MEDICAL UNIVERSITY OF VIENNA

Pharmacological perturbation of the phase-separating protein SMNDC1

Doctoral thesis at the Medical University of Vienna
for obtaining the academic degree

Doctor of Philosophy

Submitted by
Lennart Enders, MSc

Supervisor:
Stefan Kubicek, PhD
CeMM Research Center for Molecular Medicine
of the Austrian Academy of Sciences
Lazarettgasse 14, AKH BT 25.3
1090 Vienna, Austria

Vienna, 06/2023

DECLARATION

The work presented in this doctoral thesis was performed at the Research Center for Molecular Medicine (CeMM) of the Austrian Academy of Sciences in Vienna, Austria unless otherwise stated below.

The results detailed in Section 2 have been submitted as a manuscript to Nature Communications and a revised version is currently under revision:

Pharmacological perturbation of the phase-separating protein SMNDC1

Lennart Enders[1], Marton Siklos[1], Jan Borggräfe[2,3], Stefan Gaussmann[2,3], Anna Koren[1], Monika Malik[1], Tatjana Tomek[1], Michael Schuster[1], Jiří Reiniš[1], Elisa Hahn[1], Andrea Rukavina[1], Andreas Reicher[1], Tamara Casteels[1,$], Christoph Bock[1,4], Georg E. Winter[1], J. Thomas Hannich[1], Michael Sattler[2,3], and Stefan Kubicek[1,*]

[1]CeMM Research Center for Molecular Medicine of the Austrian Academy of Sciences, Lazarettgasse 14, 1090 Vienna, Austria.
[2]Helmholtz Munich, Molecular Targets and Therapeutics Center, Institute of Structural Biology, Neuherberg, 85764 München, Germany.
[3]Technical University of Munich, TUM School of Natural Sciences, Department of Bioscience, Bavarian NMR Center, Garching, 85748 München, Germany.
[4]Medical University of Vienna, Institute of Artificial Intelligence, Center for Medical Data Science, Vienna, Austria.

*Correspondence: Stefan Kubicek, skubicek@cemm.oeaw.ac.at.
[$]Present Address: Sloan Kettering Institute, 1275 York Avenue, New York, NY 10065, USA

L.E. and S.K. planned the study and designed the experiments; **L.E.** performed most of the experiments with help from A.K., M.M., T.T., A.Re. and T.C.; M.Si. performed chemical syntheses and SAR studies; J.B. and S.G. performed NMR experiments; E.H. and A.Ru. performed biotin-purification and MS sample preparation; **L.E.**, J.R., M.S., A.K., M.M., and A. Ru. analyzed the data; **L.E.** and S.K. wrote the manuscript with input from all co-authors; C.B., G.E.W., J.T.H., M.Sa., and S.K. supervised the work.

TABLE OF CONTENTS

Declaration ... **ii**

Table of Contents ... **iii**

List of Figures ... **v**

Abstract ... **vi**

Zusammenfassung ... **viii**

Publications Arising From This Thesis ... **x**

Abbreviations ... **xi**

Acknowledgements ... **xiv**

1. Introduction ... **1**

1.1 Phase-separation as a novel paradigm for the control of gene expression ... 1

1.1.1 General concepts of phase-separation ... *1*

1.1.2 Phase-separation in transcriptional control ... *3*

1.1.3 Post-transcriptional gene regulation via phase-separation ... *5*

1.1.4 Nuclear speckles: an example of a phase-separated compartment within the nucleus involved in gene regulation ... *6*

1.1.5 Phase-separation in disease ... *9*

1.1.6 Chemical perturbation of phase-separated compartments ... *12*

1.1.7 SMNDC1 and SMN – new phase-separating Tudor domain proteins as drug targets? ... *15*

1.2 Methyl readers and their inhibitors ... 19

1.2.1 Post-translational modification of proteins – methylation ... *19*

1.2.2 Domains that recognize methylated proteins ... *21*

1.2.3 Methyl-binding domains besides Tudor domains ... *22*

1.2.4 Tudor domains ... *23*

1.2.5 Inhibitors of methyl-reader domains or proteins ... *25*

1.3 Aims ... 35

2. Results ... **36**

3. Discussion 100

3.1 General Discussion 100

3.1.1 SMNDC1's phase separation behavior 100

3.1.2 Pharmacological perturbation of SMNDC1 101

3.1.3 SMNDC1 inhibition – immediate and later effects 101

3.1.4 Comparison of SMNDC1 inhibitors to other Tudor domain inhibitors 102

3.1.5 Identification of off-targets 103

3.1.6 Further development and derivatization of the inhibitor 103

3.1.7 Potential therapeutic applications 104

3.2 Conclusion and Future Prospects 106

4. Materials and Methods 109

5. References 110

6. Curriculum Vitae 134

LIST OF FIGURES

Figure 1: Phase-separated compartments in the nucleus. 2
Figure 2: Gene expression regulation happening in nuclear speckles. 8
Figure 3: Illustration of the potential scenarios in which disease may emerge as a result of anomalous phase-separation. 10
Figure 4: Aims (a) and approaches (b) of condensate-modifying therapeutics. 15
Figure 5: Structure of SMN1's and SMNDC1's Tudor domains and SMNDC1's role in the regulation of insulin expression. 19
Figure 6: Methylations of Lysine and Arginine. 20
Figure 7: Examples of histone methylation readers. 22
Figure 8: Inhibitors against Tudor domain protein Spindlin1. 28
Figure 9: Inhibitors against TP53BP1 and UHRF1. 30
Figure 10: Inhibitors against TDRD3, KDM4A, SETDB1, PHF1 and SMN1. 33
Figure 11: Human tudor domain phylogenetic tree. 34
Figure 12: A mouse model of SMA shows changes in the cell type composition of pancreatic islets. 105
Figure 13: SMNDC1 expression compared across human tissues. 106
Figure 14: A model for the mechanisms behind SMNDC1's phase separation and its pharmacological perturbation. 108

ABSTRACT

In recent years, a novel concept for cellular organization has emerged besides compartments formed by lipid bilayer membranes. It is based on the separation of different phases from a homogenous liquid mixture within an organelle in a process that cells appear to utilize to optimize their complex biochemical reactions. These phase separated compartments regulate processes surrounding the central dogma of molecular biology – the flow of information from DNA to protein via RNA. An increasing body of evidence suggests that gene expression is influenced and controlled by liquid-liquid phase separation.

Proteins containing Tudor domains and their binding partners, proteins containing methylated arginines, appear to be particularly involved in the formation of phase separated compartments. SMNDC1, an essential splicing factor, contains a Tudor domain and was shown to regulate insulin expression in α-cells, thus representing a potential therapeutical target in diabetes.

To better understand SMNDC1's role in gene regulation we generated cell lines with a fluorescent tag fused to the endogenous protein. We found that SMNDC1 localizes to the phase separated compartment of nuclear speckles and can rapidly diffuse within as shown by fluorescence recovery after photobleaching. Furthermore, SMNDC1 can undergo liquid-liquid phase separation on its own *in vitro*, in a process that is enhanced by RNA. Proximity labeling to characterize SMNDC1's interactome confirmed the protein's association with other speckle factors.

To be able to perturb SMNDC1's function in a doseable and time-dependent manner we aimed at developing pharmacological inhibitors of the protein. We developed an AlphaScreen with SMNDC1's Tudor domain and a binding peptide containing symmetrically dimethylated arginines. Screening a 90,000-compound library for inhibitors of this interaction we identified a number of promising lead compounds. Testing a large number of analogues of the most promising hits led to a thorough understanding of structure activity relationships and compounds with selectivity towards SMNDC1 compared to the structurally very similar Tudor domain of SMN. In a collaboration with the Sattler lab from the Technical University of Munich we could confirm the binding of a fragment of the inhibitor to the Tudor domain by a nuclear magnetic resonance structure and elucidate intermolecular interactions.

Finally, we characterized the effect of the inhibitor on cells. SMNDC1's distribution changed from a speckled to a more diluted localization. The nuclear speckle marker SRRM2 changed its localization accordingly. A majority of SMNDC1's interactions were lost upon inhibitor treatment, especially to nuclear speckle proteins and proteins containing symmetrically dimethylated arginines. Additionally, global changes to alternative splicing patterns were observed upon treatment with the novel SMNDC1 inhibitor.

Overall, I have expanded the understanding of SMNDC1's function and present novel SMNDC1 Tudor domain inhibitors which can be developed further, potentially leading to therapeutic application.

ZUSAMMENFASSUNG

In den letzten Jahren hat sich neben den durch Lipid-Doppelschichtmembranen gebildeten Kompartimenten ein neues Konzept für die Zellorganisation herauskristallisiert. Es basiert auf der Trennung verschiedener Phasen aus einem homogenen Flüssigkeitsgemisch innerhalb von Organellen, und Zellen nutzen dies, um ihre komplexen biochemischen Reaktionen zu optimieren. Phasenseparierung reguliert Prozesse rund um das zentrale Dogma der Molekularbiologie - den Informationsfluss von der DNA zum Protein über die RNA. Immer mehr Hinweise deuten darauf hin, dass die Regulierung der Genexpression durch Flüssig-Flüssig-Phasentrennung beeinflusst und gesteuert wird.

Proteine, die Tudor-Domänen enthalten, und ihre Bindungspartner, Proteine mit methylierten Argininen, scheinen besonders häufig an der Bildung von phasenseparierten Kompartimenten beteiligt zu sein. SMNDC1, ein Tudor-Domänen-Protein und essentieller Spleißfaktor, ist an der Regulierung der Insulinexpression in α-Zellen beteiligt und stellt somit ein potentielles therapeutisches Target bei Diabetes dar.

Um die Rolle von SMNDC1 bei der Genregulation besser zu verstehen, generierten wir Zelllinien, die endogenes SMNDC1 mit einem Fluoreszenz-Tag fusionieren. Wir fanden heraus, dass SMNDC1 im phasengetrennten Kompartiment der Nuclear Speckles lokalisiert ist und dort schnell diffundieren kann, wie fluorescence recovery after photobleaching zeigt. Darüber hinaus erfährt SMNDC1 *in vitro* ohne andere Proteine eine Flüssig-Flüssig-Phasentrennung, die durch RNA verstärkt wird. Das Interaktom von SMNDC1, das wir durch Proximity Labeling charakterisierten, bestätigte die Assoziation mit anderen Speckle Proteinen.

Um die Funktion von SMNDC1 in dosierbarer und zeitabhängiger Weise zu perturbieren, entwickelten wir Inhibitoren für das Protein. Wir etablierten einen AlphaScreen mit der Tudor-Domäne von SMNDC1 und einem bindenden Peptid, das symmetrisch dimethylierte Arginine enthält. Durch Screening einer Bibliothek mit 90.000 Verbindungen nach Inhibitoren dieser Interaktion konnten wir eine Reihe vielversprechender Leitstrukturen identifizieren. Die Prüfung einer großen Anzahl von Analoga der vielversprechendsten Substanzen führte zu einem genauen Verständnis der Struktur-Wirkungsbeziehung und zu Verbindungen mit Selektivität für SMNDC1 im Vergleich zur strukturell sehr ähnlichen Tudor-Domäne von SMN. In einer Zusammenarbeit mit dem Sattler-Labor der Technischen Universität München konnten wir die Bindung eines Fragments des Inhibitors an die Tudor-Domäne durch eine Kernspinresonanz-Struktur bestätigen und intermolekulare Wechselwirkungen aufklären.

Schließlich charakterisierten wir die Wirkung des Inhibitors auf die Zellen. Die Verteilung von SMNDC1 änderte sich von Nuclear Speckles zu einer diffuseren Lokalisierung. Auch der Nuclear-Speckle-Marker SRRM2 änderte seine Lokalisierung entsprechend. Ein Großteil der

Interaktionen von SMNDC1 ging nach der Behandlung mit dem Inhibitor verloren, insbesondere die Interaktionen mit Nuclear-Speckle-Proteinen und Proteinen, die symmetrisch dimethylierte Arginine enthalten. Darüber hinaus wurden bei der Behandlung mit dem neuen SMNDC1-Inhibitor globale Veränderungen der alternativen Spleißmuster beobachtet.

Insgesamt erweitere ich mit dieser Arbeit das Verständnis der Funktion von SMNDC1 und stelle neuartige SMNDC1-Tudor-Domänen-Inhibitoren vor, die weiterentwickelt werden und zu therapeutischen Anwendungen führen können.

PUBLICATIONS ARISING FROM THIS THESIS

The results detailed in Section 2 have been submitted as a manuscript to Nature Communications and a revised version is currently under revision:

Pharmacological perturbation of the phase-separating protein SMNDC1
Lennart Enders, Marton Siklos, Jan Borggräfe, Stefan Gaussmann, Anna Koren, Monika Malik, Tatjana Tomek, Michael Schuster, Jiří Reiniš, Elisa Hahn, Andrea Rukavina, Andreas Reicher, Tamara Casteels, Christoph Bock, Georg E. Winter, J. Thomas Hannich, Michael Sattler, and Stefan Kubicek

ABBREVIATIONS

		Abbreviation	Meaning
A	….	aDMA	ω-N^G,N^G-asymmetric dimethylarginine
		ALS	amyotrophic lateral sclerosis
		ATM	ataxia telangiectasia mutated
C	….	CB	Cajal body
		CDK	cyclin-dependent kinase
		CDY	chromodomain Y chromosome
		CHD	chromo-ATPase/helicase-DNA-binding
		CHEK1	check point kinase 1
		CHEK2	check point kinase 2
		CTD	C-terminal domain
D	….	DNMT1	DNA (cytosine-5)-methyltransferase 1
F	….	FRAP	fluorescence recovery after photobleaching
H	….	HP1	heterochromatin protein 1
		HLB	histone locus bodies
		hnRNP	Heterogeneous nuclear ribonucleoprotein
I	….	IAPP	Islet Amyloid Polypeptide
		IDR	intrinsically disordered region
		IGC	interchromatin granule cluster
		IP	immunoprecipitation
		IRS-1	insulin receptor substrate 1
		ITC	isothermal titration calorimetry
J	….	Jmj	Jumonji
K	….	KDM	lysine-specific histone demethylase
		KMT	lysine-specific methyltransferase
L	….	LC	low-complexity
		LLPS	liquid-liquid phase separation
		lncRNA	long noncoding RNA
		LOF	loss-of-function
M	….	m^6A	N^6-methyladenosine
		MBT	Malignant Brain Tumor
		miRNA	microRNA

		MMA	ω-N^{G}-monomethylarginine
N	….	NMR	Nuclear Magnetic Resonance
P	….	PcG	Polycomb-group
		PCL	Polycomb-like
		PDB	Protein Data Bank
		PHD	plant homeodomain
		piRNA	Piwi-interacting RNA
		PML	promyelocytic leukaemia
		Pol II	RNA polymerase II
		PRC2	Polycomb repressive complex 2
		PRMT	Protein Arginine methyltransferase
		PROTAC	Proteolysis-Targeting Chimeric Molecule
		PTM	Post-translational modification
R	….	RBP	RNA-binding protein
		RING	Really Interesting New Gene
		RISC	RNA-induced silencing complex
		RNP	Ribonucleoprotein
S	….	SAM	*S*-Adenosylmethionine
		SAR	structure-activity relationship
		sDMA	ω-N^{G},N'^{G}-symmetric dimethylarginine
		SET	Su(var)3-9, Enhancer of Zeste, Trithorax
		SICC	senescence-induced CTCF cluster
		siRNA	Small interfering RNA
		SMA	spinal muscular atrophy
		SMN1	Survival of motor neuron 1
		SMNDC1	Survival motor neuron domain containing 1
		SMNrp	SMN related protein
		SPF30	Survival of motor neuron-related-splicing factor 30
T	….	T2D	Type 2 Diabetes
		TAF14	TBP associated factor 14
		TDRD6	Multi-Tudor domain-containing protein
		TP53BP1	Tumor protein p53-binding protein 1
		TSA-Seq	Tyramide signal amplification-sequencing
		TTD	tandem Tudor domain
U	….	UHRF1	ubiquitin-like, containing PHD and RING finger domains 1

W	….	WDR	WD40 repeat
0-9	….	5A-DMP	5-amino-2,4-dimethylpyridine

ACKNOWLEDGEMENTS

Looking back at the past years that I have spent pursuing a PhD and living in Vienna fills me with an overwhelming gratefulness. I am thankful to all the people who have supported and accompanied me on this journey, and I would like to mention a few here below.

First of all, I want to thank my supervisor Stefan. Thank you for believing in me and giving me the opportunity to start this adventure almost 6 years ago. You have truly supported me along the whole way with your optimism and you are a great teacher. Whenever I needed something, be it feedback for presentations, manuscripts, and ideas or just a quick question on experiment design – you took the time. You make the Kubicek lab a fantastic place to work!

Of course, the lab wouldn't be nearly as great without its past and present members. Thank you for all the help in the lab, lunches I'm looking forward to every day, bouldering and ice skating, hiking, spicy dinners, Aperol on the terrace,… A special thank you to Tamara, Rob and Brenda as the original beta cell team for all the fun we had at work and play.

Even such a great group as the Kubicek lab cannot create the great atmosphere alone that I enjoyed at work. Thank you to the whole CeMM community for making this institute so cooperative, stimulating, and well-organized, and not to forget fun!

This research project would not have been possible without the hard work of other people, most of them are co-authors on the manuscript. I would like to especially thank Marton for all his efforts with synthesizing and choosing molecules for yet another batch of analogs and derivatives.

Last but definitely not least, I want to thank my family and friends for sharing this adventure with me. After 6 years and even though some dear people have moved on already, I feel like being part of a community in Vienna which makes this great city my home.
Jana – you make me happy every single day and I'm excited to see what will come next on our joint ride.
To my parents, thank you for your never-ending unconditional love and support that I can feel no matter where I go.

To all of you, thank you.

1. INTRODUCTION

1.1 Phase-separation as a novel paradigm for the control of gene expression

1.1.1 General concepts of phase-separation

Cells must form compartments to efficiently organize their complex biochemical reactions. Within these compartments, molecules can diffuse freely to meet their reaction partner. Compartment boundaries limit diffusion in order to prevent molecules meeting the “wrong” reaction partner, while enhancing reaction rates within the compartment. In a classical understanding, these boundaries are formed by lipid bilayers creating organelles such as the nucleus, mitochondria, etc. In recent years, it has been shown that cells organize their biochemical functions not only in the well-known membrane-separated organelles, but also in liquid-liquid separated phases without a membrane (Hyman *et al*, 2014).

In contrast to cross-linked aggregates like glycogen granules (Stubbe *et al*, 2005) that behave more like a solid, within such phase separated compartments diffusion is still possible to ensure that biochemical reactions can take place. This is described as a liquid-like state. Liquid-liquid phase separation (LLPS) describes a change in the distribution of a certain molecule (in biology typically a protein) in which there exists a dense phase and a less dense phase with very different local concentrations of the molecule and potential interactors. The dense phase still behaves like a liquid with relatively free diffusion of the individual molecules. These distribution changes are typically observed in a phase diagram in which concentration and other parameters like temperature, salt concentration or pH change. Further changing parameters can result in a change of material properties e.g., the transition from liquid-like to gel-like which behaves more like a solid or even crystal-like with a high degree of order (Alberti & Hyman, 2016).

