipitation approaches (Annexin A1, HHR23A, PSMD2, Ephexin5, p53) (21–29). We developed an ubiquitination assay for neuronal cell culture using a highly denaturing protocol, and showed that proteasome regulating proteins Rpn10 and Uch-L5 are ubiquitinated by Ube3a (30), both being reported to be essential for mammalian brain development (31,32). Recently, evidence is accumulating in regards to UBE3A regulating proteasomal activity (28–30,33), suggesting that UBE3A might indirectly be affecting the regulation of many other proteins targeted to the proteasome by other E3 ligases.

Interaction between an ubiquitin ligase and its substrates is transient and hard to capture *in vivo*. Additionally, the low stoichiometry at which ubiquitin modified proteins are found within the cell hinders the identification of ubiquitination substrates *in vivo*. Antibodies that specifically recognize the ubiquitin diGlycine (diGly) signature have been employed to isolate, and subsequently identify putative ubiquitination sites by mass spectrometry (MS) (34–39). Such methodology has also been recently used in a screen for putative UBE3A substrates in HEK293 cells (29). However, this approach requires digestion of proteins by trypsin prior to their isolation, preventing any orthogonal validation, which is essential, as the diGly signature is also a remnant of other ubiquitin-like (UBL) modifications, such as Nedd8 or ISG15 (35), and can even be an experimental artefact under certain conditions (40). We have developed in our lab two methodologies that have proven to be suitable for the *in vivo* analysis of ubiquitinated proteins (41). The ^{bio}Ub strategy, based on the *in vivo* biotinylation of ubiquitin (42), has recently been used in combination with quantitative shotgun proteomics to identify substrates of the E3 ligase Parkin involved in Parkinson's disease (43). The second strategy favours the isolation of GFP-tagged proteins under denaturing conditions, and was first used to screen for *Drosophila* Ube3a substrates (30).

In the present study, we have combined the ^{bio}Ub strategy with the over-expression of Ube3a to identify 79 putative neuronal Ube3a substrates. Amongst those, we noted the presence of 13 proteasome subunits or proteasome interacting proteins. We validated that Ube3a ubiquitinates the proteasomal ubiquitin receptor Rings lost (Rngo) in *Drosophila* photoreceptor neurons *in vivo*. Furthermore, we have confirmed that ubiquitination of DNA damage-inducible protein 1 homologue 1 (DDI1), the human orthologue of Rngo, is enhanced upon UBE3A over-expression in neuroblastoma cells, without being targeted for degradation. The observation that Ddi1 is highly expressed in the developing mice brain suggests that this protein has a yet uncharacterized biological function in neuronal development.

Results

Unbiased identification of Ube3a substrates in *Drosophila* neurons *in vivo*

In order to identify by MS analysis the proteins whose ubiquitination depends on Ube3a, we used the following fly lines: BirA, ^{bio}Ub, ^{bio}15B and ^{bio}A3. ^{bio}Ub flies express the $(^{bio}Ub)_6$-BirA precursor (Fig. 1A) in the *Drosophila* photoreceptor neurons under the control of the eye-specific GMR-GAL4 driver, which has been shown to be the most suitable neuronal(-like) driver for identifying low abundance proteins, and for optimizing reproducibility across samples (44). BirA control flies express just the bacterial biotinylating enzyme BirA (44). ^{bio}A3 flies are ^{bio}Ub flies over-expressing the Ube3a E3 ligase (Fig. 1B), as confirmed by immunoblotting (Fig. 1C). On the other hand, ^{bio}15B flies are ^{bio}Ub flies carrying a loss of function *Ube3a* deletion (*Ube3a^{15B} allele*) in heterozygosis. Homozygous mutant *Ube3a^{15B}* flies lack any detectable Ube3a protein (Supplementary Material, Fig. S1A), but flies carrying this allele could not be expanded in homozygosis. The *Ube3a^{15B}* heterozygous ^{bio}15B flies used for this study only show a partial reduction of Ube3a protein levels (Fig. 1C). Free BirA, indicating appropriate processing of the $(^{bio}Ub)_6$-BirA precursor, was observed for all three genotypes (Fig. 1D); no undigested forms of the precursor were found above the expected molecular size of BirA (35 kDa). Biotin

immunoblotting confirmed biotinylation and incorporation into conjugates of the GMR-GAL4-driven ectopic biotin-tagged ubiquitin in all ^{bio}Ub, ^{bio}A3 and ^{bio}15B flies (Fig. 1E). Expression of the ^{bio}Ub construct in those three fly lines did not significantly alter total ubiquitin levels, when compared to the BirA control (Supplementary Material, Fig. S1B).

Comparison of the ubiquitinated proteome of ^{bio}A3 flies and ^{bio}Ub flies should allow the identification of proteins whose ubiquitination is enhanced by Ube3a. Conversely, we would expect to have a reduction on the ubiquitination of Ube3a substrates on ^{bio}15B flies in respect to ^{bio}Ub flies (Fig. 2A). Biotinylated ubiquitin conjugates formed within the fly photoreceptor neurons were isolated using neutravidin beads, those pulldowns being performed on three biological replicates for each of the three conditions. Despite collecting whole heads, the isolated material is expected to originate just from the GMR-GAL4 expressing cells. Similar amounts of ubiquitinated proteins were eluted from the three genotypes (Fig. 2B and Supplementary Material, Fig. S2A and B). After fractionation by SDS-PAGE, each gel lane, corresponding to one sample, was cut into several slices as indicated in Supplementary Material, Fig. S2B). Protein loads from each individual gel slice were in-gel digested with trypsin and subsequently analysed by LC-MS/MS.

Similar number of ubiquitinated proteins and a high correlation of the label-free quantification (LFQ) intensity values were detected, both between replicas and across the different genotypes (Supplementary Material, Fig. S3A and B). Random LFQ values from a distribution meant to simulate expression below the detection limit (45) were imputed to those proteins for which LFQ values were not reported by the MaxQuant software on that given experiment (Supplementary Material, Fig. S3C). The experimental design was successful as evidenced from the LFQ intensity values obtained for Ube3a, which appeared highly enriched in the ^{bio}A3 sample, and significantly reduced in the ^{bio}15B sample (Fig. 2C). Western blot analysis confirmed those results and revealed that Ube3a is mostly isolated in an unmodified form (Fig. 2D), as reported previously (44).

We compared the LFQ intensities of ubiquitinated proteins on the ^{bio}A3 sample with the corresponding ^{bio}Ub values and plotted their fold-changes (*X*-axis) and significance *P*-values (*Y*-axis) as a Vulcano plot (Fig. 3). As expected, most of the proteins detected in this study displayed a ratio close to one, including endogenously biotinylated proteins acetyl-CoA carboxylase (ACC), pyruvate carboxylase (PCB) and CG2118 (46) (shown with filled squares in Fig. 3), indicating that the pull-down process was equally efficient in the different samples. From the 751 protein groups identified across all genotypes, 79 were significantly ($P < 0.05$) enriched at least 2-fold in the ^{bio}A3 sample relative to the ^{bio}Ub control, and can therefore be defined as putative Ube3a substrates (Supplementary Material, Table S1). In order to focus on the highest confidence Ube3a substrates, the 79 proteins regulated by Ube3a were analysed at the peptide level to comply with the following requirements: (1) Average ratio between common peptides identified in both conditions should be at least two after the subtraction of the standard error of the mean (S.E.M.); and (2) individual peptide intensities should globally follow the same tendency as the protein LFQ intensity and show in average a 2-fold enrichment. Out of the 79 proteins identified as enriched at the protein level, 39 candidate Ube3a substrates (labelled with filled circles in Fig. 3) appeared homogenously enriched also at the peptide level (Supplementary Material, Fig. S3D), most of them having human orthologues (Table 1). Among those 39 high confidence candidate Ube3a substrates, eight proteins regulate protein degradation through the ubiquitin-proteasome system, including proteasomal subunits Prosα1, Prosα3, Rpt2, Rpt4, Rpn3 and Rpn8, and the proteasomal shuttling proteins Rngo and Rpn10, the latest being previously identified as target of Ube3a in neuronal cell culture (30). Besides, two proteins related to autophagy, Atg8a and Ref(2)P (47), and three chaperone proteins, CCT3, CCT7 and CCT8 (48), were also found as high confidence substrates of Ube3a. In addition, proteasomal subunits Prosα4, Prosα7, Rpn2, proteasome activator REG and the proteasome-associated deubiquitinating (DUB) enzyme Uch-L5 were significantly enriched according to LFQ values, but did not pass the requirements at peptide level.

On the other hand, a total of 55 proteins also appeared to be less ubiquitinated in ^{bio}A3 sample due to the over-expression of Ube3a (Supplementary Material, Table S1), displaying a 2-fold significant reduction in their abundance compared to ^{bio}Ub (^{bio}A3/^{bio}Ub < 0.5; $P < 0.05$). Following the same peptide analysis described before (Supplementary Material, Fig. S3E), we concluded that 14 proteins are clearly less ubiquitinated when Ube3a is over-expressed, as compared to the control sample (labelled with empty circles in Fig. 3). Interestingly, we reliably identified calcium/calmodulin-dependent protein kinase II (CaMKII)—which in *Drosophila* is coded by a single gene—as being less ubiquitinated upon over-expression of Ube3a.

We also compared the LFQ intensities of the ^{bio}15B sample with the ^{bio}Ub values; but, using the same criteria as above, Ube3a itself did not appear significantly enriched with high confidence in the ^{bio}Ub sample relative to the ^{bio}15B ubiquitinated material. However, three proteins (Rdhb, Map2015 and Axo) appeared significantly less ubiquitinated in the ^{bio}15B sample, barely above the defined thresholds (data not shown).

Ubiquitin LFQ values and ubiquitin chain linkages indicate proteasomal perturbation upon Ube3a over-expression

Based on the experimental design, we expected ubiquitin levels to be unaltered by Ube3a over-expression. Furthermore, no significant differences were observed on the hard-to-quantify smears of the silver stained gels (Fig. 2B), as well as ubiquitin (Supplementary Material, Fig. S1B) and biotin (Supplementary Material, Fig. S2A) blots. However, based on LFQ values, ubiquitin levels were quantified to have a significant increase of 1.66-fold in the ^{bio}A3 flies, relative to the ^{bio}Ub sample (Supplementary Material, Table S1), this is, a >60% increase of ubiquitin was detected on the collected ubiquitinated material of ^{bio}A3 flies. In order to elucidate the type of ubiquitin chains enriched upon over-expression of Ube3a, we compared the intensity of all detected diGly-containing peptides across the samples. Digestion of a complex mixture of ubiquitinated proteins is expected to result in a complex mixture of ubiquitin chain linkages. Even though the isolated ubiquitinated material is composed of substrates of hundreds of ligases, this analysis indicated that K48 and K63 linkages were significantly more abundant on the ^{bio}A3 sample, while K33 linkages were reduced (Supplementary Material, Fig. S3F). The 60% increase on ubiquitin levels together with the significant changes in chain linkages suggests a global perturbation of the ubiquitin proteasome system upon Ube3a over-expression.

Rngo is ubiquitinated by Ube3a *in vivo* in *Drosophila* neurons

Amongst the highest confidence candidates as Ube3a substrates (Table 1), eight of the identified proteins are proteasome integral or regulatory subunits. We decided to focus on Rngo, a predicted proteasomal shuttling factor for which working antibodies were available (49), and which was already observed in our first neuronal ubiquitome studies (42). According to the label-free quantitative MS-based analysis performed, Rngo is 5-fold more enriched in Ube3a over-expressing flies in respect to control flies. In agreement with that, western blot analysis showed that ubiquitination of Rngo is enhanced when Ube3a is over-expressed and reduced in its absence (Fig. 4A), making Rngo the first Ube3a substrate validated in any type of neuron *in vivo*. Interestingly, Rngo total levels were not altered, suggesting that it is not being targeted for degradation. As a control, the same membrane was used to detect the presence of Fax, a protein that we found by MS to be less ubiquitinated upon Ube3a over-expression (Fig. 3). Indeed, Fax displayed the opposite trend, and its ubiquitination was dramatically reduced by the over-expression of Ube3a (Fig. 4A), indicating that the increase ubiquitination seen for Rngo is specific to the protein and not to a more efficient general isolation of proteins in the ^{bio}A3 sample. We also analysed by western blot the ubiquitination of another proposed UBE3A substrate (27), the Na^+/K^+-ATPase α-subunit (Atpα), but as with the MS analysis (Fig. 3), western blot did not show any enhancement by Ube3a over-expression on its ubiquitination (Fig. 4B).

Ddi1 and Ddi2 are expressed throughout development in the mouse brain

Mammalians have two protein homologues to *Drosophila* Rngo: DNA damage-inducible protein 1 homologue 1 (DDI1) and homologue 2 (DDI2). Partial expression data for DDI2 are available at the Human Protein Atlas (www.proteinatlas.org; date last accessed March 19, 2018), but

information about mammalian DDI1 is very scarce. We tried to assess the expression of DDI1 using commercially available antibodies, but, since DDI1 and DDI2 proteins have very high amino acid sequence identity (72%), all the commercially available antibodies recognize DDI2 protein, or do simply not work (data not shown). We therefore tested whether mouse Ddi1 and Ddi2 are expressed in the brain. Analysis of Ddi1 gene expression in mice brain at different ages performed by qRT-PCR revealed a drastic and significant increase of Ddi1 mRNA levels at embryonic stage E16.5 (Fig. 5A). The Ddi1 mRNA expression peak was rapidly reduced from E17.5, returning to basal levels at E19.5-P1, remaining relatively stable during all the tested adult time points. In contrast, Ddi2 mRNA levels, which were also detectable in the mice brain, only fluctuated slightly during development (Fig. 5B).

