The regulation of the ADP-ribosyl-hydrolase MacroD2 upon DNA damage

Barbara Golia

Medizinischen Fakultät

Ludwig-Maximilian-Universität München

München

2016

Aus dem Adolf-Butenandt-Institut
im Biomedizinischen Centrum (BMC)
der Ludwig-Maximilians-Universität München
Lehrstuhl : Physiologische Chemie.
Vorstand : Prof. Andreas G. Ladurner, PhD

The regulation of the ADP-ribosyl-hydrolase MacroD2 upon DNA damage

Dissertation
zum Erwerb des Doktorgrades der Naturwissenschaften
an der Medizinischen Fakultät
der Ludwig-Maximilians-Universität München

vorgelegt von Barbara Golia
aus Neapel, Italien
2016

Gedruckt mit Genehmigung der Medizinischen Fakultät
der Ludwig-Maximilians-Universität München

Betreuer: Prof. Andreas Ladurner, Ph.D.

Zweitgutachter: Prof. Dr. Axel Imhof

Dekan: Prof. Dr. med. dent. Reinhard Hickel

Tag der mündlichen Prüfung: 04.05.2017

Eidesstattliche Versicherung

Golia, Barbara

Name, Vorname

Ich erkläre hiermit an Eides statt,

dass ich die vorliegende Dissertation mit dem Thema

The regulation of the ADP-ribosyl-hydrolase MacroD2 upon DNA damage

selbständig verfasst, mich außer der angegebenen keiner weiteren Hilfsmittel bedient und alle Erkenntnisse, die aus dem Schrifttum ganz oder annähernd übernommen sind, als solche kenntlich gemacht und nach ihrer Herkunft unter Bezeichnung der Fundstelle einzeln nachgewiesen habe.

Ich erkläre des Weiteren, dass die hier vorgelegte Dissertation nicht in gleicher oder in ähnlicher Form bei einer anderen Stelle zur Erlangung eines akademischen Grades eingereicht wurde.

München, 07.09.2016

Ort, Datum

Barbara Golia

Unterschrift Doktorandin/Doktorand

Abstract

To cope with the rapid changes in the environment, all the cellular processes must be carefully regulated and this is often achieved by means of the post-translational modifications. These chemical modifications added to the side chains of amino acids in protein sequences have a variety of consequences, ranging from the stability to the change in protein-protein interactions. One of these post-translational modifications is ADP-ribosylation. ADP-ribosylation is mainly associated to the regulation of DNA damage repair, although it is important for the modulation of a large variety of cellular processes. Being a complex signaling pathway, ADP-ribosylation requires a likely complex metabolism and a number of transferases and hydrolases orchestrate its spatial and temporal dynamics. Among the latter, MacroD2 removes the most proximal ADP-ribosyl-moiety from the glutamate of a substrate protein.

My host lab is interested in the understanding of the modulation of the DNA damage repair response by focusing on the ADP-ribosyl-hydrolases. This brought my direct supervisor, Dr. G. Timinszky, to discover a particular phenomenon associated to MacroD2 protein: when the protein is tagged with EGFP, upon DNA damage it recruits to the DNA lesion, while its nuclear signal decreases over time. The work of my PhD aimed to understand and characterize the strange DNA damage-dependent behavior of MacroD2.

For my work, I use a combination of microscopy and biochemical technics. Thus, I show that the decrease in MacroD2 nuclear protein levels is due to its regulated nuclear export. This behavior of EGFP-tagged MacroD2 construct is also performed by the endogenous protein. I then characterize the stimulus that triggers MacroD2 nuclear export. Upon DNA damage, mostly

upon formation of double-strand breaks, the kinase ATM, master regulator of genotoxic stress response, is activated. ATM induces the phosphorylation of MacroD2 on two specific serine residues located in the MacroD2 intrinsically disordered C-terminus region. This event triggers the nuclear export of MacroD2. I also show that MacroD2 nuclear export is able to affect its own recruitment dynamics at the DNA lesion, suggesting a potential role in the regulation of the DNA damage response.

Although I defined the events inducing MacroD2 nuclear export, to understand the mechanism with which MacroD2 crosses the nuclear envelope, I performed a co-immunopurification experiment associated to peptide mass fingerprinting. By comparing genotoxic stress condition to the control, and a series of MacroD2 constructs (full-length protein, macrodomain or C-terminus fragments), it is possible to draw the interactome of the protein for the specific constructs in the specific conditions. The experiment does not indicate any candidate that could drive MacroD2 into the cytoplasm. On the other hand, the co-immunopurification experiment suggests few hypotheses about MacroD2 functional roles in the cells. In fact, although little is known about MacroD2 functions, by combining the analyses on the enriched proteins with published studies, I formulate testable hypotheses able to connect the enzymatic activity of MacroD2 to its genomic association with autism syndrome.

Zusammenfassung

Alle zellulären Prozesse müssen genau reguliert werden, um auf plötzliche Umwelteinflüsse reagieren zu können. Dies wird häufig durch posttranslationale Modifikationen erreicht. Hierbei werden chemische Modifikationen an Aminosäureseitenketten hinzugefügt. Dies kann viele, verschiedene Auswirkungen auf Zellen haben, von Änderungen der Proteinstabilität bis zu unterschiedlichen Proteininteraktionen. Eine dieser Modifikationen ist die ADP-ribosylierung von Proteinen. ADP-ribosylierung wird hauptsächlich mit der Reparatur von DNA-Schäden verbunden, obwohl diese Modifikation für eine Vielfalt von weiteren zellulären Prozessen von Bedeutung ist. Da

ADP-ribosylierung ein komplexer Signaltransduktionsweg ist, wird ein extensiver Metabolismus benötigt, welcher aus mehreren Transferasen und Hydrolasen besteht. Diese Enzyme bestimmen die örtliche und zeitliche Dynamik dieses Prozesses. Eines der Enzyme ist MacroD2, das den ADP-ribose-Rest, welcher über einen Glutamatrest proximal an ein Substratprotein gebunden ist, entfernt.

Der Arbeitskreis von Dr. G. Timinszky am Department für physiologische Chemie der LMU ist daran interessiert herauszufinden, wie die Reparatur von DNA-Schäden modifiziert wird, insbesondere durch ADP-Ribosyl- Hydrolasen. Dies führte zur Entdeckung eines besonderen Phänomens von MacroD2 durch meinen direkten Betreuer Dr. G. Timinszky. EGFP-markiertes MacroD2 rekrutiert zu induzierten DNA-Läsionen, wobei das nukleäre EGFP Signal über die Zeit geringer wird. Das Ziel meiner Doktorarbeit war es, dieses bemerkenswerte Verhalten von MacroD2 nach der Induktion von DNA-Schäden zu verstehen und zu charakterisieren.

In meiner Doktorarbeit habe ich eine Kombination von Mikroskopie und biochemischen Methoden verwendet. Damit konnte ich zeigen, dass die Reduktion des nukleären MacroD2s durch dessen regulierten Zellkernexport hervorgerufen wird. Dieses Verhalten von EGFP-markiertem MacroD2 wurde auch bei endogenem MacroD2 nachgewiesen. Ich konnte auch den Stimulus charakterisieren, der dazu führt, dass MacroD2 aus dem Zellkern exportiert wird. Durch DNA-Schäden, hauptsächlich durch DNA-Doppelstrangbrüche, wird die Kinase ATM aktiviert, welche Signalkaskaden in der Zelle startet, um auf genotoxischen Stress zu antworten. ATM phoshoryliert MacroD2 an zwei spezifischen Serinresten in dessen ungeordneter, C-terminaler Region. Dies hat den Export von MacroD2 aus dem Zellkern zur Folge. Ich konnte des Weiteren zeigen, dass der Zellkernexport von MacroD2 dessen Rekrutierungsdynamik an DNA-Läsionen beeinflusst. Dies spricht für eine potentielle Rolle von MacroD2 in der Antwort auf DNA-Schäden.

Nachdem ich den Mechanismus definieren konnte, der dazu führt, dass MacroD2 aus dem Zellkern exportiert wird, war es wichtig herauszufinden,

durch welchen Mechanismus MacroD2 die Kernmembran passieren kann. Hierfür habe ich Co-Immunopräzipitationen mit anschließender Analyse der Peptid-massenfingerprints durchgeführt. Durch den Vergleich von genotoxischen Stressbedingungen mit Kontrollbedingungen und durch die Verwendung mehrerer MacroD2 Konstrukte (Volllängenprotein, Macrodomäne und C-terminale Domäne) ist es möglich das Interaktom der Proteine mit MacroD2 für verschiedene Konstrukte und Bedingungen zu definieren. Diese Experimente zeigten jedoch keinen Kandidaten, welcher MacroD2 aus dem Zellkern in das Zytoplasma exportieren könnte. Andererseits konnten mehrere Hypothesen über funktionelle Aufgaben von MacroD2 in Zellen aufgestellt werden. Obwohl wenig über Funktionen von MacroD2 bekannt ist, konnte ich, durch die Kombination der Analyse von angereicherten Proteinen und publizierten Ergebnissen, prüfbare Hypothesen aufstellen. Diese Hypothesen verbinden die enzymatische Funktion von MacroD2 mit genomweiten Assoziationsstudien über das Autismus Syndrom.

To Nonna Maria,
whose fierce spirit has prompted me
to aim to self-improvement
and to finally complete this journey.

Acknowledgements

I would like to thank Prof. Andreas Ladurner for granting me the opportunity to work in his department. The stimulating atmosphere under his direction provided many opportunities for the development of my critical thinking. Additionally, by considering the PhD an important step to forge our skills asset, Prof. Ladurner supported me in participating to scientific events and in developing soft skills useful for my future career.

I thank Gyula Timinszky for being my direct supervisor. Hours of productive discussion and occasionally direct contrasts taught me both on the professional and human level. May the next students be as lucky as me in growing by his side and uncovering their potentialities. And may they be able to understand his sense of humor.

I then want to thank all the members of the Timinszky group, who helped me with scientific discussion, practical help and a truly enjoyable atmosphere: Rebecca Smith, Giuliana Möller, Anna Hegele, Julia Preißer and Mai Ly Tran, thank you all.

I thank Andrew Bowman and Gytis Jankevicius for teaching me, willingly or less willingly, how to survive in the lab: professionalism and craziness at the same time. I also want to thank Aurelio Nardozza, Marta Forn, Ramon Barrales and Ava Handley for the many questions you were asked and the patience with which you were answering them. Additionally, I would like to thank Andreas Schmidt for the help in designing and performing the mass spectrometry experiments.

I would like to thank Anton Eberharter, Christine Werner and Corey Laverty for having taken care of all the bureaucratic and management issues, as well as the long hours of English editing. To all the other members of the department of Physiological Chemistry, thank you for being my "enlarged" family for these four years. I also thank Sébastien Huet and his group for making the visiting time at the University of Rennes both educative and enjoyable.

I want to personally thank Rebecca, Giuliana, Ava and Corey for the precious comments to this thesis, the psychological support and the German editing.

I then would like to thank my mother, my father and my brother Renato for the support and the several “self-esteem” shots. My dear friends Giovanni and Elena, close to me like brothers and sisters, even if far away in kilometers. And the wonderful people who help me make Munich my home: Andy, Sabine, Marta, Bex, Ava, Corey, Gytis, Lisa, Marie, Daria, Tamas and Sebastian. Thanks for making me not feel like a stranger in a strange land.

My last thank goes to Thomas, for your support in the sad moments and your lucid advice; for your patience in listening to my enraged monologues; for the optimism with which you look to the future, and the courage you are able to infuse in my veins.

As small as this piece of work is, it would have never seen the light without you all. Thank you.

Contents

List of Figures

LIST OF FIGURES

LIST OF FIGURES

List of Tables

1

Introduction

Life demands adaptation to new stimuli, from the outside world as well as from the inside. Every cellular process consists of a series of enzymatic reactions performed by a large number of proteins. All these processes need a great deal of coordination: a multi-step pathway must take place in a controlled manner, in the right proportion, at the right time and in the right place. It is indeed a complex task for the cell to organize everything. Therefore, for each pathway, a regulatory system that accurately fine tunes the different steps in that pathway has evolved.

Each process needs to be adjusted according to the variable conditions inside and outside of the cell. When the environment changes rapidly, the cell response needs to be equally fast. This fast response cannot rely on the synthesis of new proteins, since all together transcription, transcript processing and translation are relatively slow. To generate an appropriate quick response, other types of regulation have evolved. For example, post-translational modifications, which are covalently-attached chemical groups, can affect the activity or the stability of the target protein. Another means to achieve the regulation of a pathway is to modulate the localization of the proteins within the cell. Either alone or in combination, these two strategies of regulation can change the activity of a specific pathway and adjust its functionality to the momentary needs of the cell.

1.1 Post-translational modifications in modulating cellular processes

Many eukaryotic proteins are regulated by means of a chemical modification. The amino acid side chains of a protein can be modified covalently by a chemical group or even a small protein, which affects its activity or stability (1). Since the modification occurs after the synthesis of the protein, it is called a post-translational modification. There are many classes of post-translational modifications and virtually all proteins could be regulated through them **(Figure 1.1)**. The most well-studied post-translational modifications are phosphorylation, methylation, acetylation, ubiquitination and ADP-ribosylation. But many more are at work in the cells.

The first evidence of reversible post-translational modification is described in the work on the glycogen phosphorylase (2), whose phosphorylation state dictates the activation levels of the protein (3). Since then, our knowledge of post-translational modifications has expanded rapidly and there is an increasing number of types of modification and mechanisms regulating protein activity. Despite their diversity, most post-translational modifications are modulated through three types of protein components: a "writer", an "eraser" and a "reader". The "writer" is an enzyme, like a transferase, that attaches the modification to the target protein. The "eraser" removes the modification. The "reader" is a protein that is able to decode the modification and induce the effect that the modification is calling for **(Figure 1.2)**. This theory was first elaborated for the modifications occurring on histone proteins (4). However, this theory is generally applicable in describing other post-translational modification-based signaling pathways. Nonetheless, this view may suffer of extreme generalization, as, for example, the reader function can be dispensable. In fact, in many cases the post-translational modification is able to produce an effect by changing the structural conformation of the substrate, in a way that is independent from "reader" factor.

So far, the main body of research describes how one specific post-translational modification regulates one biological process. But it is becoming clear that the different pathways do not work as separated blocks and the survival of the cell is accomplished only when all the processes are coordinated (5). This is the reason why the interactions of different pathways are the real key to understand how life is so apt to adapt. Only

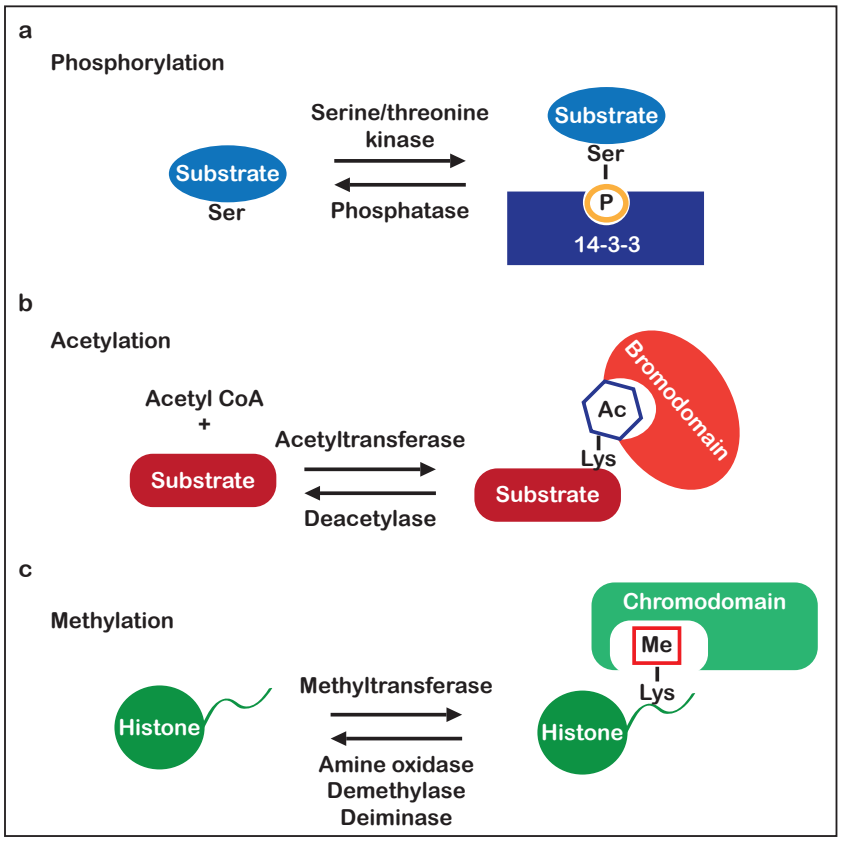


Figure 1.1: Different post-translational modifications present similar functional components - A small selection of post-translational modifications in the cell: phosphorylation (a), acetylation (b), methylation (c). Simplified representation of the reactions that add or remove the post-translational modification. Also shown as an example is one typical domain that binds the protein substrates only when the specific post-translational modification is present. P, phosphorylation; Ac, acetylation; Me, methylation. Adapted from Deribe et al., 2010 (1).

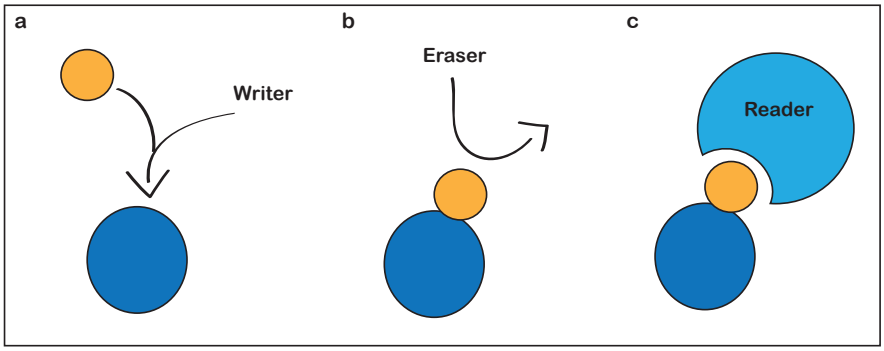


Figure 1.2: Post-translational modifications dynamics result from the combination of three different functions - As originally formulated in the "histone code hypothesis", the structure of each post-translational modification is based on three functions: the "writers" set the signal; the "erasers" remove the signal; the "readers" decode the signal. Adapted from Baker et al., 2008 (4).

by knowing how the network of different processes is interconnected, we can actually understand the system, and maybe one day control it.

1.2 ADP-ribosylation: regulation and functions

ADP-ribosylation is a post-translational modification discovered already in 1963 (6). Although in the early years it was mainly referred as enzymatic activity of bacterial toxins, like diphtheria toxin, it then became clear that ADP-ribosylation is a reversible post-translational modification implicated in the regulation of several cellular processes, such as genomic stability (7), chromatin structure (8), transcription (6), circadian rhythm (9), and RNA metabolism (10). ADP-ribosylation is a very dynamic post-translational modification, whose high turnover is due to the presence of several "writers" and "erasers" **(Figure 1.3)**. ADP-ribosylation is very complex, since the modification can comprise of a single moiety up to 200-long-units chain, which can be linear or branched. So far, the single ADP-ribosyl-modification works in a distinct manner, when compared to the long-chain regulatory activity (11). However, the possible combinations for the long chain, such as the number of moieties in the chain and the presence of branching, may exponentially increase the coding abilities of this post-translational modification.

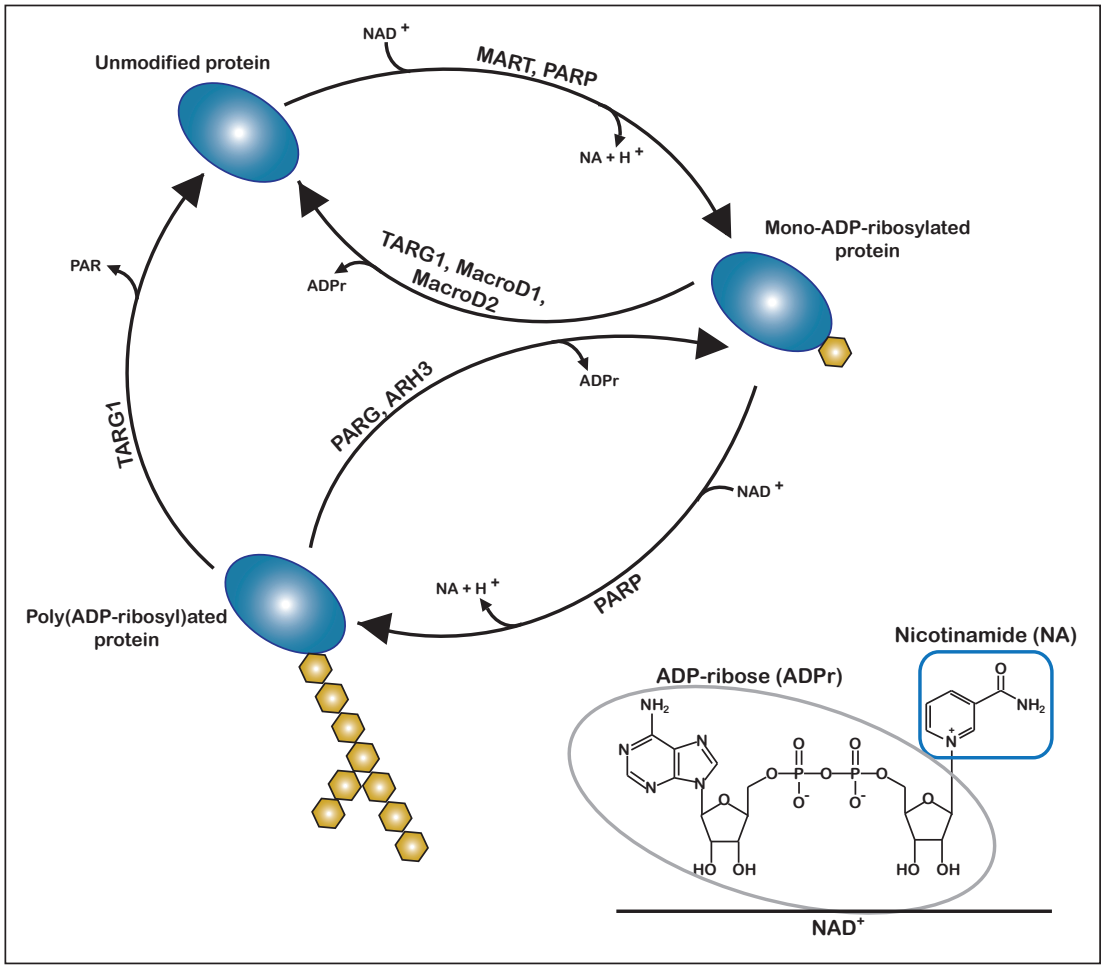


Figure 1.3: ADP-ribosylation shows a complex metabolism - Protein substrates are initially modified with the addition of one ADP-ribose moiety. This signal can be reversed or used as starting point for a long chain that is linear or branching. Several enzymatic activities ensure that the protein is reverted to the mono-ADP-ribosylated or unmodified state. Adapted from Barkauskaite et al. 2013 (12).

1.2.1 Human ADP-ribose transferases are diverse in structure and function

In humans, ADP-ribosylation is carried out by different classes of transferases. According to the recently proposed nomenclature, ADP-ribose transferases (ARTs) are divided into two groups (13), depending on which protein structure they adopt: the first and most studied group consists of the eighteen ADP-ribosyltransferases diphtheria toxin-like proteins (ARTDs), most of which were previously known as poly(ADP-ribose) polymerases (PARPs); the second group consists of the five C2 and C3 toxin-like ADP-rybosyltransferases (ARTCs). Besides the ARTDs and ARTCs, SIRT2, SIRT4 and SIRT6, belonging to the family of the de-acetylase sirtuins, were also reported to act as ADP-ribosyl-transferases (14, 15, 16, 17).

The ARTs transfer an ADP-ribose moiety from nicotinamide adenine dinucleotide (NAD^+) onto the target protein and release nicotinamide as by-product **(Figure 1.4;** (6)**)**. Many residues that accept ADP-ribose have been described so far. These are: glutamate (18), aspartate (18), lysine (19), arginine (20), asparagine (21), cysteine (22), diphthamide (23) and phospho-serine (24) residues.

The complexity of ADP-ribosylation signaling is due to the fact that the modification itself shows a certain degree of complexity. While most enzymes can only perform the transfer of one ADP-ribose moiety, and therefore are called mono(ADP-ribose) transferases (MARTs), some can attach other moieties to the first, granting the possibility to create a chain up to 200 units, thus the former name of PARPs (6). The chain can be linear, if the new moiety is attached to the adenine-proximal ribose unit through a $2'$-$1''$-*O*-glycosidic bond. Also, the PARPs can attach the new moiety at the nicotinamide-proximal ribose, thus creating a $2''$-$1'''$-*O*-glycosidic bond and forming a branch (25). The branching happens every 20-50 ADP-ribose units per chain in both *in vitro* and *in vivo* experiments (25, 26, 27, 28).

The eighteen human ARTDs are further classified according to their enzymatic activity **(Figure 1.5)**. ARTD1, ARTD2, ARTD5 (TNKS1) and ARTD6 (TNKS2) generate poly(ADP-ribose) (PAR) (11). The others can only transfer a single mono-ADP-ribose (MAR) onto target proteins, with the exception of ARTD9 and ARTD13 that are inactive. Additionally, the sirtuins SIRT2, SIRT4 and SIRT6 have been shown to be themselves mono-ADP-ribosyl-transferases (14, 15, 16, 17).

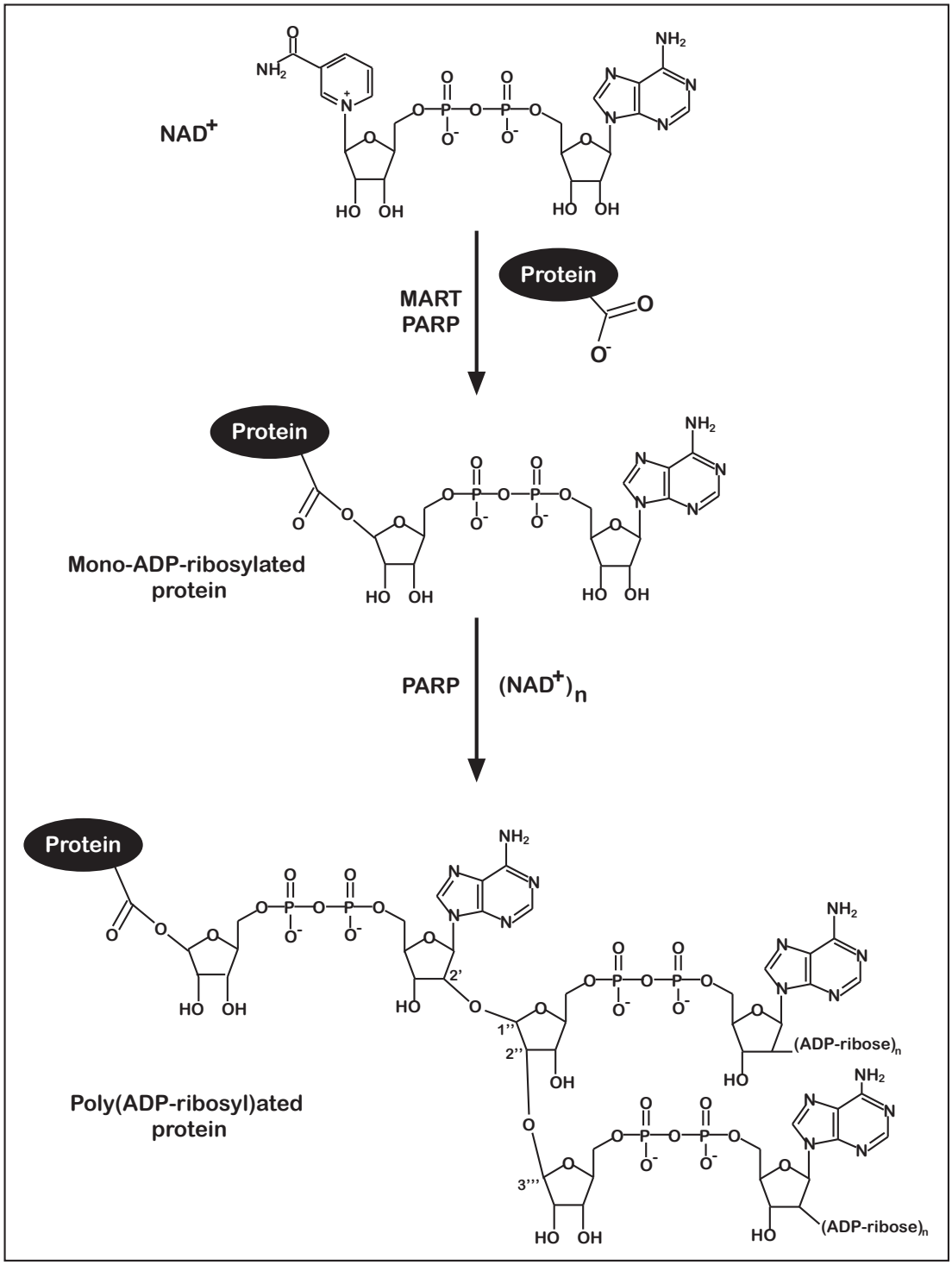


Figure 1.4: ADP-ribose can be attached in a linear or a branched fashion - MARTs or PARPs use NAD^+ as donor for both the MARylation and PARylation. For a linear chain, two ADP-ribose molecules are connected through a $2'$-$1''$-*O*-glycosidic bond. Every 20-50 ADP-ribose units, the branching of the chain takes place with the formation of a $2''$-$1'''$-*O*-glycosidic bond.

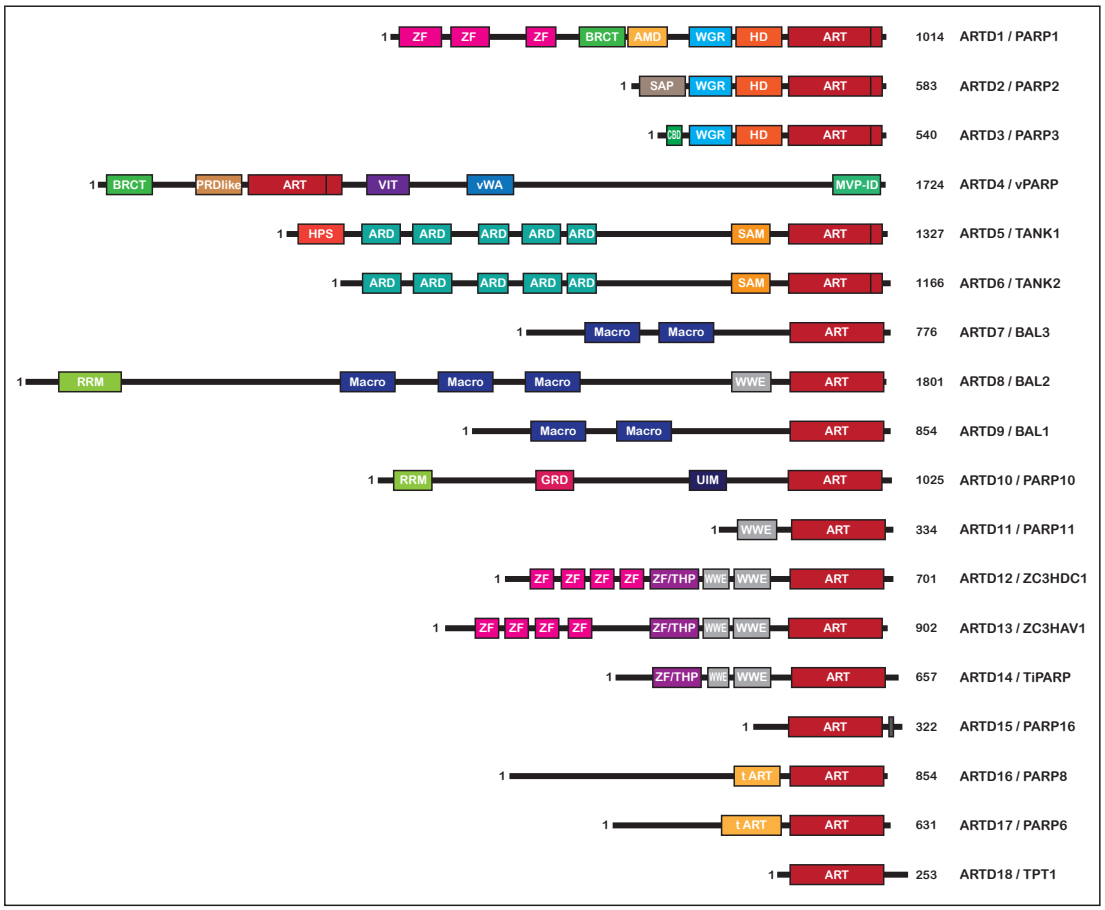


Figure 1.5: Domain architecture of the human ARTDs - List of the eighteen ARTDs with their domain composition. ART: catalytic core for ART activity; HD: linking helical subdomain; WGR: conserved central motif (W-G-R); BRCT: BRCA1 carboxy-terminal domain; SAM: sterile alpha motif; ARD: ankyrin repeat domains; VIT: vault protein inter-alpha-trypsin; vWA: von Willebrand type A; WWE: named after three conserved residues (W-W-E); Macro: macrodomain; ZF: zinc finger domains; SAP: SAF/Acinus/PIAS-DNA-binding domain; MVP-ID: Major-vault particle interaction domain; HPS: Histidine-proline-serine region; RRM: RNA-binding/recognition motif; UIM: ubiquitin interaction motif; TPH: Ti-PARP homologous domain. GRD: glycine-rich domain; CBD: caveolin-binding domain. Adapted from Hottiger et al., 2010 (13)

ARTD1/PARP1 is the most studied member of the ARTDs. ARTD1 a 116 kDa enzyme, is one of the few ARTs that is activated by binding specific DNA-damage related structures. ARTD1 is composed of many structural domains: two homologous zinc-finger domains (ZnF1 and ZnF2); a third zinc finger domain (ZnF3) diverged from the previous two; a BRCA1 C-terminus (BRCT) motif with the auto-modification domain; a WGR (Trp-Gly-Arg) motif; a catalytic domain containing the linking helical subdomain (HD) and the conserved catalytic signature (HYE) in the real catalytic fragment (CAT) **(Figure 1.6** (29)**)**. So far, crystallization of the whole protein has been elusive, but through crystallization of ARTD1 fragments it has been possible to define how the enzyme binds the damaged DNA and how the binding enhances the enzymatic activity (30, 31, 32, 33, 34, 35). According to several studies, the ZnF1, ZnF3, WGR and CAT domains are necessary for DNA-dependent activity on DSB (29). However, Ali et al., 2012 showed that *in vivo* recruitment of the ARTD1 fragment to DNA lesions is dependent on a functional ZnF2 upon UV-laser microirradiation (33). This observation is explained by the role of ZnF2 in the recognition of DNA single-strand-breaks, largely induced by UV laser (30).

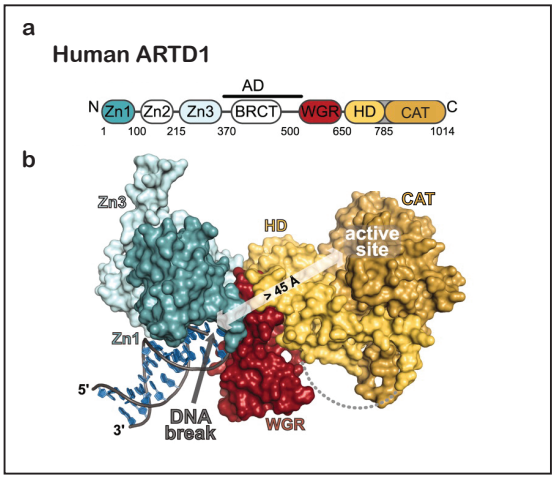


Figure 1.6: Domain architecture and activation of ARTD1 - ARTD1 is the best characterized ARTD. In *a*, the domain composition of the human ARTD1. In *b*, surface rapresentation of ARTD1/DNA structure. Adapted from Langelier et al., 2012 (34).

The most recent mechanism for ARTD1 activation suggests that the ZnF3 and the WGR domain relay the DNA damage binding to the catalytic domain (34, 35). When the ZnF domains bind a DNA lesion, a conformational switch in the WGR domain

makes the HD subdomain to move away from the CAT domain, which then increases its catalytic processivity. ARTD2 and ARTD3 have the same allosteric regulation between the HD and CAT domains, although the binding domain and the activating DNA molecule are different from ARTD1.

Since ARTD1 is so well studied, many examples are described, where ARTD1 activity is modulated by post-translational modifications from other enzymes: SET7/9-dependent methylation (36), SIRT6-induced MARylation (37, 38) and cyclin-dependent kinase 2 (CDK2)-dependent phosphorylation (39) are all able to induce ARTD1. ARTD1 is also phosphorylated by DNA-PK (40, 41) and by ataxia-telangiectasia-mutated (ATM) (42, 43, 44), although it is not clear if there is a functional role for these modifications.

Little is known about the structural details of the other ARTD enzymes (45). And even less is known about the structure of the MARTs. ARTD10 has been the first MART to be studied in detail and the definition of its catalytic mechanism suggested that the sequence requirement for the MARylating activity and for the polymerase activity are different (46).

The MART activity of the sirtuins has been described in a number of paper, but little is known about their functional role or their substrates. SIRT2 was the first human sirtuin found to have MART activity in *in vitro* experiments (14), although other enzymatic activities, such as deacetylation and demyristoylation, might be prevalent in the cellular environment (47). SIRT4 is localized in the mitochondria and has been shown to MARylate the glutamate dehydrogenase (GDH) and modulate in this way its enzymatic activity (16). Last but not least, SIRT6 has been shown to be recruited to DNA lesions and to MARylate ARTD1 (37, 38). This modification would increase the polymerase activity of ARTD1.

1.2.2 The ADP-ribose recognition modules relay the encoded message

Post-translational modifications are often recognized by specific recognition modules. So far, seven ADP-ribose recognition modules have been identified **(Figure 1.7** (8, 12)**)**. These are able to recognize specific or common features of MAR and PAR. The first identified were the linear PAR-binding motifs (PBMs) (48). Then, in rapid succession, more modules have been discovered: WWE domain (49), macrodomain (50) and PAR binding-zinc finger (PBZ) (51). The following domains to be discovered, FHAs and BRCTs, were previously associated to the recognition of phosphorylation.

These specific domains seems to have diverged from the canonical phosphorylation-binder FHA and BRCT domains and a combination of *in vitro* and *in vivo* assays showed for them a better affinity for ADP-ribosylation (52, 53). Similar story occurred for the single-strand DNA binding domain OB. The authors were able to show binding for the ADP-ribosylation in place of DNA molecules (54).

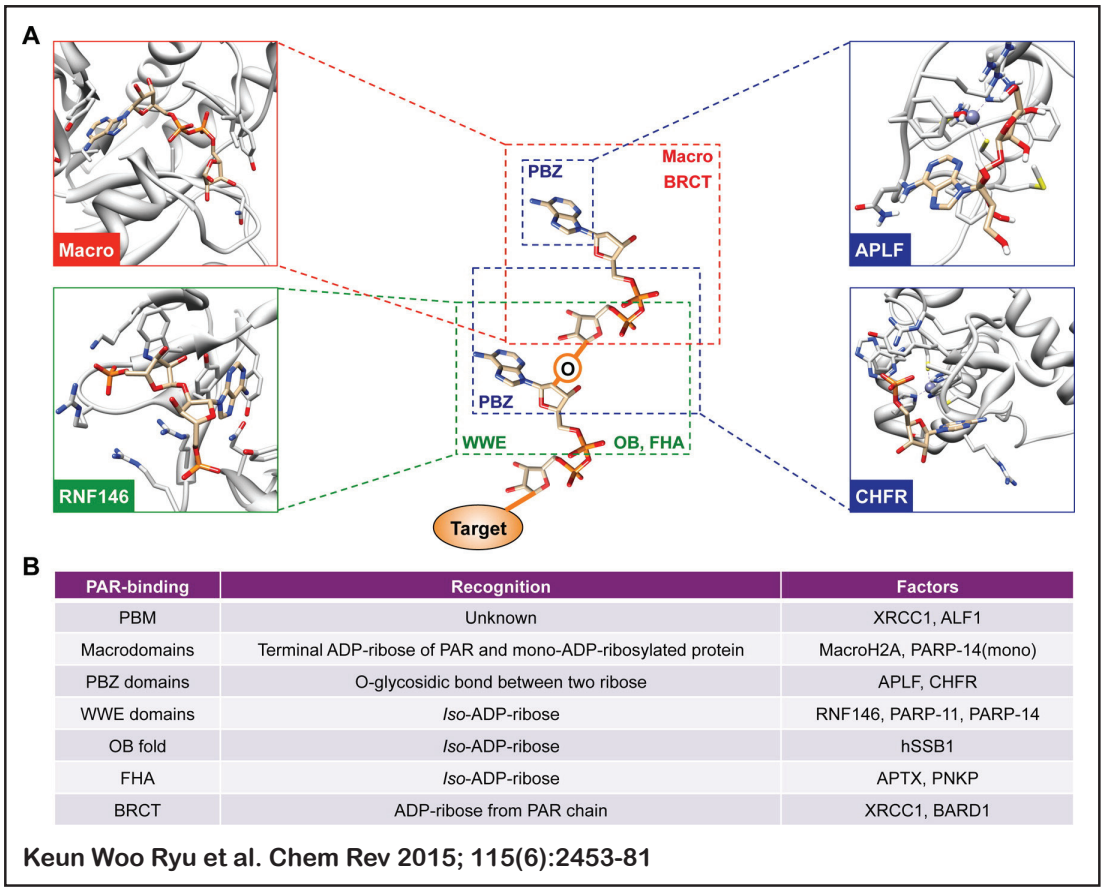
Being very diverse binding modules, they can recognize different features of the ADP-ribose molecule, as different immunoglobulins from a polyclonal antibody recognize several epitopes. For example, WWE domain binds at the linkage between two different ADP-ribose moieties, making it the perfect reader for PAR (55). On the other hand, macrodomain is binding the most terminal potion of the ADP-ribosyl-residue. Therefore, it can bind either the end of a PAR chain or a single MAR (56).

1.2.3 PARG and ARH3 degrades the poly-ADP-ribose into mono-ADP-ribose

In every type of signaling, the timely removal is essential for proper regulation. Therefore, in humans there are a number of enzymes that negatively regulate the levels of ADP-ribosylation, even though they are much fewer compared to the pool of the transferases (12).

The PAR glycohydrolase (PARG) and ADP-ribose hydrolase 3 (ARH3) degrade PAR by hydrolyzing the O-glycosidic ribose-ribose bonds and release ADP-ribose (57, 58). In particular, PARG removes preferentially one moiety at the time, but can also remove the whole chain all together (59). Conversely, the crystal structure of ARH3 suggests that it can only remove the last ADP-ribose of the chain (60, 61). However, they are ineffective on the protein-linked mono-ADP-ribose, so that their final products are mono-ADP-ribosylated proteins.

Surprisingly, the two proteins show different structures in the catalytic domain, since the first possess a catalytic macrodomain, while the latter belongs to the dinitrogenase reductase-activating glycohydrolase-related protein family (62, 63). PARG activity is essential to a proper regulation of ADP-ribosylation, due to the lethality at the embryo stage in $PARG^{-/-}$ mice and to the pronounced neural overproduction of PAR associated with neurodegeneration in *Drosophila* (64, 65). ARH3 activity is also important for proper cell function upon oxidative stress (66).



PAR-binding	Recognition	Factors
PBM	Unknown	XRCC1, ALF1
Macrodomains	Terminal ADP-ribose of PAR and mono-ADP-ribosylated protein	MacroH2A, PARP-14(mono)
PBZ domains	O-glycosidic bond between two ribose	APLF, CHFR
WWE domains	*Iso*-ADP-ribose	RNF146, PARP-11, PARP-14
OB fold	*Iso*-ADP-ribose	hSSB1
FHA	*Iso*-ADP-ribose	APTX, PNKP
BRCT	ADP-ribose from PAR chain	XRCC1, BARD1

Keun Woo Ryu et al. Chem Rev 2015; 115(6):2453-81

Figure 1.7: The different PAR-binding modules - Each PAR binding module can recognize distinct features of the PAR molecule. *a*) The PBZ domain (blue) uses a zinc-coordinated fold that recognise the 2″-1′-O-glycosidic bond between two ribose units, shown in the boxes with crystal structures of APLF and CHFR binding analogues. The macrodomain (red) binds the terminal ADP-ribose residue of PAR or mono-ADP-ribosylated protein. The WWE domain (green) recognize the iso-ADP-ribose residue. Human RNF146 WWE domain in complex with iso-ADP-ribose is shown as example. *b*) A table summarizing all the different PAR binding modules, the PAR element recognized and some example factors. From Ryu et al., 2015 (6).

Until recently, ADP-ribosylation was not a completely reversible modification, since the enzymes removing the most proximal ADP-ribose moiety attached to the acidic residues were still unknown. In 2013, three different papers characterized the enzymatic activity of MacroD1, MacroD2 and TARG1, the human mono-ADP-ribosylhydrolases (56, 67, 68).

1.2.4 MacroD1, MacroD2 and TARG1: the ultimate "erasers"

ADP-ribosyl-arginine hydrolases in animal species were already known since the middle 1980s (69). These activities, however, were only able to remove the final moiety from other amino acid residues a part from arginine. Therefore, the mono-ADP-ribosyl hydrolase activity acting on acidic residues was still missing. MacroD1, MacroD2 and TARG1 can reverse protein MARylation by hydrolyzing the ester bond between ADP-ribose and glutamate or aspartate (56, 67, 68). These three proteins belong to the same protein family, since they share the same catalytic domain, the macrodomain **(Figure 1.8)**. The same domain is present also in PARG, even though modified in key residues so that PARG cannot perform this last removal step **(see 1.2.3;** (62)**)**.

The macrodomain is a globular domain presenting a combination of 5 α-helices and 6 β-sheets **(Figure 1.9;** (50, 70)**)**. The macrodomain is a very conserved module, spread throughout the whole evolutionary arc, from archeobacteria, bacteria to vertebrates. It is also represented in some virus classes, arguing for a primordial role in the anti-viral response and a successively exchange between hosts and viruses (70). Also, the number of macrodomain-containing genes within an organism well correlates with its complexity, even though it is not always true. Among the model organisms in research, macrodomains are present in mouse, frog, fruit fly, chicken, zebrafish and roundworm, as well as many others. Remarkably, budding yeast has not ARTDs but two macrodomains, probably successively reintroduced (70).

The macrodomain has been first found as an unknown sequence of the histone variant MacroH2A (71). Independently, by an unbiased genomic approach in yeast, it was found that this same sequence could hydrolyze a byproduct of tRNA splicing, the ADP-ribose-1″-phosphate (72). But at the time there was no connection with MacroH2A, as well as no crystal structure nor protein family. In 2005 the crystal structure of the macrodomain from the *Archaeoglobus fulgidus* protein Afl1521 was published, matching the folding of the previously unknown sequence with the yeast

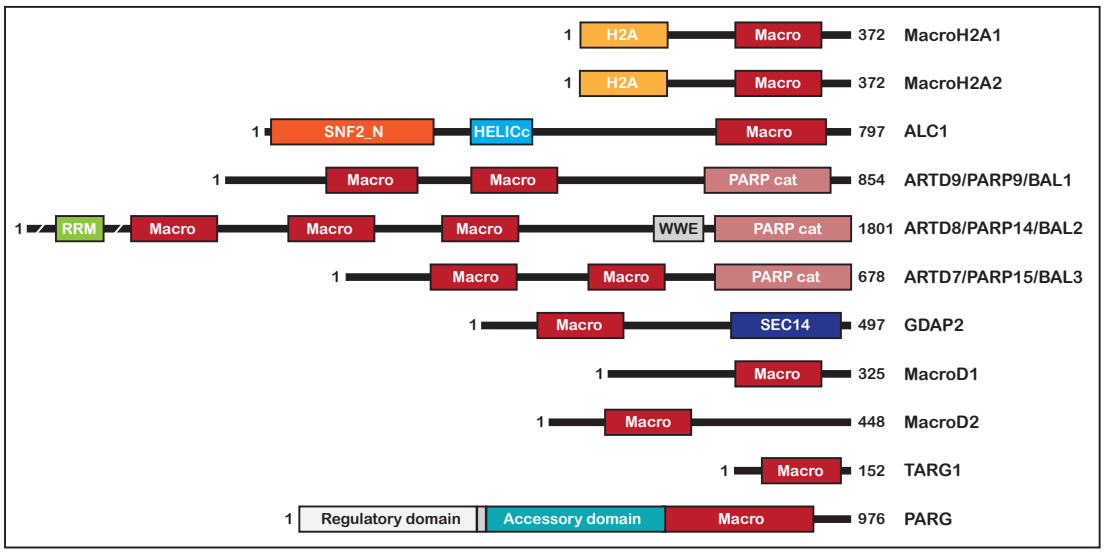


Figure 1.8: Macrodomain family members in humans - Domain structure of the human macrodomain proteins. Domain abbreviations: H2A, core histone H2A-like domain; HELICc, helicase conserved C-terminal domain; Macro, macrodomain; PARP cat, poly(ADP-ribose) polymerase catalytic domain; RRM, RNA recognition motif; SEC14, named after yeast SECretory protein 14 (Sec14p); SNF2-N, SNF2 helicase family N-terminal domain; WWE, domain is named after a conserved tryptophan/glutamate-containing motif. Adapted from Rack et al., 2016 (70).

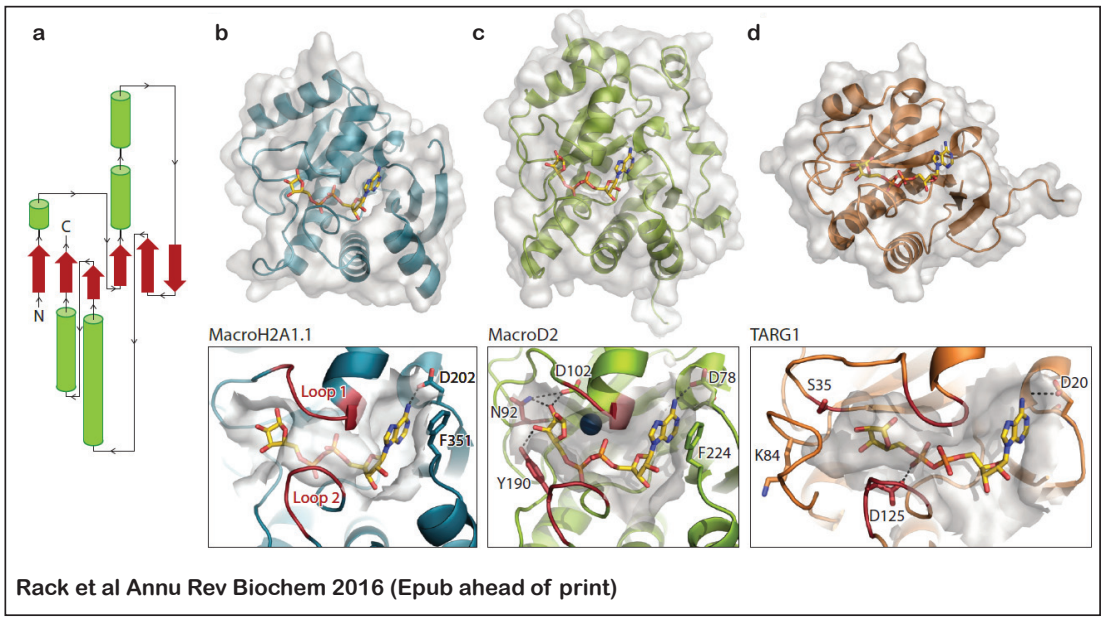


Figure 1.9: How the macrodomain folds and binds to the ADP-ribose - *a*)The topological representation of the macrodomain shows the organization of the central six-stranded -sheet (red) flanked on both sides by five α-helices (green). However, some macrodomains have additional α-helices or β-strands. *b-d*) The ribbon and surface representations of three human macrodomains, MacroH2A1.1 (*b*), MacroD2 (*c*) and TARG1 (*d*). The structures show the binding of ADP-ribose (*b*,*c*) or the lysyl-ADPr intermediate (*d*). In the magnifications, details on the coordination of the adenosine moiety by a conserved phenylalanine and/or asparagine residue as well as by the substrate-binding loops 1 and 2 (red) is depicted in the magnifications. From Rack et al., 2016 (70).

protein (50). Along with the folding, the macrodomain was identified as a ADP-ribose-1″-phosphate hydrolase, as well as a PAR-binding module. MacroH2A was next shown to bind O-acetyl-ADP-ribose, a byproduct of Sirtuins de-acetylase activity, and that this ability was lost according to the splicing isoform of MacroH2A (73).

It took some time to show that macrodomains from the human MacroD1 and MacroD2 were able to hydrolyze O-acetyl-ADP-ribose as well, suggesting a role of macrodomain proteins in the regulation of histone acetylation (74). In the meantime, the crystal structure and the O-acyl-ADP-ribose hydrolase activity of C6orf130/TARG1 were confirmed (75). Two years later, several groups finally showed the mono-ADP-ribose hydrolase activity of MacroD1, MacroD2 and TARG1 (56, 67, 68).

The three mono-ADP-ribose hydrolases share the macrodomain structure. But they greatly differ in other elements. TARG1 is the smallest of the three proteins, only 17 kDa. It has most of the protein sequence overlapping with the macrodomain and it is more related to the ALC1-like folding than to the MacroD protein folding (70). When tagged with EGFP, TARG1 is found in the whole cell, therefore it seems there is no restriction in its localization (68). The small size should not affect the free movement of TARG1 within the cell. The catalytic activity of TARG1 requires the formation of a stable intermediate of the target protein bound to the enzyme. This mode of action is slightly different from the other two enzymes of the group, consistent with the fact that TARG1 is in a different phylogenic branch (68, 70, 75). The only described interactors of TARG1 are ARTD1 and ARTD10, so far (68). Also, TARG1 is shown to recruit to DNA damage sites upon UV-laser microirradiation and it is dependent on the ADP-ribose binding ability of the macrodomain. TARG1 activity is associated with neurodegeneration (68), similarly to PARG (65). Thus, the regulation of ADP-ribosylation appears to be particularly necessary in the brain.

MacroD1 is a slightly bigger protein, with 35.5 kDa molecular weight. Aside the macrodomain, MacroD1 presents an N-terminal sequence with a mitochondrial targeting sequence (MTS) (76). Its mitochondrial location suggests that the MacroD1 enzymatic substrates could be all mitochondrial proteins. Studies have shown that MARylation occurs in mitochondria: for example, SIRT4 negatively regulates GAPDH activity by MARylation (16). Also, one of PARG isoforms presents a MTS (12). Therefore, regulation of mitochondrial processes could be mediated by ADP-ribosylation. However, *in vitro* tests on the MAR-hydrolase activity of MacroD1 were performed on

ARTD1 and ARTD10, whose main location is not in the mitochondria. Additionally, the work of W. Han group shows that MacroD1 (called by them LRP16) is co-activator of the estrogen receptor α and androgen receptor, as well as being involved in the NF-κB regulation (77, 78, 79). It is still unknown how these functional interactions can be consistent with the mitochondrial localization.

The catalytic mechanism proposed for MacroD1 and MacroD2 is different compared to the mechanism of TARG1. MacroD1 and MacroD2 are proposed to perform a substrate-assisted catalysis (56). The macrodomain would adapt its conformation to better accommodate the substrate, so that the pyrophosphate of the ADP-ribose would activate a water molecule for a nucleophilic attack against the glutamate of the substrate, leading to the release of the ADP-ribose.

1.2.5 The elusive nature of MacroD2

MacroD2 is the third member of the macrodomain mono-ADP-ribosyl hydrolases. MacroD2 is the biggest of the three, with 448 amino acids length and 50 kDa molecular weight. As the other two members, it has a macrodomain that is located in the N-terminal portion of the protein **(see Uniprot page** `http://www.uniprot.org/uniprot/A1Z1Q3`**)**. The C-terminal portion is instead not annotated and includes a compositional biased sequence for glutamic acid between the aa 247 and 388. In the evolution of macrodomains, MacroD1 and MacroD2 functions were probably performed by one gene only, until when there has been a duplication event at the beginning of the vertebrate clade (80).

Previous molecular studies on MacroD2 have focused on the macrodomain alone (56, 67). In fact, before my PhD, I worked on the recruitment of MacroD1, MacroD2 and TARG1, as shown in Jankevicius et al., 2013 and Sharifi et al., 2013 (56, 68). Thus, I showed that the MacroD2 macrodomain is able to recruit to the DNA damage sites upon laser microirradiation, as it is also true for other macrodomains, namely MacroH2A1.1 variant, MacroD1 and TARG1 (56, 68, 81). The recruitment of MacroD2 was dependent on ADP-ribosylation, since both ARTD inhibition and mutation in the MacroD2 ADP-ribose-binding pocket (MacroD2 G188E mutant) prevent MacroD2 from localizing at DNA lesions (56). Also, the recruitment of MacroD2 was different from MacroH2A1.1, a sole reader of the mark. In fact, while MacroH2A1.1 recruited in a sharp peak, MacroD2 showed a bimodal curve **(Figure 1.10)**. These results suggested

that MacroD2 recruited in a first peak, but only upon conversion of the PAR into MAR by PARG it was able to reside longer at the DNA damage site. To confirm the hypothesis, when I depleted PARG by RNAi-mediated knock-down, the first peak was lost and the overall recruitment was strongly decreased (56).

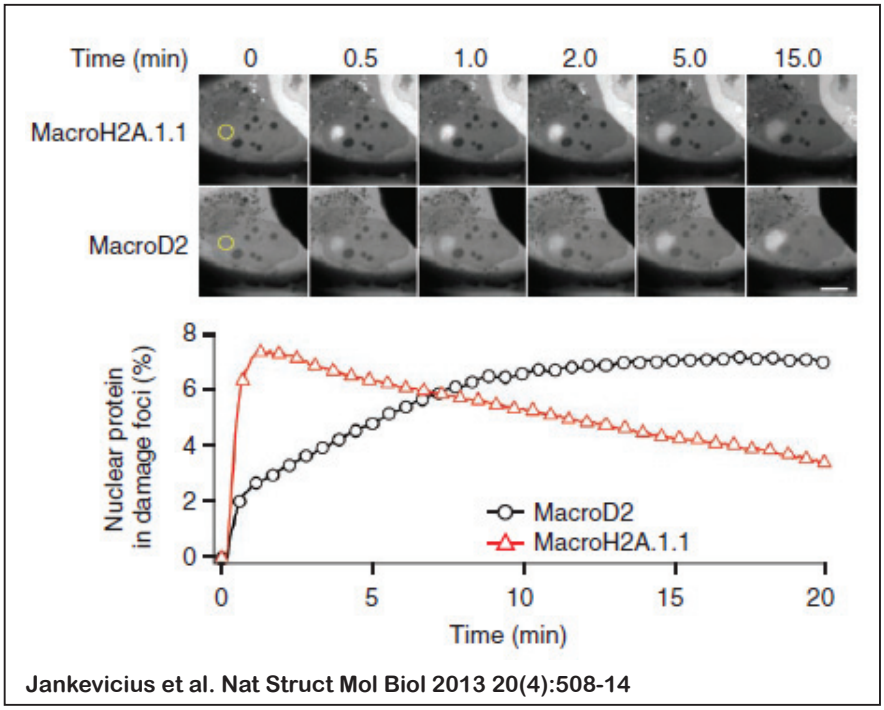


Figure 1.10: MacroD2 macrodomain recruits to DNA lesions - Recruitment of EGFP-MacroD2 and mCherry-MacroH2A.1.1 macrodomains to sites of laser-induced DNA damage. The focus of laser microirradiation is indicated with a yellow circle. Scale bar, 10 μm. Quantification of accumulation to the damage site. From Jankevicius et al., 2013 (56).

The characterization of MacroD2 macrodomain activity defined as possible substrates ARTD1, ARTD10 and Histone H1 **(Figure 1.11;** (56, 67)**)**. As mentioned before, MacroD2 enzymatic activity adopts the substrate-assisted catalyst mechanism described for MacroD1 (56).

Additionally, MacroD2 successfully removes the modification from glycogen synthase kinase β (GSK3β) both in *in vitro* and *in vivo* experiments **(Figure 1.11; (67))**. GSK3β is an enzyme involved in the regulation of several pathways, involved in proliferation and apoptosis, cell morphology and motility, as well as to diabetes and several

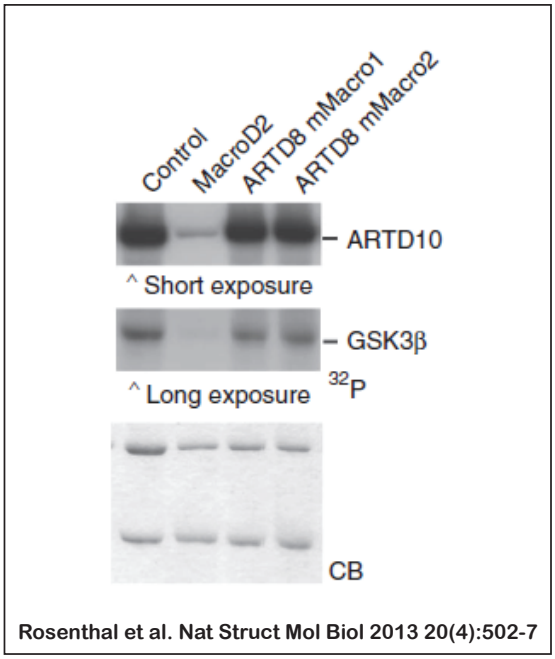


Figure 1.11: MacroD2 removes MAR from ARTD10 and GSK3β *in vitro* - Coomassie blue (CB) staining and autoradiography (32P) results of *in vitro* assays with tandem-affinity purification (TAP)-tagged ARTD10 and GST-GSK3β coupled to beads. MacroD2 macrodomain removes MAR from both purified ARTD10 and GSK3β, while the macrodomains 1 and 2 from mouse ARTD8 do not. From Rosenthal et al., 2013 (67).

neuropathologies (82, 83). GSK3β has been shown to be *in vitro* and *in vivo* MARylated by ARTD10 and this modification results in the decreased activity of the kinase (84). Thus, ARTD10 and MacroD2 could represent an independent pathway, important for the fine-tuning of GSK3β activation levels. But more experiments awaits to test this hypothesis.

Overall, MacroD2 cellular function is still unclear. There are many hints suggesting MacroD2 to be involved in many processes, but no real evidence. For example, initial MacroD2 association with Kabuki syndrome has been proved wrong (85, 86). Still with some controversy, MacroD2 is often associated with autism syndrome, so that it has also been referred as one of the autism-associated prototypical genes in genome-wide association studies (87, 88, 89, 90, 91, 92, 93). In addition, MacroD2 has been associated to other neurological related diseases, as attention-deficit hyperactivity disorder (ADHD) or major depressive disorder (89, 94, 95). Other association studies were linking MacroD2 to the temporal lobe volume (96) and to brain connectivity in patients with neurological and psychiatric disorders (97). Therefore, MacroD2 function might have an important role in the development of brain functions, as well as in their protection. As confirm, the MacroD2 knockout mouse shows a number of behavioral and neurological phenotypes (International Mouse Phenotyping Consortium), firmly indicating the functional role of MacroD2 in the brain. What would be its molecular role in the brain is still to define.

MacroD2 overexpression was shown to induce tamoxifen resistance in estrogen receptor α positive breast cancers (98). Considering the abovementioned interaction between MacroD1 and the same pathway (77), the authors did actually test and exclude that MacroD1 overexpression is involved in the formation of the resistance. Multiple studies found the deletion of MacroD2 gene in association with colorectal cancer, suggesting that MacroD2 could be a tumor-suppressor gene (99, 100, 101, 102). However, it has also been proposed that the MacroD2 gene locus is fragile (94, 103), thus maybe erroneously found in diverse genome-wide association studies (94, 104, 105, 106, 107, 108, 109). In conclusion, if these many phenotypes associated with MacroD2 indicate a true relationship and not a coincidence, it is still unknown. And how MacroD2 enzymatic activity could explain such phenotypes, is still an open question.

1.2.6 ADP-ribosylation regulates several aspects of DNA damage repair

Since its discovery, most efforts in studying ADP-ribosylation signaling function have focused on its involvement in DNA damage response (DDR) **(Figure 1.12;** (8)**)**. DNA-depentent ARTDs, namely ARTD1, ARTD2 and ARTD3, are differentially involved in this response, even though most studies characterize just ARTD1 activity in the different DDR pathways (110).

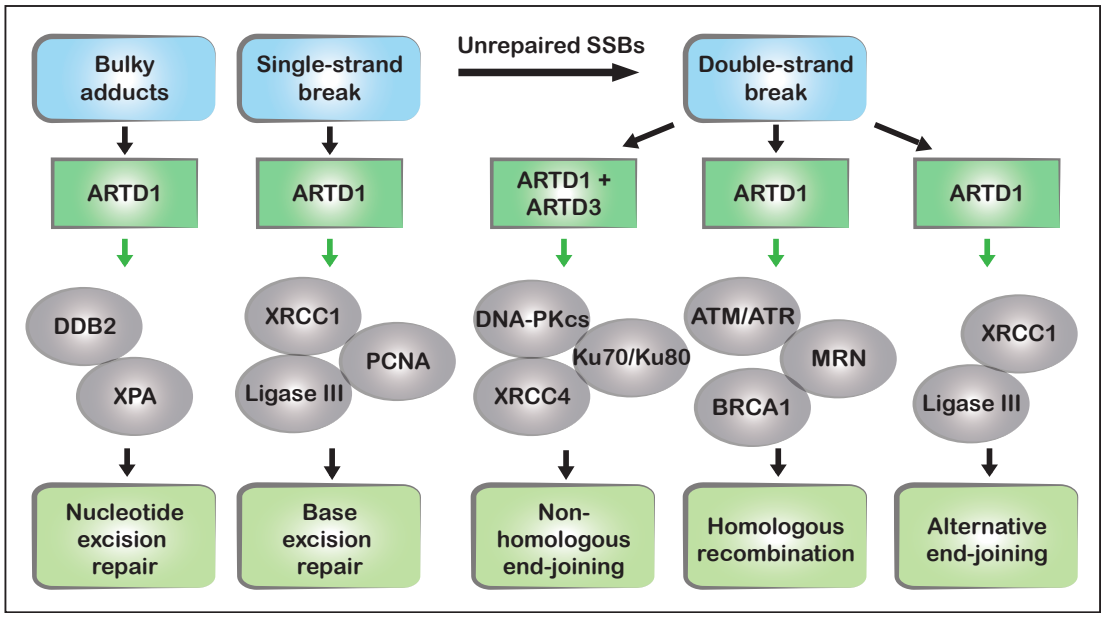


Figure 1.12: ADP-ribosylation is involved in several DNA repair pathways - Different types of DNA damage types activares ARTD1 and others transferases. In grey, some factors that are activated by ADP-ribosylation or recruited to DNA lesions in a PAR-dependent manner. In the lower boxes, the output repair pathways that are influenced by ADP-ribosylation signalling.

Many different types of lesions occur on DNA. Single-strand breaks (SSB) arises either because of endogenous factors, like reactive oxygen species (ROS), or because of failure in the enzymatic activity of type I topoisomerases (111). These enzymes are involved in DNA replication and transcription and their role consists in allowing the denaturation of the DNA double strand, in order to access to the information kept within: thus, type I topoisomerases nick the sugar-phosphate backbone of one strand while creating the covalent intermediate of a phospho-tyrosine. As the whole

DNA molecule has relaxed, adjusting its physical tensions, the topoisomerases ligate the ends back together. But when there is an abortive ligation or a malfunction in topoisomerase I activity, the DNA molecule gets an unresolved SSB.

SSBs are also formed as natural process in the repair of other types of lesions. Oxidative damage, alkylation, methylation, deamination or hydroxylation of a DNA base are recognized as lesions, since they could interfere with the DNA- or RNA-polymerase processing and create a base-pairing mismatch. For this reason, the base excision repair pathway (BER) is initialized. The BER starts with the detection and excision of the damaged base by DNA glycosylases, which remove from one to eight nucleotides of the same strand and generate a SSB (112, 113).

Even if ARTD2 is also involved to SSB repair (114), ARTD1 has been described as an important factor in the induction of BER response, although its actual contribution to the repair efficiency is controversial (115, 116). Depletion of ARTD1 affects repair of lesions derived from uracil or 8-oxoguanine, inducers of abasic sites. Also ARTD1 and PARG depletion were relevant in SSB repair upon treatment with H_2O_2 in chicken DT40 or A549 human lung carcinoma cells (117). The importance of ARTD1 activity is probably due to its function in recruiting downstream BER factors due to the formation of the PAR scaffold. In fact, many proteins involved in the pathways have been found enriched in PAR-binding proteins or as ADP-ribosylated proteins (53, 118, 119, 120, 121, 122, 123, 124, 125, 126). Therefore, depletion of ARTD1 is affecting the speed of the repair process but not its successful completion (127, 128, 129).

Other types of DNA lesions that require a more invasive repair pathway, the nucleotide excision repair (NER), do still belong to the SSBs: cyclobutane pyrimidine dimers (CPDs), (6-4) pyrimidine-pyrimidones photoproducts (6-4PPs) are induced by UV-C light and consists of bulky adducts distorting the DNA molecule, so that the entire nucleotide is required to be exchange (130). These DNA lesions are shown to be strong activators of ADP-ribosylation, arguing for its role in the NER pathway (131). ARTD1 interacts with damaged DNA binding 2 (DDB2), the factor that recognizes the malignant photoproducts (130, 132, 133, 134). ARTD1 activity on one side helps the rearrangement of the chromatin by the recruitment of the chromatin remodeler ALC1. On the other side, PARylation helps DDB2 to retain at the DNA lesions, although it is not required for its initial recruitment (133). ARTD1 strongly interacts also with xeroderma pigmentosum-complementation group A (XPA), another factor for one NER

sub-pathway (48, 135, 136). ARTD1 activity is important for the recruitment of XPA to DNA lesions, but the interaction of the two factors might provoke the release of both ARTD1 and XPA from the DNA.

Although less frequent than the SSBs, the double-strand break (DBS) is the most dangerous type of lesion, since it can generate important rearrangements in the genome if not repaired in time (137). DSBs can have an endogenous origin, like in case of collapse of DNA replication fork or in case of cell-specific genome editing, as in meiosis in the gamete formation or in V(D)J recombination and class-switch recombination for the immune cell maturation (138). In addition, exogenous factors can also induce DSBs, such as ionizing radiation, radiomimetic chemicals and type II topoisomerase inhibitors (137, 138, 139).

DSBs are repaired by a variety of pathways, whose activation is defined in the very early onset of the response. The two most studied DSB repair pathways are the homologous recombination (HR) and the classical non-homologous end-joining (c-NHEJ). But other pathways can also take place, like the single-strand annealing (SSA) and the alternative end-joining (alt-EJ) (140). The c-NHEJ pathway consists of the fast re-ligation of the broken DNA ends. It occurs throughout the whole cell-cycle but it is the preferential repair in G0/G1 and late G2 phase (141). Even if c-NHEJ inhibits big translocations and chromosome rearrangements, it brings micro-mutations at the re-ligation point, defining it as an error-prone pathway (142, 143, 144, 145). The main factors involved in this pathway are the DNA-dependent protein kinase catalytic subunit (DNA-PKcs), Ku70 and Ku80, XRCC4 and the Ligase IV (140).

If the broken DNA ends are immediately resected and present 3 single-strand overhangs, c-NHEJ is blocked and one of the other three pathways can take place. The homology-driven repair, another name for the HR, is generally activated after replication, because it requires a homologous sequence for the restoration of the original information. Since the most preferred homologous sequence is the sister chromatid, HR occurrence peaks in mid-S phase (141). Main factors of this slow pathway are the replication protein A (RPA), which avoids degradation of DNA single strand by coating it, and the recombinase RAD51.

The last two pathways, SSA and alt-EJ, are highly mutagenic and also required a different degree of DNA resection. SSA is a homology-driven repair process exploiting nucleotide repeats (140). Since one of the repeats is deleted during the repair, the loss

of genetic information makes it an error-prone pathway. The interaction between RPA and RAD52 allow the repair by SSA. Lastly, the alt-EJ was at first described as a backup pathway for the c-NHEJ. The alt-EJ employs a different set of factors: in fact, the repair is mediated by most of the factors that works in the BER pathway, such as ARTD1, XRCC1 and Ligase III (138).

ARTD1 recognizes DSBs and its clear that ADP-ribosylation is interacting with different DSB repair pathways, even though the underlying mechanism is not totally clear (8, 29). Also, the study of ARTD1 for DSB repair is tightly connected with the story of the PARP inhibitors (PARPi) (146). In 2005 two studies in parallel showed that inhibition of ARTD1 in cancer cells defected in BRCA genes − genes involved mainly in HR completion, causing these cell lines to be called *BRCAness* cells − leads to their death (147, 148). These studies envisioned for the first time a synthetic-lethality strategy for treatment in cancer with tumor-suppressor gene deficiency. Since then, more and more sophisticated PARPi molecules were designed and tested mainly in ovarian and breast cancer patients with *BRCAness* background, alone or in combination with other cytotoxic compounds (146). Unfortunately, it is not clear what is the real meaning of this synthetic lethality. More recent studies suggest that PARPi induces a stalled ARTD1-DNA intermediate that can be processed only by HR pathway (149, 150).

ARTD1 negative regulates HR by counteracting the accumulation of RAD51 (151). Such an activity of ARTD1 affects the sister chromatid exchange (152), the intra- and the extrachromosomal recombination (151, 153, 154). However, some proteins involved in HR are found enriched in pull-downs of PARylated proteins in the presence of genotoxic stress (118, 119). Even more factors present a PAR-binding module, suggesting that their recruitment could be dependent on ARTDs activity, at least at some stage of the damage response (51, 52, 53, 54, 118, 119, 155, 156, 157, 158, 159).

A similar concept is true also for the c-NHEJ pathway: some studies showed interaction between repair factors and PARylation, mainly because of the presence of PAR-binding modules (48, 53, 118, 119). Because of their PBMs, DNA-PKcs and Ku70 recruit to DNA damage in a PAR-dependent manner (48). They form a complex together with ARTD1 at the DNA lesion (40, 41, 48, 160, 161, 162, 163). However, it is not clear what are the effects of such an interaction. Maybe the PARylation of DNA-PK stimulates the kinase activity (41, 160). Similarly, the DNA-PK-mediated

phosphorylation of ARTD1 fails to have a clear effect on the activity of the ADP-ribose trasferase (40, 41). However, in many studies, ARTD1 and DNA-PK activities cooperates, for example in double knock-out mice for DNA repair or in the case of V(D)J recombination in B cells (164, 165, 166). Also, DNA-PK and ARTD1 interaction is shown to be relevant in the unresected stalled DNA-replication fork (167), as well as in the inhibition of the ribosomal RNA synthesis (168).

But the most important role of ARTD1 in DNA damage response is associated with the alt-EJ pathway (138). First defined as backup pathway, it has now emancipated as an alternative pathway, whose activation is mainly due to the location of the DNA damage and the poor accessibility of other pathways factors. In fact, the alt-EJ is activated in the lack of c-NHEJ factors, like the Ku proteins (162), given the shear competition that ARTD1 and the Ku proteins have for the DNA ends (169, 170, 171, 172). The alt-EJ is particularly prone to cause translocations and major genomic rearrangements and requires XRCC1 and Ligase III for the accomplishment of the repair (173, 174).