A number of suborganelle structures have long been known based on the patterns they form in microscopy imaging, but only recently it was recognized that they are formed by LLPS. The earliest example shown to undergo LLPS is the P granule formed at multiple spots during early *C. elegans* development (Brangwynne *et al*, 2009). Another phase-separated assembly is the nucleolus, one of many sub-compartments in the nucleus and the site of ribosome formation (Brangwynne *et al*, 2011; Sleeman & Trinkle-Mulcahy, 2014; Feric *et al*, 2016). An overview of the different phase-separated compartments within the nucleus can be found in Palikyras & Papantonis, 2019 (Figure 1).

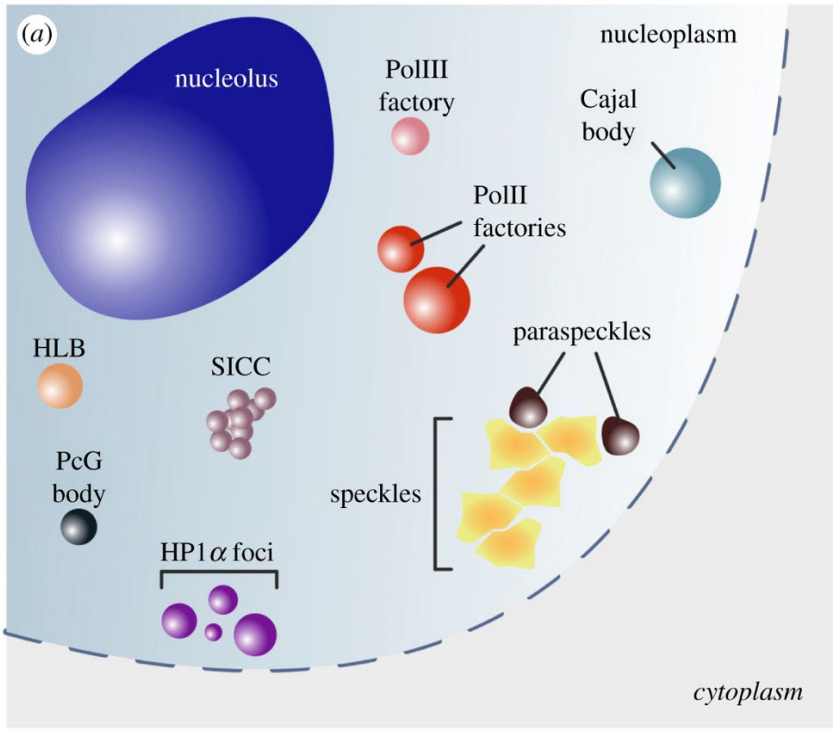


Figure 1: Phase-separated compartments in the nucleus.
Taken from Palikyras & Papantonis, 2019. HLB: Histone locus body, PcG body: Polycomb-group body, SICC: senescence-induced CTCF cluster.

The assemblies can consist of proteins, nucleic acids, and other molecules and are found both in the cytoplasm and the nucleus (Boeynaems *et al*, 2018; Strom & Brangwynne, 2019). Many proteins that were found to undergo LLPS share an important feature: the so-called intrinsically disordered regions (IDRs). IDRs enable multiple and multivalent interactions instead of folding into a clearly defined structure binding specifically to another structure (Jonas & Izaurralde, 2013; Malinovska *et al*, 2013). Many RNA-binding proteins (RBPs) were shown to phase-separate together with their bound RNAs. The phase-separation behavior can be initiated by RNA (Garcia-Jove Navarro *et al*, 2019) and regulated by the secondary structure of RNAs and the ratio of RNA to RBPs (Banerjee *et al*, 2017; Langdon *et al*, 2018; Maharana *et al*, 2018). A rather detailed understanding of the formation of stress granules exists, a type of RNA-protein condensates (Ribonucleoprotein (RNP) granule). Stress granule formation depends on the concentration of RNA and its interaction with one of the IDRs of RNA binding protein G3BP1 (Guillén-Boixet *et al*, 2020; Sanders *et al*, 2020; Yang *et al*, 2020). Given the fact that the nucleus and its sub-compartments are enriched in IDR-containing proteins (IDPs) (Frege & Uversky, 2015) and the obvious abundance of electrically charged nucleic acids (both DNA and RNA) the nucleus is primed for LLPS (Aumiller & Keating, 2016).
Obviously, many questions are not answered yet in this young field of research.

1.1.2 Phase-separation in transcriptional control

Recently, the influence of phase-separation on transcription and ultimately protein expression has gained attention (Hnisz *et al*, 2017; Palikyras & Papantonis, 2019). Biophysical models can predict how chromatin subcompartments can be formed by phase-separation in a self-organizing manner (Erdel & Rippe, 2018).

The aforementioned LLPS of the nucleolus as an example of a nuclear subcompartment was postulated to be driven by the enrichment of protein-mediated, dynamic chromosomal crosslinks (Hult *et al*, 2017). This suggests that the nucleolus, besides the well-known function in ribosome formation, could also directly influence positioning and structure of chromosomes (Palikyras & Papantonis, 2019). Moreover, the nucleolus was shown to have a role in protein quality control, another important step in the control of protein expression (Frottin *et al*, 2019).

There are several other nuclear subcompartments with association to chromatin that were shown to undergo LLPS, such as Cajal bodies (CBs), histone locus bodies (HLBs) and promyelocytic leukaemia (PML) bodies (Zhu & Brangwynne, 2015). Paraspeckles, containing the known transcription modulating long noncoding RNAs (lncRNA) as a scaffold for further RNA-protein and protein-protein interactions, were suggested to be phase-separated, too (Fox *et al*, 2018; Galganski *et al*, 2017). Nuclear speckles, discovered as early as 1910 (Ramón y Cajal, 1910) and later primarily associated with splicing factors, contain transcription and chromatin-remodeling factors as well (Avşar Ilık and Aktaş, 2021; Galganski et al., 2017, see 1.1.4 for a more detailed description).

One focus of research has been on the formation of heterochromatin by phase-separation (Larson & Narlikar, 2018). A central player in this process is heterochromatin protein 1 (HP1) (Larson *et al*, 2017; Strom *et al*, 2017). Human HP1α demixes into droplets when saturated DNA is added or when its N-terminal extension is phosphorylated. Furthermore, HP1α then leads to compaction of DNA and could serve as a hub recruiting further heterochromatin proteins via specific interactions, which could be shown *in vitro* (Larson *et al*, 2017). Analysis of the known heterochromatic proteins shows that indeed these are significantly more disordered than random or even nuclear proteins in general (Guthmann *et al*, 2019).

In addition, specific histone modifications were shown to depend on phase-separation as an underlying mechanism. H2B ubiquitination along gene bodies is augmented by what the authors call a phase-separated, chromatin-associated ‘reaction chamber’ consisting of layered condensates of histone-modifying enzymes (Gallego *et al*, 2020).

Transcription is very much controlled by the assembly of the transcription machinery at specific loci. With their sequence-specific DNA-binding domains and activation domains, transcription factors are key players in this process (Green, 2005). Very often, their activation domains also contain IDRs (Tantos *et al*, 2012). Together with the large Mediator complex and RNA

polymerase II, transcription factors can form phase-separated condensates to promote transcription (Boija *et al*, 2018; Cho *et al*, 2018; Chong *et al*, 2018b). Specifically, transcriptional co-activators were shown to form phase-separated condensates at super-enhancers, concentrating the transcription apparatus at genes essential for cell-identity (Sabari *et al*, 2018). This was also shown for transcription factors containing prion-like domains in Candida albicans (Frazer *et al*, 2020) or for the yeast transcriptional regulator TBP associated factor 14 (Taf14) which serves as a phase-separated hub for multiple transcriptional regulators (Chen *et al*, 2020a). Also transcription factors specific for certain pathways like TAZ employ phase-separation for specific activation of transcription (Lu *et al*, 2020). However, other authors have shown that transcriptional activation can also be enhanced by multivalent interactions alone, while the formation of phase-separated droplets had no influence on the activation (Trojanowski *et al*, 2022).

An example where the influence and control of phase-separation on function and transcription is very well understood is RNA polymerase II (Pol II). Pol II, responsible for transcription of mRNA, needs to be hyperphosphorylated at its C-terminal domain (CTD) in order to be able to elicit its elongation purpose (Kwon *et al*, 2013). This hyperphosphorylation depends on several factors that create a phase-separated functional compartment (Lu *et al*, 2018). The CTD as an IDR enabling phase-separation also appears to be responsible for the clustering of multiple Pol II units together with activators for initiation of transcription before the release for elongation by CTD phosphorylation (Boehning *et al*, 2018). Finally, the number of phosphorylations even controls in which phase-separated condensates Pol II acts (Guo *et al*, 2019). A hypophosphorylated CTD leads to incorporation to condensates with Mediator for transcription initiation, whereas a hyperphosphorylation of the CTD by regulatory cyclin-dependent kinases (CDKs) leads to incorporation to condensates with splicing factors for RNA processing. Furthermore, the FET-family transcriptional regulators tend to undergo phase-separation. One of them, TAF15, also interacts with Pol II's CTD leading a positive feed-back loop to create a localized phase-separated compartment of transcriptional components (Wei *et al*, 2020). Overall, the phase-separation appears to be important to accelerate the finding of and binding to the target site of transcriptional regulators, exemplified by CBX2 (Kent *et al*, 2020).

More recently, it was shown that the IDR of MED1 can bring Pol II into specific condensates with positive, but not negative allosteric regulators (Lyons *et al*, 2023). This compartmentalization is dependent on a pattern of alternating blocks of charged amino acids within the sequence of transcriptional regulators and sufficient to activate transcription, even leading to a cell-state transition.

1.1.3 Post-transcriptional gene regulation via phase-separation

The influence of phase-separation on gene expression does not stop at the level of transcription, but membraneless compartmentalization continues to be relevant also for the following steps of gene regulation at the level of RNA and translation of RNA to protein. An important step towards mature mRNA is splicing which can happen co-transcriptionally already (Hsin & Manley, 2012; Braunschweig *et al*, 2013). Several publications could show that splicing occurs in phase-separated compartments. As mentioned before, the switch from initiation- to co-transcriptional splicing-related condensates was shown to depend on phosphorylation of Pol II's CTD (Guo *et al*, 2019). Not only co-transcriptional splicing depends on phase-separation. Splicing factor TDP-43, for which mutations are involved in amyotrophic lateral sclerosis (ALS) (see also 1.1.5 phase-separation in disease), was shown to phase-separate, and variants in a partly helical region in its C-terminal IDR can influence its material properties and splicing function. Heterogeneous nuclear ribonucleoproteins (hnRNPs) contain IDRs and form a large family of RNA-binding proteins. Alternative splicing events were shown to be dependent on phase-separated multivalent assemblies containing hnRNPs (Gueroussov *et al*, 2017). Two low-complexity (LC) domains (similar to IDRs) of hnRNPH1 serve separate purposes (Kim & Kwon, 2021), whereby phase-separation of LC1 appears to be important for the alternative-splicing function and binding to other RBPs, while LC2 leads to transcriptional activation together with the DNA-binding domain. A proposed model (Liao & Regev, 2020) for the role of the interface of the phase-separated compartment of nuclear speckles (see also 1.1.4) with the surroundings in splicing could be validated in important parts experimentally later (Paul *et al*, 2022). The deubiquitylase USP42 also localizes to nuclear speckles and its downregulation leads to reduced nuclear speckle foci, and deregulation of multiple splicing events (Liu *et al*, 2021).

Phase-separation stays relevant also for the next step of mRNA maturation, 3' end processing. In Arabidopsis, coiled-coil protein FLL2 is responsible for the phase-separation of a number of 3' end processing factors including FCA in nuclear bodies, leading to correct formation of the 3' end of specific mRNAs (Fang *et al*, 2019).

mRNA modification and its role in phase-separation gained more and more attention. The best studied and most abundant modification in mammals is N^6-methyladenosine (m^6A) (Dominissini *et al*, 2012) which is involved in export, splicing, degradation and translation of mRNA and processing of miRNA (Yu *et al*, 2018; Meyer *et al*, 2012). More recently, the modification's relevance for phase-separation of proteins from the YTHDF family of reader proteins with the m^6A recognition (YTH) domain was shown (Fu & Zhuang, 2020; Gao *et al*, 2019; Ries *et al*, 2019; Wang *et al*, 2020a), nicely reviewed in Su *et al*, 2021. This is also true

for the modification of enhancer RNA, which leads to transcriptional condensate formation and consequent gene activation (Lee *et al*, 2021).
Last but not least, translational control can be modulated by phase-separation. Two examples are mRNAs in stress granules which can still be translated (Mateju *et al*, 2020), and the translational heat shock response in yeast which depends on the phase-separation of the essential translation initiation factor Ded1p (Iserman *et al*, 2020).
Thus, multiple layers of gene regulation are controlled and affected by phase-separation, but a plethora of phenomena have not been understood yet. One particular phase-separated compartment - nuclear speckles - that appears to be involved especially in splicing, but potentially also in transcriptional gene regulation and emerged as relevant for the study of SMNDC1 shall be discussed in further detail in the next section.

1.1.4 Nuclear speckles: an example of a phase-separated compartment within the nucleus involved in gene regulation

The substructures in the nucleus for which I will use the term “nuclear speckles”, also known as interchromatin granule clusters (IGCs), were observed already in 1910 by Ramón y Cajal (Ramón y Cajal, 1910), who is also the eponym of Cajal bodies (Ramón y Cajal, 1903). The first electron microscopy experiments which also identified RNA as a component were published in 1957 (Swift, 1959), and shortly after, the term “speckles” was used for the first time in the description of these subnuclear structures (Beck, 1961). Two recent reviews describe the current knowledge about structure and function (Avşar Ilık & Aktaş, 2021; Galganski *et al*, 2017). A typical eukaryotic interphase cell nucleus contains 20-50 nuclear speckles with irregular shape and varying size (Spector & Lamond, 2011). Nuclear speckles are dynamic structures that can vary in size and shape depending on various factors such as cellular ATP levels or inhibition of RNA polymerase II transcription and splicing (Zhang *et al*, 2016; Sinclair & Brasch, 1978). Typical for a phase-separated compartment, proteins containing IDRs are enriched amongst known nuclear speckle proteins (Galganski *et al*, 2017), and the specific function of IDRs for the localization to nuclear speckles could be shown for example for SPOP (Marzahn *et al*, 2016).
An overview of the different proteins localizing to nuclear speckles which are involved in nuclear gene expression regulation, namely initiation of transcription (e.g. RNA Pol II subunits) (Saitoh *et al*, 2004), splicing, m6A-modification (e.g. METTL3/14, YTHDC1) (Xiao *et al*, 2016; Ping *et al*, 2014; Bokar *et al*, 1997), 3’ end modification (e.g. CPSF proteins) (Saitoh *et al*, 2004), or preparation of mRNA export from the nucleus (e.g. ALYREF, PABPC1) (Teng & Wilson, 2013; Copeland *et al*, 2013) can be found in Figure 2.

The splicing function of nuclear speckles seems to be the evolutionarily oldest (Avşar Ilık & Aktaş, 2021) and is supported by many studies locating splicing factors to nuclear speckles (Wahl *et al*, 2009). While introns increase in length during evolution it becomes more and more complicated for consecutive exons to come together for the process of splicing in the crowded nuclear environment. Nuclear speckles as a structure integrating SR proteins, a family of splicing proteins containing repeats of serine (S) and arginine (R), could facilitate this exon bridging process (Liao & Regev, 2020; Paul *et al*, 2022).

Recent studies conducted by the Belmont group have provided evidence that nuclear speckles play a direct role in regulating transcription, besides splicing. This was demonstrated by the general correlation between spatial proximity to nuclear speckles and gene expression amplification (Kim *et al*, 2019), as first observed through live-cell imaging of heat-shock responsive genes (Khanna *et al*, 2014). Furthermore, the group compared the distance between nuclear speckles and endogenously expressed genes in four different cell lines using tyramide signal amplification-sequencing (TSA-Seq) (Chen *et al*, 2018), which further supports the finding that genes positioned closer to nuclear speckles are more highly transcribed (Zhang *et al*, 2020).

In addition, spatial organization of chromatin into active and inactive domains is another discussed function of nuclear speckles (Brown *et al*, 2008; Quinodoz *et al*, 2018; Hu *et al*, 2019). Nuclear speckles contain several components that regulated gene expression via epigenetic mechanisms. These include histone variants (Soboleva *et al*, 2017), histone acetyltransferases (Chakraborty *et al*, 2001; Papoutsopoulou & Janknecht, 2000), histone methyltransferases (Deng *et al*, 2004; Yano *et al*, 1997), histone deacetylases (Singh *et al*, 2010; Baertschi *et al*, 2014), and HP1 protein (Saitoh *et al*, 2004). Despite the fact that these proteins can localize to nuclear speckles, the potential function of nuclear speckles in their regulation has not been directly studied. As a result, there is a possibility for discovering novel mechanisms of gene expression regulation in the context of epigenetic factors and nuclear speckles.

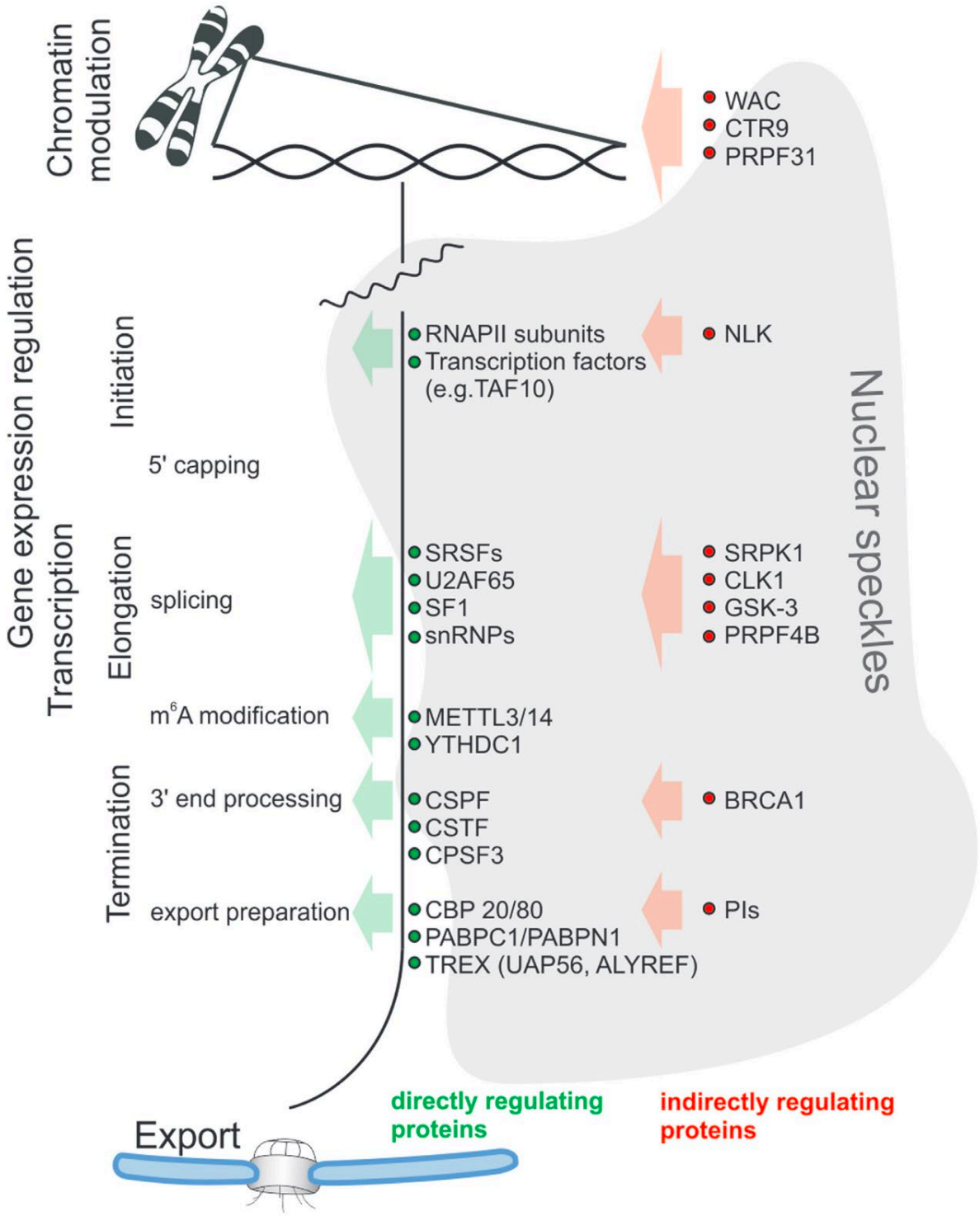


Figure 2: Gene expression regulation happening in nuclear speckles.
Taken from Galganski *et al*, 2017.