Additionally, we analysed the expression profile of Ddi1 in the developing brain by performing RNA *in-situ* hybridization experiments in CD-1 mouse embryos at four different stages of development—E9.5, E10.5, E14.5 and E16.5. As shown in Figure 6, Ddi1 is expressed in all parts of developing brain (telencephalon, diencephalon, mesencephalon and rhombencephalon) at the two younger developmental stages. Similar Ddi1 expression pattern was observed for these two stages in both whole mount samples (Fig. 6A and C) and sagittal paraffin sections (Fig. 6B and D). At the stage E14.5, the expression is located in neurons of mesencephalic and telencephalic structures, with signal accumulation in upper hill (colliculus tectum) and ventricular zone of pallium (Fig. 6E–G). At further developmental stage E16.5, which shows highest mRNA level expression according to our qRT-PCR screen, Ddi1 is expressed in particular telencephalon parts, mainly the ventricular layer and cortical plate of isocortex and the ventricular layer of olfactory bulb (Fig. 6H and I). This could represent the dividing neuroblasts and their migration toward the superficial layer of isocortex. The positive staining of neuronal cells in the isocortex (Fig. 6J) is clear evidence of Ddi1 expression in neuronal tissue, however, up to date no single report has described the role DDI1 could exert in the brain.

Human DDI1 is ubiquitinated by UBE3A, but is not targeted for degradation

Once identified and confirmed that Ube3a ubiquitinates Rngo in flies, we aimed to test whether its homologs DDI1 and DDI2 are substrates of human UBE3A. For that purpose, we employed an *in cellulo* ubiquitination assay (30,50) in SH-SY5Y neuroblastoma cells. Wild-type UBE3A ($UBE3A^{WT}$) induced the ubiquitination of DDI1-GFP in SH-SY5Y cells (Fig. 7 and Supplementary Material, Fig. S5A), but did not seem to increase the ubiquitination status of DDI2-GFP (Supplementary Material, Fig. S5B). We also confirmed the specificity of UBE3A-dependent DDI1 ubiquitination by proving that Parkin, the E3 ligase involved in Parkinson's disease, could not mediate ubiquitination of DDI1 (Supplementary Material, Fig. S5B). Altogether, the data presented here demonstrates for the first time that Rngo human homologue DDI1 is a UBE3A substrate. Despite earlier reports suggested that UBE3A generates degradation-leading K48 ubiquitin linkages *in vitro* (15), at least in SH-SY5Y neuroblastoma cells, UBE3A-mediated ubiquitination of human DDI1 does not lead to its degradation, as indicated by the intensity of the DDI1-GFP bands, which is independent of the activity of UBE3A (Fig. 7).

Discussion

Our unbiased proteomic analysis for the identification of differentially ubiquitinated proteins in *Drosophila* photoreceptor neurons upon Ube3a over-expression has resulted in a list of 79 putative Ube3a substrates, out of which proteasomal proteins appear very highly enriched. To our knowledge, this is the first time a list of candidate Ube3a substrates, whose ubiquitinated fraction is enhanced upon Ube3a over-expression, is reported in neurons *in vivo*. Several of the putative Ube3a substrates identified (Arc1, Chc, Gclc, GlyRS, Path, Tig and SesB) play a role either in axon and dendrite morphogenesis (51–54) or synaptic transmission (55–57). Interestingly, upon Ube3a over-expression we also found a significant reduction of the ubiquitination of CaMKII, a key kinase known in humans to regulate neurotransmitter synthesis and release, modulation of ion channel activity, neurite extension, synaptic plasticity, learning and memory (58). Reduced activity and protein levels of CaMKII at the postsynaptic density have been described in a mouse

model of AS (59). It still remains to be explained how a reduction of UBE3A levels enhances the ubiquitination and/or degradation of CaMKII.

Analysis of the putative Ube3a substrates by g:Profiler analysis of GO terms and KEGG pathways indicate a highly significant enrichment of the proteasome (data not shown). Furthermore, comparison of the data obtained in this work to the dataset of Parkin substrates (43), further confirms a deregulation of the proteasome upon Ube3a over-expression. Indeed, UBE3A is known to be a proteasome-associated protein (28,60), as well as being capable of ubiquitinating several proteasomal subunits in cell culture (28–30). However, it is controversial whether UBE3A inhibits (28) or stimulates (33) the proteolytic activity of the proteasome. Ube3a over-expression results in an increase in total ubiquitin levels, while over-expression of Parkin did not. Having identified several proteasomal subunits as ubiquitinated by Ube3a, the simplest explanation is that Ube3a-driven proteasomal deregulation results in the accumulation of substrates ubiquitinated by Ube3a or other E3 ligases. This would also explain the reduction observed here in the ubiquitination of many other proteins. It is known that the specific deregulation of the proteasome by over-expressing a dominant negative Rpn10 subunit reduces the ubiquitination levels of Fax and other mono-ubiquitinated proteins (44). Deubiquitination of mono-ubiquitinated proteins is also seen when inhibiting the proteasome pharmacologically (35), and is explained by a reduction of the free ubiquitin available pool concomitant to the accumulation of proteasome targeted poly-ubiquitinated proteins (36,61–63). The decrease on Fax ubiquitination upon Ube3a over-expression, detected here by both MS (Fig. 3) and immunoblotting (Fig. 4A), could therefore be caused by an Ube3a-induced proteasomal inhibition. Significant changes in global ubiquitin chain linkages (Supplementary Material, Fig. S3F) are also in line with this interpretation.

In this work, we have validated by western blot that Rngo is a direct Ube3a substrate in photoreceptor neurons. Rngo contains both an UBL and an Ubiquitin-binding domain (UBD), a hallmark of proteasomal shuttles (49), but can also bind directly another proteasomal shuttle protein, Rpn10 (49). The ubiquitination of Rpn10 is also increased in Ube3a over-expressing flies (Fig. 3), and was already identified as an Ube3a substrate in *Drosophila* cells, as well as being shown to interact genetically with Ube3a in neurons *in vivo* (30). Ubiquitination of such proteasomal regulators by Ube3a can be predicted to severely interfere with proteasomal function. Considering that the proteasome regulates dendritic development (32), long-term potentiation (64), long-term depression (65), synaptic plasticity (66), synaptic strengthening (67), memory consolidation (68), circadian rhythms (69) and many other aspects of neuronal function, an UBE3A-dependent proteasomal regulation could easily explain how a single E3 ligase mutation can cause a disorder as complex as AS. And indeed, it should be noted that UBE3A has been shown to regulate most aspects of neuronal function listed above (70–74). Non-degradative ubiquitination of proteasomal receptors could alter their function to a similar extent as their degradation, since the activity of those receptors is dependent on their UBDs. An ubiquitinated ubiquitin-binding subunit is likely to prioritirily bind its own ubiquitin moieties, therefore blocking its normal function. Our prediction would be that—based on the increased ubiquitination of those subunits upon UBE3A over-expression—proteasome activity should be increased in the brain of AS patients, on which UBE3A levels are reduced.

Interestingly, Rngo is also the *Drosophila* homologue of yeast Ddi1/Vsm1, a protein that binds to several Snc-interacting t-SNAREs (75), negatively regulates exocytosis (76) and also regulates protein secretion (77). Similarly, the *C. elegans* homologue DDI-1/VSM-1 has also been proposed to regulate synaptic function, with *vsm-1* mutants displaying a significant increase in synaptic density along the dorsal nerve cord (78). Thus, Rngo/DDI1 may have an additional role in synaptic transmission by controlling SNARE mediated exocytosis. In fact, over-expression of Ube3a, but not its ligase dead form, has been reported to alter neurotransmission at the neuromuscular junction in *Drosophila* (79).

Two mammalian Rngo homologues have been described, DDI1 and DDI2, but neither of those proteins have to date been functionally characterized. We have found that both are expressed in the developing brain (Figs 5 and 6). Structurally, both proteins contain the characteristic Retroviral Protease-like domain that was recently reported to cleave/activate the Nrf1

transcription factor under proteasome inhibition (80,81). Further, they both contain an additional helical HDD domain (82,83) and an N-terminal UBL domain, but lack the well-defined Ubiquitin-associated (UBA) domain. Although, a weak Ubiquitin-binding motif (UIM) is present at the C-terminus of DDI2 (82), no such a motif can be found in DDI1. It is thus feasible that the DDI1 UBL domain, which in yeast Ddi1 is capable of binding ubiquitin (84), could substitute the role of the UBA domain, a mechanisms that would be facilitated by DDI1's homodimeric conformation.

Having confirmed the ubiquitination of Rngo in photoreceptor neurons by the *Drosophila* Ube3a and the expression of its mammalian homologues in the brain, we then tested whether any of Rngo's human orthologues are ubiquitinated by UBE3A. We confirmed DDI1 to be an ubiquitination substrate of UBE3A in SH-SY5Y neuroblastoma cells. Different controls indicated that ubiquitination of DDI1 by UBE3A is specific, but does not lead to a reduction of DDI1 protein levels. It is not the first time that regulation of protein activity, even with formation of K48-linked chains, has been reported to be proteolysis independent (85). A UBD found in Met4 was proposed to cap its own K48-linked ubiquitin chain, inactivating the protein and protecting it from degradation (86), and something similar could be happening with Rngo/DDI1. Further work is required to elucidate what the functional role of DDI1 ubiquitination might be and whether the presence of UBD/UIM/UBA/UBL in ubiquitination substrates might interfere into the canonical role of K48 ubiquitin-linked chains.

Up to now, no candidate ubiquitination substrates of UBE3A had been directly validated in neurons *in vivo*. We had earlier identified a proteasomal shuttling factor, Rpn10, to be regulated by Ube3a in *Drosophila* cells (30), which we now confirmed *in vivo* by MS. Further, we have validated that in *Drosophila* photoreceptor neurons Ube3a ubiquitinates another proteasomal shuttling factor, Rngo, becoming the first Ube3a substrate to be identified and validated *in vivo* in neurons within a whole organism. More importantly, UBE3A regulates the orthologue DDI1 protein in human neuroblastoma cells. Since UBE3A ubiquitination appears to be regulating several proteasomal-associated subunits, as already indicated by ourselves and others (29,30,33), it would not be surprising to see that a highly significant number of proteins regulated downstream the proteasome will display significant changes on their abundance upon UBE3A mutation or over-expression. It is now a challenge to elucidate which Ube3a substrates are direct, in addition to the proteasomal proteins themselves.

DDI1 gene has been reported to be affected in siblings of a familial neurodegenerative disorder characterized clinically as a variant of Alzheimer's disease (87). Having now described for the first time the temporal and spatial expression of Ddi1 in the mouse brain, we next need to perform a functional characterization to uncover the neuronal role of DDI1, as this might as well bring light to our understanding of how AS is regulated. Based on studies in yeast and *C. elegans*, it is likely that DDI1 is involved in both regulation of synapses and proteasomal function (75–78,84). Given the complexity of neuronal function, and previous work aiming to identify the mechanisms regulated by UBE3A, it is likely that both processes are actually misregulated during the genesis of AS. Finally, if we take into account that Ube3a regulates the proteasome, and that UBE3A expression declines with age (88), it would not be surprising that this E3 ligase has a further role in proteostasis not yet characterized.

Materials and Methods

Drosophila stocks and sample collection

We have used in this work the BirA, ^{bio}Ub, ^{bio}A3 and ^{bio}15B flies, all of which express their corresponding constructs in the *Drosophila* photoreceptor neurons under the control of the GMR-GAL4 driver. BirA and ^{bio}Ub flies, expressing respectively the BirA enzyme alone and the $(^{bio}Ub)_6$-BirA precursor, have been described previously (44). Their genotypes are respectively *GMR-GAL4/CyO*; *UAS-BirA/TM6* and *GMR-GAL4*, *UAS-*$(^{bio}Ub)_6$*-BirA/CyO*. Ube3a gain of function (*UAS-Ube3a*^{A3}) and loss of function (*Ube3a*^{15B}) flies (89) were a gift from Professor Janice Fisher. Both *UAS-Ube3a*^{A3} and *Ube3a*^{15B} fly lines were independently mated to *GMR-GAL4*, *UAS-*$(^{bio}Ub)_6$*-BirA/CyO*; *TM2/TM6* flies to generate *GMR-GAL4*, *UAS-*$(^{bio}Ub)_6$*-BirA/CyO*;

UAS-Ube3a^{A3}/TM6 and *GMR-GAL4*, *UAS-(^{bio}Ub)$_6$-BirA/CyO*; *Ube3a^{15B}/TM6* lines. The *Drosophila* Ube3a mutants (*Ube3a^{15B}*) had been reported to be viable and fertile in homozygosis (89). When combined with ^{bio}Ub flies, it was, however, required to grow them in heterozygosis, as null Ube3a flies were viable but not fertile at 25°C (J. Ramirez and U. Mayor, unpublished data). GMR-GAL4 (BL 1104) and *OregonR* (BL 2376) flies were provided by the Bloomington *Drosophila* Stock Center (Bloomington, IN, USA). The ^{bio}Ub abbreviation is used throughout the text to refer to the *GMR-GAL4*, *UAS-(^{bio}Ub)$_6$-BirA/CyO* flies, and the ^{bio}A3 and ^{bio}15B (for *Ube3a^{A3}* and *Ube3a^{15B}* alleles) to refer to *GMR-GAL4*, *UAS-(^{bio}Ub)$_6$-BirA/CyO*; *UAS-Ube3a^{A3}/TM6* and *GMR-GAL4*, *UAS-(^{bio}Ub)$_6$-BirA/CyO*; *Ube3a^{15B}/TM6* flies, respectively.