In conclusion, ADP-ribosylation has definitely a role in the early phase of the response, helping the general recruitment of factors involved in different DSB repair pathways. However, it might be that the location of the DNA lesion in the genome might preferentially activate different pathways and, similarly, the involvement of ADP-ribosylation might change accordingly. As last comment, the role of MARylation in all DNA damage responses is still obscure. For example, ARTD10 interacts with PCNA and ARTD10-depletion induced sensitivity to hydroxyurea and UV, both inducing stalled replication forks (175). But we still miss a complete picture of the involvement of MARylation in DNA damage response.

1.2.7 ADP-ribosylation regulates also other cellular processes

Beyond the DDR, the eighteen human ARTDs are collectively involved in almost every cellular process. In a recent study that compares the location and the knock-down-induced effects on morphology of each ARTD, the authors showed, although just in a macroscopic way, how ADP-ribosylation influences different aspects of cellular life, ranging from cell cycle defects to cytoskeletal defects (176). Among the eighteen ARTDs, ARTD1 is the most studied, above all in relation to DNA damage repair. But

fragmented studies show the involvement of ARTD1 and many other transferases in a vast variety of cellular processes.

For example, many studies involving ARTD1 explore its function in gene expression, as modulator of chromatin compaction status. Since the early time of ADP-ribosylation field, PARylation on histones was shown to relax the chromatin and make it more accessible (177, 178). In fact, the generation of a PAR chain, long and rich in negative charges, is bound to have dramatic effects on the activity of the target protein or its affinity for other molecules. Thus, the modification and consequent eviction of H1 is the most favorable explanation for the chromatin relaxation (179). The eviction process is shown also for other chromatin factors, like the histone chaperone complex FACT or the histone demethylase KDM5B (180, 181). But PAR formation on chromatin induces also the recruitment of other proteins, such as the chromatin remodelers (8). ALC1 has a macrodomain that allows the recruitment to DNA and, if bound to PAR, induces the activation of the remodeling activity **(Figure 1.13,1;** (182)**)**. PARylation is also inducing the recruitment of SMARCA5 and CHD4, two major components of other chromatin remodeling complex, respectively the ISWI family complexes and the NuRD complex (157, 183, 184). Of course, by regulating the accessibility of different factor to the genomic information, ARTD1 holds the key for the proper regulation of almost all processes that take place in the cell.

Other studies also focusing on ARTD1 showed that PARylation on a target protein can induce the recruitment of PAR-binding E3 ubiquitin ligases, so that ADP-ribosylation can be read as mark for protein degradation. RNF146/Iduna and CHFR, two E3 ligases, have a WWE and a PBZ domain, respectively (159, 185). Once recruited, the E3 ligases ubiquitylate the PARylated target and mark it for degradation. It is not clear if all the PARylated proteins are actually target of this degradation pathway. It might be that the E3 ligases require an additional signal to activate their ligase activity. However, this mechanism has been shown for axin, a component of the Wnt signaling (186), mediated by ARTD5 activity **(Figure 1.13,2;** (187)**)**, as well as for ARTD1 itself, to avoid DNA damage-induced hyperactivation **(Figure 1.13,3;** (185, 188)**)**. Hyperactivation of ARTD1 and PARylation are connected to a particular form of caspase-independent programmed cell death, thus called *parthanatos* (189). The *parthanatos* is based on the release of free PAR chains, which transmits as sec-

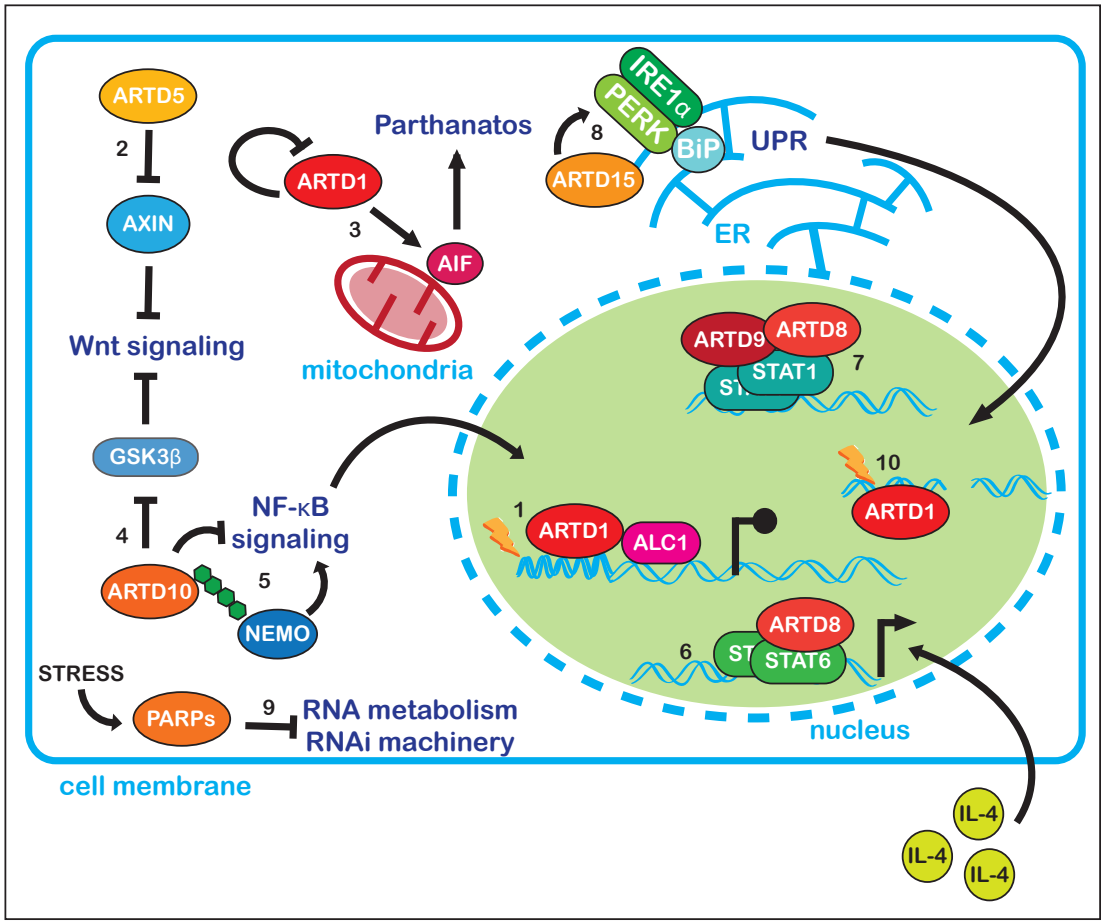


Figure 1.13: The many roles of ADP-ribosylation in the cell - MARylation and PARylation have been connected with several signaling pathways and cell processes. An overview on some described functional interactions: *1*) upon DNA damage, ARTD1 activated recruits chromatin remodelers, like ALC1, for relaxation of the chromatin and modulation of gene expression and repair; *2*) ARTD5 release the negative feedback loop of Wnt signaling activation by PARylating axin, negative regulator of the pathway, and targeting it to degradation; *3*) ARTD1 targets itself to degradation by PARylation-mediated ubiquitilation and prevents its own hyper-activation, which brings to the AIF-mediated cell-death, also known as *parthanatos*; *4*) ARTD10 MARylates GSK3β, which normally inhibits the activation of Wnt signaling by targeting β-catenin for degradation; *5*) ARTD10 also inhibits NF-κB signaling by affecting the positive signaling mediated by NEMO; *6*) upon IL-4 cytokine signaling, ARTD8 activity co-activates STAT6-responding genes; *7*) ARTD9, maybe together with ARTD8, affects STAT1-dependent gene expression; *8*) ARTD15 controls the unfolded protein response (UPR) by modifying IRE1α and PERK; *9*) when different types of stress arise, PARylation is generally involved in the modulation of RNA metabolism and temporary block of RNA interference; *10*) PARylation, mediated for example by ARTD1, modulates different DNA damage repair pathways. Adapted from Bütepage et al., 2015 (11)

ondary messenger the death signal to the mitochondria and leads to the release of the apoptosis-inducing factor (AIF) (190)

Apart from the omnipresent ARTD1, other transferases were found to be involved in interesting cellular processes. ARTD10, for example, might be involved in the regulation of the Wnt signaling, since it modifies the glycogen synthase kinase 3 β (GSK3β) *in vitro* (84). In normal condition, the GSK3β kinase is assembled in a complex located at the cellular membrane and its activity results in the degradation of β-catenin, the activating transcription factors for the Wnt-responsive genes **(Figure 1.13,4)**. When Wnt molecule binds the receptor outside the cellular membrane, β-catenin is not phosphorylated, thus not degraded, and can activate the Wnt-responsive genes, which lead to proliferation. Therefore, if MARylation is able to decrease the activity of GSK3β, the β-catenin would consequently be stabilized. Unfortunately, the authors of the previous studies did not show if the MARylation of GSK3β was actually affecting Wnt signaling. However, *in vitro* and *in vivo* experiments indicated GSK3β as one of the few substrates of MacroD2 activity (67), further suggesting a functional role of the MARylation for the regulation of this enzyme. Nonetheless, the regulation of β-catenin protein level is just one of the functions ascribed to GSK3β: for example, as the name of the protein suggests, it regulates the glycogen synthesis, therefore this could open a possible connection between MARylation and metabolism (191, 192).

ARTD10 has been also shown to negatively regulate the activation of NF-κB pathway, which activates pro-inflammatory genes in response to cytokines like IL-1β and TNFα or genotoxic stress **(Figure 1.13,5;** (11, 193)**)**. The study performed by Lüscher group showed that ARTD10 can modifies one of the positive regulators of the pathway, NF-κB essential modulator (NEMO) (194). This interaction interferes with the K36-polyubiquitination of NEMO, which helps maintain the signaling induced. The result is the renewed inhibition of NF-κB-dependent transcriptional activation. Thus, ARTD10 and MARylation could represent a checkpoint for the total activation of the inflammatory response. On the other hand, ARTD1 activity has been shown to help NF-κB activation by promoting NEMO nuclear export, as well as by acting as co-activator of the NF-κB transcription factors at the promoters (195, 196, 197). This opposite effect between MARylation and PARylation shows that MARylation is not just a step of the formation of PAR chains, but a real independent signal.

Other transferases are also involved in the regulation of signaling pathways. Few studies have described how ARTD8/PARP15/BAL2 co-activates with STAT6 the expression of specific cytokines, which are important for the humoral immunity **(Figure 1.13,6;** (198, 199, 200, 201)**)**. The macrodomains of ARTD8 are necessary for the interaction with STAT6, but the exact reason is not yet clear (199, 200). Also, the MART activity is required for ARTD8 function (198, 200). In fact, ARTD8 is proposed to function as a *transcriptional switch*, by repressing or activating the expression according to the interaction with STAT6 (201).

The inactive ARTD9/PARP9/BAL1, instead, is involved in another immune response-related pathway, the interferone-γ (INFγ)-dependent STAT1 pathway **(Figure 1.13,7;** (202, 203)**)**. In Camicia et al., 2013, the authors state (but not show) that STAT1 is modified by ARTD8 and ARTD10 *in vitro*. However, they could not reproduce such a result *in vivo*, probably due to lack of a proper tool for the detection of MARylation in the cell (202). However, they showed that in diffuse large B-cell lymphoma ARTD9 interacts via its macrodomains with both MARylated isoforms of STAT1 (STAT1α and STAT1β). As consequence, STAT1α activates two proto-oncogenes genes (interefer-one responsive factor 2 (IFR2) and B-cell CLL/lymphoma 6 (BCL6)), while STAT1β negatively regulates transcription of the INFγ-dependent tumor-suppressor IRF1.

ARTD15/PARP15 regulates the activation of the Unfolded Protein Response (UPR) **(Figure 1.13,8)**. ARTD15 is a tail-anchored protein embedded in the membrane of endoplasmic reticulum (ER) and facing the cytosol (204). When in the ER there is abundance of unfolded proteins, the UPR is activated in order to temporally block protein synthesis and to express chaperones that will assist in the folding of the already synthetized but denaturated proteins. Jwa and colleagues showed that ARTD15 MARylates IRE1α and PERK, the two kinases involved in the activation of the UPR. Such a modification could induce decreased affinity for their inhibitor, the binding immunoglobulin protein (BiP), and activate the UPR. Although BiP is ADP-ribosylated, it is not clear which ADP-ribosyltransferase might modify it, since BiP resides within the ER lumen (205).

Lastly, proteomic studies of PAR-binding proteins or PARylated proteins showed enrichments in RNA biology-related classification terms (10, 118, 118, 206). In addition, several studies showed effects of ARTDs in all the key steps of RNA metabolism (10). For example, several splicing factors, which are members of the hnRNPs, contains

PAR-binding modules (207). Upon formation of PAR due to heat-shock stress, some splicing factors are sequestered from the mRNAs under maturation (208). Therefore, induction of ADP-ribosylation can have important consequences on the maturation of RNA transcripts **(Figure 1.13,9)**.

But ADP-ribosylation affects life of RNA molecules also in the cytoplasm. Mono(ADP-ribosylation) of eukaryotic elongation factor-2 (eEF-2) dissociates the ribosomal complex and blocks protein synthesis, as performed also by the cognate A fragment of diphtheria toxin (209, 210). This regulation occurs also upon inflammatory response, since cells treated with interleukin-1 β (IL1β) show increased levels of modified eEF-2 (211). Also, ADP-ribosylation affects the translation of mRNAs also by regulating the RNA-interference machinery. Interestingly, all the members of the Ago family are modified and the ADP-ribosylation affects their affinity towards the target mRNA, leading to a negative regulation of the microRNAs (212). Interestingly, in the same paper PAR is shown to be required for the formation of cytoplasmic stress granules, where RNA-binding proteins accumulate together with mRNA molecules. Even so far away from the chromatin, PAR shows to be precious for its scaffolding properties.

In conclusion, while PARylation has been studied since 50 years, only recently MARylation has gained a spotlight. In fact, limitations in the detection of the specific MARylation on the proteins has so far limited the possibility to define the subset of physiological processes regulated by this signal (11). However, the development of techniques that combine genetics, organic chemistry and mass-spectrometry has begun to explore the network connected to specific ARTDs (213, 214). In these studies, a specific ARTD is mutated so that it can use as donor of ADP-ribose a special NAD^+ derivative. The power of this tool is then to define both the complete list of substrates of the specific transferase, but also to provide the yet unknown sequence requirements for the modification. Therefore, in the next years more comprehensive understanding of the MARylation network will arise.

1.2.8 Upon DNA damage MacroD2 shows a double behavior

Dr. Gyula Timinszky, the direct supervisor of my PhD project, has worked on ADP-ribosylation and macrodomain proteins since the professional relationship with Prof. A.G. Ladurner started. Before my arrival at the lab, Dr. Timinszky was comparing the

behavior of different macrodomains upon DNA damage, by using a live-cell imaging approach.

Most of the macrodomains are able to recruit to the DNA damage sites because of their ability to bind ADP-ribosylation. For example, when MacroH2A1.1 macrodomain is expressed as fragment tagged with EGFP, it recruits to the microirradiation-induced DNA lesions with a very fast dynamic (56, 81). Similarly, this recruiting behavior has been successively shown for the three macrodomain ADP-ribosyl-hydrolases, MacroD1, MacroD2 and TARG1 (56, 67, 68).

Aside the published work, Dr. Timinszky discovered that, in contrast of the other macrodomain proteins tested, the EGFP-tagged full-length form of MacroD2 showed the decrease of the nuclear signal upon DNA damage. This behavior was proved for few conditions, like different DNA-damage inducing compounds or upon UV-microirradiation **(Figure 1.14)**. This phenomenon might be explained either by regulated nuclear transport or by a regulated form of protein degradation.

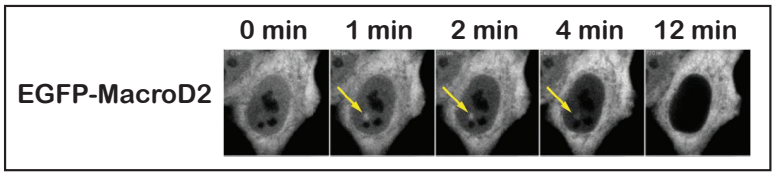


Figure 1.14: MacroD2 nuclear protein level decreases upon DNA damage - Laser micro-irradiation experiment on HeLa cells expressing EGFP-MacroD2. Yellow arrows show the focus point of the laser microirradiation. The experiment was performed by Dr. Gyula Timinszky.

Dr. Timinszky decided also to probe for possible inducers of the phenomenon. In a set of experiments, he showed the involvement of ataxia-telangiectasia-mutated (ATM) kinase in the induction of MacroD2 signal decrease. ATM is a key regulator of the DNA damage response and the activation of the cell-cycle checkpoint (215). Dr. Timinszky showed also that, if MacroD2 signal decreased in the nucleus, the inhibition of the kinase restores the initial distribution of the signal, arguing also for the reversibility of the process **(Figure 1.15)**.

The sudden discovery of the double MacroD2 behavior upon DNA damage is an intriguing open question that can add more to the regulation and function of this still unknown protein. Moreover, the interaction between ADP-ribosylation metabolism

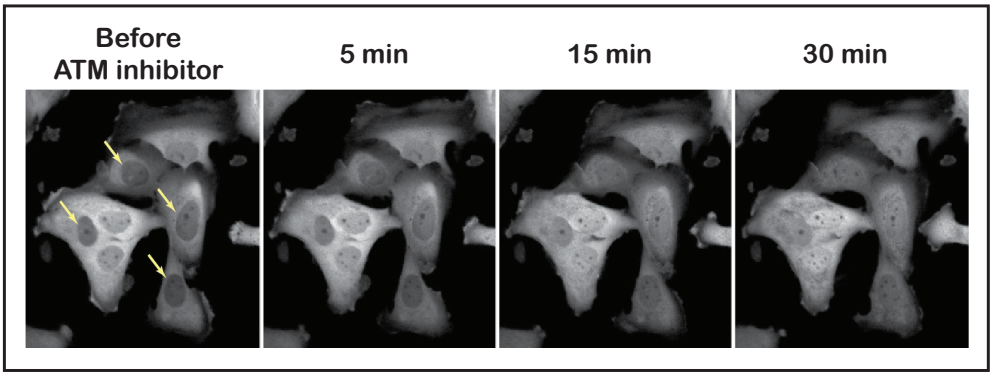


Figure 1.15: ATM inhibition results in reversal of MacroD2 signal decrease in the nucleus - Cell were pre-treated for one hour with camptothecin, yellow arrows indicate those cells that responded to the drug. Then, they were imaged upon treatment with ATM inhibitor KU55933. The experiment was performed by Dr. Gyula Timinszky.

and ATM-signaling, and its possible role in the DNA damage response, urges for a more detailed understanding of MacroD2 functions in the cell.

Since MacroD2 responds to ATM activation, I will give a summary about ATM kinase, its modes of activation and its protein family **(see Section 1.3)**. Finally, since I could show that MacroD2 new behavior is explained by regulated nuclear export, I will give an overview on the nuclear transport and the basic features of its regulation **(see Section 1.4)**.

1.3 ATM and the PI3K-like kinases

Ataxia-telangiectasia mutated (ATM) is a key regulator of the DNA damage response and it belongs to the PI3K-like kinse family. The amount of studies about this kinase family is massive, since most of the members are involved in DNA damage and tumor development.

1.3.1 ATM is the protein behind the ataxia-telangiectasia disease

ATM received its name from the autosomal recessive disorder ataxia telangiectasia (A-T) (216, 217). This disease presents several symptoms, among which the dilatation of blood vessel (telangiectasia), neurodegeneration in the cerebellum that affects movement and coordination and general immunodeficiency (215). Also, A-T patients

are predisposed to cancer, especially lymphomas, and are very sensitive to ionizing radiation (218).

ATM is a big protein, 350 kDa in weight and 3056 residue long. ATM belongs to the PI3K-like kinase (PIKKs) family, related to the lipid-kinase PI3K. Unlike their cousins, PIKKs transfer phosphate group exclusively onto serines or threonines of protein targets (219). For some members of the family, ATM among them, the modified serine or threonine must be followed by a glutamine, forming the so called SQ/TQ motif (220). ATM is located mainly in the nucleus (221), even though many studies showed presence of ATM in the cytoplasm of neuronal cells of mouse and human cerebellum or in differentiated neuronal cell lines (222, 223, 224, 225). One report suggested that ATM might export together with NEMO upon stress-dependent activation of the NF-κB pathway (226).

1.3.2 ATM is a master regulator of the DNA damage repair

Cells lacking ATM activity present increased genomic instability and sheer sensitivity to DNA damage-inducing agents, in particular to irradiation or radiomimetic compounds (227, 228). This sensitivity was soon linked to the important function that ATM has during DNA damage response, in particular upon DSB repair (215). In fact, ATM-dependent regulation is important for the overall DNA repair process, the cell-cycle checkpoint regulation and the definition of the cell fate: survival over apoptosis (229). In addition, upon DNA damage, ATM activity immediately restructures most of cellular processes through a powerful amplification of the signal, mediated by ATM itself and its associated kinases: Chk1, Chk2 and Mk2, kinases that are activated by ATM to relay the SOS signal everywhere in the cell (230).

It is not clear how ATM is activated in the very first place upon DSB **(Figure 1.16)**. ATM might have the ability to bind damaged DNA (231), but it seems to be not a strong interaction and it requires help from other proteins (232). Thus, some factor must be sensing the presence of DSB and transfer the message to ATM itself, in order to initiate the system. The best candidate is the complex MRN, formed by Mre11-Rad50-Nbs1 (233, 234, 235): in fact, Mre11 is recruiting to irradiated foci with a fast dynamics and independent from ATM (236). Patients lacking Mre11 present most of A-T symptoms, above all the sensitivity to irradiation, thus this disease is called A-T-like disorder (ATLD). The mutation of Nbs1, on the other hand, induces

the Nijmegen breakage syndrome (NBS), which among diverse symptoms shows strong radiosensitivity. The similarities are probably due to the fact that the MRN complex is important for the initial activation of ATM and for its binding to DNA (232, 237). It is suggested that the MRN complex is able to bind the DNA ends, slightly open the DNA molecule and predispose it to the interaction with ATM (232).

After the initial activation of ATM mediated by the MRN complex (232, 237), more steps are required to the full activation. In the absence of DNA damage, two ATM molecules constitute a dimer, whose positioning inhibits the binding and modification of substrates (238). Upon DNA damage, the two molecules auto-phosphorylate in *trans* on S1981 and this is enough to dissociate the dimer. In addition, the histone acetylase TIP60, constitutively binding the C-terminus of ATM, acetylates ATM on position K3016 and fully activates it (239). TIP60 is also involved in a positive feedback loop, since ATM induces the Chk2 kinase, whose activity helps TIP60 remain active (240, 241, 242).

Once ATM is activated, it can perform its main role of amplification of the signal. ATM phosphorylates the histone variant H2AX (γH2AX), which is used as recruitment platform for repair factors, forming visible irradiation-induced foci (243). One of the first factors recruited and modified by ATM is the mediator of DNA damage checkpoint protein 1 (Mdc1) (244, 245). In fact, Mdc1 is important for the setup of a positive feedback loop by binding the γH2AX with its BRCT domain, ATM with the FHA domain and interacting with Nbs1 via a Ck2-dependent phosphorylation event (246, 247, 248). Modification of Mdc1 by ATM helps its oligomerization and therefore the formation of the foci (249).

This recruiting environment enhances ATM chances to encounter and modify its almost thousand substrates (42, 44, 250, 251). Mass spectrometry approaches showed an intricate network, mainly involved in DNA repair, but also regulating transcription, mRNA processing, chromatin remodeling, cell-cycle regulation. ATM-mediated modifications on the kinases Chk1, Chk2 and Mk2 are able to achieve the maximum amplification of the response (252, 253, 254). Other described modified proteins are Nbs1, Rad50, Kap1, the cohesin subunits Smc1 and Smc3 and the tumor suppressor p53 (44).

The tumor suppressor p53 was the very first identified target of ATM activity (255, 256, 257). In cells undergoing DNA damage in G1 phase, the G1/S checkpoint

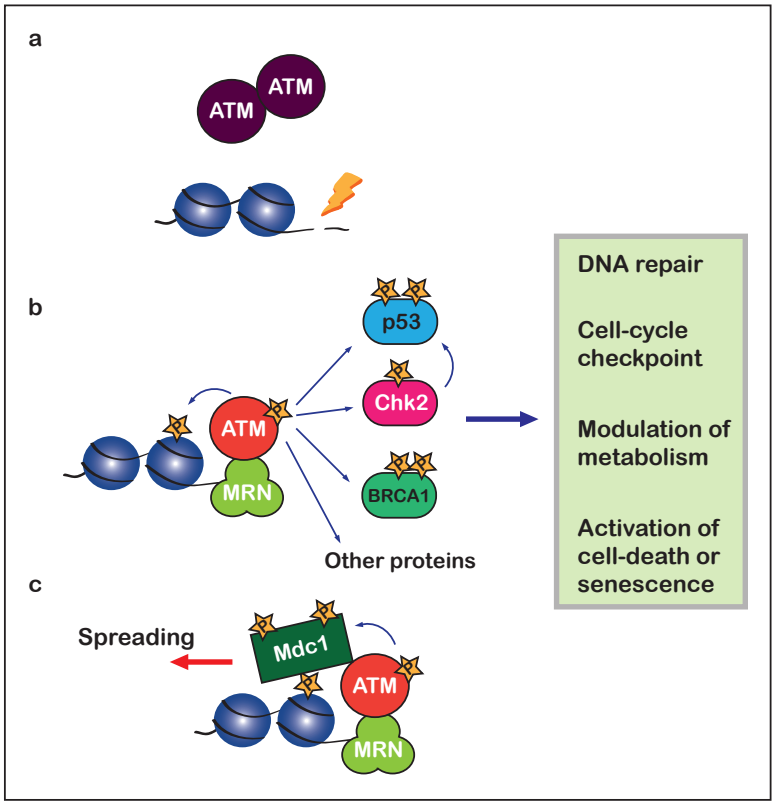


Figure 1.16: ATM is activated by double-strand breaks - *a*) In normal conditions, ATM is in a dimeric inactive form. But upon DSB, ATM becomes active. *b*) ATM activates itself by autophosphorylation, which induces the release from the dimer. ATM is then able to interact with the MRN complex at the DNA lesion- The interaction further activates ATM, that in turn modifies many effectors: histone H2AX, p53, BRCA1 and Chk2 and others. The start of the signaling cascade leads to the repair of the DNA damage, the temporary arrest of the cell-cycle and the modulation of the metabolism, with the temporary standby of energy-consuming processes. In extreme case, ATM initiates the cell-death program or the entrance to senescence. *c*) The phosphorylation of the histone H2AX recruits Mdc1, which induces the further activation of ATM and the spreading of the histone mark. Adapted from Maréchal et al. 2013 (230).

is activated and cells are prevented to start the DNA replication, due to the possible problems that the replication machinery can encounter in the presence of unrepaired DSBs (258). The DNA-induced phosphorylation on S15 mediated by ATM is one of the sequential events that lead to the stabilization of p53, which acts as transcription factor and induces the cell-cycle checkpoint (259). ATM has a guardian role throughout the cell cycle. Mutations in ATM impair both the S-phase and the S/G2 checkpoints. In both, ATM modifies the breast cancer type 1 susceptibility protein (BRCA1) in two sites, whose mutation affects either one or the other checkpoint initiation (260, 261, 262).

Recent studies have also connected ATM activation with NF-κB signaling (226). Upon DNA damage, ATM performs one of the three post-translational modifications required to target the NF-κB essential modulator (NEMO) from the nucleus to the cytoplasm. Such a regulated localization is important for the phosphorylation and dissociation of the NF-κB inhibitory complex, thus leading to the activation of the inflammatory response (193). Thus, while ARTD10 MARylates NEMO, repressing the activation of the pathway (194), ATM induces it by leading NEMO into the cytopasm. Therefore, induction of the inflammatory pathway represents another condition when several signals must be integrated to obtain an organic response.

1.3.3 ATM is activated also in MRN-independent manner

Apart from the previously described MRN-dependent activation of ATM, other activation modes have been discovered in the recent years (229, 263). Upon hypotonic stress or cloroquine treatment, the ATM INteracting protein (ATMIN) binds ATM, as shown by colocalization in immunofluorescence experiments (264). ATM and ATMIN association induces the ATM-mediated phosphorylation on p53. Also, ATMIN competes with Nbs1 for the same binding site on ATM, therefore ATMIN-dependent and MRN-dependent activation pathways are mutually exclusive (265). The interaction between ATM and ATMIN is abrogated by a ubiquitylation event performed by the E3 ligase UBR5 upon DNA damage (266). Thus, this event shifts the equilibrium of the ATM species towards the interaction with Nbs1 and the initiation of the DDR.

Hypoxia is also able to activate ATM (267). The study from the Hammond group showed that the activation of ATM is independent of Nbs1 and the foci formation, since ATM stays in a diffuse nuclear form. However, they showed that Mdc1 is required for

the amplification of the signal. Later, ATM-activation was confirmed and linked to the phosphorylation and stabilization of hypoxia-inducible factor 1-α (HIF-1α) (268).

A MRN-independent mode of activation of ATM is also achieved upon oxidative stress (263, 269). It was already clear that both A-T cell lines and mouse models showed higher levels of ROS if compared with normal cells (270, 271, 272). Therefore, it was suggested that the absence of ATM induces oxidative stress in neurons. This chronic stress could explain the degeneration in the cerebellum of A-T patients. This hypothesis is partially confirmed with the higher sensitivity that A-T cells have for oxidative agents (273, 274).

The mechanism of the activation is quite peculiar, since it involves two ATM molecules that keep the dimeric structure, which is actually enforced by the formation of disulfide bonds (275). Therefore, in contrast to the inactive, non-covalent dimer of the inactive form, Paull and colleagues found an ATM dimer form that is actually active. The most relevant disulfide bond, out of the many forming within the dimer, utilizes the C2991. Mutation of this cysteine into alanine affects only the hydrogen peroxide-induction of ATM but not the canonical MRN-dependent pathway (275).

To sum up, the activity of ATM has the critical role in coordinating many processes that take place in the cell and in driving to the most successful resolution of the crisis: by stopping activities like cell-cycle progression and DNA transcription, by easing the access to the DNA lesion with a suitable recruitment platform and relaxation of the chromatin, by readjusting the cellular metabolism for a less DNA damage-prone environment.

1.3.4 ATR and DNA-PK also belong to the PIKK family

The PI3K-like kinase family is composed in total by six members and all are specialized in the response to different stimuli (276). For the DNA damage repair response, two other members are activated beyond ATM: ATM and Rad3 related (ATR) and DNA-dependent protein kinase (DNA-PK). ATR activity is induced upon DSBs that are associated to replication fork (277). The resection of the DSBs and the formation of RPA-coated single-strand DNA induces the activation of ATR in complex with ATRIP (278). On the other hand, c-NHEJ is regulated by DNA-PK catalytic subunit (DNA-PKcs), in association with the Ku70/Ku80 dimer (279).

The DNA-PK holoenzyme is composed by three components: the DNA-PKcs, around 350 kDa in size; Ku70; Ku80 **(Figure 1.17, panel a;** (279)**)**. The latter two factors form the clamp that recognizes the DNA double strand and help the recruitment of the DNA-PKcs to the DSB. A low-resolution structure shows how the kinase binds directly the DSB ends with the help of the Ku70/Ku80 dimer (280). Also, the DNA-PKcs phosphorylates itself in several sites, behaving similarly to ATM (281). However, different groups of auto-modification induce distinct effects, like negatively regulating the interaction with the Ku proteins or modulating the enzymatic activity (282, 283).

DNA-PK modifies a series of effectors involved in the c-NHEJ (279). In addition, DNA-PK modifies the histone H2AX, even though this event is linked to particular conditions, like hypotonic stress combined to irradiation and mitosis regulation (284, 285). As I mentioned above, DNA-PK interacts with ARTD1. Even though they might interact upon DSB repair, their interaction is important for the DSB-dependent induction of transcription of estrogen-responsive genes (286). DNA-PK is also involved in the transcription of other classes of genes, like the androgen-responsive (287) or the insulin-responsive ones (288).

As third DSB-dependent kinase, ATR role is so important that the homozygous mutation of ATR brings embryonic lethality in mice and cell death in human cells (230). ATR is activated upon formation of long stretches of ssDNA, as in the case of DSBs and with stalled replication forks **(Figure 1.17, panel b)**. Due to the common SQ/TQ motif and the connection to the HR regulation, the targets of ATM activity found by mass spectrometry experiments are actually shared with ATR (42), although the two subsets of ATM and ATR targets change according to the DNA damage type (289). While ATM is more relevant for the G1 checkpoint, ATR is more active in the intra-S and G2/M checkpoints, by phosphorylating its main effector, Chk1 (252, 290, 291). However, since the role of ATM is shown in all checkpoints, it is possible that for each case the two kinases cooperate to properly amplify the response.

The other three members of the PIKKs are activated upon other stimuli. Suppressor in morphogenesis in genitalia 1 (SMG-1) is involved in the non-sense mediated decay, a quality control pathway for the mRNAs (292). The mechanistic target of rapamycin (mTOR, previously known as FRAP) is instead activated upon metabolic

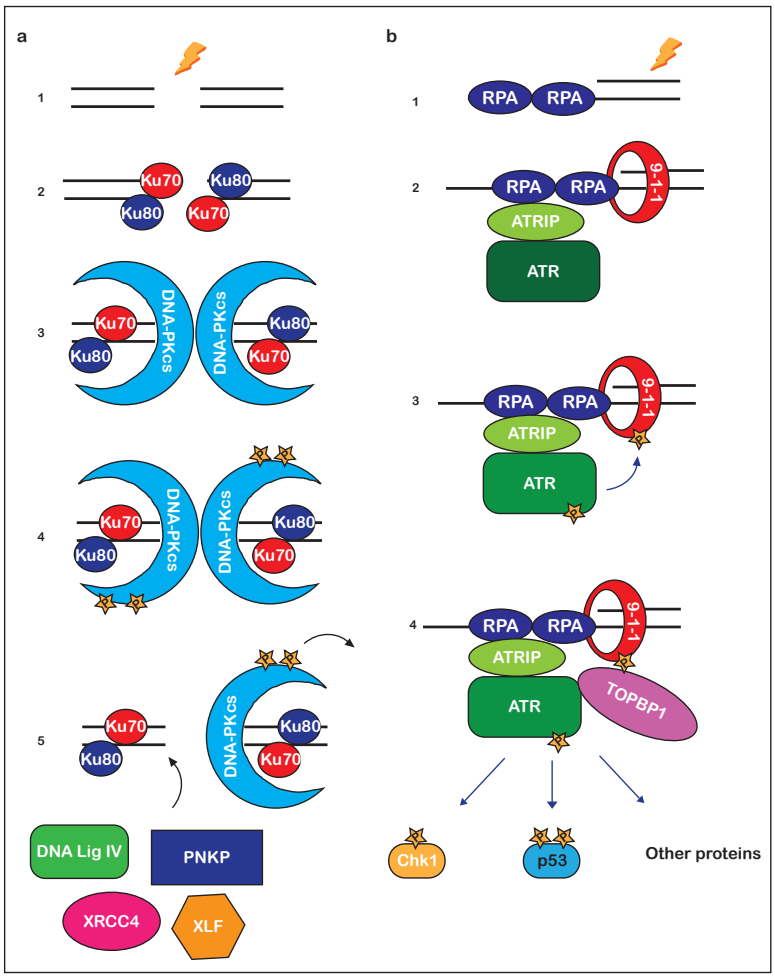


Figure 1.17: The activation of DNA-PK and ATR by double-strand breaks - *a*) When DSBs occur (1), the Ku70/Ku80 dimer is recruited to the ends of the DNA lesion. The Ku dimer recruits DNA-PKcs (3), forming the DNA-PK holoenzyme. In the complex, the Ku proteins shift along the DNA to allow DNA-PKcs to bind the DNA ends. DNA-PK is finally active (4), auto-phosphorylates and leads to the recruitment of c-NHEJ factors (as example, XRCC4, XLF, PNKP, DNA Ligase IV) (5). The activation induces also DNA-PKcs to disattach from the DNA. *b*) RPA molecules coat ssDNA, also in the proximity of the replication fork (1). RPA induces the independent recruitment of inactive ATR via ATRIP, and of the 9-1-1 complex (2). ATR gets activated and modifies the 9-1-1 complex (3). The modification leads to the recruitment of the TopBP1 protein, which further enhances ATR activity (5). ATR modifies other targets, for example Chk1 and p53. Adapted from Jette and Lees-Miller 2015 and Maréchal et al. 2013 (230, 279).

stimuli (293). mTOR is responding to different pathways, like insulin and growth factors, and induces proliferation. It is the catalytic subunit of two different complexes, with different components: mTORC1 induces ribosome genesis and lipid synthesis, in response of a favorable energetic environment; mTORC2, instead, positively regulates the actin cytoskeleton, cell size and cell progression. The last member of the PIKK family is transformation/transcription domain-associated protein (TRRAP) and, although it is catalytically inactive, it is a component of a histone acetyltransferase complex specialized in the activation of mitotic checkpoint genes (294).

1.4 The nuclear transport

When a stimulus occurs, survival requires a fast but accurate response. As mentioned above, this is easily achieved by post-translational regulation of mediators of the response. Another way to switch on or off the activity of a protein in a fast manner is to regulate its location in the cell. An enzyme, for example, might be restricted to a compartment, secluded from its substrates, and released only at the moment of need. Considering that within the eukaryotic cell there are many membrane-enclosed compartments, the regulation of the function can take advantage of this evolutionary feature.

The preliminary work performed by my supervisor Dr. Timinszky indicated that MacroD2 nuclear signal is depleted upon DNA damage. During my PhD project, I showed that this depletion is explained by the regulated nuclear export of MacroD2. The nuclear transport is not a trivial problem for the cell, but mechanisms have evolved to exploit a probably energetic-disadvantage into a regulatory opportunity.

1.4.1 The mechanism of nuclear transport

The presence of a distinct compartment where to keep the genetic information and the related processes was one of the major breakthroughs in the evolution of the eukaryotic cell. The nucleus is separated from the cytoplasm by the nuclear envelope. This is a double-bilayer membrane connected with the endoplasmic reticulum. The nuclear envelope is fenestrated by giant protein complexes, the nuclear pore complexes (NPCs), which allow the passage of molecules between the nucleus and the cytoplasm. Each

vertebrate cell has around 4000 NPCs embedded in the nuclear envelope, allowing a constant traffic between the two compartments (295).

More than simple channels, NPCs are complexes formed by multiple copies of 34 proteins, called nucleoporins (296). These make a passage in the points where the inner and the outer membranes of the nuclear envelope fuse. In *X. laevis*, the giant structure is around 120 MDa in total weight, a diameter of 126 nm and a height of 71 nm within the envelope, excluding the accessory elements on the cytoplasmic and nuclear face (297). In fact, on the nuclear side eight rod-shapes proteins connect to a distal ring and form the so called nuclear basket; likewise, on the cytoplasmic face there are eight flexible protrusions, completing the structure. The overall structure is conserved across all eukaryotic clades (298).

The conduit across the envelope is full of unstructured tails of the nucleoporins, rich in phenylalanine-glycine (FG) repeats. These repeats weakly interact with each other, forming a hydrogel in the center of the pore that form the permeability barrier for the passage (299). The diameter of the conduit is a controversial issue, since even if it is calculated to be 2.5-5 nm (300), it can accommodate particles up to 40 nm in diameter (301). The NPC allows the indiscriminate passage of ions and metabolites. Macromolecules smaller than 5nm in diameter (roughly 40 kDa in size) can also diffuse freely (300). This is a rough estimation, since slightly bigger macromolecule but with an elongated shape still show free passage. On the other hand, bigger macromolecules require assisted transport, mediated by the nuclear transport receptors, a class of proteins that bind their cargo and help the cross beyond the NPC (302).

The nuclear transport receptors bind or release the protein cargos depending on the binding with the small GTP-binding protein/GTPase Ran **(Figure 1.18)**. Ran converts guanosine triphosphate (GTP) into guanosine diphosphate (GDP): the conformation of Ran changes according to the guanosine species it is binding at the moment. However, the enzymatic activity is not efficient. To this end, Ran needs to associate with the GTPase-activating proteins (GAPs), which enhance the conversion from GTP to GDP; and with the guanine nucleotide exchange factors (GEFs), which induce the release of the GDP and the bind with the next GTP molecule. By positioning GAPs and GEFs in the cytoplasm and in the nucleus, respectively, it is possible to create a directionality in the conformational change: RanGAPs, like RanBP1, are associated to

the cytoplasmic protrusions of the NPC, while RanGEFs, like RCC1, are sitting on the chromatin (303).

The nuclear transport receptors are divided into two classes: the importins bind the cargo in the cytoplasm and release it in the nucleus upon binding with the Ran-GTP; on the other hand, the exportins bind the cargo in the nucleus in a trimeric complex with Ran-GTP. This complex dissociates in the cytoplasm when Ran is activated to convert the GTP into GDP. Ran is then transported back into the nucleus by binding with the nuclear transport receptor NTF2. This nucleocytoplasmic cycle, thus, even if it offers many regulatory options, is energy-dependent (302, 303).

The importins and exportins belong together to the karyopherin β superfamily (304). Most karyopherins bind their cargos directly, while other times adaptor proteins are required between the nuclear transport factor and the cargo. The importin β, for example, binds the cargo through the interaction with the karyopherin importin α (305). While the importin α/β complex accounts for the import of most proteins, there are other importins that have a more specialized subset of cargoes. Similarly, the exportin 1 (Xpo1/CRM1) is the one that helps the export of the vast majority of cargoes. Other exportins have few other specialized targets, for example exportin t carries the tRNAs into the cytoplasm (306).

The karyopherins recognize special sequences on the cargo protein, called nuclear localization signal (NLS), in case it induces the nuclear import, and nuclear export sequence (NES), if it is able to drive the nuclear export. These sequences are generally located in unstructured loops and the presence of such sequences helps with the prediction of the general localization of the protein, even though the accessibility of a specific sequence can be constantly modulated. Importin α recognizes an NLS sequence that consists of either one stretch of basic residues or two basic stretches connected via a short linker, in which case it is called a bipartite NLS (307). Apart from the most classical NLS, other sequences have been described, like the so called PY-NLS recognized by the adapter importin β2 (308, 309). Regarding the export signal, exportin 1 recognizes an 8- to 15- residues leucine/hydrophobic-rich motif (310). It is also worth to mention that the cargos can also interact with other partners, conferring to proteins without any localization sequence the ability to cross the nuclear envelope. This mechanism is called “piggy-backing” (302).

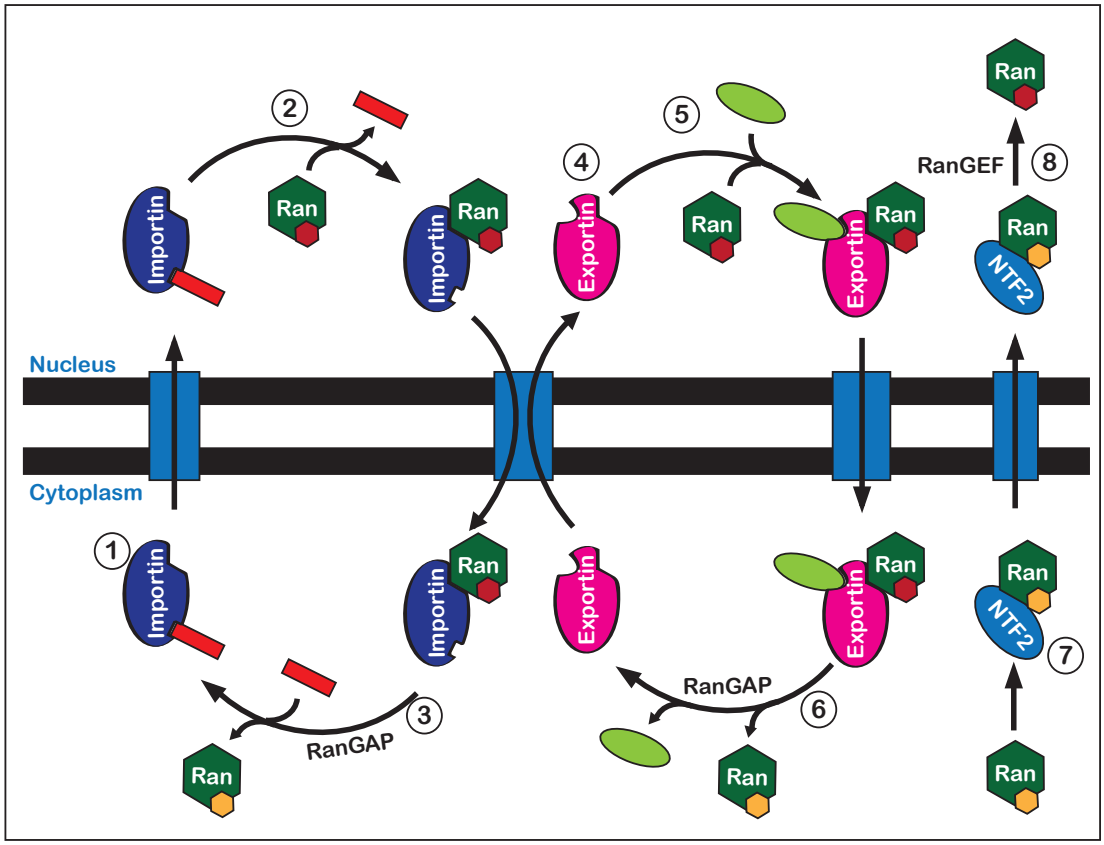


Figure 1.18: The assisted nuclear transport cycle - The nuclear tranport is illustrated: *1*) The importin binds its cargo in the cytoplasm and crosses the nuclear envelope; *2*) in the nucleus the Ran-GTP (red) molecule binds the importin, induces the release of the cargo and drives the importin back to the cytoplasm; *3*) the RanGAP enhances Ran GTPase activity and the Ran-GDP (yellow) releases the importin, which is free to bind the new cargo; *4*) in the nucleus, the exportin interacts with Ran-GTP (red) which allows the interaction with the exporting cargo; *5*) the complex formed by exportin-Ran-GTP-cargo crosses the nuclear envelope; *6*) in the cytoplasm, the conversion of Ran-GTP (red) into Ran-GDP (yellow) incudes the disassembly of the complex and the exportin is free for a new round; *7*) Ran-GDP (yellow) is transported back by the transporter NTF2; *8*) in the nucleus, a RanGEF induces the release of GDP and the binding with the new GTP, regenerating the Ran-GTP (red) for the nuclear transport.

1.4.2 Regulation of protein localization

The localization of the protein across the different compartments depends on a series of factors. If the protein is smaller than 40 kDa, or if it has a more suitable shape, it can freely pass through the permeability barrier of the NPC and equally distribute between the nucleus and the cytoplasm. However, the size will impact on the speed with which the protein will equilibrate, even though the equilibrium will be eventually reached. On the contrary, the localization can be tightly regulated if the proteins need the help of the nuclear transport receptors. Thus, the overall distribution is defined by many factors. First of all, the concentration of free karyopherins and how the transport system itself is saturated can impact on the speed of the process. On the other hand, the accessibility of the cargo localization signal is often modulated (302).

In some occasions, an interaction partner can mask the signal upon binding, affecting the interaction between the cargo and the karyopherin, like in the case of the inhibitor IκBα that masks the NLS of NF-κB (311, 312). The regulation can also be due to an allosteric change in the protein conformation, which the cargo adopts upon binding with an interactor: when the fatty-acid binding protein 5 (FABP5) binds to its ligand, it undergoes a conformational change so that a tertiary non-linear nuclear targeting signal is formed (313).

In other cases, the protein can be trapped in one specific compartment, due to a stable interaction with an immobile component, such as cytoskeleton, chromatin or the cellular membrane. This means that the pool of protein available for the nuclear transport is reduced and the modulation of the export consists mainly in modulating the interaction with the cellular component. This is the case of the Ca^{2+}-dependent induction of calmodulin nuclear import (314). On the other hand, some proteins, like Notch receptor, are generally anchored in a particular compartment and the release and translocation is triggered by a protease, which cleaves away their anchor (315).

Last but not least, the modulation can occur also co-transcriptionally, by performing an alternative splicing that might introduce (or remove) a localization signal in (or from) the protein sequence. In the case of the Nek2 kinase, the splicing of three alternative variants changes the presence and even the strength of the NLS, making Nek2C isoform mainly nuclear, Nek2B mainly cytoplasmic and Nek2A equally distributed (316).

Most often, however, the modulation occurs via post-translational modification. The post-translational modifications might affect the interaction between the cargo and the transporter. For example, upon DNA damage response ATM or ATR phosphorylate p53, next to its NES, preventing its interaction with exportin 1 and leading to its accumulation in the nucleus (256, 257, 317). Likewise, the estrogen receptor α is also phosphorylated upon p38MAPK activation and the modification affects the NES recognition, leading to nuclear accumulation (318). Even if the phosphorylation is the most common post-translational modification that affects nuclear transport, other post-translational modifications can occur as well. The p300 acetylates the NLS of the RecQ protein-like 4 (RECQL4) helicase, preventing it to localize in the nucleus (319).

Post-translational modifications could also actively induce the translocation, although these cases are not very frequent. The best known cases occur in the MAP kinases signal cascade: p38MAPK and p42MAPK are phosphorylated in two different sites and the phosphorylation induces the recognition of the NLS by the importin 7, leading to the accumulation in the nucleus and the activation of nuclear targets (320). On the other hand, post-translational modification-dependent nuclear export occurs in case of the androgen receptor (AR). While it is translocated into the nucleus upon androgen-binding, the AR nuclear export is enhanced when other signaling pathways are active: p38MAPK or c-Jun induce the phosphorylation on S650, next to the NES of the AR, thus increasing the interaction with the exportin 1 (321). Similarly, upon stress, p38MAPK induces the p38 regulated/activated protein kinase (PRAK, also known as MAPKAPK5) nuclear export (322). Several types of cellular stress induce also the phosphorylation of the kinase MK2 on the threonine 317, which induces a change in conformation and increases the accessibility of its NES, leading to the accumulation in the cytoplasm (323). Remarkably, a series of three different post-translational modifications are necessary to export NF-κB-inhibitor NEMO: ubiquitination, ATM-dependent phosphorylation and finally SUMOylation (193, 226).

Sometimes, however, the post-translational modification-dependent accumulation in the cytoplasm is caused by the fact that the protein tightly interacts with residents of the cytoplasm and its nuclear import is therefore impaired. This is true for the ubiquitin E3-ligase COP1, whose activity leads to the nuclear export and degradation of p53. COP1 is phosphorylated upon DNA damage and the mark is recognized by the chaperone 14-3-3σ in the cytoplasm (324). Also other proteins get trapped in

the cytoplasm due to the interaction with the 14-3-3 proteins, namely FOXO1 (325) or FOXO4 (326, 327). In the case of the nuclear factor of activated T-cells (NF-AT), GSK3 modifies the protein and the phosphorylated form is trapped in the cytoplasm, although it is not clear if it is due to interaction with other factors (328). Upon Ca^{2+} signaling, calcineurin dephosphorylates NF-AT and induces its nuclear import (329, 330).

1.5 Integration of the signals: the key to the success

In conclusion, to achieve fast adaptation to all the inputs that are constantly collected, cells need to be able to change rapidly the pathway usage and the specific protein activity. For example, the complex machinery that is employed for the repair of DSBs requires an accurate calibration of its activation, as well as the block of all the physiological processes that might negatively impact the repair. The easiest way to achieve such a complex result is to send out a signal that can tune the activity of many pieces of the clockwork at the same time: post-translational modifications are information theory applied to biochemistry, the clever way with which the cells have coped with the daunting task of adaptation.

2

Aims of the project

As a post-translational modification, ADP-ribosylation presents stunning dynamics and an elevated level of complexity. It is involved in virtually every process in the cell, either directly, as a regulatory mechanism of single steps of a process, or indirectly, by its involvement in the regulation of chromatin accessibility. Only recently have appropriate tools been developed to thoroughly study this modification, bringing a new Renaissance to the field.

Although the presence of tools to detect PARylation has directed most of the efforts in the understanding its functional role, still little is known about the actual extent of regulation mediated by MARylation. However, recent studies suggest a similar extensive involvement of MARylation signaling in cellular life. In 2013, it was possible to define the last enzymes that complete the whole metabolism of ADP-ribosylation, with the discovery of MacroD1, MacroD2 and TARG1 (56, 67, 68). But their actual functions are still far from being defined.

MacroD2 is the largest of the three proteins. The localization of the EGFP-tagged version of MacroD2 overlaps with that of TARG1, but MacroD2 contains an additional protein sequence that suggests a different function or regulation as compared to the smallest brother. The *MacroD2* gene, as distinct from *MacroD1*, is first found in the vertebrates, probably due to a duplication event followed by a fusion with an additional unknown sequence, which is present only within vertebrates with more than 50 % of identity. Therefore, the presence of MacroD2 as gene correlates with high complexity, and the association studies with deficiencies in neural function suggest that it might indeed have a specific function that appeared only recently in evolution.

2. AIMS OF THE PROJECT

During a short internship at the Timinszky group, before starting the PhD, I helped characterize the recruitment dynamics of the MacroD2 macrodomain, then included in the Jankevicius et al., 2013 work. Before that point, published studies on MacroD2 had mainly shown only broad associations with disease. The characterization of MacroD2 enzymatic activity has been a remarkable step forward, but it is still unclear how it is connected with the greater picture. Therefore, the serendipitous discovery of a second MacroD2 behavior upon DNA damage - the decrease of its nuclear signal - represented a thrilling opportunity to add more pieces to the puzzle.

Already at that time, preliminary experiments showed that HeLa cells transfected with mEGFP-MacroD2 full-length manifest a decrease in the fluorescent signal in the nucleus in certain conditions, like Topoisomerase poisons camptothecin and etoposide, or UV laser micro-irradiation **(Figure 1.14)**. Additionally, preliminary data showed a possible induction of MacroD2 export by ATM, and the reversibility of this process **(Figure 1.15)**.

I joined the group and this particular project with the aim of characterizing this MacroD2 behavior: the decrease of nuclear levels. In general, such a behavior could show a form of regulation for the MacroD2 function, either by activation of a degradation pathway or by active translocation across the nuclear envelope. Therefore, I thought that by dissecting the behavior that Dr. Timinszky had discovered, I would be able to understand the mechanism inducing it, which could consequently provide more insight on the cellular function of MacroD2. In addition, since the preliminary experiments showed that the induction of the MacroD2 special behavior is probably due to ATM-signaling, this project allowed me to explore a probable integration node between two massive regulatory systems upon DNA damage response, namely the ADP-ribosylation-mediated and the ATM-mediated signaling.

For these reasons, my PhD project had two main goals.

1. **Aim I: characterize the decrease of MacroD2 nuclear protein levels upon DNA damage**

   To better understand the phenomenon of decrease of MacroD2 nuclear protein levels I intend to use a combination of live-cell microscopy and biochemical approaches. Firstly, I need to define if the decrease of MacroD2 nuclear signal is due to regulated nuclear export or to the activation of a degradation pathway. Next,

I intend to define the overall conditions that robustly induce the phenomenon, as well as the contribution of ATM or other kinases of the PI3K-like kinase family. I also intend to define the requirements for the phenomenon at the level of the MacroD2 protein sequence. Lastly, the decrease of MacroD2 nuclear levels also needs to be proved for the endogenous protein.

2. **Aim II: identify candidates for the regulation of MacroD2 upon DNA damage and for MacroD2 cellular function**
   In parallel to Aim I, I intend to define the interactome of MacroD2 protein in the presence and absence of genotoxic stress. By means of an unbiased approach, this proteomic strategy can suggest testable hypotheses concerning two aspects: the protein factors that might be involved in the regulation of the decrease in nuclear levels, and the function of MacroD2 in cells. In fact, by performing the experiment under conditions that induce the new behavior of MacroD2, I can define proteins that might help MacroD2 regulation. On the other hand, both the control and genotoxic conditions could show factors involved in cellular processes that are possibly regulated by MacroD2 enzymatic activity.

3

ATM kinase induces MacroD2 nuclear export

3.1 Introduction

MacroD2 is one of the three ADP-ribosyl-hydrolases recently found in humans (56, 67). These proteins share the same catalytic domain, known as the macrodomain, which enables them to remove the most-proximal ADP-ribose from the target protein, completely reversing the ADP-ribosylation modification. The MacroD2 macrodomain has been shown to recruit to DNA lesions upon laser microirradiation experiments **(see Figure 1.10)**. However, the exact role of MacroD2 during the DNA damage response is still unknown.

As mentioned in the Section 1.2.5 of the Introduction chapter, preliminary experiments showed that the nuclear signal of EGFP-tagged MacroD2 full-length protein decreases upon DNA damage **(see Figure 1.14)**. This serendipitous discovery suggests two different possibilities: either upon DNA damage MacroD2 is degraded in a regulated manner, or the protein subcellular localization changes, with a net transport into the cytoplasm. In both cases, it proposes that MacroD2 could be regulated, either at the protein levels or in its subcellular localization. Moreover, other experiments showed that the depletion of the MacroD2 signal in the nucleus is blocked upon treatment with a specific ATM kinase inhibitor **(See Figure 1.15)**.

Although little is still known about MacroD2 function in the cell, the hypothesis that MacroD2 is regulated upon DNA damage can provide more insight into under-

standing ADP-ribosylation's function during the repair response, both at the DNA lesion and in the rest of the cell. Also, the possible involvement of ATM kinase brings ADP-ribosylation into the arena with one of the master regulators of DNA damage repair. Therefore, I decided to characterize this particular regulation of nuclear signal MacroD2, by understanding its nature and by dissecting the process of its induction.

3.2 MacroD2 exports from the nucleus upon DNA damage

A necessary first control experiment was to test if the EGFP itself could be the cause of the decrease of MacroD2 nuclear protein levels. To do this, I transfected HeLa cells stably expressing mCherry-H2B with either EGFP-tagged full-length MacroD2 or EGFP. The cells were then imaged upon UV-laser microirradiation and the images were analyzed with an automatic approach **(see Methods 6.8)**. By comparing the ratio of the nuclear signal over the cytoplasmic signal, I observed a consistent decrease of EGFP-MacroD2 signal in the nucleus. This behavior was not observed with EGFP alone, excluding that the EGFP tag could be the cause of the phenomenon **(Figure 3.1)**.

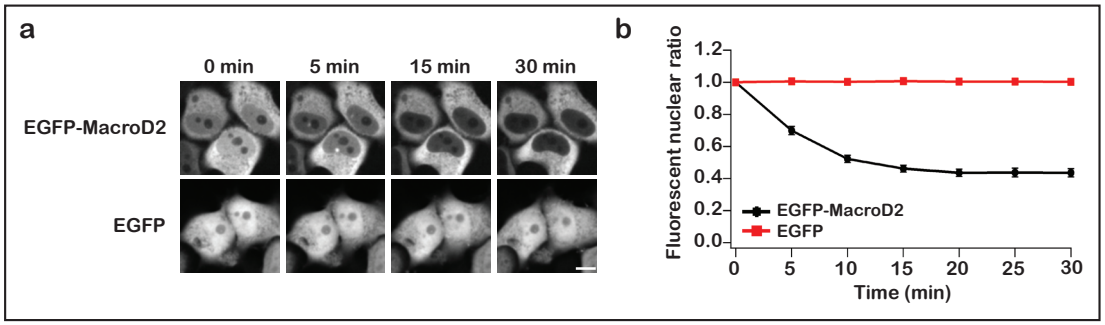


Figure 3.1: Nuclear EGFP-MacroD2 signal is depleted in the nucleus - a)UV-microirradiation experiments using HeLa cells stably expressing mCherry-H2B and transfected with mEGFP-MacroD2 or mEGFP constructs. Scale bar, 10 μm. b) 50−100 cells were quantified from three independent experiments **(see Methods 6.8)**. Nuclear/cytoplasmic ratio was calculated using CellProfiler 2.0 (331), and the mCherry signal was used for the nuclei segmentation. Error bars, 95 %CI.

I then aimed to understand whether the amount of DNA damage correlates with the extent of MacroD2 signal decrease. Thus, I tested how the increasing amount

of DNA damage that I inflicted on the cells by changing the energy of the 355 nm laser affected the extent of the MacroD2 nuclear signal decrease. Increasing amounts of DNA damage initiate the DNA damage response in a more robust manner and enhance the amplification of the signaling cascade. The increase in response is used to show that phenomenon is due to a biological process, as customary in DNA-damage research field, and not due to an artifact, which would instead show no correlation. Thus, I performed the experiment on U2OS cells stably expressing mYFP-MacroD2 full-length and mCherry-H2B and used two different energy amounts to induce DNA damage **(see Methods 6.8)**. The speed of depletion of MacroD2 nuclear signal was proportional to the amount of laser energy irradiated and, consequently, to the amount of DNA damage that was induced, confirming the presence of a DNA damage-dependent response **(Figure 3.2)**.

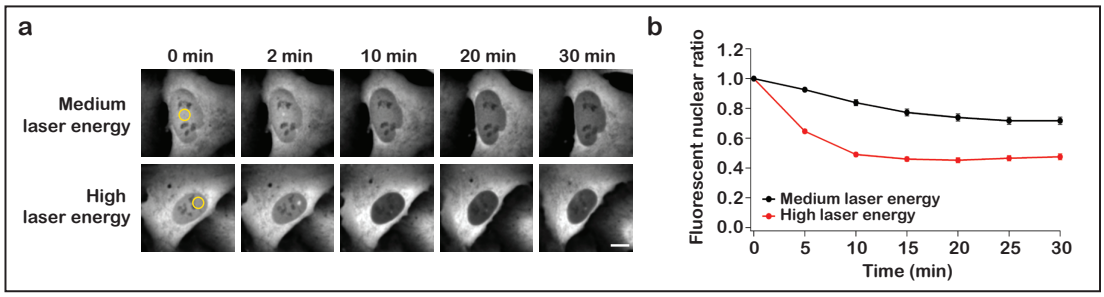


Figure 3.2: MacroD2 nuclear signal depletion occurs with a DNA damage dosage-dependent manner - a)Microirradiation live-cell microscopy using U2OS cells stably expressing the mYFP-MacroD2 full length and the mCherry-H2B constructs. Arbitrary medium and high laser energy values are used to induce different amounts of damage to the cell. Scale bar, 10 μm. b) 50−100 cells were quantified from three independent experiments **(see Methods 6.8)**. Nuclear/cytoplasmic ratio was calculated using CellProfiler 2.0 (331), the mCherry signal was used for the nuclei segmentation. Error bars, 95 %CI.

UV-laser microirradiation induces different types of DNA damage lesions according to the amount of energy delivered (332). Therefore, it is impossible to define the exact amount of the various DNA lesions that are generated with this experimental setup. To gain better insight into MacroD2's response to these different types of damage, I decided to test the MacroD2 decrease in nuclear signal in the presence of small molecule compounds, which cause specific types of DNA damage. For single-strand breaks (SSBs), I chose the drug camptothecin. Camptothecin blocks the activity of the

class I of topoisomerases, enzymes that are important for the process of replication and transcription (111). During these processes, the topoisomerase I nicks a single strand in order to release the surplus of topological energy of the DNA molecule in a controlled manner. Therefore, by not allowing the topoisomerase to perform the ligation step, the camptothecin drug induces SSBs in DNA.

I found that when camptothecin is used, MacroD2 nuclear signal depletion is limited to only a subset of cells, as shown in preliminary experiments **(see Figure 1.15)**. Since only a fraction of cells responds to the treatment, I hypothesized that the exclusive generation of SSBs does not induces the phenomenon. Since in S-phase, SSBs are converted into double-strand breaks (DSBs) upon replication (111), the few responsive cells upon camptothecin treatment might be in S-phase, when their SSBs have been converted into DSBs. Therefore, if only SSBs are generated, the cell-cycle could influence the decrease of MacroD2 nuclear protein levels.

To test this hypothesis, I synchronized U2OS stably expressing mEGFP-MacroD2 via aphidicolin treatment, which blocks cells at the G1-S interphase **(see Method 6.7.4)**. Shortly after release, the cells were treated with camptothecin and imaged. I then checked the percentage of cells responsive to the camptothecin treatment, in synchronized and not synchronized cells **(Figure 3.3)**. Indeed, the number of cells showing MacroD2 nuclear signal depletion increased when cells were synchronized and released in S-phase. Therefore, SSBs may be poor inducers of the decrease of MacroD2 nuclear levels, compared to DSBs.

To test whether DSBs are the cause of the MacroD2 nuclear depletion, I used the etoposide compound in the same assay. Unlike camptothecin, etoposide targets class II topoisomerases, whose activity consists in interrupting the DNA phosphate backbone at both strands. The drug freezes the topoisomerase molecule before the ligation, thus forming DSBs. I tested the MacroD2 nuclear signal depletion in U2OS stably expressing mYFP-MacroD2 and mCherry-H2B upon etoposide and DMSO control treatment **(Figure 3.4)**. In contrast to the camptothecin experiment, the response upon etoposide treatment was robust and strong, suggesting that DSB formation activates the signaling pathway that leads to the depletion of the nuclear MacroD2 signal. Since etoposide strongly induces the nuclear MacroD2 signal depletion in a reproducible manner, I used this treatment for further tests regarding the MacroD2 nuclear signal depletion.

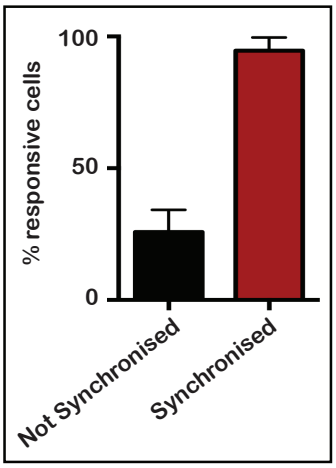


Figure 3.3: Double-strand breaks induce MacroD2 decrease in nuclear signal better than single-strand breaks - Percentage of cells showing MacroD2 nuclear signal depletion in U2OS cells stably transfected with mEGFP-MacroD2, synchronized or not with aphidicolin and treated with camptothecin 50 μM. Cells analyzed from one experiment, three imaging fields. Error bars, SD.

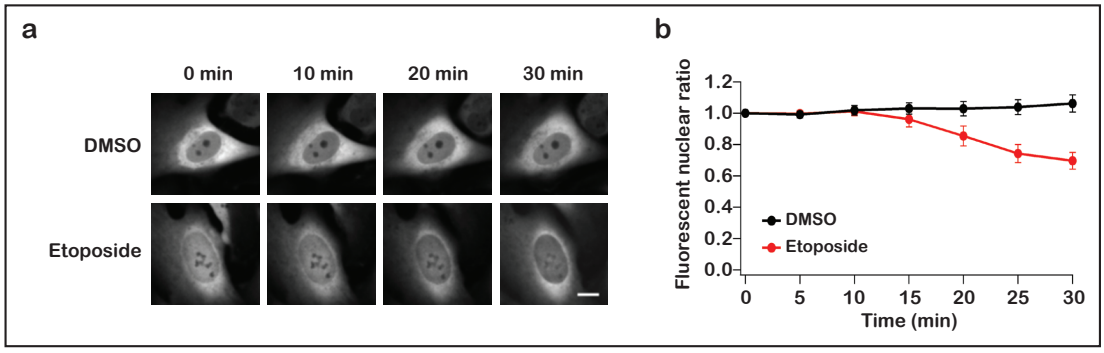


Figure 3.4: Etoposide strongly induces MacroD2 nuclear signal depletion - a)Live-cell imaging using U2OS-mYFP-MacroD2 full length+mCherry-H2B cells treated either with DMSO or etoposide 10 μM. Scale bar, 10 μm. b) 50−100 cells were quantified from three independent experiments **(see Methods 6.8)**. Nuclear/cytoplasmic ratio was calculated using CellProfiler 2.0 (331), and the mCherry signal was used for the nuclei segmentation. Error bars, 95 %CI.

The next step was to determine whether MacroD2 nuclear signal depletion is due to degradation upon DNA damage, in contrast to the hypothesis that it is caused by nuclear export. Therefore, I performed a cyclohexamide chase, to test whether the depletion of MacroD2 is due to protein degradation **(see Methods 6.10)**. Cyclohexamide is a compound that blocks the translation elongation step of the ribosomal complex, impairing the cellular protein synthesis. Once the protein synthesis is blocked by cyclohexamide, the protein levels of MacroD2 can be compared in the presence and absence of DNA damage, in order to show if there is induction of the MacroD2 protein degradation upon DNA damage response. Thus, I treated U2OS cells stably-expressing mYFP-MacroD2 with cyclohexamide or ethanol as negative control. After 30 minutes I treated the same cells either with DMSO or etoposide to induce DNA damage. Then, I collected the cells over a 24 hours time period (0, 45 minutes, 2 hours, 4 hours and 24 hours for both cyclohexamide and ethanol treatments). The following immunoblots were used for quantitative analysis **(Figure 3.5)**. Since I did not observe significant changes in EGFP-MacroD2 protein levels upon etoposide treatment in the timescale of our microscopy studies (45 minutes), I concluded that the MacroD2 nuclear depletion is not due to regulated protein degradation.

In addition to the cyclohexamide chase, to investigate if the observed nuclear depletion of MacroD2 upon DNA damage is due to nuclear export and not to protein degradation, I tracked the nuclear and cytoplasmic intensity of EGFP-tagged MacroD2 expressed in HeLa cells upon UV-laser microirradiation. Over time, the MacroD2 nuclear intensity is reduced almost to the half of the initial reading, while the cytoplasmic intensity increases by 15 % **(Figure 3.6)**. This disparity in fluorescence changes between the two set of intensities is compatible with the difference in the volume between the nuclear and the cytoplasmic compartments. This experiment further excludes that MacroD2 nuclear depletion is explained by its regulated degradation. Instead, it confirms that MacroD2 is exported from the nucleus upon DNA damage.

Since the ectopically-expressed EGFP-MacroD2 construct is reproducibly exporting, I wanted to test if the endogenous MacroD2 is also exporting. To this end, I needed a working antibody, since the commercially-available antibodies for MacroD2 fail to detect the correct protein size via immunoblot. An anti-MacroD2 serum was prepared by the in-house animal facility, where the rabbits have been immunized with purified V5-MacroD2 macrodomain prepared by G. Jankevicius. To increase the purity and the

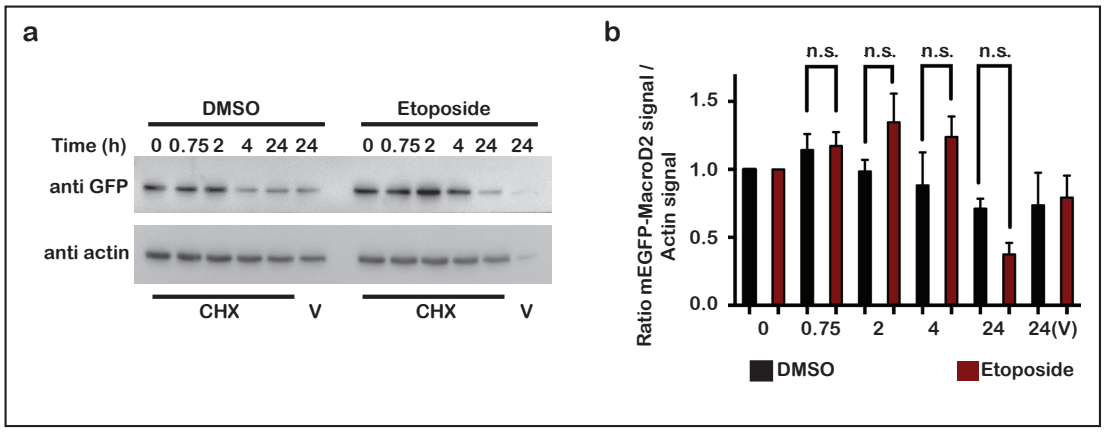


Figure 3.5: Etoposide treatment does not reduce the total mYFP-MacroD2 protein levels in 30 minutes - a)Cyclohexamide chase experiment **(see Methods 6.10)**. U2OS-mYFP-MacroD2 full length+mCherry-H2B cells were pretreated with cyclohexamide (CHX) or ethanol (V) for 30 minutes. At time 0, either DMSO or etoposide (10 μM) was added. Cells were then collected over a 24 hours time period. Immunoblot decorated with anti-GFP or anti-actin. b) Quantification of (a) from four independent experiments. Statistics performed with unpaired non-parametrical Mann-Whitney test. Error bars, SEM.

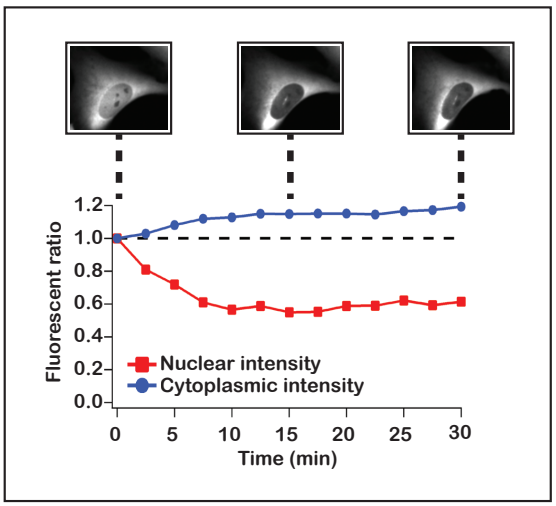


Figure 3.6: The cytoplasmic signal of EGFP-MacroD2 increases upon UV-lased microirradiation - DNA damage was induced with UV-microirradiation (high laser energy) on HeLa cells stably expressing mCherry-H2B and transfected with mEGFP-MacroD2. The quantification of nuclear and cytoplasmic intensity over time was performed by means of the software Fiji (`http://fiji.sc/Fiji`) **(see Methods 6.8.4)**. One cell is shown as example

specificity of the antibody, I performed the affinity-purification of the seventh bleed of anti-MacroD2 serum, following the standard protocol generally used in the department **(see Methods 6.11)**. The elution step was performed with three different buffers (citrate, glycine and phosphate buffer) and I tested their performance by immunoblot with different amounts of purified V5-MacroD2 macrodomain protein **(Figure 3.7)**. As expected, the antibody recognize the macrodomain of MacroD2 and the best performing batch was the one obtained by the glycine elution, which was always used for my subsequent experiments.

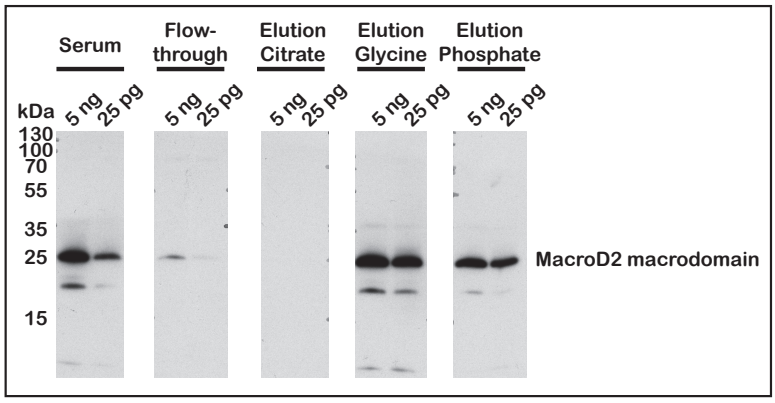


Figure 3.7: Anti-MacroD2 serum was purified by affinity-purification - V5-MacroD2 macrodomain was coupled to CNBr-Sepharose 4B beads. The affinity-column prepared was then used to purify the serum. Three types of buffers (citrate, glycine and phosphate buffer) were used for the elution steps **(see Methods 6.11)**. A western blot with 5 ng and 25 pg of purified V5-MacroD2 macrodomain, used also for raising the antibody, was used to define the purification quality. MacroD2 macrodomain runs at the height of 25 kDa.

