CHARLES UNIVERSITY

Faculty of Science

Department of Physical and Macromolecular Chemistry

BIOPHYSICAL AND FUNCTIONAL CHARACTERIZATION OF DDI1-LIKE ASPARTIC PROTEASES INVOLVED IN REPLICATION STRESS RESPONSE

Biofysikální a funkční charakterisace aspartátových proteas z rodiny proteinů podobných Ddi1, zapojených v odpovědi na replikační stres

Mgr. Michal Svoboda

Doctoral thesis

of study program Physical Chemistry

Supervisor: doc. RNDr. Jan Konvalinka, CSc.
Consultant: Mgr. Klára Grantz Šašková, Ph.D.

Prague 2020

This dissertation describes my original work except where an acknowledgement is made in the text. It is not substantially the same as any work that has been or is being submitted to any other university for any degree, diploma, or any other qualification.

Prague 05. 11. 2020 ……..……..

Michal Svoboda, MSc.

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my supervisor Honza Konvalinka, who, years ago, was brave enough to bring an unskilled, confused, but curious student to his great group. Since that very moment, I have enjoyed all the support I could wish for and spent there several amazing years. It was a hell of a ride, sometimes bumpy, but definitely worth it.

Next, I would like to thank my consultant Klára Grantz Šašková for giving me the opportunity to join the Ddi project a few years ago, introducing me to the enchanting world of ubiquitin-related research and leading me through my PhD project.

A great deal of gratitude also belongs to J.F. Trempe and members of his lab for resparking my interest in science when I needed it the most.

To Mony for everything, because this page is far too small to express even a small part of the things I have to thank her for.

A great chunk of gratitude belongs to my beloved "Quatro happiness" guys Jardič and Kuba, as well as to Mr. Adámek. First for a constant probing of my understanding of topics we worked on by their curious questions and discussions. Because, you know, once you are not able to explain it, you do not truly understand it. Nonetheless, for the always upbeat and encouraging mood, they created in our "corner of success."

In a great Werner Herzog movie Fitzcarraldo, the main protagonist, in pursue of his dream, decides to do a seemingly crazy thing and drags a massive boat over the mountain ridge. Sadly, way too often during my PhD I was not like the brave Brian Fitzgerald, but rather resembled the burden he was dragging. Fortunately, during those low times, I was surrounded by an amazing group of people always offering support. Therefore, I would like to thank those pushers and pullers, whose coordinated effort helped to move the shipwreck of my PhD over the mountain, namely Zuzka, Kačka, Bimča, Karolína, Klárka, Anež, Edita, Begoš, Sofie, Leňa, Cyril, Xenie, Markét, and the greatest whip master himself, Václav.

Big thank you belongs also to those anonymous heroes of Youtube, who produced such amazing mixes as "Three hours of gentle rain" or "10 hours of rain and stormy ocean sounds aboard a wooden ship." Thanks for your work guys, I would not make it without it.

Last, but most definitely not least bych rád poděkoval své rodině za bezbřehou podporu, které se mi od nich dostávalo po celou dobu mého nekonečného studia. Díky moc.

ABSTRAKT

Bezchybná a včasná replikace molekul DNA je jedním z rozhodujících momentů životního cyklu každého organismu. Jakýkoliv defekt, ať už v časování či lokalizaci, může tento delikátní, na přesnosti závislý mechanismus uvrhnout do stavu takzvaného replikačního stresu. Některé z překážek blokujících postup replikační vidličky obsahují proteinovou složku. V průběhu evoluce se tak vyvinuly specializované proteasy, zodpovědné právě za uvolňování replikačního stresu. Kompletní paleta těchto proteolytických enzymů, stejně jako jejich substrátový repertoár a detailní molekulární mechanismus jejich zapojení v obraně proti DNA replikačnímu stresu však dosud nebyly objasněny. Tato disertační práce se zabývá rolí evolučně konzervovaných aspartátových proteas z rodiny proteinů podobných Ddi1 v proteolytické odpovědi na DNA replikační stres. V této práci byly strukturně a biofyzikálně charakterisováni členové rodiny Ddi1 podobných proteinů kvasinkového a lidského původu, byly prozkoumány jejich interakce s ubikvitinem, stejně jako s polyubikvitinovými řetězci. V neposlední řadě byla též mutantního kvasinkového kmene, který vedle genu pro *DDI1* postrádá také gen pro DNA aktivovanou metalloproteasu *WSS1*, identifikována přecitlivělost na hydroxymočovinu, známý inhibitor replikace DNA. Detailní rozbor role kvasinkové DDI1 v zprostředkování tolerance vůči působení hydroxymočoviny odkryl zásadní roli proteolytické ativity domény podobné retrovirovým proteasám, stejně jako nepostradatelnost svazku alfa helixů v sekvenci bezprostředně předcházejících proteasovou doménu. Předkládané výsledky tak poskytují experimentální základ pro nový mechanismus odpovědi na DNA replikační stres, zprostředkovaný dvěma vzájemně zástupnými proteasami.

ABSTRACT

Accurate, timely replication of a DNA molecule is a pivotal moment in the life cycle of every living organism. Any temporal or spatial defect putting the fine-tuned replication machinery off balance causes the so-called replication stress. As the replication machinery consists mainly of enzymes and other proteins, it is not surprising that many of the obstacles most severely blocking the replication machinery progress are of protein origin. Therefore, specialized proteases responsible for relieving replication stress matured during evolution. However, neither the full repertoire of proteolytic enzymes and their particular substrates taking place in countering the DNA replication stress nor detailed molecular mechanisms involved remain unknown. This thesis describes how conserved putative aspartic proteases of the Ddi1-like family engage in countering DNA replication stress via a proteolysis dependent mechanism. We structurally and biophysically characterized yeast and human members of the Ddi1-like family, explored their interactions with ubiquitin and polyubiquitin chains, and identified hypersensitivity to DNA replication inhibitor hydroxyurea in a yeast strain double deleted for *DDI1* gene together with a DNA dependent metalloprotease *WSS1*. Detailed analysis of the DDI1 role in hydroxyurea tolerance showed a crucial role for the proteolytic activity of the aspartic retroviral protease-like domain and the essential role of the helical bundle directly preceding the protease domain. Our results provide an experimental foundation for a novel dual protease-based mechanism allowing yeast cells to counteract DNA replication stress.

ABBREVIATIONS

3D	three-dimensional
aa	amino acid
ATM	ataxia telangiectasia mutated
ATP	adenosine triphosphate
ATR	ataxia telangiectasia and Rad3 related
ATRIP	ATR interacting protein
CDC25	cell division cycle 25
CDC6	cell division cycle 6
CDC7	cell division cycle 7
CDC45	cell division control protein 45
CDC48	cell division cycle 48
Cdc48p	product of yeast *CDC48* gene
CDK	cyclin dependent kinase
CDT1	Cdc10-dependent transcript 1
CHK1	checkpoint kinase 1
CHK2	checkpoint kinase 2
CMG	Cdc45:Mcm2-7:GINS helicase complex
CPS1	carboxypeptidase yscS 1
DBF4	dumbbell former 4
DDI1	DNA damage-inducible 1
Ddi1p	product of yeast *DDI1* gene
DDI2	DNA damage-inducible 1 homologue 2
DNA	deoxyribonucleic acid
dNTP	deoxynucleotide triphosphates
DPC	DNA-protein crosslink
ER	endoplasmic reticulum
ERAD	ER associated degradation
ORC	origin recognition complex
GINS	from Japanese go-ichi-ni-san
GPA1	G protein alpha subunit 1
GPD	glyceraldehyde-3-phosphate dehydrogenase
HDD	helical domain of Ddi1-like proteins
HO	HOmothallic switching endonuclease
HPLC	highperformance liquid chromatography
HRD3	HMG-coA reductase degradation
ITC	isothermal titration calorimetry
Ku, Ku70/80	derived from the surname of the Japanese patient in which it was discovered
MAG1	3-methyl adenine DNA glycosylase
MCM 2-7	minichromosome maintenance 2–7
MCM10	minichromosome maintenance 10
MEC1	mitosis entry checkpoint 1
MS	mass spectrometry
NFE2L1	nuclear factor erythroid-2-like 1
NMR	nuclear magnetic resonance
NRF3	nuclear factor E2-related factor 3
PBS	phosphate buffer saline
PCNA	proliferating cell nuclear antigen
PCR	polymerase chain reaction
PDR3	pleiotropic drug resistance 3
PDS1	precocious dissociation of sisters 1
PNG1	peptide N-glycanase 1
PICS	proteomic identification of cleavage sites

PLK1	polo-like kinase 1
Pol α	DNA polymerase α
RAD23	radiation sensitive 23
RAD53	radiation sensitive 53
RNA	ribonucleic acid
RNAP II	RNA polymerase II
rngo	rings lost
Rpb1	RNA polymerase II subunit B1
RPN1	regulatory particle non-ATPase 1
RPN4	regulatory particle non-ATPase 4
RPN10	regulatory particle non-ATPase 10
RSC1A1	regulator of solute carriers 1
RTF2	replication termination factor 2
RVP	retroviral protease-like domain
SC	synthetic complete media
SCF	Skp1–cullin–F-box-protein ubiquitin ligase
SEC9	secretory 9
SIM	SUMO interacting motif
skn-1	skinhead-1
SMT3	suppressor of mif two 3, yeast SUMO homologue
Smt3p	product of yeast *SMT3* gene
SNARE	SNAP receptor
SNC1	suppressor of the null allele of CAP 1
SPRTN	SprT-like domain-containing protease Spartan
SUMO	small ubiquitin-like modifier
ssDNA	single-stranded DNA
SSO1	supressor of Sec one 1
Sso1p	protein product of yeast SSO1 gene
TCF11/NRF1	transcription factor 11/ nuclear factor E2-related factor 1
tSNARE	target SNAP receptor
UBA	ubiquitin-associated domain
Ube3a	ubiquitin protein ligase E3A
UBL	ubiquitin-like domain
UFO1	UV-F-box-HO 1
Ufo1p	protein product of yeast UFO1 gene
UIM	ubiquitin interacting motif
UPS	ubiquitin-proteasome system
USP9X	Ubiquitin Specific Peptidase 9 X-Linked
UV	ultraviolet radiation
VCP/p97	valosin-containing protein/protein 97
VIM	VCP interacting motif
Vsm1p	vSNARE master, product of yeast *DDI1* gene, alternative name for Ddi1p
vSNARE	vesicle SNAP receptor
WSS1	weak suppressor of SMT3 protein 1
Wss1p	protein product of yeast WSS1 gene
YPD	yeast extract – peptone – dextrose media
YPDA	yeast extract – peptone – dextrose – adenine media

TABLE OF CONTENTS

Introduction

1. INTRODUCTION

The primary objective of every living organism is to pass its genetic information, stored in the form of deoxyribonucleic acid (DNA), in the unchanged and intact form to the next generation. In every eukaryotic cell, from single-cell yeast and protozoa to cells constituting the human body, the process of passing genetic information revolves around a fundamental biological oscillator, the cell cycle. Here, in a repetitive series of events cell grows, accumulates necessary factors, building blocks, and nutrients, duplicates the complete genetic material, and then divides into two genetically identical daughter cells. The whole process, depicted in the figure 1, can be divided into four phases. In the first, G1 phase, a cell enhances protein synthesis, increases its pool of mitochondria, ribosomes, and other important components, and grows in size. The second phase, called S or synthetic, is marked by duplication of the cell's genetic material. Proteosynthesis is downregulated and energy resources of a cell are redirected toward DNA replication. The S phase is followed by another growth phase, G2, in which proteosynthesis is restarted, the cell grows rapidly and generates material for the division. The last phase of the cycle harbours the most dramatic changes in cellular architecture. In M, or mitotic, phase the previously doubled DNA condensates in tightly packed chromosomes, a microtubule-based spindle is then formed, and sister chromatids of chromosomes are separated and transported to the opposite ends of the cell. In an animal cell, mitosis is preceded by disassembly of the nuclear membrane, whereas in fungi such as S. cerevisiae mitosis occurs inside an intact nucleus. Separation of sister chromatids and eventual nuclear envelope reformation is followed by cytokinesis, where the cell divides its cytoplasmatic content roughly in halves, remodels the cytoplasmic membrane, and separates into two daughter cells. In a pathway alternative to committing to the cell cycle progression, a cell can exit the cycle in a G1 phase and transfer to the G0 phase, a resting phase in which a cell can commit to full differentiation. Progression through the cell cycle and proper separation of its phases is directed by the activity of Cyclin-Dependent Kinases (CDK) and timely expression and degradation of their interaction partners, cyclins [Vermeulen et al., 2003].

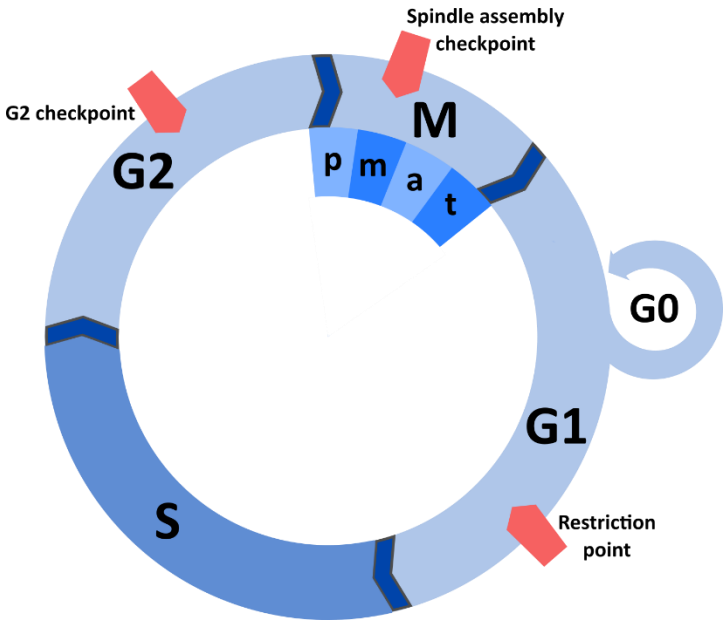


Figure 1: schematic representation of four eukaryotic cell cycle phases with a subdivision of mitotic phase. Three major cell cycle checkpoints are depicted as red wedges.

1.1. CELL CYCLE CHECKPOINTS

During the progression of the cell cycle, the cell must go through crucial decision moments, where its inner control mechanisms decide whether to commit to the progression of the cell cycle or to bail out for safety reasons. Throughout the cycle there are three such cell cycle checkpoints [Johnson and Walker, 1999, Kastan and Bartek, 2004]. Restriction checkpoint represents the first and major checkpoint. Occurring in a late G1 phase, just before entering the S phase this checkpoint is responsible for checking external stimuli, level of growth factors, DNA damage, and steady rate of protein synthesis. Once a cell passes through the restriction point it is fully committed to going through the cell cycle [Johnson and Skotheim, 2013]. The second checkpoint happens in the G2 phase and this time it specializes in scanning for DNA damage. Freshly duplicated DNA molecules are scanned for any sites of damage, lesions, or portions with incomplete replication. Similar to the restriction checkpoint, information about severe DNA damage is transferred downstream by a network of kinases, most important of which are Ataxia Telangiectasia and Rad3 related (ATR) kinase, Ataxia Telangiectasia Mutated (ATM) kinase, Checkpoint kinase 1 (CHK1), and Checkpoint kinase 2 (CHK2) and kinase targets like cell division cycle 25 (CDC25) or Polo-Like Kinase 1 (PLK1). Signaling through this network leads to an arrest in G2 with the cycle delayed until the damage is repaired [Bassermann et al., 2008]. The last important checkpoint takes place in the metaphase of a mitotic M phase and controls for proper alignment of a mitotic spindle, the correct position of all chromosomes, and adequate tension along the spindle microtubules. All the above-mentioned checkpoints represent pivotal decision moments in the cell cycle, where markers

important for non-pathological progression are checked. Once the cell fails the checkpoint, the cell cycle is arrested, the level of damage is assessed and the progression is either paused till available repair mechanisms fix the acute problem or stopped altogether and the cell is destined to following one of the pathways leading toward cell death [Pietenpol and Stewart, 2002].

1.2. DNA REPLICATION

DNA replication is a key process in the cell cycle. The whole genetic information of a given cell stored in a long, yet fragile DNA molecule must be duplicated flawlessly in this step. Due to its size (up to 100 Gb in some plants [Yang et al., 2019]) eukaryotic DNA molecules are replicated from thousands of sites at the same time. Those sites, called origins of replication, are scattered along the genome to ensure each part of every chromosome is replicated on time and correctly. In fact, there are more replication origins on the DNA molecule than is actually used in the early replication stage, but only some of the origins turn to so-called "firing origins" while others remain dormant [Heller et al., 2011]. Replication starts already in the late G1 phase with the origin licensing step, where the chosen origin is turned on by the formation of the pre-replication complex. This process must be tightly regulated to ensure that each part of the DNA molecule is duplicated only once during the replication step. The pre-replication complex is formed by sequential recruitment of its components starting with the Origin Recognition Complex (ORC), followed by licensing factors Cell Division Cycle 6 (CDC6) and Cdc10-Dependent Transcript 1 (CDT1) [Mailand and Diffley, 2005], and finished by recruiting a pair of replicative helicases complex MiniChromosome Maintenance 2–7 (MCM2-7), oriented in an opposite direction to each other [Remus et al., 2009]. The pre-replication complex in this form can then be turned into a pre-initiation complex by the activity of S phase kinases. Although many origins are licensed, just a few of them are activated, with the rest serving as a backup for cases of the stalled replication machinery. Upon origin activation, the pair replicative helicase complexes start moving away from each other, thus unwinding the DNA and providing a template for the replication machinery [Fragkos et al., 2015].

The replication machinery, the replisome, presented in figure 2, assembles on the Y-shaped structure, opened by the progression of the replicative helicase complex and called the replication fork. Replisome assembly follows the origin activation. Additional factors, most important of which is the Cell Division Control protein 45 (CDC45), are recruited to the replication fork in a process directed by Cell Division Cycle 7/ DumbBell former 4 (CDC7/DBF4) heterodimeric kinase and CDK activity [Pacek et al., 2006, Heller et al., 2011]. Upon phosphorylation, replisome binds a GINS (from the Japanese Go-Ichi-Ni-San) heterotetrameric complex, tightly associated to the first recruited DNA polymerase of the replisome, DNA polymerase ε [Gambus et al., 2009].

At this point, MCM replicative helicase complex encircles both strands of the DNA molecule at the origin site. In the following step, helicase must therefore rearrange and wrap only around the soon to be leading strand. This step is currently poorly described although it is understood that phosphorylation by both CDK and CDC7/DBF4 plays a crucial role [Bruck and Kaplan, 2015]. Final activation of the replicative helicase is facilitated by the Minichromosome Maintenance 10 (MCM10), turning helicase into active CDC45, MCM2-7, GINS (CMG) complex and accompanied by the single-strand binding Replication Protein A (RPA) association with the first appearing naked DNA single strand. Movement of the CMG helicase along the leading strand of the nascent replication forks generates single-stranded DNA (ssDNA) which is immediately decorated with RPA [Kanke et al., 2012], providing sites for priming by the DNA polymerase α (Pol α)–DNA primase complex for leading strand DNA synthesis. Primase then synthesizes a short RNA primer, necessary for the initiation of DNA synthesis [Arezi and Kuchta, 2000]. Primer is recognized by DNA polymerase and serves as a starting point for DNA polymerase processive action, copying the DNA template by synthesizing a complementary DNA strand. Because the DNA polymerase can add nucleotides into the chain only in the 5‘ to 3‘ direction, the replication process must be performed in a semidiscontinuous manner. Here, the replicative helicase-bearing leading strand is replicated continuously, while the complementary, lagging strand is replicated discontinuously in short, about 200 nucleotide fragments called Okazaki fragments which are processed afterward by primer degradation, post primer gap filling by dedicated DNA polymerase, and religation by DNA ligase [Burgers, 2009, Waga and Stillman, 1994]. DNA molecule often occurs in the super-coiled conformation. Such a state, when encountered by the progressing replication fork, generates tension in the DNA molecule and resistance for the replisome, which once unresolved, can lead even to the stalling of the replication machinery. To counter this problem, topoisomerase, an enzyme responsible for the unwinding of the super-coiled structure of the DNA in a single strand break mediated catalytic action, cleans the platform for replisome action [Champoux, 2001].

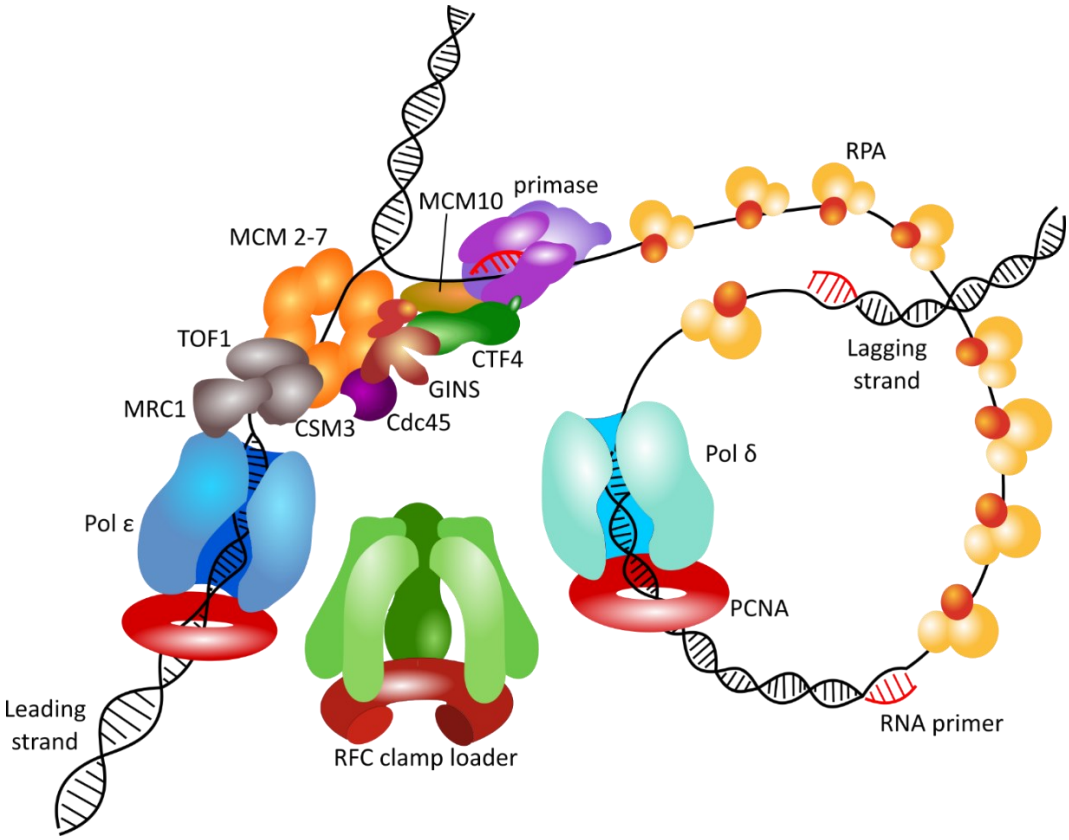


Figure 2: Organization of eukaryotic cell replication machinery on the replication fork structure. Adapted from [O'Donnell and Kurth, 2013]

1.3. REPLICATION STRESS

To successfully perform the duplication of a complete genetic material of the cell, the replication machinery encounters plenty of obstacles of either endogenous or exogenous origin, hindering its progress along the DNA molecule. Such obstacles and consequential delays in the progression of the replisome collectively contribute to a phenomenon of the so-called replication stress. Minor cases of the replication stress can be relieved on the fly by pausing and waiting till the obstacle is cleared or by repriming past the lesion. On the other hand, major incidents of replication stress, loss of key reparation pathways, or chronic replication stress can lead to severe consequences, ranging from checkpoint activation to apoptosis and in cases where the replication stress response fails systemically even to several pathologies including malignant transformation [Zeman and Cimprich, 2013].

Sources of the replication stress are not yet fully described, and the portfolio of possible causes is continually growing. Missing parts of the DNA molecule, like nicks, gaps, and extended single strands are closely connected with the replication stress as they can form both the cause of the stress as well as a symptom of resolution in progress. Nick-like structures cause pausing of the replisome before them because encounter of those structures with the replication machinery

inevitably leads to double-strand breaks [Ciccia and Elledge, 2010]. Closely related are portions of DNA molecule possessing complicated and thus stable secondary structure elements, repetitive sequences, G-quadruplexes, and other elements intrinsically challenging for the replication apparatus. Such sequences must be preprocessed by specialized helicases before the DNA replication resumes [Maffia et al., 2020]. Another source of replication stress stems from the fact that both the replication machinery and the transcription machinery are working on the same template DNA molecule. This inevitably leads to collisions between those two systems or between replisome and products of the transcription machinery. RNAses, helicases, topoisomerases, and targeted translocation of the gene product through the nuclear pore are utilized to resolve replication – transcription conflicts [Bermejo et al., 2012].

The most common and often severe sources of the replication stress origin in the unrepaired structural DNA damage, known as DNA lesions. They constitute an unreplicable barrier, inaccessible for the replicative DNA polymerase or other components of the replisome, and thus cause stalling of the replication fork. Replication blocking lesions can be of both endogenous (metabolic reactive oxygen species byproducts, products of histone code remodeling, aberrant DNA processing enzymes) and exogenous (biological and environmental toxins, drugs, ionizing radiation) origin [Mazouzi et al., 2014]. Common lesions leading to replication stress are intra- and more severely inter-strand crosslinks [Ling et al., 2016], and also bulky lesions of protein nature. Protein lesions are generating more interest in recent years and it was shown that not only covalent adduct-based lesions, termed DNA-Protein Crosslinks (DPC's) [Duxin et al., 2014] but also proteins to tightly yet noncovalently bound to DNA molecule can cause severe problems. One such example is protein Ku, which although being a member of different DNA repair pathway, once fulfills its function remains stuck in a ring-shaped conformation around the DNA double-strand and if unremoved causes the replication stress [van den Boom et al., 2016]. The severity of DNA lesions blocking the progression of replisome is corroborated by the fact that cells evolved a specialized pathway to counteract most of them. Although less complicated lesions can be bypassed by translesion synthesis [Lehmann et al., 2007], more challenging lesions usually lead to fork stalling, followed by lesion repair pathway activation and, depending on the result of the repair process or time required, either fork restart over repaired lesion or one of the available backup mechanisms (repriming, template switching, fork reversal or firing of a nearby dormant origin [Cortez, 2015].

Once the early origins are fired and the cell commits to the replication process, it needs to delicately balance fidelity, speed of replisome progression, and consumption and distribution of relevant material and metabolic resources needed to keep the replication machinery on track.

Therefore, a limited pool of free deoxynucleotide triphosphates (dNTP), insufficient synthesis of replication machinery components, or lack of histones available for DNA packaging represent another important, yet often overlooked, cause of the replication stress [Bester et al., 2011]. Hydroxyurea, an often used antineoplastic drug, is a prime example of such a mechanism, as it inhibits the enzymatic activity of free dNTP-generating enzyme, ribonucleotide reductase, thus limiting the pool of free dNTPs and in turn, slowing down the replisome progression, possibly even leading to cell cycle arrest in S phase [Petermann et al., 2010].

To detect and resolve incidents of replication stress cell exploits a network of DNA damage sensing kinases. Stalled replication fork often leads to the uncoupling of replicative CGM helicase activity from replicative polymerase activity. Helicase progressing further along the leading strand thus generates extended patches of RPA decorated single-strand DNA. Such moieties in the vicinity of nascent, freshly replicated double-stranded DNA generates a primer-template junction signal for activation of the replication stress response. ATR interacting protein (ATRIP) is then recruited to the RPA coated ssDNA patches and brings ATR and ATM kinases. ATR follows by phosphorylation of RPA, surrounding histones H2AX and downstream kinase CHK1. Signaling via ATR and its downstream factors leads to inhibition of cell cycle progression, fork stabilization, inhibition of late origin firing, inhibition of recombination, and initiation of repair of the replication block or repair of the faulty replication fork. Once the obstacle is removed, the replisome can be restarted [Yamada et al., 2013, Tkach et al., 2012].

1.4. ROLE OF PROTEOLYTIC ENZYMES IN REPLICATION STRESS RESPONSE

As many of the factors constituting replisome as well as other molecular machines working on the DNA molecule are proteins, it comes as no surprise that proteases, enzymes responsible for degradative or regulatory cleavage of proteins and polypeptide chains, represent an important group of players involved in DNA replication stress response [Barker et al., 2005]. Two types of proteolysis participating in the replication stress response can be distinguished. Indirect proteolysis involves cleavage of regulation factors and effectors or no longer needed components of replication machinery by regulatory proteases, often in a ubiquitin-dependent manner. Proteins causing replication blocking lesions could be cleaved by indirect proteolysis as well, but they require prior labeling and translocation. The other type of proteolysis involved, direct proteolysis, occurs *in situ*, on the DNA bound stalled replication fork. Here, the acting protease cleaves the lesion-causing protein component of DPC, protein noncovalently but tightly interacting with DNA strand and thus blocking replication or inhibitory inner component of replisome directly in the actual site of the collision.

Roles of indirect proteolysis processes in replication stress response are more studied and thus well established. The Ubiquitin-Proteasome System (UPS) is the key player here, participating in multiple proteolytic events. In addition to the timely degradation of cyclins and other molecules needed for the accurate onset and progression of the replication process, proteasome also directly participates in signaling networks mediating the replication stress response [Garcia-Rodriguez et al., 2016]. Particularly, it is responsible for the degradation of CHK1 kinase in both ATR phosphorylation and ubiquitination dependent manner. Degradation of phosphorylation-activated CHK1 may promote checkpoint termination to limit the duration of CHK1 signaling during low-intensity replication stress [Zhang et al., 2005]. RNA Polymerase II (RNAP II) represents another target of UPS in the replication stress response. RNAPII stalled at a DNA lesion blocks progression of the replisome and triggers a coordinated rescue mechanism that requires ubiquitylation and degradation of RNAPII [Woudstra et al., 2002]. Lastly, even proteins serving as facilitators of DNA repair could turn into replication stress-causing obstacles. Double-strand break repair protein Ku80 facilitates the nonhomologous end-joining repair mechanism of the double-strand break. Owing to its mechanism of action and sterical properties, some of the Ku80 molecules end up trapped around the now-repaired molecule of DNA in a ring-shaped conformation. Such a repair pathway byproduct represents an obstacle for progressing replisome and must, therefore, be removed. The clearance of trapped Ku80 is mediated by K48-linked polyubiquitination, followed by the recruitment of Valosin-Containing Protein/protein 97 (VCP/p97). Translocase p97 then unfolds the trapped protein and extracts it from the chromatin in an ATP-dependent manner. The whole process is finished by proteasome-mediated degradation of polyubiquitinated Ku80 [van den Boom et al., 2016].

1.5. DIRECT PROTEOLYSIS IN REPLICATION STRESS RESPONSE

In addition to the common and well-studied indirect cellular proteolytic response to replication stress, new proteases, which seem to be acting in parallel with the above mentioned proteolytic pathways, have been recently identified. Unlike the regulatory proteases, those new proteases act through direct proteolysis of proteins blocking the replication fork. In general, there is not much known about those proteolytic events nor their regulation, yet. Few substrates are known, mainly stemming from replication fork progression blocking lesions of protein nature, such as enzymes operating on DNA strand, other proteins covalently crosslinked to DNA, or replication regulating factors blocking restart of a paused fork.

Currently, there are three groups of proteases known for the direct proteolytic response to replication stress. Metalloproteases of the WSS1/Spartan family were the first identified [Stingele

et al., 2014]. The second group of proteases involved is the Ddi1-like family of aspartic proteases [Kottemann et al., 2018]. Surprisingly, recent literature shows that, in addition to its regulatory proteolysis role, 26S proteasome can be recruited directly to sites of DPC lesions and utilized as a DPC protease [Larsen et al., 2019].

1.5.1. DNA dependent metalloproteases

The first protease identified as participating in the direct proteolysis in response to replication stress was the *S. cerevisiae* Wss1p protein [Stingele et al., 2014]. First identified as a Weak Suppressor of Smt3 (*WSS1*), Wss1p is a 296 amino acid long protein belonging to the minigluzincin group of zinc endometalloproteases [Biggins et al., 2001]. Structurally, Wss1p possesses a multidomain architecture with a major part of the N-terminal half of the protein occupied by the metalloprotease domain. C-terminal half of the protein consists of various interaction-mediating domains and motifs with a major part harbouring a domain responsible for DNA binding, flanked by two peptide motifs recruiting Cell division control 48 protein (Cdc48p) ATPase/translocase (SHP and VCP-interacting motif (VIM)). The very C-terminus then provides two peptidic SUMO Interacting Motifs (SIMs), SIM1 and SIM2, binding to the Small Ubiquitin-like MOdifier (SUMO) Smt3p, or SUMO chains [Stingele et al., 2014]. A detailed three-dimensional structure of the Wss1p protease domain, shown in figure 3, was solved and provides molecular details pointing to the catalytic mechanism of the enzyme. Central to the structure is a canonical zinc-binding motif of minigluzincin family of proteases H-E-X-X-H at the catalytic center conserved among all orthologues. The glutamate residue E116 serves as a proton shuttle in the proteolytic reaction and its mutation to glutamine renders the protein catalytically inactive [Yang et al., 2017].

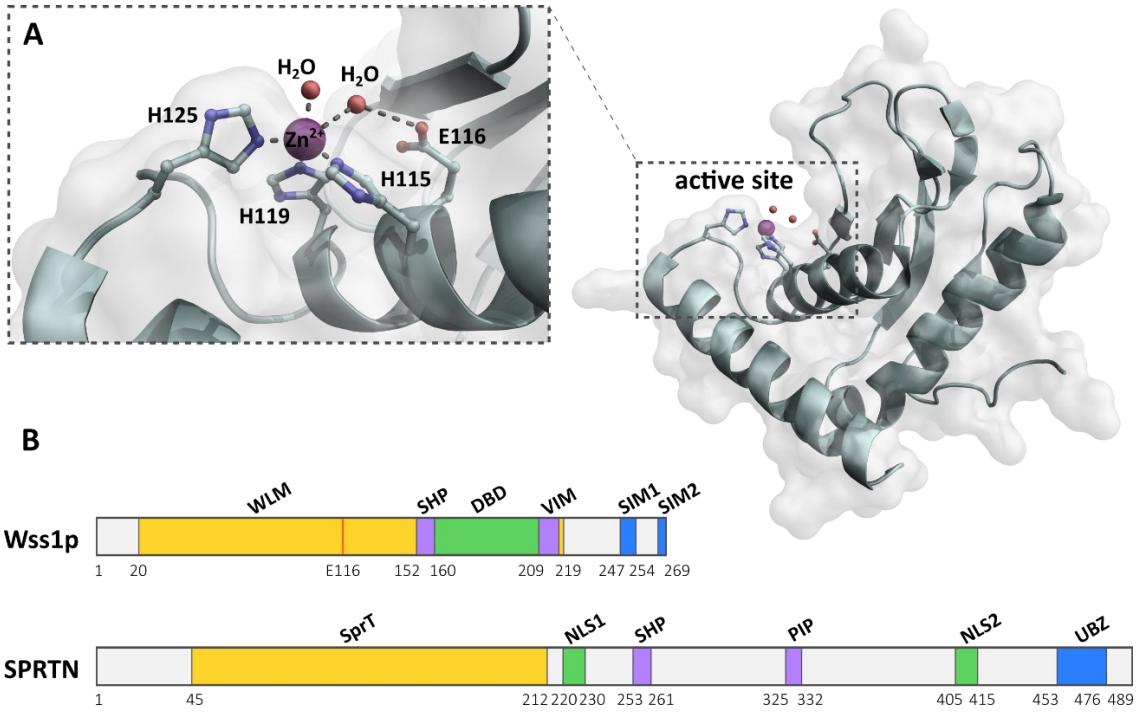


Figure 3: Structure of Wss1p/SPRTN proteins. A) 3D structure of *S.cerevisiae* Wss1p WLM protease domain. Detailed organization of the active site with Zinc ion coordinated by histidines and proton shuttling catalytic glutamate E116 shown in the inset. Adapted from PDB structure 5XBN [Yang et al., 2017]. B) Schematic depiction of domain organization in *S.cerevisiae* Wss1p and human SPRTN proteins.

High throughput genetic analysis repeatedly identified the gene encoding *WSS1* protease as being important for resistance to various stress-inducing factors, mainly those targeting DNA and cellular processes involving DNA [Pan et al., 2006, Bialkowska and Kurlandzka, 2002, Birrell et al., 2001]. The more targeted analysis revealed *WSS1* participation in stabilization or processing of a stalled replication fork, where it collaborates with the DNA damage response phosphatase Psy2p [O'Neill et al., 2004]. Interestingly, Wss1p is localized only in one sharp spot close to the nuclear membrane. In dividing budding cells this localization is exclusive to the mother cell [van Heusden and Steensma, 2008]. The role of *WSS1* in the sumoylation/desumoylation network remains unclear. Although Wss1p must be connected to Smt3p-based transactions in a cell as its overexpression was identified as suppressing phenotypes of temperature-sensitive mutants in the SMT3 gene [Biggins et al., 2001] and the protein itself harbours two SIMs, experimental evidence for the molecular mechanism remains highly conflicting. There is a publication by Mullen and colleagues suggesting Wss1p acts as a desumoylase [Mullen et al., 2011, Mullen et al., 2010], but other authors studies failed to replicate such activity [Stingele et al., 2014]. Balakirev and colleagues, on the other hand, suggest a balanced *WSS1* dependent mechanism where Wss1p acts even as a novel Smt3p ligase and at the same time cleaves oligomeric Smt3p chains but not with a typical

desumoylase activity [Balakirev et al., 2015]. Altogether, conclusions about Smt3p-related activity of Wss1p are highly conflicting and need to be refined by further research.

The importance of Wss1p protease for stabilization and rescue of replication fork was revealed by Stingele and colleagues. They provided evidence for Wss1p role as a DNA activated protease, acting on DNA-protein crosslinks [Stingele et al., 2014]. Here, as it is schematically illustrated in figure 4, Wss1p is recruited to the stalled replication fork blocked by a protein lesion covalently crosslinked to DNA molecule. Bulky DPC lesion blocks not only polymerase progression but the helicase activity as well and the cell, therefore, cannot switch on the ssDNA activated ATR-mediated DNA damage signaling. Recruited Wss1p is activated by binding to DNA and nonspecifically cleaves and degrades the bulk of the protein portion of the lesion. The small peptide lesion crosslinked to DNA left behind no longer blocks the progression of the replicative helicase and allows the helicase-polymerase uncoupling and activation of standard ATR mediated response to replication stress [Stingele et al., 2014].

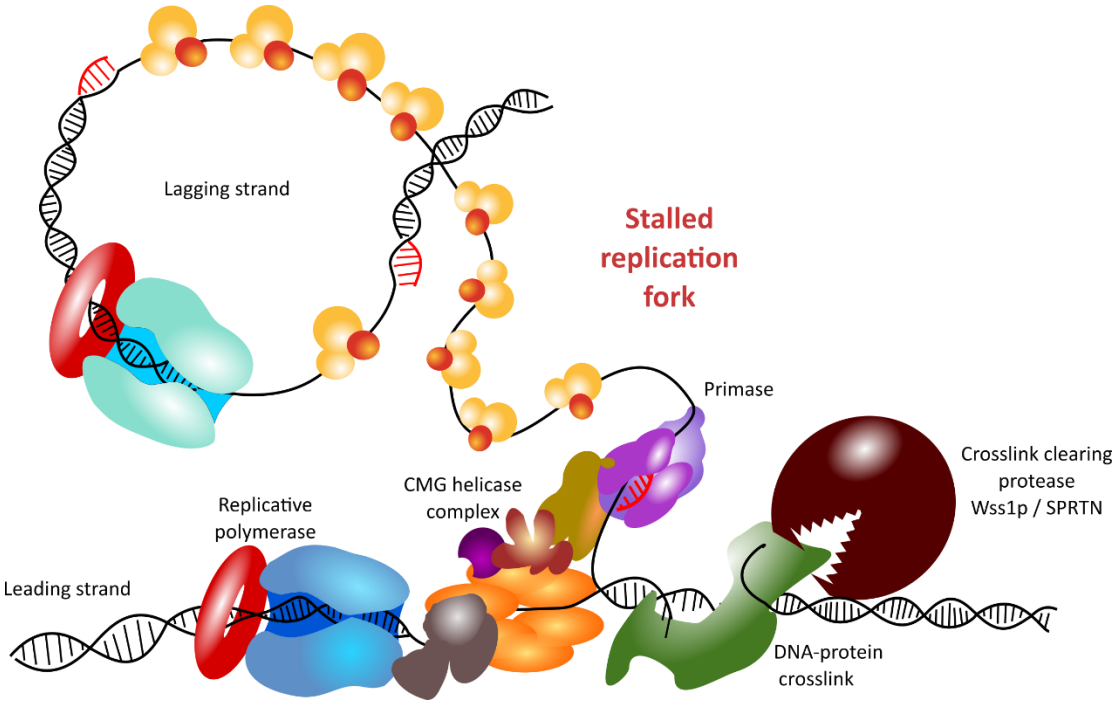


Figure 4: Schematic representation of Wss1/SPRTN protease-dependent clearance of a DNA-protein crosslink lesion in collision with the stalled replication fork.

Nonspecific active protease in the inner space of a replicating cell nucleus presents a dangerous double-edged sword and the activity of Wss1p must, therefore, be tightly regulated. Under normal conditions, Wss1p is kept at very low abundance in autoinhibited form. Proteolytic activity is released by interaction with DNA, particularly ssDNA, with a minimal DNA string length of 16 bases needed for activation. An alternative way for activation is provided by interaction with the thiol-modifying compound thiram suggesting a cysteine switch mechanism. Clearance of the

unnecessary protein, after it releases the replication stress-causing lesion, is mediated by autoproteolysis. Another layer of regulation is mediated by Wss1p interactions with Smt3p and Cdc48p segregase/translocase as it was demonstrated for cleavage of topoisomerase covalently trapped on DNA [Balakirev et al., 2015, Stingele et al., 2014].

Recently Madi and colleagues provided evidence that Wss1p not only cleaves protein trapped covalently on DNA strand but resolves also replication stress stemming from lack of resources for replication. They showed that Wss1p provides tolerance for hydroxyurea induced replication stress by the degradation of core histone subunits non-specifically and non-covalently trapped on ssDNA. Unlike Wss1p-dependent proteolysis of covalent DNA-protein crosslinks, proteolysis of histones does not require Cdc48p nor SUMO-binding activities [Maddi et al., 2020].

In mammalian cells, clearance of proteins covalently trapped on DNA is facilitated by a close relative of Wss1p protein from the same minigluzincin family of proteases, SprT-like domain-containing protease Spartan (SPRTN) [Vaz et al., 2016, Morocz et al., 2017]. SPRTN protease shares the zinc endometalloprotease activity but differs slightly in regulation. Although the activation by DNA interaction is conserved, SPRTN is autoinhibited by a monoubiquitination switch-based system and upon recruitment to the replication fork interacts with the Proliferating Cell Nuclear Antigen (PCNA) complex and moves along the leading strand together with PCNA [Li et al., 2019, Wang et al., 2016, Stingele et al., 2016]. Deficiencies in *SPRTN* manifest clinically in progeroid phenotypes and increased tumorigenesis [Lessel et al., 2014, Maskey et al., 2014, Lopez-Mosqueda et al., 2016].

1.5.2. Ddi1-like aspartic proteases

Ddi1-like proteases form a group of aspartic proteases. The first described member of this family was Ddi1p protein from buddying yeast *S. cerevisiae*. Ddi1p is a multidomain protein with various described functions inside a cell. Therefore, it comes as no surprise that this protein was first characterized on two independent occasions. First, it was described as a gene product induced by DNA damaging chemicals and named *DDI1* – DNA damage-inducible 1 [Liu and Xiao, 1997]. One year later, another group of researchers identified this protein as a regulator of SNARE proteins in vesicular transport and named it Vsm1p – vSNARE master 1 [Lustgarten and Gerst, 1999]. Although it may lead to confusion in early literature, both those names refer to the same protein, 428 amino acid long, harbouring multidomain architecture, containing N-terminal UBiquitin-Like (UBL), C-terminal UBiquitin Associated (UBA) domains, with a characteristic RetroViral aspartic Protease-like (RVP) domain in the middle [Sirkis et al., 2006]. RVP domain provides a dimerization interface and all Ddi1-like proteins homodimerize via RVP, forming a fold

structurally resembling proteases from retroviruses such as Human Immunodeficiency Virus (HIV) [Kumar and Suguna, 2018]. Structural features of Ddi1-like proteins, together with characteristic domain composition are shown in the figure 5. Ddi1-like proteins are conserved throughout the eukaryotic organisms, with RVP aspartic protease domain being the most conserved part throughout species [Krylov and Koonin, 2001].

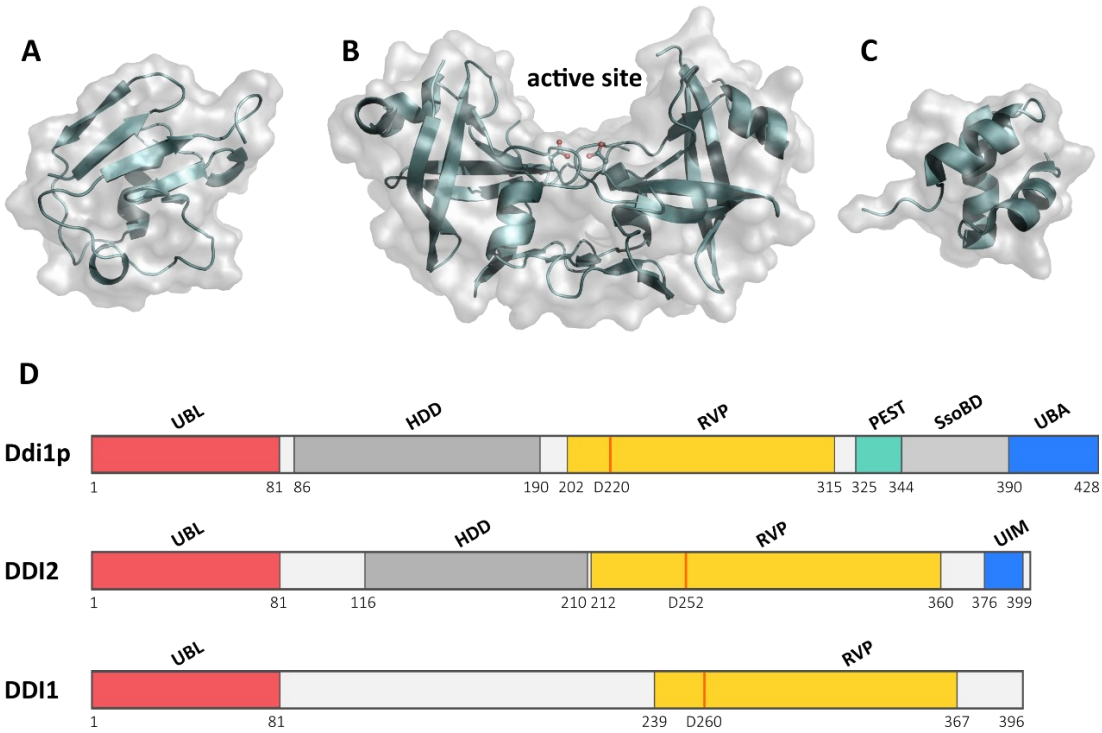


Figure 5: Structural features of Ddi1-like family of proteins. A) 3D structure of *S.cerevisiae* UBL domain adapted from PDB entry 2MRP [Nowicka et al., 2015]. B) 3D structure of *S.cerevisiae* RVP domain with highlighted catalytic aspartates Asp220, adapted from PDB entry 2I1A [Sirkis et al., 2006]. C) 3D structure of *S.cerevisiae* UBA adapted from PDB entry 2MR9 [Nowicka et al., 2015]. D) Schematic depiction of domain organization in *S.cerevisiae* Ddi1p and human DDI1 and DDI2 proteins with highlighted catalytic aspartates.