Flies were grown at 25°C in 12 h light-dark cycles in standard *Drosophila* medium (0.9% agar, 7.5% dextrose, 6% corn flour, 8.5% yeast, 2.5% Nipagin, 0.4% propionic, 0.02% benzalkonium chloride in distilled H_2O). Mixed-sex flies of 2–5 days old were flash-frozen in liquid nitrogen and shaken while still frozen to sever the heads. Frozen fly heads were then separated from the remaining body parts using a pair of sieves with a nominal cut-off of 710 and 425 μm, and then stored at −80°C. Head collections were typically performed in the morning.

Plasmids

Commercial pEGFP-N1 vector (Clontech) was used to generate DDI1-GFP and DDI2-GFP vectors (see Cloning procedures section). *FLAG-UBE3A-pCMV* (UBE3A^{WT}) and *FLAG-UBE3A^{LD}-pCMV* (UBE3A^{LD}) plasmids, expressing N-terminally FLAG-tagged versions of the wild type and catalytically inactive human UBE3A protein (90), were a gift from Dr Vjekoslav Tomaić. FLAG-tagged ubiquitin (30) in pCDNA3.1 vector (FLAG-Ub) was generously provided by Dr Jose Antonio Rodriguez Pérez (University of the Basque Country-UPV/EHU, Spain). Untagged human Parkin plasmid has been described previously (43). Empty pCDNA3.1 vector (Invitrogen) was used as control.

Cloning procedures

DDI2-pEGFP-N1 (DDI2-GFP) plasmid was generated by amplifying *DDI2* gene (Uniprot Q5TDH0) from DDI2-pET16b plasmid (82) with DDI2-Fw (5′-*AAGGTACCATGCTGCTCACCGTG-3′*) and DDI2-Rv (5′-*AAGGATCCCCTGGCTTCTGACGCTCTGC-3′*) primers and inserted between *KpnI* and *BamHI* restriction sites of pEGFP-N1 vector (Clontech). Gene for human DDI1 protein (Uniprot Q8WTU0) was synthesized by GenScript and further amplified using the *DDI1-Fw* (5′-*TATAGGTACCATGCTGATCACCGTG-3′*) and *DDI1-Rv* (5′-*TATAACCGGTATGCTCTTTTCGTCC-3′*) primers and inserted between the *Acc65I* and *AgeI* sites of the DDI2-pEGFP-N1 vector, after the *DDI2* gene had been removed using the same restriction enzymes. All PCR reactions were carried out with Phusion High-Fidelity DNA polymerase (Thermo Scientific). PCR product gel extractions and plasmid purifications were performed with the QIAGEN Gel Extraction Kit and QIAGEN plasmid mini and midi kits, respectively. Correct sequence for all plasmids was confirmed by sequencing either by the GATC Biotech Company (Köln, Germany) or the SGIKER Unit of Sequencing and Genotyping at the University of the Basque Country (Leioa, Spain).

Western blotting and silver staining

Both 4–12% Bolt Bis–Tris Plus pre-cast gels (Invitrogen) and 4–12% NuPAGE Bis–Tris gels (Invitrogen) were used for SDS-PAGE, then proteins were transferred to PVDF membranes using the iBlot system (Invitrogen). Following primary and secondary antibody incubation, membranes were developed with an ECL kit (Biorad Clarity). Dual-colour westerns were prepared by assigning independent colour channels to two independent westerns developed in the same membrane. The amount of material loaded for western blot analysis varied according to the tissue and the antibody employed. In the case of material obtained from fly biotin pulldowns, between 0.001% and 0.2% of the input samples and 5–10% of the elution samples were loaded. However, when material purified from cells was used, between 10–20% of inputs and 10–40% of the elution samples were loaded.

The following primary antibodies were used: goat anti-biotin-horseradish peroxidase (HRP)-conjugated antibody (Cell Signalling; catalogue number 7075) at 1: 1000; chicken polyclonal anti-BirA antibody (Sigma; catalogue number GW20013F) at 1: 1000; mouse monoclonal anti-GFP antibody (Roche Applied Science; catalogue number 11814460001) at 1: 1000; mouse monoclonal anti-FLAG M2-HRP conjugated antibody (Sigma; catalogue number A8592) at 1: 1000; mouse monoclonal anti-Syx1A antibody (Developmental Studies of Hybridoma Bank; DSHB; catalogue number 8C3) at 1: 100; rabbit polyclonal anti-Fax antibody, a gift from Eric Liebl (Deninson University, OH, USA) at 1: 1000; mouse monoclonal anti-Atpα antibody (DSHB; catalogue number α5) at 1: 50; rabbit polyclonal anti-Ube3a antibody (91) for the detection of *Drosophila* Ube3a protein at 1: 1000; rabbit polyclonal anti-Rngo antibody (49) at 1: 500; rabbit polyclonal anti-ubiquitin antibody (Sigma; catalogue number U5379) at 1: 100; mouse monoclonal anti-UBE3A (clone E6AP-300) antibody (Sigma; catalogue number E8655) for the detection of human UBE3A protein at 1: 1000. The following secondary antibodies were used: goat anti-mouse-HRP-labelled antibody (Thermo Scientific; catalogue number 62-6520) at 1: 4000; goat anti-rabbit-HRP labelled antibody (Cell Signalling; catalogue number 7074) at 1: 4000 and donkey anti-chicken-HRP labelled antibody (Jackson ImmunoResearch; catalogue number 703-035-155) at 1: 2000.

About 10% of the neat elution samples were used for silver staining analysis. Gels were fixed for 1 h at room temperature with 40% methanol and 10% acetic acid containing solution and then were stained using the SilverQuest kit from Invitrogen according to manufacturer's instructions.

Biotin pulldown

Biotin pulldowns (42) from *Drosophila* heads were performed as described previously (43,44,92). About 500 mg of 2–5 days old fly heads of each genotype were homogenized in 2.9 mL of Lysis buffer (8 M urea and 1% SDS in PBS) supplemented with 50 mM N-ethylmaleimide (Sigma) and a complete protease inhibitor cocktail (Roche Applied Science). Lysates were centrifuged for 5 min at 16000*g* at 4°C and supernatant applied to a PD10 desalting column (GE Healthcare) previously equilibrated with 25 ml of binding buffer (3 M urea, 1 M NaCl, 0.25% SDS and 50 mM N-ethylmaleimide). Eluates, except 50 µl that were kept for monitoring the inputs, were then incubated with 250 µl of NeutrAvidin agarose beads suspension (Thermo Scientific). Unbound material (flow through) was separated by spinning the beads at 230*g* for 2 min. Beads were then subjected to stringent washes with six different washing buffers (WB): twice with WB1 (8 M urea, 0.25% SDS), thrice with WB2 (6 M guanidine-HCl), once with WB3 (6.4 M urea, 1 M NaCl, 0.2% SDS), thrice with WB4 (4 M urea, 1 M NaCl, 10% isopropanol, 10% ethanol, 0.2% SDS), once with WB1, once with WB5 (8 M urea, 1% SDS) and thrice with WB6 (2% SDS). All buffers were prepared in PBS. Beads were then heated at 95°C for 5 min in 125 µl of elution buffer (250 mM Tris–HCl pH 7.5, 40% glycerol, 4% SDS, 0.2% BPB, 100 mM DTT) and centrifuged for 2 min at 16 000*g* in a Vivaclear Mini 0.8 µm PES micro-centrifuge filter unit (Sartorius) to recover the eluted proteins. Finally, eluates were concentrated in Vivaspin 500 centrifugal filter units (Sartorius).

Cell culture and transfection

Human neuroblastoma SH-SY5Y cells were cultured under standard conditions (37°C, 5% CO_2) in Dulbecco's modified Eagle medium/nutrient mixture F-12 (DMEM/F-12) with GlutaMAX (Thermo Scientific), supplemented with 10% fetal bovine serum (Thermo Scientific), 100 U/ml of penicillin (Invitrogen) and 100 µg of streptomycin (Invitrogen). SH-SY5Y cells (3×10^5 cells) were seeded in six well-plates for transfection experiments. Overnight incubation under serum starvation was routinely performed prior to transfections. The following day OptiMEM serum-free medium (Thermo Scientific) was replaced by fresh DMEM/F-12 and cells were co-transfected with 1 µg of *FLAG-Ub* and 1 µg of *DDI1-GFP*, or *DDI2-GFP*, for 72 h using Lipofectamine 3000 transfection reagent (Invitrogen) according to manufacturer's instructions. About 1 µg of either *pcDNA3.1* (control), $UBE3A^{WT}$, $UBE3A^{LD}$ or *Parkin* plasmids were additionally added to the transfection mixture to check the effect of UBE3A in DDI1 and DDI2 ubiquitination. Cells were washed twice in PBS and stored at −20°C until required.

GFP beads pull-down assay

Transfected SH-SY5Y cells were lysed with 500 µl of lysis buffer (50 mM Tris–HCl pH 7.5, 150 mM NaCl, 1 mM EDTA, 0.5% Triton, 1× protease inhibitor cocktail from Roche Applied Science and 50 mM N-ethylmaleimide from Sigma) and centrifuged at 14 000*g* for 10 min. Supernatants were mixed with 25 µl of GFP-Trap-A agarose beads suspension (Chromotek GmbH), which had been previously washed twice with a Dilution buffer (10 mM Tris–HCl pH 7.5, 150 mM NaCl, 0.5 mM EDTA, 1× protease inhibitor cocktail, 50 mM N-ethylmaleimde). The mixture was then incubated at room temperature for 150 min with gentle rolling and centrifuged for 2700*g* for 2 min to separate the beads from the unbound material. GFP beads were subsequently washed once with the dilution buffer, thrice with washing buffer WB5 (8 M urea, 1% SDS in PBS) and once with 1% SDS in PBS. Bound GFP-tagged proteins were eluted in 25 µl of elution buffer (250 mM Tris–HCl pH 7.5, 40% glycerol, 4% SDS, 0.2% BPB, 100 mM DTT) by heating at 95°C for 10 min.

In-gel trypsin digestion and peptide extraction

Eluates from biotin pull-down assays were resolved by SDS-PAGE using 4–12% Bolt Bis–-Tris Plus pre-cast gels (Invitrogen) and visualized with Colloidal Blue following manufacturer's instructions (Invitrogen). When processing biotin pull-down samples, each gel lane was cut into seven slices (see Supplementary Material, Fig. S2B). Based on earlier experiments and the BirA control, we excluded the intense bands corresponding to avidin monomers, dimers and an endogenously biotinylated protein from further analysis. The remaining four slices were subjected to in-gel digestion as described previously (93). Briefly, proteins were reduced and alkylated by incubating with DTT and chloroacetamide, respectively. Protein digestion was performed by saturation of the gel pieces with trypsin and overnight incubation at 37°C. Resulting peptides were extracted from the gel, dried down in a vacuum centrifuge and stored at −20°C. Peptide mixture was resuspended in 0.1% formic acid previous to the LC-MS/MS analysis.

LC-MS/MS analysis

Mass spectrometric analyses were performed on an EASY-nLC 1000 liquid chromatography system interfaced with a Q Exactive mass spectrometer (Thermo Scientific) via a nanospray flex ion source. Peptides were loaded onto an Acclaim PepMap100 pre-column (75 µm × 2 cm, Thermo Scientific) connected to an Acclaim PepMap RSLC (50 µm ×x 15 cm, Thermo Scientific) analytical column. Peptides were eluted from the column using a linear gradient of 2 to 40% acetonitrile in 0.1% formic acid at a flow rate of 300 nL min^{-1} over 45 min. The mass spectrometer was operated in positive ion mode. Full MS scans were acquired from *m/z* 300 to 1850 with a resolution of 70 000 at *m*/*z* 200. The 10 most intense ions were fragmented by higher energy C-trap dissociation with normalized collision energy of 28 and MS/MS spectra were recorded with a resolution of 17 500 at *m*/*z* 200. The maximum ion injection time was 120 ms for both survey and MS/MS scans, whereas AGC target values of 3×10^6 and 5×10^5 were used for survey and MS/MS scans, respectively. In order to avoid repeat sequencing of peptides, dynamic exclusion was applied for 45 s. Singly charged ions or ions with unassigned charge state were also excluded from MS/MS. Data were acquired using Xcalibur software (Thermo Scientific).

Data processing and bioinformatics analysis

Acquired raw data files were processed with the MaxQuant (94) software (version 1.5.3.17) using the internal search engine Andromeda (95) and searched against the UniProt database restricted to *Drosophila melanogaster* entries (release 2015_11; 43712 entries). Spectra originated from the different slices corresponding to the same biological sample were combined. Carbamidomethylation (C) was set as fixed modification whereas Met oxidation, protein N-terminal acetylation and Lys GlyGly (not C-term) were defined as variable modifications. Mass tolerance was set to 8 and 20 ppm at the MS and MS/MS level, respectively. Enzyme specificity was set to trypsin, allowing for cleavage N-terminal to Pro and between Asp and Pro with a maximum of two missed cleavages. Match between runs option was enabled with 1.5 min match

time window and 20 min alignment window to match identification across samples. The minimum peptide length was set to seven amino acids. The false discovery rate for peptides and proteins was set to 1%. Normalized spectral protein label-free quantification (LFQ) intensities were calculated using the MaxLFQ algorithm.