To test the specificity of the antibody, I pre-cleared the anti-MacroD2 antibody by incubating it with recombinant MacroD2 full-length protein. I then used the pre-cleared antibody for an immunoblot with different cell lysates (U2OS mYFP-MacroD2 full-length; U2OS; HeLa mCherry-H2B; HEK293 TREx) and purified MacroD2 full-length recombinant protein **(Figure 3.8)** in comparison to the non-cleared antibody. Upon pre-clearing, both the endogenous and recombinant MacroD2 bands showed a reduction in intensity. However, the unspecific band around 110 kDa became stronger, possibly due to a larger amount of secondary antibodies available for the recognition of those bands.

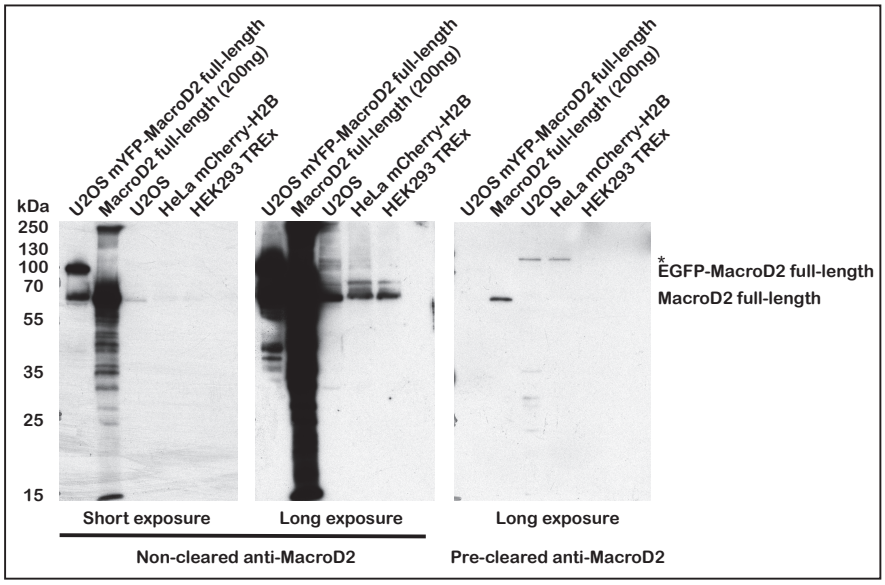


Figure 3.8: Anti-MacroD2 antibody specifically recognize endogenous and recombinant MacroD2 - Cell lysates (U2OS mYFP-MacroD2 full-length; U2OS; HeLa mCherry-H2B; HEK293 TREx; for each, 30 μg loaded) and 200 ng purified MacroD2 full-length were run for a western blot **(see Methods 6.9)**. The primary antibody (anti-MacroD2) has been previously incubated for one hour at 4 °C with 1 mg/mL purified MacroD2 full-length or water. For the non-cleared antibody section of the membrane, both short and long exposition are shown; for the pre-cleared antibody section, only long exposition is shown. * is unspecific band

To further confirm that the band at 70 kDa is specific for MacroD2, I treated U2OS cells for RNAi-mediated depletion, by using control or MacroD2-targeting siRNA **(see Methods 6.7.3)**. Then, I performed a western blot from the cell lysates **(Figure 3.9)**. The 70 kDa band is not present in the knock-down samples, indicating that this band is MacroD2.

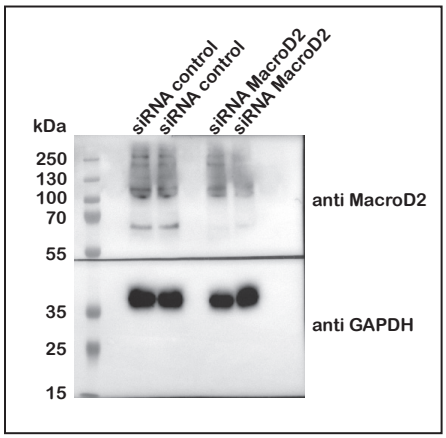


Figure 3.9: RNAi-mediated knock-down of MacroD2 confirms the partial specificity of anti-MacroD2 antibody - U2OS cells treated either with siRNA against MacroD2 or negative control. Collection of cells and lysate preparation at 72 hours post transfection. Immunoblot with anti-MacroD2 and anti-GAPDH, used as loading control.

The presence of the unspecific band around 110 kDa makes this a not ideal antibody. Nonetheless, being the best tool available, this antibody was used to test endogenous MacroD2 nuclear export by immunofluorescence microscopy **(see Methods 6.8.2)**. To this end, I compared the wild-type U2OS to the U2OS MacroD2 knock-out cell line generated by G. Möller by means of the CRISPR-Cas system (333). Cells were treated with DMSO or etoposide for 1 hour, fixed and processed according to the protocol. In wild-type U2OS, the nuclear signal showed depletion after 1 hour etoposide treatment, while the knock-out cells did not **(Figure 3.10)**. The export is clearly detectable, showing the physiological nature of the regulation of MacroD2 localization upon DNA damage. However, this antibody does not show the specific localization of MacroD2, because of the impurities still present after the purification.

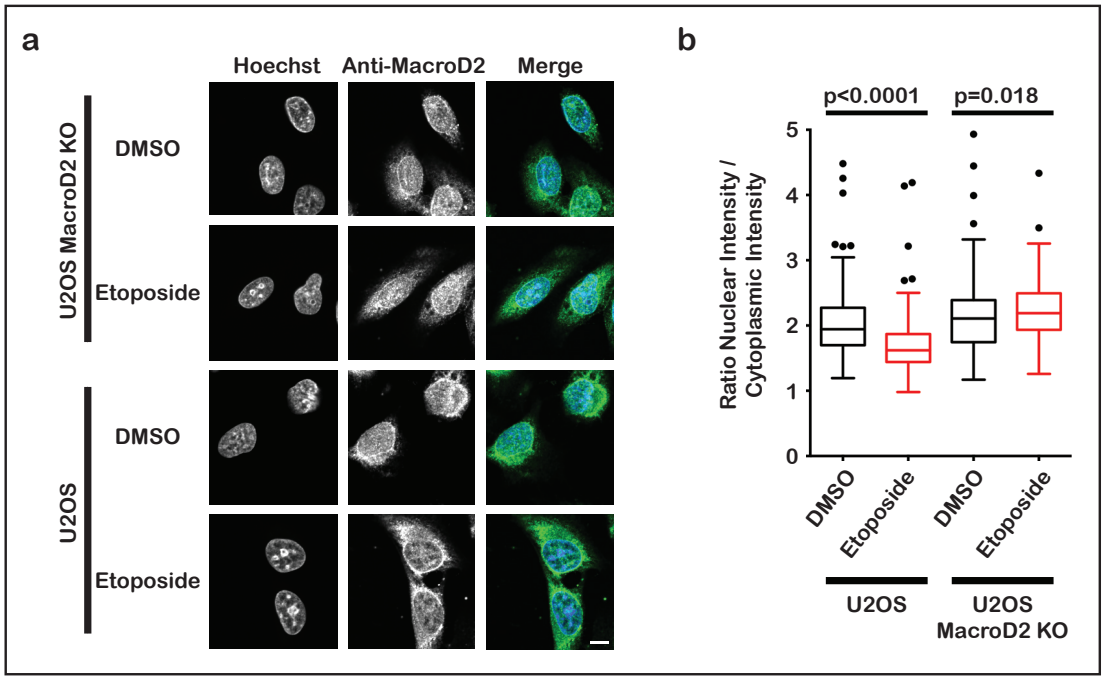


Figure 3.10: The endogenous MacroD2 exports from the nucleus upon one hour etoposide treatment - a)Immunofluorescence experiment with anti-MacroD2 antibody and Hoechst dye on U2OS and U2OS MacroD2$^{-/-}$ cells treated either with DMSO or etoposide 10 μM for 1 hour. Scale bar, 10 μm. b) Quantification of (a) in three independent experiments **(see Methods 6.8.2)**. Nuclear/cytoplasmic ratio was calculated using CellProfiler 2.0 (331), the Hoechst signal was used for the nuclei segmentation. Statistics performed with unpaired non-parametrical Kolmogorov-Smirnov test. Error bars, Tukey fences.

3.3 The C-terminal region of MacroD2 drives the nuclear export

MacroD2 has no known nuclear localization signal (NLS) nor a nuclear exporting sequence (NES). Indeed, several online predicting programs, such as NLS mapper, NL-Stradamus, SeqNLS, NES Finder 0.2, Net NES 1.1and LocNES, fail to detect any localization sequence with satisfactory confidence **(data not shown)**. Therefore, it is possible that its mode of export could differ from those of other exporting proteins.

When observing the protein sequence, MacroD2 protein consists of an annotated N-terminal globular macrodomain (aa 59-240) and a long C-terminal region with unknown function. The globular macrodomain recognizes with high affinity mono-ADP-ribosylated protein and removes the modification (56, 67). The C-terminus, instead, is not annotated and the only sequences with more than 50 % identity are from other vertebrate *MacroD2* genes via BLAST search. Therefore, I tested which part of the MacroD2 protein is sufficient for the regulated nuclear export. I expressed in HeLa cells EGFP-tagged MacroD2 full-length, macrodomain or the C-terminal region. Then, I treated the cells with etoposide and imaged using live-cell imaging **(Figure 3.11)**. Both the full-length and the C-terminal constructs exported, while the macrodomain alone did not. This indicates that the C-terminal region of MacroD2 is responsible for its nuclear export.

Due to the lack of localization signals, I decided to explore the protein sequence and structure of MacroD2. Although the structure of the MacroD2 macrodomain has been already solved (56), the C-terminus region has never been studied. Online predictions for the protein secondary structure, such as PredictProtein, indicate that the C-terminal stretch should be unstructured, with high likelihood being exposed to the cell environment **(Figure 3.12;** (334)**)**.

To confirm this prediction, I expressed and purified MacroD2, as a full-length and C-terminal region (244-488) constructs **(see Methods 6.12)**. Although the C-terminal region is predicted to be unstructured, the construct was soluble and purified without precipitation **(Figure 3.13)**. Also, the MacroD2 C-terminal region runs at the height of 50 kDa, even though its length would account for 25 kDa. This behavior is probably due to the compositional bias of the C-terminal sequence, as it is also shared by the

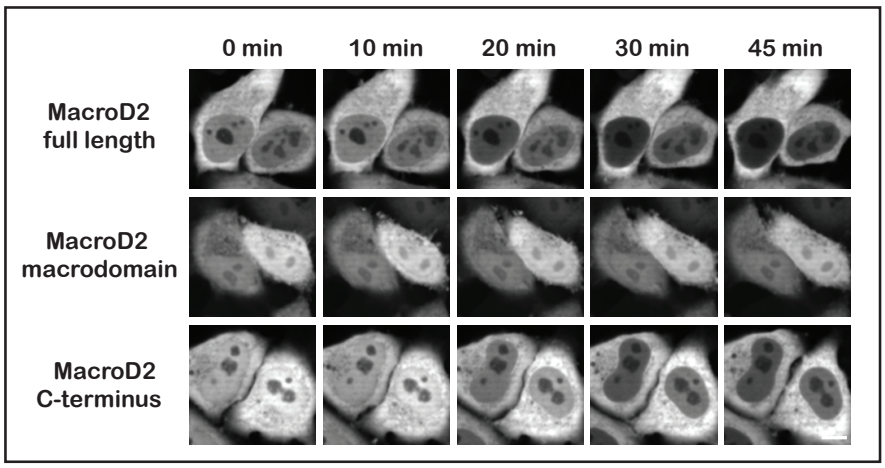


Figure 3.11: The C-terminal region is sufficient for MacroD2 nuclear export - Live-cell imaging of tagged mEGFP-MacroD2 constructs (full-length, macrodomain or C-terminal region) in HeLa-mCherry-H2B cells treated either with DMSO or etoposide 10 μM **(see Methods 6.8)**. Scale bar, 10 μm.

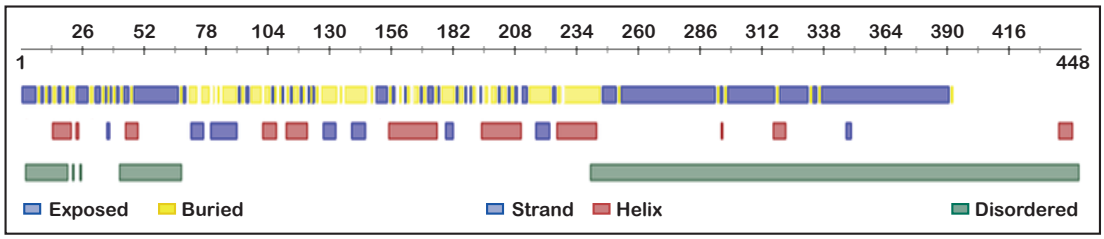


Figure 3.12: The C-terminal portion of MacroD2 is predicted to be unstructured - Predictions of MacroD2 secondary structure. Sequence of MacroD2 Isoform 1 of Uniprot was analysed with the online software PredictProtein (334). The first line shows prediction for each residue to be exposed or buried; the second line shows prediction for the presence of helices or strands of β-sheets; the third line shows prediction of disorder.

MacroD2 full-length protein. The constructs were then further purified by a size-exclusion column.

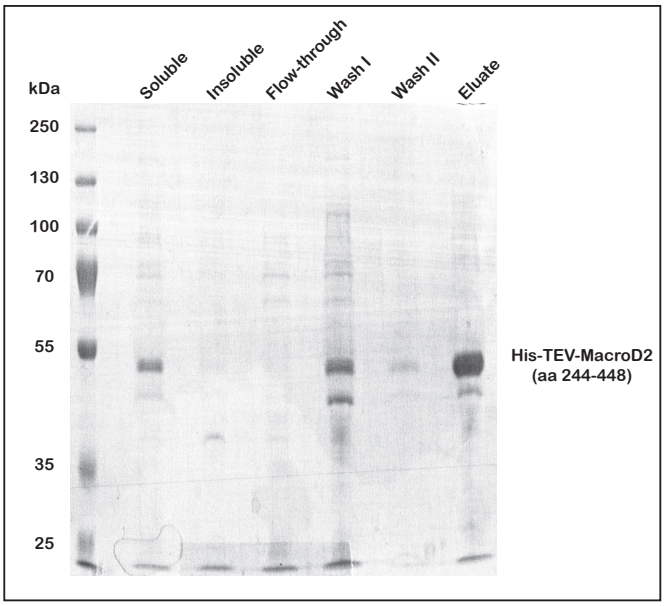


Figure 3.13: The MacroD2 C-terminal region is expressed and purified - His-TEV-MacroD2 C-terminal construct was expressed in *E. coli* and purified. The construct runs at the height of 50 kDa, even though its length would account for 25 kDa. In the Coomassie-stained gel, the different steps of the purification are shown.

I then analyzed the three purified proteins (full-length and C-terminus purified by me, macrodomain purified by G. Jankevicius) by circular dichroism, with the help of Dr. F. Kamp, LMU **(see Methods 6.13)**. Circular dichroism is generally used to investigate the secondary structure of the proteins by analyzing the change in the circularly polarized light that passes through a solution with the pure protein. Although it is not as accurate as a crystal structure, circular dichroism can show which is the predominant secondary structure of the protein that is being analyzed.

The spectrum of the macrodomain showed a predominant presence of α-helices, given to the strong positive value around 190 nm and the negative (and almost parallel to the x-axis) portion of the curve in the interval 209-222 nm **(Figure 3.14)**. This result is consistent with the percentages calculated from the crystal structure, with 37 % of α-helices and almost 15 % of β-sheets.

The result also confirmed the prediction for the structure of the C-terminus, since

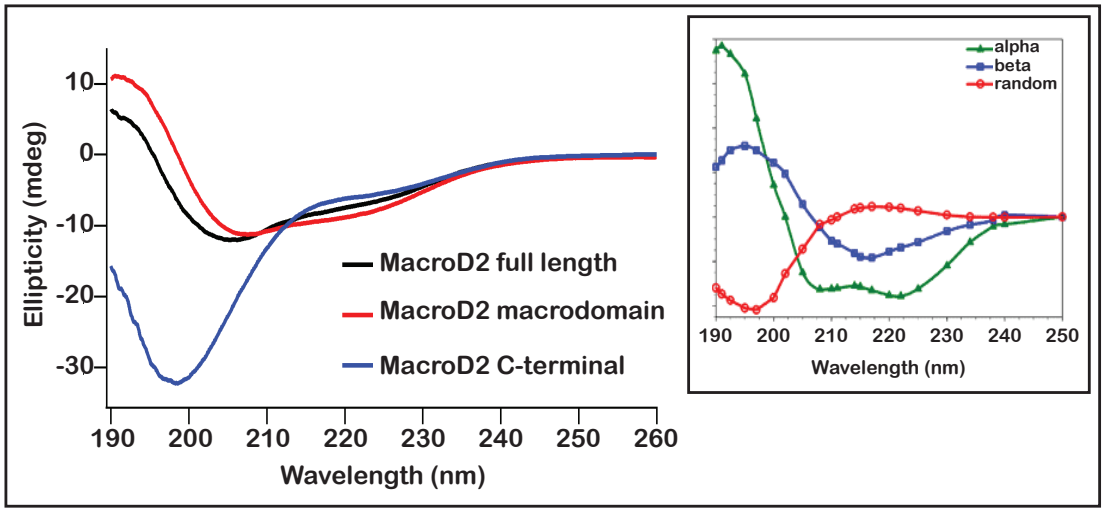


Figure 3.14: The MacroD2 C-terminal region is unstructured - Circular dichroism experiment on purified MacroD2 full-length, macrodomain and C-terminal region constructs (5 μM; 20 mM phosphate buffer pH 7.3). In the box, examples for the α-helix (alpha - green), β-sheet (beta - blue) and random coil (random - red) spectra (`http://www.fbs.leeds.ac.uk/facilities/cd/images/1.png`)

the C-terminal fragment showed a strong unstructured behavior, as shown by the deep negative valley at 198 nm and the almost positive value at 212 nm (335). The full-length construct showed mostly similarities to the macrodomain folding pattern. Nonetheless, it has a lower content of α-helices and the presence of random coils when compared to the macrodomain alone, explaining the shift towards negative values in the portion of the curve between 190 and 205 nm.

3.4 ATM kinase activity induces MacroD2 nuclear export

As previously mentioned, preliminary experiments showed that ATM kinase activity could induce MacroD2 nuclear export **(see Figure 1.15)**. Cells were pre-treated with camptothecin and then treated with an ATM inhibitor KU55399. The few cells responding to the camptothecin with MacroD2 nuclear export showed a net import of the protein upon treatment with ATM inhibitor.

ATM belongs to the PI3K-like kinases (276, 336). This class of kinases consists of six members: ATM, ATR, DNA-PK, SGM-1, mTOR and TRRAP. Of these, ATM, ATR and DNA-PK play an important role in the DNA damage response by regulating

the early phase of factor recruitment, cell-cycle checkpoints and the choice of repair pathway to activate (276). Since all three are involved in DNA damage repair regulation and might even share a subset of common targets, I wanted to define which was the specific contribution of each of the kinases to MacroD2 export, by using specific small molecule inhibitors: KU55933, VE-821 and NU7441 for ATM, ATR and DNA-PK, respectively. Therefore, I pre-treated U2OS stably expressing mYFP-MacroD2 full-length with these inhibitors for 30 minutes before performing the imaging upon laser microirradiation **(see Methods 6.8)**. The ATM inhibitor successfully inhibited MacroD2 nuclear export, while ATR and DNA-PK inhibitors did not. In fact, the inhibition of ATR minimally enhances the export **(Figure 3.15)**. This might be explained by the fact that inactivation of ATR could lead to the increased activation of ATM, since less competition would take place at the DNA damage site.

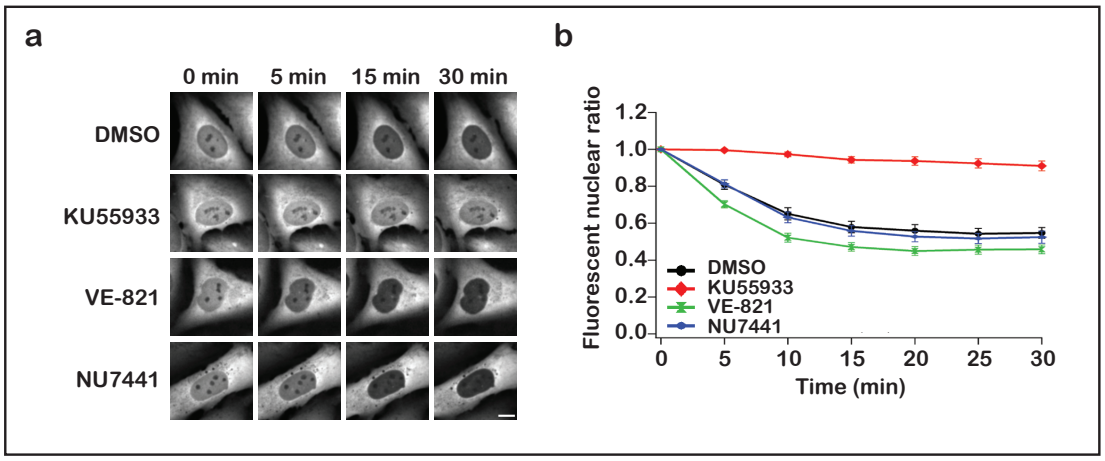


Figure 3.15: Among the PI3K-like kinases, only ATM induces MacroD2 nuclear export - a) Microirradiation live-cell microscopy of U2OS cells stably expressing mYFP-MacroD2 full length and mCherry-H2B constructs, pretreated with KU55933 10 μM, VE-821 1 μM, NU7441 1 μM or DMSO for 30 minutes. Medium laser energy was used. Scale bar, 10 μm. b) 50−100 cells were quantified from three independent experiments **(see Methods 6.8)**. Nuclear/cytoplasmic ratio was calculated using CellProfiler 2.0 (331), and the mCherry signal was used for the nuclei segmentation. Error bars, 95 %CI.

To confirm ATM involvement in the induction of MacroD2 nuclear export, I performed a similar experiment using etoposide. I pre-treated U2OS cells stably expressing mYFP-MacroD2 full-length with either DMSO or ATM inhibitor for 30 minutes. Then, I performed the experiment by adding DMSO or etoposide while imaging **(see Meth-**

ods 6.8). The experiment showed that ATM inhibition blocks MacroD2 export also when DNA damage is induced by etoposide **(Figure 3.16)**.

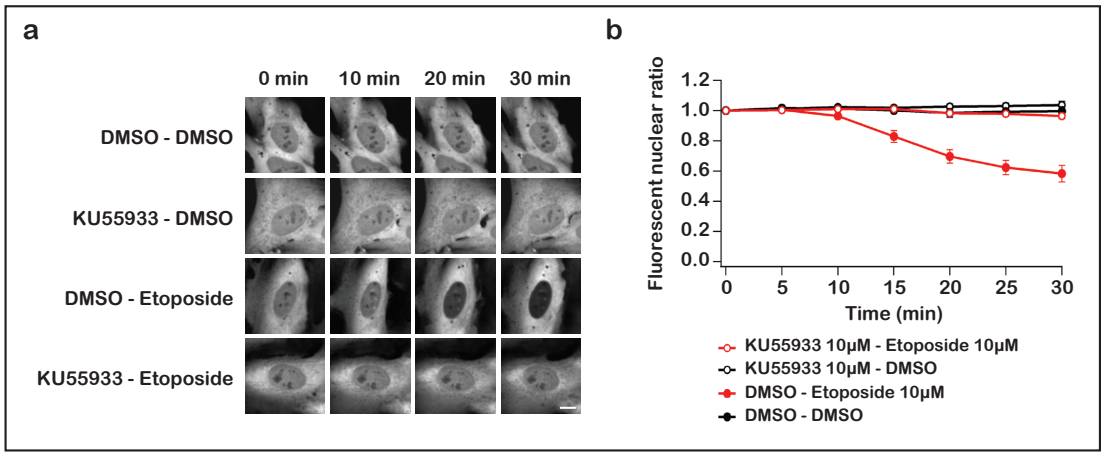


Figure 3.16: ATM inhibition blocks MacroD2 nuclear export upon etoposide treatment - a)Live-cell imaging of U2OS-mYFP-MacroD2 full length+mCherry-H2B cells pre-treated with either DMSO or KU55933 10 M for 30 minutes, and treated either with DMSO or etoposide 10 μM. Scale bar, 10 μm. b) 50−100 cells were quantified from three independent experiments **(see Methods 6.8)**. Nuclear/cytoplasmic ratio was calculated using CellProfiler 2.0 (331), and the mCherry signal was used for the nuclei segmentation. Error bars, 95 %CI.

While small molecule compounds are a very powerful tool, the main drawback for their use is the chance in affecting other proteins that might share some similarity with the main target. For this reason, it is advisable to confirm the functional interaction with at least another approach. Thus, to further validate the involvement of ATM, I wanted to perform the siRNA-mediated knockdown of ATM.

To check the efficiency of the knock-down, I performed an immunblot with whole cell lysate from U2OS cells treated with siRNA against ATM or a control siRNA for 72 hours **(see Methods 6.7.3)** and probed it with anti-ATM antibody and anti-GAPDH as loading control. The RNAi treatment drastically reduced ATM protein levels **(Figure 3.17)**. Hence, the residual expression of ATM implies that a residual ATM enzymatic activity could still be activated and may generate an intermediate phenotype, rather than the complete block of MacroD2 nuclear export.

To test the MacroD2 nuclear export upon ATM-knockdown, I imaged U2OS cells expressing EGFP-MacroD2 C-terminal region upon etoposide treatment **(see Meth-**

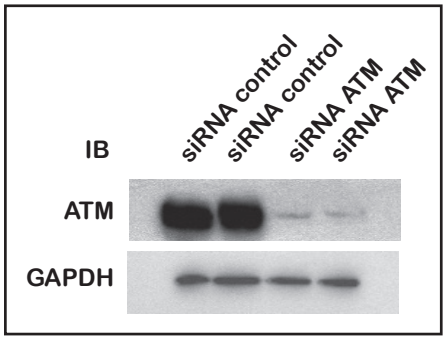


Figure 3.17: RNAi against ATM successfully reduces ATM protein levels - Immunoblotting with anti-ATM in U2OS-mYFP-MacroD2 full length+mCherry-H2B cells treated either with siRNA against ATM or negative control siRNA 72 hours post transfection **(see Methods 6.7.3)**. GAPDH is used as loading control.

ods 6.7.3 and 6.8). The export dynamics of the C-terminal region construct was reduced by half when ATM is depleted in cells **(Figure 3.18)**. As expected, given that the ATM knock-down was incomplete, the experiment showed residual export of MacroD2 C-terminus region.

I then performed the same experiment with U2OS cells stably expressing MacroD2 full-length construct. In this case, the reduction of MacroD2 nuclear export was minimal and restricted to the time-window of 15-25 minutes after etoposide addition **(Figure 3.19)**.

I then wanted to check if the reduced effect of ATM knock-down on MacroD2 full-length nuclear export is restricted to the etoposide treatment. Therefore, I tested the U2OS stably expressing MacroD2 full-length upon UV-microirradiation **(see Methods 6.8)**. Also in this condition, MacroD2 full-length nuclear export only minimally changes between the two conditions **(Figure 3.20)**. I detected a minor difference in the export speed of the ATM-depleted cells compared with the control in case of high laser energy, while such a difference was lost when I applied a modest amount of DNA damage.

So, while the ATM inhibitor successfully blocked the nuclear export of MacroD2, the RNAi-mediated treatment showed more contrasting results. These could be explained either by the residual presence of ATM, due to the incomplete depletion of the protein, or by a possible redundancy with other kinases, that appears clearly only upon specific conditions, such as the decreased activity of ATM. Thus, to further test the involvement of ATM in the induction of MacroD2 nuclear export, we acquired two

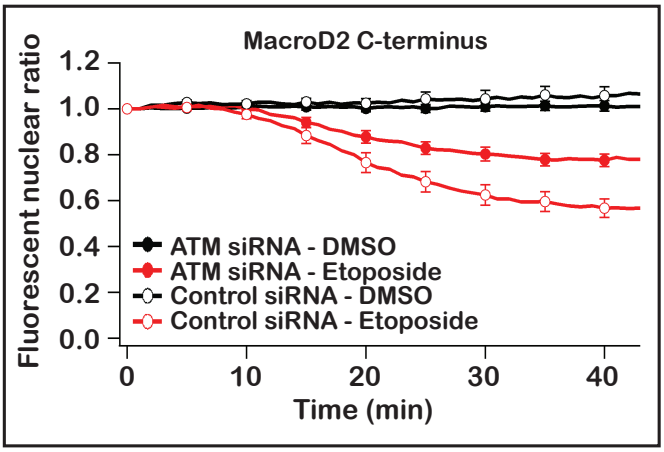


Figure 3.18: siRNA-mediated knock-down of ATM reduces the nuclear export of the MacroD2 C-terminus fragment - Quantification of live-cell imaging experiments with U2OS cells stably expressing the mEGFP-MacroD2 C-terminal region+mCherry-H2B, pretreated for 72 hours with siRNA against ATM or negative control siRNA, and treated with DMSO or etoposide 10 μM. 50−100 cells were quantified from two independent experiments **(see Methods 6.8)**. Nuclear/cytoplasmic ratio was calculated using CellProfiler 2.0 (331), and the mCherry signal was used for the nuclei segmentation. Error bars are 95 %CI.

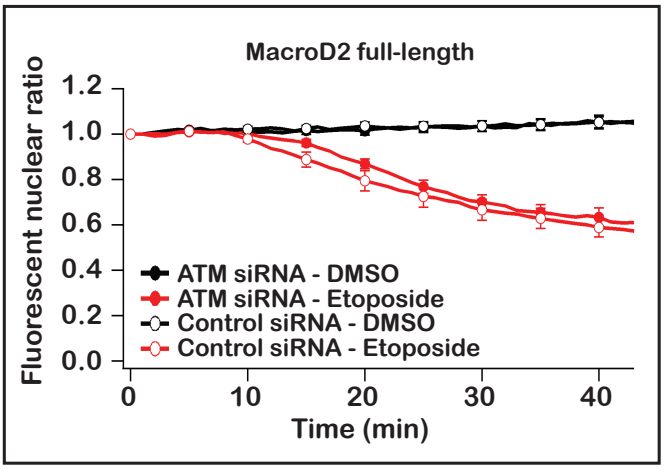


Figure 3.19: ATM-depletion minimally affects the export of the MacroD2 full-length construct - Quantification of live-cell imaging experiments with U2OS cells stably expressing the mYFP-MacroD2 full-length, pretreated for 72 hours with siRNA against ATM or negative control siRNA, and with DMSO or etoposide 10 μM. 50−100 cells were quantified from two independent experiments **(see Methods 6.8)**. Nuclear/cytoplasmic ratio was calculated using CellProfiler 2.0 (331), and the mCherry signal was used for the nuclei segmentation. Error bars are 95 %CI.

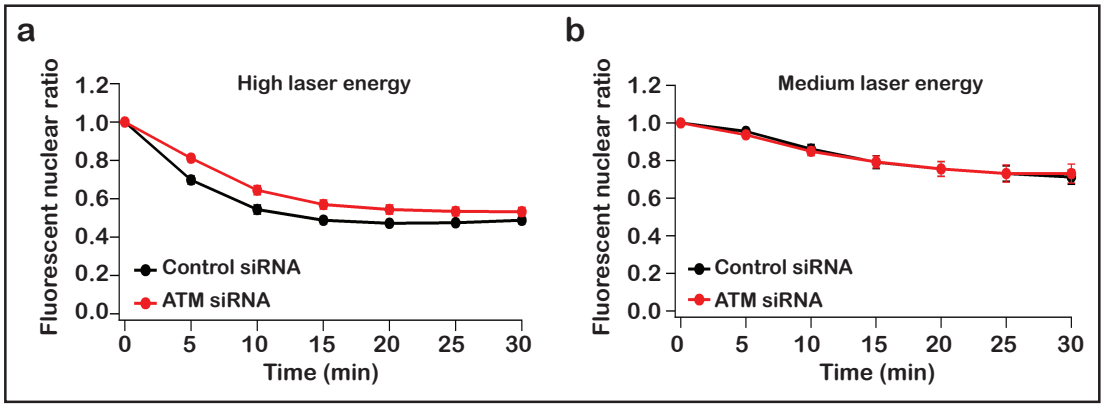


Figure 3.20: The nuclear export of the MacroD2 full-length construct upon laser microirradiation is marginally affected by the ATM knock-down - a) Quantification of live-cell imaging experiments with high UV-laser energy. b) Quantification of live-cell imaging experiments with medium UV-laser energy. For (a) and (b), 50−100 U2OS cells stably expressing the mYFP-MacroD2 full-length and mCherry-H2B, pretreated for 72 hours with siRNA against ATM or negative control siRNA, were quantified from three independent experiments **(see Methods 6.8)**. Nuclear/cytoplasmic ratio was calculated using CellProfiler 2.0 (331), and the mCherry signal was used for the nuclei segmentation. Error bars are 95 %CI.

new cell lines, the G-361 ($ATM^{+/+}$) and the HT-144 ($ATM^{-/-}$). These two cell lines are both melanoma cells with similar, although not identical, genetic background. To check for the expression of ATM, I tested their cell lysates by immunoblots with anti-ATM antibody **(see Methods 6.9)**. ATM band was clearly detected in G-361 samples, however, it was absent in HT-144 sample lanes **(Figure 3.21)**. This result confirms that HT-144 is a knock-out cell line for ATM.

Then, to test whether MacroD2 can export in the cell line without ATM kinase, I transfected G-361 and HT-144 cells with EGFP-MacroD2 full-length and imaged the cells upon UV-laser microirradiation **(see Methods 6.8)**. While in the $ATM^{+/+}$ G-361 cell line MacroD2 exports as expected, in the $ATM^{-/-}$ HT-144 cell line MacroD2 nuclear export is not blocked **(Figure 3.22)**. This result indicates that another kinase could replace ATM in the induction of MacroD2 nuclear export.

Due to similarities between the different members of the PI3K-like kinases family, it is plausible that some redundancy between these kinases is shown upon specific conditions. To determine whether one of the other PI3K-like kinases might replace ATM in the induction of MacroD2 nuclear export when ATM is absent, I transfected

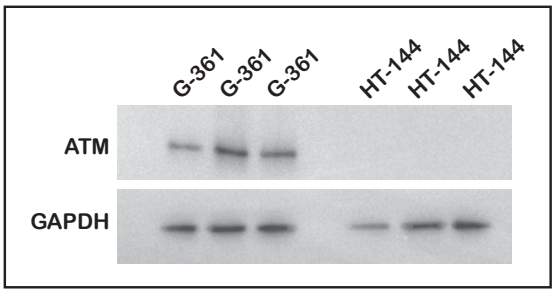


Figure 3.21: ATM$^{-/-}$ cells do not show ATM band at the immunoblot - Immunoblotting with anti-ATM using cell extracts of G-361 and HT-144 cells prepared in triplicate. GAPDH is used as loading control.

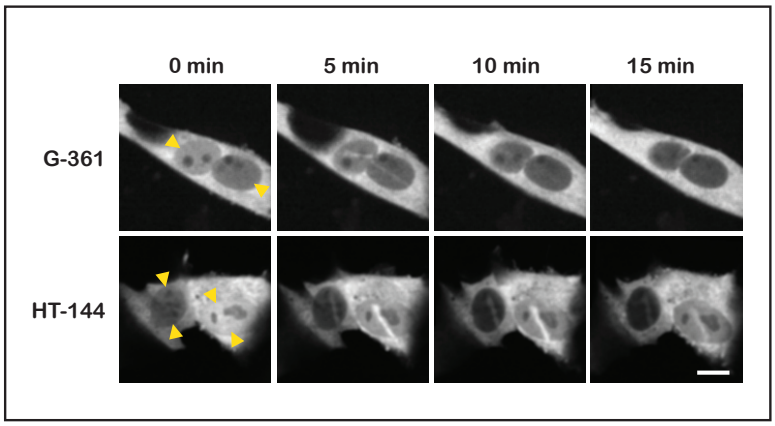


Figure 3.22: MacroD2 shows residual nuclear export in ATM$^{-/-}$ cells - Microirradiation live-cell microscopy of G-361 (ATM$^{+/+}$) and HT-144 (ATM$^{-/-}$) cells transfected with mEGFP-MacroD2 full-length construct. The focus of the laser micro-irradiation is indicated with yellow arrowheads **(see Methods 6.8)**. High laser energy was used. Scale bar, 10 μm.

HT-144 cells with EGFP-MacroD2 full-length and mCherry-H2B. I pretreated the cells with DMSO, ATM inhibitor (KU55933), ATR inhibitor (VE-821) or DNA-PK inhibitor (NU7441) for 30 minutes before inducing DNA damage by UV-microirradiation **(see Methods 6.7.3 and 6.8)**. While the ATM and the ATR inhibitors did not affect the nuclear export of MacroD2, DNA-PK inhibition resulted in the decrease of MacroD2 nuclear export. These results suggest that in ATM$^{-/-}$ HT-144 cells the role of ATM in inducing MacroD2 upon UV-laser microirradiation is taken over by DNA-PK **(Figure 3.23)**.

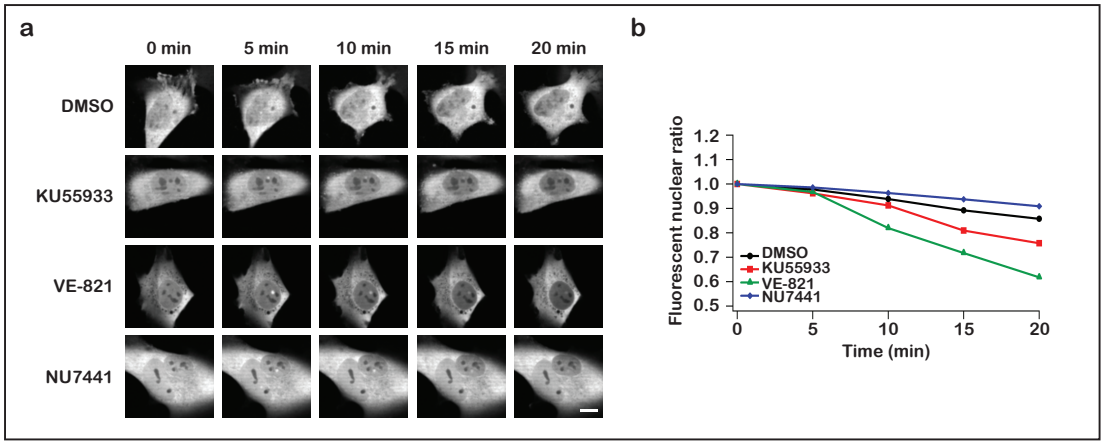


Figure 3.23: In ATM$^{-/-}$ cells DNA-PK induces MacroD2 nuclear export - a) Microirradiation live-cell microscopy of HT-144 (ATM$^{-/-}$) cells transfected with mEGFP-MacroD2 full-length and mCherry-H2B constructs. Cells were then pretreated with DMSO, KU55933 10 μM, VE-821 10 μM or NU7441 1 μM for 30 minutes **(see Methods 6.8)**. High laser energy was used. Scale bar, 10 μm. b) Quantification of an example cell. Nuclear/cytoplasmic ratio was calculated using CellProfiler 2.0 (331), the mCherry signal was used for the nuclei segmentation.

The redundancy in ATM activity could represent a back-up mechanism that is frequently activated, given the importance of the task of this kinase. On the other hand, it might occur only upon specific conditions. Therefore, to test whether this redundancy is not restricted to UV-laser microirradiation, the EGFP-MacroD2 nuclear export was tested in ATM$^{-/-}$ HT-144 upon etoposide treatment **(see Methods 6.7.3, 6.8 and 6.8.6)**. U2OS cells were used as positive control of the MacroD2 nuclear export. Unlike with the UV-laser, MacroD2 nuclear export was completely inhibited in HT-144 cells **(Figure 3.24)**. Therefore, upon etoposide treatment, MacroD2 nuclear

export is exclusively induced by ATM kinase. Such a result show that a specific DNA damage type might activate a specific pathway, which is able to induce the nuclear export of MacroD2 in a contained extent.

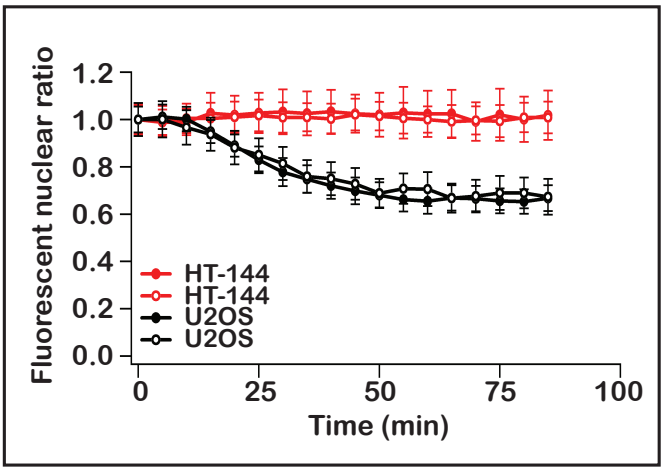


Figure 3.24: MacroD2 nuclear export is blocked when ATM$^{-/-}$ cells are treated with etoposide - Quantification of live-cell imaging experiments with U2OS and ATM$^{-/-}$ HT-144 cell, transiently transfected with EGFP-MacroD2 full-length and pre-treated with Hoechst, then treated with etoposide 10 μM. 50−200 cells were quantified from two independent transfections **(see Methods 6.7.3, 6.8 and 6.8.6)**. Nuclear/cytoplasmic ratio was calculated using CellProfiler 2.0 (331), the Hoechst signal was used for the nuclei segmentation. Error bars are 95 %CI.

3.5 MacroD2 macrodomain marginally modulates its nuclear export

In the Section 3.4, I showed that ATM induces MacroD2 nuclear export and that ATM can be replaced by other kinases according to the type of pathway that is activated. I also showed that the RNAi-mediated depletion of ATM affects the MacroD2 nuclear export in different ways, depending on the construct that has been transfected, either full-length MacroD2 or its C-terminal region **(see Figures 3.18 and 3.19)**.

Since the ATM knock-down is performed in the same way in the two experiments, the difference in MacroD2 behavior should not be caused by a difference in ATM activity. However, the two MacroD2 constructs, the full-length and the C-terminal fragment, differ by the presence (or absence) of the macrodomain. It is then important

to define whether the macrodomain affects the nuclear export of the whole protein. The macrodomain presents three features:

1. it is able to bind mono-ADP-ribose, which occurs at DNA lesions;
2. it has a mono-ADP-ribosyl-hydrolase activity, upon binding to modified targets;
3. it is a globular domain of 25 kDa, which might hinder the passage through the NPC by affecting the MacroD2 physical properties.

I then decided to test how these different features could affect MacroD2 nuclear export. Assuming that the recruitment to the DNA damage site is important for ATM to get close to MacroD2, I should see a decrease in its nuclear export dynamics by affecting MacroD2 recruitment. Thus, I pre-treated U2OS cells stably expressing MacroD2 full-length in the presence of DMSO or olaparib, a strong inhibitor of ARTD1, ARTD2 and ARTD3 (337). Olaparib treatment, thus, hinders the PAR formation at the DNA lesions. Successively, I tested the MacroD2 nuclear export dynamics upon UV-laser microirradiation **(see Methods 6.7.3 and 6.8)**. As consequence, the olaparib-treated cells showed no recruitment of MacroD2 to the DNA damage site **(Figure 3.25)**. Yet, MacroD2 nuclear export dynamics was not affected. Therefore, the recruitment of MacroD2 to the DNA lesion is not important for MacroD2 nuclear export.

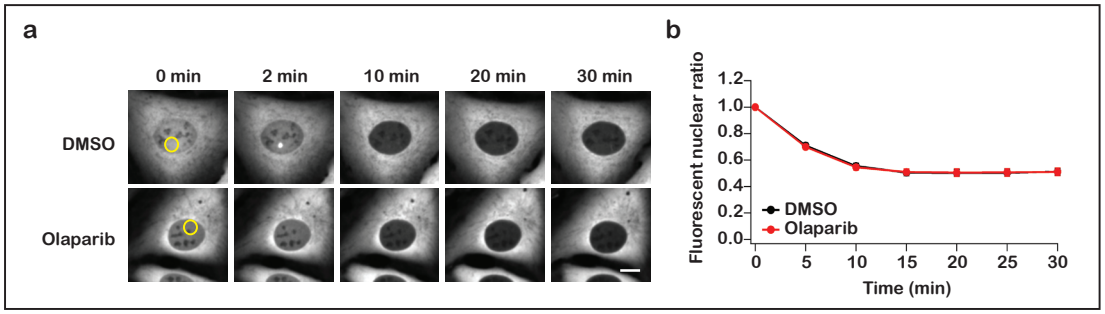


Figure 3.25: The recruitment to DNA lesions does not affect MacroD2 nuclear export - a)Microirradiation live-cell imaging of U2OS-mYFP-MacroD2 full-length+mCherry-H2B cells pretreated with either DMSO or olaparib 2 μM. Yellow circle show the microirradiation focus. High laser energy was used. Scale bar, 10 μm. b) 50−100 cells were quantified from three independent experiments **(see Methods 6.8)**. Nuclear/cytoplasmic ratio was calculated using CellProfiler 2.0 (331), and the mCherry signal was used for the nuclei segmentation. Error bars, 95 %CI.

To investigate if the abolishment of the binding to the ADP-ribose or the acquired binding to PAR could affect the dynamics of MacroD2 nuclear export, I used two published mutants. Previous studies on MacroD2 macrodomain have defined a series of mutation that might affect the binding or the catalytic activity (56). For example, the G188E mutation completely abolishes the binding of ADP-ribose. Additionally, the G100E+I189R+Y190R mutant no longer has catalytic activity, but is still able to bind the ADP-ribose. Moreover, this mutant is no longer selective for the mono-ADP-ribosylation, but gains the ability to bind PAR as well **(G. Jankevicius,** *unpublished***)**. Therefore, I transiently expressed these mutants in comparison to the wild-type in HeLa cells stably expressing mCherry-H2B and performed UV-laser microirradiation experiments**(see Methods 6.8)**.

The MacroD2 nuclear export dynamics of the three constructs were similar but not equal **(Figure 3.26)**. The MacroD2 G188E mutant that cannot bind any ADP-ribose was exported marginally faster than the wild-type construct. This minimal difference between the ADP-ribose-binding deficient mutant and the wild-type construct in the export rate is quite striking if compared to the total overlap upon olaparib treatment. This distinction could be explained by the fact that the MacroD2 macrodomain interacts with mono-ADP-ribosylation, while the main targets of olaparib are ARTD1 and ARTD2, which are poly-ADP-ribosyl-transferases. Therefore, affecting the binding of the macrodomain with any type of ADP-ribosylation induces a slightly faster export. On the other hand, the MacroD2 G100E+I189R+Y190N mutant that is capable to bind PAR as well as MAR is consistently marginally slower than the wild-type, maybe due to the increase in binding events that this protein can undergo.

The MacroD2 macrodomain is 25 kDa in size, as is EGFP. To test whether the size of the protein affects the export rate, with the help of J. Preißer, a C-terminal construct in a vector containing two EGFP in series was generated. This construct has a similar molecular weight compared to EGFP-MacroD2 full-length. Thus, by comparing the nuclear export of these two construct, two outputs are expected: one option is that the EGFP-EGFP-MacroD2 C-terminal construct behaves like the EGFP-MacroD2 full-length, meaning that the size but not the macrodomain affects the nuclear export. In the other scenario, where the EGFP-EGFP-MacroD2 C-terminal exports as the same fragment with only one EGFP attached, the macrodomain affects the nuclear export by

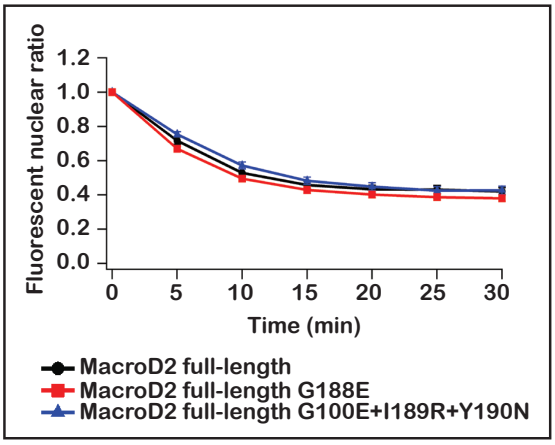


Figure 3.26: The macrodomain ability to bind ADP-ribose marginally affects the MacroD2 export rate - Quantification of microirradiation live-cell imaging of mEGFP-MacroD2 constructs (full-length wild type, full-length G188E or full-length G100E+I189R+Y190N) in HeLa cells stably expressing mCherry-H2B. 50−100 cells were quantified from three independent experiments **(see Methods 6.8)**. Nuclear/cytoplasmic ratio was calculated using CellProfiler 2.0 (331), and the mCherry signal was used for the nuclei segmentation. Error bars, 95 %CI For the microirradiation, high laser energy was used..

sourcing possible interactions. I thus transfected EGFP-MacroD2 full-length, EGFP-MacroD2 C-terminus and EGFP-EGFP-MacroD2 C-terminus constructs in HeLa cells stably expressing mCherry-H2B **(Figure 3.27)**.

The export dynamics of EGFP-EGFP-MacroD2 C-terminus shows a shape that is very similar to the one of the MacroD2 full-length construct. However, the two constructs shows a difference in the overall speed, maybe due to the fact that the EGFP-EGFP-MacroD2 C-terminus cannot bind any ADP-ribosylation and therefore it is faster in being exported. On the other hand, the EGFP-C-terminus construct shows a completely different export dynamics, suggesting that the main difference between the full-length and the C-terminus constructs is indeed just the size.

As shown previously **(see Figures 3.18 and 3.19)**, the RNAi-mediated depletion of ATM affects the dynamics of nuclear export of the MacroD2 C-terminus fragment, but only marginally the dynamics of the full-length protein. To test if these differences in behavior are indeed due to the presence of the macrodomain, four U2OS stable cell lines expressing EGFP-MacroD2 full-length, EGFP-MacroD2 full-length (G188E), EGFP-MacroD2 C-terminus or EGFP-EGFP-MacroD2 C-terminus were treated for 72

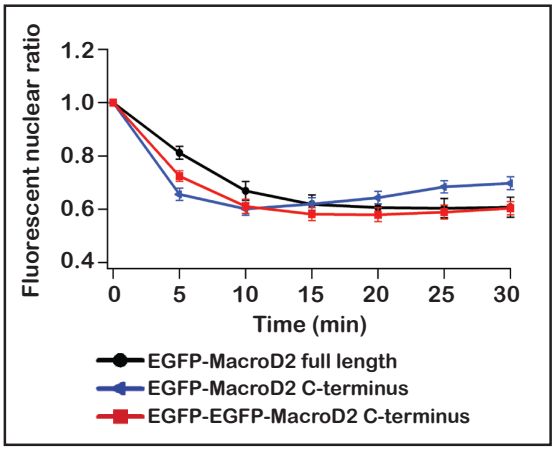


Figure 3.27: The EGFP-EGFP-MacroD2 C-terminus construct shows the same dynamic of export of the full-length protein - Quantification of microirradiation live-cell imaging of mEGFP-MacroD2 constructs (EGFP-MacroD2 full-length, EGFP-MacroD2 C-terminus or EGFP-EGFP-MacroD2 C-terminus) in HeLa cells stably expressing mCherry-H2B. Around 50 cells were quantified from three independent experiments **(see Methods 6.8)**. Nuclear/cytoplasmic ratio was calculated using CellProfiler 2.0 (331), and the mCherry signal was used for the nuclei segmentation. Error bars, 95 %CI. For the microirradiation, high laser energy was used.

hours with ATM siRNA or a negative control siRNA **(see Methods 6.7.3)**. On the third day, cells were pre-treated with Hoechst to stain the nuclei and imaged **(see Methods 6.8 and 6.8.6)**. The DNA damage was induced by etoposide treatment, since the Hoechst treatment, necessary for the nuclei segmentation, is not compatible with my settings of UV-laser microirradiation. In the control treatment, MacroD2 full-length and C-terminus construct showed the same behavior that was seen previously **(see Figure 3.28, *panel b*)**.

The EGFP-MacroD2 full-length G188E mutant also behaves as shown previously, being slightly faster than the EGFP-MacroD2 full-length. On the contrary, the EGFP-EGFP-MacroD2 C-terminus construct maintains its previously shown behavior only in the initial phase of the export dynamics, since it exports up to 20 minutes with a rate that is intermediate between the full-length and the C-terminus construct; successively, the C-terminus protein with two EGFP attached slows down in the export, becoming slower than the full-length **(see Figure 3.28, *panel a*)**. Since this experiment has been performed upon etoposide treatment and not UV-laser, it is possible that this change might have slightly changed the export rate of the different constructs, although the

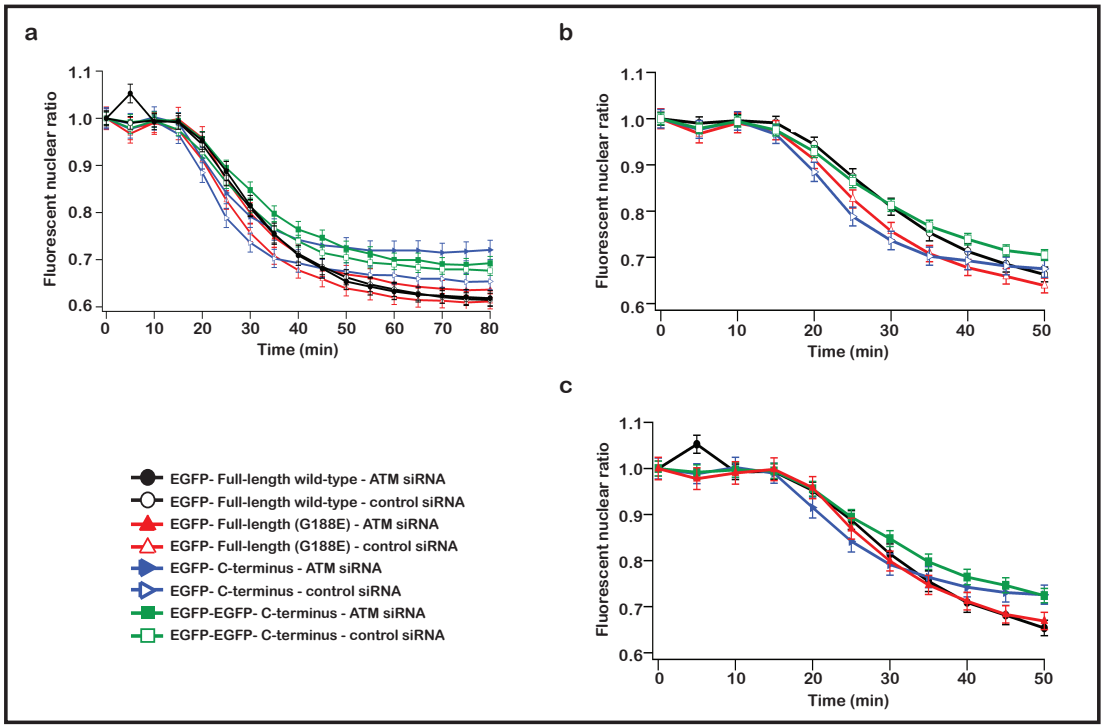


Figure 3.28: The partial depletion of ATM mostly affect the dynamics of the MacroD2 C-terminus only - Quantification of microirradiation live-cell imaging of four U2OS stable cell lines (EGFP-MacroD2 full-length, EGFP-MacroD2 full-length (G188E), EGFP-MacroD2 C-terminus and EGFP-EGFP-MacroD2 C-terminus) treated with ATM siRNA or control siRNA for 72 hours. Before imaging, cells were treated with Hoechst. Upon imaging, DNA damage was induced by using Etoposide 10 μM. 300−600 cells were quantified per each condition **(see Methods 6.8 and 6.8.6)**. *a*) The eight conditions are shown for the total length of the experiment (80 minutes). *b*) Close-up of *a*, with the four cell lines treated with control siRNA. *c*) Close-up of *a*, with the four cell lines treated with ATM siRNA. Both *a* and *b* show the first portion of the experiment, up to 50 minutes. Nuclear/cytoplasmic ratio was calculated using CellProfiler 2.0 (331), and the Hoechst signal was used for the nuclei segmentation. Error bars, 95 %CI. For the microirradiation, high laser energy was used.

two controls (full-length and C-terminus) were behaving as expected.

When comparing the control treatment with the ATM knock-down, the results show that the construct showing the most relevant decrease in export rate was the EGFP-MacroD2 C-terminus, while EGFP-MacroD2 full-length, EGFP-MacroD2 full-length (G188E) and the EGFP-EGFP-MacroD2 C-terminus constructs were less affected **(see Figure 3.28, *panels a and c*)**. The result further supports the hypothesis that the length of the construct is the main factor influencing the export dynamics of MacroD2.

3.6 ATM induces the phosphorylation of MacroD2 C-terminal region upon DNA damage

Upon DNA damage, ATM kinase modifies mediators and effectors of the repair process, creating a signaling cascade of phosphorylation events (276, 338). In particular, ATM activates other kinases, like Chk1 and Chk2, whose main role is to stop cell-cycle progression until the DNA lesion is resolved (252, 339, 340).

Therefore, I tested if MacroD2 is phosphorylated upon DNA damage as well. In HEK293 lysate I added the purified EGFP-tagged MacroD2 C-terminal fragment together with (^{32}P)γ-ATP **(see Methods 6.14)**. In control samples I also added ATM inhibitor KU55933. I then induced DNA damage in the samples by adding Benzonase nuclease and incubated for 40 minutes. I then purified EGFP-MacroD2 C-terminus protein by GFP-trap and ran the samples on a gel, which was stained with Coomassie and then analyzed for the radioactive content.

In the autoradiograph, I could see increased phosphorylation of the C-terminal fragment in the sample incubated with Benzonase, as DNA damage inducer **(Figure 3.29)**. On the contrary, when the reaction buffer was supplemented with the ATM inhibitor KU55933, the radioactive signal was reduced. This means that upon DNA damage MacroD2 is phosphorylated either directly or indirectly by ATM kinase activity.

To look at the DNA damage-induced phosphorylation of MacroD2 in more detail, we started a collaboration with Dr. A. Schmidt at the Zentrallabor für Proteinanalytik, LMU. By checking the sequence annotated on Uniprot, the MacroD2 protein sequence presents several residues that might accept phosphorylation. Thus, we adopted a mass spectrometry-based phospho-proteomics approach to measure the abundance of phosphorylated peptides from MacroD2 **(see Methods 6.15)**. U2OS cells stably expressing

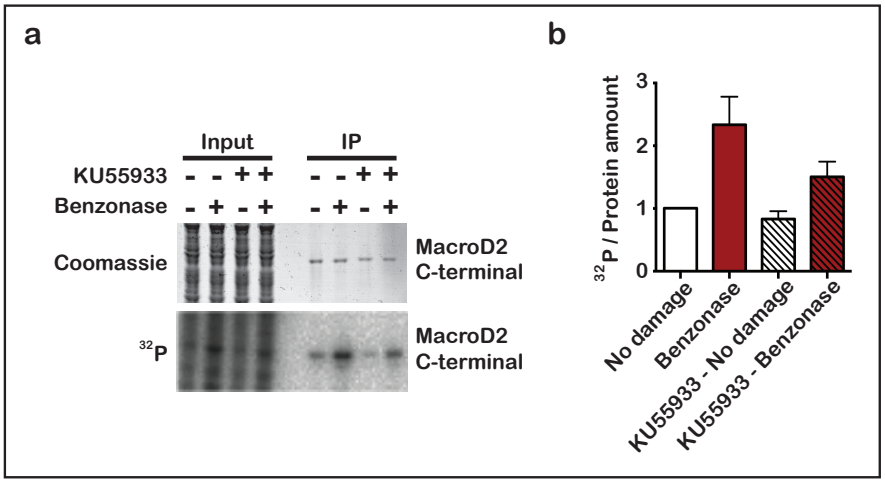


Figure 3.29: MacroD2 is phosphorylated in the C-terminal region - a)Coomassie and autoradiography of purified mEGFP-MacroD2 C-terminal domain added to HEK293 cell lysate, which has been treated or not with benzonase and KU55933, spiked with (^{32}P)γ-ATP and successively immunopurified. b) Quantification of IP results from three independent experiments **(see Methods 6.14)**. Error bars, SEM.

mYFP-MacroD2 full-length and mCherry-H2B were treated for 1 hour with DMSO or etoposide. MacroD2 proteins were then purified by using GFP-trap and treated for mass spectrometry by Dr. Schmidt. In three biological replicates we achieved a protein coverage of 70 %. Some sites showed increased phosphorylation upon etoposide treatment as compared with control, such as S292, S415 and S426 **(Figure 3.30, Table 3.1)**. Additionally, there were cases, like the phosphorylation in the aa 307-317 stretch, that is present only in the DMSO-treated cells. Other cases showed the same enrichment level in both treatments, arguing for the irrelevance of these particular modifications in MacroD2 nuclear export.

In particular, the phosphorylated version of the peptide 402-417 is more enriched upon etoposide treatment **(Figure 3.31,*panel a*)**. In fact, in the spectrum of the phosphorylated peptide **(see *panel a,1*)**, the peak is present upon etoposide treatment but absent in the untreated sample. In addition, since the peptide 402-417 is short, it is certain that the phosphorylation event occurs exclusively on S415, since the modification is present only for the y-fragments upon the peptide fragmentation **(Figure 3.31,*panel b*)**.

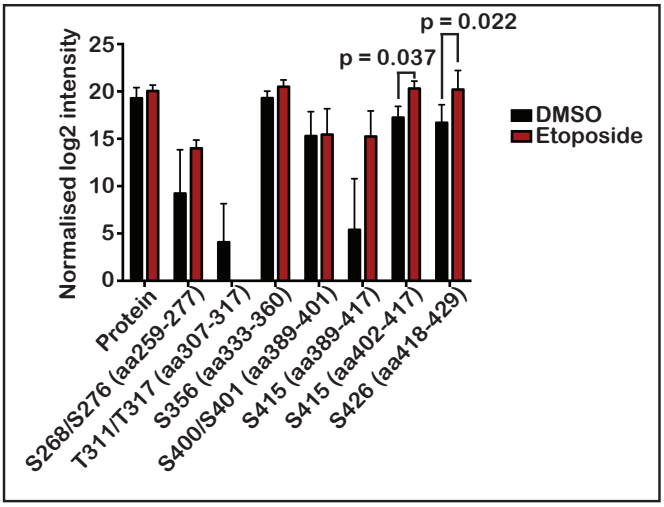


Figure 3.30: MacroD2 residues are phosphorylated upon etoposide treatment - Abundance of phosphorylated peptides of MacroD2. U2OS stably expressing mYFP-MacroD2+mCherry-H2B were treated for 1 hour with etoposide 10 μM or DMSO and immunopurified with GFP-trap. Following treatment, mass spectrometry experiment and analysis performed by Dr. A. Schmidt. Quantification of three independent experiments. Error bars, SEM. Statistics performed with paired t-Test with two-tailed distribution.

Sequence	Localization	Amino acid	Localization score
DENGPEEKQSVEEMEEQSQ	268	S	1.00
DGVNTVTVPGPASEEAVE	292	S	1.00
DENITKGGEVT	311	S	0.99
DSTKNEIKIETESQSSYMETEELSSNQE	345	S	0.65
DSTKNEIKIETESQSSYMETEELSSNQE	356	S	0.47
DSTKNEIKIETESQSSYMETEELSSNQE	352	T	0.51
DSTKNEIKIETESQSSYMETEELSSNQE	349	Y	0.59
DTPRMPGKSEGSSDLENTPGPDVEMNSQV	400	S	0.90
DTPRMPGKSEGSSDLENTPGPDVEMNSQV	401	S	0.49
DLENTPGPDVEMNSQV	415	S	1.00
DKVNDPTESQQEDQLIAGAQ	424	T	0.74
DKVNDPTESQQE	426	S	1.00

Table 3.1: Identified MacroD2 phosphorylated peptides - Overview of identified peptides and position of the phosphorylation. The localization score is the probability that the particular modification is located on the specific residue and it is calculated by dividing the (1 − pvalue) of the fragmented ion with the specific modified residues with the sum of the (1 − pvalue) of all the fragmented ions of the peptide (341).

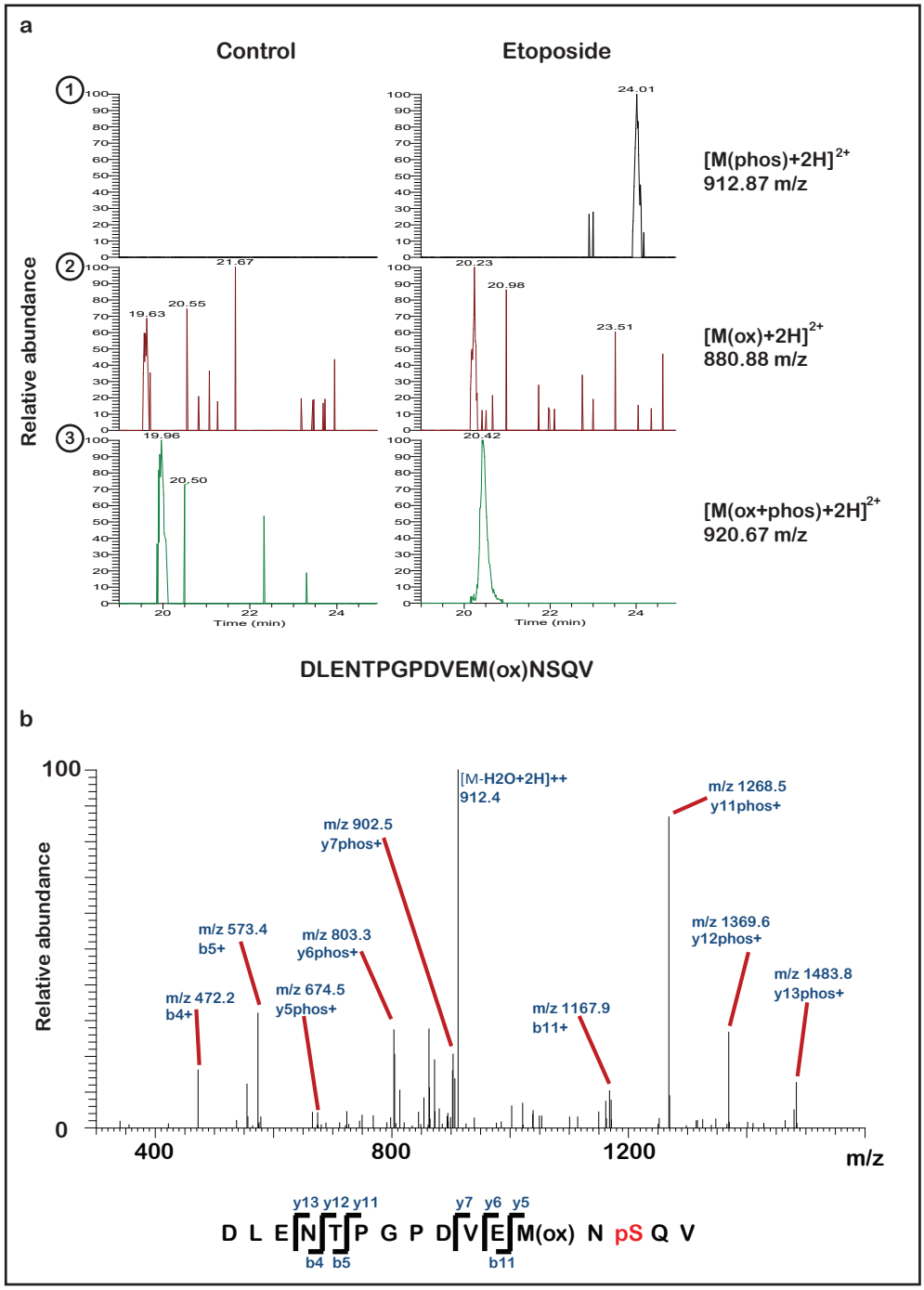


Figure 3.31: The serine in position 415 is modified upon etoposide treatment - a) Extracted ion chromatogram from experiment 3.30, showing aa 402-417 peptide with phosphorylation of the serine (1), oxidation of the methionine (2) or with both modifications (3), comparing control treatment with etoposide treatment. b) CID-MS/MS spectrum of the 402-417 peptide in the phosphorylated and oxidated form (921.37 m/z).

3.7 Phosphorylation sequence requirements indicates the direct involvement of ATM

MacroD2 is rich in residues that can accept phosphorylation, as shown for the many phosphorylation events on serines and threonines occurring on MacroD2 **(see Figure 3.30)**. Importantly, most of these amino acids are located in the C-terminal region of MacroD2, which has a biased composition towards acidic residues. This biased composition is favorable for a few classes of kinases, such as PI3K-like kinases, Chk kinases and Polo-like kinases, although these different classes have specific and more detailed consensus sequences (342, 343). In the MacroD2 C-terminal region there are four serines followed by glutamine (S276, S345, S415 and S426), a motif that is recognized by ATM, ATR and DNA-PK (220). The S415 was a well characterized site phosphorylated upon etoposide treatment, as detected by mass spectrometry **(see Figure 3.31)**, leading to the hypothesis that MacroD2 could be specifically phosphorylated by ATM and PI3K-like kinases.

To test whether these four SQ motifs are involved in the induction of MacroD2 nuclear export, I generated a EGFP-MacroD2 construct with all four C-terminal region SQ motifs mutated (mutant 4SA; S276,345,415,426A) with the help of G. Jankevicius and S. Grau. The mutant 4SA was transfected in HeLa cells stably expressing mCherry-H2B and tested for the export upon UV-laser microirradiation, in comparison to the wild-type **(see Methods 6.8)**. The EGFP-MacroD2 mutated in all SQ motifs was not able to export **(Figure 3.32)**. These motifs are thus necessary for MacroD2 nuclear export.

To address the specific contribution of each motif, we created four mutants, each with three of the four serines mutated to alanine, leaving only one SQ motif intact per construct.The constructs were then tested for the nuclear export upon UV-laser microirradiation, having been transfected in HeLa cells stably expressing mCherry-H2B **(see Methods 6.8)**. The experiment showed that the serines S276 and S426 are irrelevant for the export, because the constructs presenting only these motifs intact did not accumulate in the cytoplasm **(Figure 3.33)**. On the other hand, S345 and S415 are necessary for the nuclear export. When each of these two serines was left intact, MacroD2 nuclear export was successful.

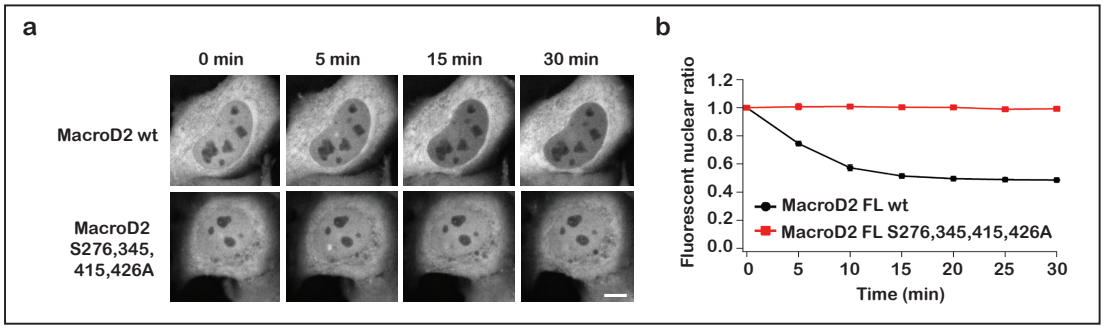


Figure 3.32: Mutation of the four SQ motifs completely abolishes MacroD2 nuclear export - a)Microirradiation live-cell microscopy of HeLa cells stably expressing mCherry-H2B and transfected with mEGFP-MacroD2 full-length or mEGFP-MacroD2 S276,345,415,426A constructs. For the microirradiation, high laser energy was used. Scale bar, 10 μm. b) 50−100 cells were quantified from three independent experiments **(see Methods 6.8)**. Nuclear/cytoplasmic ratio was calculated using CellProfiler 2.0 (331), and the mCherry signal was used for the nuclei segmentation. Error bars, 95 %CI.

Then, to further confirm the role of the S345 and S415 for the nuclear export of MacroD2, I generated a EGFP-MacroD2 full-length construct whose S276 and S426 were mutated to alanine, leaving S345 and S415 as serine. As before, I transfected the mutant and the wild-type constructs in HeLa cells stably expressing mCherry-H2B and tested their ability to export from the nucleus upon UV-laser microirradiation **(see Methods 6.8)**. The export dynamics of the MacroD2 S276,426A construct is indistinguishable from the wild-type **(Figure 3.34)**. Additionally, from Figure 3.33 it is clearly noticeable that when the S354 and S415 are intact alone in the protein, the nuclear export does occur but not at the same level of the wild-type protein. Thus, S345 and S415 are likely working in a cooperative way.