1.5.2.1. Yeast Ddi1

In *S. cerevisiae,* the single exon gene *YER143w* encoding Ddi1p protein is located on chromosome V. Product of the *DDI1* gene is transcribed from the bidirectional promoter shared with *MAG1*, 3-methyl adenine DNA glycosylase, the first enzyme involved in base excision repair pathway. Expression of both *DDI1* and *MAG1* is inducible by alkylating chemicals and other agents damaging DNA such as hydroxyurea, 4-nitroquinoline 1-oxide, or ultraviolet (UV) radiation [Liu and Xiao, 1997, Liu et al., 1997], but unlike *MAG1, DDI1* expression inducibility is regulated not only by Radiation sensitive 53 (*RAD53*) kinase dependent S-phase DNA damage checkpoint pathway but by Precocious Dissociation of Sisters 1 (*PDS1*) securin dependent DNA damage checkpoint pathway as well [Zhu and Xiao, 1998]. Overexpression of *DDI1* leads to dosage

suppression of *pds1-128* temperature-sensitive defect, specifically in the S-phase checkpoint pathway [Clarke et al., 2001]. Expression and inducibility of both *DDI1* and *MAG1* are directly controlled by the Pleiotropic Drug Resistance 3 (Pdr3p) transcriptional activator and indirectly affected by Rpn4p (Regulatory Particle Non-ATPase), a transcription factor for proteasome subunits [Zhu and Xiao, 2004] and Rad6p-Rad18p ubiquitination complex [Fu et al., 2008].

In addition to DNA damage inducibility, *DDI1* also acts as a negative regulator of constitutive exocytosis. It was identified in a yeast two-hybrid screen for Snc1p (Suppressor of the null allele of CAP) vesicle membrane receptor (vSNARE) regulators and deletion yeast strain lacking *DDI1* displayed an increased level of constitutive exocytosis, whereas *DDI1* overexpression led to decreased exocytosis [Lustgarten and Gerst, 1999]. Mechanistically Ddi1p interacts with target membrane receptor (tSNARE) protein Sso1p (Supressor of Sec One 1), thus blocking Sso1p-Sec9p (SECretory 9) interaction and formation of the tSNARE complex. Ddi1p-Sso1p interaction is a stoichiometric inhibitory interaction, independent of the Ddi1p UBA domain, and does not lead to Sso1p degradation [Marash and Gerst, 2003]. Ddi1p binds to Sso1p via amino acids 344 – 395, located between RVP and UBL domains, and phosphorylation of both proteins is important for the interaction [Gabriely et al., 2008]. The role of the catalytic activity of the RVP domain in this process was suggested as well [White et al., 2011a]. Recently, *DDI1* was found to have a role in Guanine nucleotide-binding protein Gpa1p (G Protein Alpha subunit 1) [Dixit et al., 2014] endocytosis and in selective anterograde transport of Cps1p (CarboxyPeptidase yscS 1) carboxypeptidase oligomers into endosomal multivesicular bodies [Kama et al., 2018]. Unlike in exocytosis regulation, here the Ddi1p function depends on its UBA domain and its interaction with a monoubiquitinylated partner.

The best-described function of yeast *DDI1* is its action in the delivery of polyubiquitinylated substrates to the proteasome, acting as a so-called proteasomal shuttling protein. Proteasomal shuttling proteins form a group of UBL and UBA domain-containing proteins and besides *DDI1* is represented by RADiation sensitive 23 (*RAD23)* and *DSK2* in yeast. C-terminal or internal UBA domain of shuttling proteins interacts with a polyubiquitin chain on substrate destined to proteasomal degradation and shuttling protein brings such substrate to the 26S proteasome complex. Here, the UBL domain of shuttling protein binds to one of the ubiquitin/UBL binding subunits on proteasome 19S regulatory particle, and the substrate is deubiquitinlylated, translocated to 20S proteasome catalytic particle and degraded. Shuttling protein is then released and recycled. Shuttling of specific polyubiquitinylated targets to proteasome adds another level of control to this pathway of regulatory proteolysis [Elsasser and Finley, 2005]. Affinity in all the above-mentioned interactions, together with the total amount of shuttling protein must be very fine-tuned in a cell.

High levels of proteasome shuttling protein can lead to the stabilization of ubiquitinylated substrates, due to masking of polyubiquitin chain by one fraction of shuttle protein and blocking proteasome recognition sites by the other, leading to UPS inhibition [Verma et al., 2004].

Saccharomyces cerevisiae Ddi1p fulfills all criteria for proteasome shuttling protein. UBA of Ddi1p was shown to bind both monoubiquitin and polyubiquitin chains, with a slight preference for lysine 48-linked chains [Bertolaet et al., 2001b, Saeki et al., 2002, Trempe et al., 2005, Nowicka et al., 2015]. The presence of the UBA domain stabilizes Ddi1p in yeast cell by protecting itself from proteasomal degradation [Heessen et al., 2005]. UBL domain binds to proteasome via an unconventional recognition site on the Regulatory particle non-ATPase 1 (Rpn1p) proteasome subunit [Saeki et al., 2002, Gomez et al., 2011]. Surprisingly, a recent study identified UBL of Ddi1p not only as an interaction partner for proteasome but also as an unconventional binder of ubiquitin. Here, UBL binds to ubiquitin cooperatively via a hydrophobic patch in the β-sheet area and via unconventional surface electrostatics, different from other UBLs in proteasome shuttling proteins and the Ddi1p UBL-ubiquitin interaction is strong enough for immobilized UBL to pull a fraction of polyubiquitinylated protein from yeast cell lysate [Nowicka et al., 2015]. Unlike Rad23p, Ddi1p does not facilitate degradation of artificial Ubiquitin Fusion Degradation (UFD) substrate, although it can be forced to do so by manipulating Ufd2p ubiquitin E4 ligase to recruit Ddi1p to this ubiquitin ligase complex [Kim et al., 2004]. Moreover, *DDI1* is partially functionally redundant with *RAD23* [Diaz-Martinez et al., 2006] and heterodimerizes with Rad23p via their UBL, UBA domains [Bertolaet et al., 2001a]. Ddi1p-Rad23p interaction does not inhibit ubiquitin recognition by either of the interacting proteins. On the contrary, heterodimerization facilitates the formation of the Ddi1p-Rad23p-K48-linked tetraubiquitin chain trimeric complex *in vitro*, with a suggested role in blocking unnecessary additional ubiquitination or preventing deubiquitination while transporting substrate to the proteasome. Simultaneous deletion of *DDI1* and *RAD23* leads to the accumulation of polyubiquitinylated proteins in the cell [Kang et al., 2006].

Known substrate repertoire, targeted by Ddi1p for proteasomal degradation, consists of two proteins, HOmothallic switching (HO) endonuclease, and F-box protein Ufo1p (UV-F-box-HO 1). Both substrates are involved in the DNA damage response. HO endonuclease initiates the yeast mating-type switch process by creating a DNA double-strand break at the *MAT* locus. The presence of active HO endonuclease inside a cell nucleus is tightly regulated, with expression restricted to mother cells in the late G1 cell cycle phase, once a cell is committed to the following duplication. Its cellular halftime is around 10 minutes under native conditions, where, after endonuclease cleavage, HO endonuclease is displaced from the nucleus to cytoplasm in a *MEC1* (Mitosis Entry Checkpoint 1) mediated manner, ubiquitinylated by multiprotein SCF

(Skp1–Cullin–F-box-protein) ubiquitin E3 ligase complex and degraded by the proteasome. Ddi1p was shown to copurify with polyubiquitinylated HO endonuclease via its UBA domain, whereas the UBL domain of Ddi1p is required for HO binding to proteasomal Rpn1p subunit. The deletion of the *DDI1* gene leads to the accumulation of undegraded polyubiquitinated HO endonuclease in the cytoplasm [Kaplun et al., 2005]. The second substrate targeted by Ddi1p, F-box protein Ufo1p, is actually a F-box part of the multiprotein SCF E3 ubiquitin ligase complex, responsible for HO endonuclease ubiquitination. Ufo1p is an unusual F-box protein, involved in DNA damage response, which contains four copies of C-terminal ubiquitin interacting motifs (UIM), short helical peptide sequences known to bind ubiquitin. Although the *UFO1* deletion strain is viable, deleting only the UIM part of *UFO1* leads to a dominant-negative lethality. Ectopic overexpression of *UFO1* without UIMs brings cell to G1/S arrest with a morphology dominated by long, extended buds, a phenotype similar to a temperature-sensitive mutant of *SKP1* (Suppressor of Kinetochore Protein mutant 1) member of SCF complex at the restrictive temperature. The same phenotype can be observed upon overexpression of full-length *UFO1* in yeast strain with deleted *DDI1* [Ivantsiv et al., 2006]. Ddi1p UBL interacts with Ufo1p via those UIM motifs and facilitates Ufo1p proteasome-mediated degradation. When not degraded, Ufo1p blocks an F-box exchange-driven recycling of the SCF E3 complex, causing cell cycle arrest [Kaplun et al., 2006]. Interaction between Ufo1p, HO endonuclease, and Ddi1p was confirmed by *in vitro* reconstitution of the complex [Voloshin et al., 2012].

Another cell cycle-related phenotype of *DDI1* is suppression of temperature-sensitive mutant *pds1-128* in *PDS1* securin by overexpression of *DDI1* (as well as by overexpression of *RAD23*). Pds1p securin is a separase inhibitor with a role in three cell cycle checkpoints, spindle assembly checkpoint, DNA damage checkpoint, and DNA replication checkpoint. *DDI1* role in *pds1-128* suppression is limited to the *PDS1* role in DNA replication checkpoint [Clarke et al., 2001] and is dependent on Ddi1p UBA and UBL domains [Gabriely et al., 2008]. Interestingly, unlike pds1-128 mutant, the full deletion of the *PDS1* gene can be rescued by neither *DDI1* nor *RAD23* overexpression. This observation suggests a requirement for some form of Pds1p, although mutant one, for cell cycle progression, thus pointing toward the role of UBL-UBA proteins in the stabilization of Pds1p rather than facilitating its degradation. Double deletion of both *DDI1* and *RAD23* genes leads to increased sensitivity to hydroxyurea, as does, although to a lesser extent, the deletion of both UBA domains only [Clarke et al., 2001]. The functional importance and stabilization effect of yeast Ddi1p UBA domain is highlighted by its targeting by yeast metacaspase upon glucose treatment after nutrient starvation which leads to UBA cleavage followed by Ddi1p proteolytic degradation [Bouvier et al., 2018].

1.5.2.2. Ddi1-like proteins in other eukaryotes

Functions of the Ddi1-like family of proteins were investigated in other eukaryotic model organisms as well, with human Ddi1-like proteins, especially human DNA Damage-Inducible 1 Homologue 2 (DDI2) being the most extensively studied. As mentioned above, Ddi1-like proteins are conserved throughout eukaryotes, although some changes on the domain level can be traced through eukaryotic evolution. Plants and fungi in general follow the *S. cerevisiae* domain architecture, although for example *S. pombe* apparently lost the UBL domain. In animals, the tree domain architecture known from *S. cerevisiae* is retained up to Chordata. Gene duplication from retrotransposition event apparently occurred in Chordata common ancestor, resulting in two homologues of Ddi1-like family genes present in their genomes. The original one was named *DDI2*, whereas the probably retrotransposed single-exon gene was named DNA Damage-Inducible 1 Homologue 1 (*DDI1*). Ddi1-like proteins in Chordata also lost their UBA domains, which probably migrated, via an insertion event, to *RSC1A1* (Regulator of Solute Carriers 1), a gene downstream from *DDI2*. Replacing the UBA domain, a UIM motif on the last helix of Ddi1-like proteins of Chordata was predicted.

Several protozoan Ddi1-like proteins were studied based on observation, that HIV positive patients coinfected with a protozoan parasite show improvement in protozoan infection upon treatment with HIV protease inhibitor [Lopez-Velez, 2003]. *DDI1* is one of only two aspartic proteases encoded in genomes of *Leishmania* genus parasites, the other being a membrane-bound extracellular protease of presenilin type. Indeed, HIV protease inhibitors were shown to bind and inhibit *Leishmania major* Ddi1p RVP, although in an indirect, yeast cell-based assay [White et al., 2011b]. Currently, *L. major* Ddi1p is the only Ddi1-like protein with *in vitro* measures enzyme activity of a purified protein [Perteguer et al., 2013]. The structure of *L. major* Ddi1p RVP resembles that of yeast and human Ddi1-like RVP domains [Kumar and Suguna, 2018]. Ddi1 proteins from other protozoa, namely *Plasmodium berghei* [Onchieku et al., 2018] and *Trypanosoma cruzi* [Castilho et al., 2018] were studied by homology modeling and *in silico* inhibitor docking studies with results suggesting possible binding of HIV protease inhibitors.

Drosophila melanogaster Ddi1-like gene, rngo (rings lost), is an essential gene and its disruption in the germline leads to sterility due to defects in oogenesis caused by disruption of ring canal formation. Interestingly, although Rngo protein was shown to bind ubiquitin as well as proteasome subunit Rpn10 (regulatory particle non-ATPase 10), both UBL and UBA domains of Rngo were dispensable in the rescue experiment. On the contrary, RVP and also the presence of putative catalytic aspartate were essential for ring canal formation rescue [Morawe et al., 2011]. *D. melanogaster* Rngo was also identified in neuron cells as one of the primary targets for

nondegradative K48 polyubiquitination by E3 Ubiquitin protein ligase E3A (Ube3a) defective in Angelman syndrome [Ramirez et al., 2018]. Ube3a driven Rngo ubiquitylation in *D. melanogaster* neurons is reverted by the Ubiquitin Specific Peptidase 9 X-Linked (USP9X) deubiquitinase [Elu et al., 2019].

Caenorhabditis elegans strains lacking functional *ddi-1* display increased synaptic connectivity and enhanced synaptic density [Guthmueller et al., 2011]. Moreover, *C. elegans ddi-1* is involved in the degradation of the SLO-1 BK channel, regulating locomotion, neurotransmitter release, and alcohol resistance [Oh et al., 2017]. Under proteasome inhibition or dysfunction, *ddi-1* is participating in a feedback loop leading to the expression of new proteasome subunits. Transcription of proteasome subunits is driven by transcription factor *skn-1*. Under normal conditions, SKN-1 is an endoplasmic reticulum (ER) membrane protein with a very short lifetime. After translation, SKN-1 is quickly ubiquitinylated by ER associated degradation (ERAD) ubiquitin ligase complex HRD-3 (HMG-coA Reductase Degradation 3), retrotranslocated by Cdc48, deglycosylated by Peptide N-Glycanase 1 (PNG-1), and degraded by the proteasome in usual ERAD pathway. With proteasome dysfunction, SKN-1 cannot be effectively degraded, accumulates, and is cleaved by DDI-1, most probably by the activity of its RVP aspartic protease domain. Processed SKN-1 then translocates into the nucleus where it serves as a transcription factor driving the expression of several proteasome subunits [Lehrbach and Ruvkun, 2016].

The identical function of DDI2 protein in the processing of proteasome subunit expression regulating transcription factor is conserved in mammals as well. Here, as illustrated in figure 6, transcription factor 11/ nuclear factor E2-related factor 1 (TCF11/NRF1) encoded by *NFE2L1* (nuclear factor erythroid-2-like 1) gene is an ER-resident transcription factor that is continually retrotranslocated and degraded by the proteasome in ERAD dependent manner. After proteasome inhibition, TCF11/NRF1 avoids degradation and is proteolytically processed to its active form by DDI2 RVP aspartic protease activity [Koizumi et al., 2016, Nowak et al., 2018]. In addition to TCF11/NRF1, human DDI2 activates similarly also homologous transcription factor NRF3 (nuclear factor E2-related factor 3), involved in the regulation of expression of cell proliferation factors [Chowdhury et al., 2017, Bury et al., 2019].

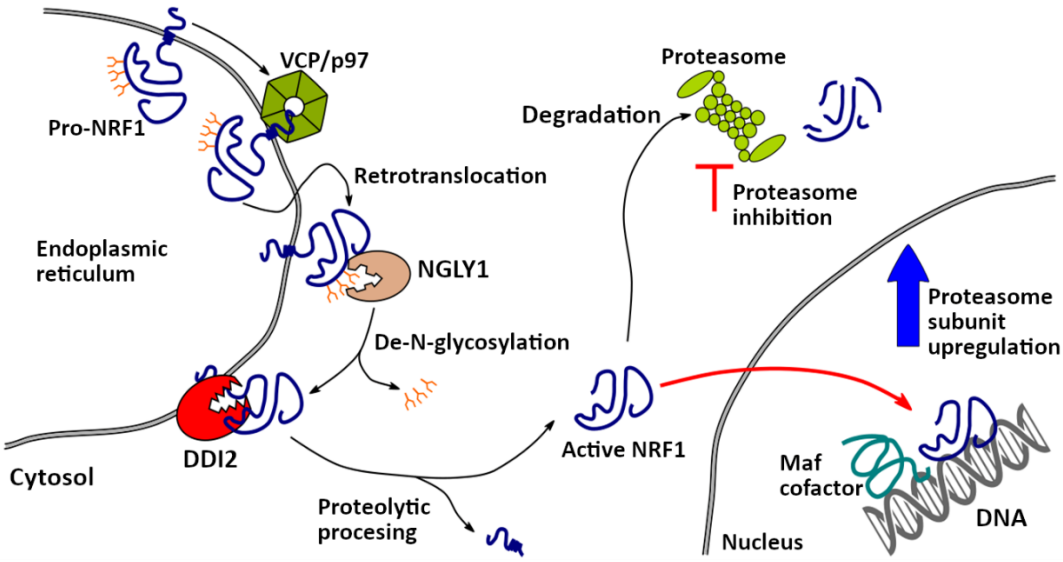


Figure 6: schematic representation of the TCF11/NRF1 activation mechanism via retrotranslocation followed by DDI2 mediated proteolytic cleavage

The most recent addition to the repertoire of mammalian Ddi1-like protein functions closes the circle and brings Ddi back to its original identification as DNA damage-inducible protein. Kotteman et. al. [Kottemann et al., 2018] identified both human *DDI1* and *DDI2* as factors involved in replication fork restart after encountering obstacles. DDI2 associates with active replisome and dual silencing of *DDI1* and *DDI2* results in increased sensitivity to replication blocking drugs hydroxyurea and aphidicolin. Based on proteomic data DDI1/2 seems to be removing replication termination factor 2 (RTF2) from stalled replication forks in a proposed proteasome shuttling mechanism and thus facilitating replication fork restart. RTF2 stabilization by dual DDI1/DDI2 depletion leads to impaired recovery from replication stress, accompanied by ssDNA accumulation, ATR kinase activation, and increased genome instability [Kottemann et al., 2018].

AIMS OF THE STUDY

2. AIMS OF THE STUDY

The main aim of this doctoral thesis was to elucidate a physiological function of a Ddi1-like class of proteases with a special focus on response to DNA replication stress, using molecular biology as well as biophysics.

Specific aims were:

1. To biophysically and structurally characterize human DDI2 protein, solve structures of its isolated domains, test *in vitro* proteolytic activity of its protease domain, and characterize DDI2 interaction with ubiquitin.
2. To biophysically and structurally characterize the Ddi1p protein of *S. cerevisiae*, solve the structure of its isolated domains, test *in vitro* proteolytic activity of its protease domain, and characterize yeast Ddi1p interaction with ubiquitin.
3. To characterize a biological function of the Ddi1p protein of *S. cerevisiae* in replication stress response, with a special focus on the role of its individual domains in this process.

Various biophysical techniques including nuclear magnetic resonance NMR, isothermal titration calorimetry, and mass spectrometry, together with molecular biology methods of classical yeast genetics and complementation were used as principle tools.

METHODS

3. METHODS

This chapter highlights only methods important for understanding the results described below. Detailed methods for all experiments performed are part of respective publications.

3.1. PROTEOMIC IDENTIFICATION OF CLEAVAGE SITES (PICS)

PICS was performed according to the previously published protocol [Schilling et al., 2011] with several modifications. Non-transfected cell cultures of either mammalian HEK293 cells or yeast *Saccharomyces cerevisiae* were grown under standard conditions, harvested, and washed with phosphate buffer saline (PBS). Cells were transferred into a hypotonic lysis buffer containing protease inhibitors and lysed by sonication (HEK293) or multiple passes through Emulsiflex high-pressure homogenizer (yeast cells), followed by cysteine reduction and alkylation. Proteins were precipitated using trichloroacetic acid and resolubilized into 200 mM HEPES, pH 7.5. Following resolubilization, the proteome-derived peptide library was prepared by cleaving isolated denatured proteins into peptides with a working protease (trypsin or GluC). After abolishing working protease activity using 1 mM PMSF, the second round of sulfhydryl reduction and alkylation was performed. Primary amines on peptide N-termini and lysine side chains were blocked by formaldehyde-cyanoborohydride reductive dimethylation. Unreacted or quenched modification reagents were removed by gel filtration, and the peptide library was purified and transferred to High Performance Liquid Chromatography (HPLC)-grade water using a C-18 solid-phase extraction cartridge. The peptide concentration was adjusted to 2 mg/ml. The integrity of the peptide library was confirmed by LC-MS/MS analysis, and aliquots were stored at −80 °C until further use.

For the PICS assay, the peptide library (final concentration 1 mg/ml) was incubated in 200 µl of the appropriate buffer with 4 µg of an assayed enzyme. The reaction was incubated for 12 h at 37 °C. Following incubation, the reaction was heat-inactivated for 30 min at 70 °C and titrated with 2 M HEPES, pH 8.5, to final pH 8.0.

Subsequently, newly formed primary amine groups (products of proteolytic cleavage) were biotinylated *in vitro* by the addition of sulfo-NHS-SS-biotin for 4 h at room temperature. Biotinylated products were then immobilized on streptavidin-agarose with mild agitation at room temperature, followed by washing. Additional washing steps (2 M urea followed by 20% isopropyl alcohol, 5% DMSO, and 5% acetonitrile, all in washing buffer (50 mM HEPES, 150 mM NaCl, pH 7.5)) were added to the original protocol, followed by 10 washes with washing buffer. Immobilized peptides were eluted with 20 mM DTT, desalted using C-18 reverse phase cartridges according to the manufacturer's protocol, and analyzed by mass spectrometry (MS).

As negative controls, catalytic aspartate mutants (yeast D220A and human D252A) and buffer alone (blank) were used. As positive controls, the HIV-1 protease cleavage profile in 100 mM sodium acetate, 300 mm NaCl, pH 4.7, was tested with wild-type and HIV-1(D25N) protease. The protease library ratio was 1:200, and the final enzyme concentration was 0.2 μM.

3.1.1. PICS data analysis

PICS data were analyzed using a series of predesigned queries in Microsoft Access database software. First, lists of identified peptides from each mass MS run were loaded into the database and filtered for peptides containing products of N-terminal modification by biotinylation. Second, peptides with confidence over 80% were picked for enzymatic reactions (active Ddi1-like protein or HIV-1 protease), whereas peptides with confidence over 10% were picked for control reactions (reactions with catalytically ineffective forms of the enzymes and blanks). To properly subtract the background signal, the list of peptides found in each enzymatic reaction was screened for peptides present in the blank reaction and the reaction with the catalytically ineffective enzyme. Such peptides were then removed from processing. Finally, the enzymatic reactions were screened for peptides identified in the original unprocessed peptide library. Such peptides were also removed from the analysis.

The final cleared list of identified peptides was then mapped against the FASTA database used for proteomics database search. By alignment of identified peptides with the database, the N-terminal portions of cleaved peptides (preceding the cleavage site) were determined. If there was more than one computationally identified amino acid for a given P position, the position was omitted from the processing, whereas the rest of the identified peptide sequence remained listed 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 position was then calculated and plotted, yielding a substrate specificity matrix for the assayed enzyme.

3.2. YEAST MANIPULATION

3.2.1. Yeast strains, growth conditions, and yeast plasmids construction

All *Saccharomyces cerevisiae* strains used were isogenic derivatives of strain S288C, in the Y7092 background (*MATα can1Δ::STE2pr-Sp_his5 lyp1Δ ura3Δ0 leu2Δ0 his3Δ1 met15Δ0*), and were obtained as a gift from professor Charles Boone (Donnelly Centre, University of Toronto). Standard yeast YPD (1% yeast extract, 2% peptone, 2% glucose) and synthetic complete (SC) or

SC dropout media were used [Sherman, 2002]. For solid media, 1.5% agar was added. All cultivations were performed at 30 °C unless explicitly stated otherwise; liquid cultures were incubated at 260 RPM in an orbital shaker.

Plasmids used for functional complementation experiments were derivatives of the centromere bound pAG416GPD-ccdB plasmid (Addgene plasmid #14148). Constructs coding Ddi1-like homologues were amplified by polymerase chain reaction (PCR) and inserted in the pDONR221 entry vector. Point mutated variants were created using Stratagene QuickChange Site-Directed Mutagenesis Kit, truncated variants were created by the Gibson assembly method. Final expression constructs were created by Gateway recombination cloning from pDONR221 to pAG416GPD-ccdB. All plasmids were verified by sequencing. Yeast cells were transformed using the standard LiAc/PEG3350 protocol.

3.2.2. Phenotypic characterization

Phenotypic characterization was performed by dilution spotting assays. Equal amounts of exponentially growing cells in dilution series were plated onto solid YPDA media in presence of the analyzed DNA damaging or replication blocking agent. Nonsuplemented YPDA solid media was used as a control. All spot tests for a given experiment were always plated on one plate. Plates were photographed after 60 h of incubation at 30 °C.

PUBLICATIONS

4. PUBLICATIONS

Publications included in the dissertation thesis

1 Monika Sivá*, **Michal Svoboda*,** Václav Veverka, Jean-Francois Trempe, Kay Hofmann, Milan Kožíšek, Rozálie Hexnerová, František Sedlák, Jan Belza, Jiří Brynda, Pavel Šácha, Martin Hubálek, Jana Starková, Iva Flaisigová, Jan Konvalinka, and Klára Grantz Šašková. Human DNA damage-inducible 2 protein is structurally and functionally distinct from its yeast ortholog. *Sci Rep.* **6**, 30443 (2016).

*equal contribution

2 Jean-François Trempe, Klára Grantz Šašková, Monika Sivá, Colin D H Ratcliffe, Václav Veverka, Annabelle Hoegl, Marie Menade, Xin Feng, Solomon Shenker, **Michal Svoboda,** Milan Kožíšek, Jan Konvalinka, and Kalle Gehring. Structural studies of the yeast DNA damage-inducible protein Ddi1 reveal domain architecture of this eukaryotic protein family. *Sci Rep.* **6**, 33671 (2016).

3 **Michal Svoboda,** Jan Konvalinka, Jean-Francois Trempe, and Klára Grantz Šašková. The yeast proteases Ddi1 and Wss1 are both involved in the DNA replication stress response. *DNA Repair.* **80**, 45–51 (2019).

Publications not included in the dissertation thesis

1 Jan Tykvart, Cyril Bařinka, **Michal Svoboda,** Václav Navrátil, Radko Souček, Martin Hubálek, Martin Hradilek, Pavel Šácha, Jacek Lubkowski, and Jan Konvalinka. Structural and biochemical characterization of a novel aminopeptidase from human intestine. *J Biol Chem.* **290**, 11321–11336 (2015).

2 Pavel Srb*, **Michal Svoboda***, Ladislav Benda, Martin Lepšík, Ján Tarábek, Václav Šícha, Bohumír Grüner, Klára Grantz-Šašková, Jiří Brynda, Pavlín Řezáčová, Jan Konvalinka, and Václav Veverka. Capturing a dynamically interacting inhibitor by paramagnetic NMR spectroscopy. *Phys Chem Chem Phys.* **21**, 5661–5673 (2019)

*equal contribution

4.1. PUBLICATION I:

Human DNA damage-inducible 2 protein is structurally and functionally distinct from its yeast ortholog

Sivá M.[*], **Svoboda M.[*],** Veverka V., Trempe J.F., HofmannK., Kožíšek M., Hexnerová R., Sedlák F., Belza J., Brynda J., Šácha P., Hubálek M., Starková J., Flaisigová I., Konvalinka J., and Grantz Šašková K. Human DNA-damage-inducible 2 protein is structurally and functionally distinct from its yeast ortholog. *Sci Rep.* **6**, 30443 (2016).

[*] Equal contribution.

My contribution: cloning of ΔUIM and catalytically inactive D252A variants of human *DDI2*; bacterial expression and purification of full length human DDI2 protein, its RVP domain and catalytically inactive variants of both; expression and purification of epitope-tagged full-length human DDI2 in mammalian cells; enzymatic synthesis of diubiquitin chains, pull-down assays for testing human DDI2 – ubiquitin interaction; PICS proteomics screen for substrates of DDI2 protease activity; BSA cleavage assay; surface electrostatics analysis.

4.1.1. Motivation of the study

Ddi1-like proteins are highly conserved through all eukaryotic organisms. Yeast and human Ddi-1 like homologues share a 30 % identity on the sequence level (56 % in the protease domain) and similar domain composition can be found in all eukaryotes. Although there is a substantial body of experimental work done on the elucidation of *DDI1* function in yeast *Saccharomyces cerevisiae*, both Ddi1-like proteins coded in the human genome (DDI2 and its homologue, DDI2) were understudied and their structure and function remained elusive.

The main goal of this study was to investigate the overall structure of human DDI2 protein, and detail structural features of its individual domains. Structural studies were followed by functional analysis of full-length protein and its isolated domains, with a special focus on putative catalytic activity and substrate profile of the RVP domain as well as on potential interactions with ubiquitin and other components of the ubiquitin-proteasome system. By shedding some light on human DDI2 structural and biophysical characteristics, we wanted to establish a stepping-stone for further studies of human DDI2 function in the biological processes in the context of the cell and the whole organism.

4.1.2. The surface electrostatic potential of the human DDI2 UBL domain differs from its yeast homologue

As Ddi1-like proteins in V*ertebrata* lost the C-terminal UBA domain to a *RSC1A1* gene, via an insertion event, a question arises, whether these proteins retained the ability to interact with ubiquitin. We did run a bioinformatical analysis and identified a peptidic UIM in the human DDI2 sequence. This motif showed weak but specific interaction with ubiquitin in nuclear magnetic resonance (NMR) experiments. New, unorthodox ubiquitin-binding moiety was recently identified in yeast Ddi1p [Nowicka et al., 2015]. In addition to ubiquitin binding via the canonical UBA domain, the UBL domain of yeast Ddi1p binds ubiquitin as well. Although the overall folds of UBL domain and ubiquitin molecule are similar, the surface charge distribution is almost inverted, and therefore yeast UBL and ubiquitin can interact via electrostatic interaction. To test whether this interaction exists for the human DDI2 UBL domain as well, we performed an NMR titration experiment with human DDI2 UBL titrated by ubiquitin and a parallel experiment where ubiquitin was titrated by human DDI2 protein lacking C-terminal UIM. Both experiments showed weak interaction in the millimolar range. To address this discrepancy between human and yeast UBLs we solved the solution structure of human DDI2 UBL using NMR (figure 7A) (Structure was solved and titration experiments were performed by Dr. Monika Sivá). In the following evaluation of the structure, I analyzed the surface electrostatic potential distribution using PDB2PQR and Adaptive Poisson Boltzmann Solver. As can be seen from figure 7B and 7C, although the secondary and tertiary structures of human and yeast UBLs are very close, the distribution of surface electrostatic potential differs, with yeast Ddi1p UBL surface being significantly more negative. Due to a strong positive charge on the ubiquitin surface (figure 7D) and the electrostatic nature of the interaction, the different surface electrostatic potential distribution could explain weaker interaction between ubiquitin and human DDI2 UBL, compared to the yeast homologue.

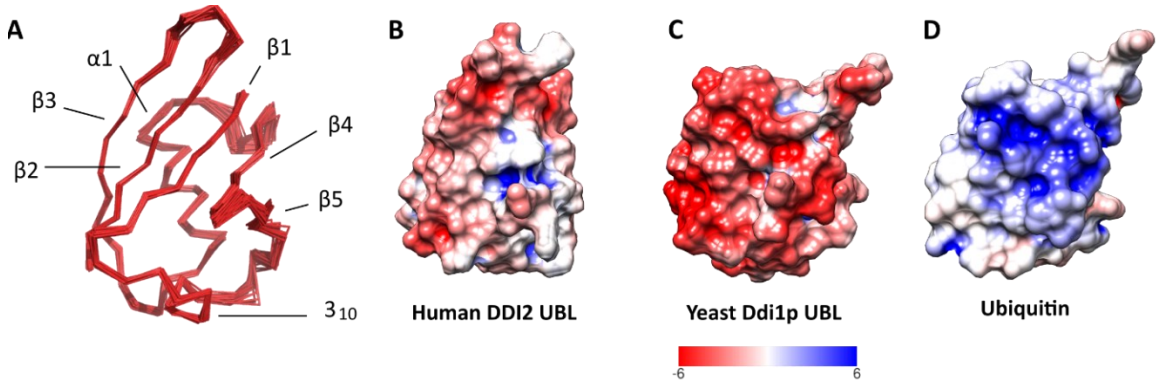


Figure 7: Solution structure of the human DDI2 UBL domain. **A)** Superimposition of 40 converged structures of the UBL domain. Comparison of the surface electrostatic potential

of **B)** human DDI2, **C)** yeast Ddi1p UBL (Publication II) and **D)** ubiquitin (from PDB code 1UBQ). 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 [Baker et al., 2001] package with structure preprocessing using the PDB2PQR tool [Dolinsky et al., 2004]. All calculations were performed using the SWANSON force field at pH 7.4; Chimera was used for final surface visualization.

4.1.3. DDI2 does not interact with diubiquitins

Although interaction with monoubiquitin can have biological significance, most of the cell regulation involving ubiquitin occurs through ubiquitin chains of various linkage architecture. To test whether human DDI2 interacts with polyubiquitin chains, with the potential for avidity mediated synergistic interaction strengthening, I performed pull-down experiments with various forms of immobilized full-length human DDI2 and diubiquitins of all eight types of covalent linkage architecture (Lys6-, Lys11-, Lys27-, Lys29, Lys33-, Lys48-, Lys63-linked and linear). As can be seen in figure 8A,B none of the eight used commercially available diubiquitin conjugates interacted with neither C-terminally (FLAG) nor N-terminally (FLAG or HA) tagged, resin immobilized, human DDI2 full-length variant. The same experiment was repeated also with in house synthesized Lys48- and Lys63-linked diubiquitins (figure 9). The data clearly show that human DDI2, in contrast to yeast Ddi1p, does not pull down any of the conjugates under physiological pH.

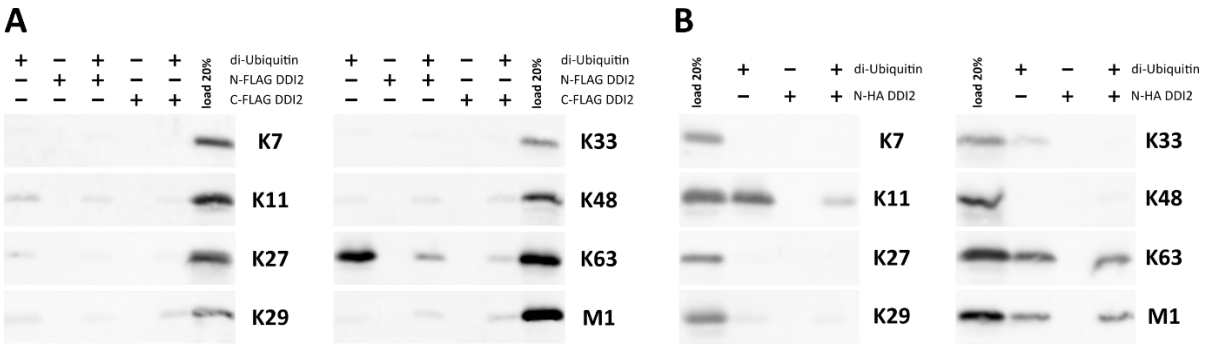


Figure 8: Polyubiquitin chain binding is not preserved in human DDI2. Western blot analysis of pull-down experiments with commercial diubiquitin conjugates of various covalent linkage architecture. Human DDI2 with a FLAG tag on either the N- or C-terminus (A) or an HA tag on the N-terminus (B) was immobilized on magnetic agarose beads. Beads were incubated with the diubiquitin conjugate of given linkage architecture, washed, eluted, and analyzed on 18% SDS-PAGE followed by immunoblotting with anti-ubiquitin antibody.

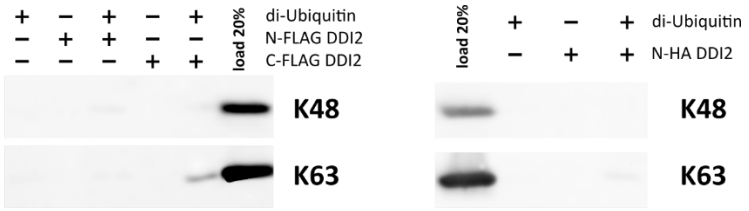


Figure 9: Human DDI2 shows no interaction with inhouse prepared diubiquitin chains. Western blot analysis of pull-down experiments with Lys48 and Lys63-linked diubiquitin conjugates prepared inhouse. Human DDI2 with a FLAG tag on either the N- or C-terminus or an HA tag on the N-terminus was immobilized on magnetic agarose beads. Beads were incubated with the diubiquitin conjugate of given linkage architecture and processed as in figure 8.

4.1.4. The protease domain of human DDI2 is not proteolytically active *in vitro*

Ddi1-like proteins are highly conserved throughout all eukaryotic organisms. Interestingly, the most conserved part across all species is the RVP domain, the putative aspartic protease of the retropepsin family. Although RVP could serve only as a homodimerization framework, its high level of conservation, as well as a presence of the typical retropepsin catalytic triad DT[S]G in all known Ddi1-like proteins, raises a question, whether RVP possesses specific catalytic activity and what is the physiological significance of such activity.

To examine the function of human DDI2 RVP we solved the three-dimensional structure of this domain. RVP structure resembles those of retroviral proteases with a C2 symmetrical dimer, multiple beta-sheet mediated dimerization interface, flexible beta hairpins, so-called flaps, covering active site cavity, and a catalytic triad coordinating water molecule, ready to serve as a nucleophile in a proteolytic reaction cycle. RVP structure was solved by Dr. Klára Grantz Šašková.

With information about structural elements being in place for enzyme catalytic action, I established an unbiased proteomic screen for the identification of peptidic substrates of the RVP domain based on the PICS (Proteomic Identification of protease Cleavage Sites) method. Peptide libraries derived from a mammalian cell culture-expressed proteome were prepared and cleaved overnight by *E. coli* expressed full-length human DDI2, using D252A putative catalytic site mutant as a negative control and structurally related HIV protease as a positive control. As can be seen in figure 10, I did not observe any peptide cleavage pattern related to DDI2 activity. This experiment was repeated with mammalian expressed full-length DDI2 as well as with the RVP domain itself. All DDI2 variants were assayed on two different variants of the peptide library (GluC and trypsin prepared) at pH 4.0, 5.0, and 7.0. None of these experiments led to protease activity detection.

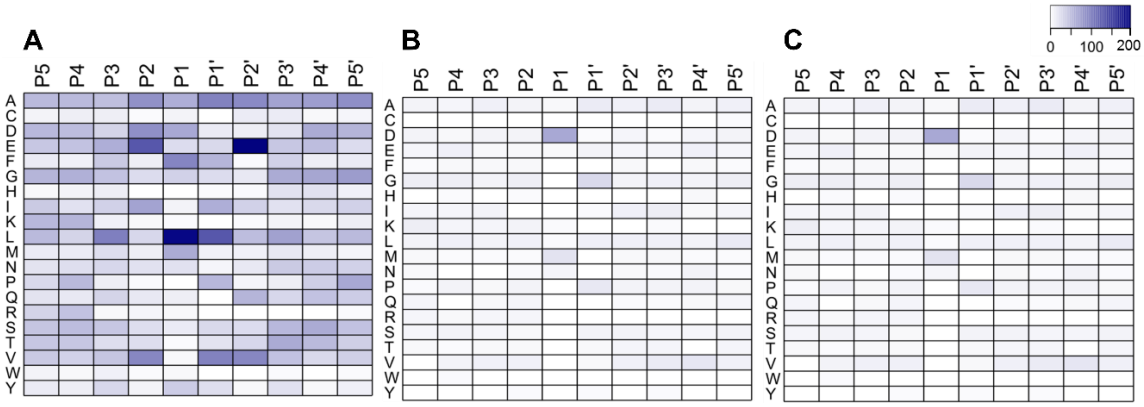


Figure 10: PICS results: Heatmap representation of protease substrate profile demonstrating the total counts of given amino acids (described by single letter code) per particular position (P5−P5´) in peptidic substrates in a HEK293 cell line-derived peptide library. The cleaved peptide bond is positioned between P1 and P1´. A) Library cleaved by HIV-1 protease (positive control) B) Library cleaved by human DDI2 protein in HEPES pH 7.0 (proteolytic activity assay). C) Background profile for the uncleaved library (negative control).

Furthermore, I tested human DDI2 for proteolytic activity against potential sole protein substrates BSA, HSA, casein, and insulin. Unlike its homologue from *Leishmania major*, human DDI2 did not cleaved any of those protein substrates.

4.1.5. Conclusion

In this publication, we reported the first structural and functional analysis of a mammalian protein from the Ddi1-like family, human DDI2. We identified two previously unknown structural elements in DDI2, C-terminal UIM, and Helical Domain of Ddi1-like proteins (HDD) located between UBL and RVP domains. We solved the X-ray structure of the DDI2 RVP domain and solution NMR structures of the HDD domain and the UBL domain. Using NMR titrations we identified weak but specific interaction between human DDI2 and ubiquitin, mediated by both UIM and UBL domains of DDI2. The surface electrostatic potential analysis showed that the considerably weaker UBL-ubiquitin interaction, compared to its yeast homologue, could be caused by the less negatively charged surface of the human DDI2 UBL. Surprisingly, although human DDI2 interacts with monoubiquitin, we did not observe interaction with any of the eight diubiquitin covalent isopeptide conjugates. Moreover, we thoroughly studied the RVP domain of human DDI2, solved its three-dimensional (3D) structure by protein crystallography, and searched for its proteolytic activity. Although we performed an extensive *in vitro* testing, we did not identify any human DDI2 related proteolytic cleavage events.

4.2. PUBLICATION II:

Structural studies of the yeast DNA damage-inducible protein Ddi1 reveal domain architecture of this eukaryotic protein family

Trempe J.F., Grantz Šašková K., Sivá M., Ratcliffe C.D.H., Veverka V., Hoegl A., Menade M., Feng X., Shenker S., **Svoboda M.,** Kožíšek M., Konvalinka J., and Gehring K. Structural studies of the yeast DNA damage-inducible protein Ddi1 reveal domain architecture of this eukaryotic protein family. *Sci Rep.* **6**, 33671 (2016).

My contribution: cloning of ΔUBA and catalytically inactive D220A variants of yeast Ddi1p; expression and purification of full-length yeast Ddi1p protein, its RVP domain and catalytically inactive variants of both; expression and purification of ΔUBA variant of yeast Ddi1p; sample preparation and assistance with ITC measurement; K48-linked diubiquitin synthesis; PICS proteomics screen for substrates of Ddi1p protease activity.

4.2.1. Motivation of the study

Yeast *DDI1* was identified by two independent teams. First, the *DDI1* gene was found to increase its expression in response to genotoxic agents, with its product involved in the UPS and cell cycle regulation. The second independent identification of yeast *DDI1* comes, originally under the name Vsm1, from its involvement in the regulation of constitutive exocytosis via SNARE protein complexes. Interestingly, all known functions of Ddi1p in yeast physiology do require intact catalytic active site of the RVP domain, with Asp220 mutations having a detrimental effect. Moreover, most of the yeast Ddi1p functions do require at least one member of the ubiquitin-related domain pair (UBL or UBA) to be present in the protein sequence.

Although the requirement for catalytic Asp220 points toward the involvement of RVP catalytic activity in yeast physiology, there was no substrate known for the protease domain neither *in vivo* nor *in vitro*. To decipher the role of yeast Ddi1p proteolytic activity in yeast physiology we decided to explore its substrate preference and structural determinants of substrate recognition. Moreover, we biophysically analyzed the interaction between yeast Ddi1p and ubiquitin.

4.2.2. The retroviral protease-like domain of yeast Ddi1p does not cleave peptide substrates *in vitro*

We determined the 3D structure of the retroviral protease-like domain of yeast Ddi1p. Unlike earlier structures, our structure contained electron densities for the flexible part of RVP, so-called flaps, covering the upper part of the active site. Thanks to specific crystal packing one of the flaps

create contacts with the N-terminus of a neighboring molecule and stabilize the N-terminus in the active site. This interaction suggests a potential substrate-binding mode for the protease domain of yeast Ddi1p. Structural studies were performed by Dr. Jean Francois Trempe.

To further explore the prospective substrate repertoire of the protease domain, I performed the unbiased proteomic assay for the identification of peptidic substrates of the RVP domain based on the PICS method. Peptide libraries derived from the expressed proteome of yeast *Saccharomyces cerevisiae* culture in an exponential growth phase were created using trypsin or GluC as preprocessing proteases. Libraries were cleaved overnight by *E. Coli*-expressed full-length yeast Ddi1p, using D220A putative catalytic site mutant as a negative control and HIV-1 protease as a positive control. As can be seen in figure 11, I did not observe any peptide cleavage pattern related to yeast Ddi1p activity. This experiment was repeated with the RVP domain itself. Both yeast Ddi1p variants were assayed on two different variants of the peptide library (GluC and trypsin prepared) at pH 4.0, 5.0, and 7.0. None of these experiments led to protease activity detection.