Data analysis and statistical tests

MaxQuant output data was analysed with the Perseus module (version 1.5.6.0) (45). Initially, proteins only identified by site, contaminants, reverse hits and proteins with no unique peptides and/or no intensity were removed. Missing LFQ intensity values were replaced with values from a normal distribution (width 0.3 and down shift 1.8), meant to simulate expression below the detection limit (45). To determine statistically significant changes in protein abundance, as well as in ubiquitin diGly peptides, two-tailed Student's *t*-test was used.

In the analysis of the biotin pulldowns, two comparisons were carried out: ^{bio}A3 versus ^{bio}Ub (i.e. Ube3a gain of function versus control) and ^{bio}15B versus ^{bio}Ub (i.e. Ube3a loss of function versus control). Proteins displaying a LFQ fold change bigger than 2 with a *P*-value smaller than 0.05 were selected for further analysis. The selected proteins were further filtered based on the intensity pattern observed for their peptides. Statistical significance in western blotting semi-quantification was evaluated using an analysis of variance (ANOVA) complemented by Tukey's honest significance difference test (Tukey's HSD) performed in GraphPad PRISM software.

RNA extraction, reverse transcription and real-time qPCR

Total RNA was isolated from 30 mg of brain from embryos, young and adult wild-type mice from different ages. RNA was extracted and further purified by using the RNeasy Mini Kit (QIAGEN) according to the manufacturer's protocol. Contaminating genomic DNA was removed by treatment with deoxyribonuclease I (QIAGEN), and cDNAs were synthesized from 1 µg RNA using the AffinityScript Multi Temperature cDNA Synthesis Kit (Agilent Technologies). Quantitative RT-PCR was performed on cDNA in the presence of Power SYBR Green PCR Master Mix (Applied Biosystems) containing preset concentrations of deoxynucleotide triphosphates and with specific primers, using the ABI Prism 7900 sequence Detection System (Applied Biosystems). PCR parameters were 50°C for 2 min, 95°C for 10 min, 40 cycles at 95°C for 15 s and 60°C for 1 min. The purity of the PCR products was assessed by dissociation curves. The amount of target cDNA was calculated by the comparative threshold (C_t) method and expressed by the $2^{-\Delta\Delta Ct}$ method according to Applied Biosystems' instructions, using glyceraldehyde-3-phosphate dehydrogenase (GAPDH) as an internal control. Expression of GAPDH mRNA was not affected by age, and the ratio of ΔC_t value did not vary with the amount of cDNA. Each primer set was used at its optimal concentration (300 nM) with maximal efficacy. It was verified that one single specific product was amplified as shown by analysis of its melting temperature value.

Primers GAPDH—Forward: 5′-ACCACAGTCCATGCCATCAC-3′

Reverse: 5′-TCCACCACCCTGTTGCTGTA-3′

Primers DDI1—Forward: 5′-TCACTGTGTATTGTGTGCGTAG-3′

Reverse: 5′-AGCTGTTCCATGTAAACGATCTG-3′

Primers DDI2—Forward: 5′-CCTCTCCGAGGTGACCTTTTC-3′

Reverse: 5′-GGCCTTTCTGCATAGACAATCT-3′

RNA *in situ* hybridization

CD-1 mice were mated overnight, and the presence of a vaginal plug indicated embryonic day (E) 0.5. Unsexed embryos at E9.5 and E10.5 (the age 9.5 and 10.5 days post coitum) were fixed in 4% PFA and used for *in situ* hybridization, both in whole mount and in paraffin sections. Heads of 14.5 and 16.5 days post coitum old CD-1 embryos were fixed in 4% PFA and further processed for paraffin sections. Embryo samples for whole mount ISH were frozen in methanol at −20°C prior tissue hydration, proteinase K treatment, acetylation and the prehybridization and

hybridization procedures. The embryos and heads used for sectioning were dehydrated, embedded in paraffin, cut to 7 µm sections and rehydrated prior further treatment. The experimental procedures were performed according to standard protocols (96).

Murine Ddi1 coding sequence (NM_027942.1) was synthesized by GenScript and cloned into pGEM-T® easy plasmid. Linearized plasmid was purified with PCR Purification Kit (QIAGEN) and used for generation of digoxigenin labelled riboprobes with digoxigenin RNA labelling kit (Roche Applied Science) by *in vitro* transcription according to the provided manual. Probes were further cleaned by RNeasy Mini kit (QIAGEN) following RNA clean-up manual. Hybridization was performed overnight at 70°C with all probes for both whole mount and sectioned embryo samples. The DIG labelled probes were detected with anti-DIG antibody conjugated with alkaline phosphatase and BM purple AP substrate precipitating solution (Roche Applied Science) was used for signal development. All samples were postfixed in 4% PFA and the slides were mounted in Aquatex. Images were taken using Zeiss ApoTome microscope.

Cell staining and microscopy

SH-SY5Y cells in Supplementary Material, Fig S5C were transfected using Lipofectamine 3000 transfection reagent (Invitrogen) and grown over a coverslip glass. After 48 h, cells were washed twice with 2× PBS, fixed in 4% PFA for 20 min at room temperature, washed with PBST (0.1% Triton X-100 in 1× PBS) for three times and nuclei were stained with NucBlue Fixed Cell Stain Readyprobes reagent (Invitrogen). After staining, cells were washed with 1× PBS three times, the cover slips were mounted into slides with ProLong Diamond antifade reagent (Invitrogen), and samples were analysed in an inverted microscope ECLIPSE TS2-FL (Nikon).

Acknowledgements

The authors are grateful of the technical support provided by UPV/EHU SGIker and Animal house staff, which are supported with European funding (ERDF and ESF). In particular we thank Arantza Alejo and the Animal Unit at UPV/EHU. We thank Janice Fisher, Fen-Biao Gao, Eric Liebl, Jose Antonio Rodríguez Pérez, Vjekoslav Tomaić and Andreas Wodarz for antibodies, flies and plasmids.

*Conflict of Interest statement***.** None declared.

Funding

This work was supported by March of Dimes [Research Grant 1-FY15–339]; by ISCIII (grants PRB2 IPT13/0001 - ISCIII-SGEFI/ERDF and PRB3 IPT17/0019 - ISCIII-SGEFI/ERDF) and Spanish MINECO [grants SAF2013–44782-P and SAF2016–76898-P, both cofinanced with FEDER funds, to UM, BFU2014–52282-P and BFU2017–84653-P to RB]. UM and RB are also part of COST action Proteostasis. This work was also partially supported by Fondation Jérôme Lejeune grant 1381-MU2015A to UM. This work was supported by the Severo Ochoa Excellence Accreditation [grant SEV-2016–0644] and Consolider Programs [grant BFU2014–57703-REDC] to RB. Support was also provided from the Government of the Basque Country (Etortek Research Programs) and the Bizkaia County (Innovation Technology Department). This work was also supported by the Ministry of Education, Youth and Sports of the Czech Republic within the National Sustainability Program II [Project BIOCEV-FAR LQ1604] and by LM2015040 Czech Centre for Phenogenomics by MEYS, CZ.1.05/1.1.00/02.0109 Biotechnology and Biomedicine Centre of the Academy of Sciences and Charles University in Vestec (BIOCEV) by MEYS and ERDF, and CZ.1.05/2.1.00/19.0395—Higher quality and capacity for transgenic models by MEYS and ERDF.

References

1. Buiting K., Williams C., Horsthemke B. (2016) Angelman syndrome - insights into a rare neurogenetic disorder. Nat. Rev. Neurol., 12, 584–593.

2. Kishino T., Lalande M., Wagstaff J. (1997) UBE3A/E6-AP mutations cause Angelman syndrome. Nat. Genet., 15, 70–73.

3. Matsuura T., Sutcliffe J.S., Fang P., Galjaard R.J., Jiang Y.H., Benton C.S., Rommens J.M., Beaudet A.L.(1997) De novo truncating mutations in E6-AP ubiquitin-protein ligase gene (UBE3A) in Angelman syndrome. Nat. Genet., 15, 74–77.

4. Huibregtse J.M., Scheffner M., Beaudenon S., Howley P.M. (1995) A family of proteins structurally and functionally related to the E6-AP ubiquitin-protein ligase. Proc. Natl. Acad. Sci. U. S. A., 92, 2563–2567.

5. Scheffner M.,Huibregtse J.M., Vierstra R.D., Howley P.M.(1993) The HPV-16 E6 and E6-AP complex functions as a ubiquitin-protein ligase in the ubiquitination of p53. Cell, 75, 495–505

6. Scheffner M., Nuber U., Huibregtse J.M.(1995) Protein ubiquitination involving an E1–E2–E3 enzyme ubiquitin thioester cascade. Nature, 373, 81–83.

7. Rougeulle C., Glatt H., Lalande M.(1997) The Angelman syndrome candidate gene, UBE3A/E6-AP, is imprinted in brain. Nat. Genet., 17, 14–15.

8. Gentile J.K., Tan W.H., Horowitz L.T., Bacino C.A., Skinner S.A., Barbieri-Welge R., Bauer-Carlin A., Beaudet A.L., Bichell T.J., Lee H.S. et al. (2010) A neurodevelopmental survey of Angelman syndrome with genotype-phenotype correlations. J. Dev. Behav. Pediatr., 31, 592–601.

9. Sadikovic B., Fernandes P., Zhang V.W., Ward P.A., Miloslavskaya I., Rhead W., Rosenbaum R., Gin R., Roa B., Fang P. (2014) Mutation Update for UBE3A variants in Angelman syndrome. Hum. Mutat., 35, 1407–1417.

10. Nawaz Z., Lonard D.M., Smith C.L., Lev-Lehman E., Tsai S.Y., Tsai M.J., O'Malley B.W. (1999) The Angelman syndrome-associated protein, E6-AP, is a coactivator for the nuclear hormone receptor superfamily. Mol. Cell. Biol., 19, 1182–1189.

11. Cooper E.M., Hudson A.W., Amos J., Wagstaff J., Howley P.M. (2004) Biochemical analysis of Angelman syndrome-associated mutations in the E3 ubiquitin ligase E6-associated protein. J. Biol. Chem., 279, 41208–41217.

12. Urraca N., Cleary J., Brewer V., Pivnick E.K., McVicar K., Thibert R.L., Schanen N.C., Esmer C., Lamport D., Reiter L.T. (2013) The interstitial duplication 15q11.2-q13 syndrome includes autism, mild facial anomalies and a characteristic EEG signature. Autism Res. Off. J. Int. Soc. Autism Res., 6, 268–279.

13. Noor A., Dupuis L., Mittal K., Lionel A.C., Marshall C.R., Scherer S.W., Stockley T., Vincent J.B., Mendoza-Londono R., Stavropoulos D.J. (2015) 15q11.2 duplication encompassing only the UBE3A gene is associated with developmental delay and europsychiatric phenotypes. Hum. Mutat., 36, 689–693.

14. Yi J.J., Berrios J., Newbern J.M., Snider W.D., Philpot B.D., Hahn K.M., Zylka M.J. (2015) An autism-linked mutation disables phosphorylation control of UBE3A. Cell, 162, 795–807.

15. Kim H.C., Huibregtse J.M. (2009) Polyubiquitination by HECT E3s and the determinants of chain type specificity. Mol. Cell. Biol., 29, 3307–3318.

16. Kühne C., Banks L. (1998) E3-ubiquitin ligase/E6-AP links multicopy maintenance protein 7 to the ubiquitination pathway by a novel motif, the L2G box. J. Biol. Chem., 273, 34302–34309.

17. Thomas M., Banks L. (1998) Inhibition of Bak-induced apoptosis by HPV-18 E6. Oncogene,

17, 2943–2954.

18. Oda H., Kumar S., Howley P.M. (1999) Regulation of the Src family tyrosine kinase Blk through E6AP-mediated ubiquitination. Proc. Natl. Acad. Sci. U. S. A., 96, 9557–9562.

19. Mani A., Oh A.S., Bowden E.T., Lahusen T., Lorick K.L., Weissman A.M., Schlegel R., Wellstein A., Riegel A.T. (2006) E6AP mediates regulated proteasomal degradation of the nuclear receptor coactivator amplified in breast cancer 1 in immortalized cells. Cancer Res., 66,

8680–8686.

20. Reiter L.T., Seagroves T.N., Bowers M., Bier E. (2006) Expression of the Rho-GEF Pbl/ECT2 is regulated by the UBE3A E3 ubiquitin ligase. Hum. Mol. Genet., 15, 2825–2835.

21.Kumar S., Talis A.L., Howley P.M. (1999) Identification of HHR23A as a substrate for E6-associated protein-mediated ubiquitination. J. Biol. Chem., 274, 18785–18792.

22. Mishra A., Jana N.R. (2008) Regulation of turnover of tumor suppressor p53 and cell growth by E6-AP, a ubiquitin protein ligase mutated in Angelman mental retardation syndrome. Cell. Mol. Life Sci., 65, 656–666.

23. Margolis S.S., Salogiannis J., Lipton D.M., Mandel-Brehm C., Wills Z.P., Mardinly A.R., Hu L., Greer P.L., Bikoff J.B., Ho H.Y.H. et al. (2010) EphB-mediated degradation of the RhoA GEF Ephexin5 relieves a developmental brake on excitatory synapse formation. Cell, 143, 442–455.