Last but not least, nuclear speckles are also involved in a number of diseases, consequently termed nuclear speckleopathies (Regan-Fendt & Izumi, 2023). While a large number of proteins localizing to nuclear speckles are involved in some form of disease, a few diseases show a more direct connection to nuclear speckles and their structure. Mutations in speckle

proteins show a tendency towards causing developmental disorders, especially of the nervous system. Examples for two central organizers of nuclear speckles (Ilik *et al*, 2020) are ZTTK (Zhu-Tokita-Takenouchi-Kim) syndrome due to mutations in SON (Kim *et al*, 2016; Zhu *et al*, 2015; Takenouchi *et al*, 2016; Tokita *et al*, 2016) or a neurodevelopmental disorder caused by mutations in SRRM2 (Cuinat *et al*, 2022).
Consequently, a lot of work is underway to better understand the architecture of nuclear speckles and their involvement in human disorders.

1.1.5 Phase-separation in disease

Nuclear speckles are not the only phase-separated compartment involved in disease, and the process of phase-separation itself has been attributed to disease mechanisms. By looking for functional IDR segments in protein sequences, more than a thousand such segments could be found in disease-related proteins, showing the potential relevance for the function of these proteins and consequently the whole spectrum of disease (Anbo *et al*, 2019). Prior to looking into examples for phase-separation as a cause for specific diseases, I will discuss the hypothetical scenarios how aberrant phase-separation may lead to disease (Figure 3). Broadly speaking, there are three underlying mechanisms through which genetic mutations or environmental stressors could impede condensate formation (Alberti & Dormann, 2019). First, by directly modifying the molecular process of condensate assembly, for example through mutations that change the valency of a client or scaffolding protein (Li *et al*, 2012) or its solubility. Second, by modifying the functionality of a central regulator of condensation, for example a kinase or the ectopic expression of a nucleator. Third, by inducing alterations in the overall physicochemical state of a cell like pH or salt concentrations. ATP levels have been shown to change the solubility of proteins on a global scale, too (Guilhas *et al*, 2020; Patel *et al*, 2017; Hayes *et al*, 2018). In the following paragraphs, I will focus on neurodegenerative disease, cancer, and endocrinological/ metabolic disease.

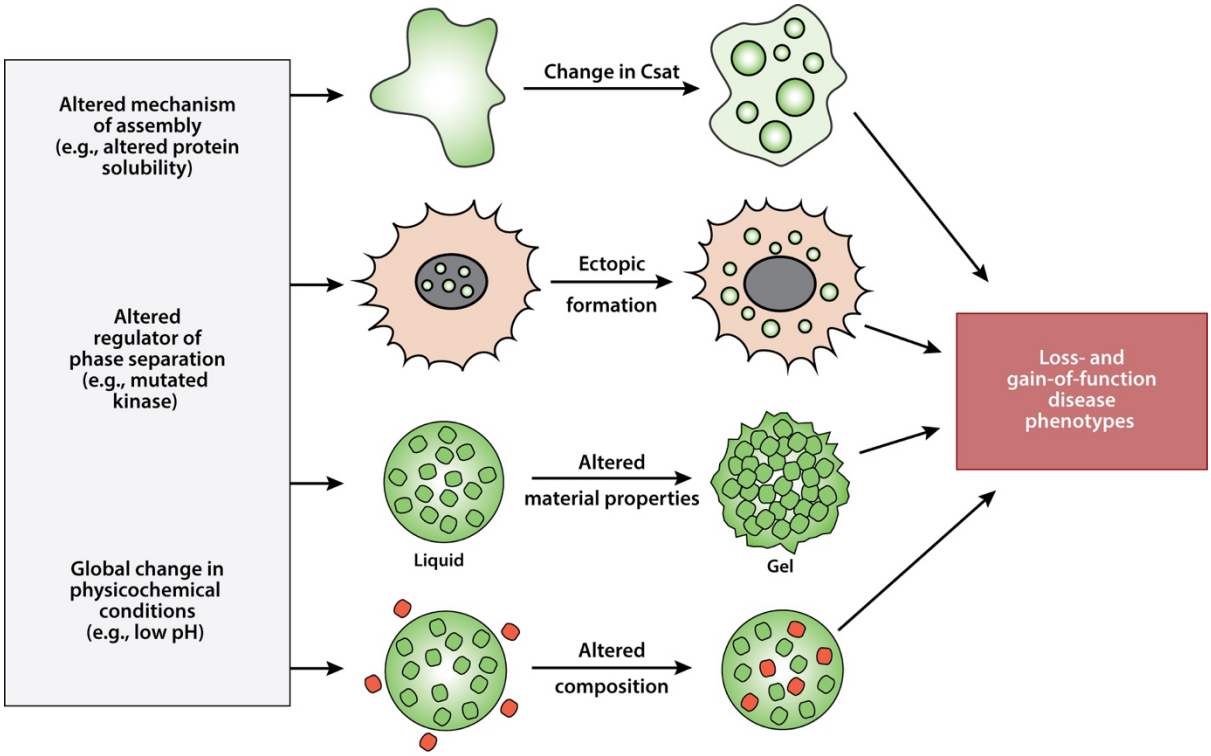


Figure 3: Illustration of the potential scenarios in which disease may emerge as a result of anomalous phase-separation.
Such ailments may stem from changes in the assembly process, regulators of phase-separation, or alterations to the physicochemical state of cells. Csat: saturation concentration for phase-separation. Taken from Alberti & Dormann, 2019.

Neurodegenerative disease is probably the form of disease best studied in the context of phase-separation (Zbinden *et al*, 2020), and neurons appear to be especially susceptible. Neurons have unique characteristics that make them highly dependent on efficient protein quality control systems, including those controlling phase-separation and RNP granule dynamics. Neurons do not undergo cell division which can help in the clearance of protein aggregates (Lim & Yue, 2015) and have a high metabolic activity, requiring large amounts of ATP (Raichle & Gusnard, 2002). They are highly polarized cells that need to transport macromolecular complexes, such as RNP granules, along microtubules into axons and dendrites. Neurons constantly undergo molecular, morphological, and functional changes through post-transcriptional gene regulation requiring elevated levels of RNA editing, localized translation and alternative splicing (Kiebler *et al*, 2013; Holt & Schuman, 2013).
Protein aggregation is a critical feature of all neurodegenerative disorders, and it is thought to be the cause of neurodegeneration (Taylor *et al*, 2002). There is now a large body of evidence linking pathological protein aggregates and liquid-to-solid phase transitions or irregular phase-separation (Zbinden *et al*, 2020). Best studied examples are FUS and TDP-43 for amyotrophic

lateral sclerosis and frontotemporal dementia (Mackenzie *et al*, 2010; Hofweber *et al*, 2018; Afroz *et al*, 2019; Qamar *et al*, 2018), Tau for Alzheimer's disease (Ainani *et al*, 2023; Boyko *et al*, 2019; Wegmann *et al*, 2018), or α-synuclein for the so-called Lewy body diseases which include Parkinson's disease (Ray *et al*, 2020).

In the next paragraph, I will focus on the implications of phase separation in the development of cancer. For cells to become cancerous they must massively increase their proliferation. Many of the famous "hallmarks of cancer" (Hanahan & Weinberg, 2000, 2011; Hanahan, 2022) can be linked to phase-separation but some have a more direct connection, illustrated by examples with direct evidence. Proliferative signaling is clearly one of these, since the clustering of membrane receptors and their signaling molecules appears to be explained by phase-separation, as shown for the T-cell receptor (Su *et al*, 2016; Ditlev *et al*, 2019). Another significant characteristic of cancer is the deregulation of transcription and as discussed before (1.1.2), transcriptional regulation is one of the best-studied cases of the role of phase-separation. Fusion-proteins, resulting from chromosomal rearrangements and more frequent due to genetic instability in cancer, can combine the IDR of one protein with a functional domain of another protein for a gain of LLPS ability. Examples include a fusion between the IDR of EWS and the FLI protein in Ewing's sarcoma (Boulay *et al*, 2017) and NUP98-HOXA9 in leukemia (Ahn *et al*, 2021). Another study could show that cancer-associated mutations in the nuclear speckle protein SPOP lead to a disruption in substrate interactions which are required for SPOP's phase-separation. Consequently, the proto-oncogene DAXX accumulates due to insufficient ubiquitination and proteasomal degradation (Bouchard *et al*, 2018).

Eventually, phase-separation is also implicated in endocrinological and metabolic diseases (Akiba *et al*, 2021; Chen *et al*, 2022). A growing body of evidence suggests that the deposition of misfolded amyloid proteins is a characteristic feature of Type 2 Diabetes (T2D). This deposition leads to the failure of pancreatic β-cells and eventually to the loss of insulin expression. Research shows that amyloid peptides and proteins undergo LLPS before the formation of amyloid fibrils (Xing *et al*, 2021). In β-cells in particular, the major amyloid component is Islet Amyloid Polypeptide (IAPP), which can phase-separate by hydrogelation followed by deleterious aggregation and formation of islet amyloid deposits, resulting in pancreatic β-cell dysfunction (Pytowski *et al*, 2020). Inhibiting amyloid aggregation and LLPS has shown initial promise for treating T2D, and IAPP is a potential therapeutic target (Sevcuka *et al*, 2022). Additional research shows that stress granule formation contributes to β-cell dysfunction and T2DM by entrapping the transcription factor PDX1, which is important for cell identity and insulin expression, through saturated fatty acids (Zhang *et al*, 2021). Furthermore, PDX1 is down-regulated through interaction with SPOP (see above for its involvement in cancer and nuclear speckles, 1.1.4) (Usher *et al*, 2021). Even though Pdx1 does not undergo

phase-separation together with SPOP, their interaction triggers SPOP to relocate from nuclear speckles to the nucleoplasm where ubiquitination of PDX1 takes place, leading to proteasomal degradation.

Down-stream of insulin expression also insulin processing and storage in secretory granules and eventually insulin signaling appear to be affected by phase-separation. Chromogranin proteins phase-separate at slightly acidic pH within the trans-Golgi network and can recruit proinsulin to these condensates (Parchure *et al*, 2022). While exocytosis and formation of secretory complexes in neurons could be shown to depend on phase-separation, reviewed in (Chen *et al*, 2020b; Hayashi *et al*, 2021), the same could not be observed for the exocytosis of insulin in β-cells (Noordstra *et al*, 2022). Formation of the secretory granules however does depend on the condensation of ICA512's RESP18 homology domain together with proinsulin and insulin (Toledo *et al*, 2019, 2023). Phase separation processes have also been associated with insulin signaling. The C-terminus of insulin receptor substrate 1 (IRS-1) undergoes LLPS, which drives the formation of insulin/IGF-1 signalosomes through self-association (Gao *et al*, 2022). Impairment of IRS-1 LLPS, e.g., through the metabolic disease-associated G972R mutation, attenuates its positive effects on insulin/IGF-1 signaling. The components of the insulin signaling cascade PI(4,5)P2, PI3K, and PDK1 are always present in the condensates of IRS-1, while insulin triggers the recruitment of PKB and the production of PIP3 (Zhou *et al*, 2022). The condensates of IRS1 serve as crucial intracellular signaling centers for insulin signaling and their formation is hampered in insulin-resistant cells.

These examples elucidate the importance of phase-separation for a multitude of underlying processes in disease and open avenues for new forms of treatments which will be discussed in the next chapter.

1.1.6 Chemical perturbation of phase-separated compartments

Due to the involvement of phase-separation in disease there are first attempts to influence these fundamental mechanisms with small molecules. One of the first tools used to disrupt membraneless compartments was 1,6-hexanediol (Kroschwald *et al*, 2017). It can dissolve liquid-like assemblies but fails to do so for solid-like protein aggregates. While it can be a useful tool to study phase-separation *in vitro* and *in cellulo,* it is also quite toxic and was shown to unspecifically inhibit kinases and phosphatases with unforeseen downstream effects (Düster *et al*, 2021). Before looking into examples of more specific chemical matter which is supposed to modulate phase-separated compartments I will discuss how small molecules might localize specifically to phase-separated compartments (Kilgore & Young, 2022). A seminal study was the paper by (Klein *et al*, 2020) showing that well-established antineoplastic drugs do not distribute uniformly, but are concentrated in specific protein condensates both *in*

vitro and in their target tumor cells. This behavior was associated with drug activity and could potentially be employed in designing novel drugs with improved activity but also reduced toxicity and side-effects. Understanding this "chemical grammar" (Kilgore & Young, 2022) which targets small molecules to condensates would advance the field tremendously. It is probable that the chemical characteristics that allow biomolecules to be selectively distributed in condensates are also responsible for the selective distribution of small molecules in particular condensates. Besides the properties that allow small molecules to concentrate in specific condensates it should also be possible to understand the chemical characteristics that let small molecules modulate these condensates in terms of their phase behavior and their material properties. Of course, these are not incompatible, and a small molecule might both concentrate in and modify a specific condensate.

An excellent overview of the field of potential therapeutic modulation of condensates can be found in two recent reviews (Mitrea *et al*, 2022; Conti & Oppikofer, 2022). The theoretical approaches to address condensate-related diseases include repairing a faulty condensate, interrupting the regular operation of a condensate linked to the illness, or obstructing the function of a target molecule by either disabling it within its original condensate or removing it from that condensate (Figure 4a). There are numerous methods that can be employed to pinpoint condensate-modifying therapeutics that accomplish these three aims (Figure 4b). The stability of biomolecular condensates can be affected by modulating their scaffold. Compounds can achieve this by changing the interaction valency or strength, blocking or stabilizing protein-protein or nucleic acid interactions without inhibiting a particular protein completely, which can be achieved by moderate changes in the weak networking interactions. To modulate composition of condensates drugs can be developed to either inhibit or promote client-scaffold interactions. This approach can be used to help an aberrantly de-partitioned protein to return to its native condensate, as seen for the nucleolar protein NPM1 in acute myeloid leukaemia (Mukherjee *et al*, 2015). The conformational and interaction landscape can be modulated by engaging with IDRs. Traditional structure-based methods cannot be used to screen for condensate-modifying therapeutics that interact with IDRs since they are conformationally dynamic. Developing drugs that target proteins in families such as transcription factors, hormone receptors, and nucleotide-binding proteins containing IDRs have been challenging but might be possible by producing high local drug concentration within a condensate, overcoming the low affinity. Another approach to remove a protein from a condensate is to degrade it using proteolysis-targeting chimera (PROTAC) or molecular glue strategies (Dale *et al*, 2021). By degrading a scaffold, one can reduce the effective concentration below the saturation concentration for phase-separation and prevent or reverse assembly of the condensate. Yet another way to regulate condensate assembly is through up-stream processes catalyzed by enzymes like chaperones and helicases, which can affect the

condensate environment and the behavior of specific proteins or nucleic acids (Snead & Gladfelter, 2019). Modifying turnover kinetics, changing post-translational states of proteins, or epigenetic and epitranscriptomic modifications of DNA or RNA can also regulate condensate composition. Phosphorylation and methylation are examples of such modifications that can affect protein and RNA condensation (Hofweber & Dormann, 2019).

Until now, all of these approaches are in the very early stages of development or even completely hypothetical. Future research will show whether these novel modes of action can lead to clinically used therapeutics.

Two interesting proteins that could be targeted by condensate-modifying molecules will be discussed in the next section.

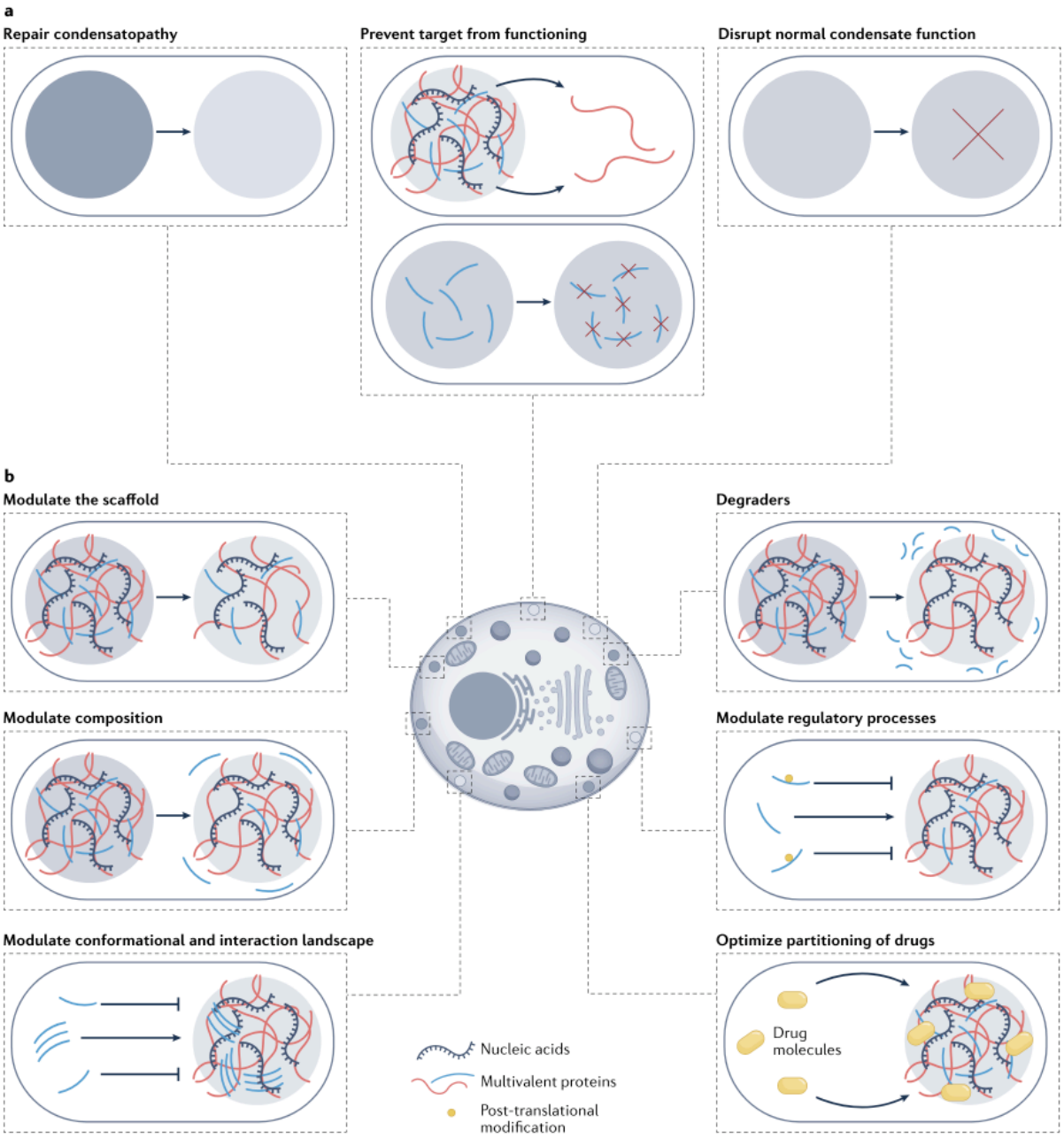


Figure 4: Aims (a) and approaches (b) of condensate-modifying therapeutics. Taken from Mitrea *et al*, 2022.