The fact that I could pinpoint two specific residues essential for MacroD2 nuclear export argues for the presence of two exporting sequences in the C-terminus region. It is likely that each sequence includes a phosphorylated serine. The serine S415 is clearly shown to be phosphorylated upon DNA damage **(see Figure 3.31)**. Also S345 was recognized as site during the identification of the phosphorylated sites with mass spectrometry, although not with the same degree of certainty **(see Table 3.1)**. Indeed, due to the length of the peptide and, thus, the presence of many residues capable of modification, S345 was not shown to be the only site able to accept the phosphorylation within the peptide. Nonetheless, in both cases the mutation of the serine blocks the

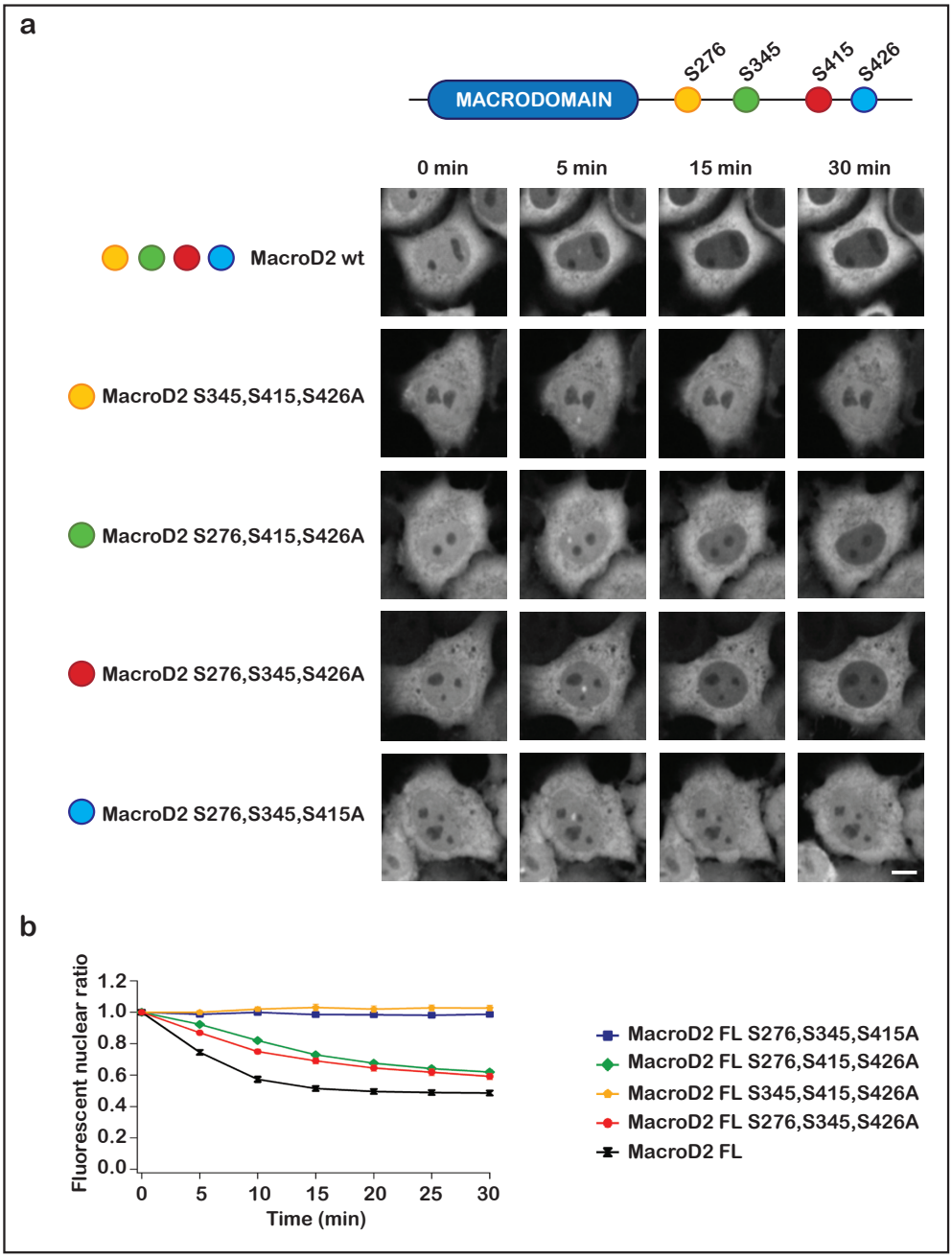


Figure 3.33: S345 and S415 are neccessary for MacroD2 nuclear export - a)Microirradiation live-cell microscopy of HeLa cells stably expressing mCherry-H2B and transfected with mEGFP-MacroD2 full-length or different serine triple mutants (S345,415,426A; S276,415,426A; S276,345,426A; S276,345,415A). For the microirradiation, high laser energy was used. Scale bar, 10 μm. b) 50−100 cells were quantified from three independent experiments **(see Methods 6.8)**. Nuclear/cytoplasmic ratio was calculated using CellProfiler 2.0 (331), and the mCherry signal was used for the nuclei segmentation. Error bars, 95 %CI.

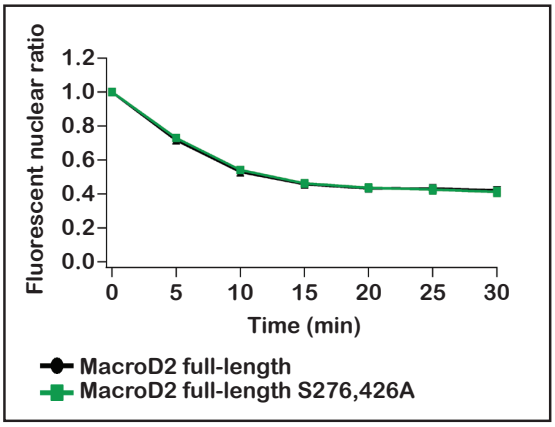


Figure 3.34: S345 and S415 are sufficient for the nuclear export of MacroD2 - Quantification of microirradiation live-cell microscopy of HeLa cells stably expressing mCherry-H2B and transfected with mEGFP-MacroD2 full-length wild type or S276,426A. 50−100 cells were quantified from three independent experiments **(see Methods 6.8)**. Nuclear/cytoplasmic ratio was calculated using CellProfiler 2.0 (331), and the mCherry signal was used for the nuclei segmentation. Error bars, 95 %CI. For the microirradiation, high laser energy was used.

nuclear export. Thus, it is highly likely that the S345 undergoes the phosphorylation event able to induce the nuclear export, as the serine S415 does. However, more sequence requirements could be necessary for the activity of these export sequences. For this reason, I wanted to characterize the contribution of the residues surrounding the phospho-serine for the export. This piece of information could be in fact very helpful to define possible interactors that are important for the accomplishment of the localization.

I initially focused on the sequence surrounding the serine 415 because it was the best characterized modification in the phospho-proteomics analysis **(see Figure 3.31)**. First, I tested a series of fragments with decreasing length to find a sequence short enough for further manipulation, yet, still be able to export. I therefore compared EGFP-MacroD2 full-length, MacroD2 C-terminal (aa 236-448) and the fragments aa 382-418, aa 402-418 and aa 410-418, all transiently expressed in HeLa-mCherry-H2B cells and measured the nuclear export upon UV-laser microirradiation **(see Methods 6.8)**. Surprisingly, just a 9-residue-long fragment, namely aa 410-418 (*DVEMNSQVD*), is still able to export residually **(Figure 3.35)**. Therefore, I defined this fragment as the core exporting sequence.

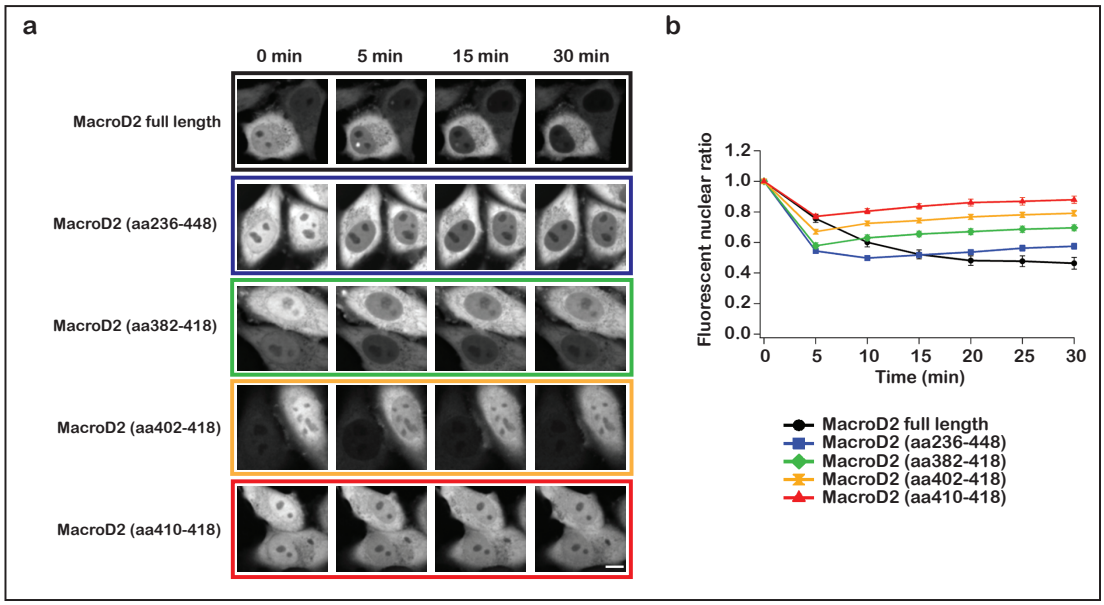


Figure 3.35: The fragment aa 410-418 is sufficient for MacroD2 nuclear export - a)Microirradiation live-cell imaging of HeLa cells stably expressing mCherry-H2B and transfected with mEGFP-tagged constructs (MacroD2 full-length, MacroD2 C-terminal region (aa 236-448), MacroD2 aa 382-418, MacroD2 aa 402-418, MacroD2 aa 410-418). For the microirradiation, high laser energy was used. Scale bar, 10 μm. b) 50−100 cells were quantified from three independent experiments **(see Methods 6.8)**. Nuclear/cytoplasmic ratio was calculated using CellProfiler 2.0 (331), and the mCherry signal was used for the nuclei segmentation. Error bars, 95 %CI.

To define to what extent each residue in the sequence is required for phosphorylation, I adopted a scanning mutagenesis approach. I mutated each residue of the 410-416 stretch to alanine (or aspartate in the case of V411, given to the close similarity between alanine and valine) within the EGFP-MacroD2 (382-418) fragment. I performed the mutations in a longer fragment to avoid impairments in the binding of the kinase due to sequence length. To assay the kinase activity, I used HEK293 cell lysates spiked with the purified proteins and treated with Benzonase nuclease **(see Methods 6.14)**. I tested the incorporation of (^{32}P)-phosphate in the His-EGFP-tagged mutagenized fragments compared to the GST-EGFP-tagged wild-type fragment, which acted as an internal positive control for phosphorylation **(Figure 3.36)**.

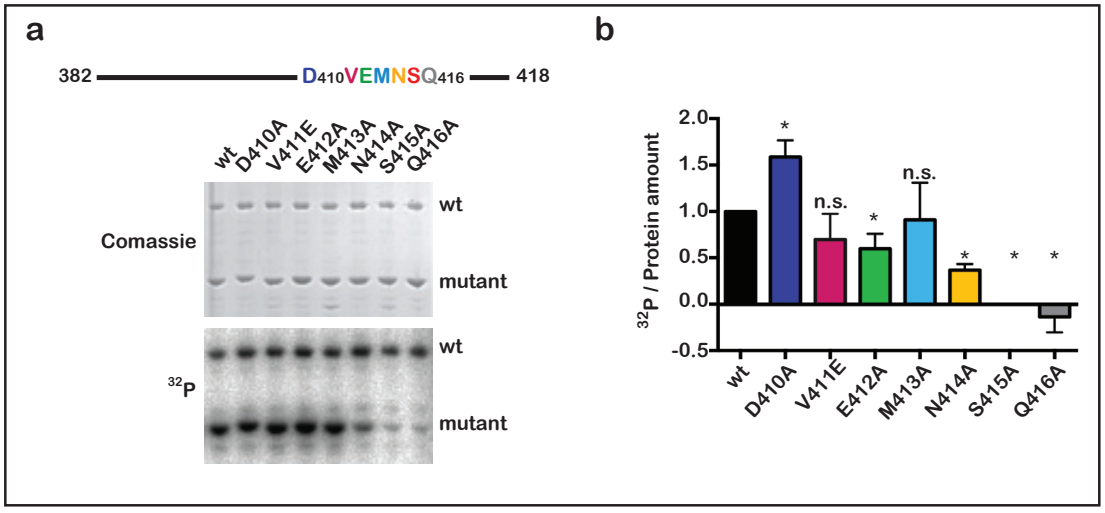


Figure 3.36: Both S415 and Q416 are required for the phosphorylation - a)Coomassie and autoradiography of purified GST-mEGFP-MacroD2 (aa382-418) wild-type and His-mEGFP-MacroD2 (aa382-418) (wild-type or mutant proteins). Purified constructs were added to HEK293 cell lysate, which was treated with benzonase and spiked with (^{32}P)γ-ATP. Successively, immunopurification with GFP-trap. b) Quantification of (a) in three independent experiments. Error bars, SEM. Statistics performed with unpaired non-parametrical Mann-Whitney test against the wild-type **(see Methods 6.14)**. * is p = 0.05; n.s. is p > 0.05.

Different residues appear to have different effects on the phosphorylation, presumably due to a variation in recognition by the kinase. As expected, the mutation of the serine 415 abolished the incorporation of radioactivity. Additionally, the Q416A mutation totally blocked the modification. This result argued for the direct involvement of

ATM in the phosphorylation of MacroD2, since ATM, ATR and DNA-PK are the only kinases that strictly require the glutamine after the serine: among them, ATM seems to be the main inducer of MacroD2 nuclear export **(see Section 3.4;** (342)**)**. On the other hand, other mutations only marginally affected the modification, thus they are not fundamental for the kinase to recognize the sequence.

After having defined the phosphorylation level of these mutant, I wanted to see their nuclear export efficiency. Thus, I transfected all the EGFP-tagged mutants of the aa 382-418 fragment in HeLa cells stably expressing mCherry-H2B and I tested them in nuclear export assays upon UV-laser microirradiation **(see Methods 6.8)**. As expected, neither S415A nor Q416A are able to export **(Figure 3.37)**. Interestingly, the M413A mutant shared the same phenotype.

Since M413A mutant shows no change in incorporation of the (^{32}P)-phosphate compared to the wild type, this result suggests that the methionine could be essential in the nuclear export in a step that is downstream of the phosphorylation of the serine, such as for the recognition of the sequence by the export machinery.

A similar approach was taken to define the sequence requirements also for the S345 site. Together with J. Preißer and M.L. Tran, we determined the shortest sequence required for the export of a fragment containing the serine 345. We found that the stretch between K340 and M350 is necessary and sufficient for the export (*KIETESQSSYM*): in comparison with the fragment 410-418 around the serine 415, the 340-350 stretch is slightly longer **(data not shown)**. Still, to again avoid impairments in the binding, we performed the scanning alanine mutagenesis in a longer fragment, namely aa 336-372: the two fragments, aa 382-418 and aa 336-372, were designed to have a comparable length.

M.L. Tran expressed all the mutants in HeLa cells stably expressing mCherry-H2B and tested their nuclear export dynamics upon laser microirradiation. Similarly to the 410-418 fragment, S345, Q346 and M350 are necessary for the nuclear export of this fragment **(Figure 3.38)**. On the other hand, few other mutations also slowed the export dynamics **(Table 3.2)**. Their residues were mainly located downstream from the SQ motif.

In both sequences, the mutation of the serine or the following glutamine affected the nuclear export. Likewise, in both sequences the mutation of a methionine into an alanine impaired the export, although their position relative to the serine is different:

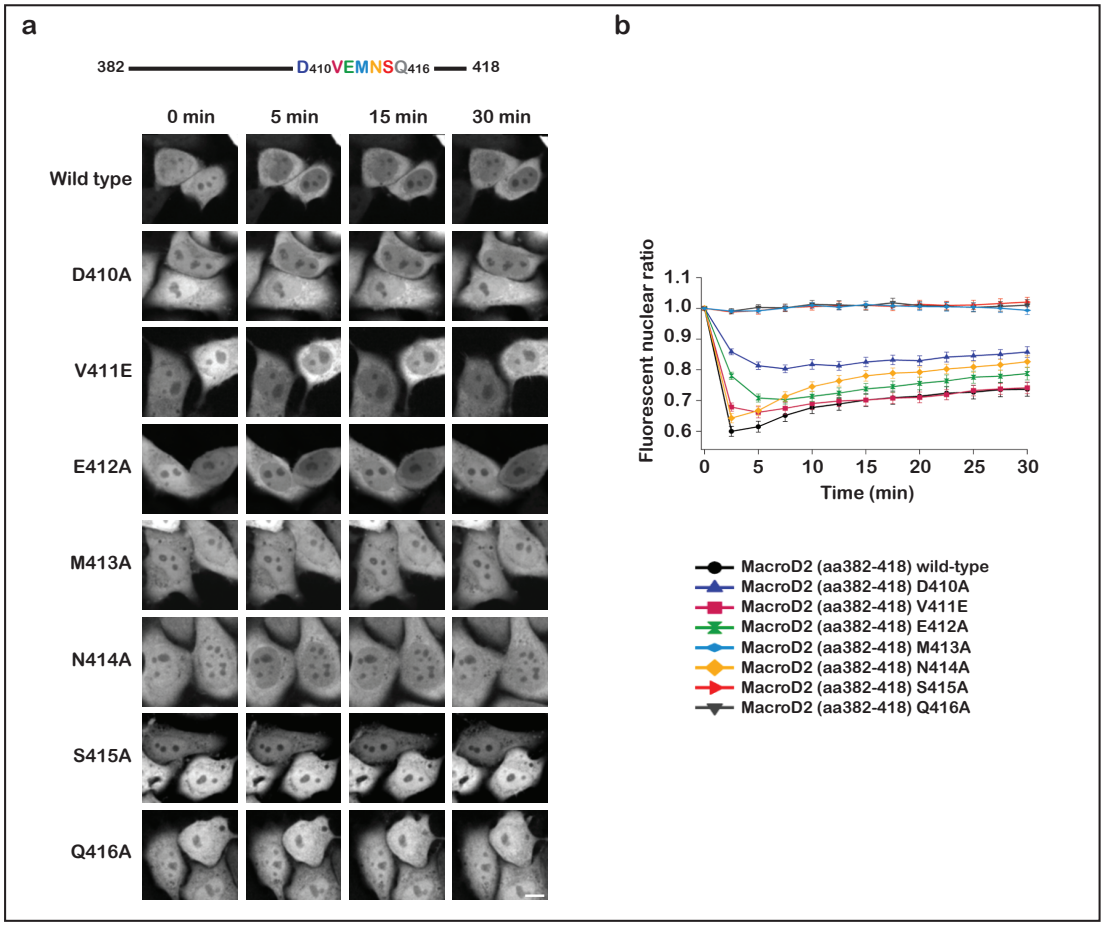


Figure 3.37: M413 is also necessary for the nuclear export of the fragment aa 382-418 - a)Microirradiation live-cell microscopy of HeLa cells stably expressing mCherry-H2B and transfected with mEGFP-MacroD2 (aa382-418) wild-type construct or mutants (D410A; V411E; E412A; M413A; N414A; S415A; Q416A). For the microirradiation, high laser energy was used. Scale bar, 10 μm. b) 50−100 cells were quantified from three independent experiments **(see Methods 6.8)**. Nuclear/cytoplasmic ratio was calculated using CellProfiler 2.0 (331), and the mCherry signal was used for the nuclei segmentation. Error bars, 95 %CI.

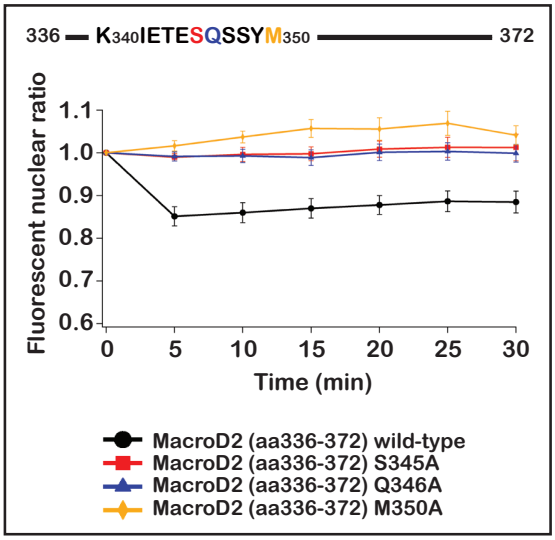


Figure 3.38: In the aa 336-372 fragment, S345, Q346 and M350 are necessary for the export - Quantification of microirradiation live-cell imaging of HeLa cells stably expressing mCherry-H2B and transfected with mEGFP-MacroD2 (aa336-372) wild-type construct or mutants S345A, Q346A and M350A. 50−100 cells were quantified from three independent experiments **(see Methods 6.8)**. Nuclear/cytoplasmic ratio was calculated using CellProfiler 2.0 (331), and the mCherry signal was used for the nuclei segmentation. Error bars, 95 %CI. For the microirradiation, high laser energy was used.

Position	Wild-type	Mutant	Export in mutant	Significance
340	K	A	unchanged	no
341	I	A	unchanged	no
342	E	A	unchanged	no
343	T	A	slow	yes
344	E	A	unchanged	no
345	S	A	no	yes
346	Q	A	no	yes
347	S	A	slow	yes
348	S	A	slow	yes
349	Y	A	slow	yes
350	M	A	no	yes

Table 3.2: Export dynamics for alanine mutant in the aa 336-372 fragment - Alanine scanning mutagenesis for the fragment aa 336-372. Each mutant is compared to the wild-type and the outcome of unpaired non-parametrical Mann-Whitney tests is shown.

M413 is in position -2 compared to the S415, while M350 is in position +5 from the serine S345. In addition, when comparing the two export sequences, all the residues of the aa 382-418 fragment that are relevant for the definition of the export dynamics were located shortly upstream from the SQ motif. Conversely, in the aa 336-372 fragment, the majority of the residues whose mutation affected the nuclear export were located downstream from the SQ motif.

The inhibition of the MacroD2 nuclear export when the methionine is mutated into alanine shows that by placing a residue of different nature in that position does impair the occurrence of the process, but does not prove that the methionine and only the methionine is required in that position. To test how essential is the methionine, I mutated the M413 to phenylalanine or isoleucine. These two residues are, in fact, belonging to the same group of strongly hydrophobic amino acids.

For the experiment, I transfected the mEGFP-MacroD2 (aa388-424) wild-type and mutants M413A, M413F and M413I constructs in HeLa cells stably expressing mCherry-H2B. I then tested the export dynamics of these construct upon UV-laser microirradiation **(see Methods 6.8)**. When comparing the wild-type construct with its related mutants, it is clear that the mutation of the methionine impairs the nuclear export, even when it is mutated to strongly hydrophobic residues **(Figure 3.39)**.

In summary, the phosphorylation of the S415 residue within the aa 382-418 fragment is dependent on the presence of the serine and the glutamine. Moreover, for both the exporting sequences, the mutation of the serine or the glutamine to alanine blocks the nuclear export. Having the serine followed by a glutamine in both motifs as sequence requirement, strongly suggests that ATM itself is interacting with MacroD2 and modifying it upon DNA damage.

In addition to the SQ motif, the two sequences share another feature: a methionine, whose mutation to alanine impairs the nuclear export of MacroD2 fragments. The mutation of the methionine does not affect the phosphorylation process, as I showed for the aa 382-418 sequence. However, the relative positions of the methionines and the SQ motifs change in the two exporting sequences.

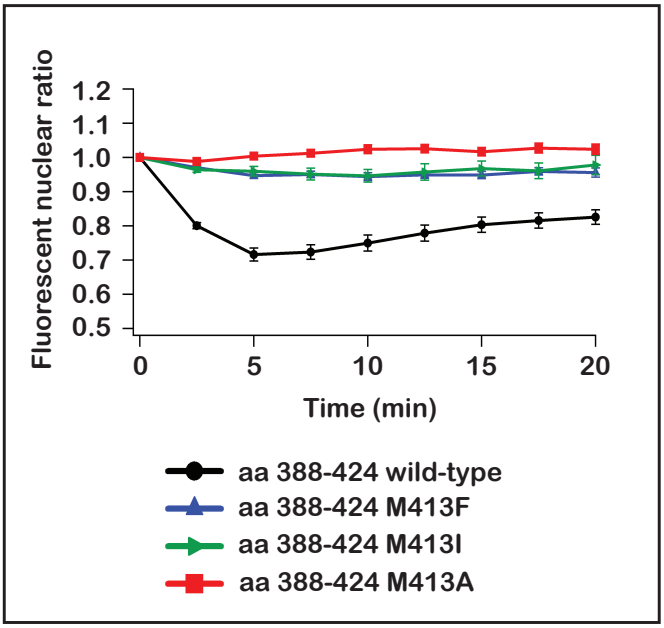


Figure 3.39: Mutation of the M413 to other hydrophobic residues does not rescue MacroD2 nuclear export - Quantification of microirradiation live-cell imaging of HeLa cells stably expressing mCherry-H2B and transfected with mEGFP-MacroD2 (aa388-424) wild-type construct or mutants M413A, M413F and M413I. 50−100 cells were quantified from two independent experiments **(see Methods 6.8)**. Nuclear/cytoplasmic ratio was calculated using CellProfiler 2.0 (331), and the mCherry signal was used for the nuclei segmentation. Error bars, 95 %CI. For the microirradiation, high laser energy was used.

3.8 14-3-3s do not bind MacroD2

As the two sequences in MacroD2 protein necessary for its nuclear export upon DNA damage have been defined **(see Section 3.7)**, it is possible to address which mechanism mediates the nuclear export of MacroD2. As shown above, this process is regulated by means of phosphorylation events, which can modulate the interaction with nuclear transport receptors or adaptor proteins. As described in Section 1.4.2, it is a common mechanism that proteins are phosphorylated and such phospho-sites can be recognized by adaptor proteins, which mediate in turn the formation of the complex with the export machinery (302, 344, 345).

The best described class of these adaptor proteins are the 14-3-3s (346, 347, 348). Preliminary analyses of the first biological replicate of the co-immunopurification experiment coupled to peptide mass fingerprinting, which I describe in Chapter 4, showed the binding of 14-3-3ϵ to MacroD2 full-length and C-terminal fragment upon etoposide treatment. This preliminary result prompted me to perform validation experiments on this possible binding candidate for the MacroD2 nuclear export.

14-3-3 proteins are small (28 kDa), abundant and very stable and preferentially bind phosphorylated serines or threonines that are placed within a consensus sequence (349, 350). There are three described consensus sequences:

- -R(S/T)XpSXP-
- -RX(F/Y)XpSXP-
- -pSXCOOH

Several proteins have been described to interact with the 14-3-3s in a phosphorylation-independent manner, such as the anti-apoptotic protein BAX (351), the transcription factor ChREBP (352) and the viral ADP-ribosyltransferase ExoS (353).

The 14-3-3s might be suitable candidates to drive MacroD2 out of the nucleus for a number of reasons. First of all, they bind phospho-serines or phospho-threonine. In the case of COP1, the phospho-site is modified by ATM itself (324). 14-3-3s have been described several times as intermediate of nuclear export, such as for COP1, Chk1, FKHRL1, FOXO1, FOXO4, BAD or Cdc25C (324, 325, 346, 354, 355, 356). However, in most cases the nuclear export is due to the fact that, upon binding, the 14-3-3 masks

the NLS of the transported protein, thus impairing the nuclear import. Often, the binding with 14-3-3s require two phospho-sites, since the 14-3-3 proteins adopt either a homo- or a heterodimer form (350). The two residues that act as binding site for the 14-3-3s are preferentially located in unstructured sequences, to help in formation of all possible contacts between the proteins (356). For these reasons, it is feasible to hypothesize a possible interaction between MacroD2 and a 14-3-3 protein upon DNA damage. As shown in the previous section, MacroD2 has two motifs undergoing phosphorylation upon DNA damage, which drive its nuclear export. Both sites are present in the unstructured C-terminus region of MacroD2.

On the other hand, some other considerations question the involvement of 14-3-3s in the export of MacroD2. The consensus sequence for the 14-3-3 binding tends to be incompatible with the one of the PI3K-like kinases, since the 14-3-3s prefers positive charges, while the PI3K-like kinases prefers sequences rich in negative charges (357). Therefore, the consensus is theoretically incompatible also with MacroD2 exporting sequences, being that MacroD2 is rich in acidic amino acids. Still, given the many elements supporting the hypothesis, it was worthwhile to test the interaction.

For this reason, with the help of J. Preißer, I expressed and purified all seven human 14-3-3 isoforms (ϵ, σ, ζ, β, η, θ and γ) with a GST tag **(see Methods 6.16.1)** Additionally, I also purified the 14-3-3ϵ with a mutation of lysine 49 into glutamate. This mutant is described as unable to bind the phospho-serine, and as such represents a suitable negative control for a pull-down assay. I used synthetic biotynilated-peptides in place of MacroD2 protein in order to test the interaction in an *in vitro* condition:

1. Biotin - G S L S Q R Q R S T S T P N V H (Control peptide - unphosphorylated)
2. Biotin - G S L S Q R Q R S T pS T P N V H (Control peptide - phosphorylated)
3. Biotin - N T P G P D V E M N S Q V D K V
4. Biotin - N T P G P D V E M N pS Q V D K V
5. Biotin - N T P G P D V E A N S Q V D K V
6. Biotin - N T P G P D V E A N pS Q V D K V

1 and *2* are control peptides that have been previously published (343). The other four peptides are designed from MacroD2 protein sequence (aa 402-420). Peptides *4* and *6* are phosphorylated at serine 415; *5* and *6* have the methionine 413 mutated to alanine. Considering that the percentage of phosphorylated MacroD2 over the total cellular pool is unknown, by adopting synthetic peptides I sought to work in a defined condition, where all the MacroD2 sequences would be modified.

I then performed the pull-down experiment of the GST-14-3-3 proteins using the biotinilated peptides conjugated to Streptavidin resin **(see Methods 6.16.2)**. The Figure 3.40 shows the pull-down experiment of the GST-14-3-3 proteins with the synthetic peptides.

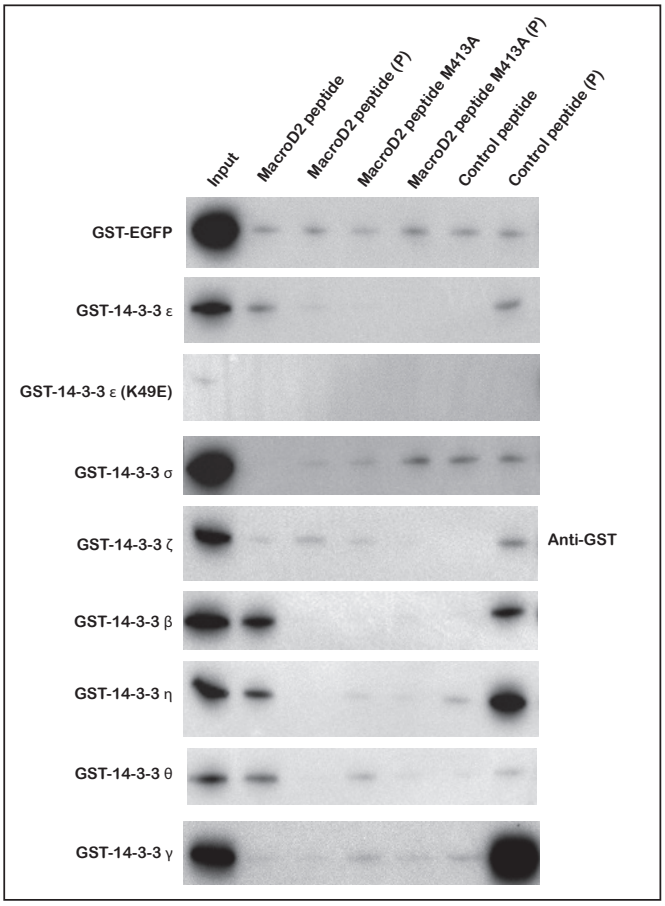


Figure 3.40: The MacroD2 phosphorylated peptides do not interact with purified 14-3-3s - Pull-down experiment of purified 14-3-3 (ϵ, ϵ(K49E), σ, ζ, β, η, θ and γ) and biotinylated-peptides, phosphorylated or not. Immunoblot of the samples with anti-GST **(see Methods 6.16.2)**.

In the negative control EGFP, the signal in most of the lanes suggested a background binding, while the pull-down of the 14-3-3 ϵ (K49E) mutant was not comparable because of its faint signal in the input. All the 14-3-3 proteins except the σ convincingly bound the phosphorylated control protein more than the unphosphorylated form. Among these pull-downs, the 14-3-3 ϵ, β, η and θ bound better the unphosphorylated MacroD2 peptide compared to the phosphorylated one, while only the 14-3-3 ζ seemed to bind more the phosphorylated peptide than the unphosphorylated form, even if very weakly.

When comparing the pull-downs performed with the peptides with the mutated methionine, the 14-3-3 σ was more enriched in the phosphorylated form compared to the unphosphorylated one. However, the equal binding for the positive control urged me to not consider this protein, since published work clearly showed that it should bind more the phosphorylated control peptide over the unposphorylated (343). This output was also present for the 14-3-3 β, even though it was a very faint signal. On the other hand, the 14-3-3 ζ, η, θ and γ showed more enrichment upon the unphosphorylated version. Finally, the 14-3-3 ϵ showed no signal for either pull-downs. According to these results, no 14-3-3 protein met these three criteria: to bind the control peptides as expected (*2* more than *1*), to bind the phosphorylated MacroD2 peptide (*4*) more than the unphosphorylated one (*3*), and to not bind any peptide mutated in the methionine (*5* and *6*).

I repeated the experiment with a subset of proteins, but the results were not reproducible, possibly due to different amounts of peptides conjugated to the beads across the different experiments **(data not shown)**. Although the control peptides always worked consistently, since I did not find a way to define the amount of MacroD2 peptides present until the end of the experiment, I decided to test the interaction with another approach: a co-immunopurification experiment between 14-3-3 proteins and the purified MacroD2 C-terminus fragment.

To this end, I added EGFP-MacroD2 C-terminal domain and a selection of the 14-3-3 proteins (ϵ, ϵ(K49E), ζ and β) to the HEK293 lysate and activated the DNA damage response by adding Benzonase to the system **(see Methods 6.16.3)**, as done previously **(see Figures 3.29 and 3.36)**. As mentioned above, the strength of the pull-down approach with synthetic peptides consisted of having at least 95 % of the molecules phosphorylated, while in this case I ignore the exact percentage of phosphorylated protein. However, since I used a large amount of purified proteins in the experiments,

it may have been possible to detect the interaction between the two partners in this condition. Nonetheless, the band for the EGFP-MacroD2 C-terminal region did not appear in any of the pull-downs from samples treated with Benzonase **(Figure 3.41)**.

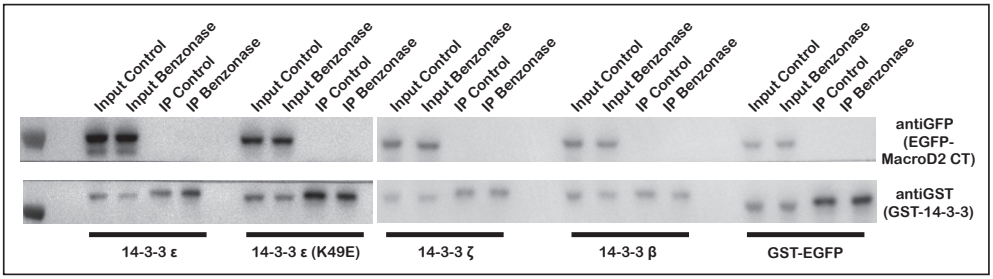


Figure 3.41: The MacroD2 C-terminal domain does not interact with 14-3-3ϵ, ζ and β *in vitro* - Immunoprecipitation experiment of EGFP-MacroD2 C-terminal domain and GST-14-3-3s (ϵ, ϵ(K49E), ζ and β) or GST-EGFP. Purified proteins were spiked into HEK293 cell-extract. DNA damage was induced by adding Benzonase nuclease. Control immunopurifications were performed with water in place of Benzonase. Co-IP performed with glutathione beads and interaction tested with anti-GFP.**(see Methods 6.16.3)**

Although I could not detect any signal, I decided to perform a variation of the same experiment, using the EGFP-MacroD2 (aa382-418) fragments previously used **(see Figure 3.36)**. In particular, I added the same above mentioned 14-3-3 proteins and the MacroD2 fragments, both the wild type and the mutants M413A and S415A to the HEK293 cell lysate **(see Methods 6.16.3)**. As before, I could not detect any enrichment of the MacroD2 fragments upon pull-down of the GST-14-3-3 proteins **(Figure 3.42)**.

As a last attempt to observe the interaction between MacroD2 and the 14-3-3s, I performed a co-immunoprecipitation experiment of the 14-3-3ϵ with EGFP-MacroD2 or EGFP alone stably expressed in HEK293 cells **(see Methods 6.16.4)**. The 14-3-3ϵ was the isoform identified in the first place in the co-immunopurification experiment coupled with peptide mass fingerprinting **(see Chapter 4)**. The cells were treated for one hour with etoposide before harvesting, reproducing the same conditions used for the co-IP experiment that is described in Chapter 4. However it was observed that even though MacroD2 was highly enriched, the 14-3-3ϵ was not detectable **(Figure 3.43)**.

To sum up, the multiple attempts at identifying an interaction were unsuccessful and, therefore, the 14-3-3 identified by the mass spectrometry approach appears to be a false positive.

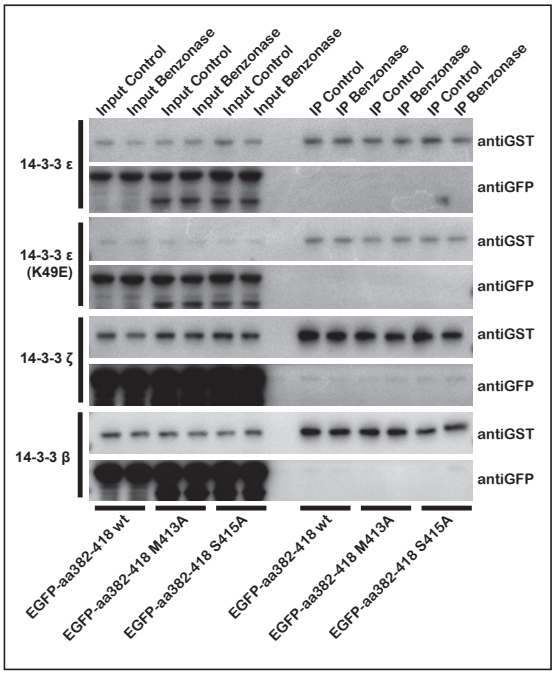


Figure 3.42: The 14-3-3s do not interact with MacroD2 fragments *in vitro* - Immunoprecipitation experiment of EGFP-MacroD2 fragment (aa382-418) wild type and mutants (M413A and S415A) and GST-14-3-3s (ϵ, ϵ(K49E), ζ and β). Purified proteins were spiked into HEK293 cell-extract. DNA damage was induced by adding Benzonase nuclease. Co-IP performed with glutathione beads and interaction tested with anti-GFP on immunoblot **(see Methods 6.16.3)**.

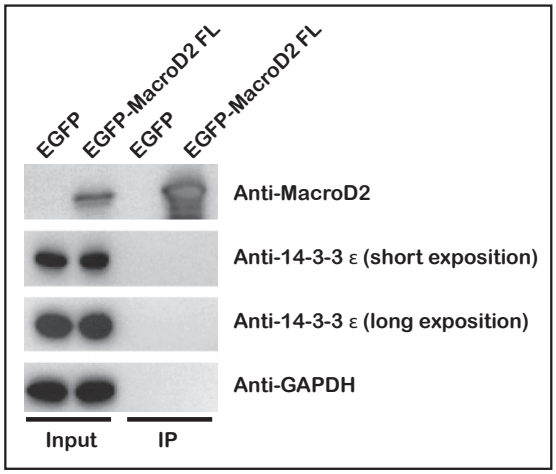


Figure 3.43: MacroD2 fails to pull down the 14-3-3ϵ *in vivo* - Immunoprecipitation experiment of 14-3-3ϵ protein from HEK293 cell extract. Cells were overexpressing EGFP-MacroD2 full-length or EGFP alone, treated for one hour with etoposide 10 μM and immuno-purified by GFP-trap. GAPDH was used as negative control **(see Methods 6.16.4)**.

3.9 ATM activity regulates the recruitment of MacroD2 to DNA lesions

After having characterized MacroD2 nuclear export, the question about the possible role of this regulation is still open. To this end, I tested if the ATM-induced MacroD2 nuclear export could affect MacroD2 recruitment to sites of DNA damage. Thus, I transfected EGFP-tagged MacroD2 full-length, MacroD2 macrodomain and MacroD2 C-terminus proteins in HeLa stably expressing mCherry-H2B and performed UV-laser microirradiation experiments to test their recruitment dynamics **(see Methods 6.8)**. The recruitment dynamics of the three protein was very different **(Figure 3.44)**. Both MacroD2 full-length and macrodomain recruited to DNA damage sites, while the EGFP-MacroD2 C-terminus fragment did not. Moreover, between the two recruiting constructs there was a striking difference in the residency time: the full-length protein dissipates from the sites of DNA damage faster than the macrodomain only construct.

To investigate if the MacroD2 nuclear export can affect the recruitment dynamics, I compared the accumulation of EGFP-tagged wild type MacroD2 with that of the export-incompetent (4SA - S276,345,415,426A) mutant at the UV microirradiated sites

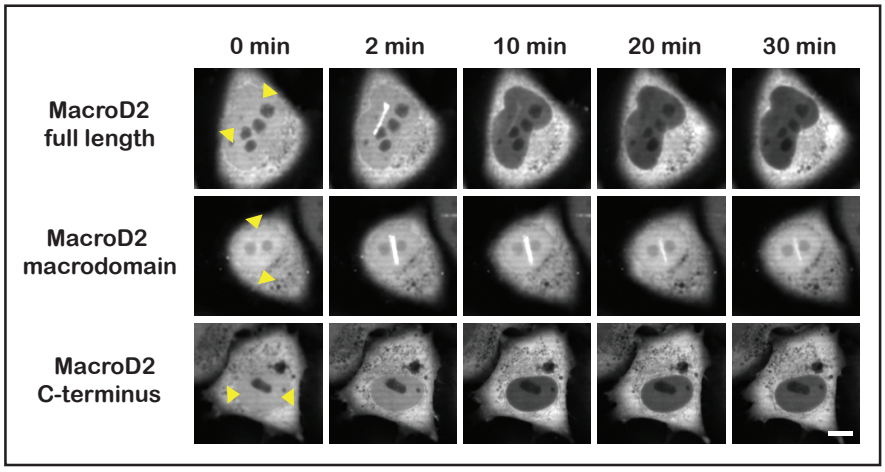


Figure 3.44: The macrodomain of MacroD2 is necessary for the recruitment of MacroD2 to DNA lesions - Recruitment of tagged mEGFP-MacroD2 constructs (full length, macrodomain or C-terminal domain) in HeLa cells stably expressing mCherry-H2B. The focus of laser micro-irradiation is indicated with yellow arrowheads. For the microirradiation, medium laser energy was used. Scale bar, 10 μm.

of DNA damage induced by UV-laser microirradiation. The two constructs were also compared in the recruitment in the presence of the ATM inhibitor, KU55933, or DMSO control **(see Methods 6.8.5)**. The recruitment of the wild type MacroD2 construct treated with DMSO peaked 2 minutes after laser irradiation **(Figure 3.45)**.

Afterwards, the protein amount at the sites of DNA damage - and within the whole nucleus - decreased, which led to the loss of MacroD2 proteins at the damage site within 5 minutes.

When the cells transfected with the wild-type construct are treated with the ATM inhibitor, MacroD2 visibly resided at the DNA lesions up to 30 minutes after the laser microirradiation, as also shown by the quantified levels of the fluorescent signal in the recruitment area that persist around 1 for the whole experiment. Likewise, the cells trasfected with the export-incompetent mutant show high recruitment and persistence of the fluorescent signal at DNA lesions until the end of the experiment (30 minutes), as shown by both the images and the quantification.

In summary, while the recruitment has no impact on the nuclear export dynamics of MacroD2 **(see Figure 3.25)**, these results show that MacroD2 nuclear export affects its recruitment dynamics.

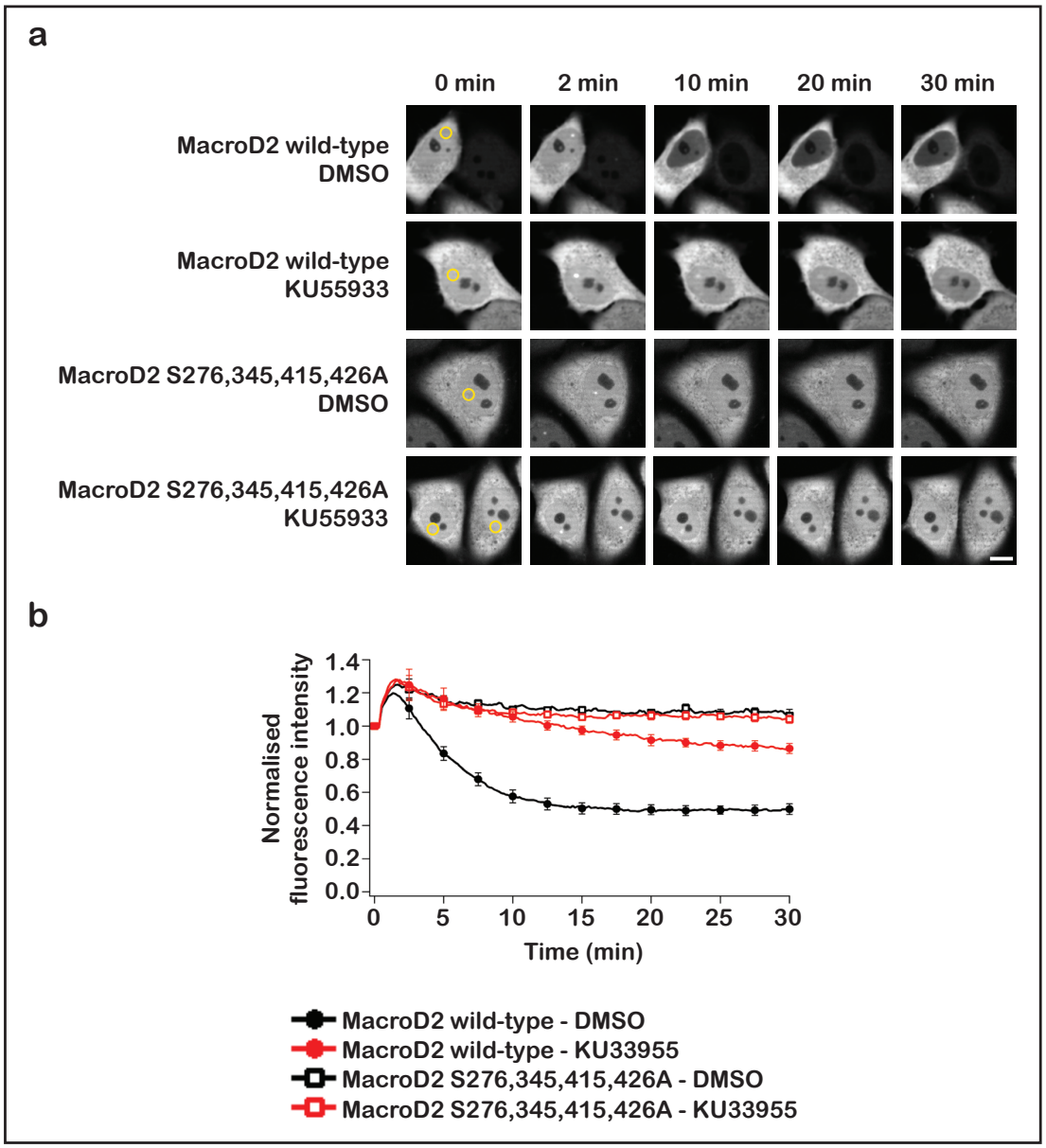


Figure 3.45: MacroD2 nuclear export affects its recruitment to DNA lesions - a) Recruitment of tagged mEGFP-MacroD2 constructs (wild-type or S276,345,415,426A mutant) in HeLa cells stably expressing mCherry-H2B, treated either with DMSO or KU55933 10 μM. The focus of laser micro-irradiation is indicated with a yellow circle. For the micoirradiation, medium laser energy was used. Scale bar, 10 μm. b) Quantification of the recruitment performed on 15 cells from three independent experiments **(see Methods 6.8.5)**. Error bars are 95 %CI.

3.10 Discussion

This chapter shows that MacroD2 nuclear signal depletion upon DNA damage is explained by its regulated nuclear export **(Figure 3.46)**. The phenomenon was also confirmed for the endogenous protein, suggesting a physiological relevance of this form of regulation.

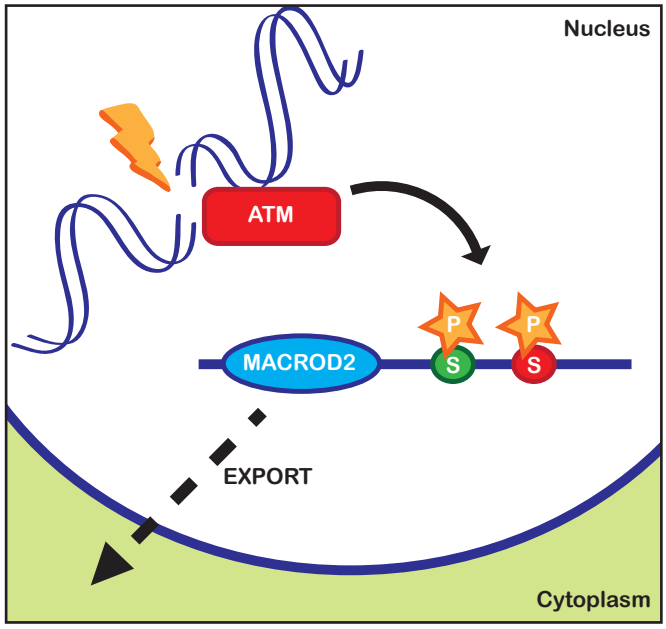


Figure 3.46: The model of MacroD2 nuclear export - Upon DNA damage, most likely via DSB formation, ATM gets activated. ATM modifies the unstructured C-terminus of MacroD2 in two serines: S345 and S415. This event induces the regulated export of MacroD2 from the nucleus.

Upon DNA damage, the activation of the ATM kinase induces the phosphorylation of two serines in the MacroD2 C-terminal unstructured region. These two serine residues (S345 and S415) are both followed by a glutamine, which is a motif that is directly recognized by ATM, ATR and DNA-PK kinases (220). Since the mutation of the Q416 showed loss of phosphorylation, a PI3K-like kinase is likely performing the modification. As long as ATM kinase is available and active, the inhibition of ATR or DNA-PK does not affect MacroD2 nuclear export. However, in other conditions, such as ATM knock-out, RNAi-mediated knock-down or activation of ATM-independent DNA damage reponse pathways, ATM kinase can be replaced in inducing MacroD2

nuclear export by other PI3K-like kinases, such as DNA-PK.

I then showed that for both exporting sites a short stretch of sequence ($10 - 15$ amino acids long) is able to target EGFP for nuclear export. These results exclude the possibility that the modification might unmask a hidden nuclear exporting sequence, but rather suggest that the phosphorylation induces the interaction of these stretches with factors that enable the nuclear export. Within each of the two exporting sequences, a methionine is essential for the export: in fact, when mutated into alanine, phenylalanine or isoleucine, the nuclear export of MacroD2 fragments cannot take place. However, when comparing the two exporting sequences, the residue is not present at the same position in relation to the phosphoserine.

I then considered the possibility that MacroD2 is exported from the nucleus via a "piggy-back" mechanism by interacting with a NES-carrying protein. Preliminary data from a co-immunoprecipitation coupled to peptide mass fingerprinting experiment suggested a 14-3-3 protein as a viable candidate as an interaction partner. These chaperone proteins can bind a phosphorylated motif in the cargo protein and cause the protein to accumulate in the cytoplasm. Although I have performed three different types of experiments to test their binding, I did not see any interaction. Indeed, a scientist can only show that a certain phenomenon happens, therefore it is not possible to show without any doubt that two proteins do not interact: I could spend years, endlessly changing the experimental procedure and never be sure that I found the right condition to see the interaction. However, considering the chaperone nature of the 14-3-3s and the amount of material I used for the experiments, it is likely that the interaction between the partners could easily have occurred. Since it did not happen, it is unlikely that the 14-3-3s are the exporting factors interacting with MacroD2.

Finally, I compared the recruitment dynamics at the DNA lesion of MacroD2 full-length, macrodomain and C-terminus fragment. While the whole protein is both able to recruit and export from the nucleus, the macrodomain-only construct remains at the DNA lesion for a much longer time, due to its inability to export. This is confirmed also by the MacroD2 full-length exporting-deficient mutant, as well as by the inhibition of ATM activity by small inhibitor. Thus, the nuclear export of MacroD2 affects its own residency at the DNA lesion and might impact the DNA damage response.

3.10.1 MacroD2 C-terminus is an intrinsically disordered region

MacroD2 is a protein with a globular N-terminal macrodomain domain (56, 67). However, the C-terminal portion of the protein has not been previously annotated. My experiments indicated that the C-terminus region is unstructured **(see Figure 3.14)**. The circular dichroism spectrum exhibits the features of a totally unstructured region, with this portion likely to be unstructured also within the full-length protein context. Surprisingly, although unstructured fragments are generally associated with insolubility in physiological conditions, the purification of the C-terminal fragment was successful **(see Figure 3.13)**, likely due to the amino acidic composition of the sequence, rich in residues that form hydrogen bonds. The composition is indeed in line with the requirements for the intrinsically disordered proteins and might allow the correct solubility of the fragment (358).

Intrinsically disordered proteins are defined as having at least a region that, although able to assume different conformation with various degrees of order, is still biologically active (358). In eukaryotic proteomes, the presence of proteins with at least a short disordered stretch seems quite important, with at least 1500 proteins showing this feature. The presence of a disordered region makes the protein able to bind interactors more dynamically, by lowering the free-energy barriers. For this reason, the disordered proteins represent important hubs in signaling pathways, with p53 being the most notorious example of disordered proteins (359). This might be connected to the fact that intrinsically disordered proteins are on average substrate to twice as many kinases as the structured ones (360).

The fact that the MacroD2 C-terminus is an intrinsically disordered region could be very interesting for two reasons. First, an intrinsically disordered domain confers augmented interacting capabilities, and could represent a fundamental source of both positive and negative regulation of the enzymatic activity. Secondly, the disordered proteins are underrepresented in the group of proteins involved in catalysis and biosynthesis (361). Therefore, MacroD2 could represent an intriguing anomaly in the world of the flexible proteins.

3.10.2 MacroD2 ratio across the nuclear envelope changes upon DNA damage

MacroD2 is a protein 448 amino acids long with a molecular weight of 50 kDa. Since almost half of the protein is disordered, the shape of MacroD2 is more similar to a 25 kDa globular protein with a long, highly hydrated tail, than a globular protein of 50 kDa. This change in shape should be taken into account for what concerns the mode of passage through the nuclear pore.

According to the laboratory of Dr. D Görlich (Max Planck Institute for Biophysical Chemistry, Göttingen), a non-spherical probe, like a GFP-GFP fusion, although having a mass of 54 kDa, crosses the nuclear pore faster than the globular maltose-binding protein, which is 43 kDa in weight (*mentioned as unpublished datum in* (300)). This observation is explained by the possibility that the protein could passively cross the membrane with a preferred orientation. Even if such a concept could seem counterintuitive in light of our knowledge about how proteins behave in gels and pores, it well explains why the EGFP-tagged version of MacroD2 can be easily found in both nuclear and cytoplasmic compartments. For this reason, when there is no DNA damage in the cell, MacroD2 might passively diffuse across the nuclear envelope.

However, even if MacroD2 is constantly redistributing itself between the nucleus and the cytoplasm in a passive manner, upon DNA damage the flux of MacroD2 towards the cytoplasm is increased. The net flux out across the nuclear envelope comes from the combination of two components: the influx and the efflux of the protein, the first factor being smaller than the second (362). The net flux of MacroD2 into the cytoplasm could be explained either by increased efflux and unchanged influx or by unchanged efflux and decreased influx. In this second case, the influx contribution can reduce so much to be null, due for example to the protein being anchored in the cytoplasm.

According to the Fick laws for diffusion applied to a permeable membrane (363), the flux of of the molecule X across the membrane (J_x, n/m^2s) is the result of the difference of concentration of the protein in the two compartments, multiplied for a constant P, so that:

$$J_x = P([X]_2 - [X]_1)$$

where $[X]_1$ is the concentration of the molecule X in the first compartment, while $[X]_2$ is the concentration in the second one. The constant P is the permeability in the membrane of the molecule X. Thus, the increase in MacroD2 flux out might be caused by a change in the gradient across the membrane, maybe because MacroD2 is released by some anchoring in the nucleus (for example from chromatin) or is anchored once it arrives in the cytoplasm. Another cause of efflux increase could be a change in the permeability coefficient: for example, MacroD2 molecules might be positively regulated to interact with the export machinery, resulting in a facilitated transport.

A fluorescence loss in photobleaching (FLIP) experiment, where cytoplasmic EGFP-MacroD2 signal is constantly photo-bleached, allowed Dr. Timinszky to determine the efflux upon control or DNA damage treatment **(Dr. G. Timinszky,** *unpublished***)**. By transforming the single-exponential curve into a linear function by change in scale, he was able to determine the P constant in the two conditions. In fact, upon DNA damage, the constant was higher. Thus, MacroD2 efflux is not due to a change in gradient across the envelope, arguing against the anchoring hypothesis. Instead, this result supports the hypothesis that MacroD2 interacts in a regulated manner with a component of the exporting machinery upon DNA damage.

In summary, since EGFP-tagged MacroD2 is distributed always between the nucleus and the cytoplasm, it is possible that such a distribution is true also for the endogenous protein. Instead of imagining MacroD2 fully nuclear or cytoplasmic, MacroD2 could show a precise ratio between the nucleus and the cytoplasm, possibly connected with the unknown function that it has (or has not) to perform at a certain physiological state. However, upon DNA damage and activation of the response, the phosphorylation shifts the equilibrium towards the cytoplasmic accumulation. Most likely it favors the occurrence of more enzymatic reactions in the cytoplasm, while it discourages those in the nucleus.

3.10.3 The mechanism behind MacroD2 passage through the nuclear envelope

According to the data that I have collected, MacroD2 accumulates in the cytoplasm upon DNA damage in a regulated manner. However, MacroD2 lacks targeting sequences, both for the import and for the export **(see Section 3.3)**. As a mechanism

underling this accumulation, the two short MacroD2 exporting sequences are phosphorylated, triggering the nuclear export.

One hypothesis is that this modification could induce the binding to an adaptor protein carrying a NES: therefore, MacroD2 would export by a "piggy-back" mechanism. Considering the different export-deficient mutants **(see Figures 3.37 and 3.38)**, the binding seems to require the presence of a methionine, possibly important to form contacts with the binding pocket of the exporting adaptor. The position of the methionine was different in the two exporting sequences. Although it is possible to argue that the mechanism regulating the two exporting sequences might be different, due perhaps to different partners, it is difficult to explain that the methionine is essential in both sequences by mere coincidence. Thus, considering the phosphorylated serines and the methionines as requirements for the binding of the exporting interactor, I tested whether the 14-3-3 proteins can meet these requirements. Even though the 14-3-3 proteins are known phospho-binders, these candidates did not show any clear binding upon *in vitro* and *in vivo* experiments **(see Section 3.8)**.

Together with the direct interactor, I lack clues as to which exportin protein might be involved in the nuclear transport. Exportin 1 is the most common transporter for the nuclear export (303). The involvement of exportin 1 is generally tested by use of leptomycin B, a compound that blocks exportin 1 interaction with cargoes by modifying the binding pocket (364). When Dr. Timinszky used this compound, MacroD2 nuclear export was only partially blocked, thus arguing against its exclusive involvement **(Dr. G. Timinszky,** *unpublished***)**. More data should be collected to have an exhaustive picture about the involvement of other exportins.

An alternate hypothesis is that the phosphorylation might confer increase MacroD2 affinity for the components of the NPC, allowing MacroD2 to cross the membrane more often. In fact, a study showed that nucleoporins contains hydrophobic residues yet are overall positively charged, which well matches with the overall negative charge of the nuclear transport receptors (365). This would be also compatible with the overall negative charge of the C-terminal portion of MacroD2.

As mentioned previously, the mutation of a methionine within both MacroD2 exporting sequences completely inhibits the nuclear export. Since the methionine is a very hydrophobic residue, it could be important for the interaction with the hydrophobic residues of the nucleoporins. After all, the loss of binding due to methionine oxidation

in few cases of protein-protein interaction studies shows how essential the methionine can be in such a situation (366, 367, 368). Also, this hypothesis would better explain why the methionine does not have the same position in the two MacroD2 exporting sequences, since the successful interaction for the export would not require a fixed consensus sequence, but only the position of hydrophobic residues in a negatively charged patch. Therefore, it could be possible that the phosphorylation of the serines shifts the balance towards an even more unfolded and negatively charged sequence, allowing the methionine to directly bind the FG repeats of the nucleoporins present in the NPC. It would probably be the first case of "regulated passive diffusion".

Therefore, according to this hypothesis, the phosphorylation would allow faster translocation through the nuclear pore, regardless of the direction of the transport. To assure the directionality from the nucleus to the cytoplasm, the phosphorylation step must be assumed to take place within the nucleus, while MacroD2 should be preferentially dephosphorylated in the cytoplasm. In support to this hypothesis, the mutation of the MacroD2 exporting serines into glutamate did not show any net nuclear export of MacroD2 **(Dr. G. Timinszky,** *unpublished***)**.

However, mutation of the methionine into other strongly hydrophobic residues (isoleucine and phenylalanine) impairs the nuclear export **(see Figure 3.39)**. Although all are hydrophobic residues, such a result argues for a specific role of the methionine, that cannot be mediated just by its hydrophobic nature. This result indeed argues against this alternative hypothesis.

Nonetheless, the methionine is an important element for the accomplishment of MacroD2 nuclear export. Such an important role of the methionine could then explain a little conundrum that involves MacroD2 and oxidative stress: at the beginning of my PhD, I tested MacroD2 nuclear export upon H_2O_2 treatment, and although being preliminary experiments, both Dr. Timinszky and I found that MacroD2 nuclear export was minimal **(Figure 3.47)**. Although more experiments and more controls are needed to transform this preliminary observation into a result, the results I have successively collected might explain this anomalous behavior. In fact, now we know that the methionine is necessary for the MacroD2 nuclear export and it is generally oxidized upon oxidative stress (366).

Oxidative stress is a powerful inducer of ADP-ribosylation and ARTD1 activity. However, a set of studies showed that, beside DBSs, oxidative stress can also activate

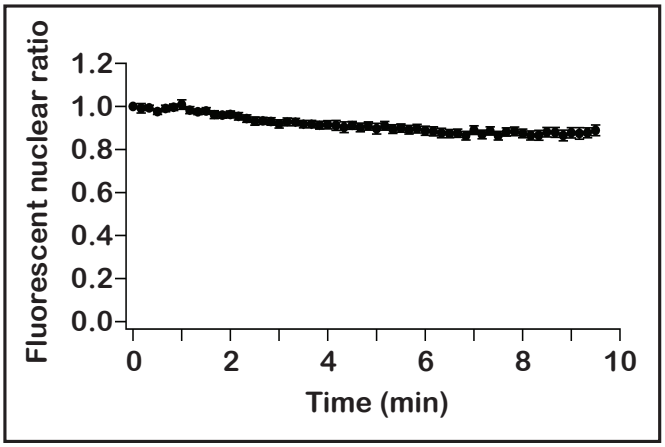


Figure 3.47: Oxidative stress fails to trigger MacroD2 nuclear export - Quantification of live-cell imaging of U2OS-mEGFP-MacroD2 C-terminus+mCherry-H2B cells treated with H_2O_2, 600 μM. 50−100 cells were quantified analyzed from one experiment. Nuclear/cytoplasmic ratio was calculated using CellProfiler 2.0 (331), and the mCherry signal was used for the nuclei segmentation. Error bars, 95 %CI.

ATM kinase (275). Therefore, the activity of ATM should lead to MacroD2 nuclear export, but this does not happen. A first explanation could be the absence of MacroD2 phosphorylation, because ATM might change its substrate subset according to the type of DNA damage type and activation mode, although it is still not clear whether it is true (263). On the other hand, the lack of export may not be due to a change in the modification of the MacroD2 C-terminus region, but because of impaired interaction with the export machinery. The oxidation of the methionine, as a consequence of the oxidative stress, could be thus a mechanism to override ATM regulation and still keep MacroD2 within the nucleus to perform its activity. Such a mechanism would be then negatively regulated by the activity of the methionine sulfoxide reductase A, which can restore the native state of the methionines (368).

3.10.4 ATM kinase regulates MacroD2 nuclear export

In my experiments, I showed that ATM is the primary kinase that induces MacroD2 nuclear export **(see Section 3.4)**. But MacroD2 may not be the only export target of ATM: ATM activation has been shown to induce changes in the localization of several proteins. In some cases, a direct intervention of ATM has been shown to affect the nuclear export of some factors (226, 324, 369). The NF-κB essential modulator, NEMO,

requires three sequential post-translational modifications for its nuclear export and ATM phosphorylation is the second; ATM is also shown to cross the nuclear envelope together with NEMO (226). COP1, instead, is the E3 ubiquitin ligase that targets p53 for degradation (369). ATM-induced nuclear export of COP1 is important to stabilize p53 and activate it.

ATM modifies a vast number of proteins (42, 44). Many of these targets have more than one site modified and contain a stretch of 100 residues with three phosphorylated SQ/TQ motifs, called a SQ/TQ cluster domain (SCD) (370). In a recent meta-study, among the proteins with SCDs, unexpected cellular processes were enriched, such as "gene regulation", "vesicle-mediated traffic", "cytoskeletal components" and "neural development" (371). However, I realized that in this meta-study, MacroD2 did not appear, even though it does meet the definition (S345+S415+S426). This is explained by the fact that the first isoform annotated on Genbank, which the authors of the study used for their study, contains only one of the three SQ motifs. The S415 and S426 are in fact included in an exon that is likely spliced in alternative ways, considering all the splice variants annotated on Uniprot **(Figure 3.48)**. The exon sequence is very conserved among vertebrates.

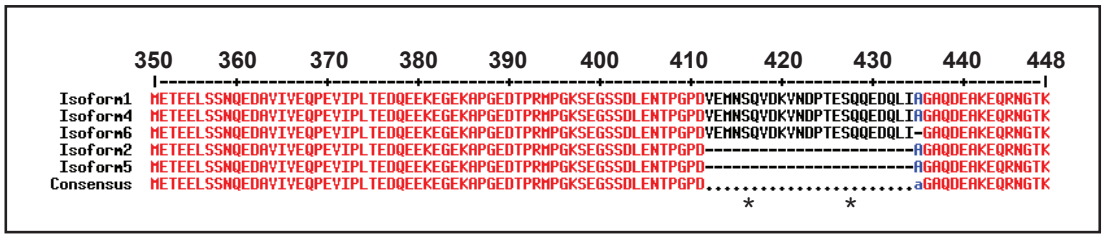


Figure 3.48: Comparison of the sequence 350-448 in the different annotated MacroD2 isoforms - The sequence between 350 and 448 amino acids of all the isoforms of MacroD2 present on Uniprot database is shown. On the bottom line is the consensus sequence for the alignment. The stars indicate the position of the two SQ motifs. Alignment performed with MultAlin (372)

Such a hypothesis would represent an interesting mode of regulation of the protein. Since the two SQ motifs S345 and S415 have an additive effect on the nuclear export of MacroD2 **(see Figure 3.33)**, the regulated exclusion of one of them could impact on the dynamics of MacroD2 nuclear export. I can thus imagine the possibility of differential splicing according to cell types or environmental stimuli.

When testing the involvement of ATM, I showed the complete blockage of MacroD2 export upon inhibitor treatment in the case of UV-laser microirradiation with medium energy and etoposide treatment, as well as in ATM$^{-/-}$ HT-144 cell line treated with etoposide **(see Section 3.4)**. The preliminary work with camptothecin also suggest the exclusive involvement of ATM **(see Figure 1.15)**. In the other conditions that I tested, such as RNAi-mediated ATM knock-down or HT-144 cells treated with UV-laser microirradiation, the inhibition of MacroD2 nuclear export was only partial.

The RNAi-mediated knock-down of ATM showed only a partial if not minimal effect on the MacroD2 nuclear export dynamics **(see Figures 3.18, 3.19 and 3.20)**. This result could be explained first of all by the partial depletion of ATM kinase, whose residual amount might still suffice to accomplish the kinase functions, although with a decreased rate. On the other hand, the partial impact of the ATM depletion on MacroD2 export rate could be due to the activation of other kinases that replace ATM when its presence in cells is drastically decreased, as shown also in the case of the ATM knock-out cell line, where DNA-PK clearly activates MacroD2 nuclear export **(see Figure 3.23)**. Nonetheless, the fact that the absence of ATM might allow other kinases to modify non-physiological substrates could represent either a back-up pathway or the occurrence of unwanted interactions with possible pathological consequences.

Lastly, the impact that the siRNA treatment has on the two MacroD2 constructs (full-length or C-terminal region) is different. While the C-terminus export is reduced by half its maximum level upon ATM depletion, the full-length protein export is affected only marginally **(see Figures 3.18, 3.19 and 3.20)**. It is possible that the presence of the macrodomain favors the interaction between the kinase and the substrate, limiting the effect of the ATM depletion in this specific case. Indeed, the presence of a functioning macrodomain shows to have an impact, even if minimal, on the nuclear export dynamics when comparing the MacroD2 full length wild-type protein with the mutant deficient in ADP-ribose binding (G188E) **(see Figure 3.26)**.

However, the EGFP-EGFP-MacroD2 C-terminus, which has a comparable size to the EGFP-MacroD2 full-length construct but lacks the macrodomain, shows likewise a minimal change in export rate upon ATM depletion, supporting the idea that the presence of the macrodomain cannot answer for this reduced effect of the knock-down **(see Figures 3.28)**. Most likely, it is the size of the protein in the first place that defines the speed of translocation across the nuclear envelope. For slower translocators,

such as EGFP-MacroD2 full-length and EGFP-EGFP-MacroD2 C-terminus, crossing the nuclear envelope, not the phosphorylation event, may be the limiting step of the process.

3.10.5 ADP-ribosylation meets ATM signaling

ADP-ribosylation is a very dynamic post-translational modification, broadly involved in the regulation of DNA damage repair. Similarly, ATM is activated by a number of genotoxic stimuli and its main function is to oversee the correct repair and survival of the cell, or to induce its programmed death, if damage is beyond repair (336). It is not a surprise that these two signaling pathways show a number of connections. Several factors involved in homologous recombination and ATM signaling, including ATM itself, recruit to DNA breaks thanks to PAR-binding modules (8, 156). For example, both Nbs1 and Mre11 initial recruitment depend on ARTD1 activation, and subsequently the Mre11-Rad50-Nbs1 (MRN) complex induces increased levels of ATM activation (155).

In general, the recruitment of the repair factors employs modules that recognize poly-ADP-ribosylation, since a vast platform of PAR molecules offers a better substrate for the formation of a three-dimensional network (6). The negative regulation of this recruiting platform is thus bestowed mainly on PARG activity, as shown by the change in recruitment dynamics in the presence and absence of RNAi-mediated PARG depletion (56).

While the role of PARylation in DNA damage response is well established **(see Section 1.2.6)**, more evidence continues to be collected about the contribution of MARylation in DNA repair regulation. ARTD3 is a MART that is involved in DNA damage repair, namely in the choice between HR and NHEJ pathways (373, 374). ARTD10, the prototypical MART, is instead involved in the regulation of trans-lesion DNA synthesis, which occurs when the replication fork encounters a DSB site (175).

In my study, while inhibition of the recruitment of MacroD2 to DNA lesions did not have any impact on its nuclear export dynamics, the nuclear export did affect the recruitment dynamics **(see Figure 3.45)**. The recruitment to the DNA damage site is a feature of the macrodomain **(see Figure 3.44)**, and it is quite clear that the fluorescent signal persists for 30 minutes at the damage focus if not promptly removed by the nuclear export **(see Figures 3.44 and 3.45)**.

Furthermore, inhibition of ATM or mutation of the SQ motifs result in the failure of MacroD2 nuclear export and its persistence within the nucleus. Indeed, a longer presence of MacroD2 in the nucleus means a longer presence at the DNA lesion **(see Figure 3.45)**, which might easily be a source of misregulation of the repair pathway. Thus, the ATM-induced nuclear export of MacroD2 could represent another case supporting the role played by MARylation in the regulation of the DNA repair process. In particular, MacroD2 might negatively regulate the DNA damage response.

3.11 Future directions

3.11.1 Regarding the location of MacroD2 and the mechanism of export

While I could show that endogenous MacroD2 still exports from the nucleus upon etoposide treatment using an immunofluorescence approach **(see Figure 3.10)**, I could not define its cellular localization in the two conditions. It would be important to address this question in order to understand whether the whole pool of protein relocates or just a fraction. As the antibody used in this study has limits, as previously discussed, it is essential to develop new immunological tools. Currently, a PhD student in our lab, G. Möller, is preparing and testing new monoclonal antibodies developed together with the "Monoclonal Antibodies" facility of the Helmholtz Zentrum, Munich. Once the tool has been developed, it is possible to test the real localization of the endogenous MacroD2 upon control and DNA damage condition. I hypothesize that MacroD2 will most likely be present in both the nuclear and cytoplasmic compartment in the control condition, showing a common behavior with the EGFP-tagged MacroD2. However, whether there will be residual MacroD2 left in the nucleus upon DNA damage or if the nuclear export will be complete, remains unclear. If the endogenous MacroD2 is generally present in both nuclear and cytoplasmic compartments, except during the DNA damage response, this result would argue for the continuous shuttling of MacroD2 across the nuclear envelope.

To rule out that MacroD2 signal is distributing across the nuclear envelope because of a form of active transport, it would be necessary to test the distribution of MacroD2 upon RNAi-mediated knock-down of importins or exportins. When such an experiment is performed in the absence of DNA damage, it can discriminate between

passive diffusion and active transport: if the depletion of at least one importin or exportin changes the nuclear-cytoplasmic ratio, it suggests that MacroD2 distribution is mediated by a form of active transport dependent on importin/exportin cycles. If the ratio is unchanged, it means that in the absence of DNA damage MacroD2 moves via passive diffusion between the two compartments.

Since the depletion of the importins and exportins using RNA-interference might be only partial, interfering with the formation of a more definite phenotype, the use of knock-out cell lines for the different factors would probably be the ideal way to define whether MacroD2 distribution is indeed due to a regulated active transport. Nonetheless, the results of such experiments might still show a confusing output, since the nuclear transport receptors are essential factors for the healthy localization of cellular components and their absence may generate pleiotropic effects.

The mechanism of transport of MacroD2 upon DNA damage could be different from the mechanism with which MacroD2 crosses the nuclear envelope in the absence of stress. If MacroD2 exports upon DNA damage by an active transport mechanism, it will most likely be through a "piggy-back" mechanism, where phosphorylation of MacroD2 enables it to bind another protein that contains a NES. MacroD2 protein, in fact, lacks annotated or predicted NES sequences. Since it is a matter of finding the right interactor, the best option is to adopt a pull-down approach coupled with mass spectrometry. This is indeed what I did and I will describe it in the next chapter **(see Chapter 4)**.

The second hypothesis I formulated consists of a "regulated passive diffusion". According to this hypothesis, MacroD2 would be intrinsically able to bind the nucleoporins without any intermediary and to cross the nuclear envelope. Upon DNA damage, the phosphorylation of its intrinsically disordered region would further favor the interaction with the nuclear pore complex. The interaction may take place due to the presence of negativly charged residues and few hydrophobic residues, positioned in strategic positions: among them, the two methionines (M350 and M413) would be important for the interaction with the FG-rich motifs of the nucleoporins. This hypothesis could be refuted by the fact that the MacroD2 fragment in which the methionine 413 has been mutated to phenylalanine or isoleucine showed no nuclear export. Nonetheless, it could still be interesting to perform the *in vitro* pull-down using the MacroD2 exporting

fragments and nucleoporin FG-rich tail fragments in the presence and absence of DNA damage, to test if the two proteins can directly interact.

Regardless from the mechanism of the translocation of MacroD2, my experiments showed the importance of the methionine for the accomplishment of the nuclear export of MacroD2. Since the oxidation of the methionine has been shown to affect protein-protein interactions, I wonder if the oxidation of MacroD2 methionines upon H_2O_2 treatment could be a way to override ATM-induced nuclear export and to keep MacroD2 within the nucleus to perform its function. First of all, more experiments are needed to define whether MacroD2 nuclear export is responsive or not to oxidative stress. Also, I would test the PAR levels in the cells upon H_2O_2 treatment, in order to test the export in a condition where the oxidative stress response is active.