24. Shimoji T., Murakami K., Sugiyama Y., Matsuda M., Inubushi S., Nasu J., Shirakura M., Suzuki T., Wakita T., Kishino T. et al. (2009) Identification of annexin A1 as a novel substrate for E6AP-mediated ubiquitylation. J. Cell. Biochem., 106, 1123–1135.

25. Greer P.L., Hanayama R., Bloodgood B.L., Mardinly A.R., Lipton D.M., Flavell S.W., Kim T.K., Griffith E.C., Waldon Z., Maehr R. et al. (2010) The Angelman syndrome protein Ube3A regulates synapse development by ubiquitinating arc. Cell, 140, 704–716.

26. Zaaroor-Regev D., de Bie P., Scheffner M., Noy T., Shemer R., Heled M., Stein I., Pikarsky E., Ciechanover A. (2010) Regulation of the polycomb protein Ring1B by self-ubiquitination or by E6-AP may have implications to the pathogenesis of Angelman syndrome. Proc. Natl. Acad. Sci. U. S. A., 107, 6788–6793.

27. Jensen L., Farook M.F., Reiter L.T. (2013) Proteomic profiling in Drosophila reveals potential Dube3a regulation of the actin cytoskeleton and neuronal homeostasis. PLoS One, 8, e61952.

28. Jacobson A.D., MacFadden A., Wu Z., Peng J., Liu C.W. (2014) Autoregulation of the 26S proteasome by in situ ubiquitination. Mol. Biol. Cell, 25, 1824–1835.

29. Yi J.J., Paranjape S.R., Walker M.P., Choudhury R., Wolter J.M., Fragola G., Emanuele M.J., Major M.B., Zylka M.J. (2017) The autism-linked UBE3A T485A mutant E3 ubiquitin ligase activates the Wnt/β-catenin pathway by inhibiting the proteasome. J. Biol. Chem., 292, 12503–12515.

30. Lee S.Y., Ramirez J., Franco M., Lectez B., Gonzalez M., Barrio R., Mayor U. (2014) Ube3a, the E3 ubiquitin ligase causing Angelman syndrome and linked to autism, regulates protein homeostasis through the proteasomal shuttle Rpn10. Cell. Mol. Life Sci., 71, 2747–2758.

31. Al-Shami A., Jhaver K.G., Vogel P., Wilkins C., Humphries J., Davis J.J., Xu N., Potter D.G., Gerhardt B., Mullinax R. et al. (2010) Regulators of the proteasome pathway, Uch37 and Rpn13, play distinct roles in mouse development. PLoS One, 5, e13654.

32. Puram S.V., Kim A.H., Park H.Y., Anckar J., Bonni A. (2013) The ubiquitin receptor S5a/Rpn10 links centrosomal proteasomes with dendrite development in the mammalian brain. Cell Rep., 4, 19–30.

33. Tomaić V., Banks L. (2015) Angelman syndrome-associated ubiquitin ligase UBE3A/E6AP mutants interfere with the proteolytic activity of the proteasome. Cell Death Dis., 6, e1625.

34. Xu G., Paige J.S., Jaffrey S.R. (2010) Global analysis of lysine ubiquitination by ubiquitin remnant immunoaffinity profiling. Nat. Biotechnol., 28, 868–873.

35. Kim W., Bennett E.J., Huttlin E.L., Guo A., Li J., Possemato A., Sowa M.E., Rad R., Rush J., Comb M.J. et al. (2011) Systematic and quantitative assessment of the ubiquitin-modified proteome. Mol. Cell, 44, 325–340.

36. Wagner S.A., Beli P., Weinert B.T., Nielsen M.L., Cox J., Mann M., Choudhary C. (2011) A proteome-wide, quantitative survey of in vivo ubiquitylation sites reveals widespread regulatory roles. Mol. Cell. Proteomics, 10, M111.013284. M111.013284.

37. Na C.H., Jones D.R., Yang Y., Wang X., Xu Y., Peng J. (2012) Synaptic protein ubiquitination in rat brain revealed by antibody-based ubiquitome analysis. J. Proteome Res., 11, 4722–4732.

38. Wagner S.A., Beli P., Weinert B.T., Schölz C., Kelstrup C.D., Young C., Nielsen M.L., Olsen J.V., Brakebusch C., Choudhary C. (2012) Proteomic analyses reveal divergent ubiquitylation site patterns in murine tissues. Mol. Cell. Proteomics, 11, 1578–1585.

39. Sarraf S.A., Raman M., Guarani-Pereira V., Sowa M.E., Huttlin E.L., Gygi S.P., Harper J.W. (2013) Landscape of the PARKIN-dependent ubiquitylome in response to mitochondrial depolarization. Nature, 496, 372–376.

40. Nielsen M.L., Vermeulen M., Bonaldi T., Cox J., Moroder L., Mann M. (2008) Iodoacetamide-induced artifact mimics ubiquitination in mass spectrometry. Nat. Methods, 5, 459–460.

41. Ramirez J., Elu N., Martinez A., Lectez B., Mayor U. (2017) In vivo strategies to isolate and characterize the neuronal ubiquitinated proteome. In Santamaría E., Fernández-Irigoyen J. (eds) Current Proteomic Approaches Applied to Brain Function, Neuromethods. Humana Press, New York, NY, pp. 179–189.

42. Franco M., Seyfried N.T., Brand A.H., Peng J., Mayor U. (2011) A novel strategy to isolate ubiquitin conjugates reveals wide role for ubiquitination during neural development. Mol. Cell. Proteomics, 10, M110.002188.

43. Martinez A., Lectez B., Ramirez J., Popp O., Sutherland J.D., Urbé S., Dittmar G., Clague .J., Mayor U. (2017) Quantitative proteomic analysis of Parkin substrates in Drosophila neurons.

Mol. Neurodegener., 12, 29.

44. Ramirez J., Martinez A., Lectez B., Lee S.Y., Franco M., Barrio R., Dittmar G., Mayor U. (2015) Proteomic analysis of the ubiquitin landscape in the Drosophila embryonic nervous system and the adult photoreceptor cells. PLoS One, 10, e0139083.

45. Tyanova S., Temu T., Sinitcyn P., Carlson A., Hein M.Y., Geiger T., Mann M., Cox J. (2016) The Perseus computational platform for comprehensive analysis of (prote)omics data.

Nat. Methods, 13, 731–740.

46. Chandler C.S., Ballard F.J. (1986) Multiple biotin-containing proteins in 3T3-L1 cells. Biochem. J., 237, 123–130.

47. Nezis I.P., Simonsen A., Sagona A.P., Finley K., Gaumer S., Contamine D., Rusten T.E., Stenmark H., Brech A. (2008) Ref(2)P, the *Drosophila melanogaster* homologue of mammalian p62, is required for the formation of protein aggregates in adult brain. J. Cell Biol., 180, 1065–1071.

48. Pavel M., Imarisio S., Menzies F.M., Jimenez-Sanchez M., Siddiqi F.H., Wu X., Renna M., O'Kane C.J., Crowther D.C., Rubinsztein D.C. (2016) CCT complex restricts neuropathogenic protein aggregation via autophagy. Nat. Commun., 7, 13821.

49. Morawe T., Honemann-Capito M., von Stein W., Wodarz A. (2011) Loss of the extraproteasomal ubiquitin receptor Rings lost impairs ring canal growth in Drosophila oogenesis. J. Cell Biol., 193, 71–80.

50. Min M., Mayor U., Lindon C. (2013) Ubiquitination site preferences in anaphase promoting complex/cyclosome (APC/C) substrates. Open Biol., 3, 130097.

51. Stevens A., Jacobs J.R. (2002) Integrins regulate responsiveness to slit repellent signals. J. Neurosci. Off. J. Soc. Neurosci., 22, 4448–4455.

52. Chihara T., Luginbuhl D., Luo L. (2007) Cytoplasmic and mitochondrial protein translation in axonal and dendritic terminal arborization. Nat. Neurosci., 10, 828–837.

53. Lin W.Y., Williams C., Yan C., Koledachkina T., Luedke K., Dalton J., Bloomsburg S., Morrison N., Duncan K.E., Kim C.C. et al. (2015) The SLC36 transporter Pathetic is required for extreme dendrite growth in Drosophila sensory neurons. Genes Dev., 29, 1120–1135.

54. Mercer S.W., La Fontaine S., Warr C.G., Burke R. (2016) Reduced glutathione biosynthesis in *Drosophila melanogaster* causes neuronal defects linked to copper deficiency. J. Neurochem., 137, 360–370.

55. Trotta N., Rodesch C.K., Fergestad T., Broadie K. (2004) Cellular bases of activity-dependent paralysis in Drosophila stress-sensitive mutants. J. Neurobiol., 60, 328–347.

56. Montana E.S., Littleton J.T. (2006) Expression profiling of a hypercontraction-induced myopathy in Drosophila suggests a compensatory cytoskeletal remodeling response. J. Biol. Chem., 281, 8100–8109.

57. Kasprowicz J., Kuenen S., Miskiewicz K., Habets R.L.P., Smitz L., Verstreken P. (2008) Inactivation of clathrin heavy chain inhibits synaptic recycling but allows bulk membrane uptake. J. Cell Biol., 182, 1007–1016.

58. Yamauchi T. (2005) Neuronal Ca2+/calmodulin-dependent protein kinase II–discovery, progress in a quarter of a century, and perspective: implication for learning and memory. Biol. Pharm. Bull., 28, 1342–1354.

59. Weeber E.J., Jiang Y.H., Elgersma Y., Varga A.W., Carrasquillo Y., Brown S.E., Christian J.M., Mirnikjoo B., Silva A., Beaudet A.L. et al. (2003) Derangements of hippocampal calcium/calmodulin-dependent protein kinase II in a mouse model for Angelman mental retardation syndrome. J. Neurosci. Off. J. Soc. Neurosci., 23, 2634–2644.

60. Martínez-Noël G., Galligan J.T., Sowa M.E., Arndt V., Overton T.M., Harper J.W., Howley P.M. (2012) Identification and proteomic analysis of distinct UBE3A/E6AP protein complexes. Mol. Cell. Biol., 32, 3095–3106.

61. Dantuma N.P., Groothuis T.A.M., Salomons F.A., Neefjes J. (2006) A dynamic ubiquitin equilibrium couples proteasomal activity to chromatin remodeling. J. Cell Biol., 173, 19–26.

62. Groothuis T.A., Dantuma N.P., Neefjes J., Salomons F.A. (2006) Ubiquitin crosstalk connecting cellular processes. Cell Div., 1, 21.

63. Kaiser S.E., Riley B.E., Shaler T.A., Trevino R.S., Becker C.H., Schulman H., Kopito R.R. (2011) Protein standard absolute quantification (PSAQ) method for the measurement of cellular ubiquitin pools. Nat. Methods, 8, 691–696.

64. Fonseca R., Vabulas R.M., Hartl F.U., Bonhoeffer T., Nägerl U.V. (2006) A balance of protein synthesis and proteasome-dependent degradation determines the maintenance of LTP. Neuron, 52, 239–245.

65. Forrest C.M., Darlington L.G., Stone T.W. (2013) Involvement of the proteasome and caspase activation in hippocampal long-term depression induced by the serine protease subtilisin.

Neuroscience, 231, 233–246.

66. Dong C., Bach S.V., Haynes K.A., Hegde A.N. (2014) Proteasome modulates positive and negative translational regulators in long-term synaptic plasticity. J. Neurosci. Off. J. Soc. Neurosci., 34, 3171–3182.

67. Zhao Y., Hegde A.N., Martin K.C. (2003) The ubiquitin proteasome system functions as an inhibitory constraint on synaptic strengthening. Curr. Biol., 13, 887–898.

68. Figueiredo L.S., Dornelles A.S., Petry F.S., Falavigna L., Dargél V.A., Köbe L.M., Aguzzoli C., Roesler R., Schröder N. (2015) Two waves of proteasome-dependent protein degradation in the hippocampus are required for recognition memory consolidation. Neurobiol. Learn. Mem., 120, 1–6.

69. van Ooijen G., Dixon L.E., Troein C., Millar A.J. (2011) Proteasome function is required for biological timing throughout the twenty-four hour cycle. Curr. Biol., 21, 869–875.

70. Jiang Y.H., Armstrong D., Albrecht U., Atkins C.M., Noebels J.L., Eichele G., Sweatt J.D., Beaudet A.L. (1998) Mutation of the Angelman ubiquitin ligase in mice causes increased cytoplasmic p53 and deficits of contextual learning and long-term potentiation. Neuron, 21, 799–811.

71. Miao S., Chen R., Ye J., Tan G.H., Li S., Zhang J., Jiang Y., Xiong Z.-Q. (2013) The Angelman syndrome protein Ube3a is required for polarized dendrite morphogenesis in pyramidal neurons. J. Neurosci., 33, 327–333.

72. Gossan N.C., Zhang F., Guo B., Jin D., Yoshitane H., Yao A., Glossop N., Zhang Y.Q., Fukada Y., Meng Q.J. (2014) The E3 ubiquitin ligase UBE3A is an integral component of the molecular circadian clock through regulating the BMAL1 transcription factor. Nucleic Acids Res., 42, 5765–5775.

73. Pignatelli M., Piccinin S., Molinaro G., Di Menna L., Riozzi B., Cannella M., Motolese M., Vetere G., Catania M.V., Battaglia G. et al. (2014) Changes in mGlu5 receptor-dependent synaptic plasticity and coupling to homer proteins in the hippocampus of Ube3A hemizygous mice modeling angelman syndrome. J. Neurosci. Off. J. Soc. Neurosci., 34, 4558–4566.