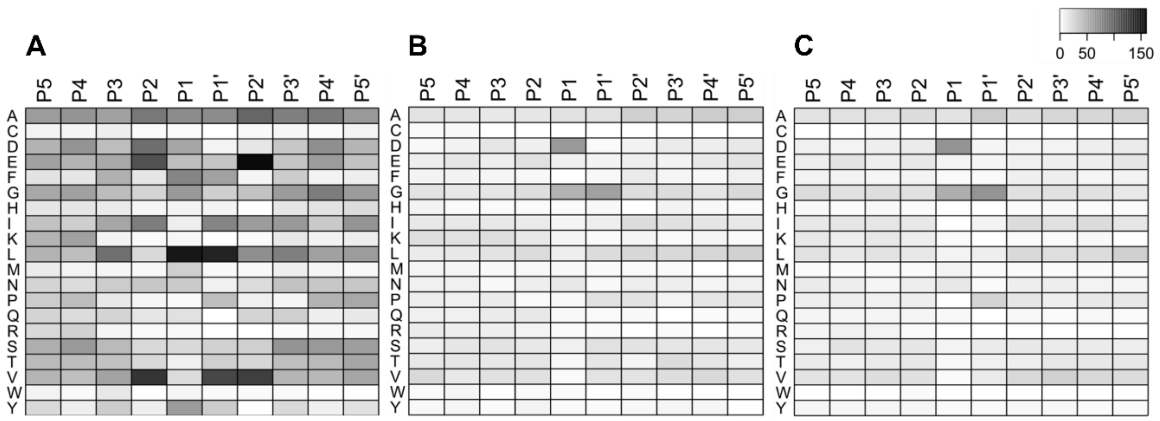


Figure 11: PICS experiment results. Heat map representation of protease substrate profile demonstrating total counts of a given amino acid (in single letter code) per particular position (P5−P5´) in a peptidic substrate on *Saccharomyces cerevisiae* derived peptide library. (A) library cleaved by HIV-1 protease (positive control) (B) Library cleaved by yeast Ddi1p protein in acetate buffer pH = 4.0 (proteolytic activity assay) (C) background - **uncleaved** library (negative control). The color key is common for all three heat maps.

4.2.3. Ubiquitin binding

To explore yeast Ddi1p interaction with ubiquitin we performed an isothermal titration calorimetry (ITC) experiment, where we titrated full-length yeast Ddi1p by ubiquitin. To our surprise, we obtained a stoichiometry ratio of 2:1 for the ubiquitin:full-length yeast Ddi1p interaction (figure 12A), although yeast Ddi1p contains only one UBA domain per molecule. Indeed, during the preparation of this manuscript, an interaction of ubiquitin with the UBL domain of yeast Ddi1p was reported by Nowicka et.al. [Nowicka et al., 2015]. To further elucidate the mechanism of ubiquitin - yeast Ddi1p interaction we prepared truncated variants of Ddi1p lacking

either UBL or UBA domain. As can be seen in figure 12B and C, both UBL and UBA show a stoichiometry of 1:1, with UBA bearing the main part of the total interaction (Kd=43 μM) and UBL providing a weaker contribution with a Kd of 310 μM. K48 covalent linkage in diubiquitin conjugate potentiates the interaction with full-length Ddi1p with a Kd of 77 μM and a 1:1 stoichiometry (figure 12D), suggesting that each Ddi1p dimer binds two K48-diubiquitin molecules

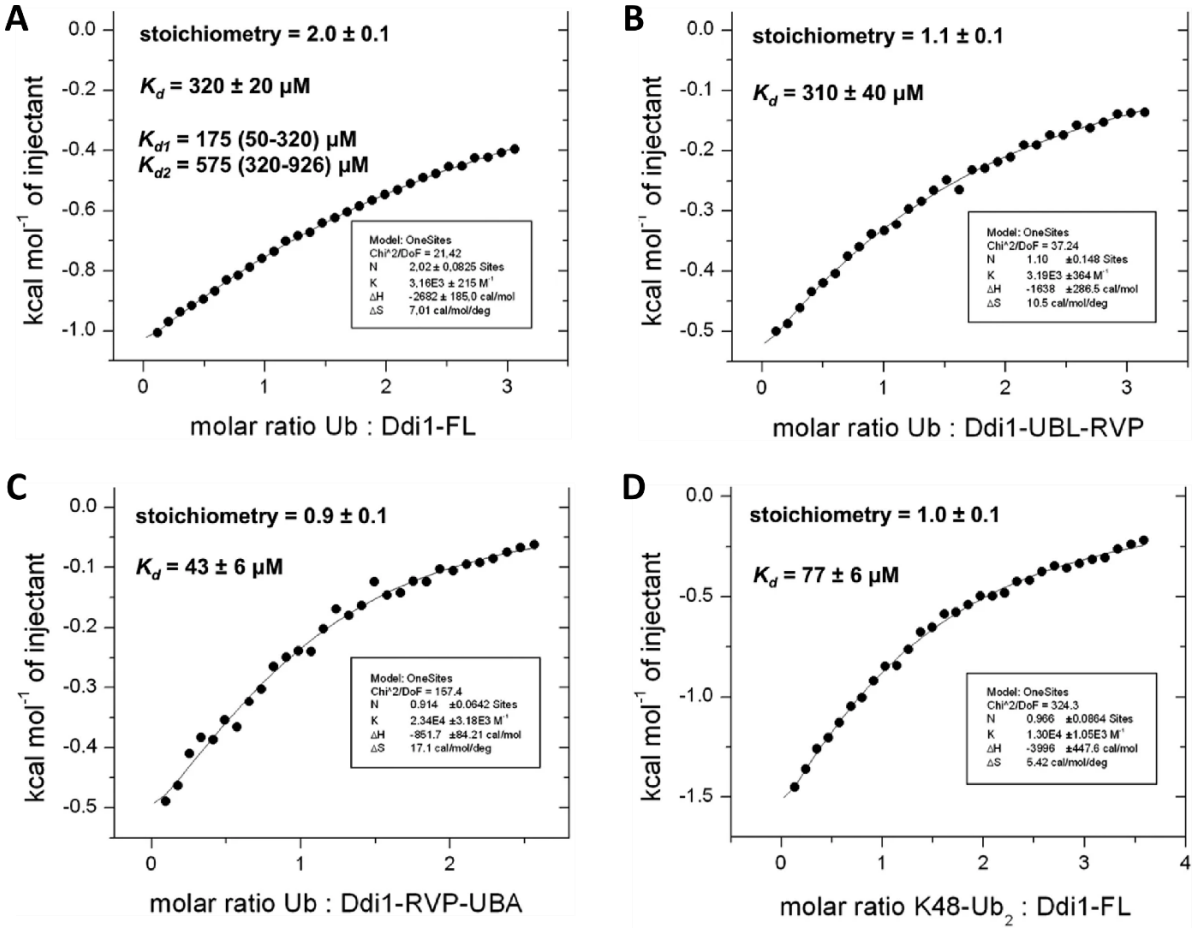


Figure 12: Yeast Ddi1p interacts with ubiquitin. Ubiquitin or K48-linked diubiquitin (Ub_2) were added by syringe to particular Ddi1p constructs in the sample cell in an ITC titration experiment. The different titrations were (A) ubiquitin to full-length yeast Ddi1p, (B) ubiquitin to UBL-RVP, (C) ubiquitin to RVP-UBA, and (D) K48-Ub_2 to full-length yeast Ddi1p. For full-length Ddi1p binding to ubiquitin (A), the dataset was also fitted to a model with two independent binding sites (K_{d1} and K_{d2}), each with a stoichiometry of 1.

4.2.4. Conclusion

In this publication, we investigated the multi-domain structure of yeast Ddi1p using X-ray crystallography, NMR, and small-angle X-ray scattering. The crystal structure of the RVP domain provides insight into a putative substrate recognition mechanism involving a conserved loop. Unlike related retroviral proteases, yeast Ddi1p RVP domain is not capable of cleavage of peptide

substrates *in vitro*. Isothermal titration calorimetry confirmed that both UBL and UBA domains bind ubiquitin, and that yeast Ddi1p binds K48-linked diubiquitin with enhanced affinity. We solved solution NMR structures of the UBL domain and a helical domain that precedes the protease. Our structural studies suggest that the helical domain could serve as a landing platform for protease substrates in cooperation with attached ubiquitin chains binding to the UBL and UBA domains.

4.3. PUBLICATION III:

The yeast proteases Ddi1 and Wss1 are both involved in the DNA replication stress response

Svoboda M., Konvalinka J., Trempe J.F. and Grantz Šašková K. The yeast proteases Ddi1 and Wss1 are both involved in the DNA replication stress response. *DNA Repair*. **80**, 45–51 (2019).

My contribution: experimental design; cloning of all used DNA constructs; yeast phenotypic assays; yeast transformation; complementation assays; protein expression analysis; data analysis; wrote the manuscript.

4.3.1. Motivation of the study

Although *DDI1* was initially identified in a high throughput screen looking for genes with the expression level responsive to treatment with chemicals causing DNA damage (*DDI1* stands for DNA Damage-Inducible 1), little was known about the molecular mechanism, by which yeast *DDI1* contributes to countering DNA damage. In addition to known options available for the cell to counter DNA damage, a novel pathway was recently identified involving direct proteolysis of protein-based lesions blocking the progression of the replication fork. In this pathway metalloprotease of Wss1p/SPRTN family is recruited to the blocked fork site and cleaves off the protein part of the lesion. Surprisingly, large screens for synthetic lethality of gene-pair deletions identified yeast *DDI1* as a gene with strong synthetic lethality/sickness once deleted together with yeast *WSS1*. With this synthetic sick pair of proteases, a hypothesis of yeast Ddi1p involvement in the clearance of protein-based replication barriers arose. Therefore, we decided to find out whether yeast Ddi1p acts in countering the so-called replication stress, whether it is the protease domain of Ddi1p that facilitates this function, which of the other Ddi1p domains are involved, and whether this function could be functionally complemented by human homologues of *DDI1*. To answer the question, we employed methods of classical yeast genetics with phenotypic assays and functional complementation assays.

4.3.2. Simultaneous deletion of *DDI1* and *WSS1* renders yeast cells hypersensitive to hydroxyurea

To test whether Ddi1p participates in replication stress countering or other DNA damage response mechanisms, I performed a phenotypic assay to assess the effects of various agents causing DNA damage on the growth of Δ*ddi1,* Δ*wss1* double-deleted yeast strain. As can be seen in figure 13, the weak effect of methyl methanesulfonate and formaldehyde on Δ*ddi1,* Δ*wss1* strain was observed out of the panel of all tested compounds. On the other hand, ribonucleotide

reductase inhibitor hydroxyurea had a strong inhibitory effect over a broad range of concentrations with a specific response in Δ*ddi1, Δwss1* strain, compared to wild type and single deletion strains.

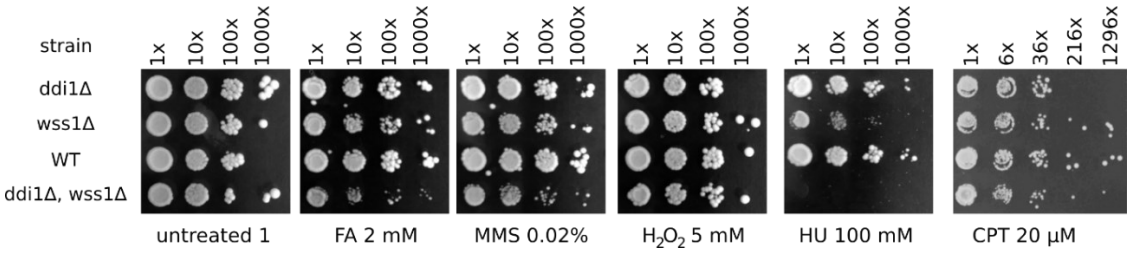


Figure 13: The Δ*ddi1, Δwss1* strain is hypersensitive to hydroxyurea. Dilution spot assays testing the viability of wild-type, single knock-out, and double knock-out strains under the influence of different DNA damage-causing chemicals. Equal amounts of exponentially growing cells were plated in serial dilutions on plates containing either YPDA media only or YPDA supplemented with formaldehyde (FA), methyl methanesulfonate (MMS), hydrogen peroxide (H_2O_2), hydroxyurea (HU), or camptothecin (CPT).

4.3.3. Hydroxyurea hypersensitivity can be rescued by *DDI1* overexpression in an active site-dependent manner

To test whether the hydroxyurea hypersensitivity phenotype is caused by the absence of Ddi1p protein rather than disruption of gene regulatory elements, I performed a rescue experiment where the Δ*ddi1, Δwss1* strain was transformed by plasmids driving the overexpression of either wild type or point mutated variants of yeast *DDI1*. Figure 14 shows that hydroxyurea hypersensitivity can be rescued by overexpression of the wild type yeast Ddi1p protein. The rescue was independent of the presence of known phosphorylation and ubiquitination sites as well as cysteine 300, described previously as residue required for proper nuclear localization. On the other hand, a single amino acid mutation of the catalytic active site aspartate 220 to either asparagine or alanine completely abolished the ability of yeast Ddi1p to counter the hydroxyurea hypersensitivity, suggesting a crucial role of the yeast Ddi1p proteolytic activity in the mechanism.

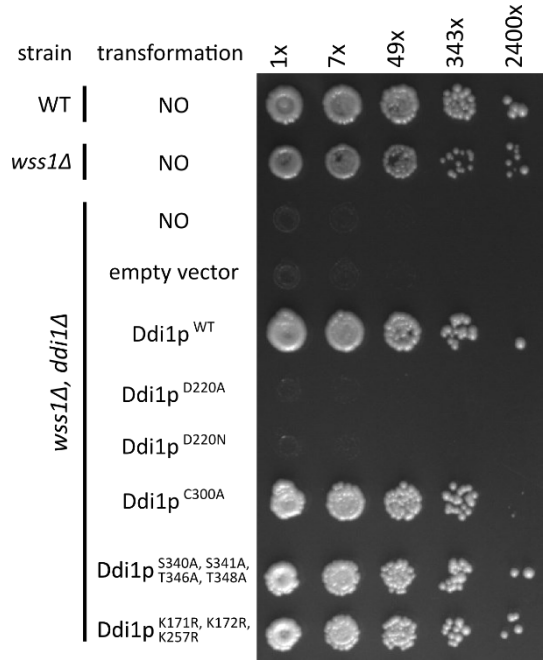


Figure 14: Retroviral protease-like (RVP) domain catalytic activity is indispensable for complementation of hydroxyurea sensitivity. Equal amounts of exponentially growing cells transformed with single-copy plasmids overexpressing point-mutated variants of Ddi1p protein under the control of the **Glyceraldehyde-3-phosphate dehydrogenase** (GPD) promoter were plated in 7-fold serial dilutions on one YPDA plate supplemented with 50 mM hydroxyurea**.**

4.3.4. HDD domain is indispensable for the Ddi1p function in replication stress response

Knowing that the hydroxyurea hypersensitivity of the Δ*ddi1,* Δ*wss1* strain can be rescued by yeast *DDI1* overexpression in an active site-dependent manner, I analyzed the requirement for particular domains of yeast Ddi1p for the rescue phenotype. I prepared truncated constructs of yeast *DDI1* lacking each known individual domain of the protein and performed a rescue experiment. Figure 15 shows that, as expected, the deletion of the RVP (aa 202-315) domain produces a phenotype identical to the mutation of catalytic aspartate 220. Surprisingly, the same phenotype was observed for the deletion of the HDD domain, more specifically its second helical bundle (aa 148-191). This observation can be partially explained by the vicinity of the HDD domain to the RVP domain and its hypothetical role as a proteolysis substrate landing platform.

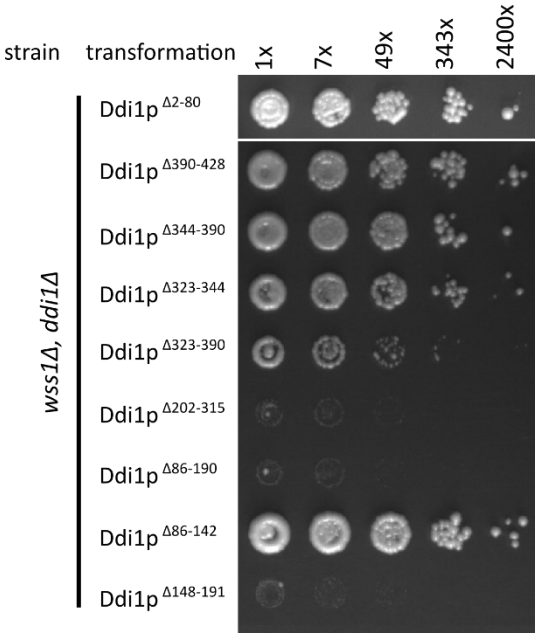


Figure 15: Retroviral protease-like (RVP) domain and a helical bundle (HDD) preceding the RVP domain are indispensable for complementation of hydroxyurea sensitivity. Equal amounts of exponentially growing cells transformed with single-copy plasmids overexpressing variants with deletions of known structured domains of Ddi1p under the control of the GPD promoter were plated in 7-fold serial dilutions on one YPDA plate supplemented with 50 mM hydroxyurea. All spots were done on the same plate.

4.3.5. The *DDI1* function in replication stress response is neither UBL nor UBA dependent

To further explore the mechanism behind yeast Ddi1p action in countering hydroxyurea-caused replication stress I identified the minimal part of yeast Ddi1p necessary and sufficient for the rescue of the hydroxyurea hypersensitivity phenotype. I prepared several truncated variants, with both N- and C-terminal truncation, and analyzed those constructs in the complementation assay. As can be seen in figure 16, both UBL and UBA domains are dispensable for yeast Ddi1p function in hydroxyurea tolerance, as the construct lacking both (aa 81-389) is capable to rescue the phenotype. The minimal construct, sufficient to fill the role of Ddi1p in the DNA replication stress response, spanned amino acids 146-322 and therefore covered the second helical bundle of HDD domain and the RVP domain.

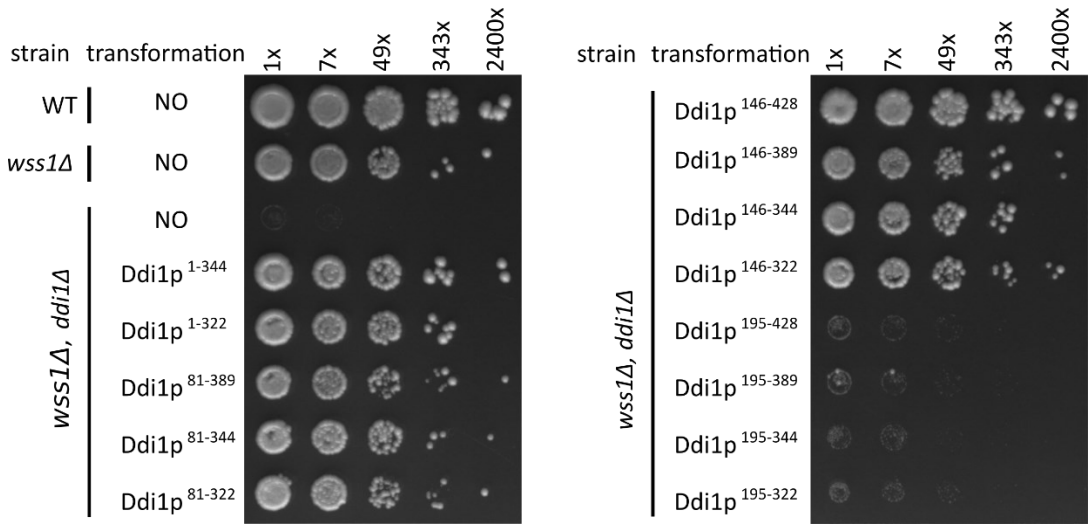


Figure 16: Complementation of hydroxyurea sensitivity by *DDI1* overexpression does not require the ubiquitin-like and ubiquitin-associated domains of Ddi1p. The minimal construct capable of complementation spans residue 146 to residue 322, covering the second multi-helical bundle of HDD domain and the RVP domain. Dilution spot assays with 7-fold serial dilutions of cells overexpressing truncated variants of Ddi1p protein plated on YPDA supplemented with 50 mM hydroxyurea are shown. All 16 spot tests were done on one plate.

Due to the loss of anti-Ddi1p antibody detection of some of the truncated variants, we performed an identical dilution spot experiment using C-terminally HA-tagged Ddi1p protein variants and obtained the same results.

4.3.6. Human Ddi1-like proteins are capable of partial complementation of *DDI1* function

To assess whether the replication stress countering ability applies also for human homologues of Ddi1-like proteins, I performed complementation experiments with overexpression of both human *DDI1* and *DDI2*. Surprisingly, both *DDI1* and *DDI2* are capable of partial functional complementation (figure 17) with the weaker effect compared to yeast *DDI1*. Nevertheless, rescue by human homologues is also dependent on the presence of the putative catalytic aspartate, and the difference between human and yeast homologues pronounces with decreasing temperature, pointing towards enzymatic reaction with different temperature optima as an underlying mechanism.

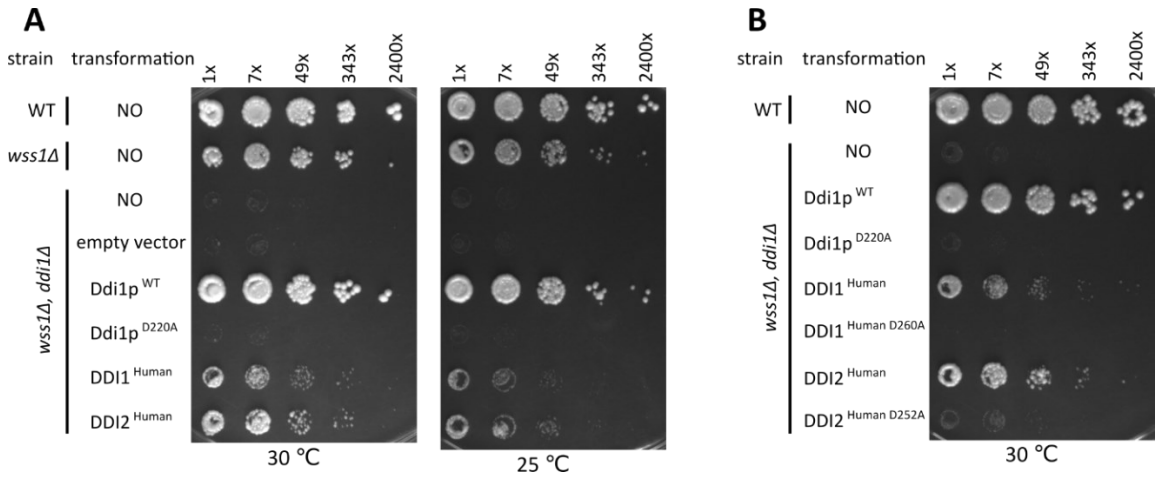


Figure 17: Human *DDI1* and *DDI2* are capable of partial, proteolysis dependent, complementation of hydroxyurea sensitivity. Dilution spot assays with 7-fold serial dilutions of cells overexpressing A) yeast and human Ddi1-like proteins, cultured under two different temperatures, and B) their respective catalytically inactivated mutant variants. All spots were plated on YPDA plates supplemented with 50 mM hydroxyurea.

4.3.7. Conclusion

In this publication, we identified a hydroxyurea hypersensitivity in double-mutated Δ*ddi1,* Δ*wss1* yeast strain. Hydroxyurea acts as a ribonucleotide reductase inhibitor and causes stalling of the replication fork and subsequent replication stress. Yeast Ddi1p protein overexpression counters hydroxyurea hypersensitivity phenotype and rescues to normal growth in Δ*ddi1,* Δ*wss1* strain. Ddi1p function in this context is protease activity-dependent and requires the presence of the second helical bundle of the HDD domain. Surprisingly, neither UBL nor UBA domains are required for the rescue, therefore excluding the so-called proteasome shuttling mechanism and pointing towards mechanism based on direct proteolysis. Taken together, our data suggest the existence of a dual protease mechanism providing yeast cells with the ability to overcome DNA-replication stress caused by hydroxyurea, the ribonucleotide reductase inhibitor and known anti cancer drug. Furthermore, we identified a first function for the recently described HDD domain.

DISCUSSION

5. DISCUSSION

Ddi-1 like proteins forms a unique family of retroviral protease domain-containing proteins conserved through all sequenced eukaryotic organisms [Krylov and Koonin, 2001]. Despite the high level of conservation, only the yeast orthologue, *DDI1*, was studied more extensively. It was identified as a gene product with expression stimulated by treatment with DNA damaging chemicals, potential cell cycle regulator, and a participant in the ubiquitin-proteasome system [Liu and Xiao, 1997, Zhu and Xiao, 1998, Kaplun et al., 2005, Clarke et al., 2001]. In parallel, another group identified yeast *DDI1* as a regulator of constitutive exocytosis, acting via phosphorylation regulated network of interactions with SNARE proteins [Lustgarten and Gerst, 1999]. Ddi1-like proteins from other eukaryotic organisms remained understudied, except for isolated reports on fruit fly rngo [Morawe et al., 2011], *C. elegans ddi-1* [Guthmueller et al., 2011], and, recently, a set of papers describing the role of human *DDI2* in the regulation of expression of proteasome subunits [Koizumi et al., 2016, Chowdhury et al., 2017]. Surprisingly, despite the identification of the first member of the Ddi1-like family, yeast *DDI1* as a gene induced by DNA damage more than twenty years ago [Liu and Xiao, 1997], the molecular mechanism behind the inducibility, or the particular function of Ddi-1 like proteins in response to DNA damage remained elusive. We, therefore, decided to shed some light on Ddi1-like proteins function in this process.

First, we focused on the understudied human DDI2 protein. Combining biophysical and biochemical methods, we reported the first structural and functional analysis of this protein. Ddi1-like protein family is an integral part of the UPS. Interestingly, human DDI2 did lose its UBA domain to the downstream *RSC1A1* gene in a gene insertion event. We, therefore, explored other possible interaction interfaces enabling a functional connection to the UPS via interaction with ubiquitin. We identified a UIM motif on the very C-terminal alpha-helix of human DDI2, analyzed its interaction with ubiquitin, and found the interaction is weak, yet specific. During the preparation of this manuscript, Nowicka and colleagues published a paper describing a new paradigm in ubiquitin binding structural elements, showing that the UBL domain of yeast Ddi1p protein is binding to ubiquitin via complementary surface electrostatic charges [Nowicka et al., 2015]. We solved the NMR structure of the human DDI2 UBL and showed that the surface charge distribution for this domain differs significantly from its yeast homologue, with the overall charge on the potential interaction interface more neutral. Different electrostatic potential distribution, compared to the yeast homologue, explains significantly weaker, but still specific, interaction with ubiquitin, we observed in NMR titrations. In contrast, we did not observe any interaction with polyubiquitin chains, tested in the pull-down experiment with diubiquitin conjugates of all eight native linkages. Reports of others published later on show the human DDI2 role in the processing

of TCF11/NRF1, transcription activator responsible for driving proteasome subunit transcription upon UPS inhibition [Koizumi et al., 2016]. Here, in contrast with our results, human DDI2 activity appears to be ubiquitination dependent. The disagreement between our results and results from Murata lab could be explained by the minimum length requirement for the polyubiquitin chain. This *in vivo* observation can be explained by an orchestrated action of four weak-affinity ubiquitin-binding modules present in the DDI2 homodimer that might bring sufficient avidity to mediate a productive polyubiquitin interaction as observed by Koizumi et. al. In our experiments, we used only diubiquitin chain *in vitro*. Indeed, a recent publication by Dirac-Svjestrup and colleagues elucidated this situation, showing that cell lines with *DDI2* gene knocked-out accumulate high molecular, highly polyubiquitinylated protein conjugates and that human DDI2 can directly process TCF11/NRF1, but requires very long polyubiquitin chain decoration, probably as a substrate recognition motif [Dirac-Svejstrup et al., 2020]. Human DDI2 does, therefore, bind polyubiquitin chain, but preferentially long ones. The exact molecular mechanism behind human DDI2 recognition and preference for long polyubiquitins remains to be clarified.

Besides the UBL-UBA (or, in mammalian Ddi1-like proteins, the UBL-UIM) architecture, another characteristic feature of Ddi1-like proteins is the presence of retroviral aspartic protease domain in the central portion of their sequence [Krylov and Koonin, 2001]. We solved the three-dimensional structure of human DDI2 RVP using X-ray crystallography. Our structure revealed a canonical fold of retropepsin family of aspartic proteases very similar to the previously published structure of yeast Ddi1p RVP [Sirkis et al., 2006], structure of yeast Ddi1p RVP from our back to back publication, unpublished structural genomics consortia structure of human DDI1 RVP (PDB code 3S8I) or, more generally, HIV-1 protease and other homodimeric aspartic proteases from retroviruses. On the structural level, human DDI2 RVP harbours all important determinants of aspartic protease activity. The bottom of the structure consists of the multiple beta-sheet-formed dimerization interface and the active site cavity is covered from the top by a pair of flexible beta-hairpins, so-called flaps. At the bottom of the active site cavity resides the catalytic Asp-Ser-Gly triad in a typical "fireman's grip" coordination [Strisovsky et al., 2000], binding a putative catalytic water molecule that serves as a nucleophile in the prospective proteolytic reaction. The only striking difference between human DDI2 RVP (and RVPs of Ddi1-like proteins in general) and aspartic proteases from retroviruses is a more spacious active site cavity, capable of accommodation of bulkier substrates and thus pointing toward possible cleavage of partially unfolded proteins instead of peptides. With all determinants of proteolytic activity present, it took us by surprise that we were unable to detect any *in vitro* proteolytic activity for neither isolated RVP nor the full-length human DDI2 protein under none of the various conditions we tested. We

employed an expressed proteome-based peptide library approach (PICS), as well as nonspecific cleavage of proteins and a large panel of known HIV-1 protease and other aspartic protease peptide substrates, all assayed under various pH and salt concentrations. None of these experiments led to a measurable proteolytic activity that we could ascribe to RVP. Indeed, although there was plenty of data suggesting the functional importance of catalytic activity for cell physiology in different species [Diaz-Martinez et al., 2006, Gabriely et al., 2008, Morawe et al., 2011, White et al., 2011a], no one, except for a solitary publication on Ddi1-like protein from protozoa *Leishmania major* under nonphysiological, very acidic conditions [Perteguer et al., 2013], ever reported a biochemical assay for *in vitro* enzymatic activity of Ddi1-like protein. Even in an article focused on the human DDI2 role in the processing of TCF11/NRF1, proteasome subunits transcription activator, with solid evidence for the requirement for RVP catalytic activity, authors explicitly stated that despite multiple attempts they were not able to biochemically reconstitute the RVP proteolytic activity *in vitro* [Koizumi et al., 2016]. The above-mentioned recent publication by Dirac-Svejstrupup and colleagues shed light on this problem as well, proving that human DDI2 does have a proteolytic activity that can be assayed even *in vitro*, but the substrates must be polyubiquitinylated, with a strong preference for high molecular weight polyubiquitin conjugates. In agreement with our results, Dirac-Svejstrupup did not observe any deubiquitinylating activity of human DDI2 [Dirac-Svejstrup et al., 2020]. The detailed molecular mechanism of RVP proteolysis, particularly the process of long polyubiquitin recognition and passing of the information on the proteolytic part of the protein with the following activation of proteolysis though remains elusive.

The second part of this dissertation is focused on the Ddi1p protein of budding yeast *Saccharomyces cerevisiae*. We started with a detailed structural and biochemical characterization of the protein, solving three-dimensional structures of the UBL, HDD, and RVP domains and investigating interactions between the UBL domain and its suspected binding partners. Surprisingly, while analyzing the interaction between yeast Ddi1p and ubiquitin, we detected an unexpected stoichiometry of 2:1 for the interaction, although yeast Ddi1p harbours only one UBA domain. During our work on this manuscript, a publication by Nowicka and colleagues revealed a novel, unorthodox ubiquitin binding mode for yeast Ddi1p [Nowicka et al., 2015]. In their experiments, the UBL domain interacted with ubiquitin via a complementary charged surface, explaining the higher stoichiometry we observed. Based on the publication, we analyzed the UBL - ubiquitin interaction in an ITC experiment. However, our results differ from data from Nowicka et. al. Although we observed a ubiquitin interaction for both UBL and UBA, in our measurement UBA was the tighter binder of ubiquitin. This discrepancy could be explained by the different experimental setup, as we used a different method (ITC versus NMR), different buffer

(HEPES pH 7.4 versus sodium phosphate pH 6.8), and the UBL construct containing the initiator Met1, which alters the position of Asp2 involved in the interaction with ubiquitin. Moreover, compared to Nowicka's NMR titrations on isolated domains, our ITC experiments were carried out in the context of full-length, dimeric proteins. Despite minor discrepancies, our data are consistent with the yeast Ddi1p protein harbouring two independent ubiquitin interacting moieties.

We reported a structure of yeast Ddi1p RVP as well. In general, our structure is similar to the earlier published structure of yeast Ddi1p RVP by Sirkis [Sirkis et al., 2006] and RVP structures from other Ddi1-like proteins. Our structure indicates interaction of the yeast Ddi1p RVP active site with the N-terminus of the adjacent molecule as observed in the crystal lattice. The interaction is mediated by extended beta conformation formed with a conserved loop of one of the flaps, adjacent to the active site. The extensive network of hydrogen bonds keeps the N-terminus in the active site, in the position expected to be occupied by a substrate of the RVP domain. Although the observed interaction probably originates from a crystallization artifact and there is little to none physiological importance for it, mechanistically, such observation points toward the molecular details of proteolysis by RVP and likely represents a model of substrate engagement. The pseudo-substrate N-terminal segment here adopts a conformation similar to HIV-1 protease peptide substrates.

Encouraged by a promising result from the structural study, we explored a putative yeast Ddi1p RVP proteolytic activity in biochemical *in vitro* assay. Here, we employed a PICS method used also in our back to back publication with human DDI2. We used full-length yeast Ddi1p protein, as well as isolated RVP domain, with exponentially growing yeast culture-derived peptide libraries under several assay conditions. Like its human orthologue, yeast Ddi1p did not exhibit any protease activity in our proteomic screen. While this could be interpreted as 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 during its physiological function. The newly identified HDD domain could serve as a substrate landing platform in this context. Our hypotheses on yeast Ddi1p activation were later confirmed by Yip and colleagues in a recent publication, where they described activation of yeast Ddi1p by substrates bearing long polyubiquitin chains [Yip et al., 2020]. The mechanism behind polyubiquitin-caused Ddi1-like protein activation in proteolysis thus appears conserved throughout species.

In a follow-up publication, we focused on the identification of yeast *DDI1* function in response to treatment with DNA damaging agents. Despite its identification as a gene product with expression induced by treatment with methyl methanesulfonate, a genotoxin [Liu and Xiao, 1997], the function of yeast *DDI1* in DNA damage response remained unknown. We decided to take

advantage of the simple setup for gene modification in yeast and analysis of the phenotype of individual mutants or their combinations. We used a yeast Synthetic Genetic Array as a starting point. In this method, exploring a synthetic lethality between double mutated strains with an array of deleted gene pairs, *DDI1* exhibited a strong negative genetic interaction with *WSS1*, replication stress countering protease [Costanzo et al., 2016]. By testing several typical DNA damage-inducing chemicals on Δ*ddi1,* Δ*wss1* yeast strain as well as on the wild type and single mutated strains, we found the double deleted strain to be highly sensitive to hydroxyurea. Hydroxyurea acts as an inhibitor of ribonucleotide reductase, an enzyme responsible for converting ribonucleotide triphosphates to deoxyribonucleotide triphosphates and thus providing building blocks for DNA replication [Elford, 1968]. Inhibition of ribonucleotide reductase by hydroxyurea results in replication stress, replication fork stalling, and, if not resolved, even to replication fork collapse [Petermann et al., 2010]. We observed a mild hydroxyurea sensitivity in *WSS1* deleted single mutant strain as described previously [O'Neill et al., 2004]. The synthetic effect of both mutations substantially potentiated hydroxyurea sensitivity, beyond simple additivity. This suggests that Ddi1p and Wss1p act in two independent pathways, each counteracting the replication stress caused by hydroxyurea. Such observation is in agreement with results from Kottemann and colleagues, who identified human DDI1 and DDI2 as proteins involved in the degradation of the replication termination factor RTF2, facilitating restart of stalled replication forks after replication inhibition by, for example, hydroxyurea [Kottemann et al., 2018]. Interestingly, Wss1p was identified earlier as a protease clearing the protein part of DNA-protein crosslinks [Stingele et al., 2014]. Hydroxyurea, though, is not known to cause DPC's and our observation thus raises a couple of questions. It remains to be clarified which DNA damage response pathway, triggered by hydroxyurea, leads to growth arrest when both *DDI1* and *WSS1* are lost. The double deletion strain may experience the upregulation of checkpoint pathways in the S phase, which induces cell cycle arrest upon DNA damage. Other important questions we did not address by our results, such as which substrate(s) are cleaved by the Ddi1p protease domain in this context and whether the target molecule is the same for Wss1p, remain to be answered.

Next, we focused on a more detailed understanding of yeast Ddi1p function in the process. By overexpressing variants of the protein in the double knock-out background, followed by spot test phenotypic assay, we found out that although the hydroxyurea hypersensitivity phenotype is fully rescued by yeast Ddi1p reintroduction by overexpression, a single point mutation in putative catalytic aspartate renders the protein completely ineffective, leading to the phenotype of background double knockout strain. This represents yet another function of a protein from the Ddi1-like family, dependent on protease activity, and, together with other identified proteolysis

dependent functions, highlights the importance of the proteolytic activity of Ddi1-like proteins in cell physiology. The following experiment with complementation by truncated variants of the protein, lacking the whole individual folded domains corroborated our results. The protein variant missing the RVP presented the same phenotype as did the catalytic aspartate point mutant. Surprisingly, another construct we used, a protein lacking the HDD domain, or more specifically its second helical bundle, was also incapable of functional complementation. In a reverse experiment, aimed at the identification of a minimal yeast Ddi1p construct capable of functional complementation, we found that a protein construct spanning only the second helical bundle of the HDD domain and the RVP domain (aa 146-322) is sufficient to rescue the phenotype. This came as a surprise, as many publications focused on yeast Ddi1p pinpoint its role as the putative proteasomal shuttling protein [Voloshin et al., 2012, Diaz-Martinez et al., 2006]. The proteasomal shuttling mechanism was also proposed as a basis of the mammalian orthologues, human DDI1, and human DDI2 role in the degradation of RTF2 and subsequent restart of replication forks [Kottemann et al., 2018]. Our data suggest that in this case, we can exclude not only the proteasomal shuttle but as well the so-called alternative shuttle mechanism, proposed by Nowicka [Nowicka et al., 2015], as we observed full rescue even with constructs lacking any ubiquitin-binding moieties whatsoever. We, therefore, proposed a dual protease mechanism for hydroxyurea-caused DNA replication stress tolerance.

Interestingly, almost all of our results were independently corroborated in a manuscript published a couple of months after our publication [Serbyn et al., 2020]. Serbyn and colleagues addressed a set of questions very similar to the one we asked. They show that yeast *DDI1* is crucial for one of three partially redundant pathways for clearing proteins covalently trapped on DNA strand the other two being centered around *WSS1* protease and 26S proteasome. In agreement with our data, they observed hypersensitivity of Δ*ddi1,* Δ*wss1* yeast strain (although in the different genetic background) to hydroxyurea, as well as the requirement of RVP catalytic activity and the presence of HDD domain for successful Ddi1p complementation. However, Serbyn et. al. went a step further showing Ddi1p localization in the nuclear fraction, its direct association with chromatin, and, most interestingly identified the core Pol II component Rpb1p (RNA polymerase II subunit B1) as a probable substrate involved in yeast Ddi1p proteolysis mediated replication stress response. Interesting information could be found also in the supplementary data accompanying Serbyn's publication. Here they showed that although UBL is not required for rescue under overexpression of yeast Ddi1p, where the cell is flooded with the overexpressed protein, once expressed from a native promotor, UBL suddenly becomes essential for the rescue. More details, filling some blanks left in ours as well as Serbyn's papers come from a recent

publication by Yip and colleagues. Here they found that, like its human orthologue, yeast Ddi1p is an active protease that requires very long polyubiquitin chain decorated substrates. They established an *in vitro* biochemical assay and analyzed various truncated constructs of yeast Ddi1p. In agreement with our data, they showed that construct lacking HDD domain is incapable of proteolysis, construct spanning HDD and RVP domains possesses rather weak proteolytic activity as does a construct lacking the UBL domain. Full-length and UBA lacking constructs are on the other hand fully active [Yip et al., 2020]. This provides a possible explanation for the discrepancy between our data and results published by Serbyn. In *DDI1* overexpressing cells, a high concentration of overexpressed protein substitutes for the low proteolytic activity of truncated variants. In cells expressing from the native promotor, propper activation/localization is paramount due to the normal physiological level of protein expression.

CONCLUSION

6. CONCLUSION

The focus of this doctoral thesis was to investigate the function of the Ddi1-like family of proteins with an emphasis on their role in DNA damage related processes and response to the replication stress. To fulfill those goals, we employed a combination of biochemical and biophysical methods as well as tools of the classical yeast genetics.

First, we biophysically and structurally characterized human DDI2 protein. We analyzed the structure of the UBL domain, assessed its surface charge distribution, and analyzed the interaction of the UBL domain with ubiquitin and ubiquitin chains. Furthermore, we performed proteolytic activity assays on the RVP domain of human DDI2 but were not able to detect any proteolytic activity we could attribute to the protein.

In the second part, we focused on *Saccharomyces cerevisiae* orthologue, yeast Ddi1p protein. First, we performed a detailed analysis of its interaction with ubiquitin and K48-linked diubiquitin, revealing higher stoichiometry for the interaction and corroborating the earlier reported unorthodox ubiquitin binding mode, mediated by the UBL domain. Despite promising results from yeast Ddi1p RVP structural studies, we were unable to detect any proteolytic activity for this domain in the *in vitro* biochemical assay.

In the third part, we focused on the characterization of Ddi1-like proteins function in the cellular response to DNA replication stress. We decided to utilize the yeast model since it provides facile methodology for the genetic screens of mutant combinations that would be much more labourious and time consuming in a mammalian model. We observed hypersensitivity to the DNA replication blocking drug hydroxyurea in the yeast strain with double deletion of *DDI1* combined with a DNA dependent metalloprotease *WSS1*. Further analysis showed that the hydroxyurea hypersensitivity phenotype can be rescued by overexpression of yeast Ddi1p as well as its mammalian orthologues, human DDI1, and human DDI2. Ddi-1 like proteins function in this context is dependent on the functional RVP domain and presence of newly identified HDD domain, directly preceding the RVP domain in the protein sequence.

Taken together, we structurally characterized both human and yeast orthologues from the Ddi1-like protein family analyzed their interaction with ubiquitin and discovered yeast Ddi1p function in a dual protease mechanism of preventing DNA replication stress. Unexpectedly, by the time we started to study this class of proteins, it appeared that at least in higher organisms, the abbreviation Ddi1 represents purely a denomination and that it does not reflect the biological role of the protein family. We spent several years looking for the evasive biological function of proteins from the „DNA damage“ family only to find out that they indeed do have a role in DNA damage.

REFERENCES

7. REFERENCES

[Arezi and Kuchta, 2000] Arezi, B. and Kuchta, R. D. (2000). Eukaryotic DNA primase. *Trends Biochem Sci*, 25:572–576.

[Baker et al., 2001] Baker, N. A., Sept, D., Joseph, S., Holst, M. J., and McCammon, J. A. (2001). Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc Natl Acad Sci U S A*, 98:10037–10041.

[Balakirev et al., 2015] 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.

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

[Bassermann et al., 2008]Bassermann, F., Frescas, D., Guardavaccaro, D., Busino, L., Peschiaroli, A., and Pagano, M. (2008). The Cdc14B-Cdh1-Plk1 axis controls the G2 DNA-damage-response checkpoint. *Cell*, 134:256–267.

[Bermejo et al., 2012] Bermejo, R., Lai, M. S., and Foiani, M. (2012). Preventing replication stress to maintain genome stability: resolving conflicts between replication and transcription. *Mol Cell*, 45:710–718.

[Bertolaet et al., 2001a] Bertolaet, B. L., Clarke, D. J., Wolff, M., Watson, M. H., Henze, M., Divita, G., and Reed, S. I. (2001a). UBA domains mediate protein-protein interactions between two DNA damage-inducible proteins. *J Mol Biol*, 313:955–963.

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

[Bester et al., 2011] Bester, A. C., Roniger, M., Oren, Y. S., Im, M. M., Sarni, D., Chaoat, M., Bensimon, A., Zamir, G., Shewach, D. S., and Kerem, B. (2011). Nucleotide deficiency promotes genomic instability in early stages of cancer development. *Cell*, 145:435–446.

[Bialkowska and Kurlandzka, 2002] Bialkowska, A. and Kurlandzka, A. (2002). Proteins interacting with Lin1p, a putative link between chromosome segregation, mRNA splicing and DNA replication in Saccharomyces cerevisiae. *Yeast*, 19:1323–1333.

[Biggins et al., 2001] Biggins, S., Bhalla, N., Chang, A., Smith, D. L., and Murray, A. W. (2001). Genes involved in sister chromatid separation and segregation in the budding yeast Saccharomyces cerevisiae. *Genetics*, 159:453–470.

[Birrell et al., 2001] Birrell, G. W., Giaever, G., Chu, A. M., Davis, R. W., and Brown, J. M. (2001). A genome-wide screen in Saccharomyces cerevisiae for genes affecting UV radiation sensitivity. *Proc Natl Acad Sci U S A*, 98:12608–12613.

[Bouvier et al., 2018] 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*, 285:1097–1110.

[Bruck and Kaplan, 2015] Bruck, I. and Kaplan, D. L. (2015). The Dbf4-Cdc7 kinase promotes Mcm2-7 ring opening to allow for single-stranded DNA extrusion and helicase assembly. *J Biol Chem*, 290:1210–1221.

[Burgers, 2009] Burgers, P. M. J. (2009). Polymerase dynamics at the eukaryotic DNA replication fork. *J Biol Chem*, 284:4041–4045.

[Bury et al., 2019] Bury, M., Le Calve, B., Lessard, F., Dal Maso, T., Saliba, J., Michiels, C., Ferbeyre, G., and Blank, V. (2019). NFE2L3 controls colon cancer cell growth through regulation of DUX4, a CDK1 inhibitor. *Cell Rep*, 29:1469–1481.e9.