1.1.7 SMNDC1 and SMN – new phase-separating Tudor domain proteins as drug targets?

A growing body of evidence links arginine methylation, especially in RGG/RG motifs, and phase-separation, for example of the RNA-binding protein FUS (Fused in sarcoma) (Hofweber *et al*, 2018; Qamar *et al*, 2018), recently reviewed by Chong *et al*, 2018. *In vitro* and *in vivo* results indicate that arginine methylation inhibits phase-separation or weakens intermolecular

interactions (Liu *et al*, 2011; Nott *et al*, 2015; Hofweber *et al*, 2018; Ryan *et al*, 2018), but can promote it in other cases (Arribas-Layton *et al*, 2016).

There is also a number of studies linking the domains that bind methylated arginines, Tudor domains, and Tudor domain containing proteins to phase-separation, reviewed by (Šimčíková *et al*, 2023). These include Tudor-SN (Su *et al*, 2017; Gao *et al*, 2010) and the multi-Tudor domain-containing protein (Tdrd6) which has been shown to fine-tune aggregating properties of the phase-separated Germ plasm in Zebrafish via interaction with Buc (Roovers *et al*, 2018).

The Tudor domain protein Survival Motor Neuron (SMN) is well studied due to its involvement in the genetic neuromuscular disorder spinal muscular atrophy (SMA). Its diverse functions in cellular homeostasis were nicely summarized by (Chaytow *et al*, 2018). An important function is the assembly of RNPs, especially the spliceosome (Liu *et al*, 1997; Liu & Dreyfuss, 1996; Fischer *et al*, 1997). Furthermore, SMN is involved in RNA and protein trafficking, especially in axons of mature neurons (Giavazzi *et al*, 2006); connected to this is its function in translational control through directing subcellular localization of mRNAs (Fallini *et al*, 2016), but also by binding to polyribosomes (Sanchez *et al*, 2013) and through the mTOR pathway (Kye *et al*, 2014). Also connected to its function in trafficking is SMN's role in the cytoskeleton, exemplified by modulating β-actin (Rossoll *et al*, 2003). In addition, dysfunctional SMN appears to impair endocytosis (Dimitriadi *et al*, 2016) and mitochondrial metabolism (Ripolone *et al*, 2015; Acsadi *et al*, 2009). More recently, a study claimed that SMN binds H3K79me1 hinting a role in epigenetic regulation (Binda *et al*, 2023a) and a more direct evidence of transcriptional control was provided by a study showing that SMN regulates termination through the interaction with symmetrically dimethylated Arginine R1810 of Pol II with its Tudor domain (Zhao *et al*, 2016).

Finally, SMN's Tudor domain was also shown to be required for the regulation of stress granules (Chitiprolu *et al*, 2018), a phase-separated compartment, and SMN was shown to form droplets *in vitro* (Binda *et al*, 2023b). Another study could show that SMN's Tudor domain promotes condensation through interaction with dimethylarginine and that the specific levels of aDMA and sDMA fine-tune these interactions, also for other Tudor domains (Courchaine *et al*, 2021).

SMNDC1 (Survival Motor Neuron Domain Containing 1), which is also known as SPF30 (splicing factor 30) or SMNrp (SMN related protein), was identified first by proteomic analysis of purified spliceosomes (Neubauer *et al*, 1998) and by looking for homologous genes of Survival Motor Neuron (SMN) (Talbot *et al*, 1998). Its involvement in assembling the mature spliceosome complex was discovered shortly afterwards (Rappsilber *et al*, 2001; Meister *et al*, 2001). As its name and its discovery suggests, SMNDC1 shares a homologous part with SMN. Both SMN and SMNDC1 contain a conserved Tudor domain that recognizes symmetrically dimethylated Arginines (sDMA) (Tripsianes et al., 2011) (Figure 5A). The Tudor

domain structures binding sDMA and aDMA were solved and show that an aromatic cage mediates the recognition (Figure 5B). In a study with the yeast ortholog it was shown that Spf30 is a splicing factor which also has a role in the exosome pathway of heterochromatin silencing (Bernard *et al*, 2010). SMNDC1 is conserved in in vertebrates and plants, and a phylogenetic analysis in plants found 82 SMNDC1 genes in 64 plant species, but did not look into their function in more detail (Zhang *et al*, 2019). A more recent interactome analysis showed that SMNDC1 associates with MTR4 and the nuclear exosome RNA-decay machinery and other proteins involved in multiple RNA metabolic pathways (Ishida *et al*, 2021). Another study focusing on essential splicing factors in hepatocellular carcinoma could show that patients with high SMNDC1 expression have worse survival rates (Zhu *et al*, 2023). In addition, a knock-down of SMNDC1 suppressed migration and proliferation in hepatocellular carcinoma cells. Overall, there are not many studies focusing on SMNDC1 and its function.

To identify silencers of insulin in α-cells, we set up an RNAi screen with a focus on proteins containing a chromatin-binding domain (Casteels *et al*, 2022). Successful knock-down of SMNDC1 induced insulin expression by at least 3-fold (Figure 5C) which sparked our interest in this protein. We combined a number of methods like RNA sequencing, ATAC-Seq, proteomics and SMNDC1 interaction profiling to understand the mechanism of induced insulin expression upon SMNDC1 knockdown (Figure 5D). As we could show, SMNDC1 regulates the levels of the chromatin remodeler ATRX via direct binding and splicing, leading to a loss of ATRX mRNA and protein when SMNDC1 is knocked down. In turn, this influences accessibility and expression of pluripotency and β-cell genes like the transcription factor PDX1 which is responsible for the upregulation of insulin.

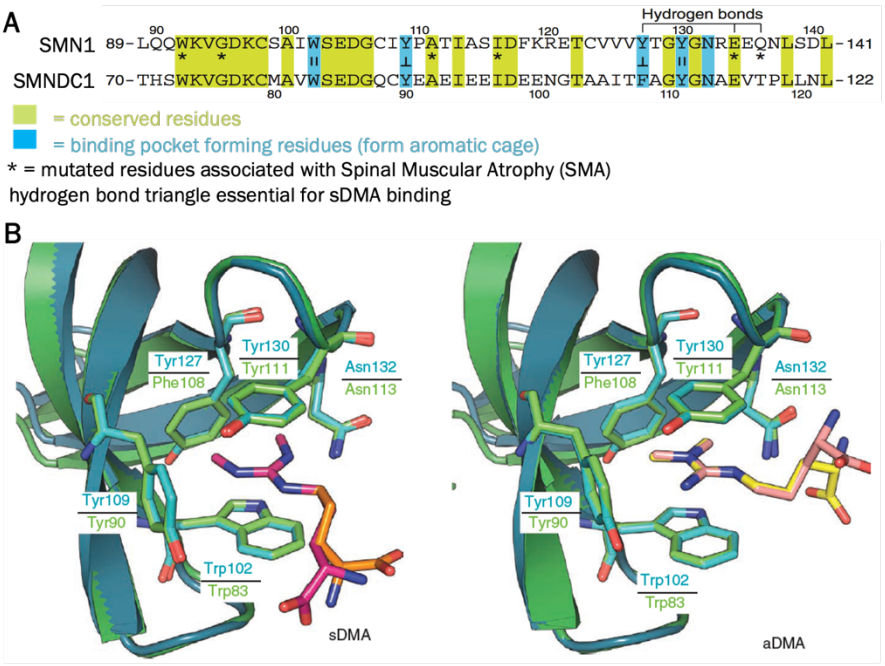

A
Hydrogen bonds
SMN1 89-LQQWKVGDKCSAIWSEDGCIYPATIASIDFKRETCVVVYTGYGNREEQNLSDL-141
SMNDC1 70-THSWKVGDKCMAVWSEDGQCYEAEIEEIDEENGTAAITFAGYGNAEVTPLLNL-122
= conserved residues
= binding pocket forming residues (form aromatic cage)
* = mutated residues associated with Spinal Muscular Atrophy (SMA)
hydrogen bond triangle essential for sDMA binding
B
Tyr127
Phe108
Tyr130
Tyr111
Asn132
Asn113
Tyr109
Tyr90
Trp102
Trp83
sDMA
aDMA
C
Smndc1
Insulin (Ins2)
Percent of Control
p<0.0001
NTC 94 95 96 97 98
Short Hairpins
D
Alpha Cell
SMNDC1-Deficient Alpha Cell
SMNDC1
Atrx mRNA
STOP
ATRX
H3K9me3
Pluripotency Genes
Beta Cell Genes
BAF Complex
Pdx1
Pdx1 mRNA
Insulin

Figure 5: Structure of SMN1's and SMNDC1's Tudor domains and SMNDC1's role in the regulation of insulin expression.
A: Amino acid sequence of the Tudor domains of SMN1 and SMNDC1. Light green are conserved residues, blue are the residues that form the aromatic cage. Marked with an asterisk are residues that when mutated are associated with SMA. Adapted from Tripsianes et al., 2011. **B:** 3D model of the overlaid Tudor domains of SMN1 (blue) and SMNDC1 (green) binding sDMA (left) and aDMA (right) with the residues forming the aromatic cage. Adapted from Tripsianes et al., 2011. **C:** qPCR results for Smndc1 and Insulin after knock-down with numbered hairpins (94-98) or non-targeting control (NTC). Taken from Casteels *et al*, 2022. **D:** Graphical summary of the findings by Casteels *et al*, 2022. Taken from Casteels *et al*, 2022.

From the examples of other Tudor domains and Tudor domain proteins, we hypothesize that SMNDC1's Tudor domain could play a key role in regulating phase-separation, too. Since it is also involved in regulating insulin expression, SMNDC1 is an excellent starting point to discover fundamental mechanisms of transcriptional control in α-cells and beyond. A way to target phase-separated compartments could be via classical small molecule inhibition of organizing proteins (see also section 1.1.6). SMN1 and SMNDC1 with their Tudor domain pocket binding methylated Arginines are an attractive target from this point of view. Previous examples of methyl reader domains targeted by small molecule inhibitors will be discussed in the following section.

1.2 Methyl readers and their inhibitors

1.2.1 Post-translational modification of proteins – methylation

Post-translational modification (PTM) of proteins is an essential regulator of their activity, stability, and interaction with other proteins. The addition of a functional group, typically to a nucleophilic part of the amino acid side chains, is one of these post-translational modifications. There are numerous functional groups that can be added to the amino acid side chains. We will focus on methyl-groups, which are usually introduced at lysine and arginine residues. Lysine's NH_3^+ group can be mono-, di-, or trimethylated whereupon the hydrogens are sequentially replaced by a methyl-group (Paik & Kim, 1975) (Figure 6a). Arginine residues can be more diversely methylated, since both a NH_2 and a NH_2^+ group are available. A monomethylation can occur only at the NH_2 group (ω-N^G-monomethylarginine (MMA)), whereas a subsequent dimethylation can occur "symmetrically" at the NH_2^+ group (ω-N^G,N'^G-symmetric dimethylarginine (sDMA)) or "asymmetrically" at the NH_2 group (ω-N^G,N^G-asymmetric dimethylarginine (aDMA)) (Blanc & Richard, 2017) (Figure 6b).

The modifying enzymes are specific to lysine or arginine, but all depend on *S*-Adenosylmethionine (SAM) as a methyl-group donor. Particularly in the context of histones, these methyltransferases are called “writers”, which complement the functions of “erasers” (demethylases), and “readers” which contain a domain recognizing and binding the modified amino acid. In the case of lysine-specific methyltransferases (KMTs), one distinguishes between SET (**S**u(var)3-9, **E**nhancer of Zeste, **T**rithorax) domain containing or non-SET domain containing enzymes. It is estimated that the human genome contains around 50 KMTs (Arrowsmith *et al*, 2012).

Protein arginine methyltransferases (PRMTs) are divided into 3 types. Type I (PRMT1, PRMT2, PRMT3, PRMT4, PRMT6, and PRMT8) carries out monomethylation and subsequently asymmetric dimethylation. Type II (PRMT5 and PRMT9) on the other hand can only produce symmetrical dimethylation after an intermediate monomethylation (Branscombe *et al*, 2001; Yang *et al*, 2015). PRMT7 is the only member of type III and is able to monomethylate arginines (Feng *et al*, 2013).

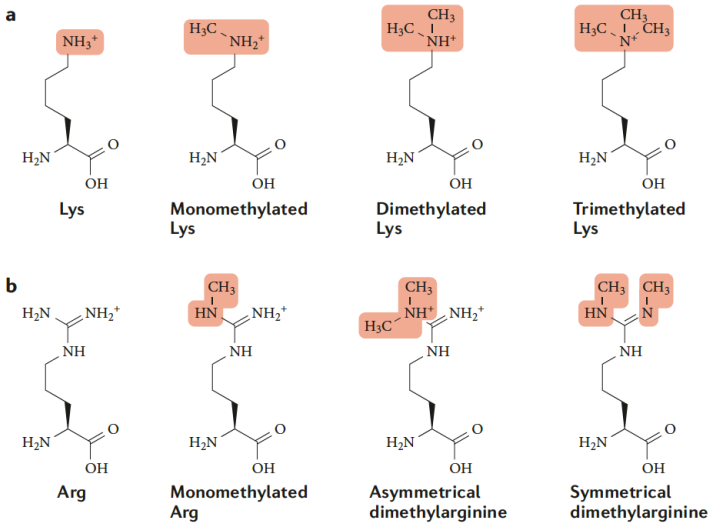


Figure 6: Methylations of Lysine and Arginine.
Taken from Biggar & Li, 2015. a: Lysine (Lys) can be mono-, di- or trimethylated at its NH3+ group. b: Arginine (Arg) is monomethylated at its NH2 group, and subsequently dimethylated either asymmetrical also at the NH2 group, or symmetrical at its NH2+ group.

It was a long time after the first discovery of protein methylation in a biological setting (Ambler & Rees, 1959), until the first biological role for protein methylation was identified (Shen *et al*, 1998). Until today, most that is known about the function of protein methylation is related to histone modifications and their influence on transcription which can be either activating or repressive. This depends on the site, but also the degree of methylation. For lysines, typical examples for a repressive histone mark are trimethylation of histone 3, lysine 9 (H3K9me3) or

H3K27, whereas methylations of H3K4 and H3K36 are considered activating (Barski *et al*, 2007). For arginines, activating marks are H4R3me2a, H3R2me2s, H3R17me2a, and H3R26me2a, while H3R2me2a, H3R8me2a, H3R8me2s, H4R3me2s are seen as repressive histone marks. Often these assignments are also context dependent, as there is widespread crosstalk between different modifications.

There is a growing body of evidence that not only histones, but a large proportion of the mammalian proteome is methylated (Larsen *et al*, 2016). Methylation of non-histone proteins draws increasing attention and a variety of functions are attributed to methylation of non-histone proteins, e.g. in signaling (Biggar & Li, 2015).

If protein methylation serves an important function, then it stands to reason that it must also be tightly regulated. Consequently, in addition to methyltransferases, there are also demethylases, or “erasers”. The human genome encodes around 30 lysine-specific histone demethylases (KDMs), of which several (KDM3A, KDM4E, KDM5C) were shown to possess arginine demethylase activity *in vitro* (Walport *et al*, 2016).

Several lysine-specific demethylases have been identified, but not yet for arginine (Yang & Bedford, 2013). There is, however, some evidence that does point to the existence of arginine-specific demethylase(s). For example the very dynamic changes of H3R17me2a in response to signaling (Le Romancer *et al*, 2008; Métivier *et al*, 2003) or throughout the cell cycle (Sakabe & Hart, 2010). A different strategy antagonizing arginine methylation marks by preventing methylation is the conversion of arginine to citrulline by deiminases like PADI4 (Cuthbert *et al*, 2004).

1.2.2 Domains that recognize methylated proteins

Even though the terms “writer”, “eraser”, and “reader” are widely used to distinguish three different types of proteins, they can be ambiguous. A “reader” is sometimes seen as a protein that recognizes a specific mark, but then only serves to recruit other proteins without changing the mark itself. An “eraser” on the other hand, technically must also contain a domain that recognizes the mark that it removes. In this section, I will therefore not use these distinctions to refer to proteins as a whole, but instead to describe the different “reader *domains*” that can recognize a methylated lysine or arginine.

Proteins that bind methylated proteins can be categorized according to the methyl-binding domain they contain, although some proteins contain multiple different domains, a famous example being UHRF1, containing Tudor, plant homeodomain (PHD), and Really Interesting New Gene (RING) finger domains (Xie *et al*, 2012). For an overview of the different domains, specifically for the proteins binding methylated lysines on histones, and which modified lysines they bind to, see Figure 7. In plants, the Agenet, Chromo, PWWP, MBT and Tudor domains

all share a three β-stranded core region which may point to a common protein ancestor. Together, these are called the Tudor “Royal family” (Maurer-Stroh *et al*, 2003) and including the PHD finger proteins, there are 202 known members in the human genome (Milosevich *et al*, 2016b). They all share a structural feature called an “aromatic cage” in their respective methyl binding pockets. Since my main focus lies on Tudor domains, I will give a short overview over the other methyl-binding domains before going more into detail about the Tudor domains themselves.