Once the anomalous behavior of MacroD2 is confirmed, I would test if the MacroD2 exporting fragments would be phosphorylated also upon oxidative stress. In adopting the same protocol that I used for the experiments with the radioactive γ-ATP **(see Figures 3.29 and 3.36)**, the Benzonase nuclease and the ATM inhibitor could be used as positive and negative controls, respectively. If these fragments are still phosphorylated by ATM, I would then test if the methionine is oxidized. To this end, it could be possible to analyze the same fragments with mass spectrometry, looking just for the oxidation of these residues. Since the oxidation of the methionine residues frequently occurs during sample preparation for mass spectrometry, it is vital to perform the experiment also with the control sample, without DNA damage, and to evaluate the difference in oxidation. Then, once it is determined if MacroD2 exports by interacting with a specific adaptor protein or with the nucleoporins, if upon H_2O_2 the methionines are indeed more oxidized, the efficiency of the interaction between MacroD2 exporting fragments and the counterpart should be tested in the presence of either the unmodified or oxidized methionine. Addition of reducing agents in the unmodified samples should help increase the difference in binding between the two conditions.

3.11.2 Regarding ATM and the regulation of MacroD2 nuclear export

Upon DNA damage, ATM is activated and phosphorylates MacroD2. This event triggers the nuclear export of the ADP-ribosylhydrolase. While the most studied activation mode of ATM is the induction by DSBs, the literature shows many other ways by which ATM can be activated **(see Section 1.3.3)**. It would be interesting to screen

for MacroD2 nuclear export in many other conditions: hypotonic stress, cloroquine treament for the ATMIN-dependent activation, or hypoxia.

Additionally, many targets of ATM have more than one modifiable site, such as MacroD2, as it has been defined by this study. However, two of the three SQ motifs in MacroD2 that could form a SQ/TQ cluster domain are present on a putative alternatively spliced exon **(see Figure 3.48)**. The alternative splicing of this exon might confer an additional layer of regulation to MacroD2 nuclear export and its dynamics, since it could define how fast MacroD2 should be exported from the nucleus. However, to test this hypothesis it is necessary to know the function of MacroD2, in order to define the conditions to compare and to design the experiments accordingly.

For example, one of the probable functions of MacroD2 is the negative regulation of PARylation at DNA lesions, which avoids hyper-activation of ARTDs and subsequent *parthanatos*. Thus, MacroD2 splicing could be tested in a context prone to *parthanatos*. Indeed, the generation of a MacroD2 isoform that exports more slowly from the nucleus would generate a bigger barrier against this type of programmed death. There are indeed physiological situations where high intrinsic production of reactive oxygen species is reported, such as activation from quiescence in hematopoietic stem cells or neural development (375, 376). Since these tissues suffer from substantial oxidative stress, the activation of PARylation is expected. Moreover, oxidative stress is also able to activate ATM, although it would be important to confirm the activation of the kinase also in these study models. Then, specific probes for the whole protein and for the specific exon should be designed in order to perform real-time quantitative PCR, in order to define both the expression levels of MacroD2 in these tissues and to answer for the possible exon exclusion.

3.11.3 Regarding the role of MacroD2 upon DNA damage

The nuclear export of MacroD2 and its ATM-dependent regulation upon DNA damage response suggests that MacroD2 might be involved in the DNA damage response. Moreover, this behavior negatively correlates with the residency time of MacroD2 at the DNA lesions **(see Figures 3.44 and 3.45)**. For this reason, it would be interesting to test if MacroD2 is able to affect the DNA repair.

To test this hypothesis, one could compare the wild-type and MacroD2 knock-out cell lines previously used **(see Figure 3.10)** in specific DNA repair pathway assays. By

transiently transfecting GFP-based reporter plasmids, it is possible to easily monitor the efficiency of specific repair pathways, such as homologous recombination (HR), c-NHEJ, single-strand annealing (SSA) and alt-EJ (377). Such reporter plasmids contain sequence for a GFP molecule divided into two parts and, between them, the site for a very rare restriction enzyme, Isce-I. The reporters are designed so that upon cutting with the specific enzyme, the GFP gene is successfully generated if the specific DNA repair pathways has been used. To this end, it is possible to monitor the amount of cells that underwent the specific repair by cell sorting.

If there is any difference between the control and the knock-out cell lines in a specific pathway, the experiment should be completed with the rescue of the phenotype. Transfecting the MacroD2 knock-out cells with MacroD2 wild-type, MacroD2 ADP-ribose binding-deficient or MacroD2 export-deficient constructs will help defining the function of MacroD2 needed for the repair pathway. I would expect that when wild-type or export-deficient mutant MacroD2 are transfected, the pathways that require ATM activation, like HR, would show a decrease in their repair efficiency, since the export is likely a mechanism to decrease the presence of MacroD2 at the DNA lesion and the export is activated by ATM itself. The binding-deficient mutant, instead, should act like the negative control. However, for those pathways where MacroD2 seems not to be exported from the nucleus and its activity could be welcome at the DNA lesion, like c-NHEJ (dependent on DNA-PK) or alt-EJ (dependent on ARTD1), the repair efficiency might not be affected.

The only caveat in this plan is that the transfection of MacroD2 constructs represent a strong over-expression compared to the endogenous: MacroD2 is a protein of lowly abundance, whose expression is regulated at transcriptional and possible post-transcriptional level, given the anti-sense non-coding RNA overlapping its gene (*MacroD2* gene surroundings: `http://www.ncbi.nlm.nih.gov/gene/140733`). Since MacroD2 has an enzymatic activity, I cannot exclude that its over-expression might shift the system towards other deleterious effects. Therefore, the output of the experiment should be carefully evaluated considering these two aspects. Ultimately, this problem could be solved by restoring the endogenous promoter in the mCherry-construct, although this might only ameliorate the transcriptional regulation but not the possible post-transcriptional regulation.

However, if MacroD2 shows an impact on some DNA repair pathways, the next step would be to determine which factor can be ADP-ribosylated and are a target of MacroD2 enzymatic activity. The increasing number of publications on mass spectrometry-associated pull-downs of ADP-ribosylated proteins could be a further source of candidates to test. The recently generated online database of ADP-ribosylated proteins could be the suitable starting point for this search (378). *In vitro* experiments to test the functional activity or structural status of the substrate, with and without the modification, will be necessary to characterize the role of MacroD2 regulatory activity in the DNA repair pathway.

4

MacroD2 interactome in control and genotoxic condition

4.1 Introduction

Chapter 3 detailed the nuclear export of MacroD2 upon DNA damage. While phosphorylation by ATM and PI3K-like kinases was shown to induce the export, the exporting mechanism and the related machinery were not identified. If the most likely "piggy-back" mechanism for the nuclear transport indeed occurrs, the name of an interactor able to drive MacroD2 across the nuclear envelope is missing. Instead of painstackingly searching for the whole genome looking for possible candidates on hypothesis-driven experiments, the most time and cost-effective strategy at this point would consist of a unbiased approach on proteomic scale.

Moreover, MacroD2's real cellular function is still unidentified and the current list of MacroD2 interactors is so short that it is not helpful in answering this question. The few interactors of MacroD2 already published are mostly shown to bind in *in vitro* experiments: ARTD1 (56, 67), ARTD10 (56, 67), GSK3β (67). Other studies involving MacroD2 suggest functional roles in carcinogenesis and neural development, but they are mainly based on genome-wide association studies and fail to show any mechanistic insight behind the association **(see Section 1.2.5)**. Thus, knowing the complete interactome of MacroD2 and possibly the substrates of its enzymatic activity could be an important step to unlock its cellular role.

To this end, I chose to perform co-immunoprecipitation (co-IP) experiment combined with peptide mass fingerprinting to define the MacroD2 interactome in the presence and absence of DNA damage. I therefore employed two different treatments: DMSO and etoposide. Also, I used three different baits: MacroD2 full-length, macrodomain and C-terminal region fragments, in order to define differential enrichments caused by the two halves of MacroD2. The experiment was performed in collaboration with Dr. A. Schmidt from the Zentrallabor für Proteinanalytik, LMU.

By defining the interactome of MacroD2 with and without genotoxic stress, I aimed to address two points: first, I aimed to identify the proteins involved in MacroD2 regulated export. Morover, I aimed to define the areas of cellular life that are regulated by MacroD2 enzymatic activity.

4.2 Setup of the co-immunopurification experiment

4.2.1 Generation of stable cell lines

For the co-IP, I decided to use stable cell lines expressing the constructs of interest, thus avoiding possible inconsistencies due to transfection efficiency **(see Methods 6.17.1)**. To this end, we used HEK293 Flp-In^{TM} T-REx^{TM} by Thermofisher. There are several advantages associated with using this cell line. Firstly, HEK293 achieve much higher yield of cell-lysate relate to the amount of growing surface compared to our other cell-lines of use, U2OS and HeLa. Also, these commercial cells allowed us to obtain stable cell-lines with the gene integrated in the same spot, avoiding differences in expression or genetic background. Lastly, the promoter is normally repressed by a TetR protein, allowing inducible expression upon administration of doxycyclin.

I created four cell lines with the help of S. Grau and J. Preißer, each expressing one of the following:

- EGFP-MacroD2 full-length
- EGFP-MacroD2 macrodomain (aa 1-243)
- EGFP-MacroD2 C-terminus (aa 237-448)
- EGFP

The use of the EGFP tag allows for easy screening of colonies that successfully integrated the genes by directly checking with a bench-top microscope, followed by immunoblotting. EGFP is 25 kDa, which is quite large for being used as a tag. This is the same approximate size of the macrodomain or the C-terminus alone. Therefore, I cannot exclude that such a tag might impair possible binding events, even if EGFP is considered to be an inert protein and less prone to interactions. While the size of EGFP may prevent interactions due to steric impairment, EGFP-tagged MacroD2 constructs are still able to recruit to the DNA lesion and to export from the nucleus, suggesting that the binding ability is not affected. Moreover, one of the generated cell lines expresses only the EGFP alone, in order to exclude possible interactors accounted to the tag and not to MacroD2.

The Flp-In^{TM} T-REx^{TM} system is very convenient. The HEK293 cells obtained directly from Thermofisher present two integrated constructs, one with the FRT site, selected through a zeocin-selection marker, and one with the TetR gene expressing the repressor and presenting the blasticidin-selection marker. Growing cells in zeocin-blasticidin selection medium ensures the maintenance of the FRT site in the cell population. The genes of interest were cloned into a pcDNA^{TM}5/FRT/TO plasmid, which has a hygromycin-selection marker. After seeding the cells, the hygromycin-selected plasmid is transfected together with the pOG44 plasmid containing the Flp recombinase: in the cells that acquire both plasmids, the transiently expressed Flp recombinase leads to the integration of the gene of interest to the specific point in the genome where the FRT site is located. Also, the integration makes the cells resistant to hygromycin but sensitive to zeocin, thus allowing the proper selection of the cells that have correctly performed the integration. After two weeks in selection media with hygromycin and blasticidin, single-cells colonies showing expression of green fluorescence were isolated before amplifying them for immunoblots and functional testing.

Before each experiment, the cells were tested for induction of protein expression upon doxycyclin treatment **(Figure 4.1)**. In order to confirm the identity of the cells used for the pull-down experiments, induced cells were also tested for their functional behavior upon laser microirradiation. Each induced protein preserved the same behaviors and even kinetics in HEK293 cells. EGFP-expressing cells did not show any change in the protein localization upon UV-laser microirradiation. EGFP-MacroD2 full-length showed both export from the nucleus and recruitment to the DNA lesion

with a peak at 2.5 minutes, similar to what previously shown for HeLa cells **(see Figure 3.44)**. Likewise, EGFP-MacroD2 macrodomain recruits to the DNA lesion while the EGFP-MacroD2 C-terminus exports from the nucleus.

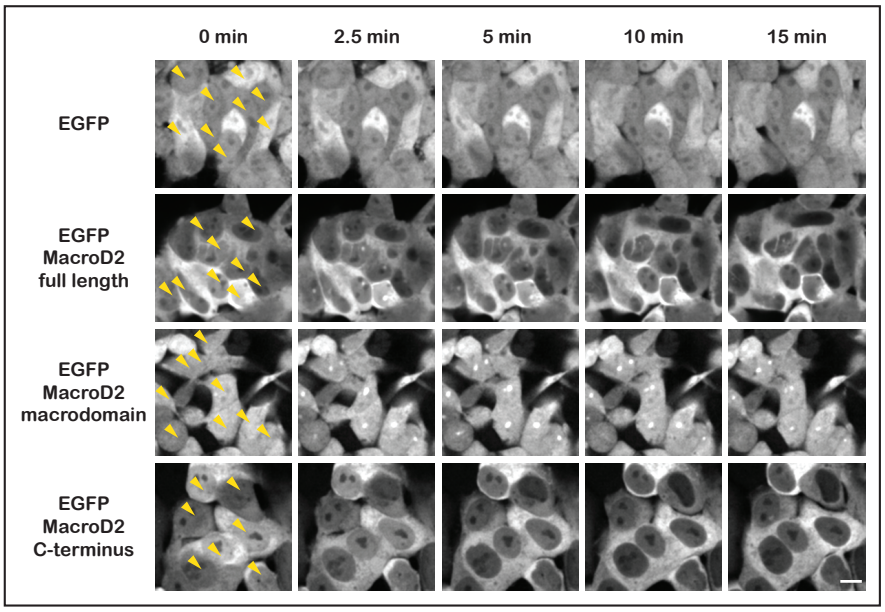


Figure 4.1: Induction and functional check of HEK293 stable cell lines - The four HEK293 stable cell lines (EGFP; EGFP-MacroD2 full-length; EGFP-MacroD2 macrodomain; EGFP-MacroD2 C-terminus) are induced with doxycyclin 1mg/mL. After one day of incubation, cells were checked for the proper induction of the constructs expression and tested upon UV laser microirradiation for recruitment and/or export. The focus of laser micro-irradiation is indicated with yellow arrowhead. Scale bar, 10 μm.

4.2.2 Protocol for the co-immunopurification

The co-immunopurification was optimized and the final conditions were as follows. RIPA buffer was used for the generation of whole-cell lysates, as described in Methods 6.17.2. The incubation of the cell lysate with the GFP-trapping resin was performed for 2 hours at 4 °C and the wash of the resin was performed once with RIPA buffer and twice with PBS. All buffers were supplemented with protease inhibitor and phosphatase inhibitors cocktails, to prevent protein degradation and dephosphorylation, which may induce loss of interactors. Moreover, the whole procedure was performed at 4 °C, so that the enzymatic activity of MacroD2 was inhibited (56).

In order to increase the chances of detecting also weak interactors, I chose to use large amount of cell lysate. Therefore, the experiment required similarly large amounts of GFP-trap beads to achieve a satisfactory pull-down. Therefore, I evaluated whether to use the homemade agarose-coupled GFP-trap, prepared by G. Jankevicius and Dr. A. Bowman, or a commercially available product (agarose-coupled GFP-trap by Chromotek). I tested the two for the efficiency of pull-down of purified EGFP (10 μg) **(Figure 4.2)**. Although the commercial resin showed its superiority in the efficiency of pull-down, possibly due to a much higher density of coupled GFP-trap, I decided to use the home-made GFP-trap for the first replicate, with an increased amount of beads used in each sample.

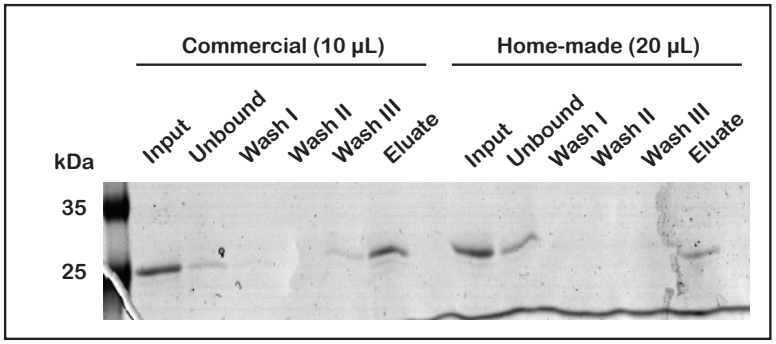


Figure 4.2: Comparison in pull-down efficiency between commercial and home-made GFP-trap - 10 μg of purified EGFP was used to a compared purification using 10 μL of commercial GFP-trap (Chromotek) or 20 μL of in-house made GFP-trap (G. Jankevicius and A. Bowman). Elution was performed with Laemmli buffer and purification efficiency was tested by Coomassie staining (loaded 2.5 % of the material in each lane). For the purification steps, see Methods 6.17.2.

Finally, to test the efficiency of the purification, I performed the co-IP with 6 mg of cell-lysate from cells expressing EGFP-MacroD2 full-length and 50 μL of home-made GFP-trap. This setup allowed an enrichment of bait by 1.6-fold compared to the input **(Figure 4.3)**. The Coomassie staining of the same samples, showed the sequential reduction of the unspecific bands within the series of washes. In the elution with Laemmli buffer, the patterns of residual bands is very different from the one present in the washes, suggesting the purification of specific interactors of MacroD2.

The elution step of my preliminary experiments was performed with Laemmli buffer in order to collect all the material attached to the beads and to have a better estimation

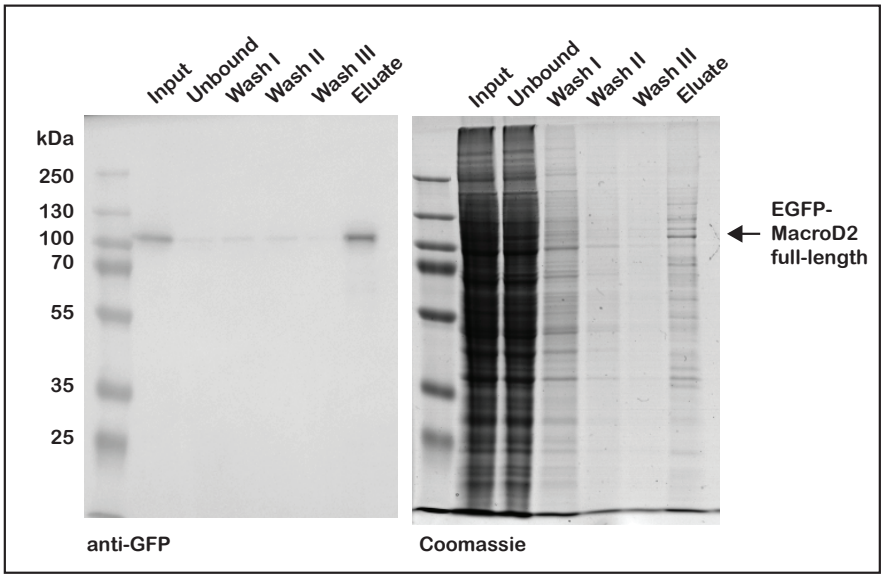


Figure 4.3: Pull-down efficiency test with EGFP-MacroD2 full-length - HEK293 EGFP-MacroD2 full-length cells were plated and induced with doxycyclin 1mg/mL. After one day, cells were collected and cell lysate was prepared. 6 mg of cell lysate was then purified with 50 μL of in-house made GFP-trap. Elution was performed with Laemmli buffer and purification efficiency was tested by western-blot with anti-GFP (loaded 2.5 % of the material in each lane) and Coomassie staining. For the purification steps, see Methods 6.17.2

of the immuno-purification efficiency. When the co-immunopurification experiments were coupled to mass spectrometry detection, the initial steps in the preparation for the mass spectrometry run were performed directly on beads by Dr. A. Schmidt. For the detailed protocol, see Methods 6.17.2.

4.3 Co-immunopurification experiment

Four biological replicates of co-immunipurification experiment were performed. For each replicate, the doxycyclin-mediated induction of the protein expression in the four cell line was tested using live-cell imaging coupled with UV-laser microirradiation, as shown in Figure 4.1. Once the correct induction was confirmed, the four cell lines were seeded in non-selective media. The three MacroD2 cell lines were plated in two distinct dishes, while the EGFP was plated once. Doxycyclin (1 mg/mL) was added to the plates 4 hours after plating. After 24 hours from induction, I treated the MacroD2 cell lines with DMSO or etoposide 10 μM for one hour in order to obtain the following conditions:

1. EGFP
2. EGFP-MacroD2 full-length DMSO
3. EGFP-MacroD2 full-length etoposide
4. EGFP-MacroD2 macrodomain DMSO
5. EGFP-MacroD2 macrodomain etoposide
6. EGFP-MacroD2 C-terminus DMSO
7. EGFP-MacroD2 C-terminus etoposide

The co-IP was complated as described in Methods 6.17.2. The first biological replicated was performed with home-made GFP-trap beads, while the commercially available resin was used for the remaining three replicates, due to the certified reproducibility in their performance compared to the home-made beads. The preparatory steps for the mass-spectrometry run and the runs themselves were performed by Dr. Schmidt **(see Methods 6.17.2 and 6.17.3)**.

4.4 Analysis and quality controls

The collected the data from the mass spectrometer were treated in order to transform a series of spectra into a usable output **(see Methods 6.17.4)**. The final output of a co-IP-peptide mass fingerprinting experiment should be a list of protein names, which represent the group of most probable interactors of the bait protein in a particular condition. From these protein lists it is then possible to perform statistical analysis to determine interesting patterns in the cellular functions and pathways.

For the peptide mass fingerprinting, the identification step is essential, yet delicate. Each peptide that flew in the machine has a specific mass/charge ratio (m/z). The MaxQuant software that we used is able to identify the protein of origin for each peptide by using the mass/charge ratio. The algorithm compares the mass/charge ratio of one experimentally-measured peptide with the values of all the peptides generated *in silico* from the human proteome. Therefore, the software can assign each detected peptide to a protein peptide. Such a process is bound to generate some error, which can be reduced by the optimized definition of the parameters available in the software. Dr. Schmidt chose the best parameters according to his personal experience and the knowledge of the instrument. The identification process is also performed at the same time for all the samples, so that the parameters used are exactly the same.

The identification step generates a table in which associates each peptide with its protein of origin and its specific signal intensity. The MaxQuant-embedded algorithm calculates the intensity of each protein in a sample by considering the peak intensity of all the found peptides related to the protein in the mass spectrometry spectra (379).

To have a first view of the quality of the different replicates, Dr. Schmidt compared the average of the signal intensities and the missing values in the four replicates. In fact, as error is normally generated upon mass spectrometry detection and/or peptide identification, from the total subset of protein detected in the whole experiment, not all the proteins are found in all the samples. These proteins that have not been detected are defined as missing values and are generated either by the loss of the signal at the detection or for a possible incorrect identification. For example, the protein A is present in replicate 1 and 3 but not in replicate 2 of the same condition. Therefore, in the replicate 2 a missing value is present for that protein. This initial analysis allowed Dr. Schmidt to realize that the first biological replicate was behaving very differently

compared to the others in terms of signal intensities and number of peptides associated per each protein (which are also known as spectral counts) **(Figure 4.4)**. Also, all the samples of replicate 1 presented many more missing values than the other replicates. Since this replicate was the only one performed with the home-made beads, we decided to exclude it from the analysis.

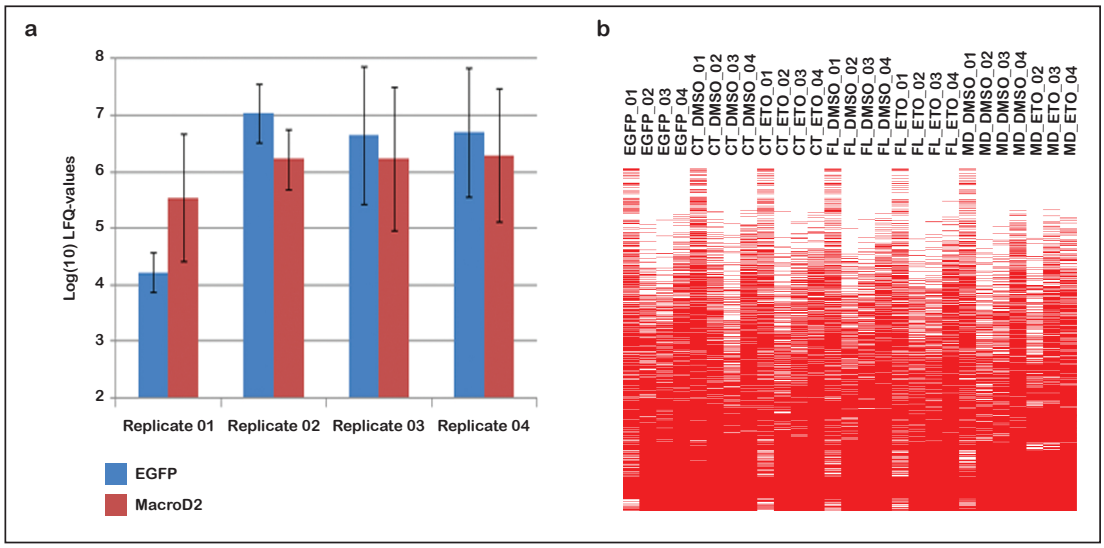


Figure 4.4: Comparison in intensities and missing values between the four biological replicates - a) Intensities of the EGFP and MacroD2 protein signals showed as LFQ (label-free quantitation) value, which are raw intensities normalized on multiple levels to reflect the relative amounts of the proteins (380); b) Missing values (red) of each protein (row) in every condition (column).

We thus proceeded in the analysis of the last three biological replicates. The signal intensity of all proteins needed to be corrected, as the overall amount of detected protein might differ between one sample and the other due to technical error generated upon the co-IP step or the mass spectrometry run.

First of all, we chose to work with the intensity-based absolute quantification (iBAQ) values: for each protein in a specific sample, the iBAQ signal is the sum of all the peptides intensities divided by the number of the peptides of the protein that could be observed. Therefore, the iBAQ represents a weighted intensity, which introduce a measure of reliability of this signal in the quantification. Then, to even out the technical error due to sample processing, the iBAQ value of every protein in each sample was normalized according to a coefficient defined per each sample. To obtain

the sample-specific coefficient, Dr. Schmidt identified the proteins that are present in most samples. He then calculated the median iBAQ of these proteins for each sample and each median was normalized against the highest median value among all samples.

Due to the very nature of the mass spectrometry technique, where random peptides are sampled, the normalized iBAQ signals showed a number of missing values as expected. Missing values are always generated, since peptides might be not present, not detected or not correctly identified. In order to avoid considering proteins with too many missing values in the statistical tests, which might generate additional noise, all the proteins that were not present in all three biological replicates for at least one condition were filtered out. For example, protein X is present in the three replicates of EGFP condition but nowhere else, thus protein X is retained in the analysis; protein Y is present only in two replicates of the EGFP condition, but nowhere else, thus protein Y is excluded from further analyses. This step was important in removing from the analysis noise generated from stochastic sampling of the mass spectrometry run.

Once the proteins were filtered, in order to determine the similarities of the different conditions across the biological replicates, I performed a hierarchical clustering test as an additional quality control. The clustering algorithm that I adopted performed a single-linkage approach: with this algorithm, the clustering proceeds by reiteratively combining two groups that contains the closest pair of elements, still not close enough to be part of the same group. This method is also called the *nearest neighbor clustering*. In the ideal case, all the samples of one condition in different replicates should cluster together. However, the test on my data showed differences between different biological replicates **(Figure 4.5)**.

Most strikingly, all the samples from the fourth biological replicate clustered together. This feature is shared also using other clustering algorithm outputs **(data not shown)**. Such a behavior indicates that the fourth biological replicate is in same ways different from the other two, since each condition of this replicate is more similar to different conditions of the same replicate than to the same condition in the other replicates. On the other hand, the other two biological replicates mostly cluster according to the condition. The only exception is the C-terminus-associated samples, which are located quite far away from each other.

In order to totally eliminate missing values from further analyses, after having filtered the proteins I imputed the iBAQ signal in place of the missing values by mean

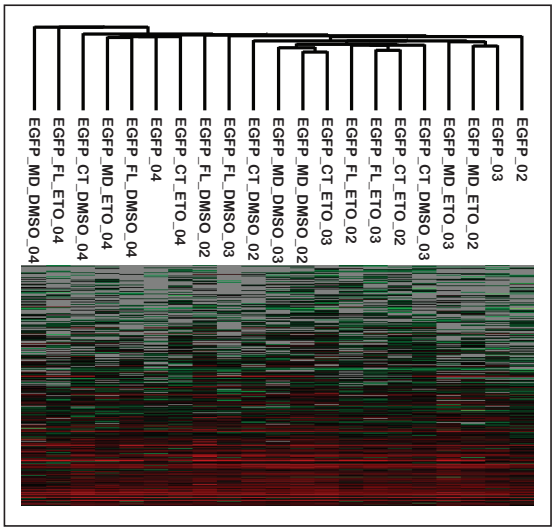


Figure 4.5: Hierarchical clustering of the different co-IP samples - Hierarchical clustering performed with single-linkage algorithm on Perseus 1.5.3.2 software.

of a normal distribution. The imputation step, thus, substitutes the null values with values that are similar to background signal so that they do not generate an important bias in the analysis, but yet can be used to perform the enrichment tests.

4.5 Enrichments and generation of protein lists

In the previous steps, the iBAQ signals have been corrected and filtered. At this point they are ready to be used for the generation of the protein lists for my six conditions of interest:

1. “MacroD2 full-length DMSO”
2. “MacroD2 full-length etoposide”
3. “MacroD2 macrodomain DMSO”
4. “MacroD2 macrodomain etoposide”
5. “MacroD2 C-terminus DMSO”
6. “MacroD2 C-terminus etoposide”

The following steps were performed in order to determine the list of significantly enriched proteins for each condition. To this end, I first compared the protein intensities of all the samples, two by two. In each comparison, I performed Student's t-tests in order to identify the proteins that are specifically enriched in one of two samples. Each t-test was performed by using 0.1 as value for the S0 coefficient and 0.08 for the false discovery rate (FDR) on the Perseus software. The FDR value of 0.08 is more lenient than the default 0.05, which was too stringent for some enrichments. For example, by comparing *MacroD2 full-length DMSO* and *MacroD2 full-length etoposide* protein intensities, I could define the $10 - 20$ protein enriched in the first sample and the similar number of proteins enriched in the second **(Figure 4.6)**. The comparison between two samples with a T-test and the generation of specific protein lists is referred as enrichment.

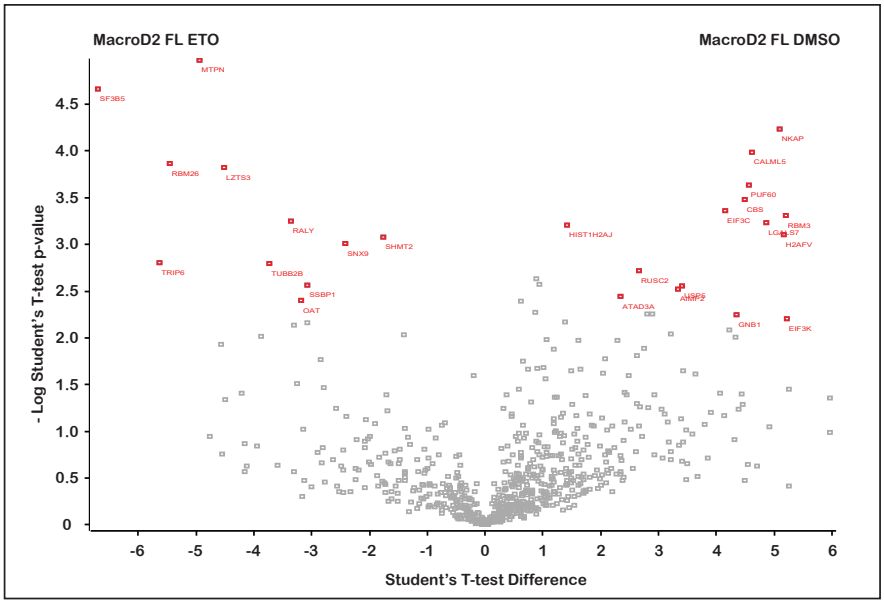


Figure 4.6: An example of enrichment between two different samples - Graphic representation of the Student's t-test between two samples, in this example MacroD2 full-length DMSO versus MacroD2 full-length etoposide. The t-test was performed with 0.1 as S0, which defines the importance of the pvalue over the difference in enrichment, and 0.08 as FDR. In red, the genes that are significantly enriched in the two samples. On the right part of the graph, the genes are enriched in the MacroD2 full-length DMSO sample, on the left side the genes are more present in the MacroD2 full-length etoposide sample. In gray are all the genes that are not significantly enriched in either of the two samples. The graph was generated with Perseus 1.5.3.2 software.

Per each condition that I wanted to define, such as "MacroD2 full-length DMSO" or "MacroD2 C-terminus etoposide", I perfomed an enrichment of the omonimous sample against each of those samples that could be considered negative controls: the negative control sample should present proteins that are not enriched in the main sample, and viceversa. Therefore, for each condition, there are as many protein enrichments as negative controls **(Table 4.1)**. For example, for the "MacroD2 full-length DMSO" condition, I performed one enrichment between the *MacroD2 full-length DMSO* and the *EGFP* sample and another enrichment between the the *MacroD2 full-length DMSO* and the *MacroD2 full-length etoposide* sample. Or, in case of the "MacroD2 C-terminus etoposide" condition, the negative controls were *EGFP*, *MacroD2 C-terminus DMSO*, *MacroD2 macrodomain DMSO* and *MacroD2 macrodomain etoposide*, and thus for this condition four protein enrichments were generated.

Condition	**Negative control**
MacroD2 full-length DMSO	EGFP
	MacroD2 full-length etoposide
MacroD2 full-length etoposide	EGFP
	MacroD2 full-length DMSO
MacroD2 macrodomain DMSO	EGFP
	MacroD2 macrodomain etoposide
	MacroD2 C-terminus DMSO
	MacroD2 C-terminus etoposide
MacroD2 macrodomain etoposide	EGFP
	MacroD2 macrodomain DMSO
	MacroD2 C-terminus DMSO
	MacroD2 C-terminus etoposide
MacroD2 C-terminus DMSO	EGFP
	MacroD2 C-terminus etoposide
	MacroD2 macrodomain DMSO
	MacroD2 macrodomain etoposide
MacroD2 C-terminus etoposide	EGFP
	MacroD2 C-terminus DMSO
	MacroD2 macrodomain DMSO
	MacroD2 macrodomain etoposide

Table 4.1: List of enrichments - Overview of the different enrichments, divided according to the condition and showing the relative negative controls.

Once created these protein enrichments, I compared all the enrichments of the same condition and defined which proteins were present in at least two out of four enrichments (or in case of the full-length conditions, in both enrichments). These recurring proteins are considered *common proteins* and represent the most reliable interactors of MacroD2 in that specific condition. However, in a few cases the list of common proteins was extremely short, which hampered the following analyses. In addition, sometimes different components of a same complex were present in two different enrichments of the same condition. Such a phenomenon also hinders the finding of relevant biological terms. For this reason, I also looked in the combined list of the enriched proteins (*all proteins*).

The protein lists generated from the above analysis, both *common proteins* and *all proteins*, are shown in Appendix A. For each list, I checked the annotations of every entry on Uniprot. This manual revision was important to define the presence of specific proteins that could be easily connected to MacroD2 regulation by nuclear transport, such as kinases, exportins or chaperones. However, neither kinases or exportins were present in any of the protein lists.

Instead, in order to have a clearer understanding about the overall protein network in each condition and to formulate hypotheses on the cellular function of MacroD2, I analyzed the lists by looking for biological patterns. To this end, each list was analyzed by means of the online database STRING and its associated software. This software scrutinizes the protein lists and classifies the proteins according to the biological tags that each protein has. Consequently, STRING finds the enriched biological terms for functional processes, pathways or cellular localizations within a group of proteins. Each biological term in the list is associated to a FDR value, which represents the likelihood that this enriched term is false: the lower the FDR is, the more likely this term is genuinely enriched in the list of terms. In my analysis only biological terms with FDR lower than 0.05 are taken into account. The enriched biological terms lists are shown in Appendix B.

4.6 General overview of MacroD2 interactors and enriched biological terms

4.6.1 MacroD2 full-length treated with DMSO

The complete list of interactors of "MacroD2 full-length DMSO" is shown in Appendix A **(Table A.1)**. For the MacroD2 full-length conditions, I considered as negative controls the *EGFP* and the MacroD2 full-length protein with the opposite compound treatment, in this case *MacroD2 full-length etoposide* sample.

When checking the common proteins to the two enrichments, only one protein remained. This was EIF3K, an initiation factor for protein translation **(Table A.2)**. This means that detection of enriched biological terms was impossible.

Therefore, to have a glimpse of what might be the overall network of the MacroD2 protein in the DMSO condition, I analyzed all the proteins present in the three lists. When analyzing all the proteins combined, the enriched biological terms showed clearly the strong contribution of RNA-binding proteins within the MacroD2 interactome **(see Table B.1)**. These biological terms reflect the presence of both components of the ribosomal complex and proteins involved in RNA splicing.

Indeed, recently the role of ADP-ribosylation in the regulation of RNA metabolism has become clearer. Proteomic studies of PAR-binding proteins or PARylated proteins showed enrichments in RNA biology-related classification terms (10, 118, 118, 206). In addition, several studies showed effects of ARTDs in all the key steps of RNA metabolism, from transcription to translation, but also in stress-dependent alternative splicing or RNA-interference regulation (10). Thus, the presence of factors necessary for RNA splicing among the MacroD2 interactors may indicate a direct involvement of MacroD2 in this new branch of ADP-ribosylation-mediated regulation.

4.6.2 MacroD2 full-length treated with etoposide

The complete list of interactors of "MacroD2 full-length etoposide" is shown in Appendix A **(Table A.3)**. The enrichments were performed against *EGFP* and *EGFP-MacroD2 full-length DMSO*. The number of proteins that are common to the two enrichments (*common proteins*) was also small **(Table A.4)**. These proteins fail to show any significant enrichment in the biological terms.

When I combined all the proteins belonging to this condition, the only significant biological terms all pertained to the binding of RNA molecules **(see Table B.2)**. This result is in line with the enriched biological terms of the MacroD2 full-length DMSO condition, however it does not address the possible role of MacroD2 upon genotoxic stress.

4.6.3 MacroD2 macrodomain treated with DMSO

The complete list of interactors of "MacroD2 macrodomain DMSO" is shown in Appendix A **(Table A.5)**. As negative controls, I considered *EGFP*, *MacroD2 macrodomain etoposide*, *MacroD2 C-terminus DMSO* and *MacroD2 C-terminus etoposide* samples.

The "MacroD2 macrodomain DMSO" was the condition with the highest number of interactors **(Table A.6)**. Consequently, the number of filtered *common proteins* was 54. The annotated interactions show an intricate network **(Figure 4.7)**. Some members are involved in RNA metabolism and translation. Other interactions, instead, involve proteins that have a role in transport across the different compartments, such as 14-3-3ζ (YWAHZ) and Ran-GTPase (RAN).

The presence of the 14-3-3ζ in this enrichment was unexpected. I tested the 14-3-3s as possible candidates for MacroD2-regulated nuclear export, since the 14-3-3ϵ protein appeared in the first biological replicate of the co-immunopurification experiment **(see Section 3.8)**. My experiments, however, did not show any interaction. Although I have never tested the binding of the 14-3-3ζ with full-length MacroD2, the fact that the protein is enriched with the macrodomain fragment indicates that, even when it binds, the 14-3-3ζ is not likely to be involved in the regulated nuclear export of MacroD2.

Likewise, the C-terminus-dependent MacroD2 nuclear export would suggest that a possible interaction with the Ran-GTPase should occur either with the full-length protein or the C-terminus fragment. Instead, the Ran-GTPase is immunoprecipitated by the MacroD2 macrodomain. However, this result is consistent with the fact that Ran-GTPase is MARylated by bacterial toxins as well as by ARTD10 *in vitro* (381). Therefore, this interaction represents another indirect validation for the co-immunopurification approach.

When checking the enriched biological terms for the list of common interactors in the MacroD2 macrodomain DMSO condition **(see Table B.3)**, the presence of many membrane-associated terms, among them the endoplasmic reticulum (ER), can

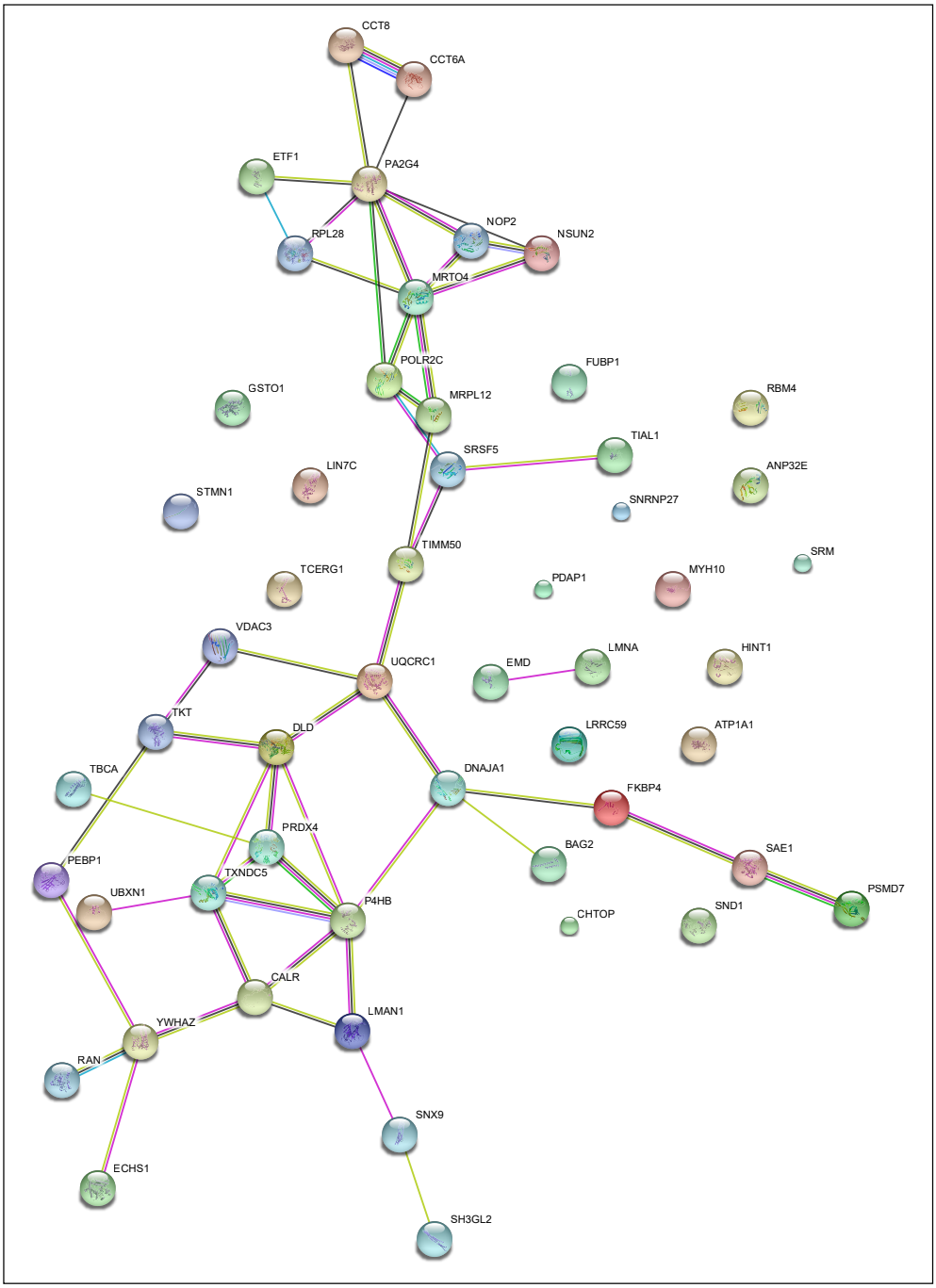


Figure 4.7: Annotated interactions within the common proteins present in the four EGFP-MacroD2 macrodomain DMSO enrichments - Annotated interactions shown with the help of STRING software. Edges of different colors according to the type of interaction and the experimental proof: *cyan*) known interaction from curated database; *fucsia*) known interaction experimentally determined; *green*) gene neighborhood; *red*) gene fusions; *blue*) gene co-occurrence; *yellow*) text-mining; *black*) co-expression; *light blue*) protein homology.

be connected to the role of MARylation in the regulation of the Unfolded Protein Response by ARTD15/PARP16 **(see Section 1.2.7)**.

Moreover, the "intracellular steroid hormone receptor signaling pathway" term suggests that MacroD2 might be involved in the signaling pathway responsive to steroid hormones. Indeed, MacroD2 itself was shown to interact with the estrogen-receptor α (ERα) in a functional way (98). Even before MacroD2, MacroD1 was connected to both the ERα and the androgen receptor as well (77, 78), while ARTD1 regulates the transcription of estrogene-responsive genes (382). Thus, it is possible to imagine a general role of ADP-riboylation in the steroid hormones-dependent signaling pathway.

I then performed the biological term analysis on the list of *all proteins* of the four enrichments **(Table B.4)**. When all the proteins are analyzed, the enriched biological term "intracellular steroid hormone receptor signaling pathway" disappeared, having lost its significance. However, other biological terms were introduced. For example, some additional biological terms refer to carbon metabolism and energy production, like "glycolysis/gluconeogenesis", "citrate cycle", "pyruvate metabolism" and even some amino acid biosynthetic pathways.

ADP-ribosylation in general, and MacroD2 in particular, might easily have a regulatory role on the cellular energy production and metabolism in general. ADP-ribosylation requires NAD^+ as donor of the ADP-ribose and consumption of this reagent can indirectly affect the activation level of glycolysis versus the citrate cycle. On the other hand, considering a potential role of sirtuins and MacroD1 in the regulation of the respiratory chain activity, it is possible that the regulation of the cytoplasmic enzymes falls upon MacroD2. However, as long as we do not have a complete list of MARylated proteins in the cell, it is difficult to formulate more precise hypotheses.

4.6.4 MacroD2 macrodomain treated with etoposide

The complete list of interactors of "MacroD2 macrodomain etoposide" is shown in Appendix A **(Table A.7)**. The "MacroD2 macrodomain etoposide" condition includes much fewer proteins compared to the DMSO condition. Therefore, the common proteins are only 16 **(Table A.8)**. ARTD1/PARP1 is present among these selected proteins. As ARTD1 is one of the few proteins whose interaction with MacroD2 has been published

(56, 67), its presence in the "MacroD2 macrodomain etoposide" condition provides validity of the whole experimental setup.

Moreover, the Ras GTPase-activating protein-binding protein 2 (G3BP2) is a factor with a RNA-binding domain, thus generally involved in the mRNA nuclear export, and also appears in the list. Interestingly, G3BP2, together with G3BP1, has been shown to be able to induce ADP-ribosylation-based cytoplasmic stress granules (212, 383). When I then analyzed this list of 16 proteins on STRING, the only enriched biological terms are related to RNA biology **(Table B.5)**. The short list of terms is probably determined by the short list of common interactors.

The complete list of proteins generated a longer list of enriched biological terms **(see Table B.6)**. A group of biological terms indicates an enrichment in factors involved in the focal adhesion foci and adherens junctions formation, which might indicate a possible function of MacroD2 in the regulation of fibril based structures.

Similarly, the neural progenitors-specific chromatin remodeling complex (npBAF complex) and the neuron-specific chromatin remodeling complex (nBAF complex) are other interesting entries of this list of biological terms. Although the presence of such terms is quite surprising if we consider that HEK293 cells derive from kidney tissues, these cells are indeed cancer cell lines with unstable genomes. Thus, the malignant nature of these cells might have allowed the unexpected expression of neuron-specific genes.

4.6.5 MacroD2 C-terminus treated with DMSO

The list of interactors for "MacroD2 C-terminus DMSO" is shown in Appendix A **(Table A.9)**. There are only 10 proteins, when the four enrichments of "MacroD2 C-terminus DMSO" condition are combined **(Table A.10)**. Four out of these ten proteins are involved in mRNA processing. In fact, the biological terms associated to this list are strictly related to RNA biology **(Table B.7)**.

The biological terms derived from the all the protein from the four enrichments combined involve mainly nuclear and nucleolar compartments **(see Table B.8)**. At the top of the list there is an enrichment of "RNA binding" terms. Other strongly represented terms in this condition are the "mRNA processing" and "spliceosome" keywords. While in other conditions the presence of ribosomal proteins might have been the major component of the RNA-binding interactors, most MacroD2 C-terminus

interactors that bind RNA molecules may be mainly involved in RNA processing, such as mRNA splicing and export from the nucleus, with a preferential localization within the nuclear envelope.

4.6.6 MacroD2 C-terminus treated with etoposide

Last but not least, I performed the analysis for the "MacroD2 C-terminus etoposide" condition. The complete list of proteins enriched in *MacroD2 C-terminus etoposide* sample over the four negative controls is shown in Appendix A **(Table A.11)**. There are 29 proteins in common to the four enrichment lists for this condition **(Table A.12 and Figure 4.8)**.

Among the annotated interactions, 3 proteins are connected to the DNA damage response: p53 (TP53), RPA2 and SSBP1. The p53 protein is an important factor of the DDR, since upon activation mediated by ATM and/or ATR, it blocks the progression of the cell cycle and decides whether the cell survives and undergoes repair or undergoes programmed death (258). In addition, RPA2 is one of the factors involved in the recruitment and activation of the heterodimer ATR-ATRIP (278). SSBP1 is a mitochondrial protein, involved in the mitochondrial DNA replication and repair, although localization of MacroD2 within the mitochondria is mostly excluded (384). The enriched biological terms for the common list prevalently regard extracellular vesicles, as cellular waste or functional secretion, and RNA binding **(see Table B.9)**.

The analysis of the whole list of proteins, instead, greatly increases the range of biological terms, many related to functional processes **(see Table B.10)**. Compared to the C-terminus DMSO condition, the localization of the factors is much more heterogeneous, with proteins located in compartments other than the nucleus, including the cytosol, the mitochondria, the ER and other vesicle-based structures. The range of functional processes is also widened, now including "mismatch repair", "DNA replication", "contractile fiber", "protein regulation of protein folding", "regulation of helicase activity" and "carbon metabolism".

It seems, thus, that the etoposide treatment has radically changed the types of interactions that the same fragment is able to have. The nature of the change might indeed be linked to the number of post-translational modifications that the C-terminus undergoes upon genotoxic stress. This modification might augment the possible contacts

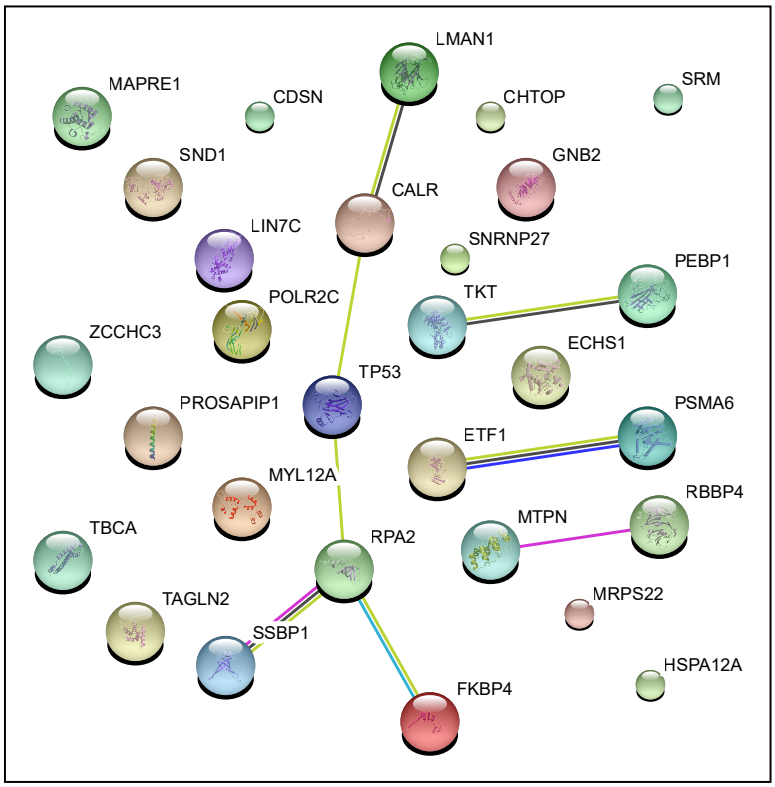


Figure 4.8: Annotated interactions within the common proteins present in the four EGFP-MacroD2 C-terminus etoposide enrichments - Annotated interactions shown with the help of STRING software. Edges of different colors according to the type of interaction and the experimental proof: *cyan*) known interaction from curated database; *fucsia*) known interaction experimentally determined; *green*) gene neighborhood; *red*) gene fusions; *blue*) gene co-occurrence; *yellow*) text-mining; *black*) co-expression; *light blue*) protein homology.

that the C-terminus fragment can perform and possibly expand the cellular localizations that the fragment can adopt.

4.7 Discussion

This co-immunopurification (co-IP) approach was the first attempt to uncover the interactome of the MacroD2 full-length protein in normal physiological conditions and under genotoxic stress. The goal of the experiment was to generate candidates that drive MacroD2 into the cytoplasm with the regulated nuclear export. Moreover, by uncovering the MacroD2 interactome it was possible to suggest testable hypotheses about the functional role of MacroD2 in the cell in order to fill the gap between the molecular activity of MacroD2 and the genome-wide association studies that concern it **(see Section 1.2.5)**. To achieve this, we designed the experiment with different constructs in different conditions. In the lists of proteins false positive candidate are further filtered out by cross-comparing the different samples.

On the other hand, the hierarchical clustering revealed problems in the quality of the data, namely that the fourth replicate did not cluster with the others **(see Figure 4.5)**. This is generally solved by performing another independent co-IP experiment in the same condition and repeating the set of analyses in order to verify that the enriched proteins and enriched biological terms do not change. However, since my host lab is setting up other unbiased approaches to define MacroD2's functional role, including a BioID-based pull-down coupled with peptide mass fingerprinting (385), we chose to use my dataset as a complementary approach to define the recurrent interactors and functional target of MacroD2, instead of investing more time and resources on a new replicate.

The experiment that I have presented has some limitations. First of all, MacroD2 might engage a vast array of transient interactions. The classical co-IP is not suitable to detect interactions with labile affinity and this is possibly why I could not find any kinases in my enrichments. To this end, the BioID approach, provided it is properly controlled against false positives, may identify these transient interactions. Secondly, the experiment could have been further improved by also using a cell line expressing MacroD2 macrodomain ADP-ribose binding-deficient (G188E) mutant. This could

have helped define the pool of interactors that are actual MacroD2 substrates by discriminating the proteins able to interact with the overall structure of the macrodomain, from those that bind MacroD2 by means of ADP-ribosylation. However, this is a minor limitation, since this can be easily tested during validations of the candidates.

The presence of ARTD1/PARP1 as an interactor in the "MacroD2 macrodomain etoposide" condition acts as validation of the experimental setup. MacroD2 macrodomain has been shown to interact with ARTD1 *in vitro* and *in vivo*, most likely mediated by ADP-ribosylation (56, 67). Published literature also shows that etoposide treatment can induce the activation of ARTD1 (386). Even if etoposide is not as good an inducer of ADP-ribosylation as oxidative stress is, the treatment can still activate ARTD1 slightly more than in the untreated state, allowing an enrichment of ARTD1 upon etoposide treatment.

4.7.1 MacroD2 full-length protein has less interactors than its two fragments

When comparing the number of interactors in the different conditions, the fragment showing the largest number of interactors upon DMSO treatment was the macrodomain; conversely, upon etoposide, the best binder was the C-terminus fragment **(Figure 4.9)**.

MacroD2 full-length showed less interactors than the two shorter fragments in both treatments. Considering how the experiment was designed, the pool of proteins binding MacroD2 full-length should include both the interactors of the macrodomain and those of the C-terminus. Therefore, I would have expected more binders for the MacroD2 full-length, both as an absolute number and in comparison to the two fragments. I also performed all the enrichments of the six conditions against EGFP with much more lenient parameters (0.05 as S0 and 0.5 as FDR), in order to be more inclusive **(data not shown)**. The result was a dramatic increase of protein shared by the macrodomain and the C-terminus fragment. However, this probably represents increased noise in the analysis.

The reason why MacroD2 full-length is a poor binder could be explained either by technical error or by a proper physiological feature of the protein. In the first case, the poor quality of the hierarchical clustering shows a problem in reproducibility across the different biological replicates, which surely impacts on the statistical tests that I perform for the enrichment. It might be that the full-length samples suffer

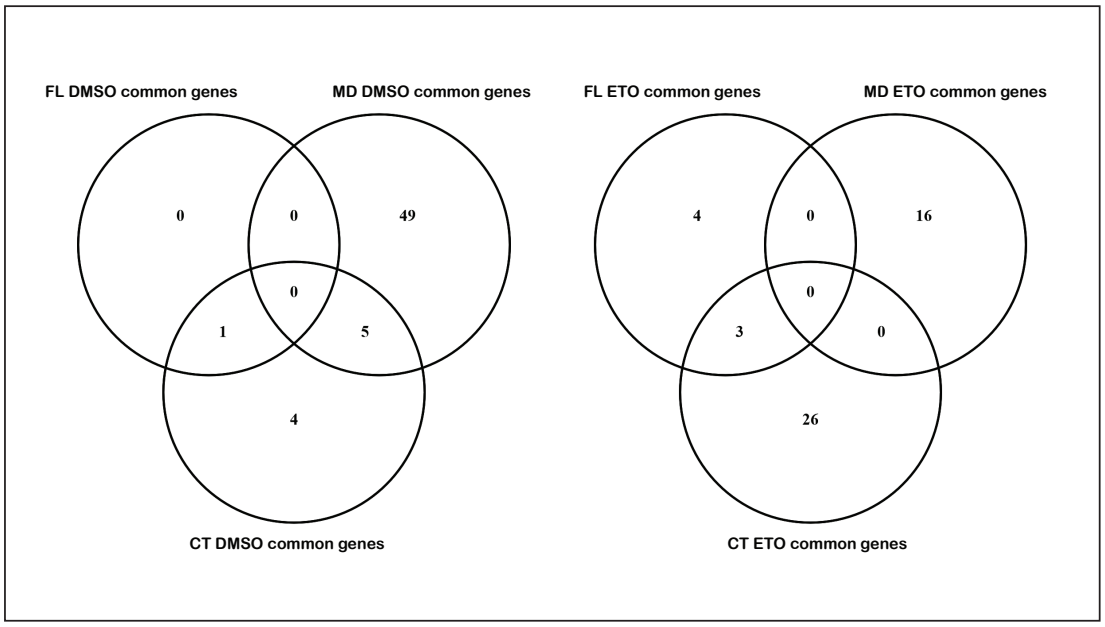


Figure 4.9: Overview of the most reliable binders for the six different conditions - Venn diagram of the most reliable binders for the six conditions. The lists are generated by the overlap of different enrichment of the same condition against two or four negative controls, defined thus as common proteins. FL is full-length; MD is macrodomain; CT is C-terminus. Generated with Venny 2.1.0.

more than the others in the issue of reproducibility. On the other hand, the two MacroD2 domains, the macrodomain and the unstructured C-terminus, might hinder the binding of reciprocal interactors when they are together in the same polypeptide. This hypothesis argues for an even stronger role in the crosstalk of the two parts of MacroD2 in regulating the protein function.

In addition, almost 50 % of genes selected for the "MacroD2 C-terminus DMSO" condition were shared with the "MacroD2 macrodomain DMSO". This overlap was not expected. However, considering that the MacroD2 macrodomain DMSO condition presents the biggest number of proteins after the filtering, I can think of two explanations: either the higher number of enriched proteins simply generated some background, which was not corrected upon normalization, or there is a biological reason. It might be that these interactors are part of complexes that might bind both the MacroD2 macrodomain and the C-terminus fragment, although I do not know of any published cases in the current literature.

4.7.2 Filtering the protein shared in the macrodomain and C-terminus conditions

Subsequently, I wanted to verify how many proteins are shared between the macrodomain and C-terminus conditions. Considering that the common lists are generated by combining the enrichments against the other three conditions, I would expect that no background gene is selected. Ideally all the final lists of common proteins should not share any protein. However, some proteins are still shared between two or three conditions **(Figure 4.10)**.

"The MacroD2 macrodomain DMSO" is the condition that shares proteins with all the other conditions, the highest commonality with "MacroD2 C-terminus etoposide" condition. Considering that I generated the lists of common proteins by taking only those that were enriched against at least two different conditions, these factors are still likely to be proper interactors of MacroD2, with the shared proteins belonging to large complexes. MacroD2 might interact with these complexes by exploiting both the macrodomain and the C-terminus.

I then checked for the enrichments in biological keywords for the four conditions. For each condition, the list has been filtered of all those factors that are shared with

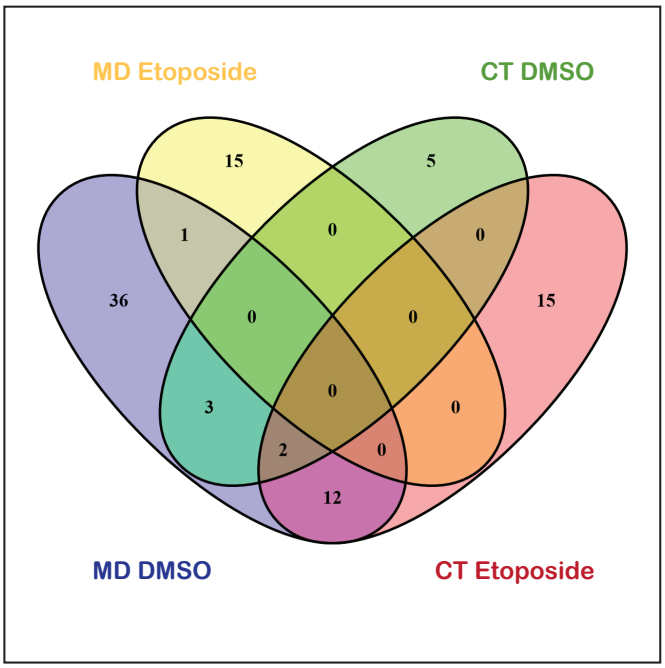


Figure 4.10: Proteins shared in the four conditions - Venn diagram of the most reliable binders for four conditions: MD DMSO, MD etoposide, CT DMSO, CT etoposide. The input lists are the common proteins for each condition. FL is full-length; MD is macrodomain; CT is C-terminus. Generated via Venny 2.1.0.

other conditions. Thus, the biological terms of the interactors specific for that certain condition are shown **(Table 4.2)**.

Category	Term	FDR
MacroD2 macrodomain DMSO		
GO.0043230	extracellular organelle	3.00E-07
GO.0065010	extracellular membrane-bounded organelle	3.00E-07
GO.0070062	extracellular exosome	3.00E-07
GO.1903561	extracellular vesicle	3.00E-07
GO.0031988	membrane-bounded vesicle	1.18E-06
GO.0031982	vesicle	1.72E-06
GO.0044444	cytoplasmic part	0.000134
GO.0032991	macromolecular complex	0.000421
GO.0006457	protein folding	0.000528
GO.0044421	extracellular region part	0.000548
GO.0044822	poly(A) RNA binding	0.000712
GO.0003723	RNA binding	0.000878
GO.0044446	intracellular organelle part	0.0011
GO.0031625	ubiquitin protein ligase binding	0.00142
GO.0044389	ubiquitin-like protein ligase binding	0.00142
GO.0044422	organelle part	0.0016
GO.0016023	cytoplasmic membrane-bounded vesicle	0.00281
GO.0042470	melanosome	0.00281
GO.0005576	extracellular region	0.00388
GO.0031410	cytoplasmic vesicle	0.0047
GO.0070013	intracellular organelle lumen	0.0047
GO.0005829	cytosol	0.00474
GO.0043233	organelle lumen	0.00515
GO.0044428	nuclear part	0.00516
GO.0031974	membrane-enclosed lumen	0.00555
GO.0005634	nucleus	0.00603
GO.0005737	cytoplasm	0.0061
GO.0044297	cell body	0.0061
GO.0005832	chaperonin-containing T-complex	0.00665
GO.0002199	zona pellucida receptor complex	0.00798
GO.0012505	endomembrane system	0.00857
GO.0043209	myelin sheath	0.00857
GO.0019953	sexual reproduction	0.0111
GO.0031967	organelle envelope	0.0116
GO.0031975	envelope	0.0116

Continues on next page

Table 4.2 – *Continued from previous page*

Category	Term	FDR
GO.0005515	protein binding	0.0118
GO.0045454	cell redox homeostasis	0.0172
GO.0005635	nuclear envelope	0.0199
GO.0061024	membrane organization	0.021
GO.0044703	multi-organism reproductive process	0.0215
GO.0043234	protein complex	0.0241
GO.0000003	reproduction	0.0243
GO.1990204	oxidoreductase complex	0.0266
GO.0043228	non-membrane-bounded organelle	0.0274
GO.0043232	intracellular non-membrane-bounded organelle	0.0274
GO.0043231	intracellular membrane-bounded organelle	0.0315
GO.1903047	mitotic cell cycle process	0.0329
GO.0005488	binding	0.0343
GO.0005739	mitochondrion	0.0365
GO.0042995	cell projection	0.0376
GO.0015630	microtubule cytoskeleton	0.0381
GO.0022402	cell cycle process	0.0414
GO.0044702	single organism reproductive process	0.0414
GO.0035770	ribonucleoprotein granule	0.0459
GO.0036464	cytoplasmic ribonucleoprotein granule	0.0459
GO.0097223	sperm part	0.0459
KEGG:04141	protein processing in endoplasmic reticulum	0.047
MacroD2 macrodomain etoposide		
GO.0003723	RNA binding	0.0117
GO.0044822	poly(A) RNA binding	0.0128
MacroD2 C-terminus DMSO		
GO.0022618	ribonucleoprotein complex assembly	0.0302
GO.0071826	ribonucleoprotein complex subunit organization	0.0302
MacroD2 C-terminus etoposide		
KEGG:03430	mismatch repair	0.025
KEGG:03440	homologous recombination	0.025
KEGG:03030	DNA replication	0.0282

Continues on next page

Table 4.2 – *Continued from previous page*

Category	Term	FDR

Table 4.2: Biological term enrichment for the core proteins of the four conditions - Overview of enriched biological terms. Generated by STRING database

The biological terms for the "MacroD2 macrodomain DMSO" condition reveal a very heterogeneous localization, with many references to the ER. The "regulation of steroid hormone signaling pathway" has instead disappeared, probably as some filtered out factors were contributing to the significance of that pathway in the enrichment. The "MacroD2 macrodomain etoposide" condition has now only terms related to RNA binding, while the " MacroD2 C-terminus DMSO" is focused on ribonucleoprotein complexes. These ribonuclear complexes include both ribosomes and particles belonging to the spliceosome. Lastly, the "MacroD2 C-terminus etoposide" lost most of its enriched biological terms, surprisingly retaining only those referring to "DNA replication" and "mismatch repair". "Homologous recombination" was also a surprising new entry in the list.

4.7.3 Looking for the exporting interactor

Defining the interactors that are important for the nuclear export of MacroD2 upon etoposide treatment was the main motivation for performing the co-immunopurification experiment. One possibility is the interaction with a protein that present a NES, so that MacroD2 is exported via a "piggy-back" mechanism. Such an interactor should be able to bind the two MacroD2 phosphorylated serines, which are within SQ/TQ motifs modified by ATM or a PI3K-like kinase upon DNA damage. Moreover, the candidate should optimally be found among the interactors that are enriched in "MacroD2 full-length etoposide" or "MacroD2 C-terminus etoposide" conditions.

It is possible to define which phospho-binding module will more likely recognize the modification performed by a certain kinase by considering the sequence requirements of the kinases and of the phospho-binders (357). The serine or threonine modified by ATM and the PI3K-like kinases are most probably recognized by FHA and BRCT binding modules. For this reason, I downloaded the list of human proteins possessing either a FHA or a BRCT domain from Uniprot and I systematically compared these two lists

with the one I have produced by analyzing the co-IP experimental data, looking for possible factors that might be involved in the recognition of the modification.

Although 71 FHA-containing proteins and 44 BRCT-containing proteins are annotated in Uniprot, out of the hundreds of proteins present in my pull-down experiment, only ARTD1 and SMARCC1 present a BRCT domain. Both the proteins are present in the list of common proteins of the MacroD2 macrodomain etoposide condition. Considering that the factor I am looking for should bind the C-terminus of the protein, I would exclude their role in the MacroD2 assisted nuclear export.

4.8 Future directions

After having performed the experiment and completed its analysis, the next step for proteomics experiments is the validation. This is normally performed by *in vitro* and *in vivo* pull-downs, co-localization tests by immunofluorescence or more refined experiments, such as BiMolecular Fluorescence Complementation, Proximity Ligation Assay or FRET. However, after the analyses it is vital to search through the literature and define which are the most sensible candidates, both as protein and as physiological functions, to pursue. What follows are the most interesting hypotheses I formulated about possible MacroD2 cellular functions and related follow-up experiments.

4.8.1 Screening for factors affecting MacroD2 nuclear export

As mentioned above, one of the reasons why I performed the co-IP coupled to peptide mass fingerprinting upon etoposide treatment was to discover the interactors that may lead MacroD2 across the nuclear envelope into the cytoplasm. When I compared the list of MacroD2 interactors with the list of proteins possessing FHA or BRCT domains, the only proteins present in both lists are ARTD1 and SMARCC1. However, both proteins are highly enriched in the "MacroD2 macrodomain etoposide" condition. Also, both proteins are described as nuclear proteins, thus it seems difficult to imagine how these two factors may be involved. Considering instead the great number of proteins that I compared, it is feasible to find two false positive elements in my search. Nonetheless, their functional involvement in MacroD2 nuclear export is easily testable. Using an RNAi-mediated approach to knock-down ARTD1 or SMARCC1 and subsequently

performing live-cell imaging experiment with EGFP-MacroD2 C-terminus upon UV-laser microirradiation should be sufficient to define the actual involvement of these two proteins to the nuclear export of MacroD2.

The discovery of the exporting interactor may have been hindered by the limitations of the co-IP approach, as it might be not sufficient for the detection of transient interactions or interactors with small affinities. Therefore, the BioID-based mass spectrometry-coupled pull-down experiment that my host lab is planning to perform could greatly increase the pool of proteins that are interacting with MacroD2. The technique is based on the generation of a hybrid construct with the protein of interest attached to a promiscuous biotin ligase, that is able to mark with biotin the proteins that get in proximity to the bait (385). Indeed, the main drawback of the technique is the possible modification of too many proteins generating a great deal of false positives. Therefore, the experiment requires good negative controls. For example, in the case of MacroD2, the ADP-ribose-binding deficient mutant (G188E) and the export-deficient mutant (S276,345,415,426A) could be useful to screen the interactor list and to find proteins that are strictly related to these two protein features of MacroD2. This approach might hopefully be able to enlarge the list of interactors that I have shown in a constructive manner.

Lastly, in order to discover the interactors that allow MacroD2 nuclear export, another strategy would take advantage of the well-known phenotype of MacroD2: the nuclear export itself. By employing libraries of siRNAs, it could be possible to screen for those factors whose absence negatively affects MacroD2 nuclear export. This experiment would be then based on RNAi-mediated strategy coupled to cell-imaging on U2OS-MacroD2 C-terminus + mCherry-H2B stable cell line. Essential genes, whose lack might impair the cellular life, should not be included in the siRNA library. Most of the membrane proteins should also be discarded, since our interactor is most probably soluble and itself able to shuttle between the nuclear and cytoplasmic compartments. Moreover, the library could be mainly targeting proteins carrying a phospho-binding domain. Consequently, these requirements should greatly decrease the number of genes of interest, making the library generation and the experiment itself more feasible. An alternative to the classical siRNA library could be instead a genome-wide CRISPR/Cas-mediated knock-out strategy (387). By generating the all possible knock-out mutants

from a specific cell line, this approach would definitely solve the interferences generated by the partial depletion of target proteins that the siRNA strategy produces.

Considering the microscopy-related aspects, such a high-throughput experiment excludes the adoption of a live-cell imaging strategy. However, it is advisable to perform not just an end-point assay, since the kinetics of the export might be affected more in the initial phase and somehow recovered in the late stage, when maybe the activation of the kinase decreases. This is even more true when multiple factors redundantly cooperate for the nuclear export of MacroD2. Lastly, if the knock-out strategy would solve the problem of the incomplete depletion, the siRNA strategy instead will only strongly decrease the expression of the candidate factor. Therefore, an imaging assay with four time points (0, 5, 30 and 60 minutes) should be enough to indicate perturbations in different phases of the nuclear export of MacroD2 and avoid false negative cases even with the siRNA strategy approach.

4.8.2 RNA biology regulation

By checking the most common biological terms found in my co-IP experiment, RNA-related keywords appeared for every condition analyzed. While RNA-binding proteins are generally regarded as particularly "sticky" in pull-down experiments (388), ADP-ribosylation has been associated to many processes where RNA molecules are involved **(see Section 1.2.7)**.

As mentioned above, factors involved in both constitutive and alternative splicing contain ADP-ribose binding modules, and ADP-ribosylation of splicing factors can impact the maturation of the transcripts (118, 206, 389). Indeed, both constitutive and alternative splicing factors have been found in my enrichment lists. Moreover, among the enriched proteins from my co-IP experiments, many ribosomal proteins appeared. These proteins are also classified as RNA-binding factors. Translation itself can be modulated by ADP-ribosylation, as shown by the physiological mono-ADP-ribosylation of eukaryotic elongation factor-2 (eEF-2) (209, 210, 211). Therefore, the question is whether these interactions are just due to sample preparation or whether MacroD2 has a physiological role in the regulation of RNA biology.

The validation of the interaction is essential and requires the additional control with the addition of RNAse, as RNA-binding proteins tend to be considered as contaminant in pull-down experiments. Such a control would exclude the possibility that large

complexes containing RNA molecules have been indeed purified and question a direct role of MacroD2 in the regulation of splicing. If however, the binding of MacroD2 and the splicing factors is validated and is indeed mediated by an RNA molecule, it could be interesting to define which fragment of MacroD2 sequence possesses this binding property. Performing *in vitro* pull-down experiments with purified MacroD2 fragments and the candidate splicing factor, in the presence and absence of a RNA molecule will be useful in assessing this hypothesis. Moreover, if MacroD2 binds the splicing factor, it might perform its enzymatic activity on the factor as well: thus, the splicing factor should be tested to see whether it undergoes any MARylation or PARylation, If the factor is modified, the modification might be able to hinder its splicing performance, which could be tested by means of an *in vitro* splicing experiment in cell-extract.

Lastly, PAR is shown to be required for the formation of cytoplasmic stress granules, where RNA-binding proteins accumulate together with mRNA molecules and Argonaute proteins in order to temporally block their translation (212). G3BP2, which was enriched in the "MacroD2 macrodomain etoposide" condition, was shown to be an inducer of the cytoplasmic stress granules (383).

The localization of the MacroD2 can be tested with arsenite treatment, the most classical method to induce the stress-granule formation. If MacroD2 is indeed able to relocate to the stress granules, a screen for MacroD2 substrates within the stress granules could define which proteins might have their ADP-ribosylation reversed by MacroD2. Afterwards, these proteins might be studied for the impact that the ADP-ribosylation has on their structure or their functional activity.