74. Pastuzyn E.D., Shepherd J.D. (2017) Activity-dependent arc expression and homeostatic synaptic plasticity are altered in neurons from a mouse model of angelman syndrome. Front. Mol. Neurosci., 10, 234.

75. Marash M., Gerst J.E. (2003) Phosphorylation of the autoinhibitory domain of the Sso t-SNAREs promotes binding of the Vsm1 SNARE regulator in yeast. Mol. Biol. Cell, 14, 3114–3125.

76. Lustgarten V., Gerst J.E. (1999) Yeast VSM1 encodes a v-SNARE binding protein that may act as a negative regulator of constitutive exocytosis. Mol. Cell. Biol., 19, 4480–4494.

77. White R.E., Dickinson J.R., Semple C.A.M., Powell D.J., Berry C. (2011) The retroviral proteinase active site and the N-terminus of Ddi1 are required for repression of protein secretion. FEBS Lett., 585, 139–142.

78. Guthmueller K.L., Yoder M.L., Holgado A.M. (2011) Determining genetic expression profiles in *C. elegans* using microarray and real-time PCR. J. Vis. Exp., 10.3791/2777.

79. Valdez C., Scroggs R., Chassen R., Reiter L.T. (2015) Variation in Dube3a expression affects neurotransmission at the Drosophila neuromuscular junction. Biol. Open, 4, 776. 10.1242/bio.20148045.

80. Lehrbach N.J., Ruvkun G. (2016) Proteasome dysfunction triggers activation of SKN-1A/Nrf1 by the aspartic protease DDI-1. eLife, 5, e17721.

81. Koizumi S., Irie T., Hirayama S., Sakurai Y., Yashiroda H., Naguro I., Ichijo H., Hamazaki J., Murata S. (2016) The aspartyl protease DDI2 activates Nrf1 to compensate for proteasome dysfunction. eLife, 5, e18357.

82. Sivá M., Svoboda M., Veverka V., Trempe J.-F., Hofmann K., Kožíšek M., Hexnerová R., Sedlák F., Belza J., Brynda J. et al. (2016) Human DNA-damage-inducible 2 protein is structurally and functionally distinct from its yeast ortholog. Sci. Rep., 6, 30443.

83. Trempe J.F., Šašková K.G., Sivá M., Ratcliffe C.D.H., Veverka V., Hoegl A., Ménade M., Feng X., Shenker S., Svoboda M. et al. (2016) Structural studies of the yeast DNA damage-inducible protein Ddi1 reveal domain architecture of this eukaryotic protein family. Sci. Rep., 6, 33671.

84. Nowicka U., Zhang D., Walker O., Krutauz D., Castañeda C.A., Chaturvedi A., Chen T.Y., Reis N., Glickman M.H., Fushman D. (2015) DNA-damage-inducible 1 protein (Ddi1) contains an uncharacteristic ubiquitin-like domain that binds ubiquitin. Struct. Lond. Engl., 23, 542–557.

85. Flick K., Ouni I., Wohlschlegel J.A., Capati C., McDonald W.H., Yates J.R., Kaiser P. (2004) Proteolysis-independent regulation of the transcription factor Met4 by a single Lys 48-linked ubiquitin chain. Nat. Cell Biol., 6, 634–641.

86. Flick K., Raasi S., Zhang H., Yen J.L., Kaiser P. (2006) A ubiquitin-interacting motif protects polyubiquitinated Met4 from degradation by the 26S proteasome. Nat. Cell Biol., 8, 509–515.

87. Alexander J., Kalev O., Mehrabian S., Traykov L., Raycheva M., Kanakis D., Drineas P., Lutz M.I., Ströbel T., Penz T. et al. (2016) Familial early-onset dementia with complex neuropathologic phenotype and genomic background. Neurobiol. Aging, 42, 199–204.

88. Williams K., Irwin D.A., Jones D.G., Murphy K.M. (2010) Dramatic loss of Ube3A expression during aging of the mammalian cortex. Front. Aging Neurosci., 2, 18.

89. Wu Y., Bolduc F.V., Bell K., Tully T., Fang Y., Sehgal A., Fischer J.A. (2008) A Drosophila model for Angelman syndrome. Proc. Natl. Acad. Sci. U. S. A., 105, 12399–12404.

90. Tomaić V., Pim D., Banks L. (2009) The stability of the human papillomavirus E6 oncoprotein is E6AP dependent. Virology, 393, 7–10.

91. Lu Y., Wang F., Li Y., Ferris J., Lee J.A., Gao F.B. (2009) The Drosophila homologue of the Angelman syndrome ubiquitin ligase regulates the formation of terminal dendritic branches. Hum. Mol. Genet., 18, 454–462.

92. Ramirez J., Min M., Barrio R., Lindon C., Mayor U. (2016) Isolation of ubiquitinated proteins to high purity from in vivo samples. Methods Mol. Biol., 1449, 193–202.

93. Osinalde N., Sánchez-Quiles V., Akimov V., Blagoev B., Kratchmarova I. (2015) SILAC-based quantification of changes in protein tyrosine phosphorylation induced by Interleukin-2 (IL-2) and IL-15 in T-lymphocytes. Data Brief., 5, 53–58.

94. Cox J., Mann M. (2008) MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nat. Biotechnol., 26, 1367–1372.

95. Cox J., Neuhauser N., Michalski A., Scheltema R.A., Olsen J.V., Mann M. (2011) Andromeda: a peptide search engine integrated into the MaxQuant environment. J. Proteome Res., 10, 1794–1805.

96. Wilkinson D.G., Nieto M.A. (1993) Detection of messenger RNA by in situ hybridization to tissue sections and whole mounts. Methods Enzymol., 225, 361–373.

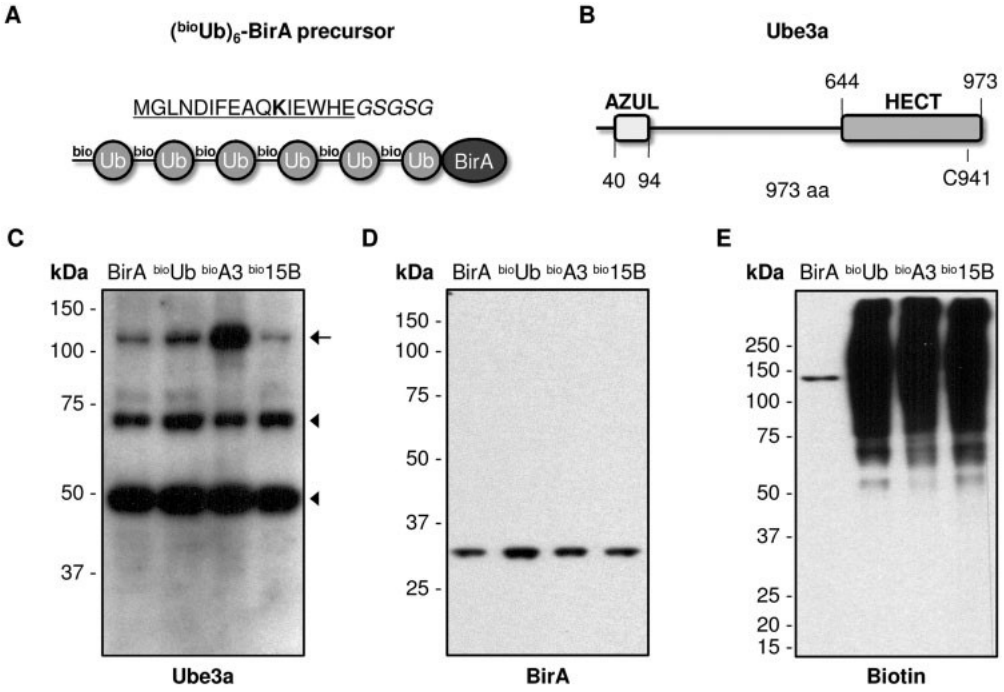


Figure 1: *Drosophila* Ube3a mutant and over-expressing flies expressing the biotin-tagged ubiquitin. (**A**) Schematic representation of the (^{bio}Ub)$_6$-BirA construct, which is expressed as a poly-ubiquitin chain fused to BirA. This precursor polypeptide is digested by endogenous DUBs, so the ubiquitin moieties and the BirA enzyme are released. Each ubiquitin bears a 16 amino-acid long biotinylatable motif at their N-terminal part (bio) that is recognized by BirA. The sequence for biotinylation added to each ubiquitin is shown underlined, followed by the five amino-acid linker (italicized). The lysine where the biotin is attached is highlighted in bold. (B) Schematic representation of the domain structure of *Drosophila* Ube3a. Only two domains have been characterized for Ube3a so far: the AZUL domain (Amino-terminal Zn-finger of Ube3a E3 Ligase), which is thought to play a role in substrate recognition, and the HECT domain (Homologous to the E6-AP Carboxyl Terminus) that provides the E2-binding platform and a catalytic cysteine residue (C941) to which ubiquitin associates via a thioester linkage. Flies over-express Ube3a without any tag and under the control of a UAS sequence. (**C**) Anti-Ube3a immunoblot on head extracts. The specific band for Ube3a is indicated with an arrow. Unspecific bands are indicated with arrowheads. (**D**) Anti-BirA immunoblot on head extracts. Appropriate processing of the (^{bio}Ub)$_6$-BirA precursor was observed for all three genotypes (^{bio}Ub, ^{bio}A3 and ^{bio}15B). Flies over-expressing just BirA were used as control. (**E**) Anti-biotin western blot performed for each of the genotypes. In the control sample (BirA) only endogenously biotinylated proteins are detected.

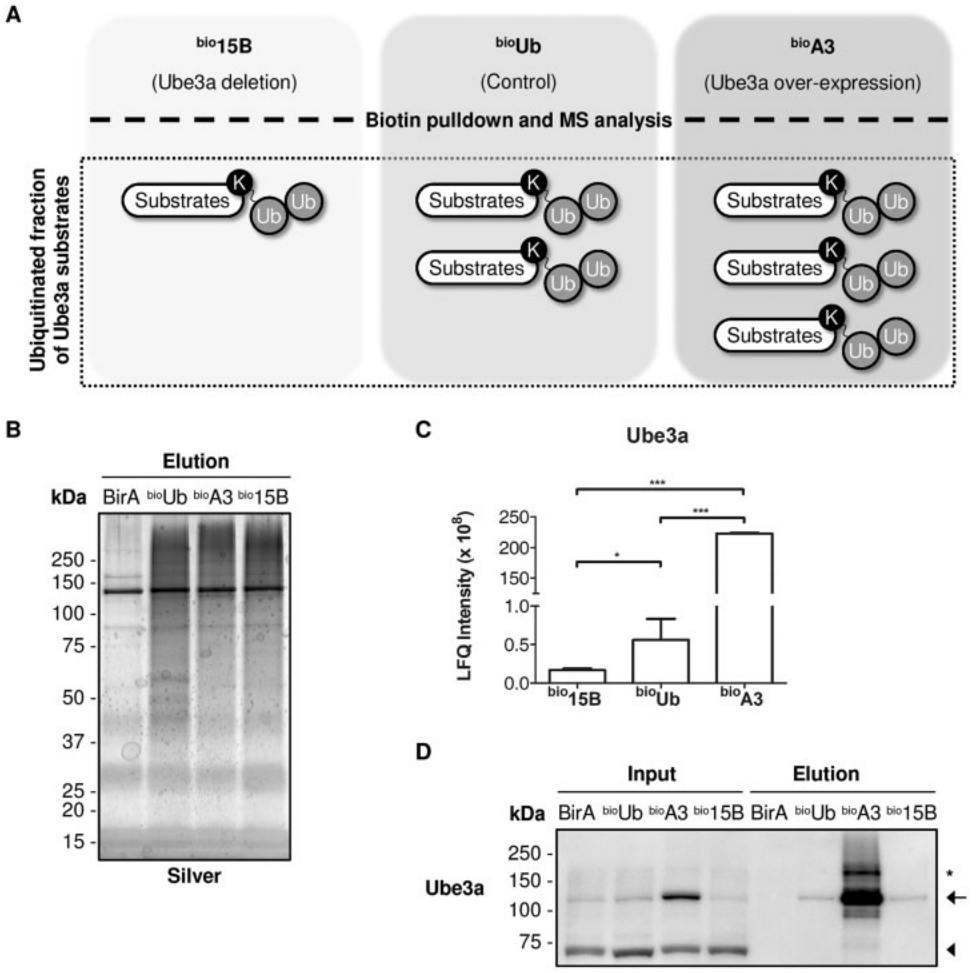


Figure 2: Isolation of candidate Ube3a substrates in *Drosophila* neurons. (**A**) Workflow for the identification of *Drosophila* Ube3a substrates. Flies over-expressing the (^{bio}Ub)$_6$-BirA precursor in the photoreceptor cells under the control of the GMR-GAL4 driver (^{bio}Ub) were compared with heterozygous Ube3a mutant (^{bio}15B) or with Ube3a gain of function flies (^{bio}A3) in order to identify proteins whose ubiquitination increases in a Ube3a dose-dependent manner. Fly heads of each of the genotype were subjected to biotin pulldown and MS analysis. Proteins whose ubiquitination is regulated by Ube3a should be found in more abundance in ^{bio}A3 flies, as compared with ^{bio}Ub and ^{bio}15B controls. (B). Silver staining of the eluted material from biotin pulldowns. Only endogenously biotinylated proteins are detected on the BirA control sample. (C) LFQ intensities of Ube3a obtained from MS analysis of eluted samples. One asterisk indicates *P*-value < to 0.05; three, *P* < to 0.0001. (**D**). Western blot to Ube3a indicates that it is mostly purified in its unmodified form (arrow), which is bound to the avidin beads due to ubiquitin bound to its active-site cysteine. This ubiquitin is removed from the active site by DTT-treatment of the samples on the elution step. A small fraction of Ube3a was also found conjugated to ubiquitin (asterisk). A non-specific band is observed in all input samples (arrowhead).