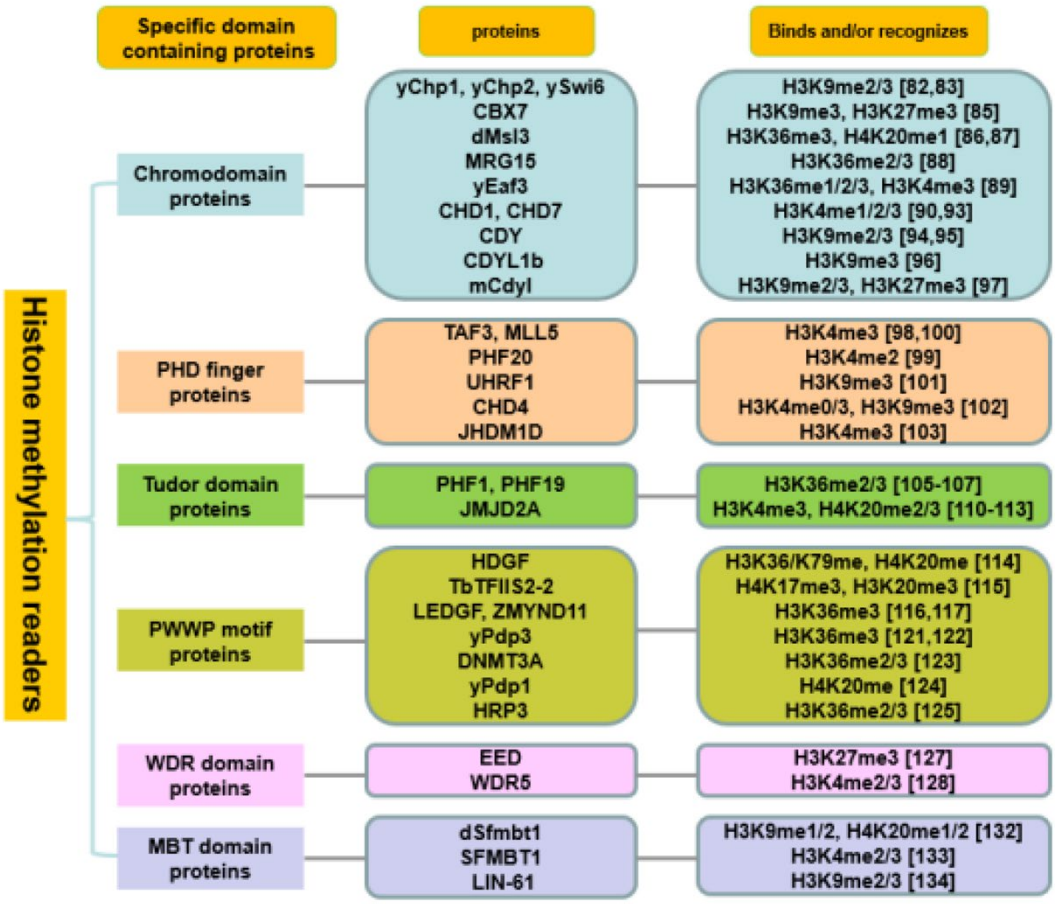


Figure 7: Examples of histone methylation readers.
Taken from Zhu *et al*, 2020.

1.2.3 Methyl-binding domains besides Tudor domains

Chromodomains (chromatin-organization-modifier domain) usually contain 30-70 amino-acids which form a three-stranded β-sheet followed by an adjacent C-terminal α-helix (Ball *et al*, 1997). They can be sub-divided further into the heterochromatin (HP1)/polycomb family, chromo-ATPase/helicase-DNA-binding (CHD) family, chromobarrel domain family, and the chromodomain Y chromosome (CDY) family (Eissenberg, 2012).

Plant homeodomain (PHD) finger domains contain a small zinc finger structural fold formed by 50–80 amino acid residues. They not only recognize the histone-modifications (also including acetylations) themselves, but do so in a sequence-dependent manner (Sanchez & Zhou, 2011).

Named after the conserved Pro-Trp-Trp-Pro motif, PWWP domains typically have a length of 100-130 amino acids (Wu *et al*, 2011). They are composed of a β-barrel core, an insertion motif between the second and third β-strands and a C-terminal α-helix bundle. PWWP domains were shown to bind not only methylated histones, but also DNA (Qiu *et al*, 2002).

WD40 repeat (WDR) domains, also known as β-propeller domains due to their centralized arrangement of β-sheet "blades", are often part of protein scaffolds, such as EED in the PRC2 complex. Since they often have a deep central cavity finding a small molecule inhibitor appears to be feasible (Schapira *et al*, 2017).

Malignant brain tumor (MBT) domains contain a 30–50 amino acid long N-terminal "arm" formed by helices, and a longer (60–80 amino acids) C-terminal beta subunit "core" (Bonasio *et al*, 2010). They all show a preference for mono- and dimethylated lysine which can be explained by the structure of the conserved aromatic cage (Milosevich *et al*, 2016b).

1.2.4 Tudor domains

Finally, there are Tudor domains, part of my target proteins SMNDC1 and SMN1. The eponym of the domain was identified in 1985 in *Drosophila melanogaster*, where the *tudor* gene was identified as required for germ plasm (Boswell & Mahowald, 1985). The resulting Tudor protein was found to contain 11 repeats of a domain that was later named Tudor domain. It wasn't until 1997 that this domain was found to be conserved among many other proteins (Callebaut & Mornon, 1997; Ponting, 1997).

As a first example, the structure of the Tudor domain protein SMN was solved (Selenko *et al*, 2001). A Tudor domain has around 60 amino acid residues which form an anti-parallel β-sheet composed of five β-strands folding into a barrel-like shape. At the same time, it was also shown that SMN's Tudor domain binds dimethylated arginine (Brahms *et al*, 2001; Friesen *et al*, 2001), with specificity for symmetrical (Liu *et al*, 2010b; Tripsianes *et al*, 2011) or asymmetrical dimethylation (Sikorsky *et al*, 2012). A bit later it became clear that Tudor domains can bind methylated lysines as well (Kim et al., 2006). In the same year the structure of methyl-lysine bound Tudor domains was solved already (Botuyan *et al*, 2006; Huang *et al*, 2006). The same was achieved for a methyl-arginine in 2010 (Liu *et al*, 2010a, 2010b), and specifically for SMN1 and SMNDC1, too (Tripsianes *et al*, 2011). The methylarginine is bound by an aromatic cage, in which cation-π interactions between the aromatic residues in the binding pocket and the cationic carbon of methylarginine occur, leading to an electrostatic stabilization (Figure 5B).

More recently, methyl-arginine independent binding modes were also identified, in this case for TDRD2 (Zhang *et al*, 2017). In humans, other known methyl-arginine binding proteins are Tudor domain-containing protein TDRD 1/2/3/6/9/11, while ~36 proteins are known to contain at least one Tudor domain (Blanc & Richard, 2017; Gayatri & Bedford, 2014). For an overview of all human Tudor domain proteins, see Figure 11.

There are a number of known functions for Tudor domains and the proteins that contain them, nicely summarized in a review by Pek et al., 2012. Overall, the Tudor domain appears to be important for the formation of protein complexes at specific membraneless compartments (Friberg *et al*, 2009; Liu *et al*, 2010a, 2010b; Tripsianes *et al*, 2011); also within one protein, the Tudor domain often serves as an adaptor and another domain works as the effector.

A number of Tudor domain proteins have functions in RNA metabolism, either by interacting with RNA directly, or by binding to other RNA-binding proteins. SMN1 and SMNDC1 are notable Tudor domain proteins involved in mRNA splicing, see the dedicated section “1.1.7 SMNDC1 and SMN – new phase-separating proteins?” for details.

Tudor domain proteins also play an important role in small RNA pathways. In the microRNA (miRNA) and Small interfering RNA (siRNA) pathways, endogenous (miRNA and siRNA) or exogenous (siRNA) hairpin transcripts are matured. The resulting 20 to 25-nucleotide single strand molecules bind to specific mRNAs and lead to reduced protein expression, either through repression of transcription, translation, or degradation of mRNA. SND1 (TDRD11) is part of the RNA-induced silencing complex (RISC) (Caudy *et al*, 2003) important in this pathway. Furthermore, it was shown that SND1 localizes to stress granules (Gao *et al*, 2010) containing mRNAs that are stalled during translation (Thomas *et al*, 2011). Another Tudor domain protein shown to localize to stress granules is TDRD3 (Goulet *et al*, 2008; Linder *et al*, 2008). In another publication, not only symmetric arginine dimethylation but even SMN1’s Tudor domain was shown to be required for the regulation of stress granules (Chitiprolu *et al*, 2018) (see also section 1.1.7 SMNDC1 and SMN – new phase-separating proteins?).

Another small RNA pathway in which several Tudor domain proteins are involved is the Piwi-interacting RNA (piRNA) pathway, best studied in *Drosophila melanogaster*. It is also involved in repression of expression; in contrast to miRNA and siRNA, however, transposons are the primary target of piRNAs (Khurana & Theurkauf, 2010; Saito & Siomi, 2010; Senti & Brennecke, 2010). Primary processing of piRNAs happens in the Yb body, another phase-separated membraneless organelle, and involves the two Tudor domain proteins Yb (mouse TDRD12) (Hirakata *et al*, 2019; Szakmary *et al*, 2009) and Vreteno (Handler *et al*, 2011; Zamparini *et al*, 2011). The secondary amplification cycle happens in another membraneless organelle called nuage, possibly also formed through phase-separation (Nott *et al*, 2015). A number of Tudor proteins (TDRD1/2/4/5/6/7/9) are required for the formation of nuage and some are also directly involved in the actual processing of piRNAs (Pek *et al*, 2012).

Besides (small) RNA processing, other functions have been attributed to Tudor domain containing proteins. The human Tumor protein p53-binding protein 1 (TP53BP1) is a major player in DNA damage response. It binds to H4K20 found at sites of DNA damage (Botuyan *et al*, 2006), recruiting effector such as check point kinase 1 (CHEK1), check point kinase 2 (CHEK2) or ataxia telangiectasia mutated (ATM) to the site of DNA damage (FitzGerald *et al*, 2009).

Finally, Tudor domains of course also bind methylated lysines and arginines on histones, again both directly modifying other histone marks or recruiting other proteins which do. I will focus on a few examples from the different outcomes on gene transcription. TDRD3 (also part of the stress granule, see above) is a so-called co-activator recruiting other proteins to H3R17me2a and H4R3me2a to activate expression (Morettin *et al*, 2017; Yang *et al*, 2010). Binding of other Tudor domain proteins results in repression of the respective loci. Polycomb-like (PCL) protein recruits components of Polycomb repressive complex 2 (PRC2) depositing H3K27me3 (Casanova *et al*, 2011), a repressive mark. The Jumonji (Jmj) domain-containing (JHDM3/JMJD2) family of proteins harbors demethylase domains (JmjN/C) besides its Tudor domains, leading to a removal of the mark it binds to (Cloos *et al*, 2006; Fodor *et al*, 2006; Huang *et al*, 2006; Klose *et al*, 2006; Whetstine *et al*, 2006). Last, but not least, Tudor domain proteins can influence methylation levels of DNA, leading to repression, too. Prime example is ubiquitin-like, containing PHD and RING finger domains 1 (UHRF1). This protein binds the repressive mark H3K9me3 and recruits DNA (cytosine-5)-methyltransferase 1 (DNMT1) to chromatin to deposit methyl-marks directly to DNA, linking the two levels of epigenetic control of gene expression (Bostick *et al*, 2007; Nady *et al*, 2011; Rottach *et al*, 2010). More recently, the structure of UHRF2 bound to H3K9me3 has been solved, too (Konopka *et al*, 2022).

1.2.5 Inhibitors of methyl-reader domains or proteins

Because of a growing interest in methyl-reader proteins and because they've been implicated in various diseases (Park *et al*, 2022), there are ongoing efforts towards finding specific small-molecule inhibitors targeting them and their pathways. Compared to a knock-out of a protein, small molecules are doseable, rapid, and reversible, and if targeting a specific domain, leave the rest of the protein and its function intact. Methyltransferases and demethylases have been targeted to a greater extent than reader domains. Also, within reader domains, bromodomain-containing proteins which recognize acetylated lysine residues have been targeted a lot more, and there specific inhibitors for several bromodomain proteins available (Kaniskan *et al*, 2015; Thinnes *et al*, 2014; Zaware & Zhou, 2019). In principle, the methyl-lysine or methyl-arginine binding pockets offer an obvious drug target site. The reviews by Milosevich et al., 2016, updated by Arrowsmith & Schapira, 2019 and Zhu et al., 2020 give an overview over small-

molecules targeting methyl-lysine readers and all histone-mark readers in cancer, respectively.

Since chromodomain protein CBX7 is a well-known player in cancer (Bernard *et al*, 2005), it has been an early target of drug development. Several inhibitors have been reported (Simhadri *et al*, 2014), including some that show efficacy in cells (Ren *et al*, 2015, 2016; Stuckey *et al*, 2016). Other CBX-family member successfully targeted so far are CBX6 (Milosevich *et al*, 2016a) and CBX8 (Wang *et al*, 2020b).

The first PHD finger protein that was targeted was JARID1A (Wagner *et al*, 2012), and apart from that Pygo has been targeted (Miller *et al*, 2014), both showing only weak affinity and not developed further. More recently, the PHD fingers of BAZ2A and BAZ2B were targeted (Amato *et al*, 2018). In addition, the natural product Berberine was shown to exhibit its anti-tumor effect by binding in the in the tandem Tudor domain (TTD) and PHD-domain of UHRF1 (Gu *et al*, 2020).

The first reported chemical probe for a PWWP domain is directed against NSD3's PWWP1 domain and occupies its aromatic cage. It is active in cells and was shown to reduce proliferation in MOLM-13 cells (Böttcher *et al*, 2019). NSD2's PWWP domain has been targeted by two independent approaches as well. One small molecule led to inhibition of the binding to H3K36me2 in cells (Ferreira de Freitas *et al*, 2021), while another influenced NSD2's localization within nuclear subcompartments (Dilworth *et al*, 2021).

For WDR-domains as drug targets, there is a nice review by Schapira *et al*, 2017. Focus of research has been on WDR5 and its interaction with MLL and on EED. First reports of a WDR5 inhibitor were published in 2013 (Senisterra *et al*, 2013; Bolshan *et al*, 2013) focusing specifically on its interaction with MLL, which is important for proliferation of lymphomas. Since then, they have been continuously worked on and improved (Grebien *et al*, 2015; Getlik *et al*, 2016; Ye *et al*, 2019; Aho *et al*, 2019; Tian *et al*, 2020). Another line of studies have used and improved peptide mimetic antagonists, that were also shown to be cell permeable and repressed cell growth (Karatas *et al*, 2013, 2017).

WD40-containing protein EED is the regulatory subunit of the PRC2 complex which plays a key role in gene regulation and has consequently been an interesting drug target. In 2017, a number of papers could present highly potent inhibitors against EED's WDR domain (Li *et al*, 2017; Huang *et al*, 2017; He *et al*, 2017; Qi *et al*, 2017). More recently their different binding modes were analyzed and compared, allowing more rational drug design and improvement in the future (Huang *et al*, 2020). Furthermore, there has been a report of agonistic ligands re-activating EED harboring loss-of-function (LOF) mutation I363M (Suh *et al*, 2019).

MBT domains were predicted to be very druggable (Santiago *et al*, 2011) and have known roles in oncogenesis. Nevertheless, inhibitors have only been developed against L3MBTL1 and L3MBTL3. The inhibitor UNC669 against L3MBTL1 was the first small-molecule directed

against any methyl reader protein (Herold *et al*, 2011), identified using ALPHA-screen technology like the one used in my screening approach (Wigle *et al*, 2010). However, it shows no selectivity between L3MBTL1 and 3 (Herold *et al*, 2012) which is difficult to achieve due to the similarity of their aromatic cages. Further work in the group did lead to specific inhibitors though (Camerino *et al*, 2013; James *et al*, 2013a) which were optimized (James *et al*, 2013b) and resulted in UNC1215 as the most potent compound.

Eventually, there are some small molecules targeted against Tudor domains, too. For an overview of all Tudor domains and the subset with published inhibitors, see Figure 11. Several studies dealing with inhibitors against Spindlin1's Tudor domain have been published, the first in 2016, also establishing a screening platform starting with an ALPHA-screen (Wagner *et al*, 2016). Their screening hit A366 (Figure 8A) had a potent IC_{50} of ~ 180 nM and showed cellular target engagement in a cellular thermal shift assay (CETSA). It was originally published as an inhibitor of KMT G9a, and is therefore not selective (Sweis *et al*, 2014). Employing protein microarrays and using UNC1215, a potent L3MBTL3 inhibitor (James *et al*, 2013a) as a starting point for analogues, (Bae *et al*, 2017)) could present EML631 as a specific Spindlin1 inhibitor (Figure 8B). It was shown to bind with a K_d of ~3 µM, engage SPIN1 in cells, block its ability to "read" H3K4me3 marks, and inhibit its transcriptional coactivator activity. The authors could also show a co-crystal structure in which you can see that EML631 interacts with the binding pocket of Tudor domain 1, but also with Tudor domain 2 (Figure 8C). Basically combining the two Spindlin1 inhibitors found before and to increase specificity and potence, bidentate inhibitors targeting two of the three Tudor domains were designed, synthesized and tested (Fagan *et al*, 2019). This resulted in "VinSpinIn" (Vinnie's Spindlin Inhibitor) (Figure 8D), which was shown to be potent (isothermal titration calorimetry ITC K_d = 9.9 nM), also cellularly active (EC_{50} =270 nM), and had no nonspecific toxicity. Crystal structure of the compound in complex with SPIN1 showed the intended bidentate binding mode, with the ethylpyrrolidine moiety occupying the aromatic cage of domain 1 and the isoindoline binding to domain 2 (Figure 8E). Another approach by the same group of authors that identified A366 as a non-specific Spindlin1 inhibitor led to the development of a specific Spindlin1 inhibitor MS31 (Figure 8F) (Xiong *et al*, 2019). Starting from UNC0638, also a G9a inhibitor, as their initial screening hit, they could achieve specificity for the Spindlin family and potency with an ITC K_d value of 91±4nM. In an X-ray co-crystal structure, it was shown that the isoindolinyl group occupies the aromatic cage of Tudor domain 2 with the protonated amino group forming a hydrogen bond with Y179 but also interacting with Y170, W151, and F141 in the aromatic cage through cation−π interactions (Figure 8G). Furthermore, there is an interaction of the phenyl ring of the isoindoline group with W151 through π−π-stacking. MS31 was also shown to be active in cells; a bidentate binding mode does not seem to be necessary.