4.8.3 The neuron-specific remodeling complex, nBAF

Among the enriched biological terms for the "MacroD2 macrodomain etoposide" condition, the "SWI/SNF superfamily-type complex", "npBAF complex", "SWI/SNF complex" and "nBAF complex" terms were present. These are all terms referring to the BRG1- and BRM-associated factor (BAF) remodelling complex, an ATP-dependent SWI/SNF-like complex (390). BAF complexes are formed with up to 15 factors, where the ATP-ase subunit is generally either BRG1 or BRM. The composition of the complex changes according to the cell type, creating hundreds of different possible complexes in humans. The only complexes that are referred to in my list are the neural progenitor BAF (npBAF) and neural BAF (nBAF). Two proteins, SMARCC1 and SMARCE1,

are found in different enrichments of the same condition and the two factors are present in both npBAF and nBAF.

npBAF and nBAF complexes are essential for the development of the nervous system, by helping the maintenance or the establishment of the required expression pattern (390). In neural progenitors, npBAF activity maintains the expression of the pluripotency factors and represses the neuron-specific genes in order to preserve the identity of the progenitors. However, upon commitment for the neural differentiation, some factors of npBAF are replaced by nBAF-specific proteins and the switch from npBAF to nBAF complex is tightly regulated (391).

Mutations in the different members of the complexes affect the neural development: loss of npBAF specific factors, for example, impairs the proliferation of the pool of progenitors, leading to microcephaly (392). However, *de novo* mutations on npBAF/nBAF factors are associated with non-familial autism (393, 394). In fact, since autism causes defects in neural circuit formation, authors have suggested that a failure in npBAF or nBAF-mediated regulation of dendritic outgrowth could be one of the causes of this syndrome (395, 396). Therefore, an interaction between np/nBAF complex and MacroD2 could be a starting point to explore the mechanism behind MacroD2's association to autism. While the co-IP has been performed in kidney-derived HEK293 cells, where neural-specific genes are not expected to be expressed, as it is a cancer cell-line, the possible ectopic expression of neural factors might be an explanation for this incongruity.

The association between MacroD2 and npBAF/nBAF is interesting because of the amount of studies that associate MacroD2 to autism and other neural development syndromes, although without offering any mechanistic insight **(see Section 1.2.5)**. If the physical interaction between MacroD2 and the np/nBAF complexes is confirmed, it would be interesting to proceed in this line of research. In fact, I can imagine two scenarios that might link MacroD2 to this chromatin remodeler complex. In the first scenario, the ADP-ribosylation is important for the recruitment of the complexes on the correct promoters. Thus, MacroD2 might interact with the remodelling complex due to their common recruitment to the same region through ADP-ribosylated factors present on the promoters of neural genes. In the second scenario, MacroD2 could have a role in the modulation of the catalytic activity of the npBAF/nBAF complex or even its assembly.

In support of the first scenario, the ATP-ase unit BRG1 was already found associated to ARTD1 at the chromatin, even though in a completely different context (397). ChIP-sequencing experiments could maybe show that BRG1 interacts on the chromatin with MacroD2 as well, provided having an anti-MacroD2 antibody with high specificity. However, the ultimate outcome of these experiments would be to test whether ADP-ribosylation is important for the targeting of the npBAF and nBAF complexes to the correct neural gene promoters.

In support of the second scenario, SMARCC1 was shown to increase the stability of SMARCE1 by impairing its ubiquitination (398). The authors indicated TRIP12 to be the E3-ligase enzyme, which contains a WWE domain. Therefore, ARTDs might regulate SMARCE1 levels by PARylation-dependent degradation. Thus, the levels of SMARCE1 protein could be tested in the presence and the absence of oxidative stress, when ARTD1 should be mostly active. I would also employ a siRNA-mediated approach and look for the increase of expression levels of SMARCE1 in order to test if other ARTDs are involved. Neuronal cell lines, such as differentiated SH-SY5Y or NT2, should be used to be sure that SMARCE1 is properly expressed. If the hypothesis holds true and an ARTD is found to regulate SMARCE1 protein levels, the role of MacroD2 in the system can be tested, by checking the influence that the over-expression or its depletion have in SMARCE1 levels.

Moreover, MARylation may be a modification that regulates the activation level or the structure of the whole complex. By screening for ADP-ribosylated or even MARylated proteins in neuronal cell lines, it may be possible to find proteins whose activity is modulated by such a modification, and, thus, to test possible interactions of MacroD2 with neuronal proteins.

4.8.4 Regulation of the cytoskeletal dynamics

Various types of cytoskeletal components were enriched with all the three MacroD2 constructs, mainly upon etoposide treatment. Two proteins associated to the cytoskeletal structure of the focal adhesion (FA) foci, TRIP6 and SNX9, were found to be enriched in the "MacroD2 full-length etoposide" condition (399, 400) **(see Table A.4)**. "Microtubule cytoskeleton" and "cytoskeleton" biological terms were enriched in the "MacroD2 macrodomain DMSO" condition **(see Table B.4)**. Also more biological terms related to cytoskeletal structures appeared in the "MacroD2 macrodomain

etoposide" condition, such as "focal adhesion", "cell-substrate adherens junction", "cell-substrate junction", "intermediate filament cytoskeleton", "adherens junction" and "anchoring junction" **(see Table B.6)**. Lastly, the enriched biological terms in the "MacroD2 C-terminus etoposide" condition included "sarcomere", "contractile fiber part", "myofibril", "cuticular plate", "I band", "anchoring junction", "spectrin ", "Z disc" and "adherens junction" **(see Table B.10)**. Although the structures defined by these different term are functionally distinct, they all present as common feature the presence of protein filaments, like F-actin, intermediate filaments or tubulin-made microtubules.

FAs and other cytoskeletal elements can be very important for generating morphological structures with special functions. These structures are even more important when set in the context of special cell-types that have a peculiar morphology and function, such as the neurons. Recent literature shows that the interaction between cytoskeleton and neural function is an expanding research field (401). F-actin are, for example, necessary for the correct formation of dendritic spines, the membrane structure forming one part of the synapse (402). Disruption of the morphology brings a disruption in the neural functionality. How everything is ultimately connected with neurological disorders is still an open question.

The above-mentioned association of MacroD2 with autism and other neurological disorders **(see Section 1.2.5)** could also be explained by a dysfunction in cytoskeletal organization. The imperfect formation of synapses may affect the general neural network. In fact, impairment of brain connectivity has been associated also with a locus nearby to the MacroD2 gene (97). In my co-IP experiment, the MacroD2 constructs purified cytoskeletal factors above all upon etoposide treatment, arguing that the hypothetical regulation of the cytoskeletal dynamics by MacroD2 is activated upon genotoxic stress. Moreover, as described in Chapter 3, ATM substrates possessing the SQ/TQ cluster domain (SCD) in the sequence were analyzed for biological term enrichment. In this case cytoskeleton-related terms were surprisingly overrepresented (371).

To proceed in this direction, I could take advantage of microscopy techniques and analyze the co-localization of MacroD2 with cytoskeletal components. As the focal adhesion biological terms appeared frequently in the analysis, examining the changes in foci formation in MacroD2 knock-out cells compared to wild-type cells would give an indication into MacroD2's role. Alternatively, testing neural cells for dendritic spine

formation in the presence and the absence of MacroD2 would also indicate MacroD2 function in this process. When a clear phenotype is linked to MacroD2 depletion, it is possible to study the mechanism behind it. It would be important to understand what protein directly interacts with MacroD2 and whether this protein is ADP-ribosylated. In fact, I imagine that even a single ADP-ribosylation moiety, at a specific site on the protein surface, can have important consequences on its activity and function.

4.8.5 NF-κB signaling pathway

The NF-κB pathway is activated either upon interleukines binding at the cell membrane or from the nucleus in response to genotoxic stimuli. Its activation increases the expression of the genes involved in inflammation by the release of the transcription factors NF-κB from the cytoplasmic inhibitory complex and by their translocation into the nucleus. The NF-κB pathway is connected with both ATM signaling and mono-ADP-ribosylation. In fact, upon genotoxic stress, ATM induces the activation of the NF-κB pathway by modifying and escorting the positive regulator of NF-κB signaling, NEMO, out of the nucleus (226). On the other side, NEMO activity is inhibited by ARTD10-mediated MARylation (194).

Several factors involved in the NF-κB signaling in different ways were found in several of my protein lists: NKAP, TRIP6, PRDX4 and G3BP2. The NF-κB activating protein (NKAP) is present in the "MacroD2 full-length DMSO" condition **(see Table A.1)** and has been shown to induce the NF-κB pathway (403). The peroxiredoxin-4 (PRDX4), instead, was present in the "MacroD2 macrodomain DMSO" condition **(see Table A.5)**, and it is a thioredoxin specifically able to activate the NF-κB pathway (404). In this study the authors first showed the cytoplasmic localization of the PRDX4 and suggested that it is involved in NF-κB regulation by affecting the phosphorylation of the IκBα inhibitor. The thyroid receptor-interacting protein 6 (TRIP6) was enriched in the "MacroD2 macrodomain etoposide" condition **(see Table A.8)**. It has been shown to reside at the promoters of the NF-κB-responsive genes and to mediate either the activation, when the NF-κB transcrition factors are binding, or the repression, when the glucocorticoid receptor is binding those same genes (405). Last but not least, the Ras GTPase-activating protein-binding protein 2 (G3BP2) was also enriched in the "MacroD2 macrodomain etoposide" condition **(see Table A.8)**. G3BP2 is able to

retain the NF-κB-IκBα complex in the cytoplasm, acting as repressor of the pathway (406).

Therefore, MacroD2 could accumulate in the cytoplasm upon ATM activation and help remove the last "break", the MARylation performed by ARTD10 on NEMO, which would be necessary for the full activation of the NF-κB pathway. In order to test this hypothesis, *in vitro* assays for the removal of MARylation on NEMO by MacroD2 are the necessary starting point. Subsequently, the activation of the NF-κB pathway must be determined *in vivo*, by comparing wild-type cells to MacroD2$^{-/-}$ cells. This can be performed by testing the phosphorylation levels of the inhibitor IκBα. IκBα's inhibitory action is mediated through trapping the transcription factors within the cytoplasm. After its ATM-induced nuclear export, NEMO activates the complex of the IκB kinase (IKK), which phosphorylates IκBα and forces the disassembly of the inhibitory complex (226). Thus, if NEMO is a substrate of MacroD2, the NF-κB translocation into the nucleus and activation of the gene expression should be less in MacroD2$^{-/-}$ cells than their wild-type counterpart. In case of decreased activation of the NF-κB pathway, the phenotype should be further verified by a rescue experiment, for example, by transiently transfecting MacroD2 wild-type, MacroD2 ADP-ribose binding-deficient mutant (G188E) or MacroD2 export-deficient mutant (S276,345,415,426A). The ADP-ribose binding-deficient mutant should not be able to rescue the phenotype and the export-deficient mutant should do it only partially.

In conclusion, the possible interaction between MacroD2 and the NF-κB pathway would be particularly interesting as it would give more relevance to the regulatory role of the nuclear export of MacroD2. In normal conditions, the NF-κB factors are cytoplasmic, blocked by the inhibitory complex, while MacroD2 presence in the nucleus avoid its overwhelming presence in the cytoplasm. Upon genotoxic stress, ATM would shift the balance towards the cytoplasm for both NEMO and MacroD2. MacroD2 would then allow the full release of the NF-κB transcription factors by ensuring the total activity of NEMO. In such a story, many passages across the nuclear envelope would hinder the overexpression of genes involved in the inflammatory response.

5

Outlook

The work I have performed in these years of research contributed to a deeper understanding of the ADP-ribosyl-hydrolase world. In fact, I show in this thesis that the ADP-ribosyl-hydrolase MacroD2 exports from the nucleus upon DNA damage. Although most of my experiments are performed with EGFP-tagged version of the protein, I could also show that the endogenous MacroD2 shares this behavior. The nuclear export of MacroD2 is induced by double-strand breaks via ATM activation. The kinase phosphorylates two serines that are located in the intrinsically disordered region of MacroD2.

MacroD2 nuclear export can be easily considered a mechanism to regulate its enzymatic activity, as shown by the change in recruitment dynamics at the DNA lesion caused by the block of the nuclear export. However, the regulation might have other possible outcomes. To help define the range of consequences, as well as the exact mechanism that allows MacroD2 to cross the nuclear envelope, I decided to start uncovering the MacroD2 interactome, by means of a co-IP experiment coupled with protein mass fingerprinting. The different hypothesis I proposed could be considered when proceeding with the validation step.

This work underlines the importance of the integration between the different signaling pathways, above all upon extraordinary conditions, such as for DNA damage. Many times, ATM signaling and ADP-ribosylation have been shown to cross each others' paths regarding modulation of the DNA damage response. But never before had the interaction shown such a clear and incontrovertible effect, like the nuclear export of an enzyme. Moreover, although ATM and ARTD1 are suggested to be directly in

contact, my work shows in a definitive way that the interaction between these two systems does happen and it can have a visible output. Now, uncovering MacroD2 cellular functions, its interacting network and its specific subset of substrates, is the next step needed to really understand the reasons for the interaction between the two signaling pathways.

With this study, the uniqueness of MacroD2 is revealed. MacroD2, in fact, is the only human ADP-ribosyl-hydrolase so far described to be regulated in its localization. Experiments with both full-length MacroD1 and TARG1 show localization at the mitochondria or only recruitment to the DNA lesion, respectively. Neither of them is exported from the nucleus upon DNA damage. This specific feature of MacroD2 suggests that the activity of this enzyme requires a fast and tight regulation, which is not required for the other two proteins. And enzymatic activity undergoing a complicated regulation could easily point to important functions or interactions so far neglected.

Another interesting finding is the double nature of MacroD2. Although it is an enzyme, MacroD2 has a C-terminus fragment that is an intrinsically disordered region. Intrinsically disordered proteins are underrepresented among enzymes, and still MacroD2 belongs to both groups. However, these two features may be combined in MacroD2 because they are well confined within the two portions of the proteins sequence: the N-terminus macrodomain can perform its enzymatic activity and keep its ordered folding, while the C-terminus region keeps its disordered structure, more apt to address its signaling functions. This unique "schizophrenic" nature of MacroD2 probably underlies the importance that it has in the regulation of cellular functions.

Remarkably, MacroD2 has been associated several times to autism syndrome and other neurodevelopmental diseases. However, we don't know yet how this phenotype can be connected to its enzymatic activity. We are surely getting closer to solve the riddle, but more effort should be put into filling the gap between MacroD2 molecular activity and its consequences on the whole organism, by understanding the one or many functions that this protein might have. Considering the consequences that disinformation on the alleged connection between autism and vaccination has caused for the public health of many western countries, the understanding of a possible mechanism inducing this syndrome should be one of the top priorities of research on MacroD2.

Another conclusion I draw from my work is that we need to develop a more comprehensive understanding of the different signaling mechanisms ongoing in the cell. So far,

partially due to technical limitations, the research on ADP-ribosylation has focused on the transferases and their activation. It is now important to pay attention to the whole cycle, by understanding also the mechanisms that induce the destruction of the specific signal when the emergency phase is over. ADP-ribosylation is a very dynamic signal, and the accurate regulation of its mechanisms must be a challenging task. But, as this is true within the regulation of the sole ADP-ribosylation signaling, it is also true when considering the many signaling pathways in action at the same time all together. Upon each condition, several signaling pathways are activated to various extents.

Although it is quite a daunting task, our knowledge of what is happening in the cell at every moment is still very fragmented. But the real understanding comes with the knowledge of all the factors contributing to a system and the ability of generating successful predictions. With my experience, I can say that in biology we are not there yet. Still, the path in front of us does not seem so long anymore.

6

Material and methods

6.1 Reagents

DMSO (Dimethyl sulfoxide) (D2438; Sigma)
Etoposide (E1383; Sigma)
Aphidicolin (A0781; Sigma)
Camptothecin (7689-03-4; Sigma)
PARP inhibitor Olaparib (AZD2281; Selleckchem)
ATM kinase inhibitor (KU55933; Selleckchem)
ATR kinase inhibitor (VE-821; Selleckchem)
DNA-PK kinase inhibitor (NU7441; Selleckchem)
(^{32}P)γATP (FP-301; Hartmann Analytic)
GFP-trap homemade (from Dr. A. Bowman and G. Jankevicius)
GFP-trap (gta; Chromotek)
bisBenzimide H 33342 trihydrocloride (Hoechst; Sigma)
Peptides (Metabion)
Oligos (Metabion)

6.2 Antibodies

Anti-MacroD2 (#494-7, rabbit polyclonal, from in-house facility)
Anti-GFP antibody (goat polyclonal, lab stock)
Anti-GST antibody (rat monoclonal from Dr. Kremmer of Helmholtz Zentrum Munich)

Anti-ATM [Y170] (ab32420, rabbit monoclonal, Abcam)
Anti-GAPDH (rat monoclonal #5C4 from Dr. Kremmer of Helmholtz Zentrum Munich)
Anti-beta actin antibody (rabbit polyclonal, ab8227, Abcam)
Alexa-488 goat anti-rabbit (Invitrogen)
Goat anti-rabbit IgG-HRP conjugated (BIO-RAD)
Goat anti-rat IgG-HRP conjugated (Jackson ImmunoResearch)

6.3 Buffers

RIPA buffer
50 mM Tris-Cl, pH 7.5
150 mM NaCl
0.1% Sodium Deoxycholate
1 mM EDTA
1% NP-40

SDS protein sample loading buffer 4x
250 mM Tris-HCl pH 6.8
10% dithiothreitol
5% SDS
40% glycerol
0,005% Bromphenol Blue

4x upper/stacking gel SDS buffer
500 mM Tris-HCl pH 6.8
0.1% SDS

4x lower/separating gel SDS buffer
1.5 M Tris-HCl; pH 8.7
0.1% SDS

1x Western Blot transfer buffer

25 mM Tris buffer

192 mM Glycine

20% Methanol

0.0375% SDS (added only for transferring ATM protein)

Ponceau S

0.1% Ponceau S (Sigma)

1% Acetic acid

1x Coomassie blue R-250

40% Methanol

10% Acetic acid

0.1% Coomassie Brilliant Blue R-250 (Sigma)

for destaining, the buffer was prepared without Coomassie dye

6.4 Cell culture media

Normal culture medium

DMEM - high glucose (#5671 - Sigma)

10 % FBS (Gibco)

2 mM L-glutamine (Sigma)

100 U mL^{-1} penicillin + 100 g mL^{-1} streptomycin (Sigma)

McCoy's-based culture medium

McCoy's 5a (Sigma)

10 % FBS (Gibco)

2 mM L-glutamine (Sigma)

100 U mL^{-1} penicillin + 100 g mL^{-1} streptomycin (Sigma)

Imaging medium

Phenol Red-free Leibovitz's L-15 (21083027 - Life Technologies)

10 % FBS (Gibco)

2 mM L-glutamine (Sigma)
100 U mL^{-1} penicillin + 100 g mL^{-1} streptomycin (Sigma)

6.5 Kits

The list of commercial kits used during the study is shown in the Table 6.1.

Item	Manifacturer
Plasmid preparation Miniprep	Metabion
Plasmid preparation Midiprep	Promega
PCR purification	Metabion
Gel extraction Kit	Metabion
NuPAGE (precast gels and buffers)	Invitrogen

Table 6.1: Kits - Overview of the commercial kits used during the study.

6.6 Plasmids

The list of plasmid used during the study is shown in the Table 6.2.

Plasmid	Database location
pmEGFP-C1	CL1036
pmEGFP-C1-MacroD2 full-length	CL2251
pmEGFP-C1-MacroD2 macrodomain (aa 1-234)	CL1977
pmEGFP-C1-MacroD2 C-terminus (aa 236-448)	CL1978
pmEGFP-C1-MacroD2 G188E	CL3521
pmEGFP-C1-MacroD2 G100E+I189R+Y190N	CL3760
pmEGFP-C1-EGFP-MacroD2 C-terminus (aa 236-end)	CL3758
pmEGFP-C1-MacroD2 S276,345,415,426A	CL3714
pmEGFP-C1-MacroD2 S276,345,415A	CL3715
pmEGFP-C1-MacroD2 S276,345,426A	CL3716
pmEGFP-C1-MacroD2 S276,415,426A	CL3717
pmEGFP-C1-MacroD2 S345,415,426A	CL3718
pmEGFP-C1-MacroD2 S276,426A	CL3759
pmEGFP-C1-MacroD2 (aa410-418)	CL3503
pmEGFP-C1-MacroD2 (aa402-418)	CL3742
pmEGFP-C1-MacroD2 (aa382-418)	CL3552
pmEGFP-C1-MacroD2 (aa382-418) D410A	CL3540

Continues on next page

Table 6.2 – *Continued from previous page*

Plasmid	Database location
pmEGFP-C1-MacroD2 (aa382-418) V411E	CL3543
pmEGFP-C1-MacroD2 (aa382-418) E412A	CL3544
pmEGFP-C1-MacroD2 (aa382-418) M413A	CL3546
pmEGFP-C1-MacroD2 (aa382-418) N414A	CL3547
pmEGFP-C1-MacroD2 (aa382-418) S415A	CL3548
pmEGFP-C1-MacroD2 (aa382-418) Q416A	CL3549
pmEGFP-C1-MacroD2 (aa336-372)	CL3729
pmEGFP-C1-MacroD2 (aa336-372) S345A	CL3736
pmEGFP-C1-MacroD2 (aa336-372) Q346A	CL3737
pmEGFP-C1-MacroD2 (aa336-372) M350A	CL3741
pETMCN-6His-TEV-MacroD2 full-length	CL3528
pETMCN-6His-TEV-V5-MacroD2 macrodomain (aa 1-234)	CL3053
pETMCN-6His-TEV-MacroD2 C-terminus (aa 244-448)	CL3529
pETMCN-6His-TEV-EGFP-MacroD2 (aa 1-448)	CL3720
pETMCN-6His-TEV-EGFP-MacroD2 C-terminus (aa 236-448)	CL3743
pGEX-6P1-GST-EGFP-MacroD2 (aa 382-418)	CL3721
pETMCN-6His-TEV-EGFP-MacroD2 (aa 382-418)	CL3745
pETMCN-6His-TEV-EGFP-MacroD2 (aa 382-418) D410A	CL3722
pETMCN-6His-TEV-EGFP-MacroD2 (aa 382-418) V411E	CL3723
pETMCN-6His-TEV-EGFP-MacroD2 (aa 382-418) E412A	CL3724
pETMCN-6His-TEV-EGFP-MacroD2 (aa 382-418) M413A	CL3725
pETMCN-6His-TEV-EGFP-MacroD2 (aa 382-418) N414A	CL3726
pETMCN-6His-TEV-EGFP-MacroD2 (aa 382-418) S415A	CL3727
pETMCN-6His-TEV-EGFP-MacroD2 (aa 382-418) Q416A	CL3728
pGEX-4T1-GST-14-3-3 ϵ	CL3747
pGEX-4T1-GST-14-3-3 ϵ K49E	CL3748
pGEX-2TK-GST-14-3-3 σ	CL3749
pGEX-4T1-GST-14-3-3 ζ	CL3750
pGEX-2TK-GST-14-3-3 β	CL3751
pGEX-6P1-GST-14-3-3 η	CL3752
pGEX-6P1-GST-14-3-3 τ	CL3753
pGEX-4T2-GST-14-3-3 γ	CL3754
pcDNA5-FRT/TO-EGFP	CL3553
pcDNA5-FRT/TO-EGFP-C1-MacroD2 (aa 1-448)	CL3538
pcDNA5-FRT/TO-EGFP-C1-MacroD2 (aa 1-243)	CL3761
pcDNA5-FRT/TO-EGFP-C1-MacroD2 (aa 237-448)	CL3554
pOG44 Flippase	CL3640

Continues on next page

Table 6.2 – *Continued from previous page*

Plasmid	Database location

Table 6.2: Plasmids - Overview of the plasmids used during the study.

6.7 Cell culture and treatments

6.7.1 Cell lines used

The lists of the used cell lines is shown in the Table 6.3.

Cell lines	Media
U2OS	Normal
U2OS MacroD2$^{-/-}$ (GM09)	Normal
U2OS mYFP-MacroD2 (aa 1-448) + mCherry-H2B	Normal
U2OS mEGFP-MacroD2 (aa 237-448) + mCherry-H2B	Normal
U2OS mEGFP-MacroD2 (aa 1-448)	Normal
U2OS mEGFP-MacroD2 (aa 237-448)	Normal
U2OS mEGFP-MacroD2 (aa 1-448) G188E	Normal
U2OS mEGFP-mEGFP-MacroD2 (aa 237-448)	Normal
HeLa Kyoto mCherry-H2B	Normal
G-361	McCoy's-based
HT-144	McCoy's-based
HEK293 Flp-In^{TM} T-REx^{TM}	Normal
HEK293 Flp-In^{TM} T-REx^{TM} EGFP	Normal
HEK293 Flp-In^{TM} T-REx^{TM} EGFP-MacroD2 (aa 1-448)	Normal
HEK293 Flp-In^{TM} T-REx^{TM} EGFP-MacroD2 (aa 1-243)	Normal
HEK293 Flp-In^{TM} T-REx^{TM} EGFP-MacroD2 (aa 237-448)	Normal

Table 6.3: Cell lines - Overview of the cell lines used during the study.

6.7.2 Cell culture

Human U2OS stable cell lines were grown in normal cell culture medium supplemented with 200 μg mL^{-1} G418 (Gibco). For maintenance of the mCherry-H2B construct, 2 μg mL^{-1} Puromycin (Sigma) may have been added to the medium. HEK293 stable cell lines were grown in normal cell culture medium supplemented with 15 μg mL^{-1} blasticidin and 100 μg mL^{-1} hygromycin. For transient transfections, HeLa Kyoto cell

line stably expressing mCherry-H2B were mainly used. Transient transfections were also performed on U2OS and HT-144 cell lines.

Cells were plated in borosilicate 8-well Lab-Tek chambered cover glasses (Thermo-Scientific) for imaging experiments. For other types of experiments, different plastic supports were used: 6-well plates, 12-well plates. 24-well plates, 60-cm^2 dishes or 150-cm^2 dishes (TPP).

6.7.3 Transfections and treatments

For transient transfections, cells were plated in a number that was adjusted according to the cell line, the day of use and the surface of plating (for example, 20000 U2OS cells were plated in one Lab-Tek well to be used on the next day).The next day, DNA transfection was performed by means of Xfect (Clontech) according to the manufacturer's recommendations. For example, regarding the most used setup, for one single Lab-Tek well, 500 ng of plasmid was used together with 0.3 μL of X-fect reagent. Cells were tested the day after transfection.

For siRNA-mediated MacroD2 and ATM depletion, Negative Control #2, s44380 and s1708 Silencer Select siRNA oligonucleotides (Ambion) were used in combination with Lipofectamine RNAiMAX (ThermoFisher) according to the manufacturers protocol. Cells were tested 72 hours post transfection.

Small molecule inhibitors were added to the culture medium 30 minutes before imaging in the concentration indicated. For DNA damage-induction upon microscopy experiments, camptothecin (50 μM) or etoposide (10 μM) were added during the imaging session. For protein lysates, cells were incubated with etoposide (10 μM) for 1 hour before collection.

6.7.4 Synchronization with Aphidicolin

To synchronize the cells, a protocol with Aphidicolin (Sigma) was used. On the first day, 20000 cells of U2OS mEGFP-MacroD2 (aa 237-448) + mCherry-H2B were plated in each 8-well Lab-Tek well. After 4 hours, Aphidicolin to a final concentration of 5 μg mL^{-1} was added. After 16 hours, the reactivation process was performed by means of a wash in PBS, a wash in normal medium and an incubation of 10 minutes in the incubator. These three steps were performed twice. Then imaging medium was added and cells were ready for the experiment.

6.8 Microscopy experiments

6.8.1 Microscope setup and imaging

During the experiments, cells were kept at 37 °C in Imaging medium. Imaging was performed on a Zeiss AxioObserver Z1 confocal spinning-disk microscope equipped with an AxioCam HRm CCD camera (Zeiss) through a Zeiss C-Apochromat 40x/1.2 water-immersion objective or a Zeiss Plan-Apochromat 63x/1.2 water-immersion objective. For laser microirradiation, we used a 355-nm wavelength, diode-pumped, solid-state, pulsed laser (DPSL-355/14, Rapp OptoElectronics). DNA damage was induced by focusing the laser to a point or a line in the nucleus or by adding camptothecin or etoposide. Cells were imaged every 10 seconds for 30 minutes.

6.8.2 Detection of endogenous MacroD2 by immunofluorescence

U2OS cells were plated in borosilicate 8-well Lab-Tek chambered cover glasses (ThermoScientific). Treatment with etoposide 10 M was performed for one hour. Cells were then fixed and permeabilised with Methanol/Acetone 1:1 solution for 10 minutes at -20 °C. Blocking was performed with PBS-Tween 0.1% with BSA 5%. Cells were incubated with the antibody anti-MacroD2 antibody (1:500), diluted in PBS-Tween 0.1% - BSA 3%, for one hour at room temperature. Incubation with Alexa-488-conjugated antibody (1:500) was performed for one hour at room temperature. Lastly, cells were incubated with Hoechst 33342 (Sigma) at 200 ng mL^{-1} final concentration for 10 minutes.

6.8.3 Analysis of MacroD2 export

The time-lapse and the immunofluorescence images were analysed with CellProfiler 2.0 (331). The mCherry-H2B or Hoechst images were used for the segmentation of the nucleus. To quantify protein export upon treatment, background was first subtracted and the ratio of nuclear intensity over cytoplasmic intensity of MacroD2 in each cell was calculated. For the time-lapse images, the ratio of each time point was normalized to the first time point. Only the cells properly segmented in the first time-point of the time laps were selected for the calculation of the average. Igor Pro (WaveMetrics) was used for analyzing and plotting the data.

6.8.4 Comparison between nuclear and cytoplasmic signal change

To compare nuclear signal depletion with cytoplasmic signal increase, HeLa cells stably expressing mCherry-H2B and transfected with mEGFP-MacroD2 were used for UV-laser experiments, as described before. The analysis was performed in Fiji (`http://fiji.sc/Fiji`) and consisted of manually segmenting the whole cell, the nucleus and the cytoplasm. The mean intensity of each object was subtracted of the background. The nuclear mean intensity over the total cell mean intensity was compared to the cytoplasm mean intensity over the total cell mean intensity. Igor Pro (WaveMetrics) was used for plotting the data.

6.8.5 Analysis of MacroD2 recruitment

To quantify the amount of protein recruited to DNA damage site, cells were registered by means of the MultiStackReg plugin (`http://bradbusse.net/sciencedownloads.html`) in Fiji (`http://fiji.sc/Fiji`). The mean intensity of the recruitment area was subtracted with the background and normalized over the mean intensity of the total cell. Then, the obtained ratio at each time point was normalized to the intensity of the initial time point. Igor Pro (WaveMetrics) was used for analyzing and plotting the data.

6.8.6 Automatized analysis of MacroD2 export

For the experiments 3.24 and 3.28, changes to the normal setup have been introduced. Cells were plated in black cycloolefin-covered 96-well plates (Greiner) with a number of cells per well adjusted according to the experiment and the cell line. Before experiment, cells were treated for 10 minutes with Hoechst, in order to allow nuclei segmentation upon analysis. During the experiments, cells were kept at 37 °C in Imaging medium. Imaging was performed on a Zeiss AxioObserver Z1 confocal spinning-disk microscope equipped with an AxioCam HRm CCD camera (Zeiss) through a Zeiss Plan-Apochromat 20x/0.8 lens. Cells were imaged while treated with etoposide 10 μM, every 5 minutes for 80 or 85 minutes. Each well was imaged in 18 positions.

The time-lapse and the immunofluorescence images were analysed with CellProfiler 2.0 (331). The Hoechst images were used for the segmentation of the nucleus. To quantify protein export upon treatment, background was first subtracted and the average nuclear intensity and the average cytoplasmic intensity of all MacroD2 cells in the condition were calculated. The ratio of nuclear intensity over cytoplasmic intensity of MacroD2 was then calculated. The ratio of each time point was normalized to the first

time point. Igor Pro (WaveMetrics) was used for analyzing and plotting the data.

6.9 Immunoblotting

RIPA buffer was used to obtain whole cell lysate. Samples were also sonicated to shear genomic DNA and incubated at 4 °C for 30 minutes. Afterwards, supernatant was cleared from cellular debris by centrifugating for 30 minutes at 11000 g. The cleared lysate concentration was after determined via a Bradford assay on strongly diluted samples. Proteins were resolved either on 12% or 6%+12% SDS-PAGE, if detecting MacroD2 or ATM, respectively. The transfer was performed onto nitrocellulose membrane (Whatman Protran). Transfer efficiency was checked with Ponceau S solution.

The membrane was blocked in PBS-Tween 0.05% buffer supplemented with 5% non-fat milk. Proteins were detected with the appropriate primary antibodies (anti-MacroD2 antibody 1:1000; anti-ATM antibody 1:1000; anti-GAPDH antibody 1:10; anti-GST 1:10; anti-14-3-3ϵ 1:1000) and secondary antibodies coupled to horseradish peroxidase (1:10000). Detection was performed with Immobilon Western Blotting detection reagent (GE Healthcare).

6.10 Ciclohexamide chase

U2OS mYFP-MacroD2 full-length + mCherry-H2B cells were plated in two 6-well plates (TPP) with 300000 cells per well. On the next day, ciclohexamide (300 μg mL^{-1}) was added in all the wells, but the two control wells with ethanol only. After one hour, the timepoint 0 cells were collected and placed on ice. In the meantime, the chase was started by adding DMSO or etoposide (10 μM) in either of the two plates. Then, following timepoints were collected at the times: 45 minutes, 2 hours, 4 hours, 24 hours. Also the control with ethanol was collected at 24 hours.

All the samples were prepared with RIPA buffer, as described before. For the western blot, anti-GFP (1:1000) and anti-actin (1:5000) were used. The different signal bands were quantified. The signal of the GFP band was normalized over the actin band and each time point was normalized over time 0. The ratios were then averaged and analyzed with unpaired non-parametrical Mann-Whitney.

6.11 Affinity-purification of anti-MacroD2 serum

Column buffer
100 mM $NaHCO_3$, pH 8.3
500 mM NaCl

Acid buffer
100 mM acetate, pH 4.0
500 mM NaCl

Basic buffer
100 mM Tris, pH 8.0
500 mM NaCl

Loading buffer
10 mM Tris, pH 7.5
1mM PMSF

Washing buffer
10 mM Tris, pH 7.5
500 mM NaCl

This protocol has been provided by Dr. D. Mokranjac. Firstly, the purified 6His-TEV-V5-MacroD2 macrodomain (1-243) was digested by a TEV protease during the first dialysis over night. Afterwards, the protein was further purified by mean of a size-exclusion column loaded on the FPLC. 5 mg of V5-MacroD2 macrodomain (1-243) was then dyalized into column buffer by means of a PD-10 column (17-0851-01; GE-Healthcare).

For the preparation of the resin, CNBr-Sepharose 4B beads powder (GE Healthcare) was used. The powder was incubated for 30 minutes with 1 mM HCl, for hydration. Further washes were performed with 1 mM HCl. Afterwards, the beads were transfered into a disposable plastic column (Bio-Rad). The protein solution was added, a part from an aliquot that works as input for the coupling efficiency. The column was then rolled for 1 hour at room temperature. Unbound material was collected for checks of the coupling efficiency. The column was then washed with column buffer and the remaining

active groups were blocked by loading 1M ethanolamine pH 8.0. The ethanolamine was incubated for 2 hours at room temperature. Then, the column was washed with three cycles of two buffers at different pH: acetate at pH 4.0 and Tris at pH 8.0. Lastly, the columns is washed with 0.05% NaN_3 and stored at 4 °C.

For the affinity-purification, the column was equilibrated with 10 mM Tris, pH 7.5. The serum was prepared by diluting it with the loading buffer (6 mL + 24 mL). Then, the diluted serum was loaded onto the column. The flow-through was collected to test the purification efficiency. The column was then washed with 10 mM Tris, pH 7.5 and with the washing buffer. Finally, the elution consisted with three steps.

- 0.1 M citrate buffer, pH 4.0
- 0.1 M glycine buffer, pH 2.5

For each of these two elution steps, 1 mL fractions were collected and immediately neutralized with 100 μL 1 M Tris, pH 8.8. Then, the column was equilibrated with 0.1 M Tris, pH 8.8 and prepared for the third elution.

- 0.1 M phosphate buffer, pH 12.5

Aso in this case, the 1 mL fractions were collected and immediately neutralized with 100 μL 1 M glycine, pH 2.0.

For further usage, the column was washed with 0.1 M Tris, pH 8.8, then washed and stored with 0.05% NaN_3 at 4 °C.

6.12 Protein purification

Lysis/wash buffer
50 mM Tris-Cl, pH 7.5
500 mM NaCl
20 mM imidazole
1mM DDT
protease inhibitor cocktail (Roche)

Elution buffer
50 mM Tris-Cl, pH 7.5
500 mM NaCl
500 mM imidazole
1mM DDT

protease inhibitor cocktail (Roche)

Storage buffer
25 mM Tris-Cl, pH 7.5
250 mM NaCl
1 mM DTT

For His-tagged proteins, protein expression constructs were expressed in *E. coli* Rosetta (DE3)pLysS cells at 18 °C for 16 hours after 0.2 mM IPTG induction. Cell pellets were snap-frozen in liquid nitrogen and stored at -80 °C.

For histidine-tagged protein purification, the thawed pellet was resuspended in a lysis/wash buffer. Lysates were sonicated for 20 x 18 seconds at medium setting (Branson) until the lysate was not viscous and centrifuged for 45 minutes at ∼45,000 g at 4 °C.

The supernatant was exposed to a column with Ni-NTA agarose (Qiagen), washed 3 times with 30 mL wash buffer and eluted with 5 mL the elution buffer. The proteins were dialyzed in storage buffer with either dialysis membranes (ZelluTrans - Roth; MWCO 3500) over-night or PD-10 columns (17-0851-01; GE-Healthcare).

Peak fractions were confirmed by SDS-PAGE and Coomassie staining. Concentrations were determined by Bradford assay (Biorad).

For GST-tagged protein purification, the procedure was as above, with the exception that no imidazole was used, glutathione Sepharose 4 Fast Flow (GE-Healthcare) was used as column resin and 20 mM reduced-glutathione elution buffer (GE Healthcare) was used.

6.13 Experimental determination of MacroD2 secondary structure composition

6His-TEV-MacroD2 full-length, 6His-TEV-V5-MacroD2 macrodomain (aa 1-234) and 6His-TEV-MacroD2 C-terminus (aa 244-448) proteins were purified as previously explained with the Ni-NTA agarose (Quiagen). Upon dialysis into storage buffer, proteins were digested over-night with a TEV protease to remove the His-tag. Then, with the help of G. Jankevicius and A. Bowman, the proteins were further purified by mean of a size-exclusion column loaded on the FPLC.

The proteins were successively dyalized with 20 mM phosphate buffer (pH 7.3). For the circular dichroism test, a sample of each protein at the concentration of 5 μM was

used by Dr. F. Kamp for the run of the machine.

6.14 Kinase assay on MacroD2 fragments

Lysis buffer
25 mM Tris-Cl, pH 7.5
150 mM NaCl
0.5% NP-40
phosphatase inhibitor cocktail (PhosSTOP - Roche)
protease inhibitor cocktail (Roche)

Dilution buffer
25 mM Tris-Cl, pH 7.5
150 mM NaCl
0.01% NP-40
phosphatase inhibitor cocktail (PhosSTOP - Roche)
protease inhibitor cocktail (Roche)

Washing buffer
25 mM Tris-Cl, pH 7.5
500 mM NaCl
0.01% NP-40
phosphatase inhibitor cocktail (PhosSTOP - Roche)
protease inhibitor cocktail (Roche)

The purified proteins His-mEGFP-MacroD2 C-terminal (aa236-448) fragment (5 μg (200 pmol) per sample); GST-mEGFP-tagged MacroD2 (aa382-418) fragment (52 μg (1 nmol) per sample) and His-mEGFP-tagged MacroD2 (aa382-418) fragments (wild-type and mutants) (35 μg (1 nmol) per sample) were used.

HEK293 Flp-In^{TM} T-REx^{TM} cells (ThermoFisher) were collected on ice. Cells were resuspended in lysis buffer, then incubated on ice for 30 minutes while mechanically lysed by syringe. Where appropriate, KU55933 at the final concentration of 10 μM was added. The cell lysate was then supplemented with 250 U Benzonase nuclease (Sigma), MgCl2 to a final concentration of 4 mM and 2 μCi (^{32}P)γATP (FP-301 Hartmann Analytic).

The lysate was incubated at 37 °C for 40 minutes. It was then cleared from the chromatin with a maximum speed centrifugation, diluted by 1:5 with dilution buffer and the input was kept. Proper amounts of GFP-trap (Chromotek), previously washed with lysis:dilution buffer, were incubated with the supernatant for 1 hour at 4 °C with agitation. The beads were then washed two times with lysis:dilution buffer and once with washing buffer. Input and beads were boiled for 10 minutes at 95 °C in equal amounts of 2x SDS protein sample loading buffer. Beads were centrifuged for 10 minutes at maximum speed, in order to break the beads apart and release all the protein. Samples were separated on 1 mm Bis-Tris, 4-12 % acrylamide NuPAGE gel according to the manufacturers protocol (Invitrogen). the gel was stained with Coomassie Brilliant blue and imaged (ImageScanner III GE Healthcare). Afterwards, the gel was destained, dried for two hours with high temperature and exposed to imaging plates (Fuji Film).

Quantification for the kinase assay on purified C-terminal fragments was performed by defining the intensity of every band in both the Coomassie staining and in the autoradiography. In each case, the band signal was subtracted of the background signal present along the lane. Then, the autoradiography signal was normalized to the Coomassie staining signal, defining per each band the ratio radioactive signal/amount. The ratios from three different experiments were then averaged. Similarly, the quantification for the kinase assay on purified mutant MacroD2 fragments was performed by defining the intensity of every band and their ratio radioactive signal/coomassie signal, as described above. Each His-construct intensity was then subtracted of the residual intensity in the His-S415A mutant, whose radioactivity signal is generated by the degradation of the upper wild-type-containing band, when the GST-tag is lost. Then, the adjusted His-construct intensities were normalized to the GST-construct intensity of the same lane, normalized to the wild-type/wild-type ratio and averaged.

6.15 Phospho-peptide enrichment upon DNA damage

Lysis buffer
25 mM Tris-Cl, pH 7.5
150 mM NaCl
0.5% NP-40
10 mM NaF
1 mM Na3VO4
phosphatase inhibitor cocktail (PhosSTOP - Roche)
protease inhibitor cocktail (Roche)

Dilution buffer
25 mM Tris-Cl, pH 7.5
150 mM NaCl
0.01% NP-40
10 mM NaF
1 mM Na3VO4
phosphatase inhibitor cocktail (PhosSTOP - Roche)
protease inhibitor cocktail (Roche)

Washing buffer
25 mM Tris-Cl, pH 7.5
500 mM NaCl
0.01% NP-40
10 mM NaF
1 mM Na3VO4
phosphatase inhibitor cocktail (PhosSTOP - Roche)
protease inhibitor cocktail (Roche)

Storage buffer
PBS
100 mM NaF
10 mM Na3VO4
phosphatase inhibitor cocktail (PhosSTOP Roche)
protease inhibitor cocktail (Roche)

Human U2OS cell lines stably expressing mYFPMacroD2 full length+mCherry-H2B were treated with etoposide 5 M or DMSO for one hour. Cells were then collected on ice and treated with lysis buffer. Lysates were than diluted 1:5 with dilution buffer. After vigorous centrifugation, the supernatant was collected and the input was kept. Proper amounts of GFP-trap (Chromotek), previously washed with lysis:dilution buffer, were incubated with the supernatant for 1 hour at 4 °C with agitation. The beads were then washed two times with lysis:dilution buffer and once with washing buffer. Beads were then kept in storage buffer.

Further preparation, run of the machine and data analysis performed by Dr. An-

dreas Schmidt. Proteins bound to beads were treated with 5 mM DTT to reduce disulfide bonds and subsequently alkylated using 20 mM iodoacetamide. Proteins were cleaved with AspN for 12 hours at 25 °C. Following proteolytic digest, peptide mixtures were desalted on reverse-phase C18 stage tips. 90% of the peptide mixture was mixed with 2 M glycolic acid for phosphopeptide enrichment on TiO2 as described in (407), then loaded onto the prepared TiO2 resin (Glygen).

Eluted phosphopeptide mixtures and peptide mixtures were separated on an Ultimate 3000 nano-RP-HPLC coupled to an LTQ-Orbitrap mass spectrometer (both Thermo-Fisher). The mass spectrometer was operated in data-dependent acquisition mode with one survey scan for precursor mass detection and up to 6 MS/MS experiments per cycle. CID-MS/MS experiments were conducted with multi-stage activation for neutral loss masses of 32.7, 49 and 98. For MS acquisition, monoisotopic precursor selection and dynamic exclusion for 30 seconds were enabled.

Raw data were searched with MaxQuant vs. 1.5 against a human protein database (Uniprot, May 2015) and an in-house target database containing common target protein and contaminants. All phosphopeptide hits for the MacroD2 were manually curated. To quantify the phosphorylation, peak areas for the most abundant charged states for unmodified peptides and phosphopeptides of the target protein were extracted from the raw data of the tryptic peptide mixture before phosphopeptide enrichment and log2 normalized. Signal strength of unmodified peptides was used for normalization of the protein amount between control and treatment experiments of each biological replicate.

6.16 Interaction assays with 14-3-3 proteins

6.16.1 14-3-3 proteins purification

All the expression plasmid of GST-14-3-3 proteins were purchased via Addgene. For some of them, some mutagenesis steps were necessary to reproduce the sequence annotated on Uniprot. The GST-14-3-3 tagged proteins were purified according to the protocol described before **(see Methods 6.12)**, with the additional step of the size-exclusion column purification. Proteins were stored in storage buffer.

6.16.2 Pull-down with biotinilated-MacroD2 peptides

Dilution buffer
25 mM Tris-Cl, pH 7.5
150 mM NaCl

0.01% NP-40
phosphatase inhibitor cocktail (PhosSTOP - Roche)
protease inhibitor cocktail (Roche)

Washing buffer
25 mM Tris-Cl, pH 7.5
500 mM NaCl
0.01% NP-40
phosphatase inhibitor cocktail (PhosSTOP - Roche)
protease inhibitor cocktail (Roche)

Peptides were ordered from Metabion.

1. Biotin - G S L S Q R Q R S T S T P N V H (Control peptide - unphosphorylated)
2. Biotin - G S L S Q R Q R S T pS T P N V H (Control peptide - phosphorylated)
3. Biotin - N T P G P D V E M N S Q V D K V
4. Biotin - N T P G P D V E M N pS Q V D K V
5. Biotin - N T P G P D V E A N S Q V D K V
6. Biotin - N T P G P D V E A N pS Q V D K V

Peptides were dissolved following the instruction given by the manufacturer: *1* and *2* were dissolved in water; the other peptides were dissolved in 20% acetonitrile. Also, to avoid the oxidation of the methionine, water and acetonitrile used for dissolve the peptides have been previously insufflated with N_2. All peptides were dissolved with a concentration of 1 mg mL^{-1}.

The peptides were then conjugated to Streptavidin Sepharose High Performance (Sigma). 100 μg of peptide were resuspended in 400 μL PBS. In the meantime, 400 μL of beads was washed three times with PBS + 0.1% NP-40. The beads and the peptides were then combined and incubated on rotation over night at 4 °C. On the next day, the beads were washed three times with PBS + 0.1% NP-40 to remove the unbound peptide. The conjugated beads were then stored in PBS and kept at 4 °C.

Per each GST-tagged 14-3-3 recombinant protein, 2 μg of protein were resuspended in in dilution buffer. To clear any precipitated protein from the solution, the samples were centrifuged for 15 minutes at maximum speed and the supernatant was then

transferred to a fresh tube. From the cleared supernatant, an aliquot was kept as Input. At the same time, aliquots of 30 μL slurry conjugated beads were prepared and washed three times in dilution buffer.

The protein solution was finally added to the conjugated beads and incubated with rotation for 1 hour at 4 °C. Afterwards, beads were washed twice with dilution buffer and once with washing buffer. Each sample with 15 μL beads had the supernatant removed and was added with 20 μL 2x SDS protein sample loading buffer. The beads were then boiled for 10 minutes at 95 °C and centrifuged for 10 minutes at maximum speed, in order to break the beads apart and release all the protein.

Samples were separated on 12% acrylamide SDS-PAGE gel. For the western blot, anti-GST (1:10) was used to probe the amount of GST-14-3-3 proteins purified by means of the peptides.

6.16.3 Pull-down with purified MacroD2 fragments

Lysis buffer
25 mM Tris-Cl, pH 7.5
150 mM NaCl
0.5% NP-40
phosphatase inhibitor cocktail (PhosSTOP - Roche)
protease inhibitor cocktail (Roche)

Dilution buffer
25 mM Tris-Cl, pH 7.5
150 mM NaCl
0.01% NP-40
phosphatase inhibitor cocktail (PhosSTOP - Roche)
protease inhibitor cocktail (Roche)

Washing buffer
25 mM Tris-Cl, pH 7.5
500 mM NaCl
0.01% NP-40
phosphatase inhibitor cocktail (PhosSTOP - Roche)
protease inhibitor cocktail (Roche)

The purified proteins His-mEGFP-MacroD2 C-terminal (aa236-448) fragment (2 μg) were used together with a selection of GST-14-3-3 proteins (ϵ, ϵ K49E, ζ and β) (2 μg), with GST-EGFP as negative control. The same GST-14-3-3 proteins were used for the next experiment, with the His-mEGFP-tagged MacroD2 (aa382-418) fragments (wild-type, M413A and S415A) (7 μg).

HEK293 Flp-In^{TM} T-REx^{TM} cells (ThermoFisher) were collected on ice. Cells were resuspended in lysis buffer, then incubated on ice for 30 minutes while mechanically lysed by syringe. The cell lysate was then supplemented with 250 U Benzonase nuclease (Sigma) and MgCl2 to a final concentration of 4 mM or only water as control.

The lysate was incubated at 37 °C for 40 minutes. It was then cleared from the chromatin with a maximum speed centrifugation, diluted by 1:5 with dilution buffer and the input was kept. Proper amounts of glutathione Sepharose 4 Fast Flow (GE-Healthcare), previously washed with lysis:dilution buffer, were incubated with the supernatant for 1 hour at 4 °C with agitation. The beads were then washed two times with lysis:dilution buffer and once with washing buffer. Input and beads were boiled for 10 minutes at 95 °C in equal amounts of 2x SDS protein sample loading buffer. Samples were separated by 12% acrylamide SDS-PAGE gel. For the western blot, anti-GFP (1:1000) was used to probe the amount of EGFP-MacroD2 fragments purified by means of the GST-14-3-3 proteins.

6.16.4 Co-Immunopurification of 14-3-3 ϵ

HEK293 Flp-In^{TM} T-REx^{TM} stably expressing EGFP or EGFP-MacroD2 full-length were plated and the recombinant protein expression was induced by 1mg/mL doxycyclin. On the next day, cells were treated for 1 hour with etoposide 10 μM and then collected on ice. The whole cell extract was processed with RIPA buffer. In the meantime, aliquots of GFP-trap beads (Chromotek) were washed three times with RIPA buffer. A part from the input aliquot, 5 mg of cell lysates were added to the beads and incubated for 2 hours at 4 °C. Beads were then washed once with RIPA buffer and twice with PBS.

Input and beads were boiled for 10 minutes at 95 °C in 2x SDS protein sample loading buffer. The beads were centrifuged for 10 minutes at maximum speed, in order to break the beads apart and release all the protein. 2.5% input and 25% beads were loaded onto 12% acrylamide SDS-PAGE gel for separation. For the western blot, anti-MacroD2 (1:1000) and anti-14-3-3 ϵ (1:1000) were used to interact the two perspective interactors. Anti-GAPDH (1:10) was used as negative control for the pull-down.

6.17 Co-immunopurification of MacroD2 interactome

6.17.1 Generation of cell lines

HEK293 Flp-In^{TM} T-REx^{TM} were used to generate the four cell lines:

- EGFP-MacroD2 full-length
- EGFP-MacroD2 macrodomain (aa 1-243)
- EGFP-MacroD2 C-terminus (aa 237-448)
- EGFP

HEK293 Flp-In^{TM} T-REx^{TM} were plated in 6-well plates, in order to have 80% confluency on the next day. On the next day, cells were transfected with either pcDNA5-FRT/TO-EGFP, pcDNA5-FRT/TO-EGFP-C1-MacroD2 (aa 1-448), pcDNA5-FRT/TO-EGFP-C1-MacroD2 (aa 1-243) or pcDNA5-FRT/TO-EGFP-C1-MacroD2 (aa 237-448), together with pOG44 Flippase in a 1:9 ratio. In one well, pmEGFP was used as positive control for the transfection. The transfection was performed with the reagent Xfect (Clontech), according to the manufacturer's recommendations. On the next day, cells from one well were plated into seven 150-cm^2 dishes (TPP) in medium added with blasticidin (15 μg mL^{-1}) and hygromicin (100 μg mL^{-1}). This strong dilution was meant to have single colony formation in the big plates.

For two weeks, cells were grown in selecting medium, which was changed every three days to restore the selecting capacity. After two weeks, 14 colonies were collected and transfered into 96-well plates. As soon as cells became confluent, cells were transfered into a support with greater surface, in order to amplify them. In each stage, slowest-growing colonies were trashed. When reached the 6-well stage, colonies were plated and induced with 1mg mL^{-1} doxycyclin. The next day, cells were collected and tested by western blot (anti-GFP) to check for the correct size of the expressed construct. Colonies were also screened for fluorescence intensity. The three colonies left for each construct were then tested at the microscope for response to laser microirradiation. This test was also performed before each biological replicate preparation. Only the colony per each construct that was best performing for fluorescence intensity and growth rate was used for the real experiment.

6.17.2 Co-immunopurification

Each experiment was performed in four biological replicates. The four cell lines described above, HEK293 Flp-In^{TM} T-REx^{TM} expressing EGFP, EGFP-MacroD2 full-

length, EGFP-MacroD2 macrodomain and EGFP-MacroD2 C-terminus, were grown in normal medium, in two different plates. One day before the treatment and collection, all cells were induced with 1 mg mL^{-1} doxycyclin. On the next day, EGFP-MacroD2 full-length, macrodomain and C-terminus-expressing cells were treated either with DMSO or etoposide 10 μM for 1 hour.

Cells were collected on ice, washed with PBS and treated with RIPA buffer, in order to obtain the whole cell extract. In the meantime, aliquots of homemade GFP (for the first replicate) or GFP-trap beads (Chromotek, for the second to fourth replicate) were washed three times with RIPA buffer. A part from the input aliquot, 5 mg of cell lysates were added to the beads and incubated for 2 hours at 4 °C.

1. EGFP
2. EGFP-MacroD2 full-length DMSO
3. EGFP-MacroD2 full-length etoposide
4. EGFP-MacroD2 macrodomain DMSO
5. EGFP-MacroD2 macrodomain etoposide
6. EGFP-MacroD2 C-terminus DMSO
7. EGFP-MacroD2 C-terminus etoposide

Beads were then washed once with RIPA buffer and twice with PBS. The samples were than brought to Dr. Schmidt to perform the rest of the processing. Disulfide bonds were reduced with dithioteitrol at 56 °C. Then, the free cysteines were alkylated with 10 mM iodoacetamide in 100 mM Tris, 6 M urea. The urea is added to unfold the proteins and make them more accessible to both alkylation and following digestion steps. Proteins are then pre-digested with Lys-C (200 ng/sample) at 30 °C on rotation for 4 hours, in order to pre-cleave the proteins and remove the beads. After centrifugation, the supernatant was transfered into a new vial and the beads were washed two times with 100 mM ammonium bicarbonate. Each wash was added to the above-mentioned supernatant, to collect all the proteins from the beads. The collected proteins were then digested by trypsin (1 μg per sample) for 12 hours at 30 °C. The output peptides of Lys-C and trypsin digestion are exactly the same, so that for *in silico* peptide identification in the analysis, we could consider as only the trypsin have been used.

6.17.3 Mass spectrometry

After a desalting step, Dr. Schmidt performed the mass spectrometry run. Tryptic protein digests were analyzed by nano-high performance liquid chromatography mass spectrometry (HPLC-MS) on an Ultimate 3000 nanoLC system (Thermo-Fisher), coupled to a quadrupole time-of-flight (qTOF) mass spectrometer 6600 Triple TOF (Sciex). First, acidified peptides were injected onto a trapping column (5 x 0.3 mm) packed with C18 material (Thermo-Fisher, Acclaim PepMap, 5 μm, 100 Å) to remove salt and chemicals applied during chromatography at a flow rate of 20 μL min^{-1} of 0.05% TFA solution.

After the washing phase, peptides were separated on a nano-RP-C18 column (Thermo-Fisher, Acclaim PepMap, 250 x 0.075 mm x 2.4 μm, 100 Å) using a linear gradient from 4-35% ACN over a mixing time of 120 min. The column outlet was directly coupled to the electrospray-source to introduce the peptides into the mass spectrometer. The mass spectrometer was operated in data-dependent acquisition mode detecting peptides eluting from the chromatography column in a survey scan from 300-2000 m/z. Up to 40 precursor signals were subjected to fragmentation analysis by collision-induced dissociation per operation cycle of 3 seconds. Precursor ions which were fragmented in previous cycles were excluded for 30 seconds to reduce re-analysis and obtain deeper proteomic data.

6.17.4 Data analysis

Peptide identification was performed by mean of the Andromeda search engine integrated in the MaxQuant software (408). All samples were mapped against a database containing the amino acids sequences of human proteins (Uniprot, vs. Oct 2015). Compared to the other three, the first biological replicate performed very poorly and was then excluded from the analysis.

As intensities, a Max-Quant-embedded algorithm was used (379). For one protein, iBAQ signal consists of the the sum of all the peptides intensities divided by the number of theoretical peptides generated by the *in silico* digestion. For normalization of the different samples, every iBAQ signal was divided by a coefficient defined per each sample. To obtain the sample-specific coefficient, the peptides that are present in most samples were identified. The median iBAQ for each sample was then calculated and each median was normalized against the highest median value.

The normalized iBAQ signals of the different proteins were then filtered in order to keep only those proteins that were present in all the biological replicate of one

specific condition (for example, protein present in replicate 2, 3 and 4 for MacroD2 macrodomain DMSO were kept, even if missing in all others conditions). Then, in place of the missing values, the iBAQ values were inputed by mean of a normal distribution. Successively, t-tests between different samples were performed. Each t-test was performed by using 0.1 as minimal fold enrichment and 0.08 as false discovery rate (FDR). For the full-length conditions, the enrichments were performed against the EGFP negative control and the full-length fragment with the opposite treatment, for a total of two enrichments per condition. For the macrodomain and C-terminus fragments, the enrichments were performed against the EGFP negative control, the same fragment with the opposite treatment and the opposite fragment in both treatment, for a total of four enrichments per condition.

Having these series of enrichments, two protein list types were generated. The first list, the common proteins, is generated by selecting the proteins shared by at least two enrichments for the same condition. The second list is the pool of all the proteins found enriched in all the enrichments of a condition. Each of the two lists was then analyzed by means of the STRING database and its embedded software (`http://string-db.org/`). The common lists were further checked for the amount of overlap, by means of the Venny 2.1.0 (`http://bioinfogp.cnb.csic.es/tools/venny/`). The filtered common lists were again analyzed with the STRING associated software.

Appendix A

Lists of MacroD2 interactors

A.1 MacroD2 full-length DMSO protein list

Protein name	Gene name
MacroD2 full-length DMSO over EGFP	
Costars family protein ABRACL	ABRACL
Apoptotic chromatin condensation inducer in the nucleus	ACIN1
T-complex protein 1 subunit epsilon	CCT5
Cx9C motif-containing protein 4	CMC4
Elongation factor 1-beta	EEF1B2
Eukaryotic translation initiation factor 3 subunit K	EIF3K
Fragile X mental retardation syndrome-related protein 1	FXR1
Heat shock 70 kDa protein 12A	HSPA12A
Leucine-rich repeat-containing protein 59	LRRC59
O-acetyl-ADP-ribose deacetylase MACROD2	MACROD2
Microtubule-associated protein RP/EB family member 1	MAPRE1
Phosphatidylethanolamine-binding protein 1	PEBP1
Ran-specific GTPase-activating protein	RANBP1
60S ribosomal protein L31	RPL31
40S ribosomal protein S10	RPS10
Tubulin-specific chaperone A	TBCA
Treacle protein	TCOF1
Nucleolysin TIAR	TIAL1
MacroD2 full-length DMSO over full-length etoposide	

Continues on next page

Table A.1 – *Continued from previous page*

Protein name	Gene name
Aminoacyl tRNA synthase complex-interacting protein 2	AIMP2
ATPase family AAA domain-containing protein 3A	ATAD3A
Calmodulin-like protein 5	CALML5
Cystathionine beta-synthase	CBS
Eukaryotic translation initiation factor 3 subunit C	EIF3C
Eukaryotic translation initiation factor 3 subunit K	EIF3K
Guanine nucleotide-binding protein (...) subunit beta-1	GNB1
Histone H2A.V	H2AFV
Histone H2A type 1-J	HIST1H2AJ
Galectin-7	LGALS7
NF-kappa-B-activating protein	NKAP
Poly(U)-binding-splicing factor PUF60	PUF60
RNA-binding protein 3	RBM3
Iporin	RUSC2
Ubiquitin carboxyl-terminal hydrolase 5	USP5

Table A.1: MacroD2 full-length DMSO - Overview of enriched proteins in *MacroD2 full-length DMSO* sample, when compared with *EGFP* and *MacroD2 full-length etoposide* samples.

Protein name	Gene name
MacroD2 full-length DMSO common proteins	
Eukaryotic translation initiation factor 3 subunit K	EIF3K

Table A.2: MacroD2 full-length DMSO, common proteins - Overview of the proteins that are in common among the different "MacroD2 full-length DMSO" enrichments.

A.2 MacroD2 full-length etoposide protein list

Protein name	Gene name
MacroD2 full-length etoposide over EGFP	
Apoptotic chromatin condensation inducer in the nucleus	ACIN1

Continues on next page

Table A.3 – *Continued from previous page*

Protein name	**Gene name**
Sodium/potassium-transporting ATPase (...) alpha-1	ATP1A1
Cx9C motif-containing protein 4	CMC4
Elongation factor 1-beta	EEF1B2
Protein FAM98A	FAM98A
Leucine zipper putative tumor suppressor 3	LZTS3
O-acetyl-ADP-ribose deacetylase MACROD2	MACROD2
Microtubule-associated protein RP/EB family member 1	MAPRE1
28S ribosomal protein S23 (mitochondrial)	MRPS23
Myotrophin	MTPN
Phosphatidylethanolamine-binding protein 1	PEBP1
Proteasome subunit alpha type-6	PSMA6
RNA-binding protein 26	RBM26
60S ribosomal protein L31	RPL31
40S ribosomal protein S10	RPS10
Splicing factor 3B subunit 5	SF3B5
Sorting nexin-9	SNX9
Spermidine synthase	SRM
Single-stranded DNA-binding protein (mitochondrial)	SSBP1
Thyroid receptor-interacting protein 6	TRIP6
Tumor susceptibility gene 101 protein	TSG101
MacroD2 full-length etoposide over full-length DMSO	
Leucine zipper putative tumor suppressor 3	LZTS3
Myotrophin	MTPN
Ornithine aminotransferase (mitochondrial)	OAT
RNA-binding protein Raly	RALY
RNA-binding protein 26	RBM26
Splicing factor 3B subunit 5	SF3B5
Serine hydroxymethyltransferase (mitochondrial)	SHMT2
Sorting nexin-9	SNX9
Single-stranded DNA-binding protein (mitochondrial)	SSBP1
Thyroid receptor-interacting protein 6	TRIP6
Tubulin beta-2B chain	TUBB2B

Table A.3: MacroD2 full-length etoposide - Overview of enriched proteins in *MacroD2 full-length etoposide* sample, when compared with *EGFP* and *MacroD2 full-length DMSO* samples.

Protein name	Gene name
EGFP-MacroD2 full-length etoposide common proteins	
Leucine zipper putative tumor suppressor 3	LZTS3
Myotrophin	MTPN
RNA-binding protein 26	RBM26
Splicing factor 3B subunit 5	SF3B5
Sorting nexin-9	SNX9
Single-stranded DNA-binding protein (mitochondrial)	SSBP1
Thyroid receptor-interacting protein 6	TRIP6

Table A.4: MacroD2 full-length etoposide, common proteins - Overview of the proteins that are in common among the different "MacroD2 full-length etoposide" enrichments.

A.3 MacroD2 macrodomain DMSO protein list

Protein name	Gene name
MacroD2 macrodomain DMSO over EGFP	
ATP-citrate synthase	ACLY
Acidic leucine-rich nuclear (...) 32 family member E	ANP32E
Sodium/potassium-transporting ATPase (...) alpha-1	ATP1A1
Calreticulin	CALR
T-complex protein 1 subunit epsilon	CCT5
T-complex protein 1 subunit zeta	CCT6A
T-complex protein 1 subunit eta	CCT7
Chromatin target of PRMT1 protein	CHTOP
Dihydrolipoyl dehydrogenase (mitochondrial)	DLD
DnaJ homolog subfamily A member 1	DNAJA1
Aspartyl aminopeptidase	DNPEP
Enoyl-CoA hydratase (mitochondrial)	ECHS1
Elongation factor 1-beta	EEF1B2
Eukaryotic translation initiation factor 2 subunit 1	EIF2S1
Eukaryotic peptide chain release factor subunit 1	ETF1
Peptidyl-prolyl cis-trans isomerase FKBP4	FKBP4
Fragile X mental retardation syndrome-related protein 1	FXR1
Glutathione S-transferase omega-1	GSTO1

Continues on next page

Table A.5 – *Continued from previous page*

Protein name	Gene name
Histidine triad nucleotide-binding protein 1	HINT1
Importin subunit beta-1	KPNB1
L-lactate dehydrogenase B chain	LDHB
Protein ERGIC-53	LMAN1
Leucine-rich repeat-containing protein 59	LRRC59
O-acetyl-ADP-ribose deacetylase MACROD2	MACROD2
MARCKS-related protein	MARCKSL1
39S ribosomal protein L12 (mitochondrial)	MRPL12
Myosin light polypeptide 6	MYL6
Probable 28S rRNA (cytosine(4447)-C(5))-methyltransferase	NOP2
tRNA (cytosine(34)-C(5))-methyltransferase	NSUN2
Protein disulfide-isomerase	P4HB
Proliferation-associated protein 2G4	PA2G4
28 kDa heat- and acid-stable phosphoprotein	PDAP1
Programmed cell death protein 5	PDCD5
Phosphatidylethanolamine-binding protein 1	PEBP1
ATP-dependent 6-phosphofructokinase. platelet type	PFKP
Prohibitin	PHB
DNA-directed RNA polymerases I and III subunit RPAC1	POLR1C
DNA-directed RNA polymerase II subunit RPB3	POLR2C
26S proteasome non-ATPase regulatory subunit 7	PSMD7
RNA-binding protein Raly	RALY
Ran-specific GTPase-activating protein	RANBP1
RNA-binding protein 4	RBM4
60S ribosomal protein L31	RPL31
40S ribosomal protein S10	RPS10
40S ribosomal protein S14	RPS14
40S ribosomal protein S21	RPS21
RuvB-like 1	RUVBL1
SUMO-activating enzyme subunit 1	SAE1
Endophilin-A1	SH3GL2
Phosphate carrier protein (mitochondrial)	SLC25A3
Staphylococcal nuclease domain-containing protein 1	SND1
Small nuclear ribonucleoprotein F	SNRPF
Sorting nexin-9	SNX9
Superoxide dismutase [Cu-Zn]	SOD1
Spermidine synthase	SRM
Serine/arginine-rich splicing factor 5	SRSF5
Tubulin-specific chaperone A	TBCA
Transcription elongation regulator 1	TCERG1
Nucleolysin TIAR	TIAL1
Mitochondrial import inner (...) translocase TIM50	TIMM50

Continues on next page

Table A.5 – *Continued from previous page*

Protein name	Gene name
Transketolase	TKT
Translation machinery-associated protein 16	TMA16
Thioredoxin domain-containing protein 5	TXNDC5
UBX domain-containing protein 1	UBXN1
14-3-3 protein beta/alpha	YWHAB
14-3-3 protein zeta/delta	YWHAZ

MacroD2 macrodomain DMSO over macrodomain etoposide

Multidrug resistance protein 3	ABCB4
Actin-like protein 6A	ACTL6A
Acidic leucine-rich nuclear (...) 32 family member E	ANP32E
Rho GDP-dissociation inhibitor 1	ARHGDIA
Sodium/potassium-transporting ATPase (...) alpha-1	ATP1A1
ATP synthase subunit gamma (mitochondrial)	ATP5C1
ATPase inhibitor (mitochondrial)	ATPIF1
BAG family molecular chaperone regulator 2	BAG2
Complement C1q tumor necrosis factor-related protein 8	C1QTNF8
Calreticulin	CALR
T-complex protein 1 subunit theta	CCT8
Corneodesmosin	CDSN
Chromatin target of PRMT1 protein	CHTOP
Density-regulated protein	DENR
Dihydrolipoyl dehydrogenase (mitochondrial)	DLD
Enoyl-CoA hydratase (mitochondrial)	ECHS1
Emerin	EMD
Peptidyl-prolyl cis-trans isomerase FKBP3	FKBP3
Peptidyl-prolyl cis-trans isomerase FKBP4	FKBP4
Protein FRG1	FRG1
Far upstream element-binding protein 1	FUBP1
Glutathione S-transferase omega-1	GSTO1
Histidine triad nucleotide-binding protein 1	HINT1
Ig alpha-2 chain C region	IGHA2
Protein lin-7 homolog C	LIN7C
Leucine-rich repeat-containing protein 59	LRRC59
Malate dehydrogenase (mitochondrial)	MDH2
39S ribosomal protein L12 (mitochondrial)	MRPL12
mRNA turnover protein 4 homolog	MRTO4
Myotubularin-related protein 4	MTMR4
Myosin-10	MYH10
Myosin-9	MYH9
Myosin regulatory light chain 12A	MYL12A

Continues on next page

Table A.5 – *Continued from previous page*

Protein name	Gene name
Nuclear autoantigenic sperm protein	NASP
Protein disulfide-isomerase	P4HB
Proliferation-associated protein 2G4	PA2G4
28 kDa heat- and acid-stable phosphoprotein	PDAP1
Phosphatidylethanolamine-binding protein 1	PEBP1
Serine/threonine-protein phosphatase PGAM5 (mitochondrial)	PGAM5
DNA-directed RNA polymerase II subunit RPB3	POLR2C
Peroxiredoxin-4	PRDX4
Phosphoribosyl pyrophosphate synthase-associated protein 2	PRPSAP2
GTP-binding nuclear protein Ran	RAN
Arginine–tRNA ligase. cytoplasmic	RARS
RNA-binding protein 4	RBM4
60S ribosomal protein L28	RPL28
60S ribosomal protein L37a	RPL37A
40S ribosomal protein S23	RPS23
SUMO-activating enzyme subunit 1	SAE1
Endophilin-A1	SH3GL2
Staphylococcal nuclease domain-containing protein 1	SND1
U4/U6.U5 small nuclear ribonucleoprotein 27 kDa protein	SNRNP27
Sorting nexin-9	SNX9
Spermidine synthase	SRM
Serine/arginine-rich splicing factor 1	SRSF1
Serine/arginine-rich splicing factor 5	SRSF5
Stathmin	STMN1
Suppressor of G2 allele of SKP1 homolog	SUGT1
Transgelin-2	TAGLN2
Tubulin-specific chaperone A	TBCA
Tubulin-folding cofactor B	TBCB
Transcription elongation regulator 1	TCERG1
Nucleolysin TIAR	TIAL1
Mitochondrial import inner (...) translocase subunit TIM50	TIMM50
Transketolase	TKT
Transmembrane protein 178A	TMEM178A
DNA topoisomerase 1	TOP1
Tubulin beta-2B chain	TUBB2B
Ubiquitin-conjugating enzyme E2 S	UBE2S
Cytochrome b-c1 complex subunit 1 (mitochondrial)	UQCRC1
Voltage-dependent anion-selective channel protein 3	VDAC3
14-3-3 protein zeta/delta	YWHAZ

MacroD2 macrodomain DMSO over C-terminus DMSO

Continues on next page

Table A.5 – *Continued from previous page*

Protein name	Gene name
Acidic leucine-rich nuclear (...) 32 family member A	ANP32A
Acidic leucine-rich nuclear (...) 32 family member E	ANP32E
Sarcoplasmic/endoplasmic reticulum calcium ATPase 2	ATP2A2
T-complex protein 1 subunit zeta	CCT6A
T-complex protein 1 subunit theta	CCT8
Corepressor interacting with RBPJ 1	CIR1
DnaJ homolog subfamily A member 1	DNAJA1
Eukaryotic peptide chain release factor subunit 1	ETF1
Peptidyl-prolyl cis-trans isomerase FKBP1A	FKBP1A
Guanine nucleotide-binding protein (...) subunit beta-1	GNB1
Histidine triad nucleotide-binding protein 1	HINT1
Ig alpha-2 chain C region	IGHA2
Protein lin-7 homolog C	LIN7C
Protein ERGIC-53	LMAN1
Prelamin-A/C;Lamin-A/C	LMNA
mRNA turnover protein 4 homolog	MRTO4
Probable 28S rRNA (cytosine(4447)-C(5))-methyltransferase	NOP2
tRNA (cytosine(34)-C(5))-methyltransferase	NSUN2
Phosphatidylethanolamine-binding protein 1	PEBP1
26S proteasome non-ATPase regulatory subunit 7	PSMD7
GTP-binding nuclear protein Ran	RAN
Serine hydroxymethyltransferase (mitochondrial)	SHMT2
U4/U6.U5 small nuclear ribonucleoprotein 27 kDa protein	SNRNP27
Stathmin	STMN1
Tubulin-specific chaperone A	TBCA
Triosephosphate isomerase	TPI1
Thioredoxin domain-containing protein 5	TXNDC5
Voltage-dependent anion-selective channel protein 3	VDAC3
14-3-3 protein zeta/delta	YWHAZ

MacroD2 macrodomain DMSO over C-terminus etoposide

Acidic leucine-rich nuclear (...) 32 family member E	ANP32E
BAG family molecular chaperone regulator 2	BAG2
Calmodulin	CALM1
T-complex protein 1 subunit zeta	CCT6A
Cleavage and polyadenylation specificity factor subunit 2	CPSF2
Putative ATP-dependent RNA helicase DHX30	DHX30
DnaJ homolog subfamily A member 1	DNAJA1
Emerin	EMD
Far upstream element-binding protein 1	FUBP1
Glutamine synthetase	GLUL

Continues on next page

Table A.5 – *Continued from previous page*

Protein name	Gene name
Ig heavy chain V-III region KOL	IGHV3-23
Prelamin-A/C	LMNA
mRNA turnover protein 4 homolog	MRTO4
Myosin-10	MYH10
Phosphatidylethanolamine-binding protein 1	PEBP1
Serine/threonine-protein phosphatase PP1-gamma (...)	PPP1CC
Peroxiredoxin-4	PRDX4
26S proteasome non-ATPase regulatory subunit 4	PSMD4
Tyrosine-protein phosphatase non-receptor type 11	PTPN11
RNA-binding protein 4	RBM4
60S ribosomal protein L28	RPL28
U5 small nuclear ribonucleoprotein 200 kDa helicase	SNRNP200
Serine/arginine-rich splicing factor 5	SRSF5
Transcription elongation regulator 1	TCERG1
Thioredoxin domain-containing protein 5	TXNDC5
UBX domain-containing protein 1	UBXN1
Cytochrome b-c1 complex subunit 1 (mitochondrial)	UQCRC1

Table A.5: MacroD2 macrodomain DMSO - Overview of enriched proteins in *MacroD2 macrodomain DMSO* sample, when compared with *EGFP*, *MacroD2 macrodomain etoposide*, *MacroD2 C-terminus DMSO* and *MacroD2 C-terminus etoposide*.