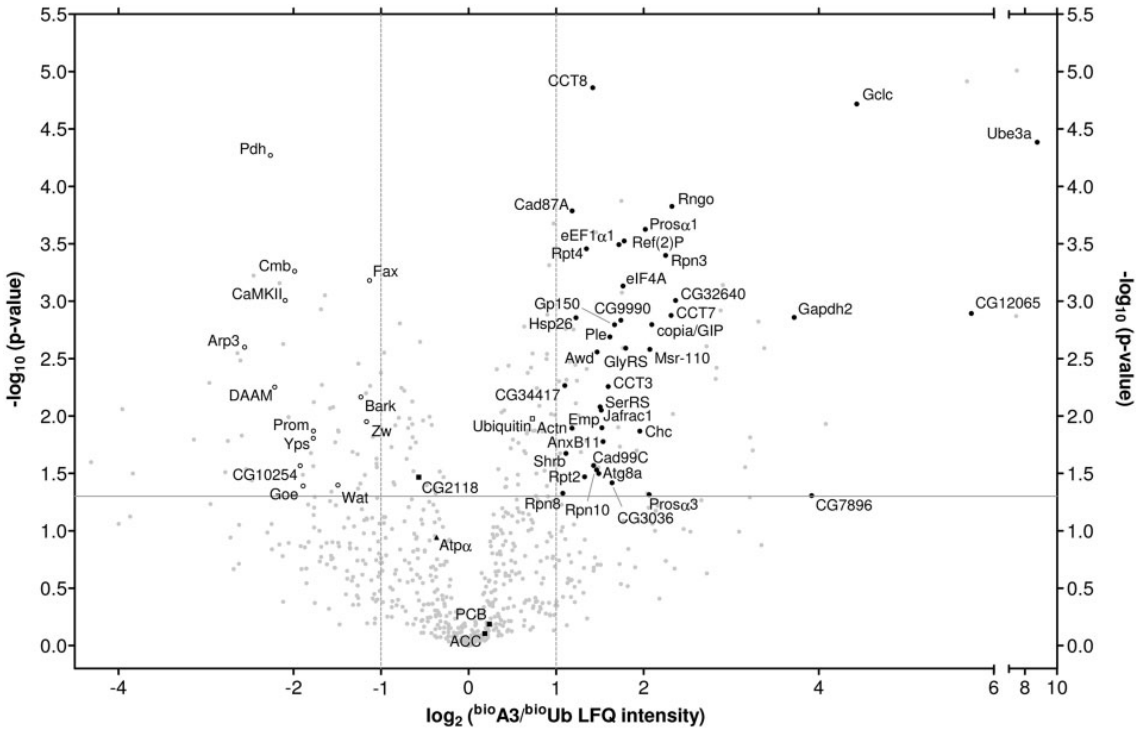


Figure 3: Identification of candidate Ube3a substrates in *Drosophila* neurons. Comparison of the abundance, determined by their LFQ intensities, of the ubiquitinated proteins identified by MS upon Ube3a over-expression relative to ^{bio}Ub flies. The Vulcano plot displays the LFQ ^{bio}A3/^{bio}Ub ratios in $\log_2$ scale (*X*-axis) and the t-*test P-values* in $-\log_{10}$ scale (*Y*-axis), determining the statistical significance ($P < 0.05$, horizontal grey lane) of the fold changes, for each protein. Labelled filled circles represent high confidence proteins found more ubiquitinated in the ^{bio}A3 sample than in the ^{bio}Ub sample. Labelled empty circles are those found less ubiquitinated in ^{bio}A3 sample. Endogenously ubiquitinated proteins (ACC, CG2118 and PCB) are shown with filled squares, and ubiquitin with an empty square. The earlier reported putative Ube3a candidate Atpα, is shown with a triangle. ^{bio}Ub: *GMR-GAL4*, *UAS-(^{bio}Ub)$_6$-BirA/CyO*; ^{bio}A3: *GMR-GAL4*, *UAS-(^{bio}Ub)$_6$-BirA/CyO*; *UAS-Ube3a^{A3}/TM6*.

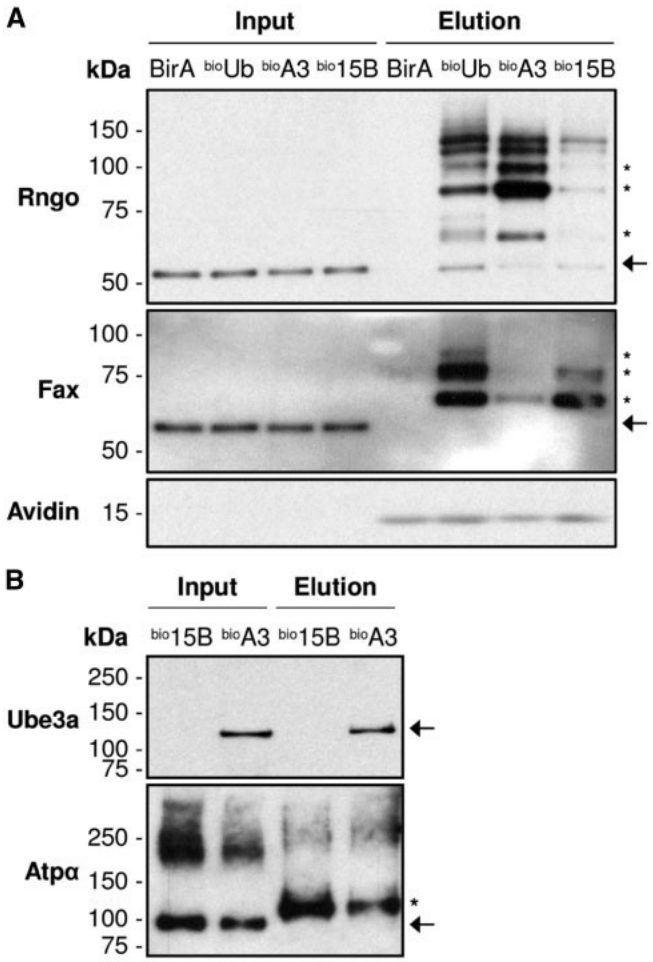


Figure 4: Rngo is ubiquitinated by Ube3a in *Drosophila* photoreceptor neurons. (**A**) Immunoblot with anti Rngo-antibody confirmed the increase ubiquitination detected by MS in ^{bio}A3 flies as compared with ^{bio}Ub flies, and more significantly to ^{bio}15B flies. This same membrane was reprobed with anti-Fax. Levels of ubiquitinated Fax were found reduced in ^{bio}A3 flies, which corroborated the MS results and confirmed that the increase ubiquitination seen for Rngo is specific to the protein and not to a more efficient general isolation of proteins in the ^{bio}A3 sample. Putative mono-, tri- and tetra-ubiquitinated forms are indicated with asterisks and unmodified or cysteine-ubiquitinated forms are indicated with an arrow. Equal levels of Avidin bands, which are non-specifically detected by Rngo antibody, are also shown. (**B**) Atpα is not a substrate of Ube3a in *Drosophila* photoreceptor cells. Western blot performed with anti-Atpα showed that its ubiquitination is not regulated by Ube3a, as levels of mono-ubiquitinated Atpα are similar, or even lower in ^{bio}A3 than in ^{bio}15B flies. The unmodified proteins are indicated with an arrow. Ubiquitinated forms are indicated with asterisks.

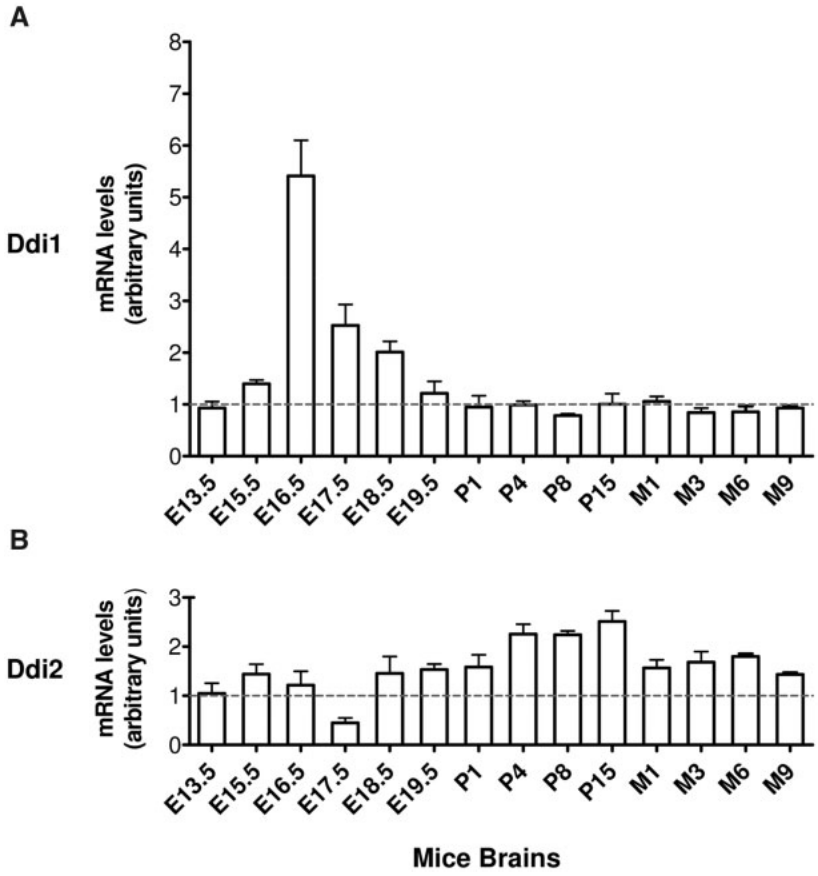


Figure 5: Ddi1 and Ddi2 gene expression temporal profiles in the mouse brain. Changes in expression of Ddi1 (**A**) and Ddi2 (**B**) mRNA levels in brains of C57BL/6J mice during aging. RNA was isolated from the brains of E13.5 to 9-month-old mice and subjected to qRT-PCR. A significant increase of Ddi1 mRNA level was observed at embryonic stage E16.5. Ddi1 and Ddi2 mRNA levels were determined and adjusted by the signal intensity of GAPDH, and the average results ($n = 3$) were calculated and expressed with respect to the values obtained in E13.5.