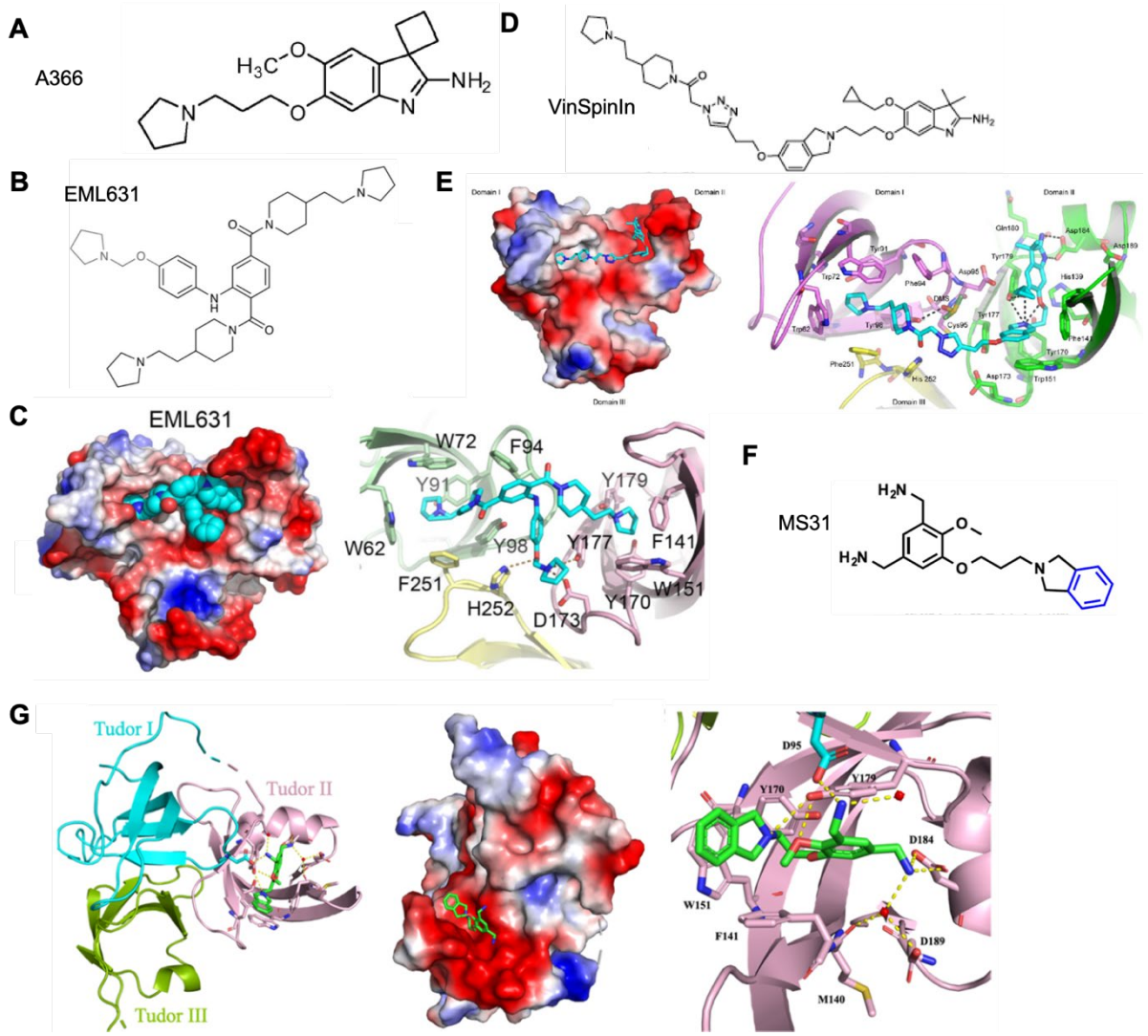


Figure 8: Inhibitors against Tudor domain protein Spindlin1.
A: compound A366. Taken from Wagner *et al*, 2016. B: compound EML631. C: compound EML631 bound to Spindlin1's tudor domains 1 (green) and 2 (pink). Taken from Bae *et al*, 2017. D: compound VinSpinIn. E: VinSpinIn bound to Spindlin1's Tudor domains 1 (pink) and 2 (green) (PDB: 6I8B). Taken from Fagan *et al*, 2019. F: compound MS31. G: MS31 bound to Spindlin 1 Tudor domain 2 (PDB: 6QPL). Taken from Xiong *et al*, 2019.

The first Tudor protein ever targeted was the previously mentioned TP53BP1, hypothesizing that inhibition could restore genomic stability in cells with BRCA1 mutations (Perfetti *et al*, 2015). They identified UNC2170 (Figure 9A) using an ALPHA-screen approach with a K_d of 22 ± 2.5 µM and could show that binding depends on hydrogen bonding between NH^+ and Asp1521 (Figure 9B). Testing analogs to get insight into structure-activity relationships (SAR) did not lead to improved potency or selectivity. In an independent approach, Sun et al. discovered a new TP53BP1 inhibitor employing AlphaScreen-based high-throughput

screening. Their best hit, DP308 (Figure 9C) had an improved K_d in the low µM range and could be modeled to bind in the same pocket as UNC2170 (Figure 9D) (Sun *et al*, 2021).
UHRF1's TTD was targeted to inhibit binding to H3K9me3 in a peptide displacement assay, resulting in hit NV01 with a K_d of 5 µM (Figure 9E) (Senisterra *et al*, 2018). Co-crystallization could prove the interaction in the binding pocket, relying on π-stacking of the oxothieno-pyrrolo-triazine moiety of NV01 in a primary aromatic cage formed by F152, Y188, and Y191 from TTD_N (Figure 9F). The authors could achieve a ~2-fold potency improvement with limited SAR, which did not allow cellular assay testing. Furthermore, two similar fragments have been found to bind to UHRF1's TTD in independent approaches. 2,4-lutidine (Figure 9G) was shown to bind at an EC_{50} of 10 µM in an AlphaScreen, and it could be shown that it actually binds at two different places, with the largest chemical shift changes induced in residues close to the interface in between the two Tudor domains (10 largest NMR chemical shift changes mapped on the X-ray co-crystal structure, Figure 9H) (Chang *et al*, 2021). 5-amino-2,4-dimethylpyridine (5A-DMP, Figure 9I) had a K_d= 19.3 µM in isothermal titration calorimetry (ITC) and a very similar binding mode compared to 2,4-lutidine could be revealed in a co-crystal structure (Figure 9J) (Kori *et al*, 2021).

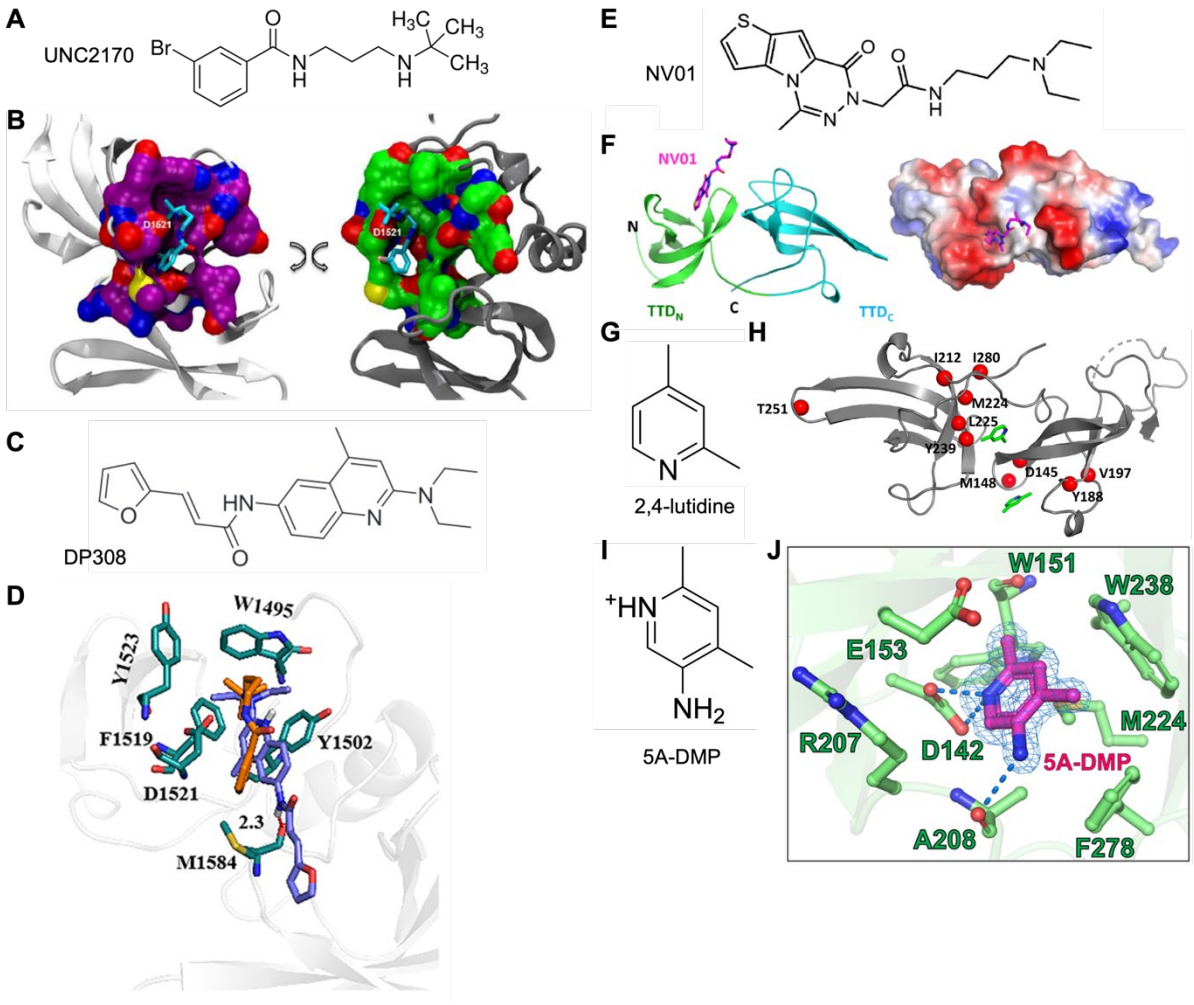


Figure 9: Inhibitors against TP53BP1 and UHRF1.
A: compound UNC2170. B: UNC2170 bound to TP53BP1 (PDB: 4RG2). Taken from Perfetti *et al*, 2015. C: Compound DP308. D: Putative binding mode of DP308 (purple), overlayed with UNC2170 (orange) in the co-crystal structure of UNC2170 with TP53BP1 (PDB: 4RG2, same as in B). E: compound NV01. F: NV01 bound to TTD_N and TTD_C of UHRF1 (PDB: 3DB4). Taken from Senisterra *et al*, 2018. G: 2,4-lutidine. H: Largest chemical shift changes mapped on the TTD X-ray co-crystal structure of UHRF1 and 2,4-lutidine (PDB ID: 6VYJ). Taken from Chang *et al*, 2021. I: Compound 5A-DMP. J: Co-crystal structure of the Arginine-binding pocket of UHRF1 with 5A-DMP, blue dashed lines indicate hydrogen bonds. Taken from Kori *et al*, 2021.

Employing Nuclear Magnetic Resonance (NMR) based screening on TDRD3, 14 fragments could be identified as hits, including structure “1” (Figure 10A), which was also successfully co-crystallized with the Tudor domain (Figure 10B) (Liu *et al*, 2018). This compound is bound by the aromatic cage through a sandwich-like π-π interaction and a hydrogen bond to N596.

Another study relying on NMR screening focused on histone demethylase KDM4A's TTD (Upadhyay *et al*, 2018). They found that fragment "1a" (Figure 10C), the enantiomer (2*S*,3*R*)-3-((dimethylamino)methyl)-1-phenylbicyclo[2.2.1]heptan-2-ol (and not enantiomer (2*R*,3*S*)) binds KDM4A's Tudor domain by forming a network of strong hydrogen bonds and hydrophobic interactions (Figure 10D).

KMT SETDB1's TTD was also targeted in a structural screening approach combining NMR and X-ray screening (Mader *et al*, 2019). SETDB1 is amplified in several human cancers and its overexpression seems to be important for their growth (Ceol *et al*, 2011; Rodriguez-Paredes *et al*, 2014; Sun *et al*, 2015); nevertheless, only one other study worked on an inhibitor, focusing on the SET domain (Park *et al*, 2017). Mader et al. reported several fragment hits that bound to the TTD's peptide binding pocket. Fragment "1" (Figure 10E) was found to bind the Kme binding pocket, but showed only weak interactions and could not be improved further (Figure 10F).

To target PHF1's Tudor domain, Engelberg et al. employed a peptidomimetic (UNC6641, Figure 10G) with a K_d= 0.96 µM and could obtain a co-crystal structure illustrating the binding mode (Figure 10H) (Engelberg *et al*, 2021).

Last but not least, also SMN1 was targeted with small molecule inhibitors. While screening for inhibitors of UHRF1's TTD, Liu et al. found a fragment inhibitor ("Compound 1", Figure 10I) which showed some selectivity for SMN1 (K_d= 2.6 µM) over UHRF1 (K_d= 16 µM) and TDRD3 (K_d= 11 µM) (Liu *et al*, 2022). Interestingly, "Compound 1" also bound SMNDC1's Tudor domain with a comparable K_d= 2.8 µM. The authors were able to co-crystallize "Compound 1" with the Tudor domains of SMN1 (Figure 10J), TDRD3 and UHRF1. Additionally, they could show that treatment of cells with "Compound 1" leads to a disruption of the interaction of SMN1 and Pol II.

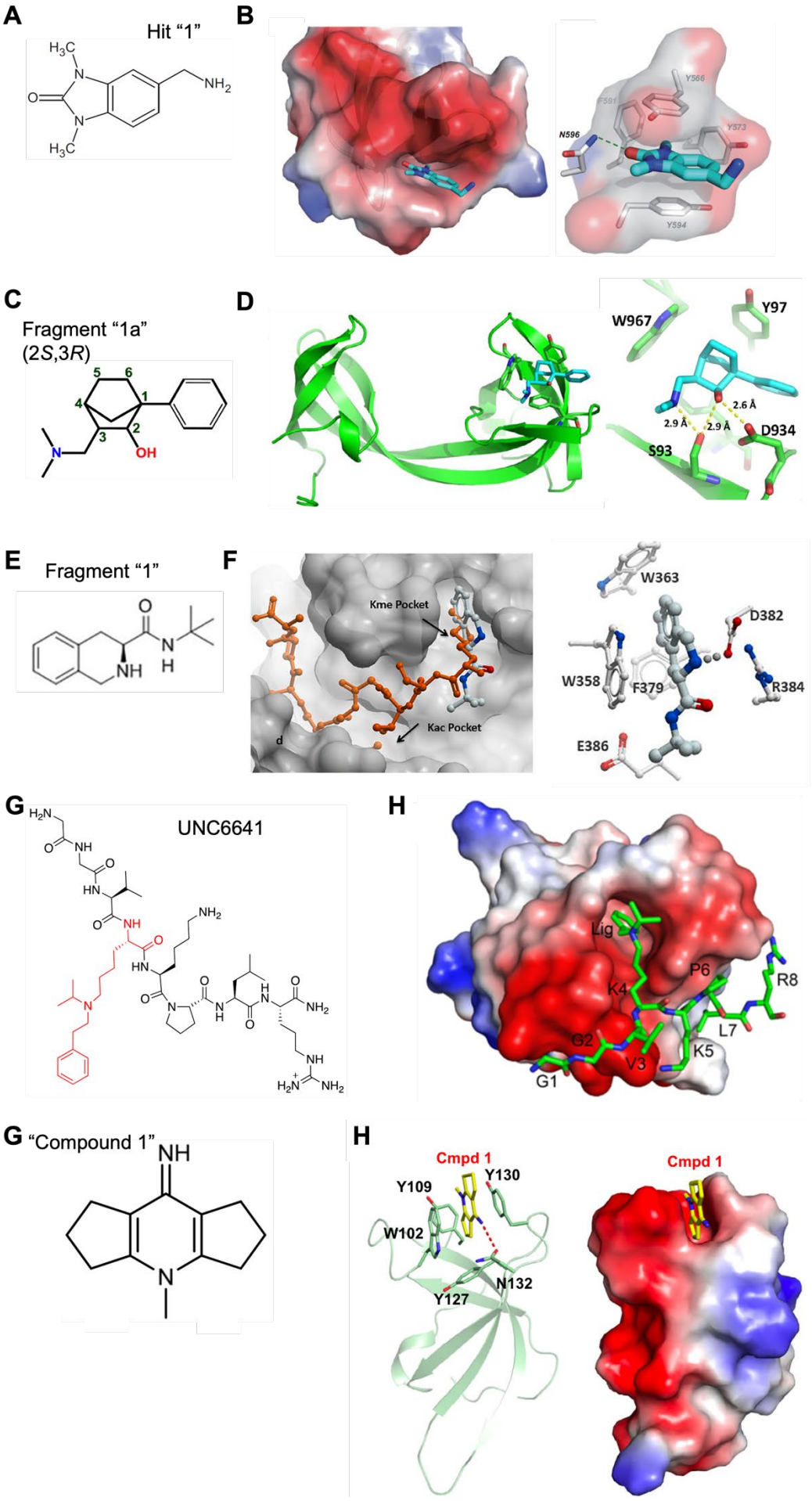

A
Hit "1"
B
C
Fragment "1a"
(2S,3R)
D
W967
Y97
2.6 Å
2.9 Å
2.9 Å
D934
S93
E
Fragment "1"
F
Kme Pocket
Kac Pocket
W363
D382
W358
F379
R384
E386
G
UNC6641
H
Lig
P6
R8
K4
L7
G2
K5
V3
G1
G
"Compound 1"
H
Cmpd 1
Y109
Y130
W102
N132
Y127
Cmpd 1

Figure 10: Inhibitors against TDRD3, KDM4A, SETDB1, PHF1 and SMN1.
A: Hit “1” from Liu *et al*, 2018. B: Hit “1” bound to TDRD3 (PDB: 5YJ8). Taken from Liu *et al*, 2018. C: Fragment “1a” from Upadhyay *et al*, 2018. D: Fragment “1a” bound to KDM4A’s Tudor domain (PDB: 5VAR). Taken from Upadhyay *et al*, 2018. E: Fragment “1” from Mader *et al*, 2019. F: Fragment “1” bound to SETDB1’s Tudor domain Kme pocket. Left: superimposed with H3K9me2K14ac (PDB: 5KE2, 6BHD). Taken from Mader *et al*, 2019. G: UNC6641. H: Co-crystal structure of UNC6641 with PHF1’s Tudor domain (PDB 7LKY). Taken from Engelberg *et al*, 2021. I: “Compound 1”. J: Cartoon and electrostatic potential surface representation of the co-crystal structure of SMN1’s Tudor domain and “compound 1” (PDB: 4QQ6). Taken from Liu *et al*, 2022.

Overall, even if still low in numbers, more and more methyl-reader domains including Tudor domains are targeted by small molecule inhibitors (Figure 11). For Tudor domains, a clear cellular effect of protein inhibition is still lacking. This could be explained by the fact that a lot of readers act in larger complexes or have multiple reader domains. Inhibition of just one domain could be insufficient for a displacement and loss of function. A promising approach could be targeting two domains or domain-interfaces in addition to the binding pocket with a bidentate inhibitor, although it doesn’t seem to be necessary to achieve good potency (Xiong *et al*, 2019).
An important Tudor domain protein which has not been targeted specifically with a small molecule inhibitor yet is SMNDC1 (Figure 11 and section 1.1.7).