Protein name	Gene name
MacroD2 macrodomain DMSO common proteins	
Acidic leucine-rich nuclear (...) 32 family member E	ANP32E
Sodium/potassium-transporting ATPase (...) alpha-1	ATP1A1
BAG family molecular chaperone regulator 2	BAG2
Calreticulin	CALR
T-complex protein 1 subunit zeta	CCT6A
T-complex protein 1 subunit theta	CCT8
Chromatin target of PRMT1 protein	CHTOP
Dihydrolipoyl dehydrogenase (mitochondrial)	DLD
DnaJ homolog subfamily A member 1	DNAJA1
Enoyl-CoA hydratase (mitochondrial)	ECHS1
Emerin	EMD
Eukaryotic peptide chain release factor subunit 1	ETF1
Peptidyl-prolyl cis-trans isomerase FKBP4	FKBP4
Far upstream element-binding protein 1	FUBP1

Continues on next page

Table A.6 – *Continued from previous page*

Protein name	Gene name
Glutathione S-transferase omega-1	GSTO1
Histidine triad nucleotide-binding protein 1	HINT1
Ig alpha-2 chain C region	IGHA2
Protein lin-7 homolog C	LIN7C
Protein ERGIC-53	LMAN1
Prelamin-A/C;Lamin-A/C	LMNA
Leucine-rich repeat-containing protein 59	LRRC59
39S ribosomal protein L12 (mitochondrial)	MRPL12
mRNA turnover protein 4 homolog	MRTO4
Myosin-10	MYH10
Probable 28S rRNA (cytosine(4447)-C(5))-methyltransferase	NOP2
tRNA (cytosine(34)-C(5))-methyltransferase	NSUN2
Protein disulfide-isomerase	P4HB
Proliferation-associated protein 2G4	PA2G4
28 kDa heat- and acid-stable phosphoprotein	PDAP1
Phosphatidylethanolamine-binding protein 1	PEBP1
DNA-directed RNA polymerase II subunit RPB3	POLR2C
Peroxiredoxin-4	PRDX4
26S proteasome non-ATPase regulatory subunit 7	PSMD7
GTP-binding nuclear protein Ran	RAN
RNA-binding protein 4	RBM4
60S ribosomal protein L28	RPL28
SUMO-activating enzyme subunit 1	SAE1
Endophilin-A1	SH3GL2
Staphylococcal nuclease domain-containing protein 1	SND1
U4/U6.U5 small nuclear ribonucleoprotein 27 kDa protein	SNRNP27
Sorting nexin-9	SNX9
Spermidine synthase	SRM
Serine/arginine-rich splicing factor 5	SRSF5
Stathmin	STMN1
Tubulin-specific chaperone A	TBCA
Transcription elongation regulator 1	TCERG1
Nucleolysin TIAR	TIAL1
Mitochondrial import inner (...) translocase subunit TIM50	TIMM50
Transketolase	TKT
Thioredoxin domain-containing protein 5	TXNDC5
UBX domain-containing protein 1	UBXN1
Cytochrome b-c1 complex subunit 1 (mitochondrial)	UQCRC1
Voltage-dependent anion-selective channel protein 3	VDAC3
14-3-3 protein zeta/delta	YWHAZ

Continues on next page

Table A.6 – *Continued from previous page*

Protein name	Gene name

Table A.6: MacroD2 macrodomain DMSO, common proteins - Overview of the proteins that are in common among the different "MacroD2 macrodomain DMSO" enrichments.

A.4 MacroD2 macrodomain etoposide protein list

Protein name	Gene name
MacroD2 macrodomain etoposide over EGFP	
Acidic leucine-rich nuclear (...) 32 family member E	ANP32E
Cx9C motif-containing protein 4	CMC4
Fragile X mental retardation syndrome-related protein 1	FXR1
Keratin. type I cytoskeletal 18	KRT18
Programmed cell death protein 6	PDCD6
60S ribosomal protein L31	RPL31
Survival of motor neuron-related-splicing factor 30	SMNDC1
MacroD2 macrodomain etoposide over macrodomain DMSO	
Alpha-aminoadipic semialdehyde dehydrogenase	ALDH7A1
Annexin A7	ANXA7
Calmodulin-like protein 5	CALML5
Coiled-coil domain-containing protein 80	CCDC80
Bifunctional coenzyme A synthase	COASY
Coronin-1C	CORO1C
dCTP pyrophosphatase 1	DCTPP1
DnaJ homolog subfamily C member 2	DNAJC2
Four and a half LIM domains protein 1	FHL1
Ras GTPase-activating protein-binding protein 2	G3BP2
Histone H2A type 1-J	HIST1H2AJ
Heat shock 70 kDa protein 12A	HSPA12A
Isocitrate dehydrogenase [NAD] subunit beta (mitochondrial)	IDH3B
Keratin. type I cytoskeletal 18	KRT18
Keratin. type II cytoskeletal 6B	KRT6B
Methylated-DNA–protein-cysteine methyltransferase	MGMT
Poly [ADP-ribose] polymerase 1	PARP1
Proteasome subunit alpha type-3	PSMA3
SWI/SNF complex subunit SMARCC1	SMARCC1

Continues on next page

Table A.7 – *Continued from previous page*

Protein name	Gene name
SWI/SNF-related (...) regulator of chromatin subfamily E 1	SMARCE1
Stomatin-like protein 2 (mitochondrial)	STOML2
Nesprin-1	SYNE1
Vacuolar protein sorting-associated protein 28 homolog	VPS28

MacroD2 macrodomain etoposide over C-terminus etoposide

Alpha-aminoadipic semialdehyde dehydrogenase	ALDH7A1
Acidic leucine-rich nuclear (...) 32 family member B	ANP32B
Sarcoplasmic/endoplasmic reticulum calcium ATPase 2	ATP2A2
T-complex protein 1 subunit delta	CCT4
Putative ATP-dependent RNA helicase DHX30	DHX30
Enoyl-CoA hydratase (mitochondrial)	ECHS1
Protein FAM98B	FAM98B
Ras GTPase-activating protein-binding protein 2	G3BP2
Glutamine synthetase	GLUL
Ig heavy chain V-III region KOL	IGHV3-23
mRNA turnover protein 4 homolog	MRTO4
Polyadenylate-binding protein 2	PABPN1
Poly [ADP-ribose] polymerase 1	PARP1
26S proteasome non-ATPase regulatory subunit 4	PSMD4
40S ribosomal protein S29	RPS29
Survival of motor neuron-related-splicing factor 30	SMNDC1
Signal recognition particle subunit SRP68	SRP68
Transcription elongation factor SPT5	SUPT5H
UBX domain-containing protein 1	UBXN1
Cytochrome b-c1 complex subunit 1 (mitochondrial)	UQCRC1
Vacuolar protein sorting-associated protein 28 homolog	VPS28

MacroD2 macrodomain etoposide over C-terminus DMSO

Acidic leucine-rich nuclear (...) 32 family member A	ANP32A
Sarcoplasmic/endoplasmic reticulum calcium ATPase 2	ATP2A2
Coiled-coil domain-containing protein 80	CCDC80
dCTP pyrophosphatase 1	DCTPP1
Protein FAM98B	FAM98B
Guanine nucleotide-binding protein (...) subunit beta-1	GNB1
Histidine triad nucleotide-binding protein 1	HINT1
Histone H2A type 1-J	HIST1H2AJ
Keratin. type I cytoskeletal 18	KRT18
Prelamin-A/C;Lamin-A/C	LMNA
Methylcrotonoyl-CoA carboxylase beta chain (mitochondrial)	MCCC2

Continues on next page

Table A.7 – *Continued from previous page*

Protein name	Gene name
mRNA turnover protein 4 homolog	MRTO4
Poly [ADP-ribose] polymerase 1	PARP1
Programmed cell death protein 6	PDCD6
26S proteasome non-ATPase regulatory subunit 7	PSMD7
Histone-binding protein RBBP7	RBBP7
Serine hydroxymethyltransferase (mitochondrial)	SHMT2
SWI/SNF-related (...) regulator of chromatin subfamily E 1	SMARCE1
Survival of motor neuron-related-splicing factor 30	SMNDC1
Transcription elongation factor SPT5	SUPT5H
Nesprin-1	SYNE1
Vacuolar protein sorting-associated protein 28 homolog	VPS28

Table A.7: MacroD2 macrodomain etoposide - Overview of enriched proteins in *MacroD2 macrodomain etoposide* sample, when compared with *EGFP*, *MacroD2 macrodomain DMSO*, *MacroD2 C-terminus etoposide* and *MacroD2 C-terminus DMSO*.

Protein name	Gene name
MacroD2 macrodomain etoposide common proteins	
Alpha-aminoadipic semialdehyde dehydrogenase	ALDH7A1
Sarcoplasmic/endoplasmic reticulum calcium ATPase 2	ATP2A2
Coiled-coil domain-containing protein 80	CCDC80
dCTP pyrophosphatase 1	DCTPP1
Protein FAM98B	FAM98B
Ras GTPase-activating protein-binding protein 2	G3BP2
Histone H2A type 1-J	HIST1H2AJ
Keratin. type I cytoskeletal 18	KRT18
mRNA turnover protein 4 homolog	MRTO4
Poly [ADP-ribose] polymerase 1	PARP1
Programmed cell death protein 6	PDCD6
SWI/SNF-related (...) regulator of chromatin subfamily E 1	SMARCE1
Survival of motor neuron-related-splicing factor 30	SMNDC1
Transcription elongation factor SPT5	SUPT5H
Nesprin-1	SYNE1
Vacuolar protein sorting-associated protein 28 homolog	VPS28

Continues on next page

Table A.8 – *Continued from previous page*

Protein name	Gene name

Table A.8: MacroD2 macrodomain etoposide, common proteins - Overview of the proteins that are in common among the different "MacroD2 macrodomain etoposide" enrichments.

A.5 MacroD2 C-terminus DMSO protein list

Protein name	Gene name
MacroD2 C-terminus DMSO over EGFP	
Uncharacterized protein C9orf43	C9orf43
Calnexin	CANX
Cx9C motif-containing protein 4	CMC4
Dihydrolipoyl dehydrogenase (mitochondrial)	DLD
Aspartyl aminopeptidase	DNPEP
Elongation factor 1-beta	EEF1B2
Eukaryotic translation initiation factor 3 subunit K	EIF3K
Fragile X mental retardation syndrome-related protein 1	FXR1
Non-POU domain-containing octamer-binding protein	NONO
Proliferation-associated protein 2G4	PA2G4
28 kDa heat- and acid-stable phosphoprotein	PDAP1
Phosphatidylethanolamine-binding protein 1	PEBP1
60S ribosomal protein L31	RPL31
Protein SET	SET
Splicing factor 1	SF1
U1 small nuclear ribonucleoprotein C	SNRPC
Small nuclear ribonucleoprotein F	SNRPF
Spermidine synthase	SRM
Tubulin-specific chaperone A	TBCA
Transcription elongation regulator 1	TCERG1
Translation machinery-associated protein 16	TMA16
NEDD8-conjugating enzyme Ubc12	UBE2M
MacroD2 C-terminus DMSO over C-terminus etoposide	
Probable ATP-dependent RNA helicase DDX11	DDX11
Eukaryotic translation initiation factor 3 subunit K	EIF3K
Far upstream element-binding protein 1	FUBP1
Ig heavy chain V-III region KOL	IGHV3-23

Continues on next page

Table A.9 – *Continued from previous page*

Protein name	Gene name
Serine/threonine-protein phosphatase PP1-gamma (...)	PPP1CC
RNA-binding protein 4	RBM4
Transcription elongation regulator 1	TCERG1
Tubulin alpha-1C chain	TUBA1C
UBX domain-containing protein 1	UBXN1
MacroD2 C-terminus DMSO over macrodomain DMSO	
Uncharacterized protein C9orf43	C9orf43
Bifunctional coenzyme A synthase	COASY
Probable ATP-dependent RNA helicase DDX11	DDX11
Protein FAM133B	FAM133B
Heterogeneous nuclear ribonucleoprotein H2	HNRNPH2
LanC-like protein 2	LANCL2
Methylated-DNA–protein-cysteine methyltransferase	MGMT
Polymerase delta-interacting protein 3	POLDIP3
Pre-mRNA-splicing factor 38B	PRPF38B
Splicing factor 1	SF1
U1 small nuclear ribonucleoprotein C	SNRPC
MacroD2 C-terminus DMSO over macrodomain etoposide	
Rho GDP-dissociation inhibitor 1	ARHGDIA
Probable ATP-dependent RNA helicase DDX11	DDX11
Density-regulated protein	DENR
Isochorismatase domain-containing protein 2 (mitochondrial)	ISOC2
39S ribosomal protein L12 (mitochondrial)	MRPL12
Myotubularin-related protein 4	MTMR4
28 kDa heat- and acid-stable phosphoprotein	PDAP1
Arginine–tRNA ligase (cytoplasmic)	RARS
RNA-binding protein 4	RBM4
Splicing factor 1	SF1
U1 small nuclear ribonucleoprotein C	SNRPC
Spermidine synthase	SRM
Tubulin-specific chaperone A	TBCA
Transcription elongation regulator 1	TCERG1
Nucleolar protein of 40 kDa	ZCCHC17
Zinc finger CCHC domain-containing protein 3	ZCCHC3

Continues on next page

Table A.9 – *Continued from previous page*

Protein name	Gene name

Table A.9: MacroD2 C-terminus DMSO - Overview of enriched proteins in *MacroD2 C-terminus DMSO* sample, when compared with *EGFP*, *MacroD2 C-terminus etoposide*, *MacroD2 macrodomain DMSO* and *MacroD2 macrodomain etoposide*.

Protein name	Gene name
MacroD2 C-terminus DMSO common proteins	
Uncharacterized protein C9orf43	C9orf43
Probable ATP-dependent RNA helicase DDX11	DDX11
Eukaryotic translation initiation factor 3 subunit K	EIF3K
28 kDa heat- and acid-stable phosphoprotein	PDAP1
RNA-binding protein 4	RBM4
Splicing factor 1	SF1
U1 small nuclear ribonucleoprotein C	SNRPC
Spermidine synthase	SRM
Tubulin-specific chaperone A	TBCA
Transcription elongation regulator 1	TCERG1

Table A.10: MacroD2 C-terminus DMSO, common proteins - Overview of the proteins that are in common among the different "MacroD2 C-terminus DMSO" enrichments.

A.6 MacroD2 C-terminus etoposide protein list

Protein name	Gene name
MacroD2 C-terminus etoposide over EGFP	
Fructose-bisphosphate aldolase A	ALDOA
Calreticulin	CALR
T-complex protein 1 subunit epsilon	CCT5
Chromatin target of PRMT1 protein	CHTOP
Spliceosome RNA helicase DDX39B	DDX39B
Dihydrolipoyl dehydrogenase. mitochondrial	DLD
Enoyl-CoA hydratase. mitochondrial	ECHS1

Continues on next page

Table A.11 – *Continued from previous page*

Protein name	Gene name
Eukaryotic peptide chain release factor subunit 1	ETF1
Fatty acid-binding protein. epidermal	FABP5
Protein FAM98A	FAM98A
Peptidyl-prolyl cis-trans isomerase FKBP4	FKBP4
Glutathione S-transferase omega-1	GSTO1
Heat shock 70 kDa protein 12A	HSPA12A
Importin subunit beta-1	KPNB1
Protein ERGIC-53	LMAN1
Leucine-rich repeat-containing protein 59	LRRC59
Leucine zipper putative tumor suppressor 3	LZTS3
Microtubule-associated protein RP/EB family member 1	MAPRE1
DNA helicase MCM9	MCM9
Myotrophin	MTPN
Phosphatidylethanolamine-binding protein 1	PEBP1
DNA-directed RNA polymerase II subunit RPB3	POLR2C
Proteasome subunit alpha type-6	PSMA6
Ran-specific GTPase-activating protein	RANBP1
Histone-binding protein RBBP4	RBBP4
Replication protein A 32 kDa subunit	RPA2
Ribosomal RNA processing protein 1 homolog B	RRP1B
Staphylococcal nuclease domain-containing protein 1	SND1
U1 small nuclear ribonucleoprotein A	SNRPA
Small nuclear ribonucleoprotein F	SNRPF
Superoxide dismutase [Cu-Zn]	SOD1
Spectrin beta chain. non-erythrocytic 1	SPTBN1
Spermidine synthase	SRM
Lupus La protein	SSB
Single-stranded DNA-binding protein (mitochondrial)	SSBP1
Tubulin-specific chaperone A	TBCA
Transketolase	TKT
Translation machinery-associated protein 16	TMA16
Cellular tumor antigen p53	TP53
NEDD8-conjugating enzyme Ubc12	UBE2M

MacroD2 C-terminus etoposide over C-terminus DMSO

Eukaryotic peptide chain release factor subunit 1	ETF1
Peptidyl-prolyl cis-trans isomerase FKBP1A	FKBP1A
Guanine nucleotide-binding protein (...) subunit beta-2	GNB2
Heat shock 70 kDa protein 12A	HSPA12A
Protein lin-7 homolog C	LIN7C
Protein ERGIC-53	LMAN1

Continues on next page

Table A.11 – *Continued from previous page*

Protein name	Gene name
28S ribosomal protein S22. mitochondrial	MRPS22
Membrane-associated progesterone receptor component 1	PGRMC1
Proteasome subunit alpha type-6	PSMA6
U4/U6.U5 small nuclear ribonucleoprotein 27 kDa protein	SNRNP27
Spermidine synthase	SRM
Single-stranded DNA-binding protein (mitochondrial)	SSBP1

MacroD2 C-terminus etoposide over macrodomain etoposide

Rho GDP-dissociation inhibitor 1	ARHGDIA
ATP synthase subunit gamma (mitochondrial)	ATP5C1
ATPase inhibitor (mitochondrial)	ATPIF1
Complement C1q tumor necrosis factor-related protein 8	C1QTNF8
Calreticulin	CALR
Corneodesmosin	CDSN
Chromatin target of PRMT1 protein	CHTOP
Enoyl-CoA hydratase (mitochondrial)	ECHS1
Eukaryotic translation initiation factor 1A. X-chromosomal	EIF1AX
Peptidyl-prolyl cis-trans isomerase FKBP3	FKBP3
Peptidyl-prolyl cis-trans isomerase FKBP4	FKBP4
Guanine nucleotide-binding protein G(s) subunit alpha	GNAS
Guanine nucleotide-binding protein (...) subunit beta-2	GNB2
Heterogeneous nuclear ribonucleoprotein H2	HNRNPH2
Heat shock protein beta-1	HSPB1
Isochorismatase domain-containing protein 2 (mitochondrial)	ISOC2
Protein lin-7 homolog C	LIN7C
Protein ERGIC-53	LMAN1
Microtubule-associated protein RP/EB family member 1	MAPRE1
28S ribosomal protein S22. mitochondrial	MRPS22
Myotrophin	MTPN
Myosin regulatory light chain 12A	MYL12A
Cytosolic Fe-S cluster assembly factor NARFL	NARFL
Proliferating cell nuclear antigen	PCNA
Phosphatidylethanolamine-binding protein 1	PEBP1
Serine/threonine-protein phosphatase PGAM5 (mitochondrial)	PGAM5
DNA-directed RNA polymerase II subunit RPB3	POLR2C
Peroxiredoxin-4	PRDX4
Histone-binding protein RBBP4	RBBP4
Replication protein A 32 kDa subunit	RPA2
60S ribosomal protein L35a	RPL35A
60S ribosomal protein L38	RPL38
SUMO-activating enzyme subunit 1	SAE1

Continues on next page

Table A.11 – *Continued from previous page*

Protein name	**Gene name**
Staphylococcal nuclease domain-containing protein 1	SND1
U4/U6.U5 small nuclear ribonucleoprotein 27 kDa protein	SNRNP27
Spectrin alpha chain. non-erythrocytic 1	SPTAN1
Spermidine synthase	SRM
Serine/arginine-rich splicing factor 1	SRSF1
Single-stranded DNA-binding protein (mitochondrial)	SSBP1
Suppressor of G2 allele of SKP1 homolog	SUGT1
Transgelin-2	TAGLN2
Tubulin-specific chaperone A	TBCA
Transketolase	TKT
Cellular tumor antigen p53	TP53
Tumor susceptibility gene 101 protein	TSG101
Voltage-dependent anion-selective channel protein 3	VDAC3
Zinc finger CCHC domain-containing protein 3	ZCCHC3

MacroD2 C-terminus etoposide over macrodomain DMSO

Annexin A7	ANXA7
Arginase-1	ARG1
Corneodesmosin	CDSN
Coronin-1C	CORO1C
Desmocollin-1	DSC1
Protein FAM133B	FAM133B
Heat shock 70 kDa protein 12A	HSPA12A
Leucine zipper putative tumor suppressor 3	LZTS3
Methylated-DNA–protein-cysteine methyltransferase	MGMT
Myosin regulatory light chain 12A	MYL12A
PITH domain-containing protein 1	PITHD1
DNA-directed RNA polymerase II subunit RPB3	POLR2C
Histone-binding protein RBBP4	RBBP4
Replication protein A 32 kDa subunit	RPA2
Staphylococcal nuclease domain-containing protein 1	SND1
Stomatin-like protein 2 (mitochondrial)	STOML2
Transgelin-2	TAGLN2
Zinc finger CCHC domain-containing protein 3	ZCCHC3

Table A.11: MacroD2 C-terminus etoposide - Overview of enriched proteins in MacroD2 C-terminus etoposide sample, when compared with *EGFP*, *MacroD2 C-terminus DMSO*, *MacroD2 macrodomain etoposide* and *MacroD2 macrodomain DMSO*.

Protein name	Gene name
MacroD2 C-terminus etoposide common proteins	
Calreticulin	CALR
Corneodesmosin	CDSN
Chromatin target of PRMT1 protein	CHTOP
Enoyl-CoA hydratase (mitochondrial)	ECHS1
Eukaryotic peptide chain release factor subunit 1	ETF1
Peptidyl-prolyl cis-trans isomerase FKBP4	FKBP4
Guanine nucleotide-binding protein (...) subunit beta-2	GNB2
Heat shock 70 kDa protein 12A	HSPA12A
Protein lin-7 homolog C	LIN7C
Protein ERGIC-53	LMAN1
Leucine zipper putative tumor suppressor 3	LZTS3
Microtubule-associated protein RP/EB family member 1	MAPRE1
28S ribosomal protein S22. mitochondrial	MRPS22
Myotrophin	MTPN
Myosin regulatory light chain 12A	MYL12A
Phosphatidylethanolamine-binding protein 1	PEBP1
DNA-directed RNA polymerase II subunit RPB3	POLR2C
Proteasome subunit alpha type-6	PSMA6
Histone-binding protein RBBP4	RBBP4
Replication protein A 32 kDa subunit	RPA2
Staphylococcal nuclease domain-containing protein 1	SND1
U4/U6.U5 small nuclear ribonucleoprotein 27 kDa protein	SNRNP27
Spermidine synthase	SRM
Single-stranded DNA-binding protein (mitochondrial)	SSBP1
Transgelin-2	TAGLN2
Tubulin-specific chaperone A	TBCA
Transketolase	TKT
Cellular tumor antigen p53	TP53
Zinc finger CCHC domain-containing protein 3	ZCCHC3

Table A.12: MacroD2 C-terminus etoposide, common proteins - Overview of the proteins that are in common among the different "MacroD2 C-terminus etoposide" enrichments.

Appendix B

Lists of enriched biological terms

B.1 MacroD2 full-length DMSO

The list for the enrichment of biological terms for *all proteins* of "MacroD2 full-length DMSO" is enlisted below **(Table B.1)**.

Category	Term	FDR
GO.0044822	poly(A) RNA binding	1.95E-05
GO.0003723	RNA binding	2.61E-05
GO.0043230	extracellular organelle	0.00688
GO.0065010	extracellular membrane-bounded organelle	0.00688
GO.0070062	extracellular exosome	0.00688
GO.1903561	extracellular vesicle	0.00688
GO.0005852	eukaryotic translation initiation factor 3 complex	0.0241
GO.0043227	membrane-bounded organelle	0.0241
GO.0005730	nucleolus	0.0366
GO.0016282	eukaryotic 43S preinitiation complex	0.0366
GO.0031982	vesicle	0.0366
GO.0031988	membrane-bounded vesicle	0.0366
GO.0033290	eukaryotic 48S preinitiation complex	0.0366
GO.0070993	translation preinitiation complex	0.0383
GO.0030529	ribonucleoprotein complex	0.0403
GO.0044421	extracellular region part	0.0403

Table B.1: Biological term enrichment for all "MacroD2 full-length DMSO" proteins - Overview of enriched biological terms. Generated by STRING database.

B.2 MacroD2 full-length etoposide

The list for the enrichment of biological terms for *all proteins* of "MacroD2 full-length etoposide" is enlisted below **(Table B.2)**.

Category	Term	FDR
GO:0003723	RNA binding	1.22E-05
GO:0044822	poly(A) RNA binding	5.58E-05
GO:0003676	nucleic acid binding	0.0174

Table B.2: Biological term enrichment for all "MacroD2 full-length etoposide" proteins - Overview of enriched biological terms. Generated by STRING database.

B.3 MacroD2 macrodomain DMSO

The list for the enrichment of biological terms for *common proteins* of "MacroD2 macrodomain DMSO" is enlisted below **(Table B.3)**.

Category	Term	FDR
GO.0044822	poly(A) RNA binding	3.93E-12
GO.0003723	RNA binding	6.30E-11
GO.0043230	extracellular organelle	2.00E-10
GO.0065010	extracellular membrane-bounded organelle	2.00E-10
GO.0070062	extracellular exosome	2.00E-10
GO.1903561	extracellular vesicle	2.00E-10
GO.0031988	membrane-bounded vesicle	2.82E-08
GO.0006457	protein folding	3.90E-08
GO.0031982	vesicle	4.95E-08
GO.0032991	macromolecular complex	1.86E-06
GO.0044421	extracellular region part	2.79E-06
GO.0070013	intracellular organelle lumen	1.63E-05
GO.0044446	intracellular organelle part	1.77E-05
GO.0043233	organelle lumen	1.86E-05
GO.0031974	membrane-enclosed lumen	2.26E-05
GO.0044428	nuclear part	3.49E-05
GO.0005634	nucleus	5.59E-05
GO.0005576	extracellular region	5.83E-05
GO.0044422	organelle part	9.07E-05
GO.0003676	nucleic acid binding	0.000128
GO.0005488	binding	0.000225
GO.0005515	protein binding	0.000244

Continues on next page

Table B.3 – *Continued from previous page*

Category	Term	FDR
GO.0042470	melanosome	0.00052
GO.0031625	ubiquitin protein ligase binding	0.000577
GO.0044389	ubiquitin-like protein ligase binding	0.000577
GO.1901363	heterocyclic compound binding	0.000577
GO.0097159	organic cyclic compound binding	0.000585
GO.0043228	non-membrane-bounded organelle	0.000708
GO.0043232	intracellular non-membrane-bounded organelle	0.000708
GO.0044444	cytoplasmic part	0.000763
KEGG:04141	protein processing in endoplasmic reticulum	0.00109
GO.0031981	nuclear lumen	0.00132
GO.0030529	ribonucleoprotein complex	0.00176
GO.0016023	cytoplasmic membrane-bounded vesicle	0.00182
GO.0005829	cytosol	0.00194
GO.0005730	nucleolus	0.00234
GO.0043234	protein complex	0.00266
GO.0043231	intracellular membrane-bounded organelle	0.00275
GO.0043209	myelin sheath	0.00291
GO.0005737	cytoplasm	0.00317
GO.0031410	cytoplasmic vesicle	0.00321
GO.0043227	membrane-bounded organelle	0.00375
GO.0016607	nuclear speck	0.00491
GO.0015630	microtubule cytoskeleton	0.00592
GO.0044297	cell body	0.00592
GO.0051087	chaperone binding	0.00632
GO.0005832	chaperonin-containing T-complex	0.0106
GO.0005654	nucleoplasm	0.0112
GO.0005874	microtubule	0.0112
GO.0035770	ribonucleoprotein granule	0.0112
GO.0036464	cytoplasmic ribonucleoprotein granule	0.0112
GO.0097223	sperm part	0.0112
GO.0003674	molecular_function	0.0114
GO.0051082	unfolded protein binding	0.0114
GO.0002199	zona pellucida receptor complex	0.0114
GO.0012505	endomembrane system	0.013
GO.0005635	nuclear envelope	0.0134
GO.0005739	mitochondrion	0.0154
KEGG:05169	Epstein-Barr virus infection	0.0183
GO.0019953	sexual reproduction	0.024
GO.0044237	cellular metabolic process	0.024
GO.0044260	cellular macromolecule metabolic process	0.024
GO.0044703	multi-organism reproductive process	0.024

Continues on next page

Table B.3 – *Continued from previous page*

Category	Term	FDR
GO.1902582	single-organism intracellular transport	0.0263
GO.0043170	macromolecule metabolic process	0.0289
GO.0031967	organelle envelope	0.0315
GO.0031975	envelope	0.0315
GO.0045454	cell redox homeostasis	0.0345
GO.0005783	endoplasmic reticulum	0.0382
GO.0044451	nucleoplasm part	0.0382
GO.0030518	intracellular steroid hormone receptor signaling pathway	0.0406
GO.0044702	single organism reproductive process	0.0406
GO.0051704	multi-organism process	0.0406
GO.0061024	membrane organization	0.0406
GO.0006810	transport	0.0428
GO.0046907	intracellular transport	0.0428
GO.0031090	organelle membrane	0.0433
GO.0001669	acrosomal vesicle	0.0448
GO.0016604	nuclear body	0.0456
GO.0005856	cytoskeleton	0.0479

Table B.3: Biological term enrichment for "MacroD2 macrodomain DMSO" common proteins - Overview of enriched biological terms. Generated by STRING database

The list for the enrichment of biological terms for *all proteins* of "MacroD2 macrodomain DMSO" is enlisted below **(Table B.4)**. Since the list of terms was extremely long, here I show just a selection, with the lowest FDR values of each super-group, such as cellular localization or molecular pathway.

Category	Term	FDR
GO.0043230	extracellular organelle	4.69E-20
GO.0065010	extracellular membrane-bounded organelle	4.69E-20
GO.0070062	extracellular exosome	4.69E-20
GO.1903561	extracellular vesicle	4.69E-20
GO.0003723	RNA binding	1.56E-16
GO.0044822	poly(A) RNA binding	1.56E-16
GO.0031988	membrane-bounded vesicle	1.05E-15
GO.0031982	vesicle	3.79E-15
GO.0044421	extracellular region part	9.04E-14
GO.0044446	intracellular organelle part	6.18E-12
GO.0006457	protein folding	2.03E-11

Continues on next page

Table B.4 – *Continued from previous page*

Category	Term	FDR
GO.0032991	macromolecular complex	2.24E-11
GO.0005576	extracellular region	3.58E-11
GO.0005829	cytosol	4.97E-10
GO.0044422	organelle part	9.91E-10
GO.0005488	binding	7.80E-08
GO.1901363	heterocyclic compound binding	8.43E-08
GO.0097159	organic cyclic compound binding	1.20E-07
GO.0003676	nucleic acid binding	6.86E-07
GO.0006458	de novo protein folding	2.70E-06
GO.0046907	intracellular transport	2.70E-06
GO.1902582	single-organism intracellular transport	3.05E-06
GO.0043933	macromolecular complex subunit organization	1.22E-05
GO.0061024	membrane organization	1.22E-05
GO.0051084	de novo posttranslational protein folding	1.71E-05
GO.0006413	translational initiation	6.71E-05
GO.0006605	protein targeting	6.71E-05
GO.0006810	transport	6.71E-05
GO.0016482	cytoplasmic transport	6.71E-05
GO.0032984	macromolecular complex disassembly	6.71E-05
GO.0044237	cellular metabolic process	6.71E-05
GO.0051649	establishment of localization in cell	6.71E-05
GO.0071822	protein complex subunit organization	6.71E-05
GO.0016071	mRNA metabolic process	6.79E-05
GO.0071704	organic substance metabolic process	8.13E-05
GO.0046483	heterocycle metabolic process	9.76E-05
KEGG:03010	ribosome	0.000208
KEGG:01200	carbon metabolism	0.000283
KEGG:04141	protein processing in endoplasmic reticulum	0.00392
KEGG:05169	Epstein-Barr virus infection	0.00777
KEGG:00010	glycolysis / gluconeogenesis	0.0222
KEGG:00020	citrate cycle (TCA cycle)	0.0257
KEGG:01230	biosynthesis of amino acids	0.0264
KEGG:03040	spliceosome	0.0264
KEGG:00270	cysteine and methionine metabolism	0.0362
KEGG:00620	pyruvate metabolism	0.0443

Table B.4: Biological term enrichment for all the "MacroD2 macrodomain DMSO" proteins - Overview of enriched biological terms. Generated by STRING database

B.4 MacroD2 macrodomain etoposide

The list for the enrichment of biological terms for *common proteins* of "MacroD2 macrodomain etoposide" is enlisted below **(Table B.5)**.

Category	Term	FDR
GO.0003723;	RNA binding	0.00139
GO.0044822;	poly(A) RNA binding	0.00139

Table B.5: Biological term enrichment for the EGFP-MacroD2 macrodomain etoposide common proteins - Overview of enriched biological terms. Generated by STRING database

The list for the enrichment of biological terms for *all proteins* of "MacroD2 macrodomain etoposide" is enlisted below **(Table B.6)**.

Category	Term	FDR
GO.0044822	poly(A) RNA binding	4.50E-06
GO.0031974	membrane-enclosed lumen	5.77E-06
GO.0044446	intracellular organelle part	5.77E-06
GO.0070013	intracellular organelle lumen	5.77E-06
GO.0043233	organelle lumen	6.13E-06
GO.0003723	RNA binding	1.96E-05
GO.0043230	extracellular organelle	2.89E-05
GO.0065010	extracellular membrane-bounded organelle	2.89E-05
GO.0070062	extracellular exosome	2.89E-05
GO.1903561	extracellular vesicle	2.89E-05
GO.0031988	membrane-bounded vesicle	4.76E-05
GO.0031982	vesicle	6.54E-05
GO.0032991	macromolecular complex	6.54E-05
GO.0044422	organelle part	6.54E-05
GO.0043227	membrane-bounded organelle	0.000252
GO.0031981	nuclear lumen	0.000363
GO.0044428	nuclear part	0.000363
GO.0044421	extracellular region part	0.000386
GO.0005634	nucleus	0.00128
GO.0043229	intracellular organelle	0.00128
GO.0005654	nucleoplasm	0.00135
GO.0044877	macromolecular complex binding	0.00155
GO.0043231	intracellular membrane-bounded organelle	0.00183

Continues on next page

Table B.6 – *Continued from previous page*

Category	Term	FDR
GO.0070603	SWI/SNF superfamily-type complex	0.00197
GO.0005737	cytoplasm	0.00205
GO.0005759	mitochondrial matrix	0.00205
GO.0043234	protein complex	0.0021
GO.0005635	nuclear envelope	0.00247
GO.0005515	protein binding	0.00315
GO.0044444	cytoplasmic part	0.00385
GO.0005576	extracellular region	0.0041
GO.0005622	intracellular	0.00498
GO.0044424	intracellular part	0.00744
GO.0043226	organelle	0.00768
KEGG:01230	biosynthesis of amino acids	0.00863
GO.0005925	focal adhesion	0.0135
GO.0043228	non-membrane-bounded organelle	0.0135
GO.0043232	intracellular non-membrane-bounded organelle	0.0135
GO.0005924	cell-substrate adherens junction	0.0141
GO.0030055	cell-substrate junction	0.0143
GO.0045111	intermediate filament cytoskeleton	0.0146
GO.0005739	mitochondrion	0.0155
KEGG:00280	valine, leucine and isoleucine degradation	0.0182
KEGG:03050	proteasome	0.0182
GO.0031965	nuclear membrane	0.019
GO.0071564	npBAF complex	0.0204
GO.0016514	SWI/SNF complex	0.0226
GO.0071565	nBAF complex	0.0226
GO.0005912	adherens junction	0.023
GO.0070161	anchoring junction	0.0271
GO.0031967	organelle envelope	0.0285
GO.0031975	envelope	0.0285
GO.0005829	cytosol	0.0292
GO.0044429	mitochondrial part	0.037
GO.0005838	proteasome regulatory particle	0.0399

Table B.6: Biological term enrichment for all the "MacroD2 macrodomain etoposide" proteins - Overview of enriched biological terms. Generated by STRING database

B.5 MacroD2 C-terminus DMSO

The list for the enrichment of biological terms for *common proteins* of "MacroD2 C-terminus DMSO" is enlisted below **(Table B.7)**.

Category	Term	FDR
GO.0003723	RNA binding	0.000113
GO.0044822	poly(A) RNA binding	0.00939
GO.0003676	nucleic acid binding	0.0335

Table B.7: Biological term enrichment for "MacroD2 C-terminus DMSO" common proteins - Overview of enriched biological terms. Generated by STRING database.

The list for the enrichment of biological terms for *all proteins* of "MacroD2 C-terminus DMSO" is enlisted below **(Table B.8)**.

Category	Term	FDR
GO.0003723	RNA binding	3.77E-17
GO.0044822	poly(A) RNA binding	4.19E-10
GO.0003676	nucleic acid binding	5.20E-10
GO.1901363	heterocyclic compound binding	5.85E-06
GO.0097159	organic cyclic compound binding	6.45E-06
GO.0030529	ribonucleoprotein complex	2.05E-05
GO.0031981	nuclear lumen	2.05E-05
GO.0044428	nuclear part	2.05E-05
GO.0070013	intracellular organelle lumen	2.05E-05
GO.0043233	organelle lumen	2.09E-05
GO.0031974	membrane-enclosed lumen	2.27E-05
GO.0005634	nucleus	3.29E-05
GO.0005730	nucleolus	0.00107
GO.0044446	intracellular organelle part	0.00139
GO.0005654	nucleoplasm	0.00213
GO.0044422	organelle part	0.00715
GO.0032991	macromolecular complex	0.0151
GO.0043231	intracellular membrane-bounded organelle	0.0199
GO.0043603	cellular amide metabolic process	0.0295
GO.0005681	spliceosomal complex	0.0297
GO.0008380	RNA splicing	0.0306
GO.0006397	mRNA processing	0.0374
GO.0016071	mRNA metabolic process	0.0374
GO.0022618	ribonucleoprotein complex assembly	0.0374
GO.0034622	cellular macromolecular complex assembly	0.0374
GO.0071826	ribonucleoprotein complex subunit organization	0.0374

Continues on next page

Table B.8 – *Continued from previous page*

Category	Term	FDR
KEGG:03040	spliceosome	0.0379
GO.0005685	U1 snRNP	0.0422
GO.0006518	peptide metabolic process	0.0442
GO.0043227	membrane-bounded organelle	0.0496

Table B.8: Biological term enrichment for all the "MacroD2 C-terminus DMSO" proteins - Overview of enriched biological terms. Generated by STRING database

B.6 MacroD2 C-terminus etoposide

The list for the enrichment of biological terms for *common proteins* of "MacroD2 C-terminus etoposide" is enlisted below **(Table B.9)**.

Category	Term	FDR
GO.0043230	extracellular organelle	0.000239
GO.0065010	extracellular membrane-bounded organelle	0.000239
GO.0070062	extracellular exosome	0.000239
GO.1903561	extracellular vesicle	0.000239
GO.0051095	regulation of helicase activity	0.00295
GO.0003723	RNA binding	0.00539
GO.0044822	poly(A) RNA binding	0.00539
GO.0044421	extracellular region part	0.00607
GO.0031988	membrane-bounded vesicle	0.00924
GO.0032991	macromolecular complex	0.00924
GO.0043228	non-membrane-bounded organelle	0.00924
GO.0043232	intracellular non-membrane-bounded organelle	0.00924
GO.0031982	vesicle	0.0103
GO.0005576	extracellular region	0.0219
GO.0003676	nucleic acid binding	0.0306

Table B.9: Biological term enrichment for "MacroD2 C-terminus etoposide" common proteins - Overview of enriched biological terms. Generated by STRING database

The list for the enrichment of biological terms for *all proteins* of "MacroD2 C-terminus etoposide" is enlisted below **(Table B.10)**.

Category	Term	FDR
GO.0043230	extracellular organelle	3.22E-12
GO.0065010	extracellular membrane-bounded organelle	3.22E-12
GO.0070062	extracellular exosome	3.22E-12
GO.1903561	extracellular vesicle	3.22E-12
GO.0044822	poly(A) RNA binding	2.35E-11
GO.0003723	RNA binding	1.56E-10
GO.0031988	membrane-bounded vesicle	1.02E-08
GO.0031982	vesicle	2.13E-08
GO.0005829	cytosol	3.59E-07
GO.0003676	nucleic acid binding	2.13E-06
GO.0005634	nucleus	2.89E-06
GO.0044421	extracellular region part	3.12E-06
GO.0032991	macromolecular complex	3.12E-05
GO.0005576	extracellular region	5.94E-05
GO.0030017	sarcomere	5.97E-05
GO.0044446	intracellular organelle part	6.94E-05
GO.0044449	contractile fiber part	8.78E-05
GO.0030016	myofibril	0.000113
GO.0031974	membrane-enclosed lumen	0.000113
GO.0043292	contractile fiber	0.000149
GO.0044428	nuclear part	0.000211
GO.0043228	non-membrane-bounded organelle	0.000215
GO.0043232	intracellular non-membrane-bounded organelle	0.000215
GO.0043233	organelle lumen	0.000215
GO.1901363	heterocyclic compound binding	0.000424
GO.0070013	intracellular organelle lumen	0.000424
GO.0097159	organic cyclic compound binding	0.000491
GO.0044422	organelle part	0.000706
GO.0043231	intracellular membrane-bounded organelle	0.000943
GO.0005739	mitochondrion	0.00122
GO.0030529	ribonucleoprotein complex	0.00159
GO.0043227	membrane-bounded organelle	0.00163
GO.0043209	myelin sheath	0.00197
GO.0043234	protein complex	0.00342
GO.0005759	mitochondrial matrix	0.00367
GO.0005488	binding	0.00412
GO.0044429	mitochondrial part	0.00441
GO.0032437	cuticular plate	0.00461
GO.0031981	nuclear lumen	0.00534
GO.0031674	I band	0.00541
GO.0070161	anchoring junction	0.0121
GO.0006457	protein folding	0.0147

Continues on next page

Table B.10 – *Continued from previous page*

Category	Term	FDR
KEGG:03040	spliceosome	0.0181
KEGG:03430	mismatch repair	0.0181
GO.0005527	macrolide binding	0.0194
GO.0005528	FK506 binding	0.0194
GO.0031967	organelle envelope	0.0195
GO.0031975	envelope	0.0195
GO.0022417	protein maturation by protein folding	0.0219
GO.0051095	regulation of helicase activity	0.0219
GO.0008091	spectrin	0.023
GO.0003684	damaged DNA binding	0.0293
GO.0005515	protein binding	0.0293
GO.0051087	chaperone binding	0.0293
GO.0019899	enzyme binding	0.0316
KEGG:03030	DNA replication	0.0329
GO.0005657	replication fork	0.0337
GO.0030018	Z disc	0.0337
KEGG:01200	carbon metabolism	0.0414
KEGG:05016	Huntington disease	0.0414
GO.0005912	adherens junction	0.0451
GO.0005730	nucleolus	0.0469
GO.0005685	U1 snRNP	0.0496
GO.0098827	endoplasmic reticulum subcompartment	0.0496

Table B.10: Biological term enrichment for all "MacroD2 C-terminus etoposide" proteins - Overview of enriched biological terms. Generated by STRING database

References

[1] Deribe, Y. L., Pawson, T. & Dikic, I. Post-translational modifications in signal integration. *Nat Struct Mol Biol* **17**, 666–672 (2010). URL http://dx.doi.org/10.1038/nsmb.1842. 2, 3

[2] Fischer, E. H. & Krebs, E. G. Conversion of phosphorylase b to phosphorylase a in muscle extracts. *J Biol Chem* **216**, 121–132 (1955). 2

[3] Krebs, E. G. & Beavo, J. A. Phosphorylation-dephosphorylation of enzymes. *Annu Rev Biochem* **48**, 923–959 (1979). URL http://dx.doi.org/10.1146/annurev.bi.48.070179.004423. 2

[4] Baker, L. A., Allis, C. D. & Wang, G. G. Phd fingers in human diseases: disorders arising from misinterpreting epigenetic marks. *Mutat Res* **647**, 3–12 (2008). URL http://dx.doi.org/10.1016/j.mrfmmm.2008.07.004. 2, 4

[5] Dantuma, N. P. & van Attikum, H. Spatiotemporal regulation of posttranslational modifications in the dna damage response. *EMBO J* **35**, 6–23 (2016). URL http://dx.doi.org/10.15252/embj.201592595. 2

[6] Ryu, K. W., Kim, D.-S. & Kraus, W. L. New facets in the regulation of gene expression by adp-ribosylation and poly(adp-ribose) polymerases. *Chem Rev* **115**, 2453–2481 (2015). URL http://dx.doi.org/10.1021/cr5004248. 4, 6, 12, 113

[7] Caldecott, K. W. Protein adp-ribosylation and the cellular response to dna strand breaks. *DNA Repair (Amst)* **19**, 108–113 (2014). URL http://dx.doi.org/10.1016/j.dnarep.2014.03.021. 4

[8] Golia, B., Singh, H. R. & Timinszky, G. Poly-adp-ribosylation signaling during dna damage repair. *Front Biosci (Landmark Ed)* **20**, 440–457 (2015). 4, 10, 21, 24, 26, 113

[9] Asher, G. & Sassone-Corsi, P. Time for food: the intimate interplay between nutrition, metabolism, and the circadian clock. *Cell* **161**, 84–92 (2015). URL http://dx.doi.org/10.1016/j.cell.2015.03.015. 4

[10] Bock, F. J., Todorova, T. T. & Chang, P. Rna regulation by poly(adp-ribose) polymerases. *Mol Cell* **58**, 959–969 (2015). URL http://dx.doi.org/10.1016/j.molcel.2015.01.037. 4, 29, 135

[11] Bütepage, M., Eckei, L., Verheugd, P. & Lüscher, B. Intracellular mono-adp-ribosylation in signaling and disease. *Cells* **4**, 569–595 (2015). URL http://dx.doi.org/10.3390/cells4040569. 4, 6, 27, 28, 30

[12] Barkauskaite, E., Jankevicius, G., Ladurner, A. G., Ahel, I. & Timinszky, G. The recognition and removal of cellular poly(adp-ribose) signals. *FEBS J* **280**, 3491–3507 (2013). URL http://dx.doi.org/10.1111/febs.12358. 5, 10, 11, 16

[13] Hottiger, M. O., Hassa, P. O., Lüscher, B., Schüler, H. & Koch-Nolte, F. Toward a unified nomenclature for mammalian adp-ribosyltransferases. *Trends Biochem Sci* **35**, 208–219 (2010). URL http://dx.doi.org/10.1016/j.tibs.2009.12.003. 6, 8

[14] Frye, R. A. Characterization of five human cdnas with homology to the yeast sir2 gene: Sir2-like proteins (sirtuins) metabolize nad and may have protein adp-ribosyltransferase activity. *Biochem Biophys Res Commun* **260**, 273–279 (1999). URL http://dx.doi.org/10.1006/bbrc.1999.0897. 6, 10

[15] Tanny, J. C., Dowd, G. J., Huang, J., Hilz, H. & Moazed, D. An enzymatic activity in the yeast sir2 protein that is essential for gene silencing. *Cell* **99**, 735–745 (1999). 6

[16] Haigis, M. C. *et al.* Sirt4 inhibits glutamate dehydrogenase and opposes the effects of calorie restriction in pancreatic beta cells. *Cell* **126**, 941–954 (2006). URL http://dx.doi.org/10.1016/j.cell.2006.06.057. 6, 10, 16

[17] Liszt, G., Ford, E., Kurtev, M. & Guarente, L. Mouse sir2 homolog sirt6 is a nuclear adp-ribosyltransferase. *J Biol Chem* **280**, 21313–21320 (2005). URL http://dx.doi.org/10.1074/jbc.M413296200. 6

[18] Cervantes-Laurean, D., Loflin, P. T., Minter, D. E., Jacobson, E. L. & Jacobson, M. K. Protein modification by adp-ribose via acid-labile linkages. *J Biol Chem* **270**, 7929–7936 (1995). 6

[19] Altmeyer, M., Messner, S., Hassa, P. O., Fey, M. & Hottiger, M. O. Molecular mechanism of poly(adp-ribosyl)ation by parp1 and identification of lysine residues as adp-ribose acceptor sites. *Nucleic Acids Res* **37**, 3723–3738 (2009). URL http://dx.doi.org/10.1093/nar/gkp229. 6

[20] Laing, S., Koch-Nolte, F., Haag, F. & Buck, F. Strategies for the identification of arginine adp-ribosylation sites. *J Proteomics* **75**, 169–176 (2011). URL http://dx.doi.org/10.1016/j.jprot.2011.07.003. 6

[21] Manning, D. R., Fraser, B. A., Kahn, R. A. & Gilman, A. G. Adp-ribosylation of transducin by islet-activation protein. identification of asparagine as the site of adp-ribosylation. *J Biol Chem* **259**, 749–756 (1984). 6

[22] McDonald, L. J. & Moss, J. Enzymatic and nonenzymatic adp-ribosylation of cysteine. *Mol Cell Biochem* **138**, 221–226 (1994). 6

[23] Oppenheimer, N. J. & Bodley, J. W. Diphtheria toxin. site and configuration of adp-ribosylation of diphthamide in elongation factor 2. *J Biol Chem* **256**, 8579–8581 (1981). 6

[24] Smith, J. A. & Stocken, L. A. Chemical and metabolic properties of adenosine diphosphate ribose derivatives of nuclear proteins. *Biochem J* **147**, 523–529 (1975). 6

REFERENCES

[25] Miwa, M., Saikawa, N., Yamaizumi, Z., Nishimura, S. & Sugimura, T. Structure of poly(adenosine diphosphate ribose): identification of 2'-[1"-ribosyl-2"-(or 3"-)(1"'-ribosyl)]adenosine-5',5",5"'-tris(phosphate) as a branch linkage. *Proc Natl Acad Sci U S A* **76**, 595–599 (1979). 6

[26] Juarez-Salinas, H., Levi, V., Jacobson, E. L. & Jacobson, M. K. Poly(adp-ribose) has a branched structure in vivo. *J Biol Chem* **257**, 607–609 (1982). 6

[27] Kiehlbauch, C. C., Aboul-Ela, N., Jacobson, E. L., Ringer, D. P. & Jacobson, M. K. High resolution fractionation and characterization of adp-ribose polymers. *Anal Biochem* **208**, 26–34 (1993). 6

[28] Martello, R., Mangerich, A., Sass, S., Dedon, P. C. & Bürkle, A. Quantification of cellular poly(adp-ribosyl)ation by stable isotope dilution mass spectrometry reveals tissue- and drug-dependent stress response dynamics. *ACS Chem Biol* **8**, 1567–1575 (2013). URL http://dx.doi.org/10.1021/cb400170b. 6

[29] Langelier, M.-F. & Pascal, J. M. Parp-1 mechanism for coupling dna damage detection to poly(adp-ribose) synthesis. *Curr Opin Struct Biol* **23**, 134–143 (2013). URL http://dx.doi.org/10.1016/j.sbi.2013.01.003. 9, 24

[30] Ikejima, M. *et al.* The zinc fingers of human poly(adp-ribose) polymerase are differentially required for the recognition of dna breaks and nicks and the consequent enzyme activation. other structures recognize intact dna. *J Biol Chem* **265**, 21907–21913 (1990). 9

[31] Eustermann, S. *et al.* The dna-binding domain of human parp-1 interacts with dna single-strand breaks as a monomer through its second zinc finger. *J Mol Biol* **407**, 149–170 (2011). URL http://dx.doi.org/10.1016/j.jmb.2011.01.034. 9

[32] Langelier, M.-F., Planck, J. L., Roy, S. & Pascal, J. M. Crystal structures of poly(adp-ribose) polymerase-1 (parp-1) zinc fingers bound to dna: structural and functional insights into dna-dependent parp-1 activity. *J Biol Chem* **286**, 10690–10701 (2011). URL http://dx.doi.org/10.1074/jbc.M110.202507. 9

[33] Ali, A. A. E. *et al.* The zinc-finger domains of parp1 cooperate to recognize dna strand breaks. *Nat Struct Mol Biol* **19**, 685–692 (2012). URL http://dx.doi.org/10.1038/nsmb.2335. 9

[34] Langelier, M.-F., Planck, J. L., Roy, S. & Pascal, J. M. Structural basis for dna damage-dependent poly(adp-ribosyl)ation by human parp-1. *Science* **336**, 728–732 (2012). URL http://dx.doi.org/10.1126/science.1216338. 9

[35] Langelier, M.-F., Riccio, A. A. & Pascal, J. M. Parp-2 and parp-3 are selectively activated by 5' phosphorylated dna breaks through an allosteric regulatory mechanism shared with parp-1. *Nucleic Acids Res* **42**, 7762–7775 (2014). URL http://dx.doi.org/10.1093/nar/gku474. 9

[36] Kassner, I. *et al.* Set7/9-dependent methylation of artd1 at k508 stimulates poly-adp-ribose formation after oxidative stress. *Open Biol* **3**, 120173 (2013). URL http://dx.doi.org/10.1098/rsob.120173. 10

[37] Mao, Z. *et al.* Sirt6 promotes dna repair under stress by activating parp1. *Science* **332**, 1443–1446 (2011). URL http://dx.doi.org/10.1126/science.1202723. 10

[38] Mao, Z. *et al.* Sirtuin 6 (sirt6) rescues the decline of homologous recombination repair during replicative senescence. *Proc Natl Acad Sci U S A* **109**, 11800–11805 (2012). URL http://dx.doi.org/10.1073/pnas.1200583109. 10

[39] Wright, R. H. G. *et al.* Cdk2-dependent activation of parp-1 is required for hormonal gene regulation in breast cancer cells. *Genes Dev* **26**, 1972–1983 (2012). URL http://dx.doi.org/10.1101/gad.193193.112. 10

[40] Ariumi, Y. *et al.* In vivo phosphorylation of poly(adp-ribose) polymerase is independent of its activation. *FEBS Lett* **436**, 288–292 (1998). 10, 24, 25

[41] Ariumi, Y. *et al.* Suppression of the poly(adp-ribose) polymerase activity by dna-dependent protein kinase in vitro. *Oncogene* **18**, 4616–4625 (1999). URL http://dx.doi.org/10.1038/sj.onc.1202823. 10, 24, 25

[42] Matsuoka, S. *et al.* Atm and atr substrate analysis reveals extensive protein networks responsive to dna damage. *Science* **316**, 1160–1166 (2007). URL http://dx.doi.org/10.1126/science.1140321. 10, 34, 38, 111

[43] Gagné, J.-P. *et al.* Proteomic investigation of phosphorylation sites in poly(adp-ribose) polymerase-1 and poly(adp-ribose) glycohydrolase. *J Proteome Res* **8**, 1014–1029 (2009). URL http://dx.doi.org/10.1021/pr800810n. 10

[44] Bensimon, A. *et al.* Atm-dependent and -independent dynamics of the nuclear phosphoproteome after dna damage. *Sci Signal* **3**, rs3 (2010). URL http://dx.doi.org/10.1126/scisignal.2001034. 10, 34, 111

[45] Barkauskaite, E., Jankevicius, G. & Ahel, I. Structures and mechanisms of enzymes employed in the synthesis and degradation of parp-dependent protein adp-ribosylation. *Mol Cell* **58**, 935–946 (2015). URL http://dx.doi.org/10.1016/j.molcel.2015.05.007. 10

[46] Kleine, H. *et al.* Substrate-assisted catalysis by parp10 limits its activity to mono-adp-ribosylation. *Mol Cell* **32**, 57–69 (2008). URL http://dx.doi.org/10.1016/j.molcel.2008.08.009. 10

[47] Kupis, W., Pałyga, J., Tomal, E. & Niewiadomska, E. The role of sirtuins in cellular homeostasis. *J Physiol Biochem* **72**, 371–380 (2016). URL http://dx.doi.org/10.1007/s13105-016-0492-6. 10

[48] Pleschke, J. M., Kleczkowska, H. E., Strohm, M. & Althaus, F. R. Poly(adp-ribose) binds to specific domains in dna damage checkpoint proteins. *J Biol Chem* **275**, 40974–40980 (2000). URL http://dx.doi.org/10.1074/jbc.M006520200. 10, 23, 24

[49] Aravind, L. The wwe domain: a common interaction module in protein ubiquitination and adp ribosylation. *Trends Biochem Sci* **26**, 273–275 (2001). 10

[50] Karras, G. I. *et al.* The macro domain is an adp-ribose binding module. *EMBO J* **24**, 1911–1920 (2005). URL http://dx.doi.org/10.1038/sj.emboj.7600664. 10, 13, 16

[51] Ahel, I. *et al.* Poly(adp-ribose)-binding zinc finger motifs in dna repair/checkpoint proteins. *Nature* **451**, 81–85 (2008). URL http://dx.doi.org/10.1038/nature06420. 10, 24

[52] Li, M. & Yu, X. Function of brca1 in the dna damage response is mediated by adp-ribosylation. *Cancer Cell* **23**, 693–704 (2013). URL http://dx.doi.org/10.1016/j.ccr.2013.03.025. 11, 24

[53] Li, M., Lu, L.-Y., Yang, C.-Y., Wang, S. & Yu, X. The fha and brct domains recognize adp-ribosylation during dna damage response. *Genes Dev* **27**, 1752–1768 (2013). URL http://dx.doi.org/10.1101/gad.226357.113. 11, 22, 24

[54] Zhang, F., Chen, Y., Li, M. & Yu, X. The oligonucleotide/oligosaccharide-binding fold motif is a poly(adp-ribose)-binding domain that mediates dna damage response. *Proc Natl Acad Sci U S A* **111**, 7278–7283 (2014). URL http://dx.doi.org/10.1073/pnas.1318367111. 11, 24

[55] Wang, Z. *et al.* Recognition of the iso-adp-ribose moiety in poly(adp-ribose) by wwe domains suggests a general mechanism for poly(adp-ribosyl)ation-dependent ubiquitination. *Genes Dev* **26**, 235–240 (2012). URL http://dx.doi.org/10.1101/gad.182618.111. 11

[56] Jankevicius, G. *et al.* A family of macrodomain proteins reverses cellular mono-adp-ribosylation. *Nat Struct Mol Biol* **20**, 508–514 (2013). URL http://dx.doi.org/10.1038/nsmb.2523. 11, 13, 16, 17, 18, 31, 47, 51, 62, 75, 105, 113, 121, 124, 139, 143

[57] Miwa, M., Tanaka, M., Matsushima, T. & Sugimura, T. Purification and properties of glycohydrolase from calf thymus splitting ribose-ribose linkages of poly(adenosine diphosphate ribose). *J Biol Chem* **249**, 3475–3482 (1974). 11

[58] Oka, S., Kato, J. & Moss, J. Identification and characterization of a mammalian 39-kda poly(adp-ribose) glycohydrolase. *J Biol Chem* **281**, 705–713 (2006). URL http://dx.doi.org/10.1074/jbc.M510290200. 11

[59] Barkauskaite, E. *et al.* Visualization of poly(adp-ribose) bound to parg reveals inherent balance between exo- and endo-glycohydrolase activities. *Nat Commun* **4**, 2164 (2013). URL http://dx.doi.org/10.1038/ncomms3164. 11

[60] Ikejima, M. & Gill, D. M. Poly(adp-ribose) degradation by glycohydrolase starts with an endonucleolytic incision. *J Biol Chem* **263**, 11037–11040 (1988). 11

[61] Mashimo, M., Kato, J. & Moss, J. Structure and function of the arh family of adp-ribosyl-acceptor hydrolases. *DNA Repair (Amst)* **23**, 88–94 (2014). URL http://dx.doi.org/10.1016/j.dnarep.2014.03.005. 11

[62] Slade, D. *et al.* The structure and catalytic mechanism of a poly(adp-ribose) glycohydrolase. *Nature* **477**, 616–620 (2011). URL http://dx.doi.org/10.1038/nature10404. 11, 13

[63] Mueller-Dieckmann, C. *et al.* The structure of human adp-ribosylhydrolase 3 (arh3) provides insights into the reversibility of protein adp-ribosylation. *Proc Natl Acad Sci U S A* **103**, 15026–15031 (2006). URL http://dx.doi.org/10.1073/pnas.0606762103. 11

[64] Koh, D. W. *et al.* Failure to degrade poly(adp-ribose) causes increased sensitivity to cytotoxicity and early embryonic lethality. *Proc Natl Acad Sci U S A* **101**, 17699–17704 (2004). URL http://dx.doi.org/10.1073/pnas.0406182101. 11

[65] Hanai, S. *et al.* Loss of poly(adp-ribose) glycohydrolase causes progressive neurodegeneration in drosophila melanogaster. *Proc Natl Acad Sci U S A* **101**, 82–86 (2004). URL http://dx.doi.org/10.1073/pnas.2237114100. 11, 16

[66] Mashimo, M., Kato, J. & Moss, J. Adp-ribosyl-acceptor hydrolase 3 regulates poly (adp-ribose) degradation and cell death during oxidative stress. *Proc Natl Acad Sci U S A* **110**, 18964–18969 (2013). URL http://dx.doi.org/10.1073/pnas.1312783110. 11

[67] Rosenthal, F. *et al.* Macrodomain-containing proteins are new mono-adp-ribosylhydrolases. *Nat Struct Mol Biol* **20**, 502–507 (2013). URL http://dx.doi.org/10.1038/nsmb.2521. 13, 16, 17, 18, 19, 28, 31, 47, 51, 62, 105, 121, 139, 143

[68] Sharifi, R. *et al.* Deficiency of terminal adp-ribose protein glycohydrolase targ1/c6orf130 in neurodegenerative disease. *EMBO J* **32**, 1225–1237 (2013). URL http://dx.doi.org/10.1038/emboj.2013.51. 13, 16, 17, 31, 47

[69] Moss, J., Jacobson, M. K. & Stanley, S. J. Reversibility of arginine-specific mono(adp-ribosyl)ation: identification in erythrocytes of an adp-ribose-l-arginine cleavage enzyme. *Proc Natl Acad Sci U S A* **82**, 5603–5607 (1985). 13

[70] Rack, J. G. M., Perina, D. & Ahel, I. Macrodomains: Structure, function, evolution, and catalytic activities. *Annu Rev Biochem* (2016). URL http://dx.doi.org/10.1146/annurev-biochem-060815-014935. 13, 14, 15, 16

[71] Pehrson, J. R. & Fried, V. A. Macroh2a, a core histone containing a large nonhistone region. *Science* **257**, 1398–1400 (1992). 13

[72] Martzen, M. R. *et al.* A biochemical genomics approach for identifying genes by the activity of their products. *Science* **286**, 1153–1155 (1999). 13

[73] Kustatscher, G., Hothorn, M., Pugieux, C., Scheffzek, K. & Ladurner, A. G. Splicing regulates nad metabolite binding to histone macroh2a. *Nat Struct Mol Biol* **12**, 624–625 (2005). URL http://dx.doi.org/10.1038/nsmb956. 16

[74] Chen, D. *et al.* Identification of macrodomain proteins as novel o-acetyl-adp-ribose deacetylases. *J Biol Chem* **286**, 13261–13271 (2011). URL http://dx.doi.org/10.1074/jbc.M110.206771. 16

[75] Peterson, F. C. *et al.* Orphan macrodomain protein (human c6orf130) is an o-acyl-adp-ribose deacylase: solution structure and catalytic properties. *J Biol Chem* **286**, 35955–35965 (2011). URL http://dx.doi.org/10.1074/jbc.M111.276238. 16

REFERENCES

[76] Neuvonen, M. & Ahola, T. Differential activities of cellular and viral macro domain proteins in binding of adp-ribose metabolites. *J Mol Biol* **385**, 212–225 (2009). URL http://dx.doi.org/10.1016/j.jmb.2008.10.045. 16

[77] Han, W.-D. *et al.* Estrogenically regulated lrp16 interacts with estrogen receptor alpha and enhances the receptor's transcriptional activity. *Endocr Relat Cancer* **14**, 741–753 (2007). URL http://dx.doi.org/10.1677/ERC-06-0082. 17, 20, 138

[78] Yang, J. *et al.* The single-macro domain protein lrp16 is an essential cofactor of androgen receptor. *Endocr Relat Cancer* **16**, 139–153 (2009). URL http://dx.doi.org/10.1677/ERC-08-0150. 17, 138

[79] Wu, Z. *et al.* Lrp16 integrates into nf-κb transcriptional complex and is required for its functional activation. *PLoS One* **6**, e18157 (2011). URL http://dx.doi.org/10.1371/journal.pone.0018157. 17

[80] Perina, D. *et al.* Distribution of protein poly(adp-ribosyl)ation systems across all domains of life. *DNA Repair (Amst)* **23**, 4–16 (2014). URL http://dx.doi.org/10.1016/j.dnarep.2014.05.003. 17

[81] Timinszky, G. *et al.* A macrodomain-containing histone rearranges chromatin upon sensing parp1 activation. *Nat Struct Mol Biol* **16**, 923–929 (2009). URL http://dx.doi.org/10.1038/nsmb.1664. 17, 31

[82] Jope, R. S. & Johnson, G. V. W. The glamour and gloom of glycogen synthase kinase-3. *Trends Biochem Sci* **29**, 95–102 (2004). URL http://dx.doi.org/10.1016/j.tibs.2003.12.004. 20

[83] Rayasam, G. V., Tulasi, V. K., Sodhi, R., Davis, J. A. & Ray, A. Glycogen synthase kinase 3: more than a namesake. *Br J Pharmacol* **156**, 885–898 (2009). URL http://dx.doi.org/10.1111/j.1476-5381.2008.00085.x. 20

[84] Feijs, K. L. *et al.* Artd10 substrate identification on protein microarrays: regulation of gsk3β by mono-adp-ribosylation. *Cell Commun Signal* **11**, 5 (2013). URL http://dx.doi.org/10.1186/1478-811X-11-5. 20, 28

[85] Maas, N. M. C. *et al.* The c20orf133 gene is disrupted in a patient with kabuki syndrome. *J Med Genet* **44**, 562–569 (2007). URL http://dx.doi.org/10.1136/jmg.2007.049510. 20

[86] Kuniba, H. *et al.* Lack of c20orf133 and flrt3 mutations in 43 patients with kabuki syndrome in japan. *J Med Genet* **45**, 479–480 (2008). URL http://dx.doi.org/10.1136/jmg.2008.058503. 20

[87] Anney, R. *et al.* A genome-wide scan for common alleles affecting risk for autism. *Hum Mol Genet* **19**, 4072–4082 (2010). URL http://dx.doi.org/10.1093/hmg/ddq307. 20

[88] Curran, S. *et al.* No association between a common single nucleotide polymorphism, rs4141463, in the macrod2 gene and autism spectrum disorder. *Am J Med Genet B Neuropsychiatr Genet* **156B**, 633–639 (2011). URL http://dx.doi.org/10.1002/ajmg.b.31201. 20

[89] Lionel, A. C. *et al.* Rare copy number variation discovery and cross-disorder comparisons identify risk genes for adhd. *Sci Transl Med* **3**, 95ra75 (2011). URL http://dx.doi.org/10.1126/scitranslmed.3002464. 20

[90] Prandini, P. *et al.* The association of rs4307059 and rs35678 markers with autism spectrum disorders is replicated in italian families. *Psychiatr Genet* **22**, 177–181 (2012). URL http://dx.doi.org/10.1097/YPG.0b013e32835185c9. 20

[91] Cheng, Y., Quinn, J. F. & Weiss, L. A. An eqtl mapping approach reveals that rare variants in the sema5a regulatory network impact autism risk. *Hum Mol Genet* **22**, 2960–2972 (2013). URL http://dx.doi.org/10.1093/hmg/ddt150. 20

[92] Tsang, K. M. *et al.* A genome-wide survey of transgenerational genetic effects in autism. *PLoS One* **8**, e76978 (2013). URL http://dx.doi.org/10.1371/journal.pone.0076978. 20

[93] Jones, R. M. *et al.* Macrod2 gene associated with autistic-like traits in a general population sample. *Psychiatr Genet* **24**, 241–248 (2014). URL http://dx.doi.org/10.1097/YPG.0000000000000052. 20

[94] Bradley, W. E. C. *et al.* Hotspots of large rare deletions in the human genome. *PLoS One* **5**, e9401 (2010). URL http://dx.doi.org/10.1371/journal.pone.0009401. 20

[95] Perlis, R. H., Ruderfer, D., Hamilton, S. P. & Ernst, C. Copy number variation in subjects with major depressive disorder who attempted suicide. *PLoS One* **7**, e46315 (2012). URL http://dx.doi.org/10.1371/journal.pone.0046315. 20

[96] Kohannim, O. *et al.* Discovery and replication of gene influences on brain structure using lasso regression. *Front Neurosci* **6**, 115 (2012). 20

[97] Jahanshad, N. *et al.* Genome-wide scan of healthy human connectome discovers spon1 gene variant influencing dementia severity. *Proc Natl Acad Sci U S A* **110**, 4768–4773 (2013). URL http://dx.doi.org/10.1073/pnas.1216206110. 20, 156

[98] Mohseni, M. *et al.* Macrod2 overexpression mediates estrogen independent growth and tamoxifen resistance in breast cancers. *Proc Natl Acad Sci U S A* **111**, 17606–17611 (2014). URL http://dx.doi.org/10.1073/pnas.1408650111. 20, 138

[99] Davison, E. J., Tarpey, P. S., Fiegler, H., Tomlinson, I. P. M. & Carter, N. P. Deletion at chromosome band 20p12.1 in colorectal cancer revealed by high resolution array comparative genomic hybridization. *Genes Chromosomes Cancer* **44**, 384–391 (2005). URL http://dx.doi.org/10.1002/gcc.20252. 20

[100] Linnebacher, M. *et al.* Single nucleotide polymorphism array analysis of microsatellite-stable, diploid/near-diploid colorectal carcinomas without the cpg island methylator phenotype. *Oncol Lett* **5**, 173–178 (2013). URL http://dx.doi.org/10.3892/ol.2012.1006. 20

[101] van den Broek, E. *et al.* High prevalence and clinical relevance of genes affected by chromosomal breaks in colorectal cancer. *PLoS One* **10**, e0138141 (2015). URL http://dx.doi.org/10.1371/journal.pone.0138141. 20

[102] Briffa, R. *et al.* Multi-scale genomic, transcriptomic and proteomic analysis of colorectal cancer cell lines to identify novel biomarkers. *PLoS One* **10**, e0144708 (2015). URL http://dx.doi.org/10.1371/journal.pone.0144708. 20

[103] Rajaram, M. *et al.* Two distinct categories of focal deletions in cancer genomes. *PLoS One* **8**, e66264 (2013). URL http://dx.doi.org/10.1371/journal.pone.0066264. 20

[104] Obeidat, M. *et al.* A comprehensive evaluation of potential lung function associated genes in the spirometa general population sample. *PLoS One* **6**, e19382 (2011). 20

[105] Slavin, T. P., Feng, T., Schnell, A., Zhu, X. & Elston, R. C. Two-marker association tests yield new disease associations for coronary artery disease and hypertension. *Hum Genet* **130**, 725–733 (2011). URL http://dx.doi.org/10.1007/s00439-011-1009-6. 20

[106] Julià, A. *et al.* A genome-wide association study on a southern european population identifies a new crohn's disease susceptibility locus at rbx1-ep300. *Gut* **62**, 1440–1445 (2013). URL http://dx.doi.org/10.1136/gutjnl-2012-302865. 20

[107] Lasho, T. *et al.* Identification of submicroscopic genetic changes and precise breakpoint mapping in myelofibrosis using high resolution mate-pair sequencing. *Am J Hematol* **88**, 741–746 (2013). URL http://dx.doi.org/10.1002/ajh.23495. 20

[108] Tan, W. J. *et al.* Novel genetic aberrations in breast phyllodes tumours: comparison between prognostically distinct groups. *Breast Cancer Res Treat* **145**, 635–645 (2014). URL http://dx.doi.org/10.1007/s10549-014-2982-y. 20

[109] Chen, C. *et al.* Next-generation-sequencing of recurrent childhood high hyperdiploid acute lymphoblastic leukemia reveals mutations typically associated with high risk patients. *Leuk Res* **39**, 990–1001 (2015). URL http://dx.doi.org/10.1016/j.leukres.2015.06.005. 20

[110] De Vos, M., Schreiber, V. & Dantzer, F. The diverse roles and clinical relevance of parps in dna damage repair: current state of the art. *Biochem Pharmacol* **84**, 137–146 (2012). URL http://dx.doi.org/10.1016/j.bcp.2012.03.018. 21

[111] Caldecott, K. W. Single-strand break repair and genetic disease. *Nat Rev Genet* **9**, 619–631 (2008). URL http://dx.doi.org/10.1038/nrg2380. 21, 54

[112] Dianov, G. L. & Hübscher, U. Mammalian base excision repair: the forgotten archangel. *Nucleic Acids Res* **41**, 3483–3490 (2013). URL http://dx.doi.org/10.1093/nar/gkt076. 22

[113] Carter, R. J. & Parsons, J. L. Base excision repair: A pathway regulated by post-translational modifications. *Mol Cell Biol* (2016). URL http://dx.doi.org/10.1128/MCB.00030-16. 22

[114] Schreiber, V. *et al.* Poly(adp-ribose) polymerase-2 (parp-2) is required for efficient base excision dna repair in association with parp-1 and xrcc1. *J Biol Chem* **277**, 23028–23036 (2002). URL http://dx.doi.org/10.1074/jbc.M202390200. 22

[115] Dantzer, F. *et al.* Base excision repair is impaired in mammalian cells lacking poly(adp-ribose) polymerase-1. *Biochemistry* **39**, 7559–7569 (2000). 22

[116] Le Page, F., Schreiber, V., Dherin, C., De Murcia, G. & Boiteux, S. Poly(adp-ribose) polymerase-1 (parp-1) is required in murine cell lines for base excision repair of oxidative dna damage in the absence of dna polymerase beta. *J Biol Chem* **278**, 18471–18477 (2003). URL http://dx.doi.org/10.1074/jbc.M212905200. 22

[117] Fisher, A. E. O., Hochegger, H., Takeda, S. & Caldecott, K. W. Poly(adp-ribose) polymerase 1 accelerates single-strand break repair in concert with poly(adp-ribose) glycohydrolase. *Mol Cell Biol* **27**, 5597–5605 (2007). URL http://dx.doi.org/10.1128/MCB.02248-06. 22

[118] Jungmichel, S. *et al.* Proteome-wide identification of poly(adp-ribosyl)ation targets in different genotoxic stress responses. *Mol Cell* **52**, 272–285 (2013). URL http://dx.doi.org/10.1016/j.molcel.2013.08.026. 22, 24, 29, 135, 152