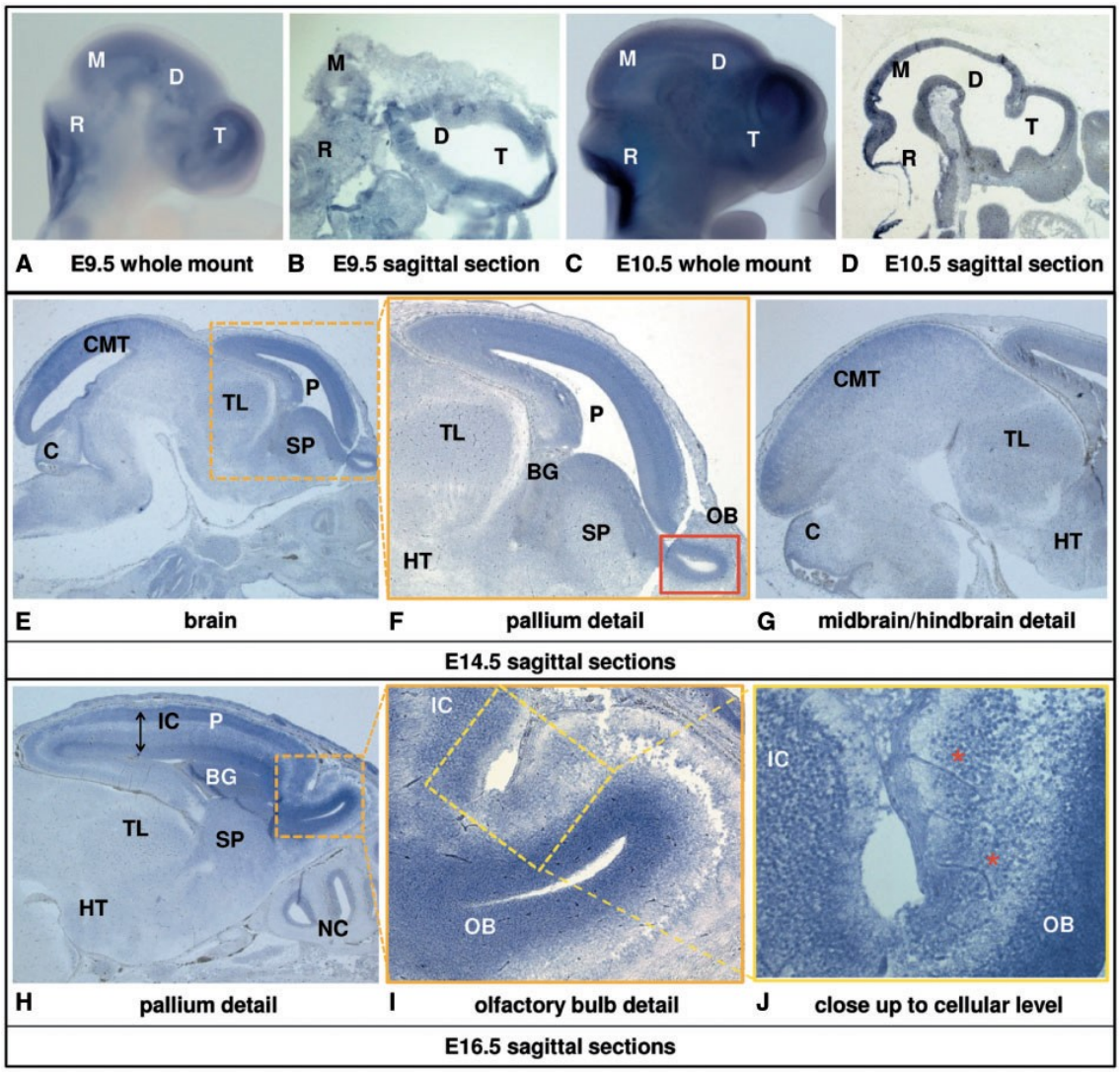


Figure 6: Spatial profile of Ddi1 gene expression in the mouse brain. Ddi1 is expressed in central nervous system during embryonic development according to RNA *in-situ* hybridization study on CD-1 mouse embryos. (**A**) and (**B**) Expression profile of Ddi1 in whole mount and paraffin sections of E9.5 embryo, respectively. (**C**) and (**D**) Ddi1 expression mapping in whole mount and paraffin section of E10.5 embryos. (**E**) Sagittal section of E14.5 embryonic brain with clearly located expression of Ddi1 in colliculus midbrain tectum and pallial part of telencephalon. (**F**) Detail of pallium: Ddi1 is expressed in isocortex and olfactory bulb. (**G**) Midbrain and hindbrain tissue shows expression of Ddi1 in culliculus tectum. (**H**) and (**I**) Ddi1 is specifically expressed in isocortex and ventricular layer of olfactory bulb of E16.5 mouse brain. (J) High resolution detail of the scan shown in (I). Red stars highlight tubular structures—probably capillaries of the central nervous system. Abbreviations: T, telencephalon; D, diencephalon; M, mesencephalon; R, rhombencephalon; C, cerebellum; CMT, colliculus midbrain tectum; TL, thalamus; P, pallium; SP, subpallium; HT, hypothalamus; BG, basal ganglia; OB, olfactory bulb; IC, isocortex; NC, nasal cavity.

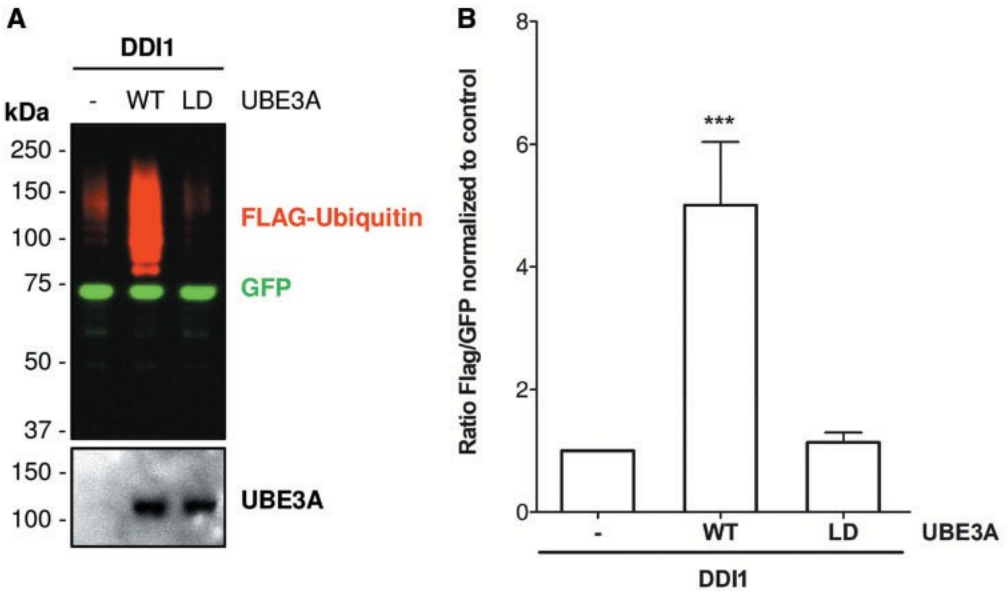


Figure 7: Human DDI1 is ubiquitinated by UBE3A in SH-SY5Y cells. (**A**) DDI1-GFP showed a significant increase in its ubiquitinated fraction in the presence of wild type UBE3A ($UBE3A^{WT}$), as illustrated by western blot to FLAG-tagged ubiquitin, compared with control (pCDNA3.1) or to ligase dead UBE3A ($UBE3A^{LD}$). The non-modified form of DDI1 was detected with anti-GFP antibody (green). The bottom panel shows levels of human UBE3A protein in the whole cell extract before the isolation of the GFP-tagged proteins. (**B**) Quantification of the ubiquitination of DDI1 was performed by calculating the FLAG: GFP ratio in panel A with Image-J. Statistical significance differences [***, $P < 0.001$ (mean ± S.E.M., $n = 5$)] were observed for the $UBE3A^{WT}$ sample relative to both control (pCDNA3.1) and $UBE3A^{LD}$ samples.

Table 1: High confidence proteins whose ubiquitination is dependent on Ube3a over-expression

Fly protein		MS data			Human orthologue[b]		
Name[a]	MW (kDa)	Fold Change	p-value	Unique Peptides	Gene symbol[c]	Gene name[c]	CC[d]
Increased ubiquitination upon Ube3a over-expression							
Ube3a	**108**	**434,43**	**4,1E-05**	**24**	**UBE3A**	**Ubiquitin protein ligase E3A**	
CG12065	71	53,55	1,3E-03	11	-	-	-
Gclc	81	21,61	1,9E-05	10	GCLC	Glutamate-cysteine ligase catalytic subunit	
CG7896	134	15,12	4,9E-02	8	IGFALS	Insulin-like growth factor binding protein acid labile subunit	
Gapdh2	35	13,16	1,4E-03	3	GAPDH	Glyceraldehyde-3-phosphate dehydrogenase	
CG32640	15	5,15	9,9E-04	7	-	-	
Rngo	**51**	**5,00**	**1,5E-04**	**11**	**DDI1/2**	**DNA damage inducible 1 homolog 1/2**	
CCT7	59	4,97	1,3E-03	4	CCT7	Chaperonin containing TCP1 subunit 7	
Rpn3	**56**	**4,76**	**4,0E-04**	**11**	**PSMD3**	**Proteasome 26S subunit, non-ATPase 3**	
Copia\GIP	48	4,26	1,6E-03	16	-	-	-
Msr-110	69	4,20	2,6E-03	11	-	-	-
Prosα3	**29**	**4,18**	**4,8E-02**	**4**	**PSMA4**	**Proteasome subunit alpha 4**	V
Prosα1	**27**	**4,05**	**2,4E-04**	**4**	**PSMA6**	**Proteasome subunit alpha 6**	
Chc	191	3,88	1,4E-02	12	CLTC/-L1	Clathrin heavy chain/Clathrin heavy chain like 1	V
GlyRS	76	3,47	2,6E-03	4	GARS	Glycyl-tRNA synthetase	
Ref(2)P	65	3,43	3,0E-04	9	SQSTM1	Sequestosome 1	V
V		3,40	7,4E-04	6	EIF4A1/2	eukaryotic translation initiation factor 4A1/2	
CG9990	84	3,34	1,5E-03	5	ABCA1-13	ATP binding cassette subfamily A member 1-13	V
eEF1α1	50	3,29	3,2E-04	3	EEF1A1/2	eukaryotic translation elongation factor 1 alpha 1/2	
Gp150	106	3,17	1,6E-03	21	TLR3	Toll like receptor 3	ER/E
CG3036	54	3,11	3,8E-02	4	SLC17A5	Solute carrier family 17 member 5	
Ple	66	3,06	2,0E-03	9	TH	Tyrosine hydroxylase	
CCT3	59	3,02	5,6E-03	11	CCT3	Chaperonin containing TCP1 subunit 3	
AnxB11	33	2,90	1,7E-02	5	ANXA11	Annexin A11	
Emp	53	2,87	1,3E-02	9	SCARB2	Scavenger receptor class B member 2	
Jafrac1	20	2,86	8,9E-03	5	PRDX1/2	Peroxiredoxin 1/2	
SerRS	56	2,83	8,3E-03	6	SARS	Seryl-tRNA synthetase	
Atg8a	14	2,80	3,2E-02	4	GABARAP	GABA type A receptor-associated protein	V
Awd	17	2,77	2,8E-03	6	NME1	NME/NM23 nucleoside diphosphate kinase 1	
Rpn10	**43**	**2,75**	**3,0E-02**	**5**	**PSMD4**	**Proteasome 26S subunit, non-ATPase 4**	
Cad99C	185	2,69	2,7E-02	19	PCDH15	Protocadherin related 15	S
CCT8	59	2,67	1,4E-05	7	CCT8	Chaperonin containing TCP1 subunit 8	
Rpt4	**45**	**2,54**	**3,5E-04**	**9**	**PSMC6**	**Proteasome 26S subunit, ATPase 6**	
Rpt2	**49**	**2,51**	**3,4E-02**	**6**	**PSMC1**	**Proteasome 26S subunit, ATPase 1**	
Hsp26	23	2,34	1,4E-03	7	HSPB1/2	Heat shock protein family B (small) member 1/2	
Cad87A	218	2,27	1,6E-04	13	CDH23	Cadherin related 23	
Actn	104	2,27	1,3E-02	9	ACTN2	Actinin alpha 2	CT
Shrb	25	2,16	2,1E-02	3	CHMP4B	Charged multivesicular body protein 4B	
CG34417	190	2,14	5,4E-03	14	SMTN/-L1	Smoothelin/Smoothelin like 1	CT
Rpn8	**38**	**2,11**	**4,7E-02**	**3**	**PSMD7**	**Proteasome 26S subunit, non-ATPase 7**	
Reduced ubiquitination upon Ube3a over-expression							
Fax	47	1/2,19	6,6E-04	29	FAXC	Failed axon connections homologue	
Zw	60	1/2,24	1,1E-02	11	G6PD	Glucose-6-phosphate dehydrogenase	
Bark	349	1/2,35	6,8E-03	20	LOXL3	Lysyl oxidase like 3	
Wat	60	1/2,81	4,0E-02	15	FAR1/2	Fatty acyl-CoA reductase 1/2	P
Prom	138	1/3,41	1,4E-02	14	PROM1/2	Prominin 1/2	ER
Yps	37	1/3,42	1,6E-02	7	YBX1-3	Y-box binding protein 1/2/3	
Goe	100	1/3,71	4,1E-02	12	PHEX[e]	Phosphate regulating endopeptidase homolog X-linked	
CG10254	154	1/3,79	2,7E-02	8	UBE2O	Ubiquitin conjugating enzyme E2 O	
Cmb	184	1/3,96	5,5E-04	10	PCM1	Pericentriolar material 1	CT
CaMKII	58	1/4,27	9,9E-04	5	CAMK2D	Calcium/calmodulin dependent protein kinase II delta	
DAAM	162	1/4,64	5,6E-03	18	DAAM1	Dishevelled associated activator of morphogenesis 1	
Pdh	30	1/4,81	5,4E-05	15	HPGD	15-hydroxyprostaglandin dehydrogenase	
Arp3	47	1/5,90	2,5E-03	10	ACTR3	ARP3 actin related protein 3 homolog	CT
CG43078	274	1/6,87	3,1E-02	6	-	-	-

Proteins whose abundance in the pulldowns is significantly altered by Ube3a, both at protein and at peptide level, are shown. The complete data set is available as Supplementary Material, Table S1. Proteasomal proteins are highlighted in bold. Cellular localization (CC) of the human proteins are indicated by graytones (dark grey rounded rectangle: nuclear; light grey cell: cytoplasmic; black surrounding square: plasma membrane). If a more specific localization within each compartment has been reported, it is further indicated with text (CT, cytoskeleton; E, endosome; ER, endoplasmic reticulum; P, peroxisomes; S, secreted; V, vesicles).

a) Given according to Flybase nomenclature.

b) Orthologues with the best Flybase score are provided.

c) Given according to HUGO Gene Nomenclature Committee.

d) Given according to Human Protein Atlas (www.proteinatlas.org) and Uniprot (www.uniprot.org).

e) Other M13 metallopeptidases are also considered orthologues (see Supplementary Material, Table S1).

Univerzita Karlova v Praze, 1. lékařská fakulta

Kateřinská 32, Praha 2

Prohlášení zájemce o nahlédnutí do závěrečné práce absolventa studijního programu uskutečňovaného na 1. lékařské fakultě Univerzity Karlovy v Praze

Jsem si vědom/a, že závěrečná práce je autorským dílem a že informace získané nahlédnutím do zpřístupněné závěrečné práce nemohou být použity k výdělečným účelům, ani nemohou být vydávány za studijní, vědeckou nebo jinou tvůrčí činnost jiné osoby než autora.

Byl/a jsem seznámen/a se skutečností, že si mohu pořizovat výpisy, opisy nebo kopie závěrečné práce, jsem však povinen/a s nimi nakládat jako s autorským dílem a zachovávat pravidla uvedená v předchozím odstavci.

Příjmení, jméno (hůlkovým písmem)	Číslo dokladu totožnosti vypůjčitele (např. OP, cestovní pas)	Signatura závěrečné práce	Datum	Podpis