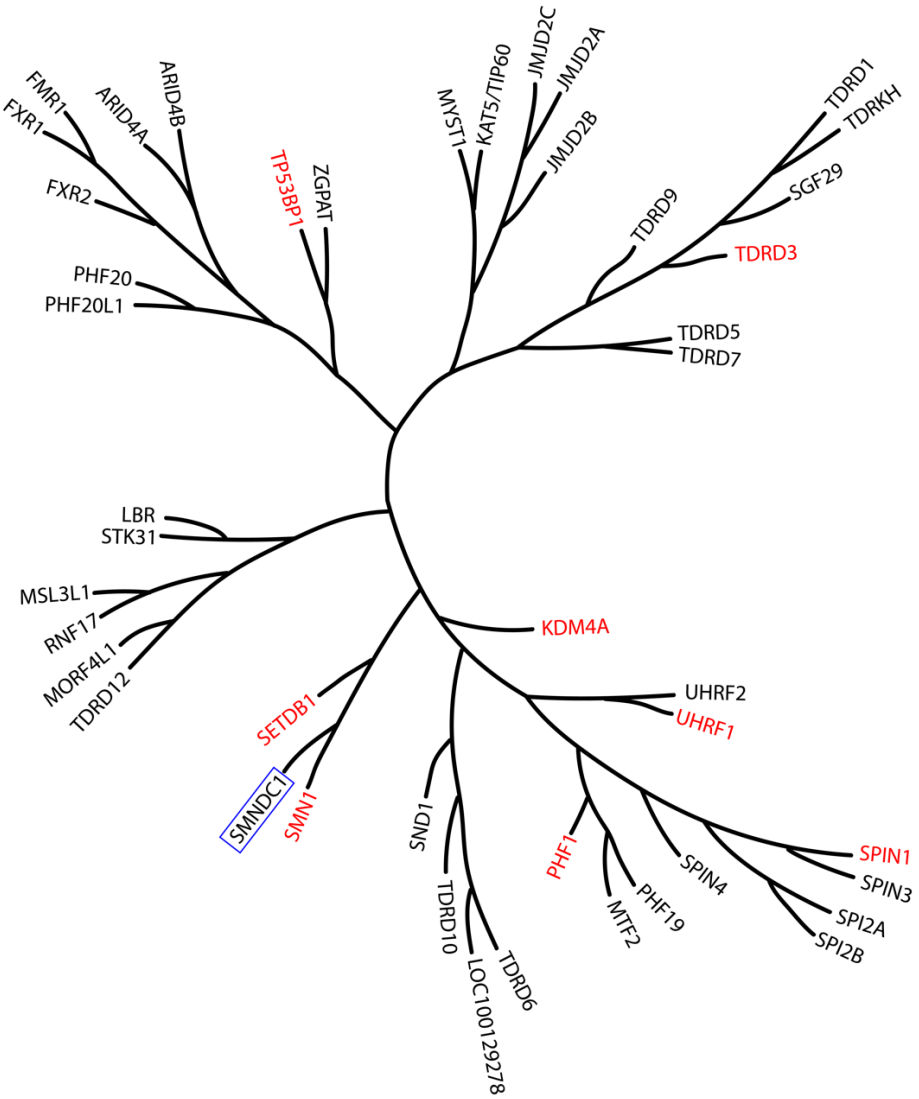


Figure 11: Human tudor domain phylogenetic tree.

Tudor domains with a published inhibitor colored red, SMNDC1 with a blue box.

1.3 Aims

As described all layers of gene regulation are controlled and affected by phase-separation. Studying this regulation with particular focus on SMNDC1, an understudied splicing factor implicated in the repression of insulin in α-cells, appears as a great opportunity.

First aim of this thesis was to characterize SMNDC1's subcellular localization by cell lines with fluorescently labeled SMNDC1. Furthermore, I addressed SMNDC1's LLPS behavior both *in vitro* using purified protein with a focus on the individual domains of SMNDC1 and *in vivo* via FRAP. To be able to study SMNDC1's function in a dose- and time-dependent manner my aim was to find small molecules inhibiting SMNDC1's Tudor domain using an AlphaScreen setup. By testing a large number of analogs, I wanted to develop a deep understanding of structure-activity relationships. Intending to fathom the binding mode of the most promising compound I established a collaboration with a laboratory from the Technical University of Munich which was able to provide an NMR structure of SMNDC1's Tudor domain with the small molecule. My final aim was the characterization of cellular effects of SMNDC1 inhibitors on localization and interactome of SMNDC1 and global splicing effects.

2. **RESULTS**

Pharmacological perturbation of the phase-separating protein SMNDC1

Lennart Enders[1], Marton Siklos[1], Jan Borggräfe[2,3], Stefan Gaussmann[2,3], Anna Koren[1], Monika Malik[1], Tatjana Tomek[1], Michael Schuster[1], Jiří Reiniš[1], Elisa Hahn[1], Andrea Rukavina[1], Andreas Reicher[1], Tamara Casteels[1,$], Christoph Bock[1,4], Georg E. Winter[1], J. Thomas Hannich[1], Michael Sattler[2,3], and Stefan Kubicek[1,*]

[1]CeMM Research Center for Molecular Medicine of the Austrian Academy of Sciences, Lazarettgasse 14, 1090 Vienna, Austria.
[2]Helmholtz Munich, Molecular Targets and Therapeutics Center, Institute of Structural Biology, Neuherberg, 85764 München, Germany.
[3]Technical University of Munich, TUM School of Natural Sciences, Department of Bioscience, Bavarian NMR Center, Garching, 85748 München, Germany.
[4]Medical University of Vienna, Institute of Artificial Intelligence, Center for Medical Data Science, Vienna, Austria.

*Correspondence: Stefan Kubicek, skubicek@cemm.oeaw.ac.at.
[$]Present Address: Sloan Kettering Institute, 1275 York Avenue, New York, NY 10065, USA

Keywords
Tudor domain, small molecule inhibitor, chemical screening, NMR, structural biology, nuclear speckles, splicing, pancreatic alpha cells

Abstract

SMNDC1 is a Tudor domain protein that recognizes di-methylated arginines and controls gene expression as an essential splicing factor. Here, we study the specific contributions of the SMNDC1 Tudor domain to protein-protein interactions, subcellular localization, and molecular function. To perturb the protein function in cells, we develop small molecule inhibitors targeting the dimethyl arginine binding pocket of the SMNDC1 Tudor domain. We find that SMNDC1 localizes to phase-separated membraneless organelles that partially overlap with nuclear speckles. This condensation behavior is driven by the unstructured C-terminal region of SMNDC1, depends on RNA interaction and can be recapitulated *in vitro*. Inhibitors of the protein's Tudor domain drastically alter protein-protein interactions and subcellular localization, causing splicing changes for SMNDC1-dependent genes. These compounds will enable further pharmacological studies on the role of SMNDC1 in the regulation of nuclear condensates, gene regulation and cell identity.

Introduction

Survival motor neuron domain-containing protein 1 (SMNDC1), also called Survival of motor neuron-related-splicing factor 30 (SPF30), is an essential splicing factor required for the formation of the spliceosome[1,2]. To promote spliceosome assembly SMNDC1 binds to methylated arginines on Sm-proteins using its Tudor domain[2,3], similar to its better-studied paralog survival of motor neuron (SMN) protein[4–6]. The Tudor domain structures of both proteins are highly conserved, revealing binding of their substrate symmetric dimethylated arginine (sDMA) in an aromatic cage through cation-π interactions[7]. Functionally, both proteins play essential and apparently opposite roles in the regulation of gene expression and cell identity in the endocrine pancreas. Patients and animal models with SMN mutations experience increased numbers of glucagon producing alpha cells and a reduction of insulin producing beta cells[8]. In contrast, for SMNDC1 we recently showed that its knock-down causes the upregulation of insulin in α-cells through splicing changes in key chromatin remodelers and induction of the beta cell transcription factor PDX1[9]. SMNDC1 further is essential for cell proliferation in different contexts, and a recent study reported worse survival in hepatocellular carcinoma patients with high SMNDC1[10]. SMNDC1 knock-down led to decreased proliferation and migration of hepatocellular carcinoma cells, establishing SMNDC1 as a potential therapeutic target.

Both SMN and SMNDC1 show distinct and focal subcellular localization patterns. The SMN Tudor domain is sufficient for formation of a phase-separated compartment dependent on the dimethylarginine (DMA) modification of binding proteins[11] and was shown to be required for the regulation of the phase-separated stress granules via symmetric dimethylarginine (sDMA)[12]. Arginine methylation in RGG/RG motifs recognized by Tudor domains can affect phase-separation of Fused in sarcoma (FUS)[13,14] and other proteins[15], and further Tudor domain containing proteins themselves have been shown to be involved in phase separation[16,17].

SMNDC1 has a speckled localization within the nucleus that – based on co-localization – was attributed to the sub-nuclear structures Cajal bodies and nuclear speckles[2], which were later defined as prime examples of membraneless organelles[18], i.e. biomolecular condensates formed by liquid-liquid phase separation (LLPS). These assemblies can consist of proteins, nucleic acids, and other molecules and are found both in the cytoplasm and the nucleus[19,20]. An important feature present in many proteins that were found to undergo LLPS are intrinsically disordered regions (IDRs), which do not adopt a well-defined globular structure. IDRs can enable multiple and multivalent interactions that mediate binding to other proteins[21]. Many RNA-binding proteins (RBPs), including SMNDC1, were found to phase-separate together with RNA, but also with chromatin[22]. Amongst other factors phase separation

behavior can be initiated by RNA[23] and regulated by the secondary structure of RNAs and the ratio of RNA to RBPs[24–26]. Given the fact that the nucleus and its sub-compartments are enriched in IDR-containing proteins (IDPs)[27] and the obvious abundance of negatively charged nucleic acids (both DNA and RNA) the nucleus is primed for LLPS[28]. Functionally, these LLPS events control gene expression within the different nuclear compartments[29] from the formation of heterochromatin[30,31] over transcription by RNA polymerase II[32] to RNA processing and (alternative) splicing[33].

Tudor domains have not been targeted extensively by small-molecule inhibitors. Only recently, a study disclosed a fragment unspecifically binding to both SMN and SMNDC1 in isothermal titration calorimetry (ITC), and with cellular specificity for SMN[34]. Similarly, specific agents perturbing biomolecular condensation events are lacking, and pharmacological approaches often rely on unspecific agents like 1,6-hexanediol[35] at concentrations of several hundred millimolar.

Here, we study the phase-separating behavior of SMNDC1 both *in vitro* and within cells and we develop specific inhibitors against its Tudor domain influencing the sub-cellular localization and phase separation of their target.

Results

SMNDC1 co-localizes with nuclear speckle markers

To identify features associated with subcellular SMNDC1 localization, we analyzed the protein sequence by comparing predictions for disordered regions by MetaDisorder[36] and for the full-length structure by AlphaFold[37,38] (Fig. 1a). The experimentally solved Tudor domain structure[7] (residues 64-128) and two interacting N-terminal alpha-helices (residues 2-25, and 30-52) are visible both in the AlphaFold prediction and in the disorder tendency plot as ordered regions. AlphaFold in addition predicts a long C-terminal alpha-helix, for which however currently no other experimental evidence exists.

We employed an endogenous tagging system that targets introns and introduces a GFP-tag as an artificial exon[39] to characterize SMNDC1's cellular functions. To rule out disrupting effects of the tag on protein localization, we targeted all of SMNDC1's introns in murine alphaTC1 cells, and then isolated clonal sublines. The targeted introns result in GFP integrations covering all regions of the protein, including one at the N-terminus (before residue 1), the N-terminal region (residue 40), the Tudor domain (residue 88), and a long stretch in the C-terminal region (residue 142, residue 193) which is predicted to be disordered[36] (Fig. 1a, b). Furthermore, we also tagged intron 2-3 in human HAP1 cells. Typically, these monoallelic tagging events resulted in cells expressing both un-tagged and GFP-tagged SMNDC1 at comparable levels as shown by western blot (WB) (Fig. 1c, quantifications and full membranes Supplementary Fig. 1b, c). The GFP-tag within the Tudor domain (intron 3-4)

showed the lowest relative expression levels, indicating possible interference with folding efficiency.

AlphaFold structure predictions[37,38] for SMNDC1 with and without GFP in the different introns revealed that the GFP-tag does not seem to disrupt the overall structure of the protein (Suppl Fig. 1a). All structural elements such as the N- and C-terminal α-helices and the Tudor domain (red) are predicted to form normally, even when the GFP-tag interrupts the Tudor domain (intron 3-4). Accordingly, all of the different intron tagged clones, including the intra-Tudor GFP integration showed consistent subcellular localization pattern (Fig. 1b). These GFP fusions showed the same speckled nuclear localization avoiding DNA-dense regions as observed for the endogenous protein by antibody-based immunofluorescence (IF) (Fig. 1d). During M-phase of the cell cycle SMNDC1 dissipated to the whole cell and formed distinct droplets called mitotic interchromatin granules[40,41] (Fig. 1e), a behavior which is typical for nuclear speckle proteins[42,43]. SMNDC1 also reacted to the overexpression of the cell-cycle dependent kinases DYRK3 and CLK1, which is known to dissolve nuclear speckles[43,44], with a loss of its focal nuclear localization (Supplementary Fig. 1d).

To further characterize SMNDC1's localization in the nucleus, we co-stained cells with antibodies against SMNDC1 and SC35, a marker for nuclear speckles. Both signals overlap to a large degree and avoid chromatin-dense regions, whereby SMNDC1 shows a wider less focal distribution (Fig. 1f, co-localization analysis Supplementary Fig. 1e). To be able to visualize nuclear speckles in live cells we RFP-tagged SRRM2 in the SMNDC1-GFP-tagged cells (Fig. 1g). SRRM2 is the target of the SC35 antibody[45] and scaffolding protein of nuclear speckles[46].

Endogenously tagged SMNDC1-GFP and SRRM2-RFP co-localized to a large degree, both in interphase and during mitosis (Fig. 1h). Even though co-localization was maintained in the mitotic interchromatin granules, there SMNDC1 showed a higher degree of diffuse localization, leading to a lower average Pearson correlation score compared to interphase cells (Supplementary Fig. 1f). Overall, we find that SMNDC1 shows behavior and localization typical for proteins in nuclear speckles, which have been described as membraneless organelles in the nucleus formed by LLPS.

SMNDC1 undergoes biomolecular condensation *in vitro* and in cellular systems

A common way to prove phase-separating behavior of a protein is to show its ability to form droplets in a purified form *in vitro*. To do so, we expressed and purified full SMNDC1 with an N-terminal GFP-tag and mixed it with PEG-8000 as a surrogate for the crowded environment of a cell. We observed droplet formation (Fig. 2a) and fusion of droplets (Fig. 2b). Subsequently we tested the influence of other biomolecules and salt concentration on droplet formation (Fig. 2c). Addition of RNA to the PEG-8000 containing buffer enhanced SMNDC1's

droplet formation while high NaCl concentrations prevented droplet formation. Digestion of RNA by RNase led to the dissolution of droplets, even after their formation (Fig. 2d). RNA also physically localized to the protein droplets (Fig. 2e).

To further understand which part of the protein is responsible for the formation of droplets, we fused different SMNDC1 truncations (Fig. 1a) to GFP and subjected them to the same treatment in buffer containing RNA and PEG-8000. These experiments clearly displayed that the C-terminal region after the Tudor domain (constructs 5 and 6), which is predicted to be intrinsically disordered[36], was sufficient to induce droplet formation with RNA (Fig. 2f, see Fig. 1a for a scheme of the truncated forms), which fit the predicted IDR scores[36] (Fig. 1a). We also confirmed that the Tudor domain alone (construct 3) cannot form droplets, consistent with previous literature[11].

To show the reversibility of phase separation *in vivo*, the aliphatic alcohol 1,6-hexanediol which interferes with weak hydrophobic interactions is often used to dissolve protein condensates[35]. SMNDC1-GFP exhibited the expected phenotype in live cells treated with 1,6-hexanediol by losing its focal localization within the nucleus (Fig. 2g). Another way to characterize the molecular dynamics and mobility of phase-separating proteins in cells is to analyze the diffusion of a fluorescently labeled protein by fluorescence recovery after photobleaching (FRAP). When bleaching SMNDC1-GFP and SRRM2-RFP, fluorescence recovered within 30 seconds (Fig. 2h), consistent with liquid-like behavior rather than protein aggregation. These data provide evidence that SMNDC1 undergoes phase separation, both *in vitro* and in membraneless organelles within the nucleus, presumably nuclear speckles.

Full-length SMNDC1 interacts with nuclear speckle proteins

We set out to characterize SMNDC1's interactome using proximity labeling by overexpressing an SMNDC1-APEX2 fusion protein (Fig. 3a). Compared to classical co-immunoprecipitation (Co-IP), this recently developed method[47] is better suited to capture weak and transient interactions as they are expected in phase-separated compartments like nuclear speckles. In addition to full-length SMNDC1 (APEX2-SMNDC1^{FL}), we also performed proximity labeling with a fusion protein of APEX2 with a truncated SMNDC1 consisting of only the Tudor domain and therefore lacking N-terminal and C-terminal regions, and the nuclear localization signal (NLS) (APEX2-SMNDC1^{TD}) (Supplementary Fig. 2a). To verify our approach, we performed proximity labeling followed by an immunofluorescence (IF) staining against SMNDC1 and biotin. APEX2-SMNDC1^{FL} caused biotinylation in the areas where SMNDC1 is localized: nuclear while avoiding chromatin-dense regions (Fig. 3b). Much less biotinylation was observed when omitting H_2O_2. The control overexpression of APEX2-SMNDC1^{TD} on the other hand showed a uniform localization throughout the cell and a corresponding biotinylation pattern. On a western blot, a ladder of biotin-labelled proteins was visible, but absent when

leaving out the H_2O_2 during the labeling. More proteins appear to be labelled by the ubiquitously localized APEX2-SMNDC1^{TD} fusion (Supplementary Fig. 2b).

Analyzing the biotinylated and enriched proteins by mass spectrometry (MS), we identified and quantified a large number of proteins (~3200) in the proximity of APEX2-SMNDC1^{FL} and APEX2-SMNDC1^{TD}. Compared to the proximity interactome of APEX2-SMNDC1^{TD}, APEX2-SMNDC1^{FL} showed overall less interactions (Fig. 3c). We attribute this to the higher specificity of interactions happening with the correctly localized full form of SMNDC1. The fact that SMNDC1 itself was enriched in APEX2-SMNDC1^{FL} over APEX2-SMNDC1^{TD} suggests that labeling in trans works better if SMNDC1 is correctly localized and concentrated in its phase-separated compartment leading to more SMNDC1 protein in its proximity. Similarly, proteins known to be localized to the nucleus were not depleted in APEX2-SMNDC1^{FL} over APEX2-SMNDC1^{TD}, reflecting the loss of correct localization when the NLS is missing (Supplementary Fig. 2c).

We then filtered for proteins enriched in APEX2-SMNDC1^{FL} over APEX2-SMNDC1^{TD} (adjusted p-value <0.1, abundance ratio >1.1) which reduced the number of proteins we considered specific interactors of SMNDC1^{FL} to 750. As expected, we found an enrichment of proteins associated with mRNA processing, and more specifically splicing, but also an enrichment of proteins associated with ribosome biogenesis and rRNA processing amongst these (Fig. 3d). When comparing these interactors to an SMNDC1 Co-IP dataset generated in our lab[9] we found a significant overlap but confirm that proximity labeling can detect more and different interactions compared to a Co-IP (Fig. 3e). A majority of APEX2-SMNDC1^{FL} interactors was also identified by SRSF7-APEX2 proximity labeling[48] (Fig. 3f), suggesting that APEX2-SMNDC1^{FL} proximity labeling did enrich for proteins localized to nuclear speckles. Furthermore, we compared the interactors to proteins identified as symmetrically di-methylated on arginine residues in a deep protein methylation profiling study[49] (Supplementary Fig. 2d). Since these interactions are expected to be mediated through the Tudor domain APEX2-SMNDC1^{TD} should bind these proteins, too. Consequently, only a small subset was enriched in APEX2-SMNDC1^{FL} over APEX2-SMNDC1^{TD}. We therefore also compared the sDMA-modified proteins to all proteins identified in our SMNDC1-APEX2 experiments and found most of them (67 out of 87 known sDMA-modified proteins). There was also an enrichment, although to a lesser degree, of proteins with asymmetrical di-methylations. These protein sets partially overlap, as the same arginine sites can often alternatively be symmetrically or asymmetrically di-methylated.