[119] Gagné, J.-P. *et al.* Quantitative proteomics profiling of the poly(adp-ribose)-related response to genotoxic stress. *Nucleic Acids Res* **40**, 7788–7805 (2012). URL http://dx.doi.org/10.1093/nar/gks486. 22, 24

[120] El-Khamisy, S. F., Masutani, M., Suzuki, H. & Caldecott, K. W. A requirement for parp-1 for the assembly or stability of xrcc1 nuclear foci at sites of oxidative dna damage. *Nucleic Acids Res* **31**, 5526–5533 (2003). 22

[121] Okano, S., Lan, L., Caldecott, K. W., Mori, T. & Yasui, A. Spatial and temporal cellular responses to single-strand breaks in human cells. *Mol Cell Biol* **23**, 3974–3981 (2003). 22

[122] Hanssen-Bauer, A. *et al.* Xrcc1 coordinates disparate responses and multiprotein repair complexes depending on the nature and context of the dna damage. *Environ Mol Mutagen* **52**, 623–635 (2011). URL http://dx.doi.org/10.1002/em.20663. 22

[123] Campalans, A. *et al.* Distinct spatiotemporal patterns and parp dependence of xrcc1 recruitment to single-strand break and base excision repair. *Nucleic Acids Res* **41**, 3115–3129 (2013). URL http://dx.doi.org/10.1093/nar/gkt025. 22

[124] Popp, O. *et al.* Site-specific noncovalent interaction of the biopolymer poly(adp-ribose) with the werner syndrome protein regulates protein functions. *ACS Chem Biol* **8**, 179–188 (2013). URL http://dx.doi.org/10.1021/cb300363g. 22

[125] Harris, J. L. *et al.* Aprataxin, poly-adp ribose polymerase 1 (parp-1) and apurinic endonuclease 1 (ape1) function together to protect the genome against oxidative damage. *Hum Mol Genet* **18**, 4102–4117 (2009). URL http://dx.doi.org/10.1093/hmg/ddp359. 22

[126] Noren Hooten, N., Kompaniez, K., Barnes, J., Lohani, A. & Evans, M. K. Poly(adp-ribose) polymerase 1 (parp-1) binds to 8-oxoguanine-dna glycosylase (ogg1). *J Biol Chem* **286**, 44679–44690 (2011). URL http://dx.doi.org/10.1074/jbc.M111.255869. 22

REFERENCES

[127] Ström, C. E. *et al.* Poly (adp-ribose) polymerase (parp) is not involved in base excision repair but parp inhibition traps a single-strand intermediate. *Nucleic Acids Res* **39**, 3166–3175 (2011). URL http://dx.doi.org/10.1093/nar/gkq1241. 22

[128] Patel, A. G., Sarkaria, J. N. & Kaufmann, S. H. Nonhomologous end joining drives poly(adp-ribose) polymerase (parp) inhibitor lethality in homologous recombination-deficient cells. *Proc Natl Acad Sci U S A* **108**, 3406–3411 (2011). URL http://dx.doi.org/10.1073/pnas.1013715108. 22

[129] Helleday, T. The underlying mechanism for the parp and brca synthetic lethality: clearing up the misunderstandings. *Mol Oncol* **5**, 387–393 (2011). URL http://dx.doi.org/10.1016/j.molonc.2011.07.001. 22

[130] Spivak, G. Nucleotide excision repair in humans. *DNA Repair (Amst)* **36**, 13–18 (2015). URL http://dx.doi.org/10.1016/j.dnarep.2015.09.003. 22

[131] Vodenicharov, M. D., Ghodgaonkar, M. M., Halappanavar, S. S., Shah, R. G. & Shah, G. M. Mechanism of early biphasic activation of poly(adp-ribose) polymerase-1 in response to ultraviolet b radiation. *J Cell Sci* **118**, 589–599 (2005). URL http://dx.doi.org/10.1242/jcs.01636. 22

[132] Luijsterburg, M. S. *et al.* Ddb2 promotes chromatin decondensation at uv-induced dna damage. *J Cell Biol* **197**, 267–281 (2012). URL http://dx.doi.org/10.1083/jcb.201106074. 22

[133] Pines, A. *et al.* Parp1 promotes nucleotide excision repair through ddb2 stabilization and recruitment of alc1. *J Cell Biol* **199**, 235–249 (2012). URL http://dx.doi.org/10.1083/jcb.201112132. 22

[134] Robu, M. *et al.* Role of poly(adp-ribose) polymerase-1 in the removal of uv-induced dna lesions by nucleotide excision repair. *Proc Natl Acad Sci U S A* **110**, 1658–1663 (2013). URL http://dx.doi.org/10.1073/pnas.1209507110. 22

[135] Fahrer, J., Kranaster, R., Altmeyer, M., Marx, A. & Bürkle, A. Quantitative analysis of the binding affinity of poly(adp-ribose) to specific binding proteins as a function of chain length. *Nucleic Acids Res* **35**, e143 (2007). URL http://dx.doi.org/10.1093/nar/gkm944. 23

[136] Fischer, J. M. F. *et al.* Poly(adp-ribose)-mediated interplay of xpa and parp1 leads to reciprocal regulation of protein function. *FEBS J* **281**, 3625–3641 (2014). URL http://dx.doi.org/10.1111/febs.12885. 23

[137] Jackson, S. P. & Bartek, J. The dna-damage response in human biology and disease. *Nature* **461**, 1071–1078 (2009). URL http://dx.doi.org/10.1038/nature08467. 23

[138] Frit, P., Barboule, N., Yuan, Y., Gomez, D. & Calsou, P. Alternative end-joining pathway(s): bricolage at dna breaks. *DNA Repair (Amst)* **17**, 81–97 (2014). URL http://dx.doi.org/10.1016/j.dnarep.2014.02.007. 23, 24, 25

[139] de Campos-Nebel, M., Larripa, I. & González-Cid, M. Topoisomerase ii-mediated dna damage is differently repaired during the cell cycle by non-homologous end joining and homologous recombination. *PLoS One* **5** (2010). URL http://dx.doi.org/10.1371/journal.pone.0012541. 23

[140] Ceccaldi, R., Rondinelli, B. & D'Andrea, A. D. Repair pathway choices and consequences at the double-strand break. *Trends Cell Biol* **26**, 52–64 (2016). URL http://dx.doi.org/10.1016/j.tcb.2015.07.009. 23

[141] Karanam, K., Kafri, R., Loewer, A. & Lahav, G. Quantitative live cell imaging reveals a gradual shift between dna repair mechanisms and a maximal use of hr in mid s phase. *Mol Cell* **47**, 320–329 (2012). URL http://dx.doi.org/10.1016/j.molcel.2012.05.052. 23

[142] Thode, S., Schäfer, A., Pfeiffer, P. & Vielmetter, W. A novel pathway of dna end-to-end joining. *Cell* **60**, 921–928 (1990). 23

[143] Mason, R. M., Thacker, J. & Fairman, M. P. The joining of non-complementary dna double-strand breaks by mammalian extracts. *Nucleic Acids Res* **24**, 4946–4953 (1996). 23

[144] Difilippantonio, M. J. *et al.* Dna repair protein ku80 suppresses chromosomal aberrations and malignant transformation. *Nature* **404**, 510–514 (2000). URL http://dx.doi.org/10.1038/35006670. 23

[145] DiBiase, S. J. *et al.* Dna-dependent protein kinase stimulates an independently active, nonhomologous, end-joining apparatus. *Cancer Res* **60**, 1245–1253 (2000). 23

[146] Lupo, B. & Trusolino, L. Inhibition of poly(adp-ribosyl)ation in cancer: old and new paradigms revisited. *Biochim Biophys Acta* **1846**, 201–215 (2014). URL http://dx.doi.org/10.1016/j.bbcan.2014.07.004. 24

[147] Bryant, H. E. *et al.* Specific killing of brca2-deficient tumours with inhibitors of poly(adp-ribose) polymerase. *Nature* **434**, 913–917 (2005). URL http://dx.doi.org/10.1038/nature03443. 24

[148] Farmer, H. *et al.* Targeting the dna repair defect in brca mutant cells as a therapeutic strategy. *Nature* **434**, 917–921 (2005). URL http://dx.doi.org/10.1038/nature03445. 24

[149] Murai, J. *et al.* Trapping of parp1 and parp2 by clinical parp inhibitors. *Cancer Res* **72**, 5588–5599 (2012). URL http://dx.doi.org/10.1158/0008-5472.CAN-12-2753. 24

[150] Murai, J. *et al.* Stereospecific parp trapping by bmn 673 and comparison with olaparib and rucaparib. *Mol Cancer Ther* **13**, 433–443 (2014). URL http://dx.doi.org/10.1158/1535-7163.MCT-13-0803. 24

[151] Schultz, N., Lopez, E., Saleh-Gohari, N. & Helleday, T. Poly(adp-ribose) polymerase (parp-1) has a controlling role in homologous recombination. *Nucleic Acids Res* **31**, 4959–4964 (2003). 24

[152] Schreiber, V. *et al.* A dominant-negative mutant of human poly(adp-ribose) polymerase affects cell recovery, apoptosis, and sister chromatid exchange following dna damage. *Proc Natl Acad Sci U S A* **92**, 4753–4757 (1995). 24

[153] Waldman, A. S. & Waldman, B. C. Stimulation of intrachromosomal homologous recombination in mammalian cells by an inhibitor of poly(adp-ribosylation). *Nucleic Acids Res* **19**, 5943–5947 (1991). 24

[154] Semionov, A., Cournoyer, D. & Chow, T. Y. Inhibition of poly(adp-ribose)polymerase stimulates extrachromosomal homologous recombination in mouse ltk-fibroblasts. *Nucleic Acids Res* **27**, 4526–4531 (1999). 24

[155] Haince, J.-F. *et al.* Parp1-dependent kinetics of recruitment of mre11 and nbs1 proteins to multiple dna damage sites. *J Biol Chem* **283**, 1197–1208 (2008). URL http://dx.doi.org/10.1074/jbc.M706734200. 24, 113

[156] Haince, J.-F. *et al.* Ataxia telangiectasia mutated (atm) signaling network is modulated by a novel poly(adp-ribose)-dependent pathway in the early response to dna-damaging agents. *J Biol Chem* **282**, 16441–16453 (2007). URL http://dx.doi.org/10.1074/jbc.M608406200. 24, 113

[157] Smeenk, G. *et al.* Poly(adp-ribosyl)ation links the chromatin remodeler smarca5/snf2h to rnf168-dependent dna damage signaling. *J Cell Sci* **126**, 889–903 (2013). URL http://dx.doi.org/10.1242/jcs.109413. 24, 26

[158] Oberoi, J. *et al.* Structural basis of poly(adp-ribose) recognition by the multizinc binding domain of checkpoint with forkhead-associated and ring domains (chfr). *J Biol Chem* **285**, 39348–39358 (2010). URL http://dx.doi.org/10.1074/jbc.M110.159855. 24

[159] Liu, C., Wu, J., Paudyal, S. C., You, Z. & Yu, X. Chfr is important for the first wave of ubiquitination at dna damage sites. *Nucleic Acids Res* **41**, 1698–1710 (2013). URL http://dx.doi.org/10.1093/nar/gks1278. 24, 26

[160] Ruscetti, T. *et al.* Stimulation of the dna-dependent protein kinase by poly(adp-ribose) polymerase. *J Biol Chem* **273**, 14461–14467 (1998). 24

[161] Galande, S. & Kohwi-Shigematsu, T. Poly(adp-ribose) polymerase and ku autoantigen form a complex and synergistically bind to matrix attachment sequences. *J Biol Chem* **274**, 20521–20528 (1999). 24

[162] Wang, M. *et al.* Parp-1 and ku compete for repair of dna double strand breaks by distinct nhej pathways. *Nucleic Acids Res* **34**, 6170–6182 (2006). URL http://dx.doi.org/10.1093/nar/gkl840. 24, 25

[163] Spagnolo, L., Barbeau, J., Curtin, N. J., Morris, E. P. & Pearl, L. H. Visualization of a dna-pk/parp1 complex. *Nucleic Acids Res* **40**, 4168–4177 (2012). URL http://dx.doi.org/10.1093/nar/gkr1231. 24

[164] Henrie, M. S. *et al.* Lethality in parp-1/ku80 double mutant mice reveals physiological synergy during early embryogenesis. *DNA Repair (Amst)* **2**, 151–158 (2003). 25

[165] Mitchell, J., Smith, G. C. M. & Curtin, N. J. Poly(adp-ribose) polymerase-1 and dna-dependent protein kinase have equivalent roles in double strand break repair following ionizing radiation. *Int J Radiat Oncol Biol Phys* **75**, 1520–1527 (2009). URL http://dx.doi.org/10.1016/j.ijrobp.2009.07.1722. 25

[166] Morrison, C. *et al.* Genetic interaction between parp and dna-pk in v(d)j recombination and tumorigenesis. *Nat Genet* **17**, 479–482 (1997). URL http://dx.doi.org/10.1038/ng1297-479. 25

[167] Ying, S. *et al.* Dna-pkcs and parp1 bind to unresected stalled dna replication forks where they recruit xrcc1 to mediate repair. *Cancer Res* **76**, 1078–1088 (2016). URL http://dx.doi.org/10.1158/0008-5472.CAN-15-0608. 25

[168] Calkins, A. S., Iglehart, J. D. & Lazaro, J.-B. Dna damage-induced inhibition of rrna synthesis by dna-pk and parp-1. *Nucleic Acids Res* **41**, 7378–7386 (2013). URL http://dx.doi.org/10.1093/nar/gkt502. 25

[169] Perrault, R., Wang, H., Wang, M., Rosidi, B. & Iliakis, G. Backup pathways of nhej are suppressed by dna-pk. *J Cell Biochem* **92**, 781–794 (2004). URL http://dx.doi.org/10.1002/jcb.20104. 25

[170] Hochegger, H. *et al.* Parp-1 protects homologous recombination from interference by ku and ligase iv in vertebrate cells. *EMBO J* **25**, 1305–1314 (2006). URL http://dx.doi.org/10.1038/sj.emboj.7601015. 25

[171] Fattah, F. *et al.* Ku regulates the non-homologous end joining pathway choice of dna double-strand break repair in human somatic cells. *PLoS Genet* **6**, e1000855 (2010). URL http://dx.doi.org/10.1371/journal.pgen.1000855. 25

[172] Paddock, M. N. *et al.* Competition between parp-1 and ku70 control the decision between high-fidelity and mutagenic dna repair. *DNA Repair (Amst)* **10**, 338–343 (2011). URL http://dx.doi.org/10.1016/j.dnarep.2010.12.005. 25

[173] Audebert, M., Salles, B. & Calsou, P. Involvement of poly(adp-ribose) polymerase-1 and xrcc1/dna ligase iii in an alternative route for dna double-strand breaks rejoining. *J Biol Chem* **279**, 55117–55126 (2004). URL http://dx.doi.org/10.1074/jbc.M404524200. 25

[174] Soni, A. *et al.* Requirement for parp-1 and dna ligases 1 or 3 but not of xrcc1 in chromosomal translocation formation by backup end joining. *Nucleic Acids Res* **42**, 6380–6392 (2014). URL http://dx.doi.org/10.1093/nar/gku298. 25

[175] Nicolae, C. M. *et al.* The adp-ribosyltransferase parp10/artd10 interacts with proliferating cell nuclear antigen (pcna) and is required for dna damage tolerance. *J Biol Chem* **289**, 13627–13637 (2014). URL http://dx.doi.org/10.1074/jbc.M114.556340. 25, 113

[176] Vyas, S., Chesarone-Cataldo, M., Todorova, T., Huang, Y.-H. & Chang, P. A systematic analysis of the parp protein family identifies new functions critical for cell physiology. *Nat Commun* **4**, 2240 (2013). URL http://dx.doi.org/10.1038/ncomms3240. 25

[177] Poirier, G. G., de Murcia, G., Jongstra-Bilen, J., Niedergang, C. & Mandel, P. Poly(adp-ribosyl)ation of polynucleosomes causes relaxation of chromatin structure. *Proc Natl Acad Sci U S A* **79**, 3423–3427 (1982). 26

[178] Huletsky, A. *et al.* The effect of poly(adp-ribosyl)ation on native and h1-depleted chromatin. a role of poly(adp-ribosyl)ation on core nucleosome structure. *J Biol Chem* **264**, 8878–8886 (1989). 26

[179] Krishnakumar, R. *et al.* Reciprocal binding of parp-1 and histone h1 at promoters specifies transcriptional outcomes. *Science* **319**, 819–821 (2008). URL http://dx.doi.org/10.1126/science.1149250. 26

[180] Heo, K. *et al.* Fact-mediated exchange of histone variant h2ax regulated by phosphorylation of h2ax and adp-ribosylation of spt16. *Mol Cell* **30**, 86–97 (2008). URL http://dx.doi.org/10.1016/j.molcel.2008.02.029. 26

[181] Krishnakumar, R. & Kraus, W. L. Parp-1 regulates chromatin structure and transcription through a kdm5b-dependent pathway. *Mol Cell* **39**, 736–749 (2010). URL http://dx.doi.org/10.1016/j.molcel.2010.08.014. 26

[182] Gottschalk, A. J. *et al.* Poly(adp-ribosyl)ation directs recruitment and activation of an atp-dependent chromatin remodeler. *Proc Natl Acad Sci U S A* **106**, 13770–13774 (2009). URL http://dx.doi.org/10.1073/pnas.0906920106. 26

[183] Chou, D. M. *et al.* A chromatin localization screen reveals poly (adp ribose)-regulated recruitment of the repressive polycomb and nurd complexes to sites of dna damage. *Proc Natl Acad Sci U S A* **107**, 18475–18480 (2010). URL http://dx.doi.org/10.1073/pnas.1012946107. 26

[184] Polo, S. E., Kaidi, A., Baskcomb, L., Galanty, Y. & Jackson, S. P. Regulation of dna-damage responses and cell-cycle progression by the chromatin remodelling factor chd4. *EMBO J* **29**, 3130–3139 (2010). URL http://dx.doi.org/10.1038/emboj.2010.188. 26

[185] Kang, H. C. *et al.* Iduna is a poly(adp-ribose) (par)-dependent e3 ubiquitin ligase that regulates dna damage. *Proc Natl Acad Sci U S A* **108**, 14103–14108 (2011). URL http://dx.doi.org/10.1073/pnas.1108799108. 26

[186] Zhang, Y. *et al.* Rnf146 is a poly(adp-ribose)-directed e3 ligase that regulates axin degradation and wnt signalling. *Nat Cell Biol* **13**, 623–629 (2011). URL http://dx.doi.org/10.1038/ncb2222. 26

[187] Huang, S.-M. A. *et al.* Tankyrase inhibition stabilizes axin and antagonizes wnt signalling. *Nature* **461**, 614–620 (2009). URL http://dx.doi.org/10.1038/nature08356. 26

[188] Kashima, L. *et al.* Chfr protein regulates mitotic checkpoint by targeting parp-1 protein for ubiquitination and degradation. *J Biol Chem* **287**, 12975–12984 (2012). URL http://dx.doi.org/10.1074/jbc.M111.321828. 26

[189] Andrabi, S. A. *et al.* Poly(adp-ribose) (par) polymer is a death signal. *Proc Natl Acad Sci U S A* **103**, 18308–18313 (2006). URL http://dx.doi.org/10.1073/pnas.0606526103. 26

[190] Wang, Y. *et al.* Poly(adp-ribose) (par) binding to apoptosis-inducing factor is critical for par polymerase-1-dependent cell death (parthanatos). *Sci Signal* **4**, ra20 (2011). URL http://dx.doi.org/10.1126/scisignal.2000902. 28

[191] Embi, N., Rylatt, D. B. & Cohen, P. Glycogen synthase kinase-3 from rabbit skeletal muscle. separation from cyclic-amp-dependent protein kinase and phosphorylase kinase. *Eur J Biochem* **107**, 519–527 (1980). 28

[192] Lochhead, P. A., Coghlan, M., Rice, S. Q. & Sutherland, C. Inhibition of gsk-3 selectively reduces glucose-6-phosphatase and phosphatase and phosphoenolypyruvate carboxykinase gene expression. *Diabetes* **50**, 937–946 (2001). 28

[193] Miyamoto, S. Nuclear initiated nf-κb signaling: Nemo and atm take center stage. *Cell Res* **21**, 116–130 (2011). URL http://dx.doi.org/10.1038/cr.2010.179. 28, 36, 45

[194] Verheugd, P. *et al.* Regulation of nf-κb signalling by the mono-adp-ribosyltransferase artd10. *Nat Commun* **4**, 1683 (2013). URL http://dx.doi.org/10.1038/ncomms2672. 28, 36, 157

[195] Stilmann, M. *et al.* A nuclear poly(adp-ribose)-dependent signalosome confers dna damage-induced ikappab kinase activation. *Mol Cell* **36**, 365–378 (2009). URL http://dx.doi.org/10.1016/j.molcel.2009.09.032. 28

[196] Hassa, P. O. *et al.* Acetylation of poly(adp-ribose) polymerase-1 by p300/creb-binding protein regulates coactivation of nf-kappab-dependent transcription. *J Biol Chem* **280**, 40450–40464 (2005). URL http://dx.doi.org/10.1074/jbc.M507553200. 28

[197] Hassa, P. O., Covic, M., Bedford, M. T. & Hottiger, M. O. Protein arginine methyltransferase 1 coactivates nf-kappab-dependent gene expression synergistically with carm1 and parp1. *J Mol Biol* **377**, 668–678 (2008). URL http://dx.doi.org/10.1016/j.jmb.2008.01.044. 28

[198] Aguiar, R. C. T., Takeyama, K., He, C., Kreinbrink, K. & Shipp, M. A. B-aggressive lymphoma family proteins have unique domains that modulate transcription and exhibit poly(adp-ribose) polymerase activity. *J Biol Chem* **280**, 33756–33765 (2005). URL http://dx.doi.org/10.1074/jbc.M505408200. 29

[199] Goenka, S. & Boothby, M. Selective potentiation of stat-dependent gene expression by collaborator of stat6 (coast6), a transcriptional cofactor. *Proc Natl Acad Sci U S A* **103**, 4210–4215 (2006). URL http://dx.doi.org/10.1073/pnas.0506981103. 29

[200] Goenka, S., Cho, S. H. & Boothby, M. Collaborator of stat6 (coast6)-associated poly(adp-ribose) polymerase activity modulates stat6-dependent gene transcription. *J Biol Chem* **282**, 18732–18739 (2007). URL http://dx.doi.org/10.1074/jbc.M611283200. 29

[201] Mehrotra, P. *et al.* Parp-14 functions as a transcriptional switch for stat6-dependent gene activation. *J Biol Chem* **286**, 1767–1776 (2011). URL http://dx.doi.org/10.1074/jbc.M110.157768. 29

[202] Camicia, R. *et al.* Bal1/artd9 represses the anti-proliferative and pro-apoptotic ifnγ-stat1-irf1-p53 axis in diffuse large b-cell lymphoma. *J Cell Sci* **126**, 1969–1980 (2013). URL http://dx.doi.org/10.1242/jcs.118174. 29

[203] Bachmann, S. B. *et al.* Dtx3l and artd9 inhibit irf1 expression and mediate in cooperation with artd8 survival and proliferation of metastatic prostate cancer cells. *Mol Cancer* **13**, 125 (2014). URL http://dx.doi.org/10.1186/1476-4598-13-125. 29

[204] Jwa, M. & Chang, P. Parp16 is a tail-anchored endoplasmic reticulum protein required for the perk- and ire1α-mediated unfolded protein response. *Nat Cell Biol* **14**, 1223–1230 (2012). URL http://dx.doi.org/10.1038/ncb2593. 29

[205] Chambers, J. E., Petrova, K., Tomba, G., Vendruscolo, M. & Ron, D. Adp ribosylation adapts an er chaperone response to short-term fluctuations in unfolded protein load. *J Cell Biol* **198**, 371–385 (2012). URL http://dx.doi.org/10.1083/jcb.201202005. 29

[206] Gagné, J.-P. *et al.* Proteome-wide identification of poly(adp-ribose) binding proteins and poly(adp-ribose)-associated protein complexes. *Nucleic Acids Res* **36**, 6959–6976 (2008). URL http://dx.doi.org/10.1093/nar/gkn771. 29, 135, 152

[207] Gagné, J.-P., Hunter, J. M., Labrecque, B., Chabot, B. & Poirier, G. G. A proteomic approach to the identification of heterogeneous nuclear ribonucleoproteins as a new family of poly(adp-ribose)-binding proteins. *Biochem J* **371**, 331–340 (2003). URL http://dx.doi.org/10.1042/BJ20021675. 30

[208] Ji, Y. & Tulin, A. V. Poly(adp-ribosyl)ation of heterogeneous nuclear ribonucleoproteins modulates splicing. *Nucleic Acids Res* **37**, 3501–3513 (2009). URL http://dx.doi.org/10.1093/nar/gkp218. 30

[209] Sitikov, A. S., Davydova, E. K., Bezlepkina, T. A., Ovchinnikov, L. P. & Spirin, A. S. Eukaryotic elongation factor 2 loses its non-specific affinity for rna and leaves polyribosomes as a result of adp-ribosylation. *FEBS Lett* **176**, 406–410 (1984). 30, 152

[210] Sitikov, A. S., Davydova, E. K. & Ovchinnikov, L. P. Endogenous adp-ribosylation of elongation factor 2 in polyribosome fraction of rabbit reticulocytes. *FEBS Lett* **176**, 261–263 (1984). 30, 152

[211] Jäger, D., Werdan, K. & Müller-Werdan, U. Endogenous adp-ribosylation of elongation factor-2 by interleukin-1β. *Mol Cell Biochem* **348**, 125–128 (2011). URL http://dx.doi.org/10.1007/s11010-010-0646-8. 30, 152

[212] Leung, A. K. L. *et al.* Poly(adp-ribose) regulates stress responses and microrna activity in the cytoplasm. *Mol Cell* **42**, 489–499 (2011). URL http://dx.doi.org/10.1016/j.molcel.2011.04.015. 30, 139, 153

[213] Carter-O'Connell, I. & Cohen, M. S. Identifying direct protein targets of poly-adp-ribose polymerases (parps) using engineered parp variants-orthogonal nicotinamide adenine dinucleotide (nad+) analog pairs. *Curr Protoc Chem Biol* **7**, 121–139 (2015). URL http://dx.doi.org/10.1002/9780470559277.ch140259. 30

[214] Carter-O'Connell, I. *et al.* Identifying family-member-specific targets of mono-artds by using a chemical genetics approach. *Cell Rep* **14**, 621–631 (2016). URL http://dx.doi.org/10.1016/j.celrep.2015.12.045. 30

[215] Shiloh, Y. Atm: expanding roles as a chief guardian of genome stability. *Exp Cell Res* **329**, 154–161 (2014). URL http://dx.doi.org/10.1016/j.yexcr.2014.09.002. 31, 32, 33

[216] Savitsky, K. *et al.* A single ataxia telangiectasia gene with a product similar to pi-3 kinase. *Science* **268**, 1749–1753 (1995). 32

[217] Ziv, Y. *et al.* Recombinant atm protein complements the cellular a-t phenotype. *Oncogene* **15**, 159–167 (1997). URL http://dx.doi.org/10.1038/sj.onc.1201319. 32

[218] Reiman, A. *et al.* Lymphoid tumours and breast cancer in ataxia telangiectasia; substantial protective effect of residual atm kinase activity against childhood tumours. *Br J Cancer* **105**, 586–591 (2011). URL http://dx.doi.org/10.1038/bjc.2011.266. 33

[219] Jung, M., Kondratyev, A., Lee, S. A., Dimtchev, A. & Dritschilo, A. Atm gene product phosphorylates i kappa b-alpha. *Cancer Res* **57**, 24–27 (1997). 33

[220] Kim, S. T., Lim, D. S., Canman, C. E. & Kastan, M. B. Substrate specificities and identification of putative substrates of atm kinase family members. *J Biol Chem* **274**, 37538–37543 (1999). 33, 83, 103

[221] Brown, K. D. *et al.* The ataxia-telangiectasia gene product, a constitutively expressed nuclear protein that is not up-regulated following genome damage. *Proc Natl Acad Sci U S A* **94**, 1840–1845 (1997). 33

[222] Barlow, C. *et al.* Atm is a cytoplasmic protein in mouse brain required to prevent lysosomal accumulation. *Proc Natl Acad Sci U S A* **97**, 871–876 (2000). 33

[223] Li, J., Han, Y. R., Plummer, M. R. & Herrup, K. Cytoplasmic atm in neurons modulates synaptic function. *Curr Biol* **19**, 2091–2096 (2009). URL http://dx.doi.org/10.1016/j.cub.2009.10.039. 33

[224] Boehrs, J. K., He, J., Halaby, M.-J. & Yang, D.-Q. Constitutive expression and cytoplasmic compartmentalization of atm protein in differentiated human neuron-like sh-sy5y cells. *J Neurochem* **100**, 337–345 (2007). URL http://dx.doi.org/10.1111/j.1471-4159.2006.04254.x. 33

[225] Lim, D. S. *et al.* Atm binds to beta-adaptin in cytoplasmic vesicles. *Proc Natl Acad Sci U S A* **95**, 10146–10151 (1998). 33

[226] Wu, Z.-H., Shi, Y., Tibbetts, R. S. & Miyamoto, S. Molecular linkage between the kinase atm and nf-kappab signaling in response to genotoxic stimuli. *Science* **311**, 1141–1146 (2006). URL http://dx.doi.org/10.1126/science.1121513. 33, 36, 45, 110, 111, 157, 158

[227] Utsumi, H. & Sasaki, M. S. Deficient repair of potentially lethal damage in actively growing ataxia telangiectasia cells. *Radiat Res* **97**, 407–413 (1984). 33

[228] Taylor, A. M., Metcalfe, J. A. & McConville, C. Increased radiosensitivity and the basic defect in ataxia telangiectasia. *Int J Radiat Biol* **56**, 677–684 (1989). 33

[229] Stracker, T. H., Roig, I., Knobel, P. A. & Marjanović, M. The atm signaling network in development and disease. *Front Genet* **4**, 37 (2013). URL http://dx.doi.org/10.3389/fgene.2013.00037. 33, 36

REFERENCES

[230] Maréchal, A. & Zou, L. Dna damage sensing by the atm and atr kinases. *Cold Spring Harb Perspect Biol* **5** (2013). URL http://dx.doi.org/10.1101/cshperspect.a012716. 33, 35, 38, 39

[231] Smith, G. C. *et al.* Purification and dna binding properties of the ataxia-telangiectasia gene product atm. *Proc Natl Acad Sci U S A* **96**, 11134–11139 (1999). 33

[232] Lee, J.-H. & Paull, T. T. Atm activation by dna double-strand breaks through the mre11-rad50-nbs1 complex. *Science* **308**, 551–554 (2005). URL http://dx.doi.org/10.1126/science.1108297. 33, 34

[233] Stewart, G. S. *et al.* The dna double-strand break repair gene hmre11 is mutated in individuals with an ataxia-telangiectasia-like disorder. *Cell* **99**, 577–587 (1999). 33

[234] Waltes, R. *et al.* Human rad50 deficiency in a nijmegen breakage syndrome-like disorder. *Am J Hum Genet* **84**, 605–616 (2009). URL http://dx.doi.org/10.1016/j.ajhg.2009.04.010. 33

[235] Zhu, J., Petersen, S., Tessarollo, L. & Nussenzweig, A. Targeted disruption of the nijmegen breakage syndrome gene nbs1 leads to early embryonic lethality in mice. *Curr Biol* **11**, 105–109 (2001). 33

[236] Mirzoeva, O. K. & Petrini, J. H. Dna damage-dependent nuclear dynamics of the mre11 complex. *Mol Cell Biol* **21**, 281–288 (2001). URL http://dx.doi.org/10.1128/MCB.21.1.281-288.2001. 33

[237] Uziel, T. *et al.* Requirement of the mrn complex for atm activation by dna damage. *EMBO J* **22**, 5612–5621 (2003). URL http://dx.doi.org/10.1093/emboj/cdg541. 34

[238] Bakkenist, C. J. & Kastan, M. B. Dna damage activates atm through intermolecular autophosphorylation and dimer dissociation. *Nature* **421**, 499–506 (2003). URL http://dx.doi.org/10.1038/nature01368. 34

[239] Sun, Y., Jiang, X., Chen, S., Fernandes, N. & Price, B. D. A role for the tip60 histone acetyltransferase in the acetylation and activation of atm. *Proc Natl Acad Sci U S A* **102**, 13182–13187 (2005). URL http://dx.doi.org/10.1073/pnas.0504211102. 34

[240] Ayoub, N., Jeyasekharan, A. D., Bernal, J. A. & Venkitaraman, A. R. Hp1-beta mobilization promotes chromatin changes that initiate the dna damage response. *Nature* **453**, 682–686 (2008). URL http://dx.doi.org/10.1038/nature06875. 34

[241] Goodarzi, A. A. *et al.* Atm signaling facilitates repair of dna double-strand breaks associated with heterochromatin. *Mol Cell* **31**, 167–177 (2008). URL http://dx.doi.org/10.1016/j.molcel.2008.05.017. 34

[242] Bolderson, E. *et al.* Kruppel-associated box (krab)-associated co-repressor (kap-1) ser-473 phosphorylation regulates heterochromatin protein 1β (hp1-β) mobilization and dna repair in heterochromatin. *J Biol Chem* **287**, 28122–28131 (2012). URL http://dx.doi.org/10.1074/jbc.M112.368381. 34

[243] Andegeko, Y. *et al.* Nuclear retention of atm at sites of dna double strand breaks. *J Biol Chem* **276**, 38224–38230 (2001). URL http://dx.doi.org/10.1074/jbc.M102986200. 34

[244] Stewart, G. S., Wang, B., Bignell, C. R., Taylor, A. M. R. & Elledge, S. J. Mdc1 is a mediator of the mammalian dna damage checkpoint. *Nature* **421**, 961–966 (2003). URL http://dx.doi.org/10.1038/nature01446. 34

[245] Lee, M. S., Edwards, R. A., Thede, G. L. & Glover, J. N. M. Structure of the brct repeat domain of mdc1 and its specificity for the free cooh-terminal end of the gamma-h2ax histone tail. *J Biol Chem* **280**, 32053–32056 (2005). URL http://dx.doi.org/10.1074/jbc.C500273200. 34

[246] Lou, Z. *et al.* Mdc1 maintains genomic stability by participating in the amplification of atm-dependent dna damage signals. *Mol Cell* **21**, 187–200 (2006). URL http://dx.doi.org/10.1016/j.molcel.2005.11.025. 34

[247] Chapman, J. R. & Jackson, S. P. Phospho-dependent interactions between nbs1 and mdc1 mediate chromatin retention of the mrn complex at sites of dna damage. *EMBO Rep* **9**, 795–801 (2008). URL http://dx.doi.org/10.1038/embor.2008.103. 34

[248] Melander, F. *et al.* Phosphorylation of sdt repeats in the mdc1 n terminus triggers retention of nbs1 at the dna damage-modified chromatin. *J Cell Biol* **181**, 213–226 (2008). URL http://dx.doi.org/10.1083/jcb.200708210. 34

[249] Liu, J. *et al.* Structural mechanism of the phosphorylation-dependent dimerization of the mdc1 forkhead-associated domain. *Nucleic Acids Res* **40**, 3898–3912 (2012). URL http://dx.doi.org/10.1093/nar/gkr1296. 34

[250] Mu, J.-J. *et al.* A proteomic analysis of ataxia telangiectasia-mutated (atm)/atm-rad3-related (atr) substrates identifies the ubiquitin-proteasome system as a regulator for dna damage checkpoints. *J Biol Chem* **282**, 17330–17334 (2007). URL http://dx.doi.org/10.1074/jbc.C700079200. 34

[251] Choi, S. *et al.* Quantitative proteomics reveal atm kinase-dependent exchange in dna damage response complexes. *J Proteome Res* **11**, 4983–4991 (2012). URL http://dx.doi.org/10.1021/pr3005524. 34

[252] Flaggs, G. *et al.* Atm-dependent interactions of a mammalian chk1 homolog with meiotic chromosomes. *Curr Biol* **7**, 977–986 (1997). 34, 38, 79

[253] Matsuoka, S. *et al.* Ataxia telangiectasia-mutated phosphorylates chk2 in vivo and in vitro. *Proc Natl Acad Sci U S A* **97**, 10389–10394 (2000). URL http://dx.doi.org/10.1073/pnas.190030497. 34

[254] Reinhardt, H. C., Aslanian, A. S., Lees, J. A. & Yaffe, M. B. p53-deficient cells rely on atm- and atr-mediated checkpoint signaling through the p38mapk/mk2 pathway for survival after dna damage. *Cancer Cell* **11**, 175–189 (2007). URL http://dx.doi.org/10.1016/j.ccr.2006.11.024. 34

[255] Siliciano, J. D. *et al.* Dna damage induces phosphorylation of the amino terminus of p53. *Genes Dev* **11**, 3471–3481 (1997). 34

[256] Banin, S. *et al.* Enhanced phosphorylation of p53 by atm in response to dna damage. *Science* **281**, 1674–1677 (1998). 34, 45

[257] Canman, C. E. *et al.* Activation of the atm kinase by ionizing radiation and phosphorylation of p53. *Science* **281**, 1677–1679 (1998). 34, 45

[258] Kastan, M. B., Onyekwere, O., Sidransky, D., Vogelstein, B. & Craig, R. W. Participation of p53 protein in the cellular response to dna damage. *Cancer Res* **51**, 6304–6311 (1991). 36, 140

[259] Martinez, J. D., Craven, M. T., Joseloff, E., Milczarek, G. & Bowden, G. T. Regulation of dna binding and transactivation in p53 by nuclear localization and phosphorylation. *Oncogene* **14**, 2511–2520 (1997). URL http://dx.doi.org/10.1038/sj.onc.1201095. 36

[260] Cortez, D., Wang, Y., Qin, J. & Elledge, S. J. Requirement of atm-dependent phosphorylation of brca1 in the dna damage response to double-strand breaks. *Science* **286**, 1162–1166 (1999). 36

[261] Xu, B., Kim St & Kastan, M. B. Involvement of brca1 in s-phase and g(2)-phase checkpoints after ionizing irradiation. *Mol Cell Biol* **21**, 3445–3450 (2001). URL http://dx.doi.org/10.1128/MCB.21.10.3445-3450.2001. 36

[262] Xu, B., Kim, S.-T., Lim, D.-S. & Kastan, M. B. Two molecularly distinct g(2)/m checkpoints are induced by ionizing irradiation. *Mol Cell Biol* **22**, 1049–1059 (2002). 36

[263] Paull, T. T. Mechanisms of atm activation. *Annu Rev Biochem* **84**, 711–738 (2015). URL http://dx.doi.org/10.1146/annurev-biochem-060614-034335. 36, 37, 110

[264] Kanu, N. & Behrens, A. Atmin defines an nbs1-independent pathway of atm signalling. *EMBO J* **26**, 2933–2941 (2007). URL http://dx.doi.org/10.1038/sj.emboj.7601733. 36

[265] Zhang, T. *et al.* Competition between nbs1 and atmin controls atm signaling pathway choice. *Cell Rep* **2**, 1498–1504 (2012). URL http://dx.doi.org/10.1016/j.celrep.2012.11.002. 36

[266] Zhang, T., Cronshaw, J., Kanu, N., Snijders, A. P. & Behrens, A. Ubr5-mediated ubiquitination of atmin is required for ionizing radiation-induced atm signaling and function. *Proc Natl Acad Sci U S A* **111**, 12091–12096 (2014). URL http://dx.doi.org/10.1073/pnas.1400230111. 36

[267] Bencokova, Z. *et al.* Atm activation and signaling under hypoxic conditions. *Mol Cell Biol* **29**, 526–537 (2009). URL http://dx.doi.org/10.1128/MCB.01301-08. 36

[268] Cam, H., Easton, J. B., High, A. & Houghton, P. J. mtorc1 signaling under hypoxic conditions is controlled by atm-dependent phosphorylation of hif-1α. *Mol Cell* **40**, 509–520 (2010). URL http://dx.doi.org/10.1016/j.molcel.2010.10.030. 37

[269] Ditch, S. & Paull, T. T. The atm protein kinase and cellular redox signaling: beyond the dna damage response. *Trends Biochem Sci* **37**, 15–22 (2012). URL http://dx.doi.org/10.1016/j.tibs.2011.10.002. 37

[270] Yi, M., Rosin, M. P. & Anderson, C. K. Response of fibroblast cultures from ataxia-telangiectasia patients to oxidative stress. *Cancer Lett* **54**, 43–50 (1990). 37

[271] Kamsler, A. *et al.* Increased oxidative stress in ataxia telangiectasia evidenced by alterations in redox state of brains from atm-deficient mice. *Cancer Res* **61**, 1849–1854 (2001). 37

[272] Quick, K. L. & Dugan, L. L. Superoxide stress identifies neurons at risk in a model of ataxia-telangiectasia. *Ann Neurol* **49**, 627–635 (2001). 37

[273] Shackelford, R. E. *et al.* The ataxia telangiectasia gene product is required for oxidative stress-induced g1 and g2 checkpoint function in human fibroblasts. *J Biol Chem* **276**, 21951–21959 (2001). URL http://dx.doi.org/10.1074/jbc.M011303200. 37

[274] Kurz, E. U., Douglas, P. & Lees-Miller, S. P. Doxorubicin activates atm-dependent phosphorylation of multiple downstream targets in part through the generation of reactive oxygen species. *J Biol Chem* **279**, 53272–53281 (2004). URL http://dx.doi.org/10.1074/jbc.M406879200. 37

[275] Guo, Z., Kozlov, S., Lavin, M. F., Person, M. D. & Paull, T. T. Atm activation by oxidative stress. *Science* **330**, 517–521 (2010). URL http://dx.doi.org/10.1126/science.1192912. 37, 110

[276] Shiloh, Y. Atm and related protein kinases: safeguarding genome integrity. *Nat Rev Cancer* **3**, 155–168 (2003). URL http://dx.doi.org/10.1038/nrc1011. 37, 65, 66, 79

[277] Lovejoy, C. A. & Cortez, D. Common mechanisms of pikk regulation. *DNA Repair (Amst)* **8**, 1004–1008 (2009). URL http://dx.doi.org/10.1016/j.dnarep.2009.04.006. 37

[278] Zou, L. & Elledge, S. J. Sensing dna damage through atrip recognition of rpa-ssdna complexes. *Science* **300**, 1542–1548 (2003). URL http://dx.doi.org/10.1126/science.1083430. 37, 140

[279] Jette, N. & Lees-Miller, S. P. The dna-dependent protein kinase: A multifunctional protein kinase with roles in dna double strand break repair and mitosis. *Prog Biophys Mol Biol* **117**, 194–205 (2015). URL http://dx.doi.org/10.1016/j.pbiomolbio.2014.12.003. 37, 38, 39

[280] Spagnolo, L., Rivera-Calzada, A., Pearl, L. H. & Llorca, O. Three-dimensional structure of the human dna-pkcs/ku70/ku80 complex assembled on dna and its implications for dna dsb repair. *Mol Cell* **22**, 511–519 (2006). URL http://dx.doi.org/10.1016/j.molcel.2006.04.013. 38

[281] Dobbs, T. A., Tainer, J. A. & Lees-Miller, S. P. A structural model for regulation of nhej by dna-pkcs autophosphorylation. *DNA Repair (Amst)* **9**, 1307–1314 (2010). URL http://dx.doi.org/10.1016/j.dnarep.2010.09.019. 38

[282] Ding, Q. *et al.* Autophosphorylation of the catalytic subunit of the dna-dependent protein kinase is required for efficient end processing during dna double-strand break repair. *Mol Cell Biol* **23**, 5836–5848 (2003). 38

[283] Uematsu, N. *et al.* Autophosphorylation of dna-pkcs regulates its dynamics at dna double-strand breaks. *J Cell Biol* **177**, 219–229 (2007). URL http://dx.doi.org/10.1083/jcb.200608077. 38

[284] Reitsema, T., Klokov, D., Banáth, J. P. & Olive, P. L. Dna-pk is responsible for enhanced phosphorylation of histone h2ax under hypertonic conditions. *DNA Repair (Amst)* **4**, 1172–1181 (2005). URL http://dx.doi.org/10.1016/j.dnarep.2005.06.005. 38

[285] Tu, W.-Z. *et al.* γh2ax foci formation in the absence of dna damage: mitotic h2ax phosphorylation is mediated by the dna-pkcs/chk2 pathway. *FEBS Lett* **587**, 3437–3443 (2013). URL http://dx.doi.org/10.1016/j.febslet.2013.08.028. 38

[286] Ju, B.-G. *et al.* A topoisomerase iibeta-mediated dsdna break required for regulated transcription. *Science* **312**, 1798–1802 (2006). URL http://dx.doi.org/10.1126/science.1127196. 38

[287] Mayeur, G. L. *et al.* Ku is a novel transcriptional recycling coactivator of the androgen receptor in prostate cancer cells. *J Biol Chem* **280**, 10827–10833 (2005). URL http://dx.doi.org/10.1074/jbc.M413336200. 38

[288] Wong, R. H. F. *et al.* A role of dna-pk for the metabolic gene regulation in response to insulin. *Cell* **136**, 1056–1072 (2009). URL http://dx.doi.org/10.1016/j.cell.2008.12.040. 38

[289] Helt, C. E., Cliby, W. A., Keng, P. C., Bambara, R. A. & O'Reilly, M. A. Ataxia telangiectasia mutated (atm) and atm and rad3-related protein exhibit selective target specificities in response to different forms of dna damage. *J Biol Chem* **280**, 1186–1192 (2005). URL http://dx.doi.org/10.1074/jbc.M410873200. 38

[290] Awasthi, P., Foiani, M. & Kumar, A. Atm and atr signaling at a glance. *J Cell Sci* **128**, 4255–4262 (2015). URL http://dx.doi.org/10.1242/jcs.169730. 38

[291] Xiao, Z. *et al.* Chk1 mediates s and g2 arrests through cdc25a degradation in response to dna-damaging agents. *J Biol Chem* **278**, 21767–21773 (2003). URL http://dx.doi.org/10.1074/jbc.M300229200. 38

[292] Denning, G., Jamieson, L., Maquat, L. E., Thompson, E. A. & Fields, A. P. Cloning of a novel phosphatidylinositol kinase-related kinase: characterization of the human smg-1 rna surveillance protein. *J Biol Chem* **276**, 22709–22714 (2001). URL http://dx.doi.org/10.1074/jbc.C100144200. 38

[293] Foster, K. G. & Fingar, D. C. Mammalian target of rapamycin (mtor): conducting the cellular signaling symphony. *J Biol Chem* **285**, 14071–14077 (2010). URL http://dx.doi.org/10.1074/jbc.R109.094003. 40

[294] Herceg, Z. & Wang, Z.-Q. Rendez-vous at mitosis: Trraped in the chromatin. *Cell Cycle* **4**, 383–387 (2005). 40

[295] Maul, G. G. & Deaven, L. Quantitative determination of nuclear pore complexes in cycling cells with differing dna content. *J Cell Biol* **73**, 748–760 (1977). 41

[296] Lin, D. H. *et al.* Architecture of the symmetric core of the nuclear pore. *Science* **352**, aaf1015 (2016). URL http://dx.doi.org/10.1126/science.aaf1015. 41

[297] Eibauer, M. *et al.* Structure and gating of the nuclear pore complex. *Nat Commun* **6**, 7532 (2015). URL http://dx.doi.org/10.1038/ncomms8532. 41

[298] Knockenhauer, K. E. & Schwartz, T. U. The nuclear pore complex as a flexible and dynamic gate. *Cell* **164**, 1162–1171 (2016). URL http://dx.doi.org/10.1016/j.cell.2016.01.034. 41

[299] Schmidt, H. B. & Görlich, D. Transport selectivity of nuclear pores, phase separation, and membraneless organelles. *Trends Biochem Sci* **41**, 46–61 (2016). URL http://dx.doi.org/10.1016/j.tibs.2015.11.001. 41

[300] Mohr, D., Frey, S., Fischer, T., Güttler, T. & Görlich, D. Characterisation of the passive permeability barrier of nuclear pore complexes. *EMBO J* **28**, 2541–2553 (2009). URL http://dx.doi.org/10.1038/emboj.2009.200. 41, 106

[301] Panté, N. & Kann, M. Nuclear pore complex is able to transport macromolecules with diameters of about 39 nm. *Mol Biol Cell* **13**, 425–434 (2002). URL http://dx.doi.org/10.1091/mbc.01-06-0308. 41

[302] Bauer, N. C., Doetsch, P. W. & Corbett, A. H. Mechanisms regulating protein localization. *Traffic* **16**, 1039–1061 (2015). URL http://dx.doi.org/10.1111/tra.12310. 41, 42, 44, 94

[303] Cook, A., Bono, F., Jinek, M. & Conti, E. Structural biology of nucleocytoplasmic transport. *Annu Rev Biochem* **76**, 647–671 (2007). URL http://dx.doi.org/10.1146/annurev.biochem.76.052705.161529. 42, 108

[304] Quan, Y., Ji, Z.-L., Wang, X., Tartakoff, A. M. & Tao, T. Evolutionary and transcriptional analysis of karyopherin beta superfamily proteins. *Mol Cell Proteomics* **7**, 1254–1269 (2008). URL http://dx.doi.org/10.1074/mcp.M700511-MCP200. 42

[305] Goldfarb, D. S., Corbett, A. H., Mason, D. A., Harreman, M. T. & Adam, S. A. Importin alpha: a multipurpose nuclear-transport receptor. *Trends Cell Biol* **14**, 505–514 (2004). URL http://dx.doi.org/10.1016/j.tcb.2004.07.016. 42

[306] Kutay, U. *et al.* Identification of a trna-specific nuclear export receptor. *Mol Cell* **1**, 359–369 (1998). 42

[307] Lange, A. *et al.* Classical nuclear localization signals: definition, function, and interaction with importin alpha. *J Biol Chem* **282**, 5101–5105 (2007). URL http://dx.doi.org/10.1074/jbc.R600026200. 42

[308] Marfori, M. *et al.* Molecular basis for specificity of nuclear import and prediction of nuclear localization. *Biochim Biophys Acta* **1813**, 1562–1577 (2011). URL http://dx.doi.org/10.1016/j.bbamcr.2010.10.013. 42

[309] Christie, M. *et al.* Structural biology and regulation of protein import into the nucleus. *J Mol Biol* (2015). URL http://dx.doi.org/10.1016/j.jmb.2015.10.023. 42

[310] Xu, D., Farmer, A., Collett, G., Grishin, N. V. & Chook, Y. M. Sequence and structural analyses of nuclear export signals in the nesdb database. *Mol Biol Cell* **23**, 3677–3693 (2012). URL http://dx.doi.org/10.1091/mbc.E12-01-0046. 42

[311] Ganchi, P. A., Sun, S. C., Greene, W. C. & Ballard, D. W. I kappa b/mad-3 masks the nuclear localization signal of nf-kappa b p65 and requires the transactivation domain to inhibit nf-kappa b p65 dna binding. *Mol Biol Cell* **3**, 1339–1352 (1992). 44

[312] Beg, A. A. *et al.* I kappa b interacts with the nuclear localization sequences of the subunits of nf-kappa b: a mechanism for cytoplasmic retention. *Genes Dev* **6**, 1899–1913 (1992). 44

[313] Armstrong, E. H., Goswami, D., Griffin, P. R., Noy, N. & Ortlund, E. A. Structural basis for ligand regulation of the fatty acid-binding protein 5, peroxisome proliferator-activated receptor β/δ (fabp5-pparβ/δ) signaling pathway. *J Biol Chem* **289**, 14941–14954 (2014). URL http://dx.doi.org/10.1074/jbc.M113.514646. 44

[314] Luby-Phelps, K., Hori, M., Phelps, J. M. & Won, D. Ca(2+)-regulated dynamic compartmentalization of calmodulin in living smooth muscle cells. *J Biol Chem* **270**, 21532–21538 (1995). 44

[315] Song, W. *et al.* Proteolytic release and nuclear translocation of notch-1 are induced by presenilin-1 and impaired by pathogenic presenilin-1 mutations. *Proc Natl Acad Sci U S A* **96**, 6959–6963 (1999). 44

[316] Wu, W. *et al.* Alternative splicing controls nuclear translocation of the cell cycle-regulated nek2 kinase. *J Biol Chem* **282**, 26431–26440 (2007). URL http://dx.doi.org/10.1074/jbc.M704969200. 44

[317] Zhang, Y. & Xiong, Y. A p53 amino-terminal nuclear export signal inhibited by dna damage-induced phosphorylation. *Science* **292**, 1910–1915 (2001). URL http://dx.doi.org/10.1126/science.1058637. 45

[318] Lee, H. & Bai, W. Regulation of estrogen receptor nuclear export by ligand-induced and p38-mediated receptor phosphorylation. *Mol Cell Biol* **22**, 5835–5845 (2002). 45

[319] Dietschy, T. *et al.* p300-mediated acetylation of the rothmund-thomson-syndrome gene product recql4 regulates its subcellular localization. *J Cell Sci* **122**, 1258–1267 (2009). URL http://dx.doi.org/10.1242/jcs.037747. 45

[320] Zehorai, E., Yao, Z., Plotnikov, A. & Seger, R. The subcellular localization of mek and erk–a novel nuclear translocation signal (nts) paves a way to the nucleus. *Mol Cell Endocrinol* **314**, 213–220 (2010). URL http://dx.doi.org/10.1016/j.mce.2009.04.008. 45

[321] Gioeli, D. *et al.* Stress kinase signaling regulates androgen receptor phosphorylation, transcription, and localization. *Mol Endocrinol* **20**, 503–515 (2006). URL http://dx.doi.org/10.1210/me.2005-0351. 45

[322] New, L., Jiang, Y. & Han, J. Regulation of prak subcellular location by p38 map kinases. *Mol Biol Cell* **14**, 2603–2616 (2003). URL http://dx.doi.org/10.1091/mbc.E02-08-0538. 45

[323] Engel, K., Kotlyarov, A. & Gaestel, M. Leptomycin b-sensitive nuclear export of mapkap kinase 2 is regulated by phosphorylation. *EMBO J* **17**, 3363–3371 (1998). URL http://dx.doi.org/10.1093/emboj/17.12.3363. 45

[324] Su, C.-H. *et al.* Nuclear export regulation of cop1 by 14-3-3σ in response to dna damage. *Mol Cancer* **9**, 243 (2010). URL http://dx.doi.org/10.1186/1476-4598-9-243. 45, 94, 110

[325] Zhao, X. *et al.* Multiple elements regulate nuclear/cytoplasmic shuttling of foxo1: characterization of phosphorylation- and 14-3-3-dependent and -independent mechanisms. *Biochem J* **378**, 839–849 (2004). URL http://dx.doi.org/10.1042/BJ20031450. 46, 94

[326] Takaishi, H. *et al.* Regulation of nuclear translocation of forkhead transcription factor afx by protein kinase b. *Proc Natl Acad Sci U S A* **96**, 11836–11841 (1999). 46

[327] Matsuzaki, H., Ichino, A., Hayashi, T., Yamamoto, T. & Kikkawa, U. Regulation of intracellular localization and transcriptional activity of foxo4 by protein kinase b through phosphorylation at the motif sites conserved among the foxo family. *J Biochem* **138**, 485–491 (2005). URL http://dx.doi.org/10.1093/jb/mvi146. 46

[328] Beals, C. R., Sheridan, C. M., Turck, C. W., Gardner, P. & Crabtree, G. R. Nuclear export of nf-atc enhanced by glycogen synthase kinase-3. *Science* **275**, 1930–1934 (1997). 46

[329] Beals, C. R., Clipstone, N. A., Ho, S. N. & Crabtree, G. R. Nuclear localization of nf-atc by a calcineurin-dependent, cyclosporin-sensitive intramolecular interaction. *Genes Dev* **11**, 824–834 (1997). 46

[330] Zhu, J. & McKeon, F. Nf-at activation requires suppression of crm1-dependent export by calcineurin. *Nature* **398**, 256–260 (1999). URL http://dx.doi.org/10.1038/18473. 46

[331] Kamentsky, L. *et al.* Improved structure, function and compatibility for cellprofiler: modular high-throughput image analysis software. *Bioinformatics* **27**, 1179–1180 (2011). URL http://dx.doi.org/10.1093/bioinformatics/btr095. 52, 53, 55, 61, 66, 67, 69, 70, 72, 73, 74, 76, 77, 78, 84, 85, 86, 87, 90, 91, 93, 110, 170, 171

[332] Kong, X. *et al.* Comparative analysis of different laser systems to study cellular responses to dna damage in mammalian cells. *Nucleic Acids Res* **37**, e68 (2009). URL http://dx.doi.org/10.1093/nar/gkp221. 53

[333] Jinek, M. *et al.* A programmable dual-rna-guided dna endonuclease in adaptive bacterial immunity. *Science* **337**, 816–821 (2012). URL http://dx.doi.org/10.1126/science.1225829. 60

[334] Rost, B., Yachdav, G. & Liu, J. The predictprotein server. *Nucleic Acids Res* **32**, W321–W326 (2004). URL http://dx.doi.org/10.1093/nar/gkh377. 62, 63

[335] Whitmore, L. & Wallace, B. A. Protein secondary structure analyses from circular dichroism spectroscopy: methods and reference databases. *Biopolymers* **89**, 392–400 (2008). URL http://dx.doi.org/10.1002/bip.20853. 65

[336] Shiloh, Y. & Ziv, Y. The atm protein kinase: regulating the cellular response to genotoxic stress, and more. *Nat Rev Mol Cell Biol* **14**, 197–210 (2013). 65, 113

[337] Wahlberg, E. *et al.* Family-wide chemical profiling and structural analysis of parp and tankyrase inhibitors. *Nat Biotechnol* **30**, 283–288 (2012). URL http://dx.doi.org/10.1038/nbt.2121. 74

[338] Blasius, M. *et al.* A phospho-proteomic screen identifies substrates of the checkpoint kinase chk1. *Genome Biol* **12**, R78 (2011). URL http://dx.doi.org/10.1186/gb-2011-12-8-r78. 79

[339] Matsuoka, S., Huang, M. & Elledge, S. J. Linkage of atm to cell cycle regulation by the chk2 protein kinase. *Science* **282**, 1893–1897 (1998). 79

[340] Bartek, J. & Lukas, J. Mammalian g1- and s-phase checkpoints in response to dna damage. *Curr Opin Cell Biol* **13**, 738–747 (2001). 79

[341] Taus, T. *et al.* Universal and confident phosphorylation site localization using phosphors. *J Proteome Res* **10**, 5354–5362 (2011). URL http://dx.doi.org/10.1021/pr200611n. 81

[342] Amanchy, R. *et al.* A curated compendium of phosphorylation motifs. *Nat Biotechnol* **25**, 285–286 (2007). URL http://dx.doi.org/10.1038/nbt0307-285. 83, 89

[343] Panni, S. *et al.* Combining peptide recognition specificity and context information for the prediction of the 14-3-3-mediated interactome in s. cerevisiae and h. sapiens. *Proteomics* **11**, 128–143 (2011). URL http://dx.doi.org/10.1002/pmic.201000030. 83, 96, 97

[344] Meier, I. & Somers, D. E. Regulation of nucleocytoplasmic trafficking in plants. *Curr Opin Plant Biol* **14**, 538–546 (2011). URL http://dx.doi.org/10.1016/j.pbi.2011.06.005. 94

[345] Macara, I. G. Transport into and out of the nucleus. *Microbiol Mol Biol Rev* **65**, 570–94, table of contents (2001). URL http://dx.doi.org/10.1128/MMBR.65.4.570-594.2001. 94

[346] Muslin, A. J. & Xing, H. 14-3-3 proteins: regulation of subcellular localization by molecular interference. *Cell Signal* **12**, 703–709 (2000). 94

[347] Tzivion, G. & Avruch, J. 14-3-3 proteins: active cofactors in cellular regulation by serine/threonine phosphorylation. *J Biol Chem* **277**, 3061–3064 (2002). URL http://dx.doi.org/10.1074/jbc.R100059200. 94

[348] Aghazadeh, Y. & Papadopoulos, V. The role of the 14-3-3 protein family in health, disease, and drug development. *Drug Discov Today* **21**, 278–287 (2016). URL http://dx.doi.org/10.1016/j.drudis.2015.09.012. 94

[349] Yang, X. *et al.* Structural basis for protein-protein interactions in the 14-3-3 protein family. *Proc Natl Acad Sci U S A* **103**, 17237–17242 (2006). URL http://dx.doi.org/10.1073/pnas.0605779103. 94

[350] Johnson, C. *et al.* Bioinformatic and experimental survey of 14-3-3-binding sites. *Biochem J* **427**, 69–78 (2010). URL http://dx.doi.org/10.1042/BJ20091834. 94, 95

[351] Nomura, M. *et al.* 14-3-3 interacts directly with and negatively regulates pro-apoptotic bax. *J Biol Chem* **278**, 2058–2065 (2003). URL http://dx.doi.org/10.1074/jbc.M207880200. 94

[352] Ge, Q. *et al.* Structural characterization of a unique interface between carbohydrate response element-binding protein (chrebp) and 14-3-3β protein. *J Biol Chem* **287**, 41914–41921 (2012). URL http://dx.doi.org/10.1074/jbc.M112.418855. 94

[353] Ottmann, C. *et al.* Phosphorylation-independent interaction between 14-3-3 and exoenzyme s: from structure to pathogenesis. *EMBO J* **26**, 902–913 (2007). URL http://dx.doi.org/10.1038/sj.emboj.7601530. 94

[354] Dunaway, S., Liu, H.-Y. & Walworth, N. C. Interaction of 14-3-3 protein with chk1 affects localization and checkpoint function. *J Cell Sci* **118**, 39–50 (2005). URL http://dx.doi.org/10.1242/jcs.01570. 94

[355] Brunet, A. *et al.* 14-3-3 transits to the nucleus and participates in dynamic nucleocytoplasmic transport. *J Cell Biol* **156**, 817–828 (2002). URL http://dx.doi.org/10.1083/jcb.200112059. 94

[356] Obsilova, V. *et al.* 14-3-3 protein interacts with nuclear localization sequence of forkhead transcription factor foxo4. *Biochemistry* **44**, 11608–11617 (2005). URL http://dx.doi.org/10.1021/bi050618r. 94, 95

[357] Mohammad, D. H. & Yaffe, M. B. 14-3-3 proteins, fha domains and brct domains in the dna damage response. *DNA Repair (Amst)* **8**, 1009–1017 (2009). URL http://dx.doi.org/10.1016/j.dnarep.2009.04.004. 95, 149

[358] Uversky, V. N. Intrinsically disordered proteins from a to z. *Int J Biochem Cell Biol* **43**, 1090–1103 (2011). URL http://dx.doi.org/10.1016/j.biocel.2011.04.001. 105

[359] Lee, H. *et al.* Local structural elements in the mostly unstructured transcriptional activation domain of human p53. *J Biol Chem* **275**, 29426–29432 (2000). URL http://dx.doi.org/10.1074/jbc.M003107200. 105

[360] Gsponer, J., Futschik, M. E., Teichmann, S. A. & Babu, M. M. Tight regulation of unstructured proteins: from transcript synthesis to protein degradation. *Science* **322**, 1365–1368 (2008). URL http://dx.doi.org/10.1126/science.1163581. 105

[361] Xie, H. *et al.* Functional anthology of intrinsic disorder. 1. biological processes and functions of proteins with long disordered regions. *J Proteome Res* **6**, 1882–1898 (2007). URL http://dx.doi.org/10.1021/pr060392u. 105

[362] Ribbeck, K. & Görlich, D. Kinetic analysis of translocation through nuclear pore complexes. *EMBO J* **20**, 1320–1330 (2001). URL http://dx.doi.org/10.1093/emboj/20.6.1320. 106

[363] Cardarelli, F., Tosti, L., Serresi, M., Beltram, F. & Bizzarri, R. Fluorescent recovery after photobleaching (frap) analysis of nuclear export rates identifies intrinsic features of nucleocytoplasmic transport. *J Biol Chem* **287**, 5554–5561 (2012). URL http://dx.doi.org/10.1074/jbc.M111.304899. 106

[364] Kudo, N., Taoka, H., Toda, T., Yoshida, M. & Horinouchi, S. A novel nuclear export signal sensitive to oxidative stress in the fission yeast transcription factor pap1. *J Biol Chem* **274**, 15151–15158 (1999). 108

[365] Colwell, L. J., Brenner, M. P. & Ribbeck, K. Charge as a selection criterion for translocation through the nuclear pore complex. *PLoS Comput Biol* **6**, e1000747 (2010). URL http://dx.doi.org/10.1371/journal.pcbi.1000747. 108

[366] Zhou, X., Mester, C., Stemmer, P. M. & Reid, G. E. Oxidation-induced conformational changes in calcineurin determined by covalent labeling and tandem mass spectrometry. *Biochemistry* **53**, 6754–6765 (2014). URL http://dx.doi.org/10.1021/bi5009744. 109

[367] Lübker, C. *et al.* Membranous adenylyl cyclase 1 activation is regulated by oxidation of n- and c-terminal methionine residues in calmodulin. *Biochem Pharmacol* **93**, 196–209 (2015). URL http://dx.doi.org/10.1016/j.bcp.2014.11.007. 109

[368] Lim, J. C., Kim, G. & Levine, R. L. Stereospecific oxidation of calmodulin by methionine sulfoxide reductase a. *Free Radic Biol Med* **61**, 257–264 (2013). URL http://dx.doi.org/10.1016/j.freeradbiomed.2013.04.004. 109, 110

[369] Dornan, D. *et al.* Atm engages autodegradation of the e3 ubiquitin ligase cop1 after dna damage. *Science* **313**, 1122–1126 (2006). URL http://dx.doi.org/10.1126/science.1127335. 110, 111

[370] Traven, A. & Heierhorst, J. Sq/tq cluster domains: concentrated atm/atr kinase phosphorylation site regions in dna-damage-response proteins. *Bioessays* **27**, 397–407 (2005). URL http://dx.doi.org/10.1002/bies.20204. 111

[371] Cara, L., Baitemirova, M., Follis, J., Larios-Sanz, M. & Ribes-Zamora, A. The atm- and atr-related scd domain is over-represented in proteins involved in nervous system development. *Sci Rep* **6**, 19050 (2016). URL http://dx.doi.org/10.1038/srep19050. 111, 156

[372] Corpet, F. Multiple sequence alignment with hierarchical clustering. *Nucleic Acids Res* **16**, 10881–10890 (1988). 111

[373] Loseva, O. *et al.* Parp-3 is a mono-adp-ribosylase that activates parp-1 in the absence of dna. *J Biol Chem* **285**, 8054–8060 (2010). URL http://dx.doi.org/10.1074/jbc.M109.077834. 113

[374] Beck, C. *et al.* Parp3 affects the relative contribution of homologous recombination and nonhomologous end-joining pathways. *Nucleic Acids Res* **42**, 5616–5632 (2014). URL http://dx.doi.org/10.1093/nar/gku174. 113

[375] Sakai, D., Dixon, J., Achilleos, A., Dixon, M. & Trainor, P. A. Prevention of treacher collins syndrome craniofacial anomalies in mouse models via maternal antioxidant supplementation. *Nat Commun* **7**, 10328 (2016). URL http://dx.doi.org/10.1038/ncomms10328. 117

[376] Walter, D. *et al.* Exit from dormancy provokes dna-damage-induced attrition in haematopoietic stem cells. *Nature* **520**, 549–552 (2015). URL http://dx.doi.org/10.1038/nature14131. 117

[377] Gunn, A. & Stark, J. M. I-scei-based assays to examine distinct repair outcomes of mammalian chromosomal double strand breaks. *Methods Mol Biol* **920**, 379–391 (2012). URL http://dx.doi.org/10.1007/978-1-61779-998-3_27. 118

[378] Vivelo, C. A., Wat, R., Agrawal, C., Tee, H. Y. & Leung, A. K. L. Adpribodb: The database of adp-ribosylated proteins. *Nucleic Acids Res* (2016). URL http://dx.doi.org/10.1093/nar/gkw706. 119

[379] Schwanhäusser, B. *et al.* Global quantification of mammalian gene expression control. *Nature* **473**, 337–342 (2011). URL http://dx.doi.org/10.1038/nature10098. 128, 185

[380] Cox, J. *et al.* Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed maxlfq. *Mol Cell Proteomics* **13**, 2513–2526 (2014). URL http://dx.doi.org/10.1074/mcp.M113.031591. 129

[381] Forst, A. H. *et al.* Recognition of mono-adp-ribosylated artd10 substrates by artd8 macrodomains. *Structure* **21**, 462–475 (2013). URL http://dx.doi.org/10.1016/j.str.2012.12.019. 136

[382] Lis, J. T. & Kraus, W. L. Promoter cleavage: a topoiibeta and parp-1 collaboration. *Cell* **125**, 1225–1227 (2006). URL http://dx.doi.org/10.1016/j.cell.2006.06.016. 138

[383] Matsuki, H. *et al.* Both g3bp1 and g3bp2 contribute to stress granule formation. *Genes Cells* **18**, 135–146 (2013). URL http://dx.doi.org/10.1111/gtc.12023. 139, 153

[384] Oliveira, M. T. & Kaguni, L. S. Functional roles of the n- and c-terminal regions of the human mitochondrial single-stranded dna-binding protein. *PLoS One* **5**, e15379 (2010). URL http://dx.doi.org/10.1371/journal.pone.0015379. 140

[385] Mehus, A. A., Anderson, R. H. & Roux, K. J. Bioid identification of lamin-associated proteins. *Methods Enzymol* **569**, 3–22 (2016). URL http://dx.doi.org/10.1016/bs.mie.2015.08.008. 142, 151

[386] Scovassi, A. I. & Poirier, G. G. Poly(adp-ribosylation) and apoptosis. *Mol Cell Biochem* **199**, 125–137 (1999). 143

[387] Koike-Yusa, H., Li, Y., Tan, E.-P., Velasco-Herrera, M. D. C. & Yusa, K. Genome-wide recessive genetic screening in mammalian cells with a lentiviral crispr-guide rna library. *Nat Biotechnol* **32**, 267–273 (2014). URL http://dx.doi.org/10.1038/nbt.2800. 151

[388] Mellacheruvu, D. *et al.* The crapome: a contaminant repository for affinity purification-mass spectrometry data. *Nat Methods* **10**, 730–736 (2013). URL http://dx.doi.org/10.1038/nmeth.2557. 152

[389] Di Giammartino, D. C., Shi, Y. & Manley, J. L. Parp1 represses pap and inhibits polyadenylation during heat shock. *Mol Cell* **49**, 7–17 (2013). URL http://dx.doi.org/10.1016/j.molcel.2012.11.005. 152

[390] Ronan, J. L., Wu, W. & Crabtree, G. R. From neural development to cognition: unexpected roles for chromatin. *Nat Rev Genet* **14**, 347–359 (2013). URL http://dx.doi.org/10.1038/nrg3413. 153, 154

[391] Choi, K.-Y., Yoo, M. & Han, J.-H. Toward understanding the role of the neuron-specific baf chromatin remodeling complex in memory formation. *Exp Mol Med* **47**, e155 (2015). URL http://dx.doi.org/10.1038/emm.2014.129. 154

[392] Santen, G. W. E. *et al.* Mutations in swi/snf chromatin remodeling complex gene arid1b cause coffin-siris syndrome. *Nat Genet* **44**, 379–380 (2012). URL http://dx.doi.org/10.1038/ng.2217. 154

[393] Neale, B. M. *et al.* Patterns and rates of exonic de novo mutations in autism spectrum disorders. *Nature* **485**, 242–245 (2012). URL http://dx.doi.org/10.1038/nature11011. 154

[394] O'Roak, B. J. *et al.* Sporadic autism exomes reveal a highly interconnected protein network of de novo mutations. *Nature* **485**, 246–250 (2012). URL http://dx.doi.org/10.1038/nature10989. 154

[395] Wu, J. I. *et al.* Regulation of dendritic development by neuron-specific chromatin remodeling complexes. *Neuron* **56**, 94–108 (2007). URL http://dx.doi.org/10.1016/j.neuron.2007.08.021. 154

[396] Tea, J. S. & Luo, L. The chromatin remodeling factor bap55 functions through the tip60 complex to regulate olfactory projection neuron dendrite targeting. *Neural Dev* **6**, 5 (2011). URL http://dx.doi.org/10.1186/1749-8104-6-5. 154

[397] Hang, C. T. *et al.* Chromatin regulation by brg1 underlies heart muscle development and disease. *Nature* **466**, 62–67 (2010). URL http://dx.doi.org/10.1038/nature09130. 155

[398] Keppler, B. R. & Archer, T. K. Ubiquitin-dependent and ubiquitin-independent control of subunit stoichiometry in the swi/snf complex. *J Biol Chem* **285**, 35665–35674 (2010). URL http://dx.doi.org/10.1074/jbc.M110.173997. 155

[399] Bai, C.-Y., Ohsugi, M., Abe, Y. & Yamamoto, T. Zrp-1 controls rho gtpase-mediated actin reorganization by localizing at cell-matrix and cell-cell adhesions. *J Cell Sci* **120**, 2828–2837 (2007). URL http://dx.doi.org/10.1242/jcs.03477. 155

[400] Yarar, D., Waterman-Storer, C. M. & Schmid, S. L. Snx9 couples actin assembly to phosphoinositide signals and is required for membrane remodeling during endocytosis. *Dev Cell* **13**, 43–56 (2007). URL http://dx.doi.org/10.1016/j.devcel.2007.04.014. 155

[401] Eira, J., Silva, C. S., Sousa, M. M. & Liz, M. A. The cytoskeleton as a novel therapeutic target for old neurodegenerative disorders. *Prog Neurobiol* (2016). URL http://dx.doi.org/10.1016/j.pneurobio.2016.04.007. 156

[402] Lei, W., Omotade, O. F., Myers, K. R. & Zheng, J. Q. Actin cytoskeleton in dendritic spine development and plasticity. *Curr Opin Neurobiol* **39**, 86–92 (2016). URL http://dx.doi.org/10.1016/j.conb.2016.04.010. 156

[403] Chen, D. *et al.* Identification of a nuclear protein that promotes nf-kappab activation. *Biochem Biophys Res Commun* **310**, 720–724 (2003). 157

[404] Jin, D. Y., Chae, H. Z., Rhee, S. G. & Jeang, K. T. Regulatory role for a novel human thioredoxin peroxidase in nf-kappab activation. *J Biol Chem* **272**, 30952–30961 (1997). 157

[405] Kassel, O. *et al.* A nuclear isoform of the focal adhesion lim-domain protein trip6 integrates activating and repressing signals at ap-1- and nf-kappab-regulated promoters. *Genes Dev* **18**, 2518–2528 (2004). URL http://dx.doi.org/10.1101/gad.322404. 157

[406] Prigent, M., Barlat, I., Langen, H. & Dargemont, C. Ikappabalpha and ikappabalpha /nf-kappa b complexes are retained in the cytoplasm through interaction with a novel partner, rasgap sh3-binding protein 2. *J Biol Chem* **275**, 36441–36449 (2000). URL http://dx.doi.org/10.1074/jbc.M004751200. 158

[407] Jensen, S. S. & Larsen, M. R. Evaluation of the impact of some experimental procedures on different phosphopeptide enrichment techniques. *Rapid Commun Mass Spectrom* **21**, 3635–3645 (2007). URL http://dx.doi.org/10.1002/rcm.3254. 179

[408] Cox, J. *et al.* Andromeda: a peptide search engine integrated into the maxquant environment. *J Proteome Res* **10**, 1794–1805 (2011). URL http://dx.doi.org/10.1021/pr101065j. 185