[Castilho et al., 2018] Castilho, V. V. S., Goncalves, K. C. S., Rebello, K. M., Baptista, L. P. R., Sangenito, L. S., Santos, H. L. C., Branquinha, M. H., Santos, A. L. S., Menna-Barreto, R. F. S., Guimaraes, A. C., and dAvila Levy, C. M. (2018). Docking simulation between HIV peptidase inhibitors and Trypanosoma cruzi aspartyl peptidase. *BMC Res Notes*, 11:825.

[Champoux, 2001] Champoux, J. J. (2001). DNA topoisomerases: structure, function, and mechanism. *Annu Rev Biochem*, 70:369–413.

[Chowdhury et al., 2017] Chowdhury, A. M. M. 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*, 7:12494.

[Ciccia and Elledge, 2010] Ciccia, A. and Elledge, S. J. (2010). The DNA damage response: making it safe to play with knives. *Mol Cell*, 40:179–204.

[Clarke et al., 2001] 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*, 21:1997–2007.

[Cortez, 2015] Cortez, D. (2015). Preventing replication fork collapse to maintain genome integrity. *DNA Repair*, 32:149–157.

[Costanzo et al., 2016] 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.

[Diaz-Martinez et al., 2006] 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.

[Dirac-Svejstrup et al., 2020] Dirac-Svejstrup, A. B., Walker, J., Faull, P., Encheva, V., Akimov, V., Puglia, M., Perkins, D., Kumper, S., Hunjan, S. S., Blagoev, B., Snijders, A. P., Powell, D. J., and Svejstrup, J. Q. (2020). DDI2 is a ubiquitin-directed endoprotease responsible for cleavage of transcription factor NRF1. *Mol Cell*, 79:332–341.e7.

[Dixit et al., 2014] Dixit, G., Baker, R., Sacks, C. M., Torres, M. P., and Dohlman, H. G. (2014). Guanine nucleotide-binding protein (GÎ±) 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.

[Dolinsky et al., 2004] Dolinsky, T. J., Nielsen, J. E., McCammon, J. A., and Baker, N. A. (2004). PDB2PQR: an automated pipeline for the setup of Poisson-Boltzmann electrostatics calculations. *Nucleic Acids Res*, 32:W665–W667.

[Duxin et al., 2014] Duxin, J. P., Dewar, J. M., Yardimci, H., and Walter, J. C. (2014). Repair of a DNA-protein crosslink by replication-coupled proteolysis. *Cell*, 159:346–357.

[Elford, 1968] Elford, H. L. (1968). Effect of hydroxyurea on ribonucleotide reductase. *Biochem Biophys Res Commun*, 33:129–135.

[Elsasser and Finley, 2005] Elsasser, S. and Finley, D. (2005). Delivery of ubiquitinated substrates to protein-unfolding machines. *Nat Cell Biol*, 7:742–749.

[Elu et al., 2019] Elu, N., Osinalde, N., Beaskoetxea, J., Ramirez, J., Lectez, B., Aloria, K., Rodriguez, J. A., Arizmendi, J. M., and Mayor, U. (2019). Detailed dissection of UBE3A-mediated DDI1 ubiquitination. *Front Physiol*, 10:534.

[Fragkos et al., 2015] Fragkos, M., Ganier, O., Coulombe, P., and Mechali, M. (2015). DNA replication origin activation in space and time. *Nat Rev Mol Cell Biol*, 16:360–374.

[Fu et al., 2008] 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*, 133:601-611.

[Gabriely et al., 2008] 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*, 19:3625–3637.

[Gambus et al., 2009] Gambus, A., van Deursen, F., Polychronopoulos, D., Foltman, M., Jones, R. C., Edmondson, R. D., Calzada, A., and Labib, K. (2009). A key role for Ctf4 in coupling the MCM2-7 helicase to DNA polymerase alpha within the eukaryotic replisome. *EMBO J*, 28:2992–3004.

[Garcia-Rodriguez et al., 2016] Garcia-Rodriguez, N., Wong, R. P., and Ulrich, H. D. (2016). Functions of ubiquitin and SUMO in DNA replication and replication stress. *Front Genet*, 7:87.

[Gomez et al., 2011] 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.

[Guthmueller et al., 2011] 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*, 53:2777.

[Heessen et al., 2005] Heessen, S., Masucci, M. G., and Dantuma, N. P. (2005). The UBA2 domain functions as an intrinsic stabilization signal that protects Rad23 from proteasomal degradation. *Mol Cell*, 18:225–235.

[Heller et al., 2011] Heller, R. C., Kang, S., Lam, W. M., Chen, S., Chan, C. S., and Bell, S. P. (2011). Eukaryotic origin-dependent DNA replication in vitro reveals sequential action of DDK and S-CDK kinases. *Cell*, 146:80–91.

[Ivantsiv et al., 2006] 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*, 26:1579–1588.

[Johnson and Skotheim, 2013] Johnson, A. and Skotheim, J. M. (2013). Start and the restriction point. *Curr Opin Cell Biol*, 25:717–723.

[Johnson and Walker, 1999] Johnson, D. G. and Walker, C. L. (1999). Cyclins and cell cycle checkpoints. *Annu Rev Pharmacol Toxicol*, 39:295–312.

[Kama et al., 2018] Kama, R., Gabriely, G., Kanneganti, V., and Gerst, J. E. (2018). Cdc48 and ubiquilins confer selective anterograde protein sorting and entry into the multivesicular body in yeast. *Mol Biol Cell*, 29:948–963.

[Kang et al., 2006] Kang, Y., Vossler, R. A., Diaz-Martinez, L. A., Winter, N. S., Clarke, D. J., and Walters, K. J. (2006). UBL/UBA ubiquitin receptor proteins bind a common tetraubiquitin chain. *J Mol Biol*, 356:1027–1035.

[Kanke et al., 2012] Kanke, M., Kodama, Y., Takahashi, T. S., Nakagawa, T., and Masukata, H. (2012). Mcm10 plays an essential role in origin DNA unwinding after loading of the CMG components. *EMBO J*, 31:2182–2194.

[Kaplun et al., 2006] Kaplun, L., Ivantsiv, Y., Bakhrat, A., Tzirkin, R., Baranes, K., Shabek, N., and Raveh, D. (2006). The F-box protein, Ufo1, maintains genome stability by recruiting the yeast mating switch endonuclease, Ho, for rapid proteasome degradation. *Isr Med Assoc J*, 8:246–248.

[Kaplun et al., 2005] 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*, 25:5355–5362.

[Kastan and Bartek, 2004] Kastan, M. B. and Bartek, J. (2004). Cell-cycle checkpoints and cancer. *Nature*, 432:316–323.

[Kim et al., 2004] Kim, I., Mi, K., and Rao, H. (2004). Multiple interactions of Rad23 suggest a mechanism for ubiquitylated substrate delivery important in proteolysis. *Mol Biol Cell*, 15:3357-3365.

[Koizumi et al., 2016] 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:e18357.

[Kottemann et al., 2018] 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*, 69:24–35.e5.

[Krylov and Koonin, 2001] 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:R584–R587.

[Kumar and Suguna, 2018] 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*, 8:1379–1394.

[Larsen et al., 2019] 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. (2019). Replication-coupled DNA-protein crosslink repair by SPRTN and the proteasome in Xenopus egg extracts. *Mol Cell*, 73:574–588.e7.

[Lehmann et al., 2007] Lehmann, A. R., Niimi, A., Ogi, T., Brown, S., Sabbioneda, S., Wing, J. F., Kannouche, P. L., and Green, C. M. (2007). Translesion synthesis: Y-family polymerases and the polymerase switch. *DNA Repair*, 6:891–899.

[Lehrbach and Ruvkun, 2016] Lehrbach, N. J. and Ruvkun, G. (2016). Proteasome dysfunction triggers activation of SKN-1A/Nrf1 by the aspartic protease DDI-1. *eLife*, 8:e44425.

[Lessel et al., 2014] Lessel, D., Vaz, B., Halder, S., Lockhart, P. J., Marinovic-Terzic, I., Lopez-Mosqueda, J., Philipp, M., Sim, J. C. H., Smith, K. R., Oehler, J., Cabrera, E., Freire, R., Pope, K., Nahid, A., Norris, F., Leventer, R. J., Delatycki, M. B., Barbi, G., von Ameln, S., Hogel, J., Degoricija, M., Fertig, R., Burkhalter, M. D., Hofmann, K., Thiele, H., Altmuller, J., Nurnberg, G., Nurnberg, P., Bahlo, M., Martin, G. M., Aalfs, C. M., Oshima, J., Terzic, J., Amor, D. J., Dikic, I., Ramadan, K., and Kubisch, C. (2014). Mutations in SPRTN cause early onset hepatocellular carcinoma, genomic instability and progeroid features. *Nat Genet*, 46:1239–1244.

[Li et al., 2019] Li, F., Raczynska, J. E., Chen, Z., and Yu, H. (2019). Structural insight into DNA-dependent activation of human metalloprotease Spartan. *Cell Rep*, 26:3336–3346.e4.

[Ling et al., 2016] Ling, C., Huang, J., Yan, Z., Li, Y., Ohzeki, M., Ishiai, M., Xu, D., Takata, M., Seidman, M., and Wang, W. (2016). Bloom syndrome complex promotes FANCM recruitment to stalled replication forks and facilitates both repair and traverse of DNA interstrand crosslinks. *Cell Discov*, 2:16047.

[Liu et al., 1997] 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*, 255:533–542.

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

[Lopez-Mosqueda et al., 2016] 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:e21491.

[Lopez-Velez, 2003] Lopez-Velez, R. (2003). The impact of highly active antiretroviral therapy (HAART) on visceral leishmaniasis in Spanish patients who are co-infected with HIV. *Ann Trop Med Parasitol*, 97 Suppl 1:143–147.

[Lustgarten and Gerst, 1999] 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*, 19:4480-4494.

[Maddi et al., 2020] Maddi, K., Sam, D. K., Bonn, F., Prgomet, S., Tulowetzke, E., Akutsu, M., Lopez-Mosqueda, J., and Dikic, I. (2020). Wss1 promotes replication stress tolerance by degrading histones. *Cell Rep*, 30:3117–3126.e4.

[Maffia et al., 2020] Maffia, A., Ranise, C., and Sabbioneda, S. (2020). From R-loops to G-quadruplexes: Emerging new threats for the replication fork. *Int J Mol Sci*, 21.

[Mailand and Diffley, 2005] Mailand, N. and Diffley, J. F. X. (2005). CDKs promote DNA replication origin licensing in human cells by protecting Cdc6 from APC/C-dependent proteolysis. *Cell*, 122:915–926.

[Marash and Gerst, 2003] 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*, 14:3114–3125.

[Maskey et al., 2014] 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.

[Mazouzi et al., 2014] Mazouzi, A., Velimezi, G., and Loizou, J. I. (2014). DNA replication stress: Causes, resolution and disease. *Exp Cell Res*, 329(1):85–93.

[Morawe et al., 2011] 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*, 193:71–80.

[Morocz et al., 2017] 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 Research*, 45:3172–3188.

[Mullen et al., 2010] Mullen, J. R., Chen, C.-F., and Brill, S. J. (2010). Wss1 is a SUMO-dependent isopeptidase that interacts genetically with the Slx5-Slx8 SUMO-targeted ubiquitin ligase. *Mol Cell Biol*, 30:3737–3748.

[Mullen et al., 2011] Mullen, J. R., Das, M., and Brill, S. J. (2011). Genetic evidence that polysumoylation bypasses the need for a SUMO-targeted Ub ligase. *Genetics*, 187:73–87.

[Nowak et al., 2018] Nowak, K., Taubert, R. M., Haberecht, S., Venz, S., and Kruger, E. (2018). Inhibition of calpain-1 stabilizes TCF11/Nrf1 but does not affect its activation in response to proteasome inhibition. *Biosci Rep*, 38(5):BSR20180393.

[Nowicka et al., 2015] 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*, 23:542–557.

[O'Donnell and Kurth, 2013] O'Donnell, M. E. and Kurth, I. (2013). *Machinery of DNA Replication*, pages 1327–1338. Springer Berlin Heidelberg, Berlin, Heidelberg.

[Oh et al., 2017] Oh, K. H., Haney, J. J., Wang, X., Chuang, C.-F., Richmond, J. E., and Kim, H. (2017). ERG-28 controls BK channel trafficking in the ER to regulate synaptic function and alcohol response in C. elegans. *eLife*, 6:e24733.

[Onchieku et al., 2018] Onchieku, N. M., Mogire, R., Ndungu, L., Mwitari, P., Kimani, F., Matoke-Muhia, D., Kiboi, D., and Magoma, G. (2018). Deciphering the targets of retroviral protease inhibitors in Plasmodium berghei. *PLoS ONE*, 13:e0201556.

[O'Neill et al., 2004] 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*, 32:6519–6530.

[Pacek et al., 2006] Pacek, M., Tutter, A. V., Kubota, Y., Takisawa, H., and Walter, J. C. (2006). Localization of MCM2-7, Cdc45, and GINS to the site of DNA unwinding during eukaryotic DNA replication. *Mol Cell*, 21:581–587.

[Pan et al., 2006]Pan, X., Ye, P., Yuan, D. S., Wang, X., Bader, J. S., and Boeke, J. D. (2006). A DNA integrity network in the yeast Saccharomyces cerevisiae. *Cell*, 124:1069–1081.

[Perteguer et al., 2013] 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*, 18:171–181.

[Petermann et al., 2010] Petermann, E., Orta, M. L., Issaeva, N., Schultz, N., and Helleday, T. (2010). Hydroxyurea-stalled replication forks become progressively inactivated and require two different RAD51-mediated pathways for restart and repair. *Mol Cell*, 37:492–502.

[Pietenpol and Stewart, 2002] Pietenpol, J. A. and Stewart, Z. A. (2002). Cell cycle checkpoint signaling: cell cycle arrest versus apoptosis. *Toxicology*, 181-182:475–481.

[Ramirez et al., 2018] 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*, 27:1955–1971.

[Remus et al., 2009] Remus, D., Beuron, F., Tolun, G., Griffith, J. D., Morris, E. P., and Diffley, J. F. X. (2009). Concerted loading of Mcm2-7 double hexamers around DNA during DNA replication origin licensing. *Cell*, 139:719–730.

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

[Schilling et al., 2011] Schilling, O., Huesgen, P. F., Barre, O., Auf dem Keller, U., and Overall, C. M. (2011). 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.

[Serbyn et al., 2020] Serbyn, N., Noireterre, A., Bagdiul, I., Plank, M., Michel, A. H., Loewith, R., Kornmann, B., and Stutz, F. (2020). The aspartic protease Ddi1 contributes to DNA-protein crosslink repair in yeast. *Mol Cell*, 77:1066–1079.e9.

[Sherman, 2002] Sherman, F. (2002). Getting started with yeast. *Methods Enzymol*, 350:3–41.

[Sirkis et al., 2006] Sirkis, R., Gerst, J. E., and Fass, D. (2006). Ddi1, a eukaryotic protein with the retroviral protease fold. *J Mol Biol*, 364:376–387.

[Stingele et al., 2016] 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*, 64:688–703.

[Stingele et al., 2014] 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*, 158:327–338.

[Strisovsky et al., 2000] Strisovsky, K., Tessmer, U., Langner, J., Konvalinka, J., and Krausslich, H. G. (2000). Systematic mutational analysis of the active-site threonine of HIV-1 proteinase: rethinking the "fireman's grip" hypothesis. *Protein Sci*, 9:1631–1641.

[Tkach et al., 2012] Tkach, J. M., Yimit, A., Lee, A. Y., Riffle, M., Costanzo, M., Jaschob, D., Hendry, J. A., Ou, J., Moffat, J., Boone, C., Davis, T. N., Nislow, C., and Brown, G. W. (2012). Dissecting DNA damage response pathways by analysing protein localization and abundance changes during DNA replication stress. *Nat Cell Biol*, 14:966–976.

[Trempe et al., 2005] Trempe, J.-F., Brown, N. R., Lowe, E. D., Gordon, C., Campbell, I. D., Noble, M. E. M., and Endicott, J. A. (2005). Mechanism of Lys48-linked polyubiquitin chain recognition by the Mud1 UBA domain. *EMBO J*, 24:3178–3189.

[van den Boom et al., 2016] van den Boom, J., Wolf, M., Weimann, L., Schulze, N., Li, F., Kaschani, F., Riemer, A., Zierhut, C., Kaiser, M., Iliakis, G., Funabiki, H., and Meyer, H. (2016). VCP/p97 extracts sterically trapped Ku70/80 rings from DNA in double-strand break repair. *Mol Cell*, 64:189–198.

[van Heusden and Steensma, 2008] van Heusden, G. P. H. and Steensma, H. Y. (2008). The Saccharomyces cerevisiae Wss1 protein is only present in mother cells. *FEMS Microbiol Lett*, 282:100–104.

[Vaz et al., 2016] Vaz, B., Popovic, M., Newman, J. A., Fielden, J., Aitkenhead, H., Halder, S., Singh, A. N., Vendrell, I., Fischer, R., Torrecilla, I., Drobnitzky, N., Freire, R., Amor, D. J., Lockhart, P. J., Kessler, B. M., McKenna, G. W., Gileadi, O., and Ramadan, K. (2016). Metalloprotease SPRTN/DVC1 orchestrates replication-coupled DNA-protein crosslink repair. *Mol Cell*, 64:704-719.

[Verma et al., 2004] 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*, 118:99–110.

[Vermeulen et al., 2003] Vermeulen, K., Van Bockstaele, D. R., and Berneman, Z. N. (2003). The cell cycle: a review of regulation, deregulation and therapeutic targets in cancer. *Cell Prolif*, 36:131–149.

[Voloshin et al., 2012] Voloshin, O., Bakhrat, A., Herrmann, S., and Raveh, D. (2012). 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.

[Waga and Stillman, 1994] Waga, S. and Stillman, B. (1994). Anatomy of a DNA replication fork revealed by reconstitution of SV40 DNA replication in vitro. *Nature*, 369:207–212.

[Wang et al., 2016] Wang, Y., Xu, M., and Jiang, T. (2016). Crystal structure of human PCNA in complex with the PIP box of DVC1. *Biochem Biophys Res Commun*, 474:264–270.

[White et al., 2011a] White, R. E., Dickinson, J. R., Semple, C. A. M., Powell, D. J., and Berry, C. (2011a). The retroviral proteinase active site and the N-terminus of Ddi1 are required for repression of protein secretion. *FEBS Lett*, 585:139–142.

[White et al., 2011b] White, R. E., Powell, D. J., and Berry, C. (2011b). HIV proteinase inhibitors target the Ddi1-like protein of Leishmania parasites. *FASEB J*, 25:1729–1736.

[Woudstra et al., 2002] Woudstra, E. C., Gilbert, C., Fellows, J., Jansen, L., Brouwer, J., Erdjument-Bromage, H., Tempst, P., and Svejstrup, J. Q. (2002). A Rad26-Def1 complex coordinates repair and RNA pol II proteolysis in response to DNA damage. *Nature*, 415:929–933.

[Yamada et al., 2013] Yamada, M., Watanabe, K., Mistrik, M., Vesela, E., Protivankova, I., Mailand, N., Lee, M., Masai, H., Lukas, J., and Bartek, J. (2013). ATR-Chk1-APC/CCdh1-dependent stabilization of Cdc7-ASK (Dbf4) kinase is required for DNA lesion bypass under replication stress. *Genes Dev*, 27:2459–2472.

[Yang et al., 2019] Yang, L., Yang, Z., Liu, C., He, Z., Zhang, Z., Yang, J., Liu, H., Yang, J., and Ji, Y. (2019). Chloroplast phylogenomic analysis provides insights into the evolution of the largest eukaryotic genome holder, Paris japonica (Melanthiaceae). *BMC Plant Biol*, 19:293.

[Yang et al., 2017] Yang, X., Li, Y., Gao, Z., Li, Z., Xu, J., Wang, W., and Dong, Y. (2017). Structural analysis of Wss1 protein from Saccharomyces cerevisiae. *Sci Rep*, 7:8270.

[Yip et al., 2020] Yip, M. C. J., Bodnar, N. O., and Rapoport, T. A. (2020). Ddi1 is a ubiquitin-dependent protease. *Proc Natl Acad Sci U S A*, 117:7776–7781.

[Zeman and Cimprich, 2013] Zeman, M. K. and Cimprich, K. A. (2013). Causes and consequences of replication stress. *Nat Cell Biol*, 16(1):2–9.

[Zhang et al., 2005] Zhang, Y.-W., Otterness, D. M., Chiang, G. G., Xie, W., Liu, Y.-C., Mercurio, F., and Abraham, R. T. (2005). Genotoxic stress targets human Chk1 for degradation by the ubiquitin-proteasome pathway. *Mol Cell*, 19:607–618.

[Zhu and Xiao, 1998] Zhu, Y. and Xiao, W. (1998). Differential regulation of two closely clustered yeast genes, MAG1 and DDI1, by cell-cycle checkpoints. *Nucleic Acids Res*, 26:5402–5408.

[Zhu and Xiao, 2004] Zhu, Y. and Xiao, W. (2004). Pdr3 is required for DNA damage induction of MAG1 and DDI1 via a bi-directional promoter element. *Nucleic Acids Res*, 32:5066–5075.

SUPPLEMENTS

8. SUPPLEMENTS

8.1. SUPPLEMENT S1: HUMAN DNA DAMAGE-INDUCIBLE 2 PROTEIN IS STRUCTURALLY AND FUNCTIONALLY DISTINCT FROM ITS YEAST ORTHOLOG

Sivá M.*, **Svoboda M.*,** Veverka V., Trempe J.F., HofmannK., Kožíšek M., Hexnerová R., Sedlák F., Belza J., Brynda J., Šácha P., Hubálek M., Starková J., Flaisigová I., Konvalinka J., and Grantz Šašková K. Human DNA-damage-inducible 2 protein is structurally and functionally distinct from its yeast ortholog. *Sci Rep.* **6**, 30443 (2016).

* Equal contribution.

My contribution: cloning of ΔUIM and catalytically inactive D252A variants of human *DDI2*; bacterial expression and purification of full length human DDI2 protein, it's RVP domain and catalytically inactive variants of both; expression and purification of epitope-tagged full-length human DDI2 in mammalian cells; enzymatic synthesis of diubiquitin chains, pull-down assays for testing human DDI2 – ubiquitin interaction; PICS proteomics screen for substrates of DDI2 protease activity; BSA cleavage assay; surface electrostatics analysis.

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

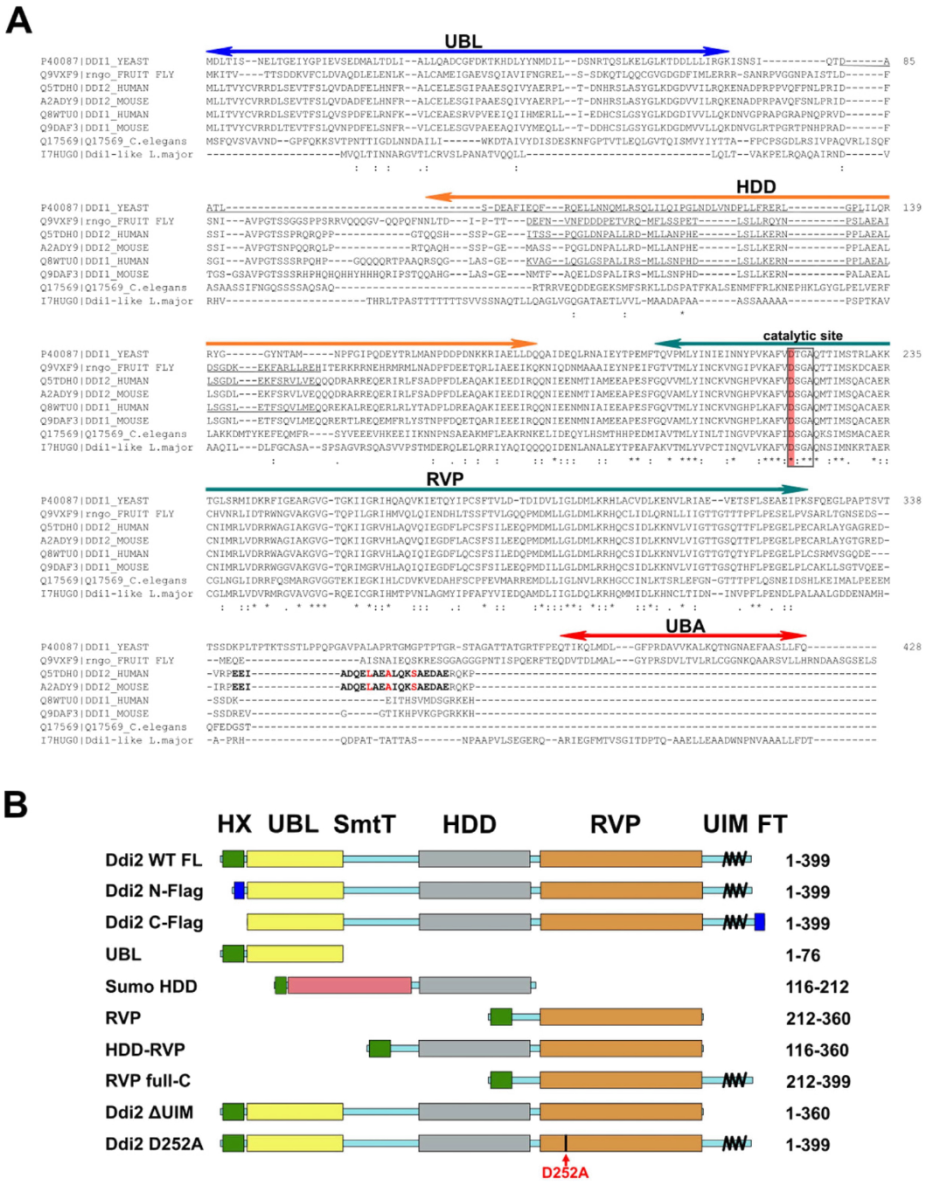


Figure 1. Sequence analysis of Ddi1 orthologs. (**A**) Sequence alignment of Ddi1-like proteins from various eukaryotic organisms. Domains are indicated with double-headed arrows. The highly conserved catalytic site of RVP is highlighted. The putative UIM motif is highlighted in bold, with residues important 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

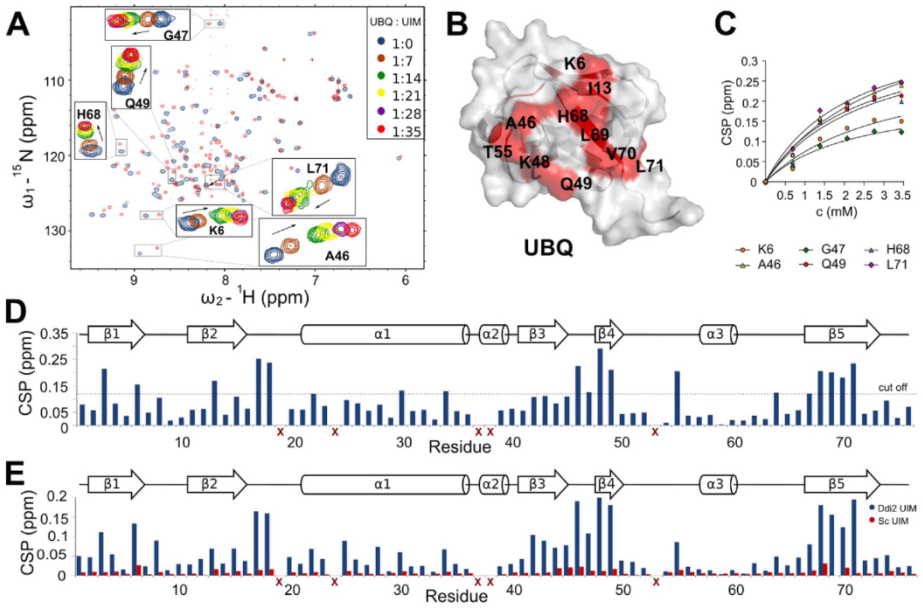


Figure 2. Mapping of the UBQ-hDdi2 interaction site. (**A**) $^{15}N/^{1}H$-HSQC titration spectra of UBQ with hDdi2-UIM peptide. (**B**) Identification of mapped residues shown on the UBQ structure (PDB entry 1D3Z)[32]. (**C**) Titration curves of selected amino acids on UBQ. (**D**) Plot of chemical 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

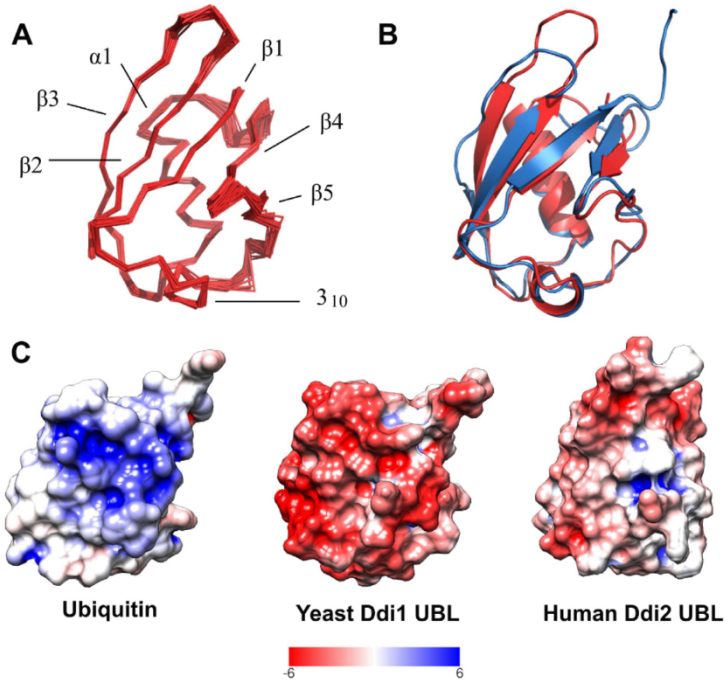


Figure 3. Solution structure of the hDdi2 UBL domain. (**A**) Superimposition of 40 converged structures of the UBL domain. (**B**) Structural alignment of solution structures of the yDdi1 UBL in blue (PDB code 2N7E) and hDdi2 UBL in orange (PDB code 2N7D). The structural alignment over 74 equivalent 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

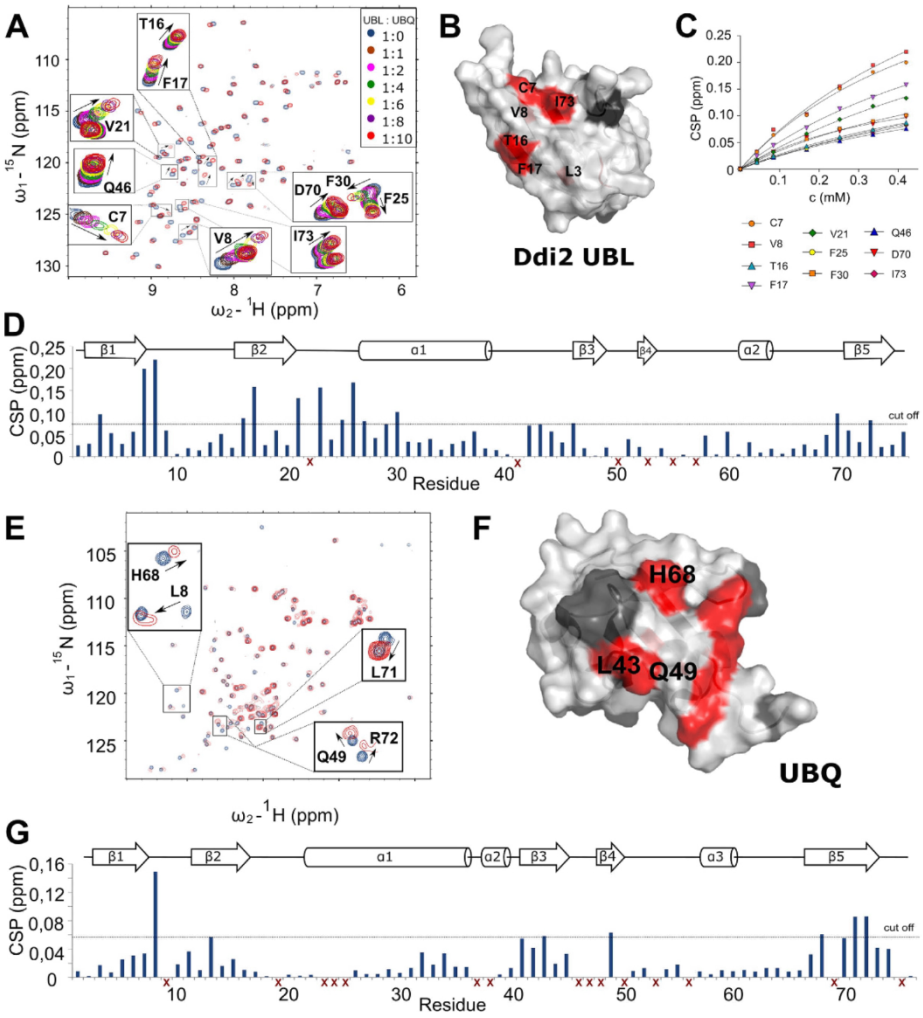


Figure 4. Characterization of the hDdi2 UBL interaction with UBQ. (**A**) $^{15}N/^{1}H$-HSQC titration spectra of Ddi2 UBL with addition of a 1-, 2-, 4-, 6-, 8-, or 10-fold molar excess of UBQ. Residues Cys7, Val8, Thr16, Phe17, Val21, Phe25, Phe30, Gln46, Asp70, and Ile73 were used for K_d calculation (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

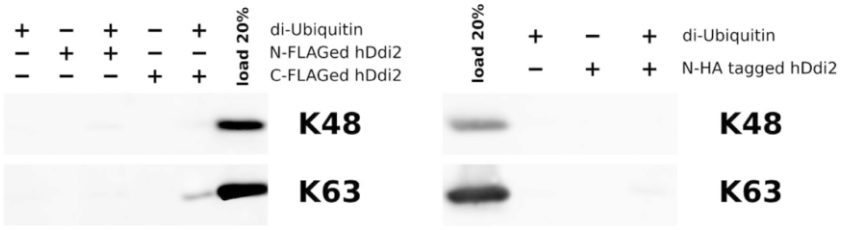


Figure 5. Human Ddi2 shows no strong interaction with di-ubiquitin chains. Western blot analysis of pull-down experiments with di-ubiquitin conjugates of Lys48 and Lys63 architecture. Human Ddi2 with a FLAG tag on either the N- or C-terminus or an HA tag on the N-terminus was immobilized on magnetic 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

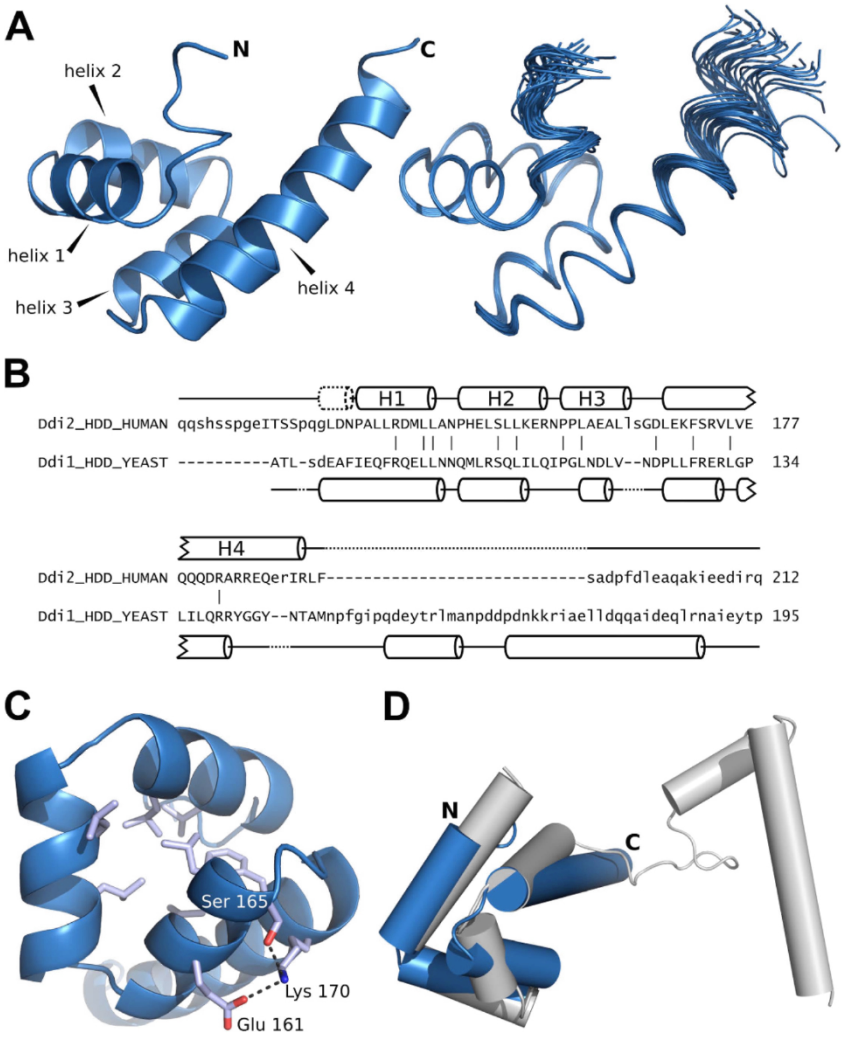


Figure 6. Solution structure of hDdi2 HDD. (**A**) Superimposition of 30 converged structures of HDD. (**B**) Structural alignment of hDdi2 HDD and yDdi1 HDD (PDB code 5KES) analyzed by Dali Pairwise comparison[37]. The Z score for these two structures is 4, and their RMSD is 5 Å. Secondary structures 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

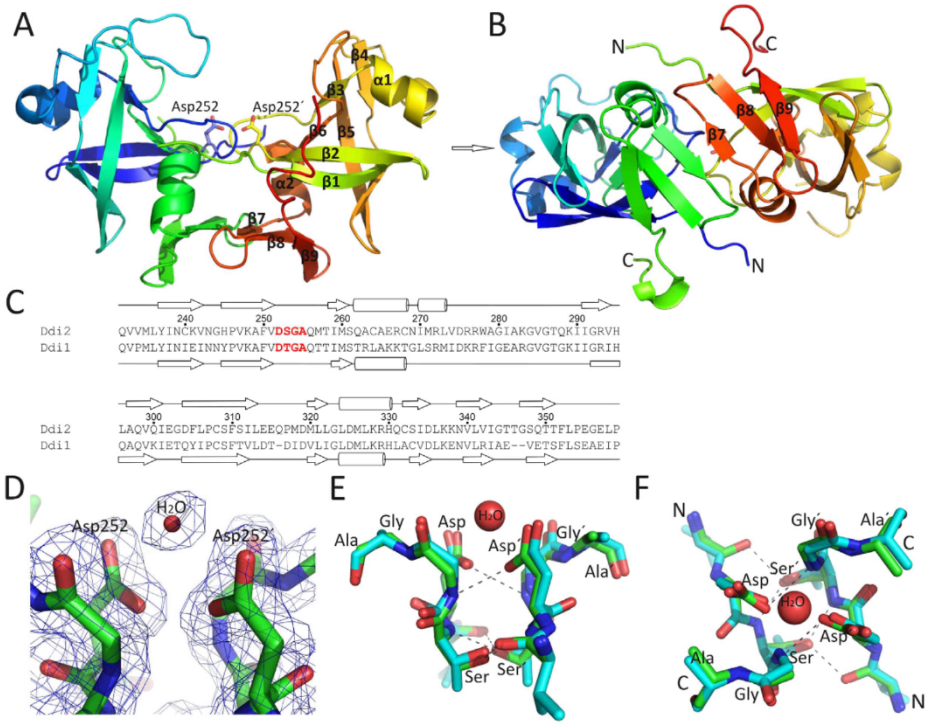


Figure 7. X-ray structure of the hDdi2 RVP domain. (**A**) A ribbon diagram of the structure of the hDdi2 RVP (residues 212–360) dimer (blue N-terminus to red C-terminus). The aspartate side chains that form the putative RVP active site are shown in stick representation. Secondary structure elements 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

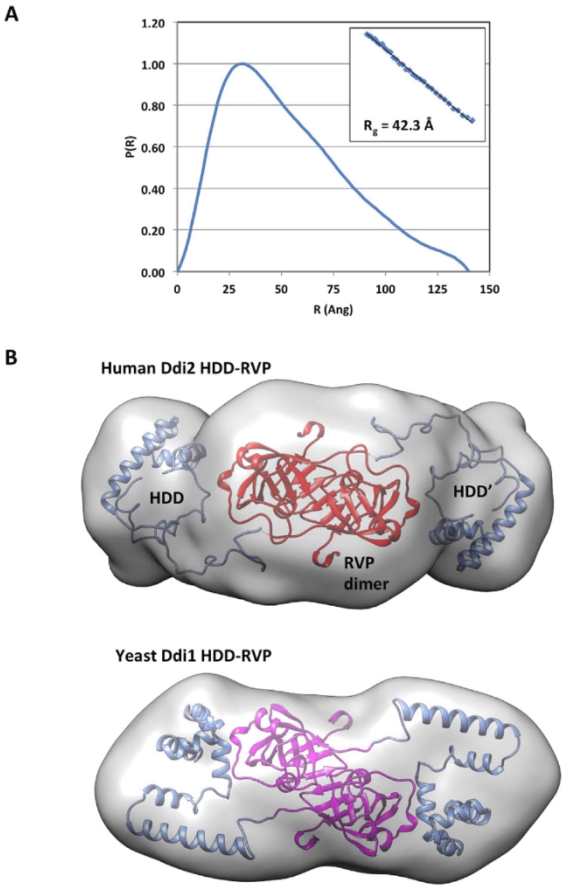


Figure 8. SAXS analysis of the HDD-RVP domains of hDdi2. (**A**) Pair-distance distribution from merged SAXS data, showing the asymmetric distribution characteristic of elongated structures. The inset shows the linearity of the Guinier plot for data collected at 5 mg/ml, indicating monodispersity. (**B**) 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).

Human DNA-Damage-Inducible 2 Protein Is Structurally and Functionally Distinct from Its Yeast Ortholog

Monika Sivá[1,2#], Michal Svoboda[1,3#], Václav Veverka[1], Jean-François Trempe[4], Kay Hofmann[5], Milan Kožíšek[1], Rozálie Hexnerová[1], František Sedlák[1,2], Jan Belza[1,6], Jiří Brynda[1], Pavel Šácha[1], Martin Hubálek[1], Jana Starková[1], Iva Flaisigová[1], Jan Konvalinka[1,6] and Klára Grantz Šašková[1,6*]

[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 Physical and Macromolecular Chemistry, Faculty of Science, Charles University, Hlavova 8, 120 00 Prague 2, Czech Republic

[4]Groupe de Recherche Axé sur la Structure des Protéines, Department of Pharmacology & Therapeutics McGill University, Montreal, QC, H3G 1Y6, Canada

[5]Institute for Genetics, University of Cologne, Zülpicher Str. 47a, 50647 Cologne, Germany

[6]Department of Biochemistry, Faculty of Science, Charles University, Hlavova 8, 120 00 Prague 2, Czech Republic

*Correspondence: saskova@uochb.cas.cz

# Authors contributed equally to this work.

Results

NMR interaction studies of hDdi2

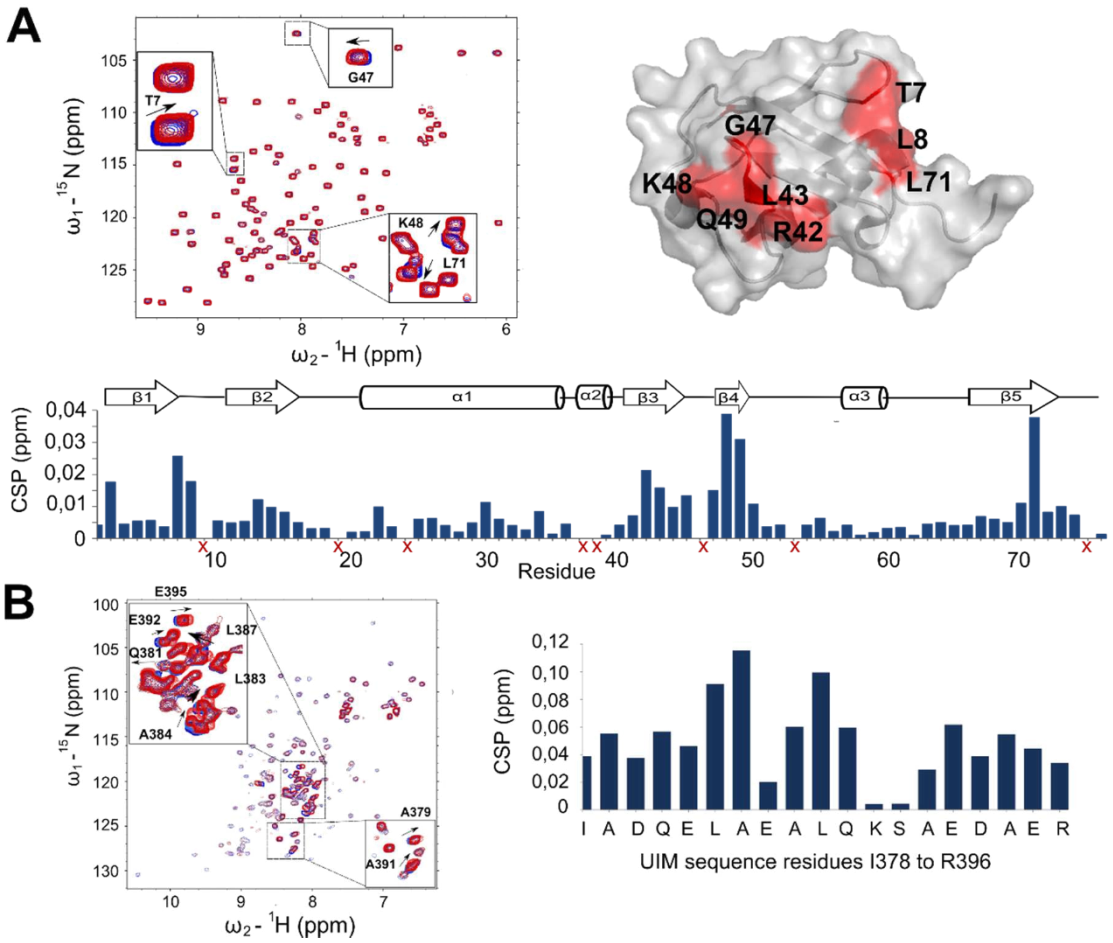


Figure S1: UBQ - hDdi2 UIM interaction verification. A) Upper left: HSQC spectra of ^{15}N-labeled UBQ (blue) with 5-fold molar addition of non-labeled hDdi2 RVP full-C (red) reveals specific but very weak CSPs of UBQ backbone amide signals (below). The shifts are mapped onto the UBQ structure (PDB entry 1D3Z)[3] (upper right). **B)** Reverse mapping of the interaction of Ddi2 UIM with UBQ on HSQC spectra of ^{15}N-labeled hDdi2 RVP full C (blue) with 5-fold molar addition of non-labeled UBQ (red) (left). The plot of CSPs of the hDdi2-UIM sequence locus are shown on the right.

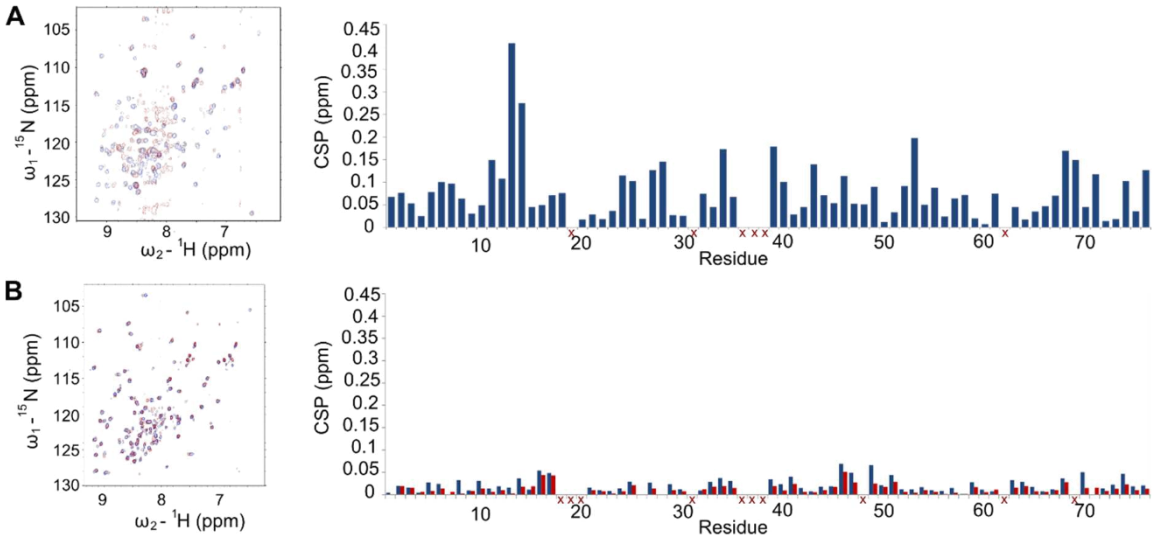


Figure S2: Interaction of Nedd8 with the C-terminal UIM motif of human Ddi2. The interaction of Nedd8 with UIM peptide was verified by acquisition of 2D HSQC spectra under two different conditions: **A)** 50 mM sodium acetate, pH 5, and **B)** 50 mM sodium phosphate, pH 7. Left: Alignments of spectra of ^{15}N-labeled Nedd8 without (blue) and with final addition of the ligand (red). Right: Plots of chemical shift perturbations upon peptide addition **A)** show remarkable shifts in several amino acids at pH 5 but **B)** no perturbation relevant for the interaction at pH 7, which is closer to physiological conditions. Moreover, addition of the scrambled peptide (red) did not change the profile of CSP on Nedd8 compared to the hDdi2 UIM peptide (blue). The interaction at pH 5 is electrostatic, and we propose that there is no physiological interaction between hDdi2 and Nedd8.

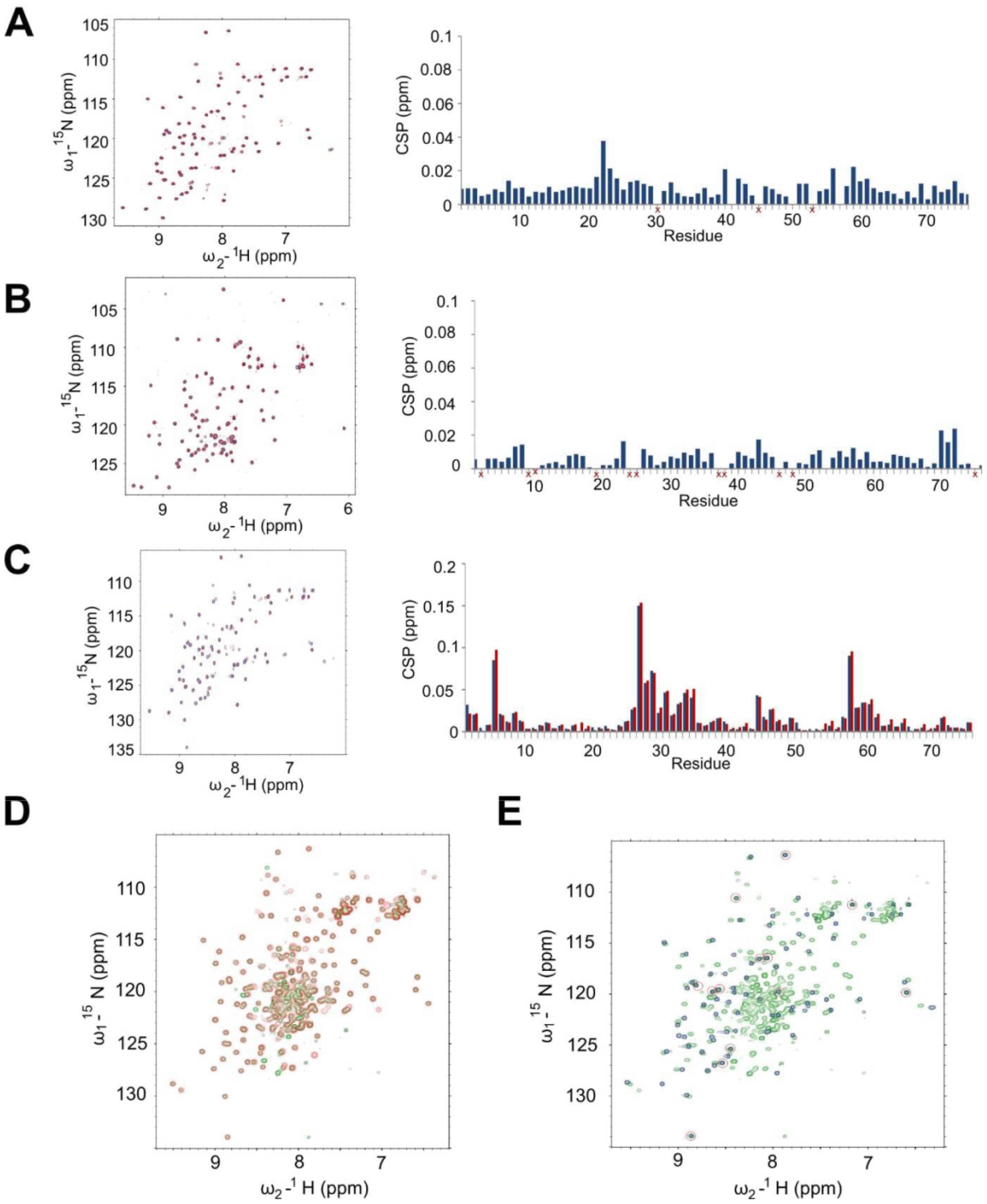


Figure S3: Human Ddi2 UBL domain interactions. A) hDdi2 UBL domain does not bind Nedd8. Left: 2D HSQC spectra of ^{15}N-labeled hDdi2 UBL domain (blue) upon 5-fold molar addition of Nedd8 (red). Right: Plot of CSPs of backbone amides observed after Nedd8 addition. **B)** UBQ does not bind to hDdi2 lacking N- and C-terminal domains. Left: No significant chemical shift perturbations were observed in 2D HSQC spectra of ^{15}N-labeled UBQ

(blue) upon 6-fold molar addition of hDdi2 HDD-RVP (red) lacking the N-terminal UBL domain and C-terminal UIM motif. Spectra monitoring changes in HSQC spectra of UBQ due to addition of hDdi2 with one of the terminal domains showed more significant CSP, assuming these domains bind UBQ *in vitro*, although very weakly. Right: Plot of CSPs of backbone amides observed in this experiment. **C)** The N-terminal UBL domain of hDdi2 does not bind C-terminal UIM peptide derived from hDdi2. Left: 2D HSQC spectra of 0.05 mM hDdi2 UBL before (blue) and after addition of the UIM peptide derived from the hDdi2 C-terminus (red) with final 1.9 mM concentration. Right: Plot of CSPs after addition of UIM-derived peptide (blue) and after addition of a scrambled peptide (red) to a final concentration of 1.2 mM. The interaction of the UBL domain with the peptides shows that the CSPs observed upon addition of UIM peptide are not sequence-specific but purely electrostatic. **D)** hDdi2 UBL domain does not bind the C-terminal UIM motif. 2D HSQC spectra of full-length hDdi2 protein (green) superimposed with 2D HSQC spectra of hDdi2 ΔUIM (red) exhibits no shifts in the signals of UBL domain, thus confirming the previous peptide titration experiments. **E)** hDdi2 UBL does not extend away from the body of the hDdi2 protein. 2D HSQC spectra of hDdi2 UBL (blue) superimposed with the HSQC 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.

Polyubiquitin chain binding is not preserved in human Ddi2

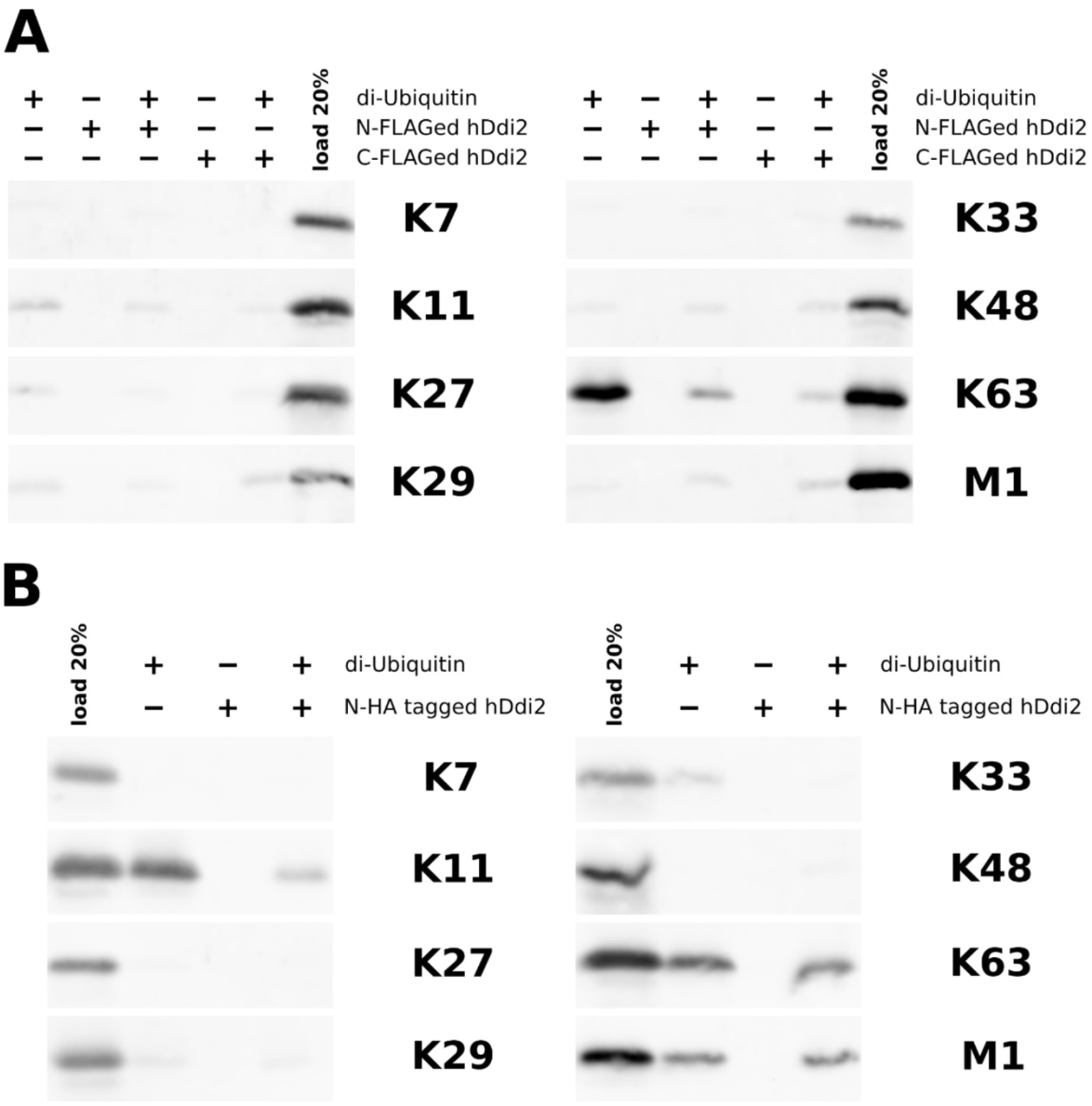


Figure S4: Polyubiquitin chain binding is not preserved in human Ddi2. Western blot analysis of pull-down experiments with di-ubiquitin conjugates of various linkage architecture. Human Ddi2 with a FLAG tag on either the N- or C-terminus **(A)** or an HA tag on the N-terminus **(B)** was immobilized on magnetic agarose beads. Beads were incubated with the di-ubiquitin conjugate of given linkage architecture (purchased from UbiQ, The Netherlands), 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 (Dako).

Sti1-like domain identification

Figure S5: Sequence alignment of the human Ddi1-like proteins, *Drosophila melanogaster* Ddi1 (Rngo), *Saccharomyces pombe* Ddi1 (Mud1), *Saccharomyces cerevisiae* Ddi1, and RAD23A. For secondary structure prediction (H-helix, in red), the most closely related sequence from a known structure of human Rad23A was used (PDB 1oqy)[1].

Search for putative proteolytic activity and small-molecular binder of the RVP domain

PICS analysis:

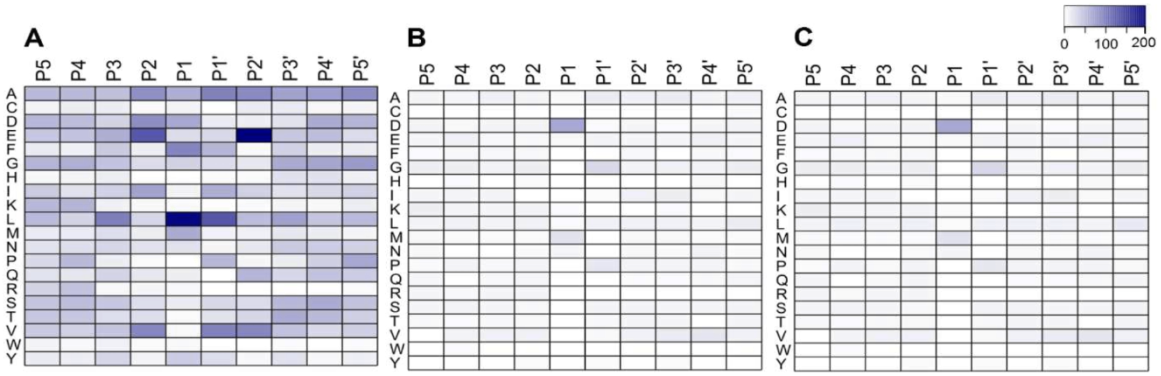


Figure S6: PICS (Proteomic Identification of Protease Substrates) results: Heatmap representation of protease substrate profile demonstrating the total counts of given amino acids (described by single letter code) per particular position (P5−P5′) in peptidic substrates in a HEK293 cell line-derived peptide library. The cleaved peptide bond is positioned between P1 and P1′. **A)** Library cleaved by HIV-1 protease (positive control) **B)** Library cleaved by hDdi2 protein in HEPES pH 7 (proteolytic activity assay). C) Background profile for the uncleaved library (negative control).

HPLC assay for human Ddi2 proteolytic activity

We used an HPLC assay to test putative hydrolysis of a complete set of HIV polyprotein-derived peptide substrates by hDdi2 RVP domain at pH 5.0 (and pH 7.0 for HIV substrates) and various salt concentrations (150 mM - 500 mM NaCl). We also used HPLC to assess putative cleavage of BSA, HSA, β-casein, and insulin by hDdi2 RVP domain.

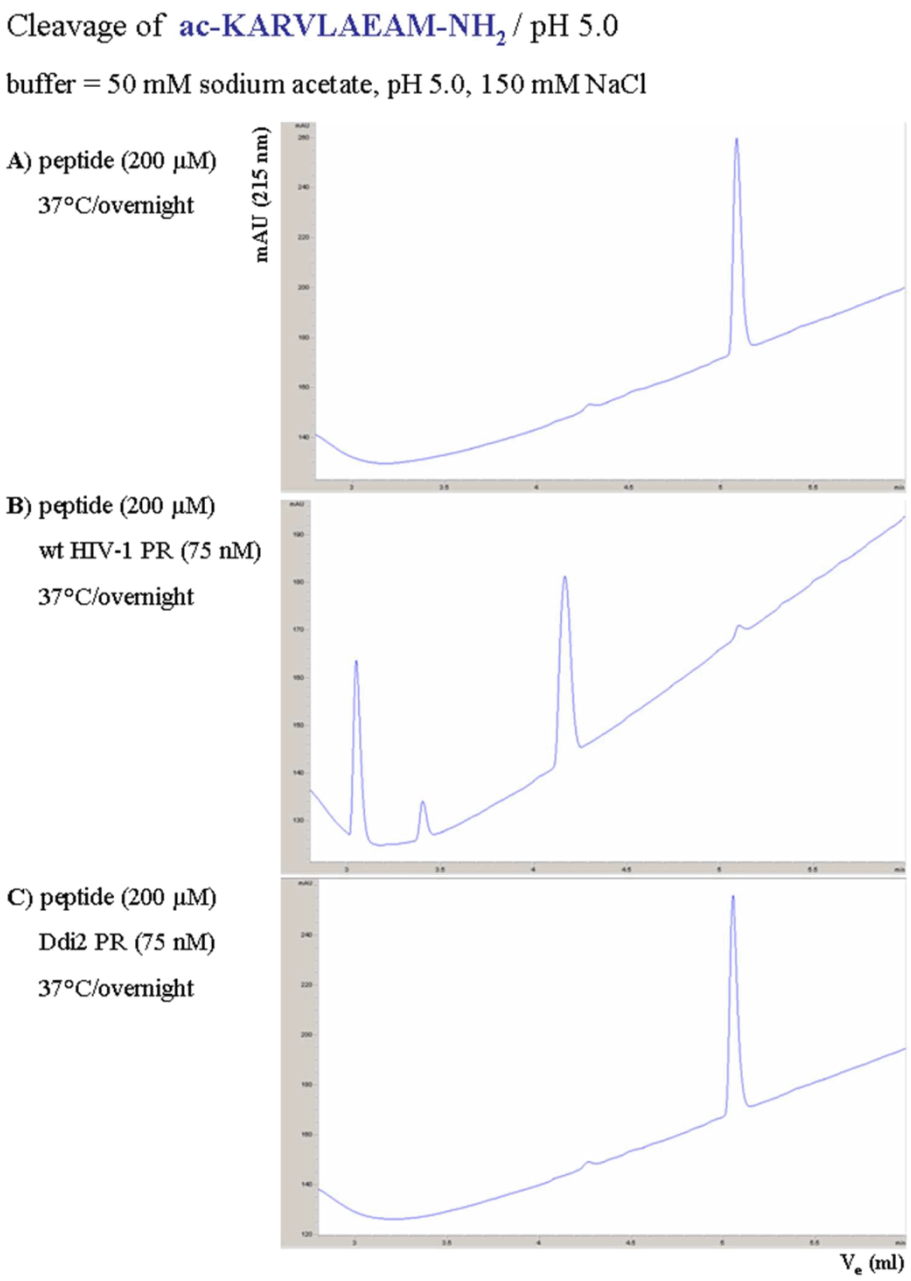


Figure S7: Testing of putative cleavage of an HIV polyprotein-derived peptide substrate with the amino acid sequence KARVLAEAM by hDdi2 RVP domain at pH 5.0 with 150 mM NaCl.

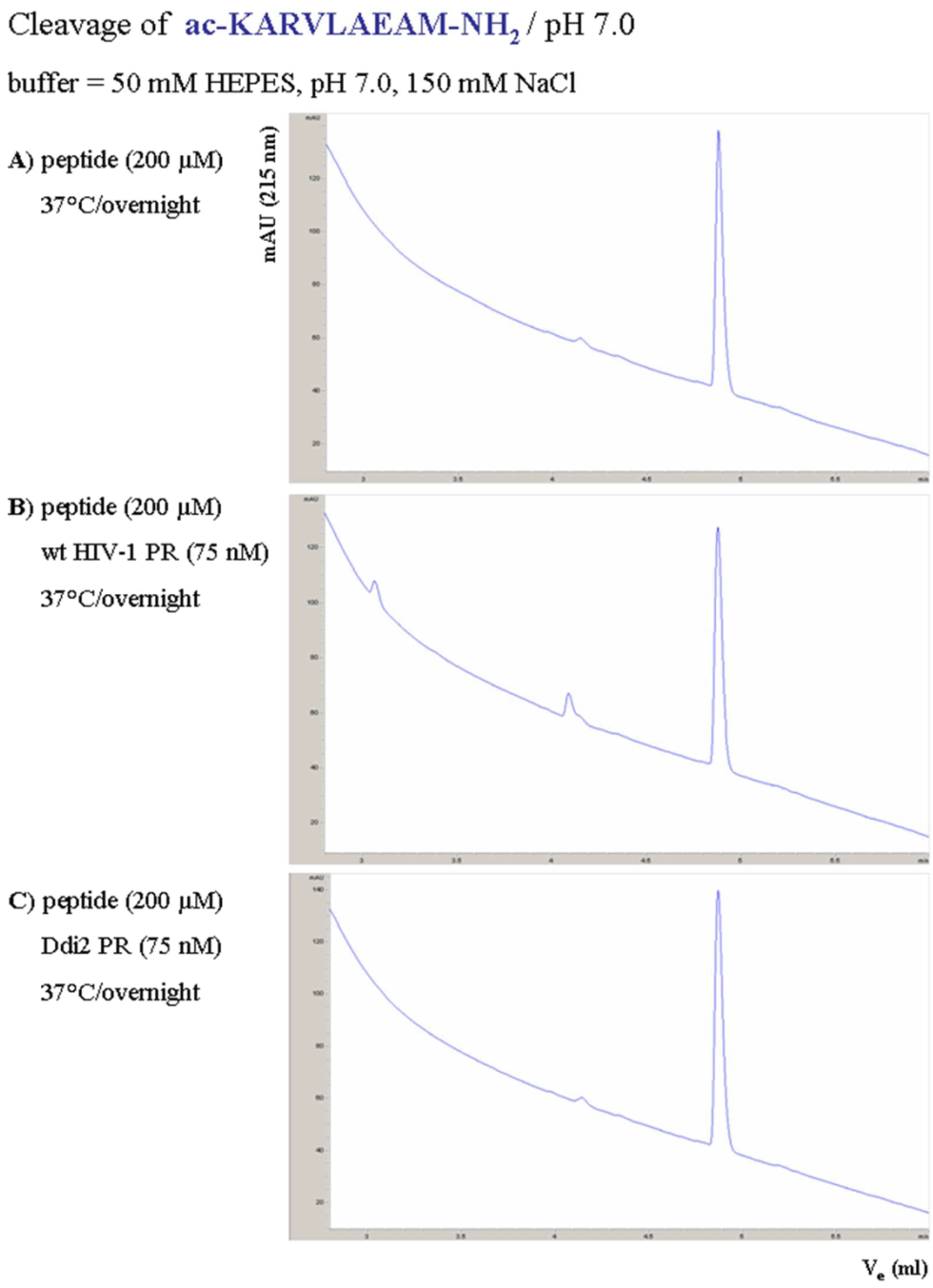


Figure S8: Testing of putative cleavage of an HIV polyprotein-derived peptide substrate with the amino acid sequence KARVLAEAM by hDdi2 RVP domain at pH 7.0 with 150 mM NaCl.

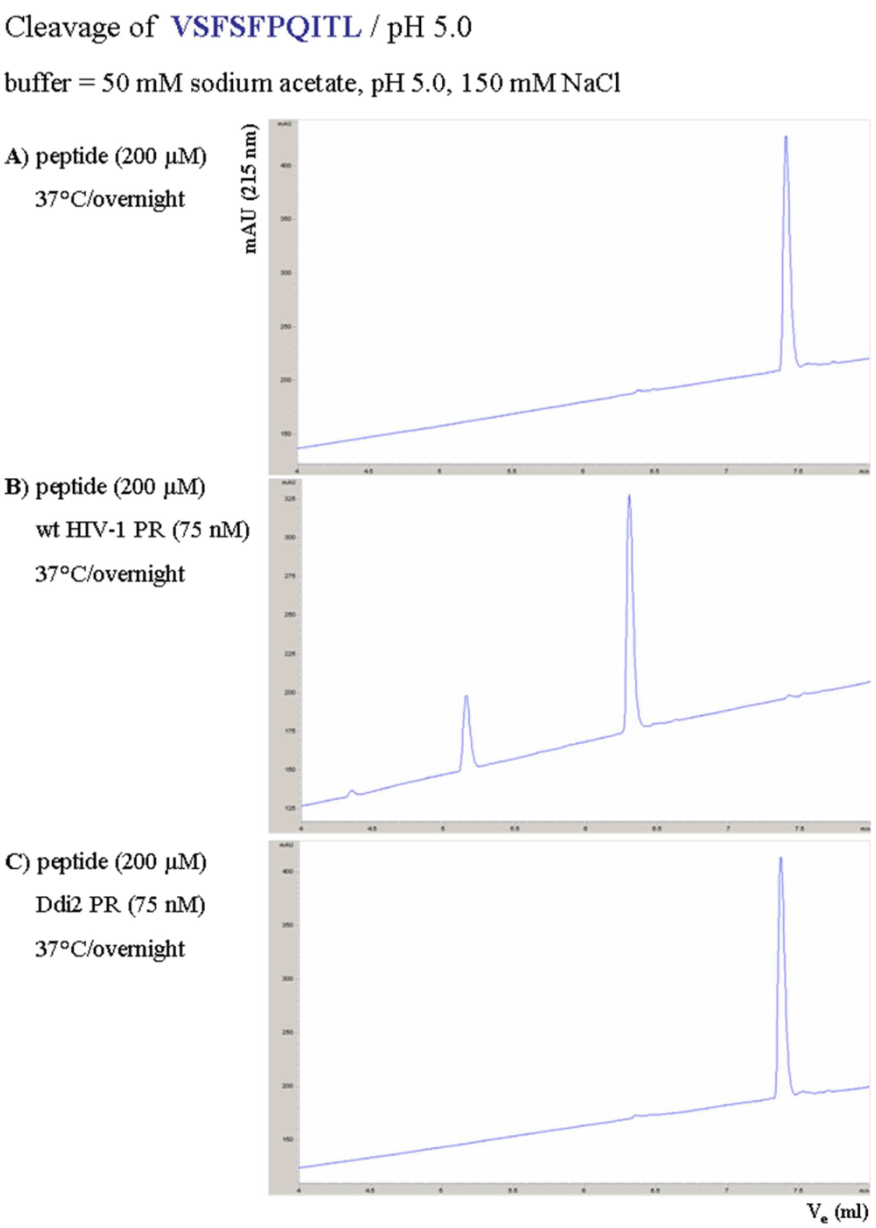


Figure S9: Testing of putative cleavage of an HIV polyprotein-derived peptide substrate with the amino acid sequence VSFSFPQITL by hDdi2 RVP domain at pH 5.0 with 150 mM NaCl.

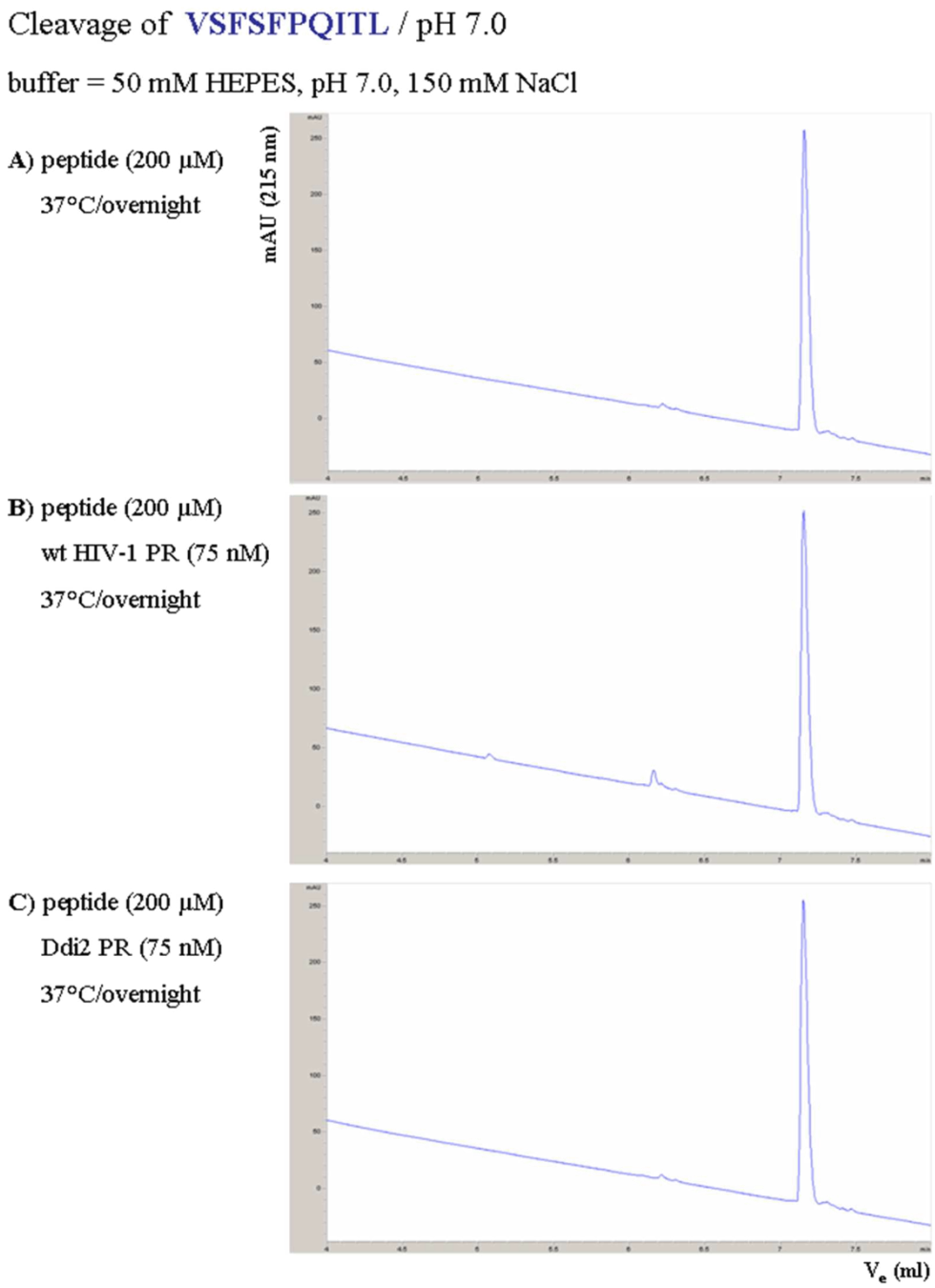


Figure S10: Testing of putative cleavage of an HIV polyprotein-derived peptide substrate with the amino acid sequence VSFSFPQITL by hDdi2 RVP domain at pH 7.0 with 150 mM NaCl.

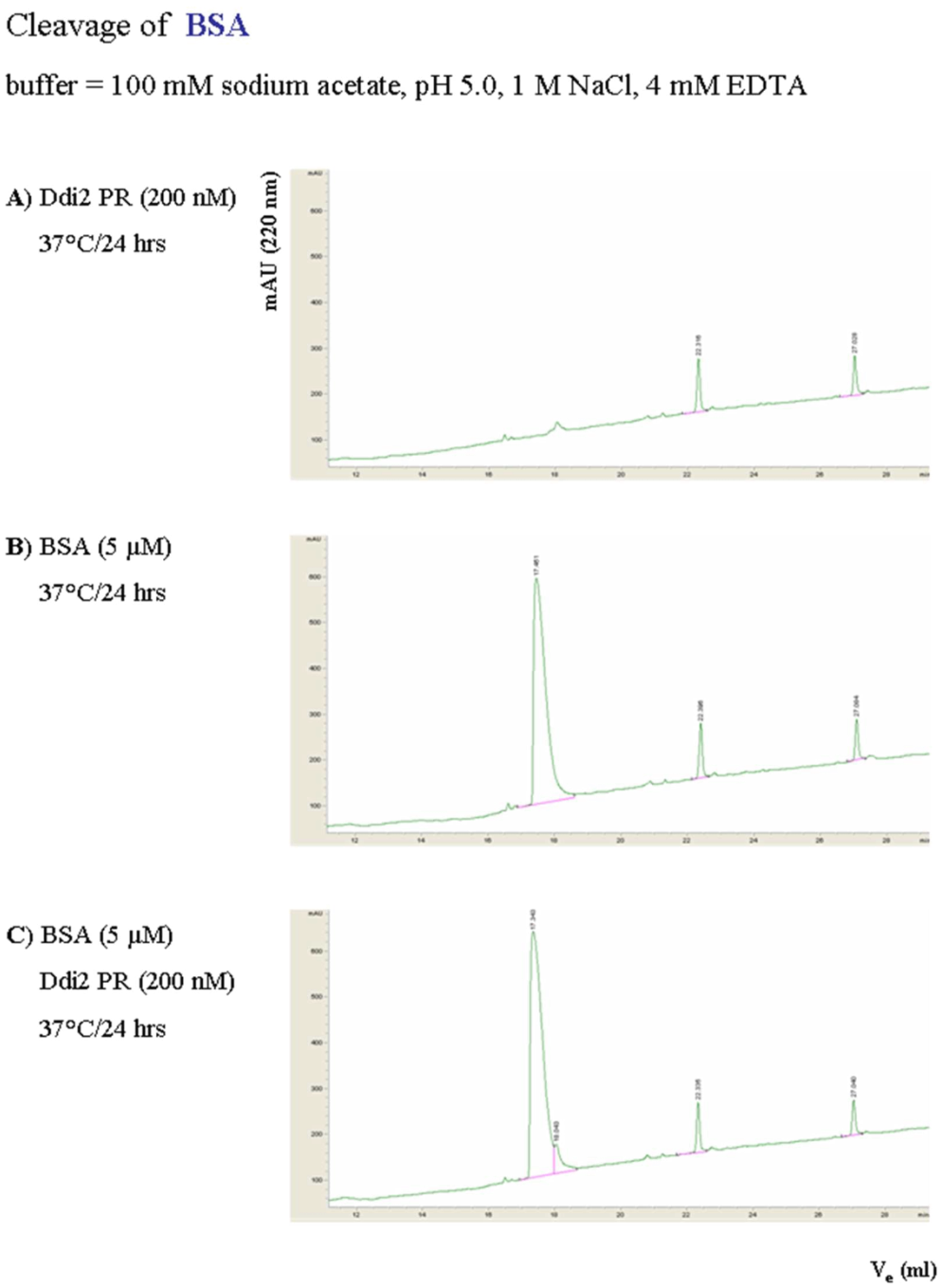


Figure S11: Testing of putative cleavage of bovine serum albumin (BSA) by hDdi2 RVP domain at pH 5.0 with 1 M NaCl.

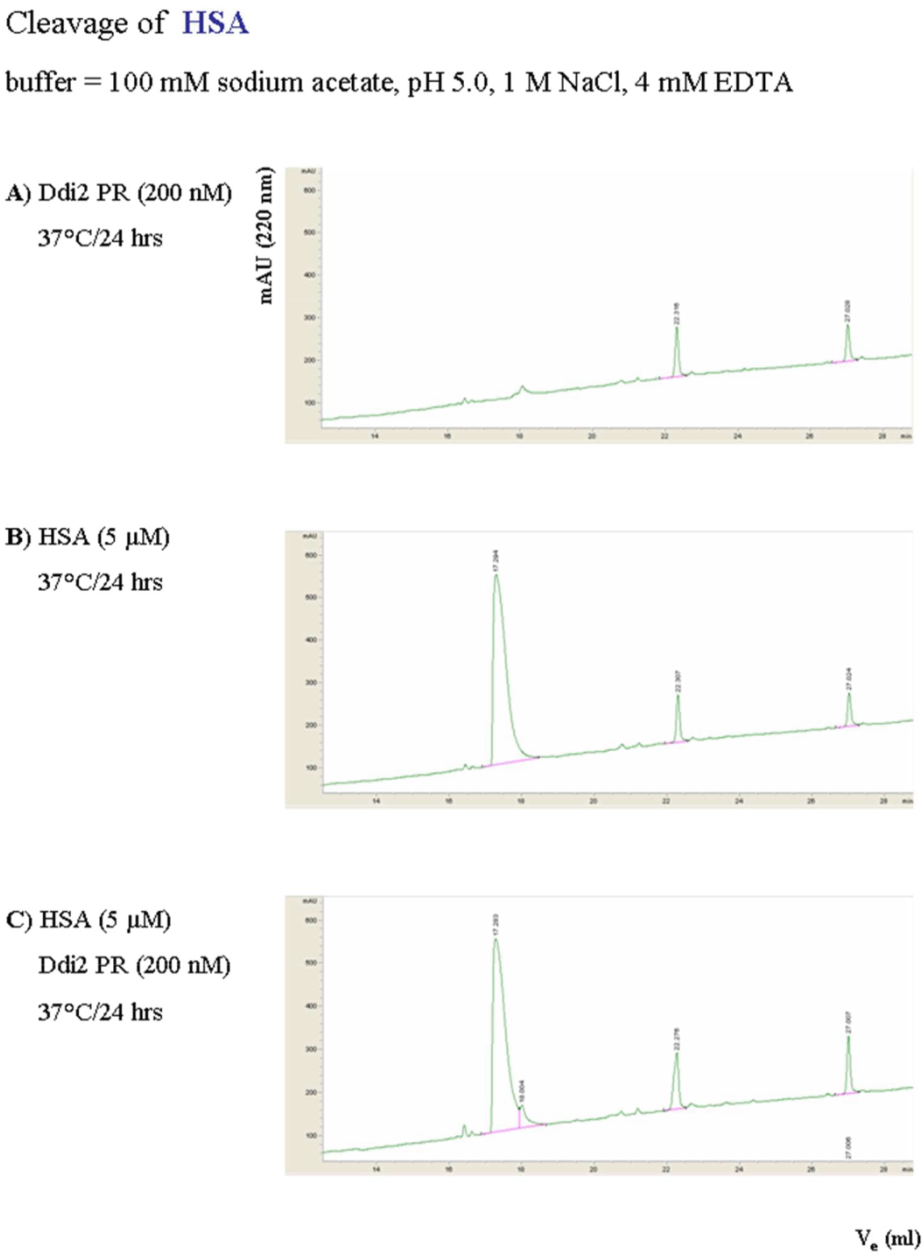


Figure S12: Testing of putative cleavage of human serum albumin (HSA) by hDdi2 RVP domain at pH 5.0 with 1 M NaCl.

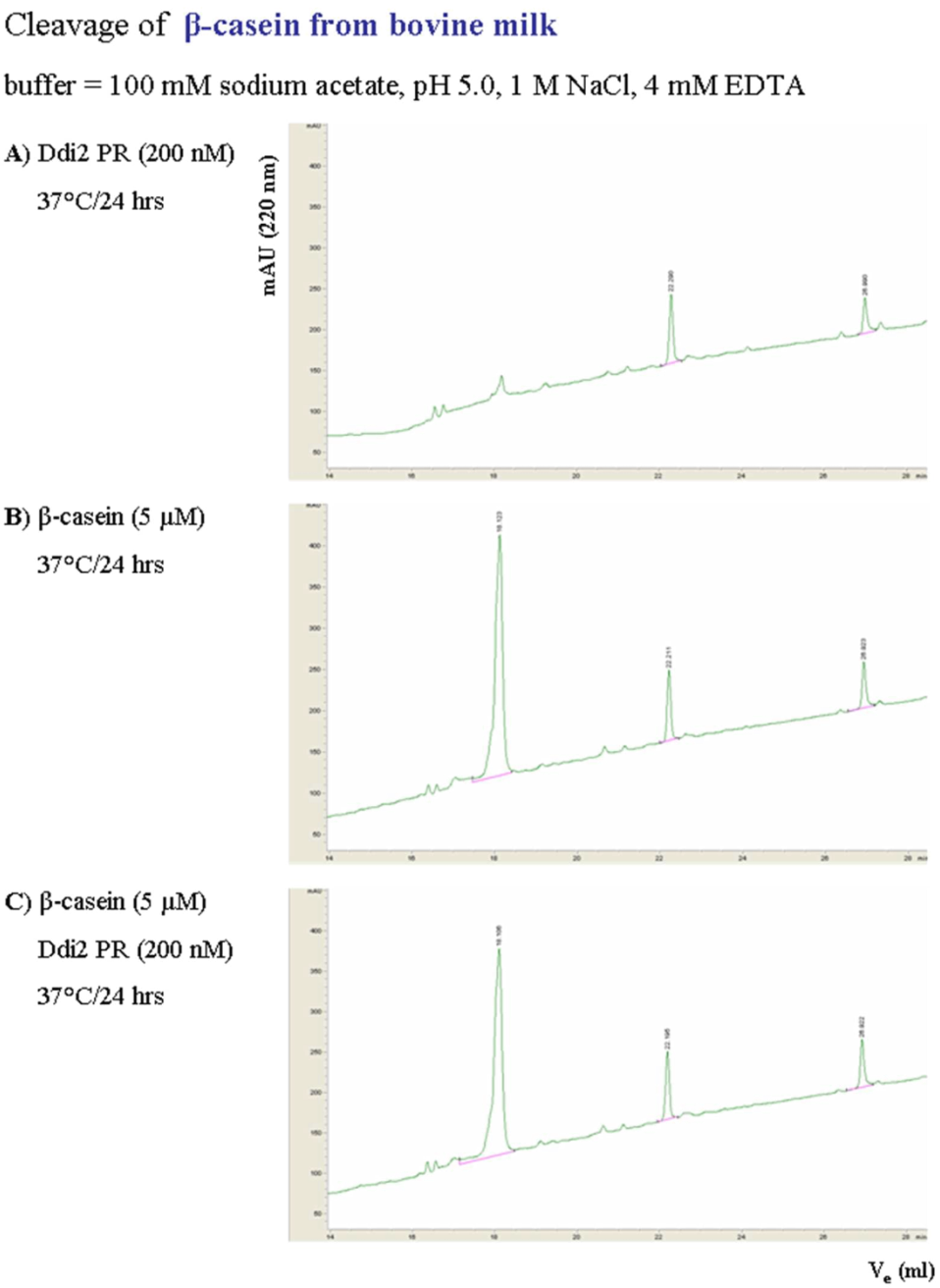


Figure S13: Testing of putative cleavage of β-casein from bovine milk by hDdi2 RVP domain at pH 5.0 with 1 M NaCl.

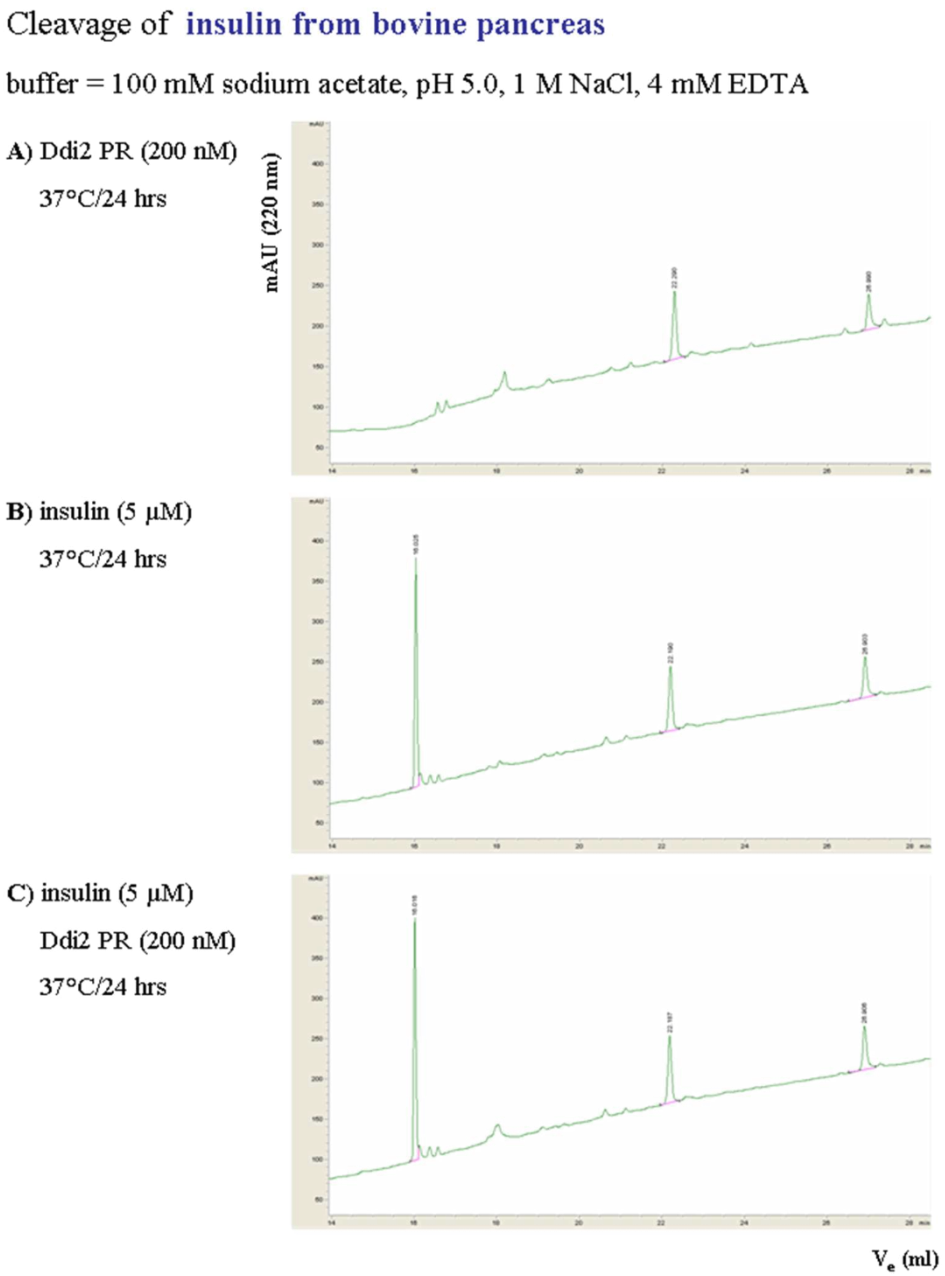


Figure S14: Testing of putative cleavage of insulin from bovine pancreas by hDdi2 RVP domain at pH 5.0 with 1 M NaCl.

Testing of putative binding of HIV protease inhibitors to the human Ddi2 RVP domain

Human Ddi2 RVP titrated with the HIV protease inhibitor darunavir[2] is used as an example.

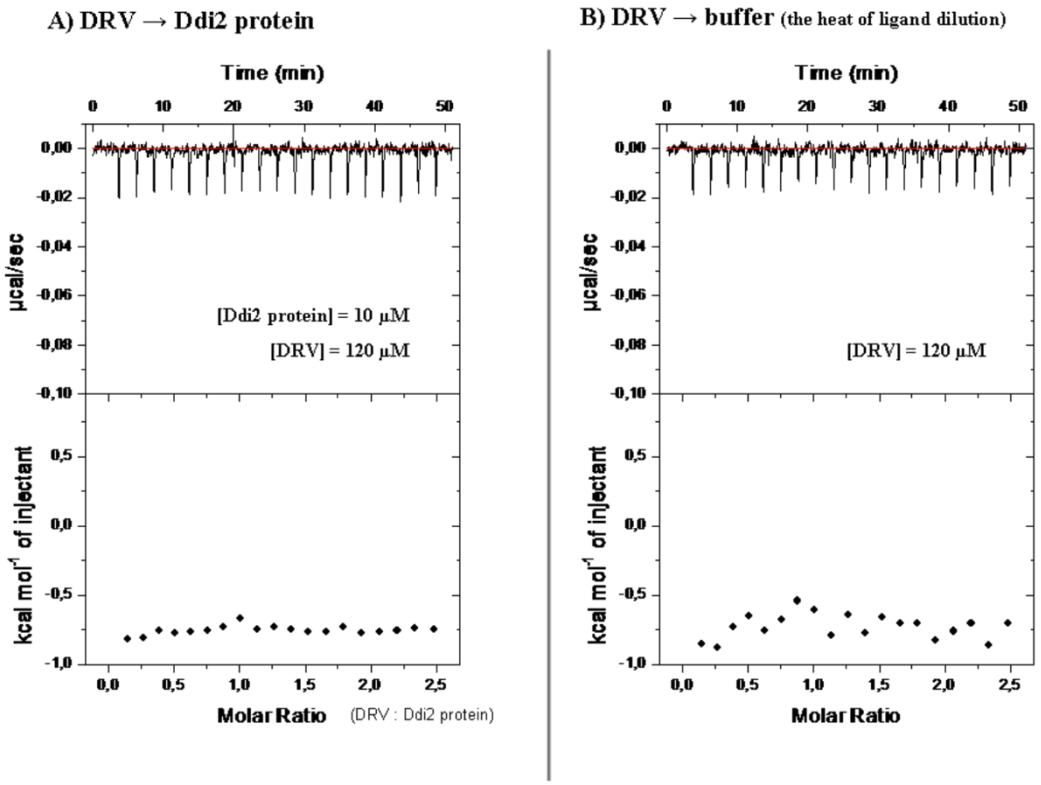


Figure S15: Testing of the putative interaction of full-length human Ddi2 with the HIV protease inhibitor darunavir. A) Calorimetric titration of darunavir with full-length hDdi2 was performed in 50 mM HEPES, pH 7.4, 150 mM NaCl at 25 ºC using an Auto-iTC$_{200}$ system (MicroCal, GE Healthcare Life Sciences). Typically, 2 µl aliquots of 120 µM HIV protease inhibitor were injected stepwise into a sample cell containing 200 µl of a 10 µM protease domain of hDdi2 protein (concentration calculated to dimer; HPLC amino acid analysis was performed). The titrations were monitored by MicroCal software implemented in Origin 7.0 (MicroCal, GE Healthcare Life Sciences). B) A control dilution experiment, in which darunavir was injected into buffer alone, was also performed.

Methods

Nuclear magnetic resonance spectroscopy

A series of double and triple resonance spectra were collected for hDdi2 UBL, HDD, and RVP full-C protein constructs and UBQ. NOEs identified in NOESY, $^{15}N/^{1}H$ NOESY-HSQC, and $^{13}C/^{1}H$ HSQC-NOESY spectra were acquired with an NOE mixing time of 120 ms. All TOCSY spectra were acquired with a mixing time of 60 ms. All spectra were processed using the program Sparky (Goddard and Kneller, UCSF).

Structural calculations

The family of converged structures for hDdi2 UBL was initially calculated using Cyana 2.1[4]. NOE-derived restraints from two-dimensional NOESY and three-dimensional ^{15}N- and ^{13}C-edited NOESY spectra were used to produce preliminary structures in Cyana. Assigned chemical shifts were used to generate backbone torsion angle constraints using the program TALOS+[5]. Subsequently, five cycles of simulated annealing combined with redundant dihedral angle constraints (REDAC) were performed[6]. A group of converged structures with no notable restraint violations (distance and Van der Waals violations <0.2 Å and dihedral angle constraint violation <5°) were further refined in explicit solvent using the YASARA software with the YASARA forcefield[7]. The 40 structures for hDdi2 UBL and 30 structures for hDdi2 HDD with the lowest energy were selected. Analysis of the family of structures was carried out using the Protein Structure Validation Software suite (www.nesg.org), Molmol[8], and PyMol (PyMOL™ 0.98) (www.pymol.org).

NMR chemical shift mapping

The assignments of triple resonance spectra of hDdi2 UBL, UBQ, and hDdi2 RVP full-C construct were performed to allow specific interaction of peptides and proteins with these proteins to be monitored by changes induced in the positions of backbone signals of ^{15}N-labeled proteins in ^{15}N/^{1}H HSQC spectra. The backbone assignment for ^{15}N/^{1}H HSQC spectra of Nedd8 was derived from a published solution structure with BMRB code 10062[9]. The most significant shifts of the backbone amide groups were used for to map the binding site. The minimal shift approach was used to assess the changes in case of one-step addition of the partner[10]. Weighted-average chemical shift perturbations were defined according to the formula ($\Delta\delta = \sqrt{(\Delta\delta_H)^2 + (\Delta\delta_N * \mathbf{0.2})^2}$). The titration curves were fitted with non-linear one site specific binding model in the program GraphPad Prism.

The mono UBQ - UIM peptide titration experiment was initialized by HSQC acquirement of 0.1 mM UBQ in 50 mM sodium phosphate buffer, pH 6.0, containing 3.9% DMSO. The conditions of the samples in the following titration steps remained identical with addition of UIM peptide to final concentrations of 0.69, 1.4, 2, 2.75, and 3.45 mM in individual samples. The cut-off for evaluation of the CSPs used for this experiment was 0.12. The 6 best-fitting curves corresponding to individual amino acids were used for calculation of the K_d. The control binding experiment was performed under the same conditions with 1.9 mM UIM scrambled peptide. The protein-protein interactions were monitored by acquiring ^{1}H/^{15}N HSQC spectra of free ^{15}N-labeled 0.1 mM UBQ in 50 mM sodium phosphate buffer, pH 7.4, 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 was studied by HSQC acquirement of 0.042 mM ^{15}N-labeled hDdi2 UBL with 1, 2, 4, 6, 8, and 10-fold molar addition of non-labeled bovine UBQ in 20 mM phosphate buffer, 0.5 mM DTT. The numerical cut-off for evaluation of the CSPs was 0.075. The 10 best-fitting curves corresponding to individual amino acids were used for calculation of the K_d. The reverse

experiment was carried out on 0.05 mM UBQ with 6-fold molar addition of hDdi2 ΔUIM protein construct in 50 mM phosphate buffer, 0.5 mM DTT. A control measurement was performed under the same conditions with 0.05 mM UBQ with 6-fold addition of Ddi2 HDD-RVP construct. The interactions of hDdi2 UBL with Ddi2 UIM peptide were studied by acquiring HSQC spectra of 0.05 mM protein in 50 mM sodium phosphate, pH 7.4, with addition of UIM peptide to a final concentration of 1.9 mM. A control experiment with one-step addition of scrambled peptide reaching 1.2 mM final concentration was performed. Verification of the Ddi2 UBL-UIM interaction was additionally performed on longer protein constructs by acquiring the HSQC spectra of 0.22 mM ^{15}N-labeled hDdi2 FL protein and HSQC spectra of 0.093 mM hDdi2 ΔUIM in 50 mM sodium phosphate buffer, pH 7.4, with 0.1 mM DTT. Other potential intramolecular interactions and UBL domain flexibility were characterized by superimposition of the HSQC spectra of full-length hDdi2 protein and HSQC spectra of 0.05 mM hDdi2 UBL domain. Nedd8 interaction with the two peptides was verified via an identical acquisition with 0.03 mM protein under two different conditions: 50 mM sodium acetate, pH 5, and 50 mM sodium phosphate, pH 7. The final addition of the UIM peptide at pH 5 was up to 1.2 mM. At pH 7, the final concentrations of the UIM peptide and the scrambled version were 1.9 and 1.2 mM, respectively. The interaction of hDdi2 UBL with Nedd8 was studied by 6-fold molar addition of Nedd8 to 0.05 mM UBL in 50 mM sodium phosphate buffer, pH 7.4, 0.5 mM DTT.

Table S1: NMR constraints and structural statistics for hDdi2 UBL and hDdi2 HDD

	hDdi2 UBL		hDdi2 HDD	
NMR distance & dihedral constraints				
Distance constraints				
Total NOE	1095		1894	
Intra-residue	287		487	
Inter-residue	808		1407	
Sequential (|i-j| = 1)	291		522	
Medium-range (|i-j| < 4)	170		593	
Long-range (|i-j > 5)	347		292	
Total dihedral angle restraints	132		108	
phi	66		54	
psi	66		54	
Structure statistics				
Violations (mean and s.d.)				
Max. dihedral angle violation (º)	5.00		5.10	
Max. distance constraint violation (Å)	0.35		0.49	
Deviations from idealized geometry				
Bond lengths (Å)	0.011		0.012	
Bond angles (º)	1.4		1.5	
Ramachandran plot summary				
Most favoured regions	91.2%		93.9%	
Additionally allowed regions	8.6%		6.0%	
Generously allowed regions	0.2%		0.1%	
Disallowed regions	0.0%		0.0%	
Average pairwise r.m.s.d. foe* (Å)	*Ordered*[a]	*all residues*	*Orderea*[b]	*all residues*
Heavy	1.0	4.8	4.8	8.1
Backbone	0.4	4.3	4.5	7.8

* 40 hDdi2 UBL and 30 hDdi2 HDD structures were used in r.m.s.d. calculations

[a] ordered residue range – residues with defined dihedral angle order parameters. S(phi)+S(psi)>=1.8

Table S2: X-ray diffraction data processing and refinement statistics for hDdi2 212-360.

Data collection	
Wavelength (Å)	0.91573
Space group	$P2_12_12$
Unit cell dimensions (Å)	66.87 86.40 52.12
(°)	90.0 90.0 90.0
Mosaicity (°)	0.9
Images	235
Oscillation angle (°)	0.5
Resolution (Å)	26.7 – 1.9 (2.0 – 1.9)
Unique reflections	21198 (2439)
Completeness (%)	87.2 (69.7) *
Multiplicity	4.6 (4.7)
R_{meas}[a] <I>/<σ(I)>	0.113 (0.422) 9.5 (3.6)
Wilson B ($Å^2$)[b]	27.6
Refinement	
Resolution (Å)	33.7 – 1.9 (1.94 -1.90)
R_{work} (%)	20.8 (39.5)
R_{free} (%)	25.6 (45.7)
Rms deviations	
Bond length (Å)	0.014
Bond angle (°)	1.75
Protein atoms	2033
Solvent atoms	86
Average B-factor[b]	26.0
Most/additional favoured regions f-y space (%)	99.2/1.2
Coordinate error (Å) [c]	0.17/0.11 (based on R value/Maxim Likelihood)

[a] R_{meas} defined in ref.[11]. [b] Wilson B by Sfcheck program[12] from CCP4 suite[13]. [c] Calculated by Refmac5[14].

* Ice ring areas were excluded from integration.

References

1 Walters, K. J., Lech, P. J., Goh, A. M., Wang, Q. H. & Howley, P. M. DNA-repair protein hHR23a alters its protein structure upon binding proteasomal subunit S5a. *P Natl Acad Sci USA* **100**, 12694-12699, (2003).

2 Koh, Y. *et al.* Novel bis-tetrahydrofuranylurethane-containin nonpeptidic protease inhibitor (PI) UIC-94017 (TMC114) with potent activity against multi-PI-resistant human immunodeficiency virus in vitro. *Antimicrob Agents Ch* **47**, 3123-3129, (2003).

3 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).

4 Herrmann, T., Guntert, P. & Wuthrich, K. Protein NMR structure determination with automated NOE-identification in the NOESY spectra using the new software ATNOS. *J Biomol Nmr* **24**, 171-189, (2002).

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

6 Guntert, P. *et al.* Structure Determination of the Antp(C39-]S) Homeodomain from Nuclear-Magnetic-Resonance Data in Solution Using a Novel Strategy for the Structure Calculation with the Programs Diana, Caliba, Habas and Glomsa. *Journal of Molecular Biology* **217**, 531-540, (1991).

7 Harjes, E. *et al.* GTP-Ras disrupts the intramolecular complex of C1 and RA domains of Nore1. *Structure* **14**, 881-888, (2006).

8 Koradi, R., Billeter, M. & Wuthrich, K. MOLMOL: A program for display and analysis of macromolecular structures. *J Mol Graphics* **14**, 51-&, (1996).

9 Sakata, E. *et al.* Direct interactions between NEDD8 and ubiquitin E2 conjugating enzymes upregulate cullin-based E3 ligase activity. *Nat Struct Mol Biol* **14**, 167-168, (2007).

10 Farmer, B. T. Localizing the NADP(+) binding site on the MurB enzyme by NMR. *Nat Struct Biol* **3**, 995-997, (1996).

11 Diederichs, K. & Karplus, P. A. Improved R-factors for diffraction data analysis in macromolecular crystallography. *Nat Struct Biol* **4**, 269-275, (1997).

12 Vaguine, A. A., Richelle, J. & Wodak, S. J. SFCHECK: a unified set of procedures for evaluating the quality of macromolecular structure-factor data and their agreement with the atomic model. *Acta Crystallogr D* **55**, 191-205, (1999).

13 Winn, M. D. *et al.* Overview of the CCP4 suite and current developments. *Acta Crystallogr D* **67**, 235-242, (2011).

14 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).

8.2. SUPPLEMENT S2: STRUCTURAL STUDIES OF THE YEAST DNA DAMAGE-INDUCIBLE PROTEIN DDI1 REVEAL DOMAIN ARCHITECTURE OF THIS EUKARYOTIC PROTEIN FAMILY

Trempe J.F., Grantz Šašková K., Sivá M., Ratcliffe C.D.H., Veverka V., Hoegl A., Menade M., Feng X., Shenker S., **Svoboda M.,** Kožíšek M., Konvalinka J., and Gehring K. Structural studies of the yeast DNA damage-inducible protein Ddi1 reveal domain architecture of this eukaryotic protein family. *Sci Rep.* **6**, 33671 (2016).

My contribution: cloning of ΔUBA and catalytically inactive D220A variants of yeast Ddi1p; expression and purification of full-length yeast Ddi1p protein, it's RVP domain and catalytically inactive variants of both; expression and purification of ΔUBA variant of yeast Ddi1p; sample preparation and assistance with ITC measurement; K48-linked diubiquitin synthesis; PICS proteomics screen for substrates of Ddi1p protease activity.

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 $\Delta ddi1\Delta rad23\Delta dsk2$ 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.

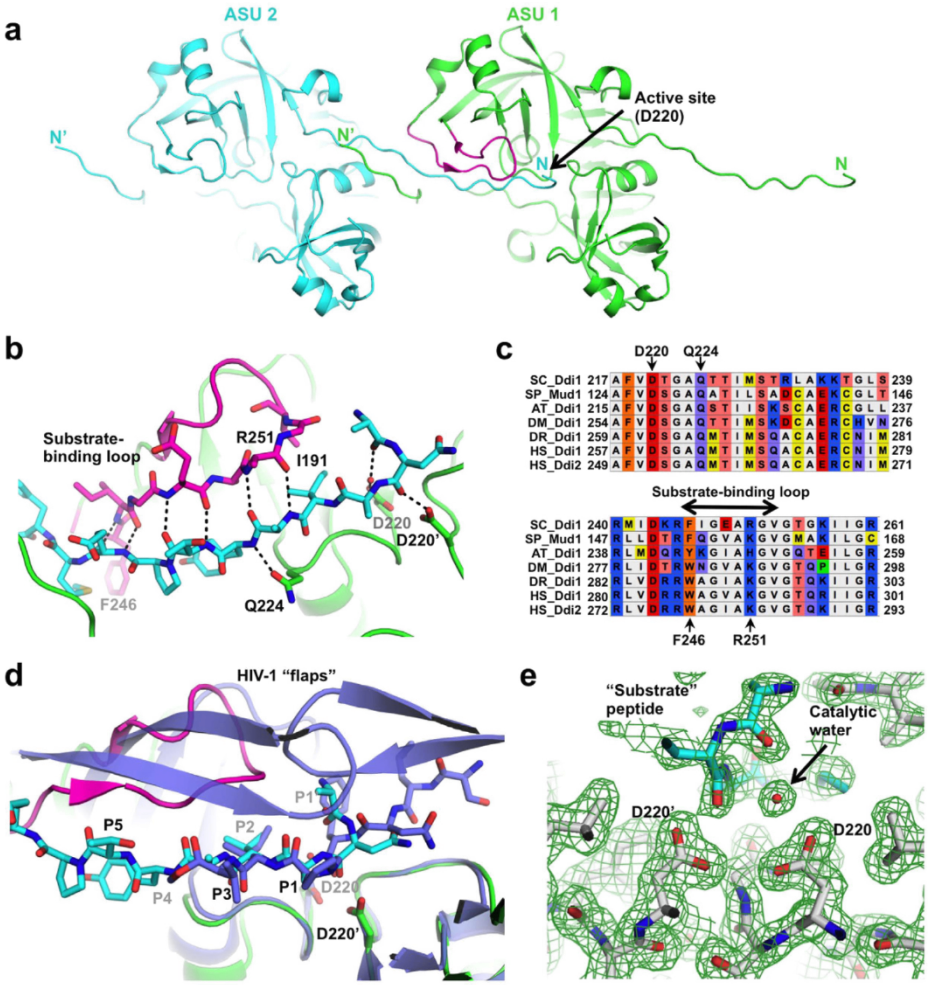


Figure 1. Crystal structure of the yeast Ddi1 protease domain reveals a potential substrate-binding mode. (**a**) Cartoon representation of two adjacent asymmetric units (ASU), showing the N-terminus of one molecule in ASU #2 (cyan) binding to the active site of a dimeric protease in ASU #1 (green). 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

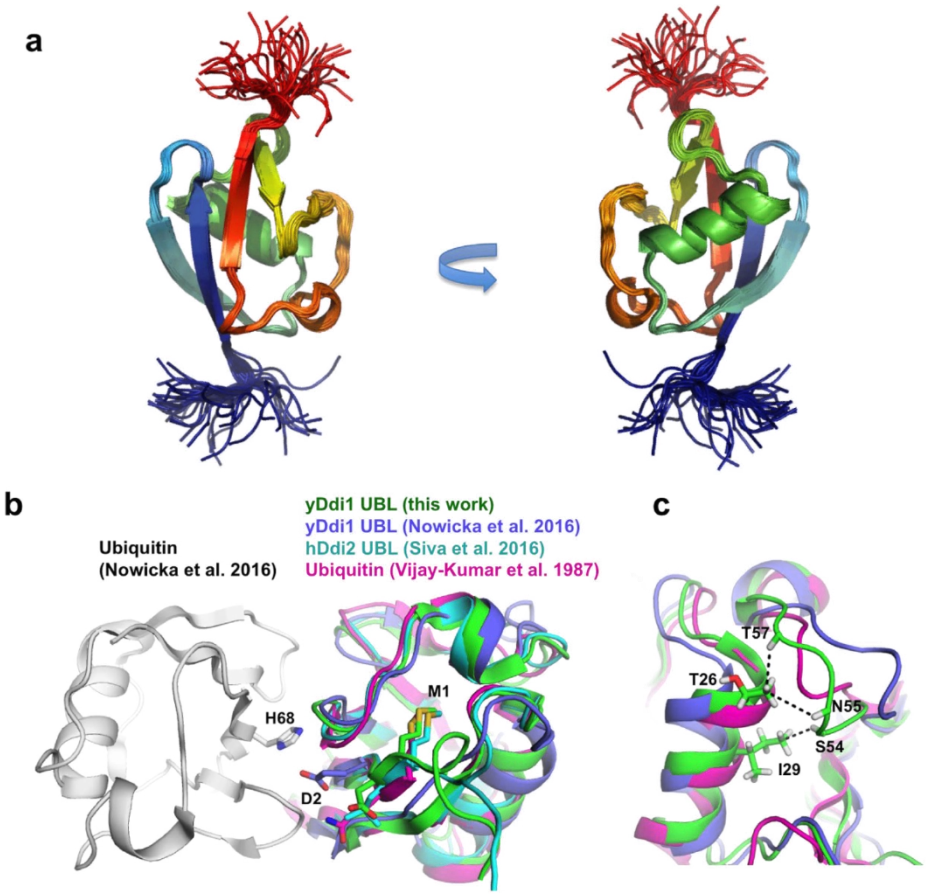


Figure 2. Solution structure of the Ddi1 UBL. (**a**) NMR solution structure of the UBL domain (a.a. 1–80; PDB 2N7E). An ensemble of 43 models is shown in cartoon representation, colored from blue to red from the N- to C-terminus. (**b**) Superposition of the yeast Ddi1 UBL NMR structure (green, pdb 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

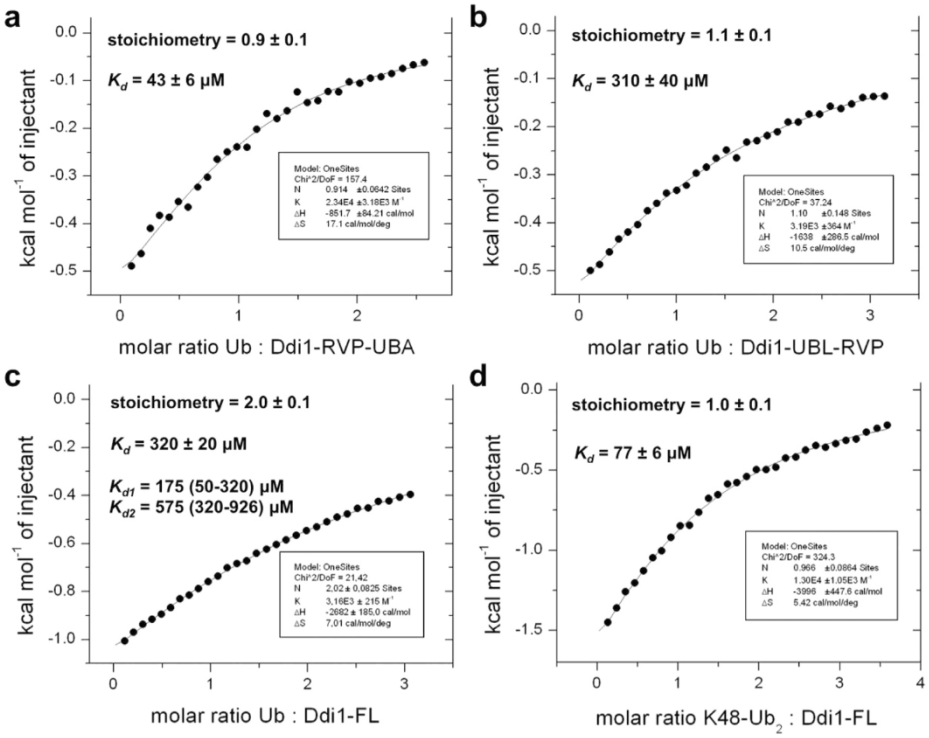


Figure 3. The UBL and UBA domains of Ddi1 bind to ubiquitin and diubiquitin. Ubiquitin or K48-linked diubiquitin (Ub_2) were added by syringe to Ddi1 constructs in the sample cell. The different titrations were (**a**) ubiquitin to RVP-UBA, (**b**) ubiquitin to UBL-RVP, (**c**) ubiquitin to full-length 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

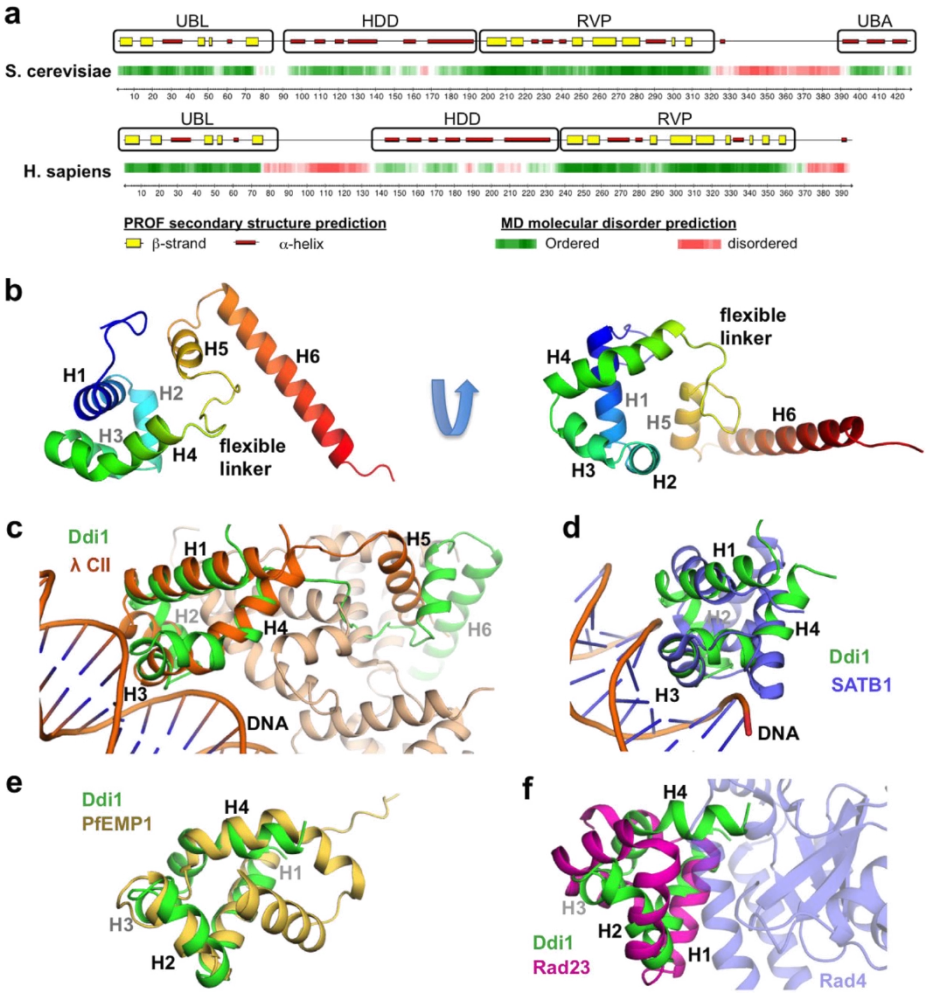


Figure 4. Ddi1 contains two helical domains preceding the protease domain. (**a**) Secondary structure and molecular disorder predictions of the yeast (top) and human (bottom) Ddi1 proteins. Predictions were performed on the PredictProtein server (http://www.predictprotein.org). (**b**) Cartoon of a 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]),

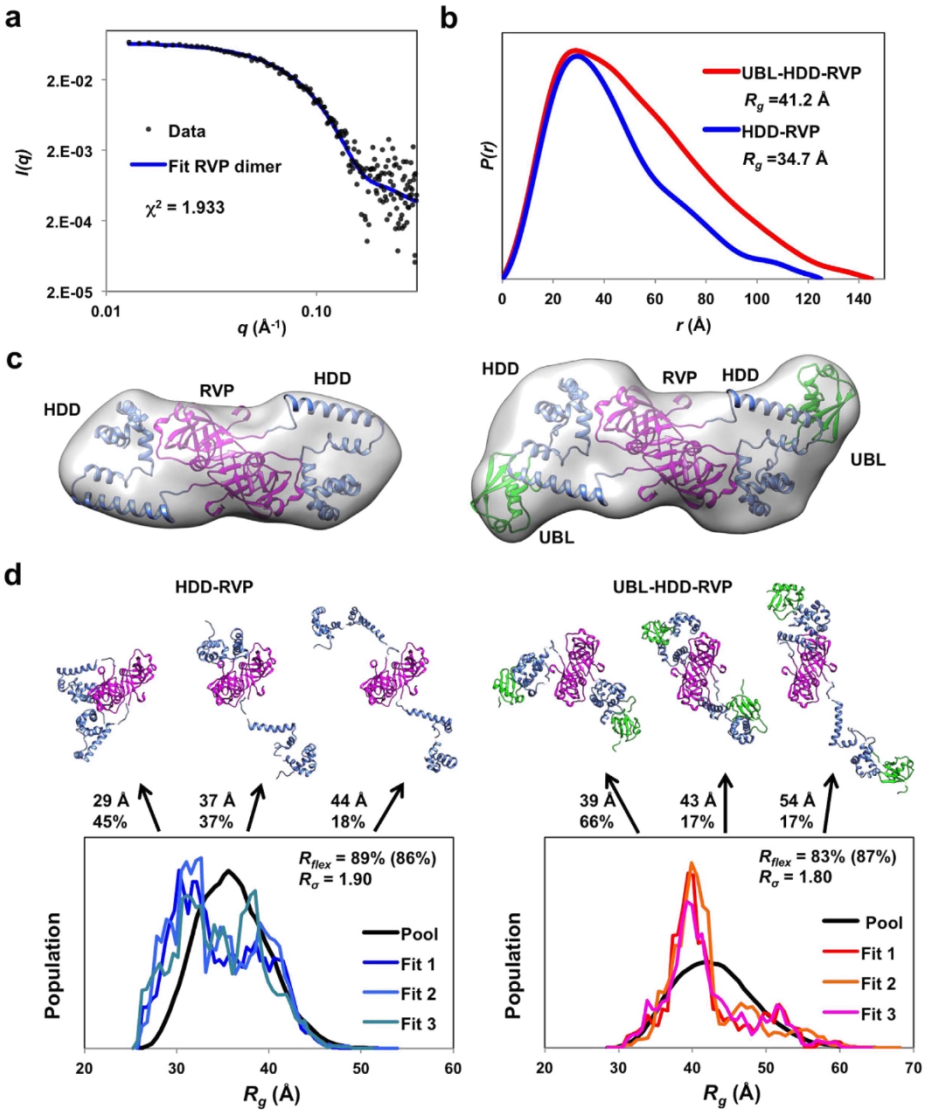


Figure 5. SAXS analysis of the Ddi1 dimer in solution. (**a**) SAXS data (circles) and calculated scattering curves (blue line) derived from the dimeric Ddi1 RVP domain crystal structure (200–325), displayed as a double-logarithmic plot. (**b**) SAXS pair-distance distribution functions for Ddi1 HDD-RVP (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 Stch. *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).

Supplementary information

Structural studies of yeast DNA damage-inducible protein Ddi1 reveal domain architecture of this eukaryotic protein family

Authors: Jean-François Trempe[1,2,*], Klára Grantz Šašková[3,4], Monika Sivá[3,4,5], Colin Ratcliffe[1], Václav Veverka[3], Annabelle Hoegl[1], Marie Ménade[1], Xin Feng[1], Solomon Shenker[1], Michal Svoboda[3,6], Milan Kožíšek[3], Jan Konvalinka[3,4], Kalle Gehring[1]

[1]Groupe de Recherche Axé sur la Structure des Protéines, Department of Biochemistry, McGill University, Montreal, QC, H3G 1Y6, Canada

[2]Current affiliation: Department of Pharmacology & Therapeutics, McGill University

[3]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

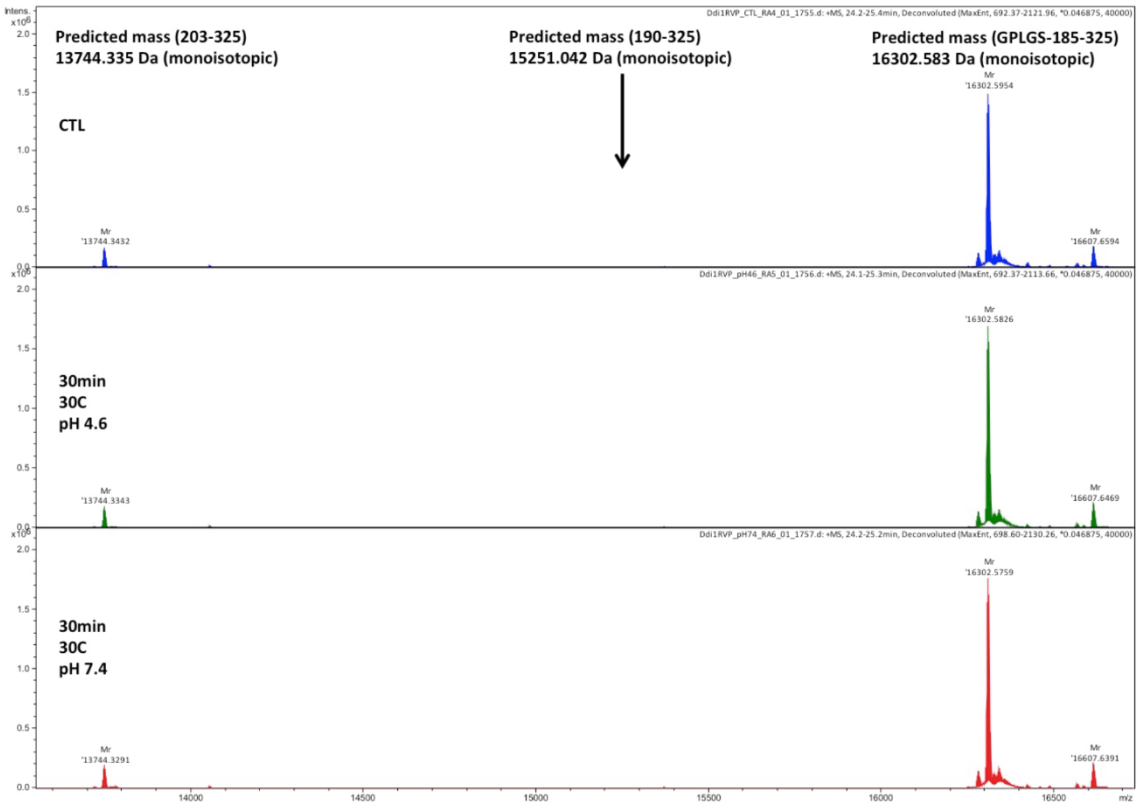
[4]Department of Biochemistry, Faculty of Science, Charles University, Hlavova 8, 120 00 Prague 2, Czech Republic

[5]First Faculty of Medicine, Charles University in Prague, Katerinska 32, 121 08, Prague 2, Czech Republic

[6]Department of Physical and Macromolecular Chemistry, Faculty of Science, Charles University, Hlavova 8, 120 00 Prague 2, Czech Republic

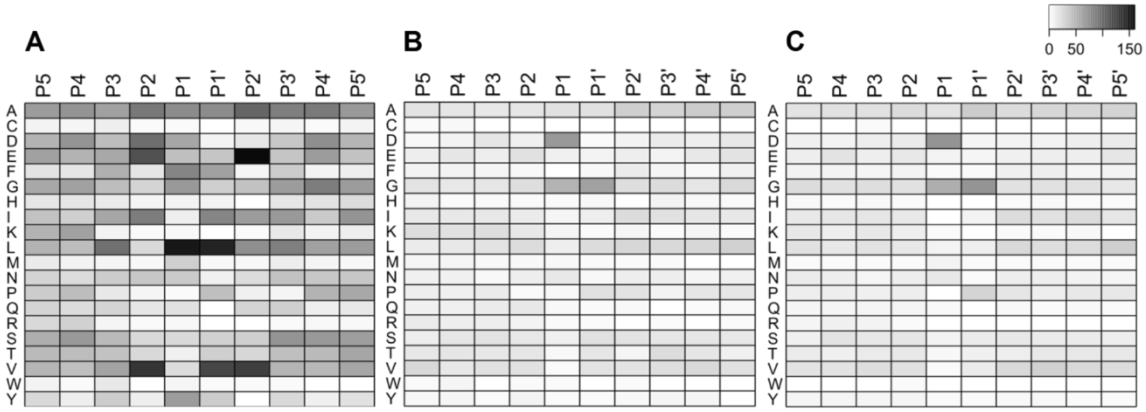
*Correspondence: jeanfrancois.trempe@mcgill.ca

Keywords: Ddi1, protease, SAXS, Sti1, UBL, UBA, ubiquitin, Lys48, crystal structure



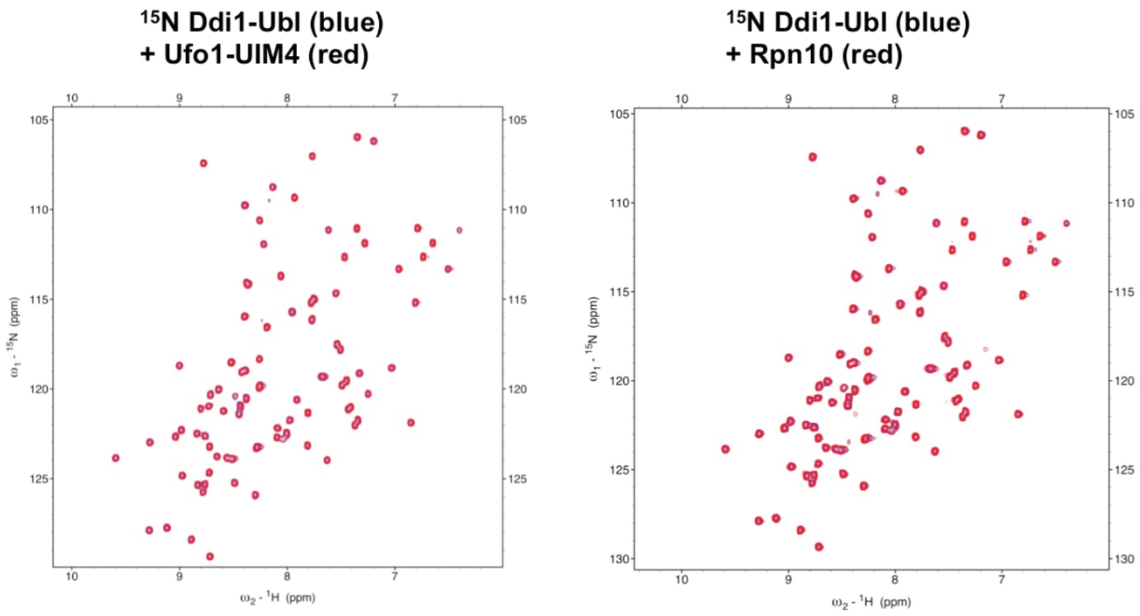
Supplementary Figure S1. The RVP domain of Ddi1 does not auto-proteolyze under neutral or acidic pH. Ddi1 (185-325) protein stock used for crystallization was diluted to 1 mg/mL in 50 mM sodium acetate pH 4.6 or 50 mM HEPES pH 7.4 and incubated 30 min at 30°C. The sample was diluted at 0.1 mg/mL in 0.05% TFA/2% acetonitrile and 20 µL was injected on a Dionex C4 Acclaim 1.0/15mm column followed by a 10 min 4-50% gradient of acetonitrile in 0.1% formic acid, with a flow rate of 40 µL/min. The eluate was analyzed on a Bruker Impact II Q-TOF mass spectrometer equipped with an Apollo II ion funnel ESI source. The multiply charged ions were deconvoluted at 40,000 resolution to yield the isotopically-resolved mass spectra shown above. Peak assignment was performed using the SNAP algorithm, which takes into account natural abundance isotope distribution to assign monoisotopic mass. The main peak in all spectra is within 0.01 Da of the predicted mass. The cleavage product predicted from the crystal structure (190-325, 15251 Da) is absent in all conditions, suggesting it is a crystallization artefact. The contaminant peak observed at 13744 Da corresponds to residues 203-325, where the entire flexible N-terminus has been digested by contaminant proteases. The peak at 16607 Da is an adduct of glutathione (306 Da), used as an eluent of the precursor GST-fusion protein.

2



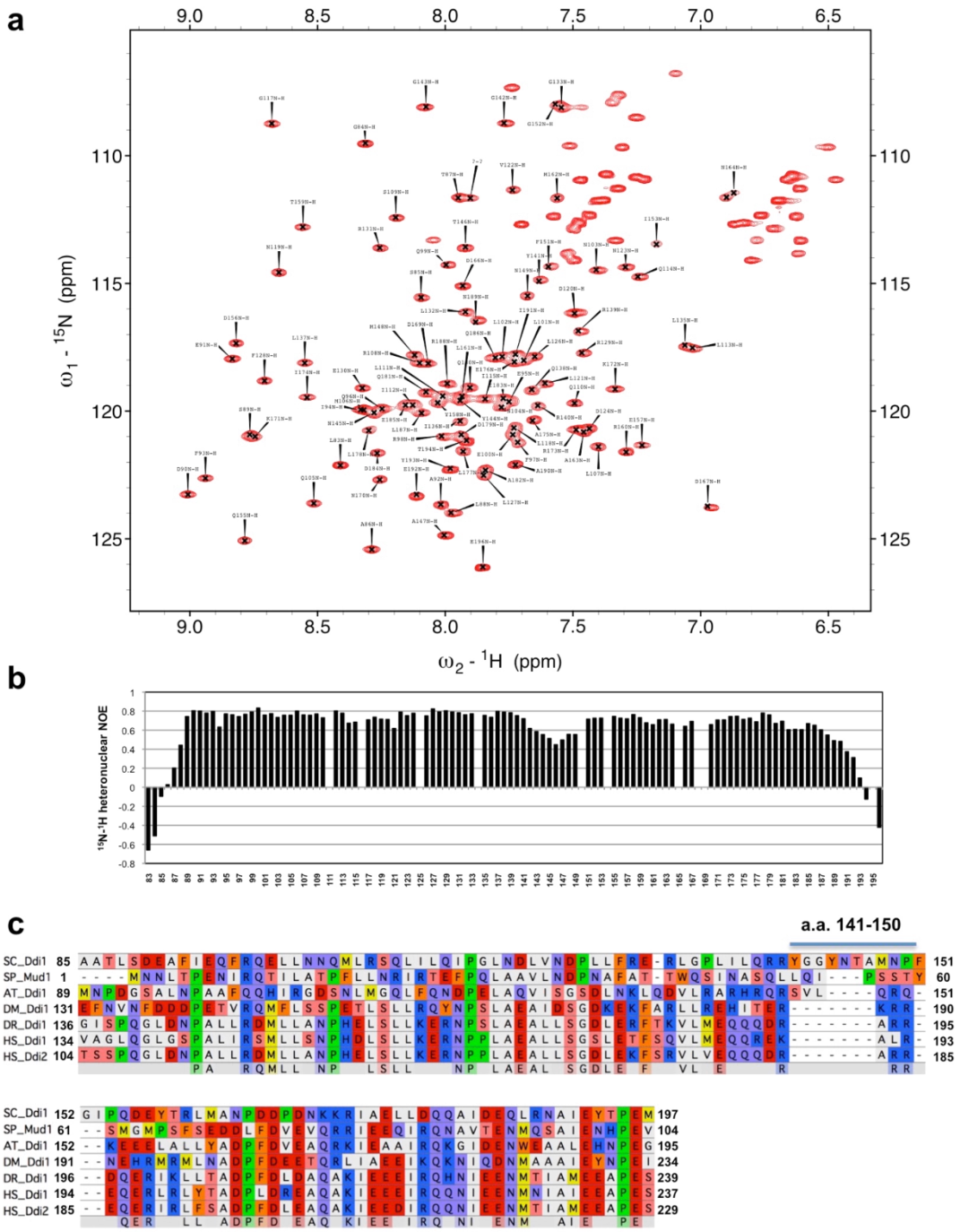
Supplementary Figure S2: PICS (Proteomic Identification of protease Substrates) experiment results. Heat map representation of protease substrate profile demonstrating total counts of given amino acid (described in single letter code) per particular position (P5–P5´) in peptidic substrate on *Saccharomyces cerevisiae* derived peptide library. (**A)** library cleaved by HIV-1 protease (positive control) (**B)** Library cleaved by yeast Ddi1 protein in acetate pH = 4.0 (proteolytic activity assay) (**C)** background profile – uncleaved library (negative control). Color key common for all three heat maps.

3

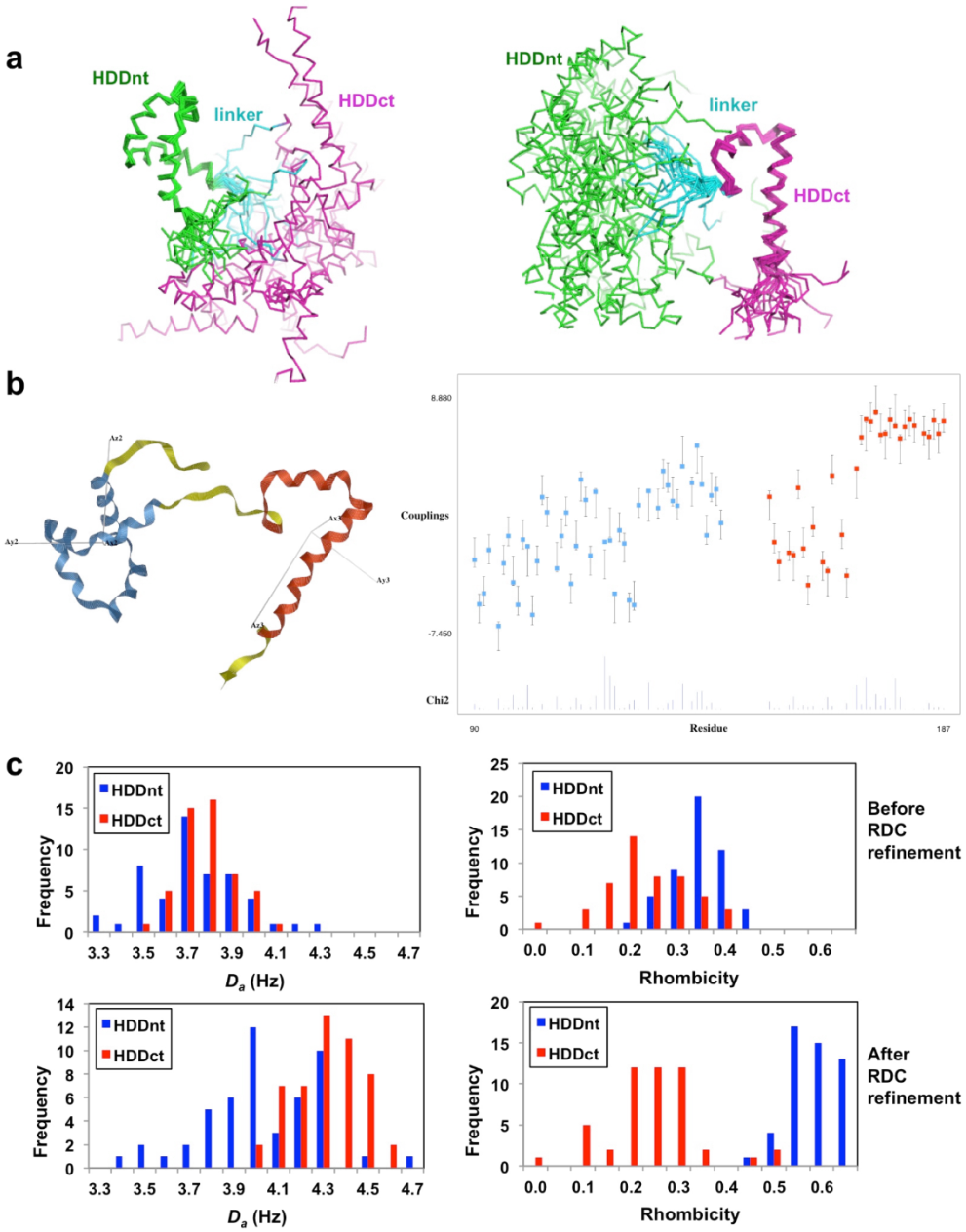


Supplementary Figure S3. The UBL domain of Ddi1 does not bind ubiquitin-interacting motifs. ^{15}N-^{1}H HSQC spectra of the Ddi1 UBL in the presence of Ufo1-UIM4 (left) and Rpn10 (right). The absence of chemical shift perturbation indicates no interaction.

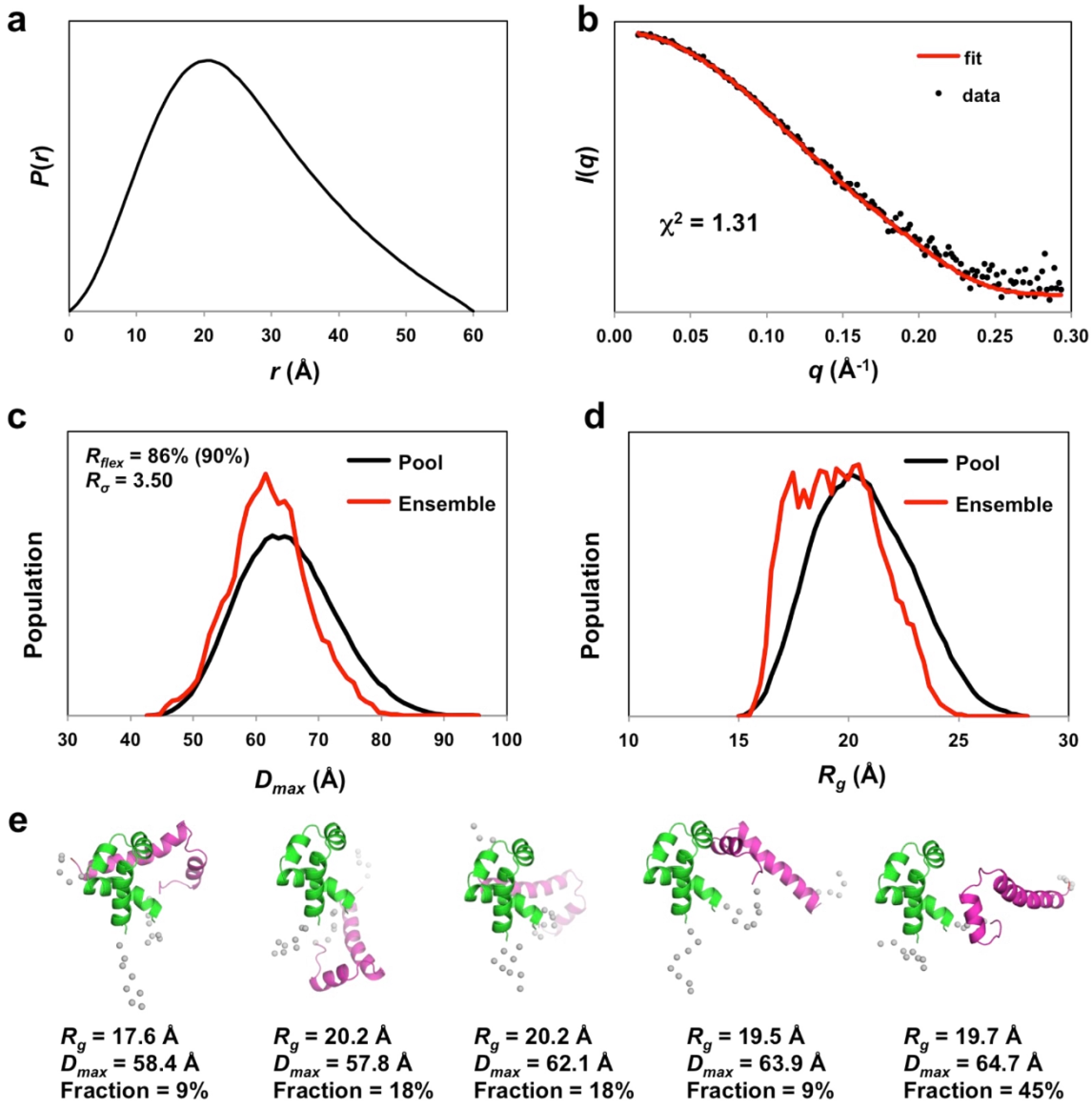
4



Supplementary Figure S4. NMR characterization of the Ddi1 HDD domain. (a) ^{1}H-^{15}N HSQC spectrum of Ddi1 86-196 (HDD), showing a good dispersion of signals in the proton dimension. **(b)** ^{1}H-^{15}N heteronuclear NOEs as a function of residue number. **(c)** Sequence alignment of Ddi1 HDD from different species. SC, *Saccharomyces cerevisiae*; SP, *Schizosaccharomyces pombe*; AT, *Arabidopsis thaliana*; DM, *Drosophila melanogaster*; DR, *Danio rerio*; HS, *Homo sapiens*. The position of the linker tethering the N- and C-terminal domains is indicated.



Supplementary Figure S5. Structure of the N-terminal and C-terminal domains of yeast Ddi1 HDD. (a) Ensemble of the solution structure of the yeast Ddi1 HDD**,** superposed on the N-terminal domain (left, green) or C-terminal domain (right, magenta). The linker is shown in cyan. **(b)** ^{15}N-^{1}H RDC tensor analysis of the HDD structure using MODULE. Left, orientation of the order tensors for the two domains of HDD. Right, back-calculated RDC values (blue and red dots) compared with the experimental data (vertical error bars, each 1.0 Hz). The bars at the bottom indicate the Chi^2 deviation for each coupling. **(c)** Monte Carlo simulation to determine the uncertainty on the magnitude (D_a) and rhombicity of the RDC alignment tensors for the two domains. 50 simulations were performed using an uncertainty of ± 0.5 Hz, using the lowest-energy structure calculated before (top) and after (bottom) refinement with RDC.



f

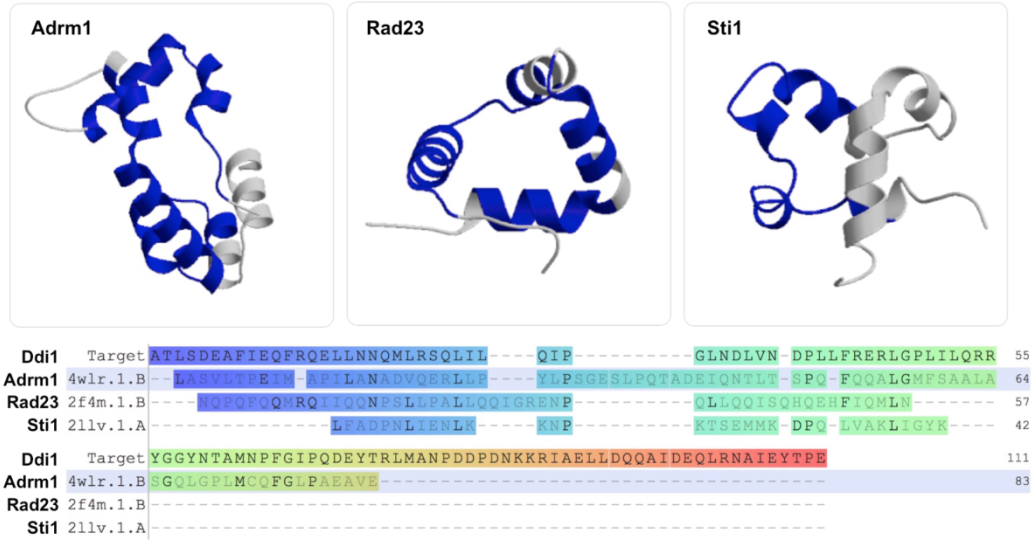
R_g (SAXS)	R_h (QELS)	M.wt (SAXS)	M.wt. (MALS)	Calc. M.wt
19.7 ± 0.2 Å	23 ± 2 Å	17 ± 2 kDa	13 ± 1 kDa	13.7 kDa

Supplementary Figure S6. SAXS analysis of the Ddi1 HDD domain. (a) SAXS pair-distance distribution function for the HDD domain. The apparent D_{max} is 60 Å. **(b)** Fit of ensemble of EOM-generated structures to experimental data. **(c)** Population distribution of D_{max} for a pool of random structures of the N-terminal and C-terminal domains of HDD (black), compared with an ensemble that best fit the experimental data (red). The ensemble and pool averages are 62.1 and 64.9 Å, respectively. **(d)** Same as in (c), except for the radius of gyration R_g. The ensemble and pool averages are 19.5 and 20.7 Å, respectively. **(e)** Models from the best ensemble solution from the EOM calculation, with R_g, D_{max} and fraction for each model indicated. **(f)** Summary of light (R_h, M.Wt) and small-angle X-ray (Rg, M.Wt) scattering parameters for Ddi1 HDD.

a

##	Scoring			RMSD	N_{align}	N_g	%seq	Query	Target (PDB entry)				
	Q	P	Z					%sse	Match	%sse	N_{res}	×	Title
1	0.31	0.3	3.6	2.22	45	3	11	60	2lkl:A	60	81	☑	STRUCTURE OF THE CORE INTRACELLULAR DOMAIN OF PFEMP1
2	0.29	0.0	1.8	3.64	50	4	16	60	2mqk:A	75	65	☑	SOLUTION STRUCTURE OF N TERMINAL DOMAIN OF THE MUB AAA+ ATPASE
3	0.25	0.1	2.9	2.84	45	2	2	60	1zs4:A	60	82	☑	STRUCTURE OF BACTERIOPHAGE LAMBDA CII PROTEIN IN COMPLEX WITH DNA
4	0.24	0.1	2.7	2.98	45	3	2	60	1zs4:C	60	79	☑	STRUCTURE OF BACTERIOPHAGE LAMBDA CII PROTEIN IN COMPLEX WITH DNA
5	0.24	0.0	0.9	3.86	49	7	12	60	1pou:A	75	71	☑	THE SOLUTION STRUCTURE OF THE OCT-1 POU-SPECIFIC DOMAIN REVEALS A STRIKING SIMILARITY TO THE BACTERIOPHAGE LAMBDA REPRESSOR DNA-BINDING DOMAIN
6	0.23	0.0	1.7	3.70	49	6	6	60	2lvk:A	60	77	☑	SOLUTION STRUCTURE OF CA-BOUND PHL P 7
7	0.23	0.0	2.1	3.45	45	4	9	60	2eo2:A	100	71	☑	SOLUTION STRUCTURE OF THE INSERTION REGION (510-573) OF FTHFS DOMAIN FROM MOUSE METHYLENETETRAHYDROFOLATE DEHYDROGENASE (NADP+ DEPENDENT) 1-LIKE PROTEIN
8	0.23	0.0	2.6	3.63	50	7	8	80	1wmg:B	67	83	☑	CRYSTAL STRUCTURE OF THE UNC5H2 DEATH DOMAIN
9	0.23	0.0	1.5	3.27	45	3	2	60	2avu:C	60	76	☑	STRUCTURE OF THE ESCHERICHIA COLI FLHDC COMPLEX, A PROKARYOTIC HETEROMERIC REGULATOR OF TRANSCRIPTION
10	0.23	0.0	1.3	4.31	51	4	20	60	4yg7:C	60	71	☑	STRUCTURE OF FL AUTOREPRESSION PROMOTER COMPLEX
11	0.22	0.0	1.7	3.68	50	5	8	60	2o4a:A	60	84	☑	CRYSTAL STRUCTURE OF THE N-TERMINAL CUT DOMAIN OF SATB1 BOUND TO MATRIX ATTACHMENT REGION DNA
12	0.22	0.0	2.5	2.92	47	6	2	60	2e6w:A	60	96	☑	SOLUTION STRUCTURE AND CALCIUM BINDING PROPERTIES OF EF- HANDS 3 AND 4 OF CALSENILIN
13	0.22	0.0	1.1	4.11	49	4	20	60	2wiu:D	60	71	☑	MERCURY-MODIFIED BACTERIAL PERSISTENCE REGULATOR HIPBA
14	0.22	0.0	1.7	3.65	48	4	17	60	2vln:A	60	80	☑	N75A MUTANT OF E9 DNASE DOMAIN IN COMPLEX WITH IM9
15	0.22	0.0	1.9	3.58	48	4	17	60	2gzf:A	60	82	☑	CRYSTAL STRUCTURE OF THE E9 DNASE DOMAIN WITH A MUTANT IMMUNITY PROTEIN IM9 (Y54F)
16	0.22	0.0	1.9	3.59	48	4	17	60	2gze:A	60	82	☑	CRYSTAL STRUCTURE OF THE E9 DNASE DOMAIN WITH A MUTANT IMMUNITY PROTEIN IM9 (Y55A)
17	0.22	0.0	1.1	4.17	49	4	20	60	3dnv:B	75	71	☑	MDT PROTEIN
18	0.22	0.0	1.8	3.61	48	4	17	60	2vlp:A	60	82	☑	R54A MUTANT OF E9 DNASE DOMAIN IN COMPLEX WITH IM9
19	0.22	0.0	1.9	3.58	48	4	15	60	2gzj:E	60	83	☑	CRYSTAL STRUCTURE OF THE E9 DNASE DOMAIN WITH A MUTANT IMMUNITY PROTEIN IM9 (D51A)
20	0.22	0.0	1.9	3.58	48	4	15	60	2gyk:E	60	83	☑	CRYSTAL STRUCTURE OF THE COMPLEX OF THE COLICIN E9 DNASE DOMAIN WITH A MUTANT IMMUNITY PROTEIN, IMME9 (D51A)

b

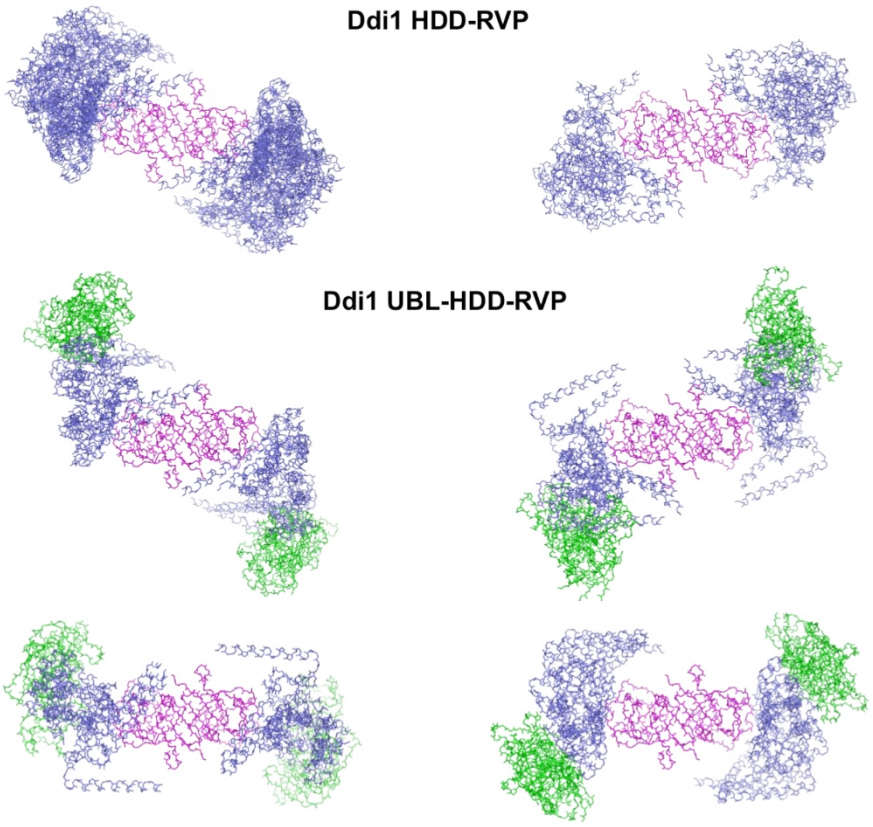


Supplementary Figure S7. Structural homology of the yeast HDD domain. **(a)** PDBeFold top hits for the HDD N-terminal domain (residue 90-141). Only structures with $N_{align} > 45$ were selected. **(b)** Output from the SWISS-MODEL web platform (http://swissmodel.expasy.org), using the yeast Ddi1 HDD sequence (86-196) as input. The three templates with the highest sequence identity (above 20%) were selected for analysis, excluding redundant structures with highly similar sequences. Top: structures of the three templates, with regions homologous to Ddi1 HDD highlighted in blue. Bottom: sequence alignment of Ddi1 HDD with the three templates, with PDB names indicated.

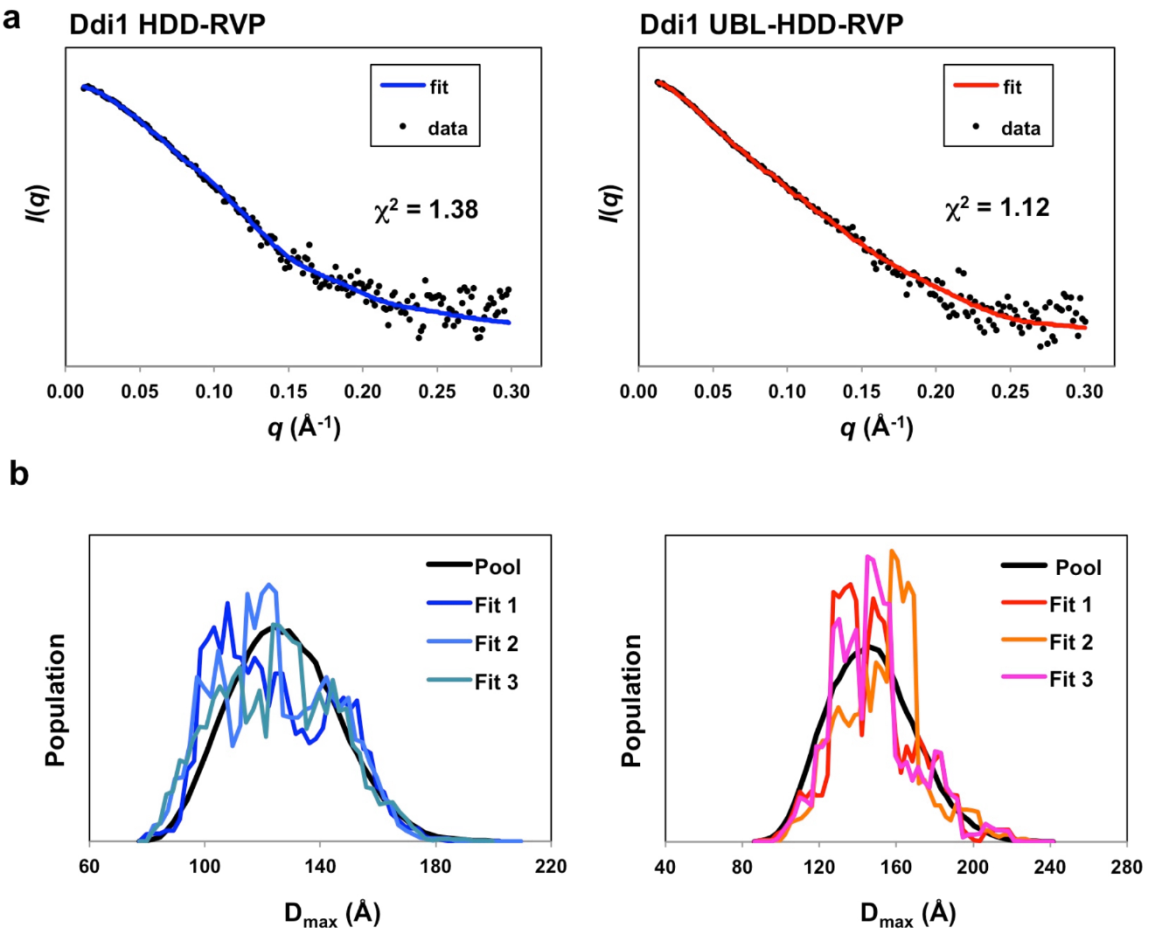
a

	Ddi1 UBL-HDD-RVP	Ddi1 HDD-RVP	Ddi1 HDD-RVP + 5% glycerol	Ddi1 RVP
R_g (SAXS)	40.7 ± 0.5 Å	40 ± 1 Å	34.6 ± 0.3 Å	25.9 ± 0.2 Å
R_h (QELS)	45 ± 3 Å	43 ± 4 Å	41 ± 4 Å	n.d.
M.wt (SAXS)	70 ± 5 kDa	65 ± 5 kDa	58 ± 5 kDa	35 ± 3 kDa
M.wt. (MALS)	72 ± 2 kDa	60 ± 4 kDa	52 ± 3 kDa	n.d.
Calc. M.wt	74.1 kDa	54.5 kDa	54.5 kDa	32.6 kDa

b



Supplementary Figure S8. SAXS analysis of Ddi1 HDD-RVP and UBL-HDD-RVP. (a) Summary of light (R_h, M.Wt) and small-angle X-ray (Rg, M.Wt) scattering parameters for different Ddi1 constructs under different conditions. N.d., not determined **(b)** Structure ensembles (twenty models) generated by the program CORAL for SAXS data acquired on Ddi1 86-325 (HDD-RVP) and 2-325 (UBL-HDD-RVP), using the UBL an HDD NMR structures, and the RVP dimer crystal structure. The UBL, HDD and RVP domains are colored in green, blue and magenta, respectively. The structures were sorted in two (HDD-RVP) or four (UBL-HDD-RVP) classes based on their similarities. The structures were superposed to the RVP domain, which was fixed during the calculations to maintain P2 symmetry. The average χ^2 for the HDD-RVP and UBL-HDD-RVP models are 1.36 and 1.19, respectively.



Supplementary Figure S9. Ensemble modeling of SAXS data for Ddi1 HDD-RVP and UBL-HDD-RVP. (a) Fit of the best ensemble of EOM-generated structures to experimental data for HDD-RVP (left) and UBL-HDD-RVP (right). **(b)** Population distribution of D_{max} for a pool of random structures of the UBL, HDD and RVP (black), compared with a ensembles that best fit the experimental data (blue and red for HDD-RVP and UBL-HDD-RVP, respectively).

Supplementary Table 1. X-ray diffraction data and refinement statistics for Ddi1 185-325

Data collection	CHESS A1
Wavelength (Å)	0.9789
Space group	$P2_12_12_1$
Unit cell dimensions	a=41.52 Å, b= 50.05 Å, c=131.56 Å
	α=β=γ=90°
Mosaicity (°)	0.30
Images	240
Oscillation angle (°)	0.5
Resolution (Å)	50.0-1.80 (1.90-1.80)
Unique reflections	25235 (2930)
Completeness (%)	96.1 (79.4)
Multiplicity	4.2 (2.6)
$<I/\sigma(I)>$	25.3 (7.0)
R_{merge}†	0.036 (0.127)
Solvent content (%)	45
No. of reflections in R_{free} set	1285
R_{work}	0.183
R_{free}	0.213
FOM	0.881
Rms deviations from ideal values‡	
Bond length (Å)	0.011
Bond angle (°)	1.5
Torsion angle (°)	6.3
Protein atoms	2090
Solvent atoms	149
Average B-factor (Å^2)	
Protein main chain	23
Protein side-chain	26
Water	33
Ramachandran outliers§	none

† $\sum_{hkl}\sum_{i}\frac{|I_i(hkl)-\langle I(hkl)\rangle|}{\Sigma_{hkl}\Sigma_i I_i(hkl)}$, where $I_i(hkl)$ is the intensity of the *i*th measurement of reflection *hkl* and $\langle I(hkl)\rangle$ is the mean value for all *i* measurements

‡ Ideal values as reported in Engh & Huber (2001)

§ Residues for which the backbone torsion angles are outside the core region of the Ramachandran plot (Kleywegt & Jones, 1996).

Supplementary Table 2: NMR constraints and structural statistics for the Ddi1 UBL domain (residues 1-80)

NMR distance & dihedral constraints		
Distance constraints		
Total NOE	1653	
Intra-residue	490	
Inter-residue	1163	
Sequential (|i-j| = 1)	409	
Medium-range (|i-j| < 4)	290	
Long-range (|i-j > 5)	464	
Total dihedral angle restraints	128	
phi	64	
psi	64	
Structure Statistics		
Violations (mean and s.d.)		
Max. dihedral angle violation (º)	7.6	
Max. distance constraint violation (Å)	0.497	
Deviations from idealized geometry		
Bond lengths (Å)	0.012	
Bond angles (º)	1.4	
Ramachandran plot summary		
Most favoured regions	94.0%	
Additionally allowed regions	6.0%	
Generously allowed regions	0.0%	
Disallowed regions	0.0%	
Average pairwise r.m.s.d. (Å)	*ordered*	*all residues*
Heavy	0.7	2.0
Backbone	0.3	1.6

Supplementary Table 3: NMR constraints and structural statistics for the Ddi1 HDD domain (residues 86-196)

NMR distance & dihedral constraints		
Distance constraints		
Total NOE	1938	
Intra-residue	589	
Inter-residue		
Sequential (|i-j| = 1)	480	
Medium-range (|i-j| < 4)	514	
Long-range (|i-j > 5)	355	
Total dihedral angle restraints	168	
phi	84	
psi	84	
Structure Statistics		
Violations (mean and s.d.)		
Max. dihedral angle violation (º)	3.6	
Max. distance constraint violation (Å)	0.49	
Max. RDC constraint violation (Hz)	1.5	
Deviations from idealized geometry		
Bond lengths (Å)	0.002	
Bond angles (º)	0.452	
Ramachandran plot summary (90-141;151-187)		
Most favoured regions	96.4 %	
Additionally allowed regions	3.5 %	
Generously allowed regions	0.1 %	
Disallowed regions	0.0 %	
Average pairwise r.m.s.d. (Å)	*90-141*	*151-187*
Heavy	1.39	1.78
Backbone	0.57	0.99
Residual dipolar couplings restraints (N)	46	32
RMS deviation (Hz)	0.68±0.03	0.60±0.06
R-factor (%)	12.2±0.6	10.2±1.1

13

8.3. SUPPLEMENT S3: THE YEAST PROTEASES DDI1 AND WSS1 ARE BOTH INVOLVED IN THE DNA REPLICATION STRESS RESPONSE

Svoboda M., Konvalinka J., Trempe J.F. and Grantz Šašková K. The yeast proteases Ddi1 and Wss1 are both involved in the DNA replication stress response. *DNA Repair*. **80**, 45–51 (2019).

My contribution: experimental design; cloning of all used DNA constructs; yeast phenotypic assays; yeast transformation; complementation assays; protein expression analysis; data analysis; wrote the manuscript.

DNA Repair 80 (2019) 45–51

Contents lists available at ScienceDirect

DNA Repair

journal homepage: www.elsevier.com/locate/dnarepair

The yeast proteases Ddi1 and Wss1 are both involved in the DNA replication stress response

Michal Svoboda[a,b,e], Jan Konvalinka[a,c], Jean-François Trempe[d], Klara Grantz Saskova[a,e,*]

[a] Institute of Organic Chemistry and Biochemistry of the Czech Academy of Sciences, Flemingovo n. 2, 16610, Prague, Czech Republic
[b] Department of Physical and Macromolecular Chemistry, Charles University, Hlavova 8, 12843, Prague, Czech Republic
[c] Department of Biochemistry, Charles University, Hlavova 8, 12843, Prague, Czech Republic
[d] Centre for Structural Biology and Department of Pharmacology & Therapeutics, McGill University, Montreal, Quebec, Canada
[e] Department of Genetics and Microbiology, Charles University, Viničná 5, 12843, Prague, Czech Republic

ARTICLE INFO

Keywords:
Yeast
Hydroxyurea
Protease
DNA replication stress
Ddi1
Wss1

ABSTRACT

Genome integrity and cell survival are dependent on proper replication stress response. Multiple repair pathways addressing obstacles generated by replication stress arose during evolution, and a detailed understanding of these processes is crucial for treatment of numerous human diseases. Here, we investigated the strong negative genetic interaction between two proteases involved in the DNA replication stress response, yeast Wss1 and Ddi1. While Wss1 proteolytically acts on DNA-protein crosslinks, mammalian DDI1 and DDI2 proteins remove RTF2 from stalled forks *via* a proposed proteasome shuttle hypothesis. We show that the double-deleted Δ*ddi1*, Δ*wss1* yeast strain is hypersensitive to the replication drug hydroxyurea and that this phenotype can be complemented only by catalytically competent Ddi1 protease. Furthermore, our data show the key involvement of the helical domain preceding the Ddi1 protease domain in response to replication stress caused by hydroxyurea, offering the first suggestion of this domain's biological function. Overall, our study provides a basis for a novel dual protease-based mechanism enabling yeast cells to counteract DNA replication stress.

1. Introduction

During genome replication in eukaryotic cells, replication forks often encounter barriers that block progress down the DNA strand. These situations are generally referred to as replication stress [1], and multiple signaling pathways and processes responsible for counteracting replication stress arose during evolution [2]. As many replication-hindering barriers contain a protein component, it is unsurprising that proteases are involved in the cellular response to replication stress. The ubiquitin proteasome system plays a central role, with 26S proteasome involvement in regulated degradation of signaling molecules [3], stalled RNA polymerases [4], and Ku70/80 proteins trapped after successful non-homologous end joining [5], as well as in general degradation of proteins trapped on the DNA strand [6].

Researchers recently identified a novel proteolytic pathway acting in parallel with the ubiquitin proteasome system. This pathway involves direct proteolysis of proteins covalently trapped on the DNA strand to remove the replication fork-blocking barrier [7,8]. The process was described in *S. cerevisiae*, in which Wss1p, a DNA-activated metalloprotease, cleaves DNA-protein crosslinks (DPCs) such as compromised topoisomerases or products of formaldehyde toxicity in a replication-coupled manner [9,10]. After removal of the bulky protein, repair can continue through the nucleotide excision repair pathway. SPRTN, the mammalian orthologue of Wss1, acts as a DPC protease in mammalian cells, where it cleaves DPCs in a tightly coupled manner regulated by DNA-dependent and deubiquitination switches [11–13].

Other proteases recently found to be involved in the replication stress response include the Ddi1-like proteins. Mammalian DDI1 and DDI2 proteins associate with nascent replication forks and are involved in restarting stalled forks by facilitating removal of the replication termination factor RTF2 [14]. DDI1- and DDI2-depleted cells fail to remove RTF2 from stalled forks, leading to accumulation of single-stranded DNA, sensitivity to replication drugs, and chromosome instability. Although both Ddi1-like proteins harbor a retroviral protease-like (RVP) domain, RTF2 removal has been proposed to proceed *via* a proteasomal shuttling mechanism. In proteasomal shuttling, proteins that contain ubiquitin-like (UBL) and ubiquitin-associated (UBA) domains serve as linkers between the 26S proteasome and poly-ubiquitinylated substrate [15]. This mechanism has been suggested for a yeast orthologue of Ddi1-like proteins, *S. cerevisiae* Ddi1p, which

* Corresponding author at: Institute of Organic Chemistry and Biochemistry of the Czech Academy of Sciences, Flemingovo n.2, 16610 Prague, Czech Republic.
E-mail address: saskova@uochb.cas.cz (K. Grantz Saskova).

https://doi.org/10.1016/j.dnarep.2019.06.008
Received 12 March 2019; Received in revised form 23 May 2019; Accepted 26 June 2019
Available online 27 June 2019

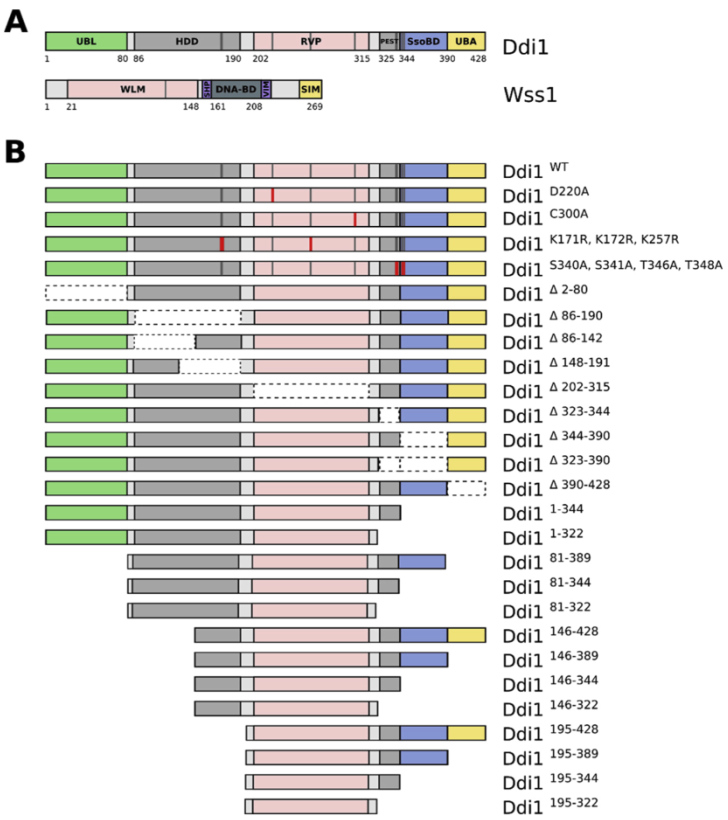


Fig. 1. A) Schematic representation of the domain organization of *S. cerevisiae* Ddi1 and Wss1 proteins. UBL – ubiquitin-like domain, HDD – helical domain of Ddi1-like proteins, RVP – retroviral protease-like domain, PEST – proline, glutamate, serine and threonine rich domain targeting proteins for rapid degradation, SsoBD – Sso tSNARE binding domain, UBA – ubiquitin-associated domain, WLM – WSS1-like metalloprotease domain, SHP – Cdc48 interacting motif found in Shp1, DNA-BD – DNA-binding domain, VIM – VCP-interacting motif, SIM – SUMO-interacting motifs B) Point-mutated, domain-deleted and truncated constructs of *S. cerevisiae* Ddi1 used in this study. Truncated variants of Ddi1 protein were also prepared with C-terminal HA-tag.

possesses UBL-UBA domain architecture (Fig. 1A) and serves as a proteasomal shuttle in degradation of HO endonuclease [16] and UFO1 protein [17,18]. Yeast *DDI1* is inducible by DNA damage and shares a bidirectional promoter with MAG1, a DNA-3-methyladenine glycosylase involved in base-excision repair [19]. Together with another proteasomal shuttle protein, Rad23, Ddi1p suppresses a temperature-sensitive mutation in Pds1 (securin), a mitotic checkpoint control protein, degradation of which by the UPS is required for mitosis [20].

A high-throughput synthetic lethality screen in yeast revealed a strong negative genetic interaction between *WSS1* and *DDI1* [21]. The strong negative genetic interaction with another protease involved in the DNA replication stress response led us to question the proteasomal shuttling hypothesis for the role of Ddi1-like proteins in this response. In this study, we tested the sensitivity of a double-deleted yeast strain lacking *WSS1* and *DDI1* to DNA-damaging chemicals and identified a hypersensitivity to hydroxyurea, a ribonucleotide reductase inhibitor that induces nucleotide depletion and arrests cells in the S phase. Based on previous studies and our structural characterization of yeast and mammalian Ddi1/2 [22,23], we performed rescue experiments with various deletions and single-site mutants of Ddi1 (Fig. 1B) to assess the involvement of particular yeast Ddi1 domains in the replication stress response.

2. Material and methods

2.1. Yeast strains and growth conditions

All *Saccharomyces cerevisiae* strains were isogenic derivatives of strain S288C, in the Y7092 background (*MATα can1Δ::STE2pr-Sp_his5 lyp1Δ ura3Δ0 leu2Δ0 his3Δ1 met15Δ0*), and were obtained as a kind gift from Prof. Charles Boone [21]. Strain genotypes are listed in Supplementary Table S1. Standard yeast YPD (1% yeast extract, 2% bactopeptone, 2% glucose) and SC or SC dropout media were used [24]. For solid media, 1.5% agar was added. All cultivations were performed at 30 °C, unless explicitly stated otherwise; liquid cultures were shaken at 260 RPM in an orbital shaker.

2.2. Plasmid construction and yeast transformation

Plasmids used for knockout phenotype rescue experiments were all derived from CEN-bound plasmid pAG416GPD-ccdB (a gift from Susan Lindquist, Addgene plasmid #14148) [25]. The wild-type *Ddi1* sequence was amplified from the plasmid pET16_wt_Ddi1 [22] using primers P1 and P2; the human *DDI2* sequence was amplified from pET16_wt_Ddi2 [23] using primers P45 and P46 (human DDI1 gene NM_001001711 was synthesized by GenScript and the coding sequence was amplified using primers P43 and P44). AttB adaptors were added by PCR with universal attB primers. Entry clones in pDONR221 were created using the standard protocol for Gateway BP clonase II (Thermo Fisher Scientific). Point mutants in the yeast *Ddi1* sequence were generated using the QuikChange Site-Directed Mutagenesis Kit (Agilent). Entry clones bearing domain deletions, truncated constructs and HA tagged constructs of yeast *Ddi1* were prepared according to the Gibson assembly procedure [26]. Final expression plasmids for knockout phenotype rescue experiments were generated by recombination cloning from the given entry clone into pAG416GPD-ccdB, using Gateway LR clonase II (Thermo Fisher Scientific). All primers and plasmids are listed in Supplementary Table S2. All plasmids were verified by sequencing. Yeast were transformed using the standard LiAc/PEG3350 protocol [27].

2.3. Phenotypic characterization

Spot assays were performed by plating equal amounts of exponentially growing cells in serial dilutions onto solid YPDA media plates in the absence or presence of DNA-damage or DNA replication stress-causing substances. All spot tests for a given experiment were always plated on one plate. Plates were photographed after 60 h of incubation at 30 °C.

2.4. Immunoblotting

Whole-cell protein extracts were prepared from exponentially growing cells as described [28]. Proteins were separated in 12% Tris-Glycine SDS PAGE and wet-transferred onto a nitrocellulose membrane. Immunoblotting was performed using anti-PGK1 (1:5000, Novex 459250), anti-Ddi1 (1:5000, rabbit polyclonal, kind gift from Prof. Gerst, Weizmann Institute of Science [29]) and C29F4 anti-HA (1:1000, Cell Signaling 3724) antibodies. Immunoblots were developed using near-infrared fluorophore-conjugated goat anti-mouse and anti-rabbit secondary antibodies (LI-COR 926-68070, 926-32211) and visualized with a LI-COR Odyssey CLx NIR fluorescence reader.

3. Results

3.1. Simultaneous loss of Ddi1 and Wss1 proteases leads to hypersensitivity to hydroxyurea

According to published data available in TheCellMap.org database [30], the *DDI1* gene displays severe synthetic sickness when deleted together with the *WSS1* gene [21]. As the protein products of both genes are known to be induced by DNA damage [19] or directly involved in the cellular response to DNA damage [10], we decided to test the effects of various DNA-damaging chemicals on growth of a *Δddi1, Δwss1* double-deleted yeast strain. Unlike the previously reported *WSS1* and *TDP1* double deletion, the *DDI1* and *WSS1* double deletion did not confer sensitivity to the topoisomerase inhibitor camptothecin (Fig. 2). Hydrogen peroxide, which causes oxidative damage, also had no effect. The *Δddi1, Δwss1* strain exhibited a slightly elevated sensitivity to methyl methane sulfonate and to long-term exposure to small concentrations of formaldehyde. However, the most pronounced phenotype of the double-deleted strain, compared to wild-type or single knock-out strains, was severe hypersensitivity to the ribonucleotide reductase inhibitor hydroxyurea (Fig. 2).

3.2. Hydroxyurea hypersensitivity can be rescued by overexpression of wild-type Ddi1 but not a catalytically inactive mutant

The *DDI1* gene sequence contains activating elements for the *MAG1* gene, with which it shares a bidirectional promoter [31]. To test whether hydroxyurea hypersensitivity in the double-deletion strain was caused by the loss of *DDI1* expression rather than disruption of *MAG1* expression, we performed a rescue experiment. The *Δddi1, Δwss1* strain was transformed with a single copy of CEN-bound plasmid encoding wild-type Ddi1 protein expressed under the GPD constitutive promoter. Overexpression of wild-type protein led to complete rescue of the hydroxyurea hypersensitivity phenotype (Fig. 3A). This result confirms that the hypersensitivity to hydroxyurea observed in the *Δddi1, Δwss1* strain is caused by the loss of the Ddi1 protein.

Next, we explored the roles of several point mutations at Ddi1 residues of known functional importance. Mutation of all three lysine residues targeted by ubiquitination, as well as all four residues known to be phosphorylated during Ddi1's involvement in exocytosis regulation [32], did not have an effect on the ability of the overexpressed protein to rescue hydroxyurea hypersensitivity, nor did mutation of Cys300, which was previously shown to cause Ddi1 accumulation in cytoplasmic punctuate structures [32]. On the other hand, mutation of the catalytic aspartate (Asp220) in the RVP domain [22,33,34] led to a complete loss of rescue, showing a phenotype identical to that of the original *Δddi1, Δwss1* double mutant strain (Fig. 3A). Thus, resistance to hydroxyurea requires the enzymatic activity of the Ddi1 aspartyl protease domain.

3.3. The multi-helical bundle preceding the RVP domain is essential for Ddi1's role in the DNA replication stress response

NMR and X-ray scattering experiments have shown that yeast Ddi1 is a multi-domain protein composed of multiple structured domains separated by flexible linkers [22,35]. To test whether Ddi1 works in replication stress reversal as a whole protein or if only some of its domains are involved, we created plasmids encoding truncated variants of the protein lacking one particular domain (summarized in Fig. 1B). Expression of Ddi1 variants was confirmed by immunoblotting using either anti-Ddi1 antibody or anti-HA antibody (Fig. S1). Constructs lacking either the UBL (Δ2-80), PEST (Δ323-344), Sso binding (Δ344-390) or UBA (Δ390-428) domains were capable of fully rescuing Δddi1, Δwss1 hydroxyurea hypersensitivity (Fig. 3B). A construct lacking both the PEST and Sso binding domains (Δ323-390) performed less effectively. As expected from our finding for a catalytic Asp220 mutation, a Ddi1 variant lacking the RVP domain (Δ202-315) completely lost the ability to rescue hydroxyurea hypersensitivity. The same phenotype was also observed for Ddi1 lacking the recently identified Helical Domain of Ddi (HDD), a domain of unknown function [22,23]. Based on NMR structural data, HDD appears to comprise two multi-helical bundles connected by a flexible linker [22]. We therefore also prepared constructs lacking either the first or the second helical bundle of HDD. Only the second bundle spanning residues 148–191, located just before the protease domain, seems to be essential for Ddi1's function in hydroxyurea hypersensitivity rescue.

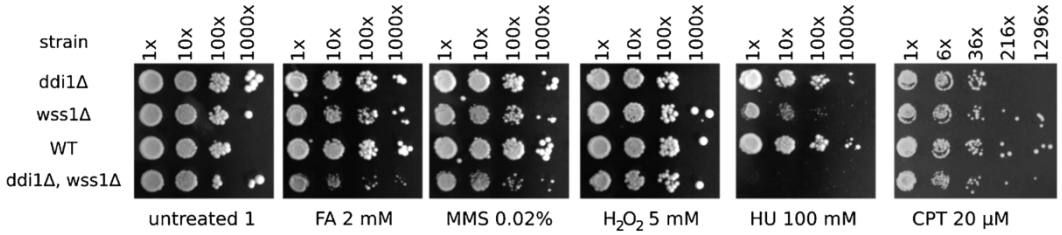


Fig. 2. The *Δddi1, Δwss1* strain is hypersensitive to hydroxyurea. Dilution spot assays testing the viability of wild-type, single knock-out and double knock-out strains under the influence of different DNA damage-causing chemicals. Equal amounts of exponentially growing cells were plated in serial dilutions on plates containing either YPDA media only or YPDA supplemented with formaldehyde (FA), methyl methane sulfonate (MMS), hydrogen peroxide, hydroxyurea (HU) or camptothecin (CPT).

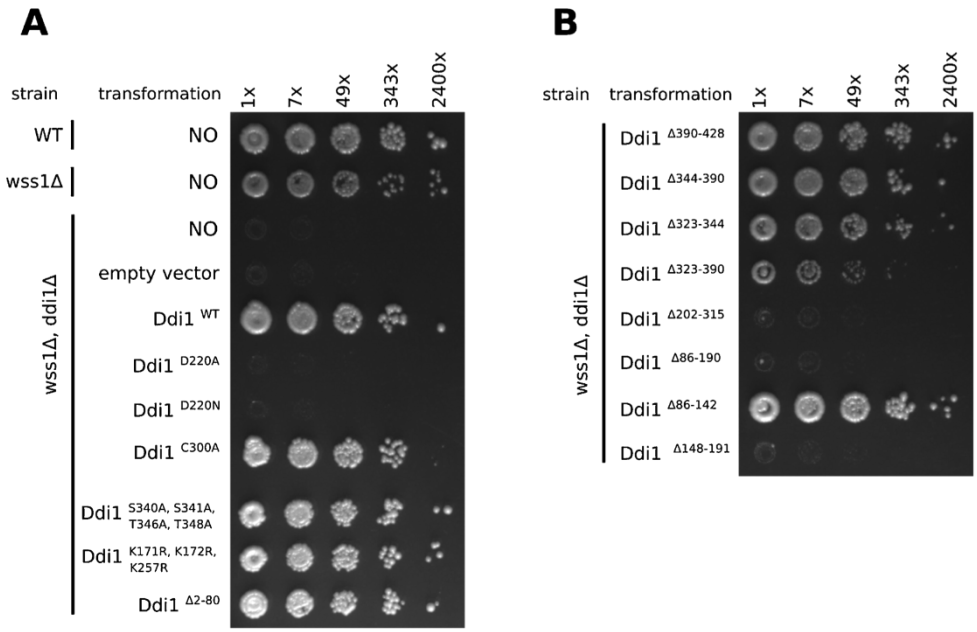


Fig. 3. Retroviral protease-like (RVP) domain catalytic activity and the presence of a helical bundle (HDD) preceding the RVP domain are indispensable for complementation of hydroxyurea sensitivity. Equal amounts of exponentially growing cells transformed with single-copy plasmids overexpressing under control of the GPD promotor A) point-mutated variants of Ddi1 protein or B) variants with deletions of known structured domains of Ddi1 were plated in 7-fold serial dilutions on one YPDA plate supplemented with 50 mM hydroxyurea.

3.4. Ubiquitin-like and ubiquitin-associated domains are not required for hydroxyurea hypersensitivity rescue

To further explore the mechanism of action of Ddi1 in the DNA replication stress response, we assessed which combination of domains is most essential for hydroxyurea hypersensitivity rescue. Neither the UBL nor UBA domain is required, as constructs lacking both (81–389) are capable of rescuing the phenotype (Fig. 4). On the other hand, constructs with N-terminal truncation beyond the second helical bundle (195-xxx, *etc.*) were incapable of rescue, highlighting the crucial role of the second helical bundle in the DNA replication stress response. We identified a construct containing the second helical bundle and RVP domain (residues 146–322) as a minimal construct sufficient to fill Ddi1's role in the DNA replication stress response, although we observed a slight decrease in effectivity. Due to loss of anti-Ddi1 antibody detection of some of the truncated variants, we performed identical dilution spot experiment using C-terminally HA-tagged Ddi1 protein variants and obtained the same results (Fig. S1).

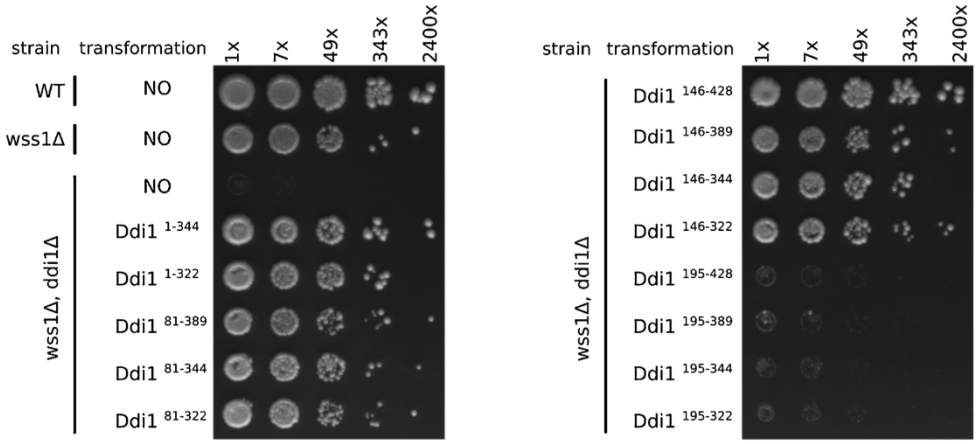


Fig. 4. Complementation of hydroxyurea sensitivity by Ddi1 overexpression does not require the ubiquitin-like and ubiquitin-associated domains of Ddi1. The minimal construct capable of complementation spans residue 146 to residue 322, covering the multi-helical bundle and retroviral protease-like domain. Dilution spot assays with 7-fold serial dilutions of cells overexpressing truncated variants of Ddi1 protein plated on YPDA supplemented with 50 mM hydroxyurea are shown. All 16 spot tests were done on one plate.

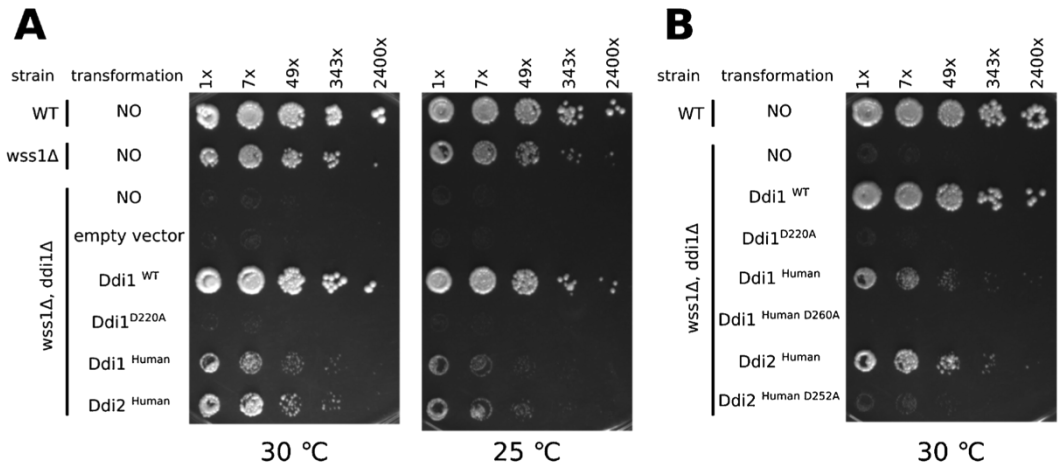


Fig. 5. Human DDI1 and DDI2 proteins are capable of partial, proteolysis dependent, complementation of hydroxyurea sensitivity. Dilution spot assays with 7-fold serial dilutions of cells overexpressing A) yeast and human Ddi1-like proteins and B) their respective catalytically inactivated mutant variants plated on one YPDA plate supplemented with 50 mM hydroxyurea are shown.

3.5. Human Ddi1-like proteins can partially rescue hydroxyurea hypersensitivity

Next, we tested whether our results also apply to human Ddi1 orthologues that contain the HDD helical domain preceding the retroviral aspartyl protease domain. We found that overexpression of both human DDI1 and DDI2 proteins can partially complement Δddi1, Δwss1 hydroxyurea sensitivity (Fig. 5), but the observed effect is weaker compared to yeast Ddi1 overexpression and the difference is greater for plates cultivated at 25 °C instead of 30 °C. Similar to the yeast DDI1, complementation by both human orthologues was dependent on the presence of the catalytic active site aspartate (Fig. 5B).

4. Discussion

The ability to counter DNA replication stress by repairing and restarting stalled replication forks is essential for successful completion of the cell cycle. Therefore, multiple pathways have evolved to respond to replication stress, including the activity of proteolytic enzymes. This study focuses on the characterization of the role of yeast Ddi1 protein in the DNA replication stress response. As part of a genome-wide synthetic lethality screen in yeast, *DDI1* exhibited a strong negative genetic interaction with *WSS1*, another replication stress-countering protease. Interestingly, there is no interaction of *DDI1* with *SGS1* or *TDP1* that are lethal in combination with *WSS1* (Supplementary Table S3) [21]. Another study identified human proteins DDI1 and DDI2, mammalian orthologues of Ddi1 that are involved in degradation of the replication termination factor RTF2, facilitating restart of stalled replication forks [14].

Inspired by the Synthetic Genetic Array in yeast, we decided to explore the interplay between Ddi1 and Wss1 with a focus on their role in response to DNA damage and DNA replication stress. By testing the sensitivity of the double-mutated *Δddi1, Δwss1* yeast strain to various DNA damage agents, we found the strain to be extremely sensitive to hydroxyurea. Hydroxyurea causes replication fork stalling by inhibiting ribonucleotide reductase and therefore depleting free nucleotide triphosphate building blocks. Unlike in mammalian cells, Ddi1 deletion in yeast does not itself cause sensitivity to hydroxyurea. Deletion of *WSS1* had been previously found to cause mild sensitivity to hydroxyurea [36], as confirmed in our experiments. The synthetic effect of both mutations substantially affected hydroxyurea sensitivity, beyond simple additivity. This suggests that Ddi1 and Wss1 act in two independent pathways, each counteracting the replication stress caused by hydroxyurea. As hydroxyurea is not known to directly induce DPCs, the main type of lesion resolved by Wss1, our observation raises a number of questions. For example, it remains unclear which DNA damage response pathway triggered by hydroxyurea leads to growth arrest when both Ddi1 and Wss1 are lost. The double deletion strain may promote upregulation of checkpoint pathways in the S phase, which induce cell cycle arrest in the presence of DNA damage. Other questions to be addressed included which substrate(s) the Ddi1 protease domain cleaves in this context and whether the target molecule is the same for Wss1.

Next, we focused on understanding the mechanism of Ddi1's function in the yeast replication stress response. First, we confirmed that the observed hydroxyurea sensitivity is indeed caused by the loss of Ddi1 expression by overexpressing Ddi1 from a plasmid to complement the phenotype. Next, we analyzed Ddi1's mode of action by complementation experiments with several Ddi1 variants derived from previous functional and structural studies. We found that the catalytic activity of the RVP domain is essential. Indeed, there were only two specific complementation variants that completely abolished rescue of the hydroxyurea sensitivity phenotype: mutation in the catalytic aspartate of RVP and deletion of the helical bundle directly preceding RVP. Our data are thus consistent with a previous observation [32] identifying the catalytically active Ddi1 protease domain as necessary for the rescue of *pds1-128* cells. Many publications describing yeast Ddi1 have focused on its role as a proteasomal shuttling protein, bringing polyubiquitinated substrates to the 26S proteasome [15–18,20]. A shuttling mechanism has also been proposed for human Ddi1-like proteins in the degradation of RTF2 and subsequent restart of replication forks [14]. Interestingly, our data shows that Ddi1 provides hydroxyurea resistance independent of its role in proteasome shuttling, as deletion of UBL, UBA or both did not impair the ability of the protein to rescue hydroxyurea sensitivity. Our findings also contrast with the so-called alternative proteasomal shuttle hypothesis proposed by Nowicka and coworkers [35], which suggests that the first Ddi1 UBL domain binds polyubiquitinated substrate and the other UBL targets the proteasome. The hydroxyurea hypersensitivity phenotype of the double-mutated *Δddi1, Δwss1* yeast strain can be partially complemented even with a minimal construct comprising the double helical bundle followed by the RVP domain (Ddi1 146–322).

To assess whether our results also apply to human orthologues, we complemented the double-deleted yeast strain with plasmids encoding

human DDI1 and DDI2 wild-type proteins as well as their catalytically inactive mutants. Both human DDI1 orthologues showed partial complementation of hydroxyurea sensitivity, dependent on the proteolytic activity. This can be explained by differences in the structure of putative substrates (RTF2, a known target of both human Ddi1-like proteins, does not have an orthologue in *S. cerevisiae*) as well as by reasons rooted in enzymology. Proteases usually have a relatively narrow temperature optimum, and the 7 °C temperature difference between the temperature optima of most human enzymes and cultivation temperature for yeast cultures may be sufficient to impair the activity of the human Ddi1 proteases. To test this hypothesis we performed the rescue experiment on even lower cultivation temperature of 25 °C and indeed we observed further decrease in complementation efficiency relative to yeast Ddi1. Nevertheless, the experiment demonstrates a conservation of function between yeast and human Ddi1 orthologues. Notably, sequence conservation in Ddi1 orthologues is higher for the second bundle of the HDD and the protease domains, and some orthologues lack either the UBL or the UBA domain [22,23]. This is consistent with the main function of Ddi1 being dependent on its protease activity, perhaps with the substrate-recognition function carried out by the HDD domain.

During revisions of our manuscript, Serbyn and co-workers released a preprint addressing questions similar to those raised in our study [37]. Their findings are complementary with our results, showing also the hypersensitivity of Δ*ddi1,* Δ*wss1* yeast strain (although in different genetic background) to hydroxyurea, as well as the requirement of RVP catalytic activity and the presence of HDD domain for successful Ddi1 complementation. Their data further support our results by showing Ddi1 localization in nuclear fraction and its direct association with chromatin. Furthermore they suggest possible Ddi1 substrates involved in replication stress response.

Taken together, our data suggest the existence of a dual protease mechanism providing yeast cells with the ability to overcome DNA-replication stress caused by the ribonucleotide reductase inhibitor hydroxyurea. Furthermore, we identified a putative function for the recently described HDD domain. The complete substrate repertoire of these proteases and the detailed mechanistic basis of this new pathway remain to be further analyzed.

Funding

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], by the project "BIOCEV" (CZ.1.05/1.1.00/02.0109) and by the European Regional Development Fund; OP RDE; Project: "Chemical biology for drugging undruggable targets (ChemBioDrug)" (No. CZ.02.1.01/0.0/0.0/16_019/0000729).

Acknowledgements

We thank Drs. Jason Tanny, Jan Šilhán and Monika Sivá for useful discussions. We thank Prof. Charles Boone for providing yeast strains, Prof. Jeffrey E. Gerst for anti-Ddi1 antibody and Prof. Susan Lindquist for pAG416GPD-ccdB plasmid. We also thank Dr. Hillary Hoffman for language editing.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.dnarep.2019.06.008.

References

[1] M.K. Zeman, K.A. Cimprich, Causes and consequences of replication stress, Nat. Cell Biol. 16 (2014) 2–9.
[2] D. Branzei, M. Foiani, The checkpoint response to replication stress, DNA Rep. (Amst.) 8 (2009) 1038–1046.
[3] Y.W. Zhang, D.M. Otterness, G.G. Chiang, W. Xie, Y.C. Liu, F. Mercurio, R.T. Abraham, Genotoxic stress targets human Chk1 for degradation by the ubiquitin-proteasome pathway, Mol. Cell 19 (2005) 607–618.
[4] E.C. Woudstra, C. Gilbert, J. Fellows, L. Jansen, J. Brouwer, H. Erdjument-Bromage, P. Tempst, J.Q. Svejstrup, A Rad26-Def1 complex coordinates repair and RNA pol II proteolysis in response to DNA damage, Nature 415 (2002) 929–933.
[5] J. van den Boom, M. Wolf, L. Weimann, N. Schulze, F. Li, F. Kaschani, A. Riemer, C. Zierhut, M. Kaiser, G. Iliakis, H. Funabiki, H. Meyer, VCP/p97 extracts sterically trapped Ku70/80 rings from DNA in double-strand break repair, Mol. Cell 64 (2016) 189–198.
[6] N.B. Larsen, A.O. Gao, J.L. Sparks, I. Gallina, R.A. Wu, M. Mann, M. Raschle, J.C. Walter, J.P. Duxin, Replication-coupled DNA-Protein crosslink repair by SPRTN and the proteasome in Xenopus egg extracts, Mol. Cell (2018).
[7] J. Fielden, A. Ruggiano, M. Popovic, K. Ramadan, DNA protein crosslink proteolysis repair: From yeast to premature ageing and cancer in humans, DNA Rep. (Amst.) (2018).
[8] J. Stingele, R. Bellelli, S.J. Boulton, Mechanisms of DNA-protein crosslink repair, Nat. Rev. Mol. Cell Biol. 18 (2017) 563–573.
[9] M.Y. Balakirev, J.E. Mullally, A. Favier, N. Assard, E. Sulpice, D.F. Lindsey, A.V. Rulina, X. Gidrol, K.D. Wilkinson, Wss1 metalloprotease partners with Cdc48/Doa1 in processing genotoxic SUMO conjugates, Elife (2015) 4.
[10] J. Stingele, M.S. Schwarz, N. Bloemeke, P.G. Wolf, S. Jentsch, A DNA-dependent protease involved in DNA-protein crosslink repair, Cell 158 (2014) 327–338.
[11] J. Lopez-Mosqueda, K. Maddi, S. Prgomet, S. Kalayil, I. Marinovic-Terzic, J. Terzic, I. Dikic, SPRTN is a mammalian DNA-binding metalloprotease that resolves DNA-protein crosslinks, Elife (2016) 5.
[12] J. Stingele, R. Bellelli, F. Alte, G. Hewitt, G. Sarek, S.L. Maslen, S.E. Tsutakawa, A. Borg, S. Kjaer, J.A. Tainer, J.M. Skehel, M. Groll, S.J. Boulton, Mechanism and regulation of DNA-Protein crosslink repair by the DNA-dependent metalloprotease SPRTN, Mol. Cell 64 (2016) 688–703.
[13] B. Vaz, M. Popovic, J.A. Newman, J. Fielden, H. Aitkenhead, S. Halder, A.N. Singh, I. Vendrell, R. Fischer, I. Torrecilla, N. Drobnitzky, R. Freire, D.J. Amor, P.J. Lockhart, B.M. Kessler, G.W. McKenna, O. Gileadi, K. Ramadan, Metalloprotease SPRTN/DVC1 orchestrates replication-coupled DNA-protein crosslink repair, Mol. Cell 64 (2016) 704–719.
[14] M.C. Kottemann, B.A. Conti, F.P. Lach, A. Smogorzewska, Removal of RTF2 from stalled replisomes promotes maintenance of genome integrity, Mol. Cell 69 (2018) 24.
[15] D. Finley, Recognition and processing of ubiquitin-protein conjugates by the proteasome, Annu. Rev. Biochem. 78 (2009) 477–513.
[16] L. Kaplun, R. Tzirkin, A. Bakhrat, N. Shabek, Y. Ivantsiv, D. Raveh, The DNA damage-inducible UbL-UbA protein Ddi1 participates in Mec1-mediated degradation of Ho endonuclease, Mol. Cell. Biol. 25 (2005) 5355–5362.
[17] Y. Ivantsiv, L. Kaplun, R. Tzirkin-Goldin, N. Shabek, D. Raveh, Unique role for the UbL-UbA protein Ddi1 in turnover of SCFUfo complexes, Mol. Cell. Biol. 26 (2006) 1579–1588.
[18] O. Voloshin, A. Bakhrat, S. Herrmann, D. Raveh, Transfer of Ho Endonuclease and Ufo1 to the proteasome by the UbL-UbA shuttle protein, Ddi1, analysed by complex formation in vitro, PLoS One (2012) 7.
[19] Y.L. Liu, W. Xiao, Bidirectional regulation of two DNA-damage-inducible genes, MAG1 and DDI1, from Saccharomyces cerevisiae, Mol. Microbiol. 23 (1997) 777–789.
[20] D.J. Clarke, G. Mondesert, M. Segal, B.L. Bertolaet, S. Jensen, M. Wolff, M. Henze, S.I. Reed, DOsage suppressors of pds1 implicate ubiquitin-associated domains in checkpoint control, Mol. Cell. Biol. 21 (2001) 1997–2007.
[21] M. Costanzo, B. VanderSluis, E.N. Koch, A. Baryshnikova, C. Pons, G. Tan, W. Wang, M. Usaj, J. Hanchard, S.D. Lee, V. Pelechano, E.B. Styles, M. Billmann, J. van Leeuwen, N. van Dyk, Z.Y. Lin, E. Kuzmin, J. Nelson, J.S. Piotrowski, T. Srikumar, S. Bahr, Y. Chen, R. Deshpande, C.F. Kurat, S.C. Li, Z. Li, M.M. Usaj, H. Okada, N. Pascoe, B.J. San Luis, S. Sharifpoor, E. Shuteriqi, S.W. Simpkins, J. Snider, H.G. Suresh, Y. Tan, H. Zhu, N. Malod-Dognin, V. Janjic, N. Przulj, O.G. Troyanskaya, I. Stagljar, T. Xia, Y. Ohya, A.C. Gingras, B. Raught, M. Boutros, L.M. Steinmetz, C.L. Moore, A.P. Rosebrock, A.A. Caudy, C.L. Myers, B. Andrews, C. Boone, A global genetic interaction network maps a wiring diagram of cellular function, Science (2016) 353.
[22] J.F. Trempe, K.G. Saskova, M. Siva, C.D.H. Ratcliffe, V. Veverka, A. Hoegl, M. Menade, X. Feng, S. Shenker, M. Svoboda, M. Kozisek, J. Konvalinka, K. Gehring, Structural studies of the yeast DNA damage-inducible protein Ddi1 reveal domain architecture of this eukaryotic protein family, Sci. Rep. (2016) 6.
[23] M. Siva, M. Svoboda, V. Veverka, J.F. Trempe, K. Hofmann, M. Kozisek, R. Hexnerova, F. Sedlak, J. Belza, J. Brynda, P. Sacha, M. Hubalek, J. Starkova, I. Flaisigova, J. Konvalinka, K.G. Saskova, Human DNA-damage-inducible protein 2 is structurally and functionally distinct from its yeast ortholog, Sci. Rep. 6 (2016) 30443.
[24] F. Sherman, Getting started with yeast, Methods Enzymol. 350 (2002) 3–41.
[25] S. Alberti, A.D. Gitler, S. Lindquist, A suite of gateway (R) cloning vectors for high-throughput genetic analysis in Saccharomyces cerevisiae, Yeast 24 (2007) 913–919.
[26] D.G. Gibson, Enzymatic assembly of overlapping DNA fragments, Methods Enzymol. 498 (2011) 349–361.
[27] R.D. Gietz, R.H. Schiestl, Quick and easy yeast transformation using the LiAc/SS carrier DNA/PEG method, Nat. Protoc. 2 (2007) 35–37.
[28] V.V. Kushnirov, Rapid and reliable protein extraction from yeast, Yeast 16 (2000) 857–860.
[29] V. Lustgarten, J.E. Gerst, Yeast VSM1 encodes a v-SNARE binding protein that may act as a negative regulator of constitutive exocytosis, Mol. Cell. Biol. 19 (1999) 4480–4494.

[30] M. Usaj, Y. Tan, W. Wang, B. VanderSluis, A. Zou, C.L. Myers, M. Costanzo, B. Andrews, C. Boone, TheCellMap.org: a web-accessible database for visualizing and mining the global yeast genetic interaction network, G3 (Bethesda) 7 (2017) 1539–1549.
[31] Y. Liu, H. Dai, W. Xiao, 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 (1997) 533–542.
[32] G. Gabriely, R. Kama, R. Gelin-Licht, J.E. Gerst, Different domains of the UBL-UBA ubiquitin receptor, Ddi1/Vsm1, are involved in its multiple cellular roles, Mol. Biol. Cell 19 (2008) 3625–3637.
[33] D.M. Krylov, E.V. Koonin, A novel family of predicted retroviral-like aspartyl proteases with a possible key role in eukaryotic cell cycle control, Curr. Biol. 11 (2001) R584–7.
[34] R. Sirkis, J.E. Gerst, D. Fass, Ddi1, a eukaryotic protein with the retroviral protease fold, J. Mol. Biol. 364 (2006) 376–387.
[35] U. Nowicka, D. Zhang, O. Walker, D. Krutauz, C.A. Castaneda, A. Chaturvedi, T.Y. Chen, N. Reis, M.H. Glickman, D. Fushman, DNA-damage-inducible 1 protein (Ddi1) contains an uncharacteristic ubiquitin-like domain that binds ubiquitin, Structure 23 (2015) 542–557.
[36] B.M. O'Neill, D. Hanway, E.A. Winzeler, F.E. Romesberg, Coordinated functions of WSS1, PSY2 and TOF1 in the DNA damage response, Nucleic Acids Res. 32 (2004) 6519–6530.
[37] Nataliia Serbyn, Audrey Noireterre, Ivona Bagdiul, Michael Plank, Agnès H. Michel, Robbie Loewith, Benoît Kornmann, Françoise Stutz, The aspartic protease Ddi1 contributes to DNA-Protein crosslink repair in yeast, bioRxiv (2019) 575860.

Supplementary materials

The yeast proteases Ddi1 and Wss1 are both involved in the DNA replication stress response.

Michal Svoboda[1,2,5], Jan Konvalinka[1,3], Jean-François Trempe[4], Klara Grantz Saskova[1,5*]

[1] Institute of Organic Chemistry and Biochemistry of the Czech Academy of Sciences, Flemingovo n. 2, 16610, Prague, Czech Republic
[2] Department of Physical and Macromolecular Chemistry, Charles University, Hlavova 8, 12843, Prague, Czech Republic
[3] Department of Biochemistry, Charles University, Hlavova 8, 12843, Prague, Czech Republic
[4] Centre for Structural Biology and Department of Pharmacology & Therapeutics, McGill University, Montreal, Quebec, Canada.
[5] Department of Genetics and Microbiology, Charles University, Viničná 5, 12843, Prague, Czech Republic

* Corresponding author at: Institute of Organic Chemistry and Biochemistry of the Czech Academy of Sciences, Flemingovo n.2, 16610 Prague, Czech Republic
E-mail address: saskova@uochb.cas.cz (Klara Grantz Saskova), Phone: +420220183518

Supplementary Table S1: *S.cerevisiae* strains used in this study

Name	Strain	Relevant Genotype (S288c)	Source
background	Y7092	*MATα, can1Δ0::STEpr-Sp_His5, lyp1Δ0, his3Δ1, leu2Δ0, ura3Δ0, met15Δ0*	(Costanzo et al. 2016)
Δddi1	YER143W_sn475	*Y7092, ddi1Δ::natMX4*	(Costanzo et al. 2016)
Δwss1	YHR134W_sn1569	*Y7092, wss1Δ::natMX4*	(Costanzo et al. 2016)
WT	Y8835	*Y7092, ura3Δ0::natMX4*	(Costanzo et al. 2016)
Δddi1, Δwss1		*Y7092, MATa, ddi1Δ::natMX4, wss1Δ::kanMX4*	(Costanzo et al. 2016)

Supplementary Table S2: Plasmids and primers used in this study

Name	Code	Relevant Sequence	Source; Primers used
empty vector	pAG14148	pAG416GPD-ccdB	(Alberti, Gitler, and Lindquist 2007)

pENTRY	vMSV001	pDONR221- Ddi1^{WT}	this study; P1, P2
Ddi1^{WT}	vMSV076	pAG416GPD-Ddi1^{WT}	this study; P1, P2
Ddi1^{D220A}	vMSV077	pAG416GPD-Ddi1^{D220A}	this study; P3, P4
Ddi1$^{\Delta 2-80}$	vMSV078	pAG416GPD-Ddi1$^{\Delta 2-80}$	this study; P5, P6
Ddi1$^{\Delta 390-428}$	vMSV079	pAG416GPD-Ddi1$^{\Delta 390-428}$	this study; P7, P8
Ddi1$^{\Delta 344-390}$	vMSV080	pAG416GPD-Ddi1$^{\Delta 344-390}$	this study; P9, P10
Ddi1$^{\Delta 323-344}$	vMSV081	pAG416GPD-Ddi1$^{\Delta 323-344}$	this study; P11, P12
Ddi1$^{\Delta 323-390}$	vMSV082	pAG416GPD-Ddi1$^{\Delta 323-390}$	this study; P13, P14
Ddi1$^{\Delta 202-315}$	vMSV083	pAG416GPD-Ddi1$^{\Delta 202-315}$	this study; P15, P16
Ddi1$^{\Delta 86-190}$	vMSV084	pAG416GPD-Ddi1$^{\Delta 86-190}$	this study; P17, P18
Ddi1$^{81-389}$	vMSV085	pAG416GPD-Ddi1$^{81-389}$	this study; P5, P6, P7, P8
Ddi1$^{K171R, K172R, K257R}$	vMSV092	pAG416GPD-Ddi1$^{K171R, K172R, K257R}$	this study; P19, P20, P21, P22, P23, P24
Ddi1$^{S340A, S341A, T346A, T348A}$	vMSV095	pAG416GPD-Ddi1$^{S340A, S341A, T346A, T348A}$	this study; P25, P26, P27, P28, P29, P30, P31, P32
Ddi1$^{\Delta 86-142}$	vMSV096	pAG416GPD-Ddi1$^{\Delta 86-142}$	this study; P33, P34
Ddi1$^{\Delta 148-191}$	vMSV097	pAG416GPD-Ddi1$^{\Delta 148-191}$	this study; P35, P36
Ddi1$^{195-322}$	vMSV098	pAG416GPD-Ddi1$^{195-322}$	this study; P6, P7, P37, P38
Ddi1$^{146-428}$	vMSV099	pAG416GPD-Ddi1$^{146-428}$	this study; P6, P40
Ddi1$^{195-428}$	vMSV100	pAG416GPD-Ddi1$^{195-428}$	this study; P6, P37
Ddi1$^{1-344}$	vMSV101	pAG416GPD-Ddi1$^{1-344}$	this study; P7, P39
Ddi1$^{1-322}$	vMSV102	pAG416GPD-Ddi1$^{1-322}$	this study; P7, P38
Ddi1$^{146-389}$	vMSV103	pAG416GPD-Ddi1$^{146-389}$	this study; P6, P7, P8, P40
Ddi1$^{195-389}$	vMSV104	pAG416GPD-Ddi1$^{195-389}$	this study; P6, P7, P8, P37
Ddi1$^{81-344}$	vMSV105	pAG416GPD-Ddi1$^{81-344}$	this study; P5, P6, P7, P39
Ddi1$^{146-344}$	vMSV106	pAG416GPD-Ddi1$^{146-344}$	this study; P6, P7, P39, P40

Ddi1$^{195-344}$	vMSV107	pAG416GPD-Ddi1$^{195-344}$	this study; P6, P7, P37, P39
Ddi1$^{81-322}$	vMSV108	pAG416GPD-Ddi1$^{81-322}$	this study; P5, P6, P7, P38
Ddi1$^{146-322}$	vMSV109	pAG416GPD-Ddi1$^{146-322}$	this study; P6, P7, P38, P40
Ddi1^{C300A}	vMSV114	pAG416GPD-Ddi1^{C300A}	this study; P41, P42
Ddi1^{human}	vMSV115	pAG416GPD-Ddi1^{human}	this study; P43, P44
Ddi2^{human}	vMSV116	pAG416GPD-Ddi2^{human}	this study; P45, P46
Ddi1_HA$^{81-389}$	vMSV125	pAG416GPD-Ddi1_HA$^{81-389}$	this study; P5, P6, P47, P48
Ddi1_HA$^{81-344}$	vMSV126	pAG416GPD-Ddi1_HA$^{81-344}$	this study; P5, P6, P47, P49
Ddi1_HA$^{146-389}$	vMSV127	pAG416GPD-Ddi1_HA$^{146-389}$	this study; P6, P40, P47, P48
Ddi1_HA$^{146-344}$	vMSV128	pAG416GPD-Ddi1_HA$^{146-344}$	this study; P6, P40, P47, P49
Ddi1_HA$^{146-322}$	vMSV129	pAG416GPD-Ddi1_HA$^{146-322}$	this study; P6, P40, P47, P50
Ddi1_HA$^{195-389}$	vMSV130	pAG416GPD-Ddi1_HA$^{195-389}$	this study; P6, P37, P47, P48
Ddi1_HA$^{195-344}$	vMSV131	pAG416GPD-Ddi1_HA$^{195-344}$	this study; P6, P37, P47, P49
Ddi1_HA$^{195-322}$	vMSV132	pAG416GPD-Ddi1_HA$^{195-322}$	this study; P6, P37, P47, P50
Ddi1$^{human, D260A}$	vMSV133	pAG416GPD-Ddi1$^{human, D260A}$	this study; P51, P52
Ddi2$^{human, D252A}$	vMSV134	pAG416GPD-Ddi2$^{human, D252A}$	this study; P53, P54

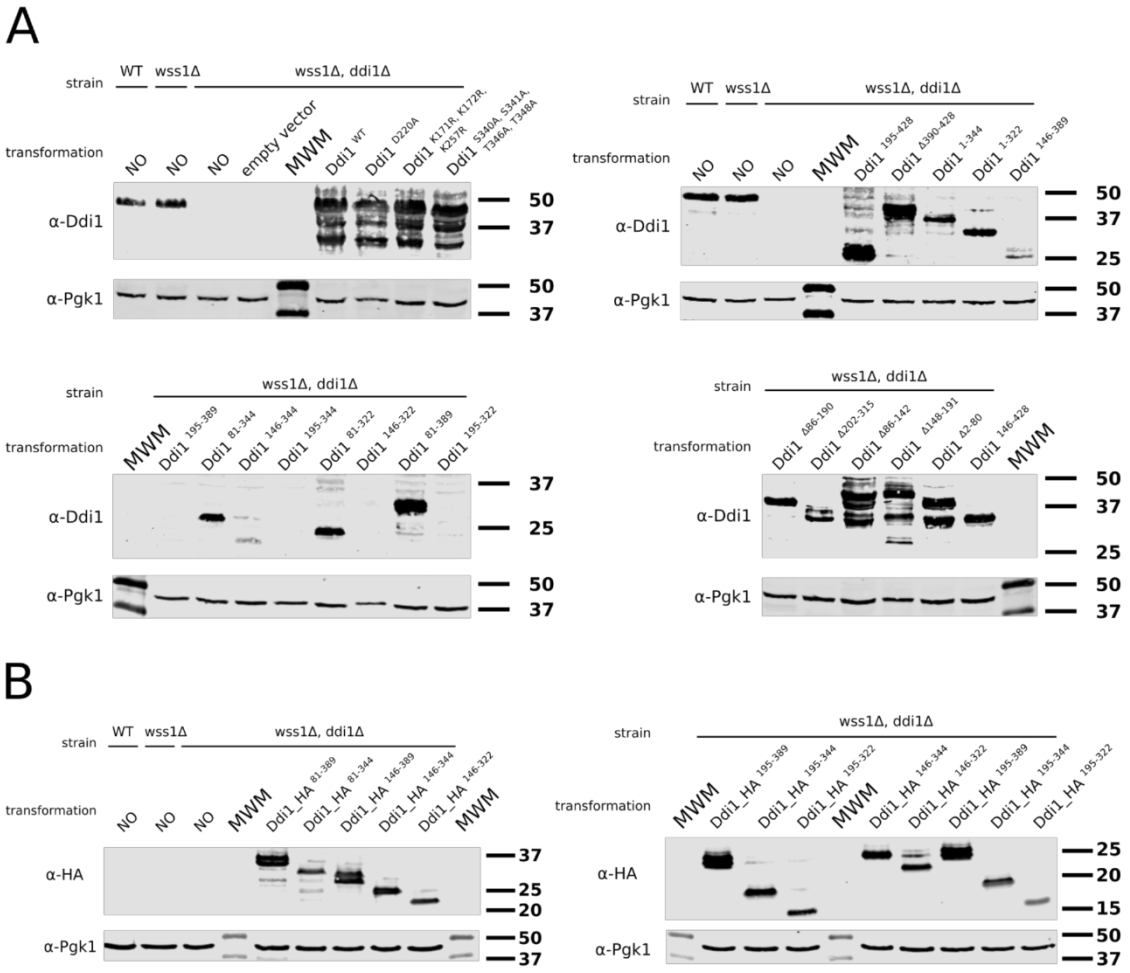
Primer	Sequence (5´-3´)
P1	AAAAAGCAGGCTACAAAATGGATTTAACAATTTCAAACG
P2	AGAAAGCTGGGTTCTATCATTGGAAAAGGAGGGATG
attB1 uni	GGGGACAAGTTTGTACAAAAAAGCAGGCT
attB2 uni	GGGGACCACTTTGTACAAGAAAGCTGGGT
P3	AAAGGCATTTGTAGCTACAGGGGCTCAAA
P4	TTTGAGCCCCTGTAGCTACAAATGCCTTT
P5	GCAGGCTACAAAATGATTCAAACAGATGCTGCTACTTTG
P6	CATTTTGTAGCCTGCTTTTTTGTAC

P7	TGATAGAACCCAGCTTTCTTGTAC
P8	AGCTGGGTTCTATCACGTTCTCCCGGTTGCCGTTG
P9	GTCTGATAAGCCCGAACAAACGATTAAACAG
P10	TCGGGCTTATCAGACGAAGTTGTAACTGAC
P11	GAAGCGGAACTAACACCCACCAAGACTAG
P12	GTGTTAGTTCCGCTTCACTCAAAAAGC
P13	GGAACCCGAACAAACGATTAAACAG
P14	GTTTGTTCGGGTTCCGCTTCACTCAAAAAGC
P15	ACCCAGGTCAGCTTTTTGAGTGAAGCGG
P16	AAAGCTGACCTGGGTAAACATTTCAGG
P17	ACAGATGCTATCGAATATACACCTGAAATGTTTACC
P18	TATTCGATAGCATCTGTTTGAATGGAATTG
P19	ATGATCCTGACAACAGGAAGAGGATTGCAGA
P20	TCTGCAATCCTCTTCCTGTTGTCAGGATCAT
P21	ATCCTGACAACAGGAGGAGGATTGCAGAGCT
P22	AGCTCTGCAATCCTCCTCCTGTTGTCAGGAT
P23	GCGTAGGAACCGGCAGAATTATTGGGAGAAT
P24	ATTCTCCCAATAATTCTGCCGGTTCCTACGC
P25	TCAGTTACAACTTCGGCTGATAAGCCCCTA
P26	TAGGGGCTTATCAGCCGAAGTTGTAACTGA
P27	TCAGTTACAACTGCGGCTGATAAGCCCCTA
P28	TAGGGGCTTATCAGCCGCAGTTGTAACTGA
P29	CTGATAAGCCCCTAGCACCCACCAAGACTAG
P30	CTAGTCTTGGTGGGTGCTAGGGGCTTATCAG
P31	CTGATAAGCCCCTAGCACCCGCCAAGACTAG
P32	CTAGTCTTGGCGGGTGCTAGGGGCTTATCAG
P33	ACAGATGCTGGCTATAACACCGCCATG
P34	GTTATAGCCAGCATCTGTTTGAATGGAATTG

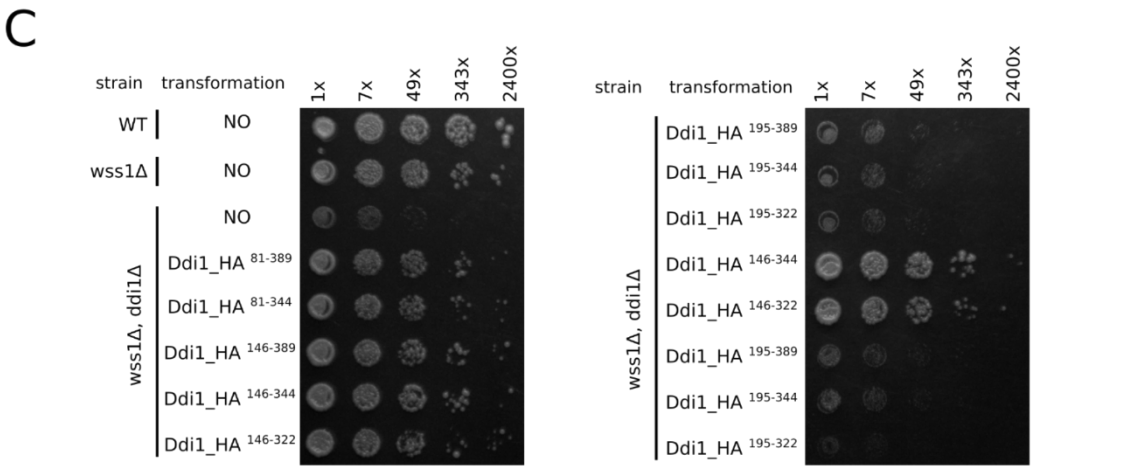
P35	CACCGCCGAATATACACCTGAAATGTTTACCCAGGTC
P36	AGGTGTATATTCGGCGGTGTTATAGCCACCATAAC
P37	GCAGGCTACAAAATGCCTGAAATGTTTACCCAGGTC
P38	AGCTGGGTTCTATCATTCCGCTTCACTCAAAAAGC
P39	AGCTGGGTTCTATCAGGGCTTATCAGACGAAGTTG
P40	GCAGGCTACAAAATGACCGCCATGAATCCTTTTG
P41	GCATTTGGCTGCTGTGGACTTAAAGGAAAAC
P42	GTTTTCCTTTAAGTCCACAGCAGCCAAATGC
P43	GCAGGCTACAAAATGCTGATCACCGTGTACTG
P44	AGCTGGGTTCTATCAATGCTCTTTTCGTCCTGAATC
P45	GCAGGCTACAAAATGCTGCTCACCGTGTACTG
P46	AGCTGGGTTCTATCATGGCTTCTGACGCTCTGC
P47	GGATCCTATCCATATGACGTTCCAGATTACGCTTGAT AGAACCCAGCTTTCTTG
P48	ATATGGATAGGATCCCGTTCTCCCGGTTGCCGTTG
P59	ATATGGATAGGATCCGGGCTTATCAGACGAAGTTG
P50	ATATGGATAGGATCCTTCCGCTTCACTCAAAAAGC
P51	TTGAAGGCTTTTGTTGCCTCGGGCGCCCAGATG
P52	CATCTGGGCGCCCGAGGCAACAAAAGCCTTCAA
P53	CTGGGCACCTGAGGCAACAAAGGCTTTCAC
P54	GTGAAAGCCTTTGTTGCCTCAGGTGCCCAG

Supplementary Table S3: Synthetic Gene Array (SGA) scores for pairwise interactions (Costanzo et al. 2016).

Query gene	Array gene	Genetic Interaction Score	P-value
DDI1	*WSS1*	-0.3660	3.904e-10
WSS1	*DDI1*	-0.2694	4.910e-24
WSS1	*SGS1*	-0.6319	1.921e-49
SGS1	*WSS1*	-0.6025	0.000e+00
WSS1	*TDP1*	-0.4032	6.302e-45
TDP1	*WSS1*	-0.5815	8.191e-18
DDI1	*SGS1*	-0.0027	4.419e-01
SGS1	*DDI1*	0.0273	2.871e-01
DDI1	*TDP1*	0.0138	3.954e-01
TDP1	*DDI1*	0.0232	3.325e-01


A
strain
WT
wss1Δ
wss1Δ, ddi1Δ
transformation
NO
empty vector
MWM
Ddi1 WT
Ddi1 D220A
Ddi1 K171R, K172R, K257R
Ddi1 S340A, S341A, T346A, T348A
α-Ddi1
α-Pgk1
50
37
Ddi1 195-389
Ddi1 81-344
Ddi1 146-344
Ddi1 195-344
Ddi1 81-322
Ddi1 146-322
Ddi1 81-389
Ddi1 195-322
25
Ddi1 195-428
Ddi1 Δ390-428
Ddi1 1-344
Ddi1 1-322
Ddi1 146-389
Ddi1 Δ86-190
Ddi1 Δ202-315
Ddi1 Δ86-142
Ddi1 Δ148-191
Ddi1 Δ2-80
Ddi1 146-428

B
strain
WT
wss1Δ
wss1Δ, ddi1Δ
transformation
NO
MWM
Ddi1_HA 81-389
Ddi1_HA 81-344
Ddi1_HA 146-389
Ddi1_HA 146-344
Ddi1_HA 146-322
Ddi1_HA 195-389
Ddi1_HA 195-344
Ddi1_HA 195-322
α-HA
α-Pgk1
37
25
20
15
50


C
strain
transformation
1x
7x
49x
343x
2400x
WT
wss1Δ
wss1Δ, ddi1Δ
NO
Ddi1_HA 81-389
Ddi1_HA 81-344
Ddi1_HA 146-389
Ddi1_HA 146-344
Ddi1_HA 146-322
Ddi1_HA 195-389
Ddi1_HA 195-344
Ddi1_HA 195-322

Supplementary Figure S1: A) Verification of overexpression of Ddi1 protein variants used in this study. The whole yeast cell extracts were separated by 12% SDS-PAGE followed by Western blotting. Ddi1 detection was performed using anti-Ddi1 antibody (1:5000, rabbit polyclonal, kind gift from Prof. Gerst, Weizmann Institute of Science (Lustgarten and Gerst 1999)); Pgk1 (anti-Pgk1, 1:5000, Novex 459250) was used as a loading control. Size markers (in kDa) are indicated on the right side of the blots. In some cases, overexpressed Ddi1 clearly undergoes partial digestion. We were repeatedly unable to detect some of the Ddi1 variants (Ddi1 $^{146-322}$, Ddi1 $^{195-344}$, Ddi1 $^{195-389}$ and Ddi1 $^{195-322}$), most likely due to the loss of the Ddi1 antibody sensitivity because of the protein truncations. **B)** Truncated Ddi1 protein variants were prepared with C-terminal HA-tag and their overexpression was confirmed by Western blotting. Ddi1 detection was performed using rabbit monoclonal C29F4 anti-HA antibody (1:1000, Cell Signaling 3724); Pgk1 (anti-Pgk1, 1:5000, Novex 459250) was used as a loading control. Size markers (in kDa) are indicated on the right side of the blots. **C)** Dilution spot assays with 7-fold serial dilutions of cells overexpressing HA-tagged truncated variants of Ddi1 protein plated on YPDA supplemented with 50 mM hydroxyurea are shown. Results are in agreement with complementation study using truncated variants of Ddi1 protein lacking the C-terminal HA-tag.

References

Alberti, S., A. D. Gitler, and S. Lindquist. 2007. 'A suite of Gateway (R) cloning vectors for high-throughput genetic analysis in Saccharomyces cerevisiae', *Yeast*, 24: 913-19.

Costanzo, M., B. VanderSluis, E. N. Koch, A. Baryshnikova, C. Pons, G. Tan, W. Wang, M. Usaj, J. Hanchard, S. D. Lee, V. Pelechano, E. B. Styles, M. Billmann, J. van Leeuwen, N. van Dyk, Z. Y. Lin, E. Kuzmin, J. Nelson, J. S. Piotrowski, T. Srikumar, S. Bahr, Y. Chen, R. Deshpande, C. F. Kurat, S. C. Li, Z. Li, M. M. Usaj, H. Okada, N. Pascoe, B. J. San Luis, S. Sharifpoor, E. Shuteriqi, S. W. Simpkins, J. Snider, H. G. Suresh, Y. Tan, H. Zhu, N. Malod-Dognin, V. Janjic, N. Przulj, O. G. Troyanskaya, I. Stagljar, T. Xia, Y. Ohya, A. C. Gingras, B. Raught, M. Boutros, L. M. Steinmetz, C. L. Moore, A. P. Rosebrock, A. A. Caudy, C. L. Myers, B. Andrews, and C. Boone. 2016. 'A global genetic interaction network maps a wiring diagram of cellular function', *Science*, 353.

Lustgarten, V., and J. E. Gerst. 1999. 'Yeast VSM1 encodes a v-SNARE binding protein that may act as a negative regulator of constitutive exocytosis', *Molecular and Cellular Biology*, 19: 4480-94.