Overall, we found a large interactome of SMNDC1 enriched for proteins interacting with RNA, localized to nuclear speckles, and with known sDMA modifications. We therefore suspected that the Tudor domain is responsible for a subset of SMNDC1's specific interactions.

A screen for small molecule SMNDC1 Tudor domain inhibitors

To pharmacologically perturb SMNDC1 function, we set out to identify small molecule inhibitors of SMNDC1's Tudor domain based on perturbing its interaction with a dimethyl-arginine peptide. To establish an AlphaScreen[50], we coupled donor beads to purified SMNDC1's (or SMN's) Tudor domain via a His-Tag and acceptor beads to a biotinylated peptide corresponding to the C-terminal region of the Small nuclear ribonucleoprotein Sm D3 containing four sDMAs (Fig. 4a). These interaction partners had previously been used in the structural study of SMNDC1 and SMN[7]. Protein domains were purified employing their His-Tag (Supplementary Fig. 3a). To identify ideal concentrations for screening, we performed a cross-titration of Tudor domains and binding peptides (Supplementary Fig. 3b). Since the AlphaScreen signal was sufficient for screening, we reduced the concentration of acceptor and donor beads to 5 µg/ml (Supplementary Fig. 3c).

Using this set-up with our in-house library of ~90,000 compounds (overview over screening strategy Fig. 4b, Supplementary Table 1), we identified 511 hits with signal <50 % of control (POC) (Fig. 4c). Since the AlphaScreen is susceptible to unspecific quenching of the singlet oxygen energy transfer we then performed a counter-screen using a crosslinking peptide which combines both affinity tags and therefore always brings donor and acceptor beads in close proximity (Fig. 4d). This led to a reduction to 40 hits that selectively inhibited the interaction between SMNDC1, and its arginine methylated binding partner. These compounds we next tested in dose response with the Tudor domains of both SMNDC1 and SMN. Several chemical scaffolds (Fig. 4e-k) of inhibitors were discovered by this screen, with IC_{50} values of 0.2 to 2 µM and different degrees of selectivity between the different Tudor domains. The molecules with the best physicochemical and structural properties were then used for the design of further analogs, aiming at improving potency and selectivity (Fig. 5).

4-arylthiazole-2-amines show clear structure-activity relationships as SMNDC1 Tudor domain inhibitors

Among the most potent hit compounds were 2-amino-4-arylthiazoles and benzoxazepines. Of these classes, benzoxazepines had undesirable physicochemical properties including very low polarity (Fig. 4f, clogP = 6.33) along with poor solubility. Therefore, we abandoned this series after testing a limited set of analogs (Supplementary Table 2).

We then selected the 4-arylthiazole-2-amine series for thorough exploration of structure-activity relationships, also due to the synthetic ease of access (Fig. 5). We found the 2-pyridyl substitution to be important for binding affinity, as its replacement with other aryl groups led to drastic loss of potency (e.g., compounds **3-7**). A 2-substituted pyrrole could be used with some loss of potency in compound **8**. Omission of the aromatic group by replacement with ethoxycarbonyl resulted in complete loss of activity (compound **9**).

In contrast, a wide variety of substituents were tolerated in the 2-position of the thiazole. Even the unmodified aminothiazole **13** showed a submicromolar IC_{50}. This compound also served as the synthetic starting point for this series and related chemical probes. The amide linkage between the thiazole and the aryl group is dispensable for activity as demonstrated by the alkylamine **17** and the sulfonamide **18**. Replacement of the aromatic amine by guanidine **19** decreased the $IC_{50.}$ The aryl amide could be substituted or replaced with a wide selection of groups, both aromatic and aliphatic rings with minor effects on potency (compounds **21-24**).
Among the most potent compounds were compound **1** and its morpholinosulfamoyl analog **2**. The arylsulfonamide could be replaced with other groups with minimal loss of potency (compounds **25-28**), whereby larger substituents as in compounds **25** and **28** increased selectivity for SMNDC1 over SMN.
The five-membered heterocycle in the core scaffold could be replaced with the isomeric scaffold 2-(pyridin-2-yl)thiazol-4-amine in compound **14**. The third possible isomer, compound **15**, had significant loss of activity and preferentially inhibited SMN over SMNDC1. When the thiazole was replaced with an analogous oxazole in the compound **16**, there was a 40-fold drop in potency. Replacement of the thiazole with 1,2,4-thiadiazoles resulted in inactive compounds. Substitution of the 5-position of the thiazole of **1** with a methyl group was tolerated without loss of potency (compound **10**) but an ethyl group or a bromine atom decreased the IC_{50} threefold (compounds **11** and **12**).
These extensive structure-activity relationships revealed features that are absolutely essential for the binding of this scaffold to Tudor domains and indicate for substructures required for achieving selectivity between SMNDC1 and SMN. In the following biological characterization, we focus on compound **1** as a potent Tudor domain inhibitor, validate findings with the SMNDC1-specific compound **28** and use the inactive compound **9** as a negative control.

2-amino-4-arylthiazoles bind the methyl-arginine pocket of the SMNDC1 Tudor domain
To prove specific binding of SMNDC1 inhibitors to the aromatic cage of SMNDC1 and to obtain structural information on the binding modes, we applied nuclear magnetic resonance spectroscopy (NMR). The Tudor domains of SMN (residues 84-147) and SMNDC1 (residues 65-128) were expressed in isotope-enriched medium and purified as described elsewhere[7]. Since organic solvents such as dimethyl sulfoxide showed non-specific binding to the Tudor domains and led to interference in the NMR experiments, we attempted to use an aqueous, buffered solution of compound **1**, which however exhibited insufficient solubility for NMR experiments. We therefore prepared an aqueous, buffered solution of the monobasic phosphate salt of compound **13** and assessed its concentration by comparing signal intensities to a DSS standard. NMR titrations of compound **13** with SMN_{84-147} and $SMNDC_{165-128}$ showed significant chemical shift perturbations (CSP) with binding kinetics reflecting fast-exchange

(gradual change of chemical shift with increasing ligand concentration[51]) for both proteins (Fig. 6a, c). CSP are highly sensitive to changes of the local chemical environment of the observed nuclear spin and therefore excellent reporters to map binding sites of a ligand and (potentially associated) conformational changes. The largest CSP are observed for the amino acids forming the aromatic cage (W83, Y90, F108, Y111) and the surrounding residues. Additionally, some parts of the β_2-strand show significant chemical shifts with increasing concentration of compound **13**. The affected residues and CSP match very well the ones published for sDMA binding[7], with exception of residues W83 and S84, suggesting a different interaction with the aromatic cage's tryptophan, as well as N113.

In order to obtain higher resolution structural information of the recognition of compound **13** by the Tudor domain we recorded ^{13}C-filtered NOESY experiments using a 1 mM ^{15}N,^{13}C-labeled SMNDC1 Tudor domain with a 20-fold excess of compound **13**. We could observe a number of contacts between the Tudor domain and the ligand by intermolecular nuclear Overhauser effects (NOE), most prominently with aromatic protons of the Tudor binding site identified by the CSP (Fig. 6b, d, Supplementary Table 3). Using the intermolecular NOEs we calculated a rigid model docking calculation using HADDOCK[52,53], which yielded one cluster with low structural deviation (Supplementary Table 4). The structure indicates that the ring nitrogens of compound **13** are in a *cis* conformation in the complex. The pyridine moiety stacks inside the aromatic cage and its aromatic protons (H2-H5) are showing multiple contacts with the protein's aromatic residues, while the thiazole proton has considerably less contacts to the protein (Fig. 6d, Supplementary Table 3, Supplementary Fig. 4a). The pyridine of compound **13** forms tight π-π stacking contacts with the aromatic rings of F83 and Y111 with distances of 3.7 Å to each, which underlines the importance of an aromatic substituent at the thiazole 4-position. The aromatic moieties of Y90 and F108 stand perpendicular to the pyridine ring while the sidechain of N113 is enclosing it from the opposite site. Overall, the structure shows high similarity to SMNDC1/sDMA (PDB: 4A4H) (Supplementary Fig. 4c) and is fully consistent with the predicted binding mode.

SMNDC1 Tudor domain inhibitors impact protein localization and splicing

We then went on to analyze the effects of the identified small molecule binders on SMNDC1's phase separation. Using the endogenously tagged cell lines, we observed strong effects on the levels and distribution of SMNDC1. Treating the cells with 50 µM of compound **1** for 12-16h leads to a loss of SMNDC1 within the nucleus (Fig. 7a, quantification Fig. 7b,). Additionally, the subnuclear distribution changed and less spots were detected within the nucleus (quantification Fig. 7c). These effects were not observed with the negative control compound **9** which lacks the 2-pyridyl crucial for the binding to SMNDC1. Co-staining nuclei with Hoechst showed that nuclear structure was not affected and that these cells were in

interphase (Supplementary Fig. 5a). Longer treatment with compound **1** resulted in cell death. While the percentage of AnnexinV and Propidium Iodide (PI) positive cells was elevated, the majority of cells was not apoptotic or undergoing other forms of cell death at this timepoint (Supplementary Fig. 5b), and cell death was only observed at later time points.

Using the cell-line in which SMNDC1 and SRRM2 are both tagged we examined the effects of compound **1** on nuclear speckles. Upon treatment with inhibitor **1**, but not compound **9**, SRRM2 and therefore general organization of nuclear speckles was also affected. The overall SRRM2 intensity upon treatment was slightly reduced, and spots appeared to dissolve into the nucleoplasm (Fig. 7d, quantifications Fig. 7e and Supplementary Fig. 5c). Treating several independent SRRM2-RFP clones replicated the results for inhibitor **1** (Supplementary Fig 5f). These results could also be confirmed using antibodies against SMNDC1 and SC35 in IF (Supplementary Fig. 5f).

To check whether inhibition of SMNDC1 is indeed responsible for the effects on nuclear speckles, we silenced SMNDC1 with and without the inhibitor (Fig. 7f, images Supplementary Fig. 5g). The knock-down of SMNDC1 also led to a reduction of SRRM2 intensity and even more pronounced to a reduction of SRRM2 spots in the nucleus, confirming the importance of SMNDC1 for the integrity of nuclear speckles. Treatment with the inhibitor could not further increase these effects, hinting that it is not unspecific effects of the inhibitor that cause the disruption of nuclear speckles. Furthermore, we tested the SMNDC1-selective compound **28** for its effects on SMNDC1 and SRRM2 localization and could confirm the effects observed for the non-selective compound **1** (Supplementary Fig. 5h), even at lower concentrations (Supplementary Fig. 5i).

To directly test the effect on SMN with its similar Tudor domain, we created cell lines in which SMN1 was endogenously tagged with RFP. Treating these cells with 50 µM of compound **1** for 16h showed effects on SMN. Overall intensity of SMN decreased while number of spots (supposedly stress granules) in the cytoplasm increased (Supplementary Fig. 5j).

Next, we analyzed the effect of compound **1** on the proximity interactome of SMNDC1. Overall, we observed that upon inhibitor treatment, more proteins showed a reduced interaction (volcano plot skewed towards down-regulated side, many more significantly down-regulated than up-regulated proteins, Fig. 7g). This indicates that the inhibitor blocks SMNDC1's function to bind to its interaction partners. Compared to APEX2-SMNDC1^{FL}, the inhibitor effects in APEX2-SMNDC1^{TD} were diminished, presumably due to less specific interactions in the truncated form at baseline (Supplementary Fig. 6a). 126 proteins were significantly depleted in APEX2-SMNDC1^{FL} treated with inhibitor vs. none in APEX2-SMNDC1^{TD} (adjusted p-value <0.05, log2 fold-change ≤-2). Proteins with known sDMA modifications identified in SMNDC1's interactome were among the most depleted upon inhibitor treatment (Fig. 7h). The same is true for proteins with assigned localization to nuclear speckles identified in the dataset,

including SRRM2 and the other main nuclear speckle organizer SON[45] (Supplementary Fig. 6b), and for proteins identified by SRSF7-APEX2[48] (Supplementary Fig. 6c). To confirm the observed effects of the inhibitor on the interactome we performed a western blot analysis of the biotin-labeled proteins after pull-down (Supplementary Fig. 6d). Also using this orthogonal technique, we see a general loss of labeled interactors upon treatment with the inhibitor, only with APEX2-SMNDC1^{FL} but not with APEX2-SMNDC1^{TD}. Furthermore, we can confirm the loss of interactions to specific proteins, e.g., the sDMA-modified splicing factor SFPQ or the loss of trans-interactions to SMNDC1 itself. Interactions to SMNDC1 itself are lost both to the endogenous protein (30 kDa band with antibody against SMNDC1) and APEX2-Fusion protein (60 kDa band with antibodies against SMNDC1 and APEX2).

As we did not observe instantaneous effects of the inhibitor on the architecture of nuclear speckles by quantifying intensity or spots per nucleus, we tested whether the inhibitor immediately influenced mobility of proteins within nuclear speckles. To this end, we applied the inhibitor **1** to live cells at a concentration of 50 µM and measured FRAP within a timeframe of 15- 45 min (Fig. 7i). Indeed, we detected a lower recovery after photobleaching for both SMNDC1 and SRRM2 when cells were treated with compound **1** while reference regions were not affected (Supplementary Fig. 5h).

To test specific cellular effects of the SMNDC1 inhibitor **1** we performed paired-end RNA sequencing and analyzed alternative splicing events using Vertebrate Alternative Splicing and Transcription Tools (VAST-TOOLS)[54]. We observed that inhibition of SMNDC1 with compound **1** led to an increased retention of introns and skipping of exons, very similar to the effects of SMNDC1 knock-down[9] (Fig. 7j). Directly comparing the differential percentage spliced-in (dPSI) values for alternative splicing events we found a significant correlation between knock-down and small-molecule inhibition of SMNDC1 (Fig. 7k). We went on to test a panel of individual events with the biggest dPSI values in both knock-down and compound **1** treatment or known to be differentially spliced upon knock-down of SMNDC1[9] (3 selected events Fig. 7l and the full panel Supplementary Fig 6f). For the majority of events, we confirmed the expected effect of the inhibition of SMNDC1 leading to the appearance of alternative spliced isoforms, comparable to SMNDC1 knock-down. Interestingly, we could also see the effect with a long-term, low-dose treatment (2 µM over 5 days) which is more comparable to the 5 days knock-down. Combining knock-down and inhibitor treatments again did not show synergistic effects.

Overall, we demonstrated show specific effects of the inhibitor on the splicing function and the localization of SMNDC1 to nuclear speckles and its proximity to interaction partners, and to the architecture of nuclear speckles in general.

Discussion

Previous work has shown that the interaction of the Tudor domain of SMNDC1's paralog SMN with dimethylarginine causes biomolecular condensation[11]. Here, using endogenous fluorescent tags and *in vitro* assays, we show that also SMNDC1 undergoes phase separation. We find that SMNDC1 localizes to phase-separated membraneless organelles within the nucleus, partially overlapping with nuclear speckles. Consistent with previous findings[11] and in contrast to SMN, the SMNDC1 Tudor domain alone is not driving this condensation behavior. Rather, the protein's C-terminal IDR is sufficient for droplet formation *in vitro*. An RNA-binding prediction algorithm, RNAbindRplus[55], suggests that the C-terminus, especially residues 177-201, interacts with RNA (Supplementary Fig. 7). We therefore hypothesize that the C-terminal IDR is binding to RNAs which in turn recruit further proteins. This model is consistent with our earlier observation[9] that the majority of SMNDC1 protein interactions are lost when RNA is hydrolyzed. Likely these RNA-mediated interactors together with arginine methyl interactions mediated by the SMNDC1 Tudor domain constitute the multivalent binding platform that is a typical prerequisite in the formation of biomolecular condensates. Only the combination of the C-terminal IDR and the Tudor domain (and the nuclear localization signal) is sufficient for the correct localization of SMNDC1 and the full spectrum of interactions to both proteins and RNA. Picking apart the individual contributions of the different parts of the proteins is challenging, as the C-terminus also harbors the NLS responsible for correct organelle localization, with the regions flanking the NLS particularly disordered (Fig. 1a). From our data it is not obvious whether SMNDC1 can form nuclear droplets on its own *in cellulo* as a scaffold or which proteins are required for SMNDC1 to localize to pre-formed membraneless organelles as a client. However, we do observe the dissolution of nuclear speckles upon SMNDC1 inhibition or knock-down. While other factors like SON and SRRM2 are known to be important for the formation of nuclear speckles[45], these data hint at an important structural role of SMNDC1 in these membraneless organelles.

To generate chemical tools for the dose- and time-dependent study of SMNDC1 function, we focused on the protein's Tudor domain. In contrast to the C-terminal region, the Tudor domain exhibits a well-defined structure with a characteristic aromatic cage that mediates specific recognition of dimethylarginine ligands. This feature potentially enables small molecule binding often referred to as druggability. We thus set out to identify small molecule inhibitors of SMNDC1's Tudor domain using an AlphaScreen set-up. Some of the hit structures from our 90,000 compound library showed selectivity in only binding the SMNDC1 but not the SMN Tudor domain and follow up studies allowed us to derive structure-activity relationships for these compounds. Interestingly, these compounds are also active in cellular assays, although relatively high concentrations are needed. Then, the most promising inhibitor led to a loss of SMNDC1 from the nucleus and nuclear speckles and diminished SMNDC1's interaction with

its partners. The kinetics we observe provide a first hint how inhibiting SMNDC1's Tudor domain might influence the architecture of nuclear speckles. While it does not immediately disrupt existing nuclear speckles, mobility and potentially inclusion of new proteins into the phase-separated compartment might be affected, leading to disruption over time. These data suggest that in cells the inhibition of the Tudor domain mediated interactions with its dimethylarginine binding partners drastically affects the protein function even with an intact C-terminal region. This perturbation then results in global splicing changes, consistent with the canonical function of the protein.

We identify the first specific inhibitors of SMNDC1's Tudor domain influencing SMNDC1's phase-separation behavior and splicing and architecture of nuclear speckles. These compounds are chemically distinct from inhibitors previously described for other Tudor domain proteins[56,57], namely TP53B1[58,59], Spindlin-1[60–63], UHRF1[64–66], TDRD3[67], KDM4A[68], SETDB1[69], PHF1[70], and SMN[34]. Our structure-activity relationships indicate that it is feasible to develop these compounds further to achieve specificity for SMNDC1 compared to its closest paralog SMN. Since the other ~34 human Tudor domain proteins in humans[71,72] are less conserved, we expect even lower affinities, but it will be important to conduct unbiased analyses of potential off-targets and their contribution to cellular phenotypes.

Overall, our findings will enable further studies to improve the potency and specificity of the compounds, and more deeply investigate further potential off-targets including other Tudor domain proteins beyond SMN. With more potent and specific compounds, the effect on cells and in vivo could be explored better and disentangled from unspecific toxic effects. Additionally, these compounds might be further derivatized to develop other classes of pharmacological SMNDC1 modulators and *in vivo* active compounds for potential therapeutic development.

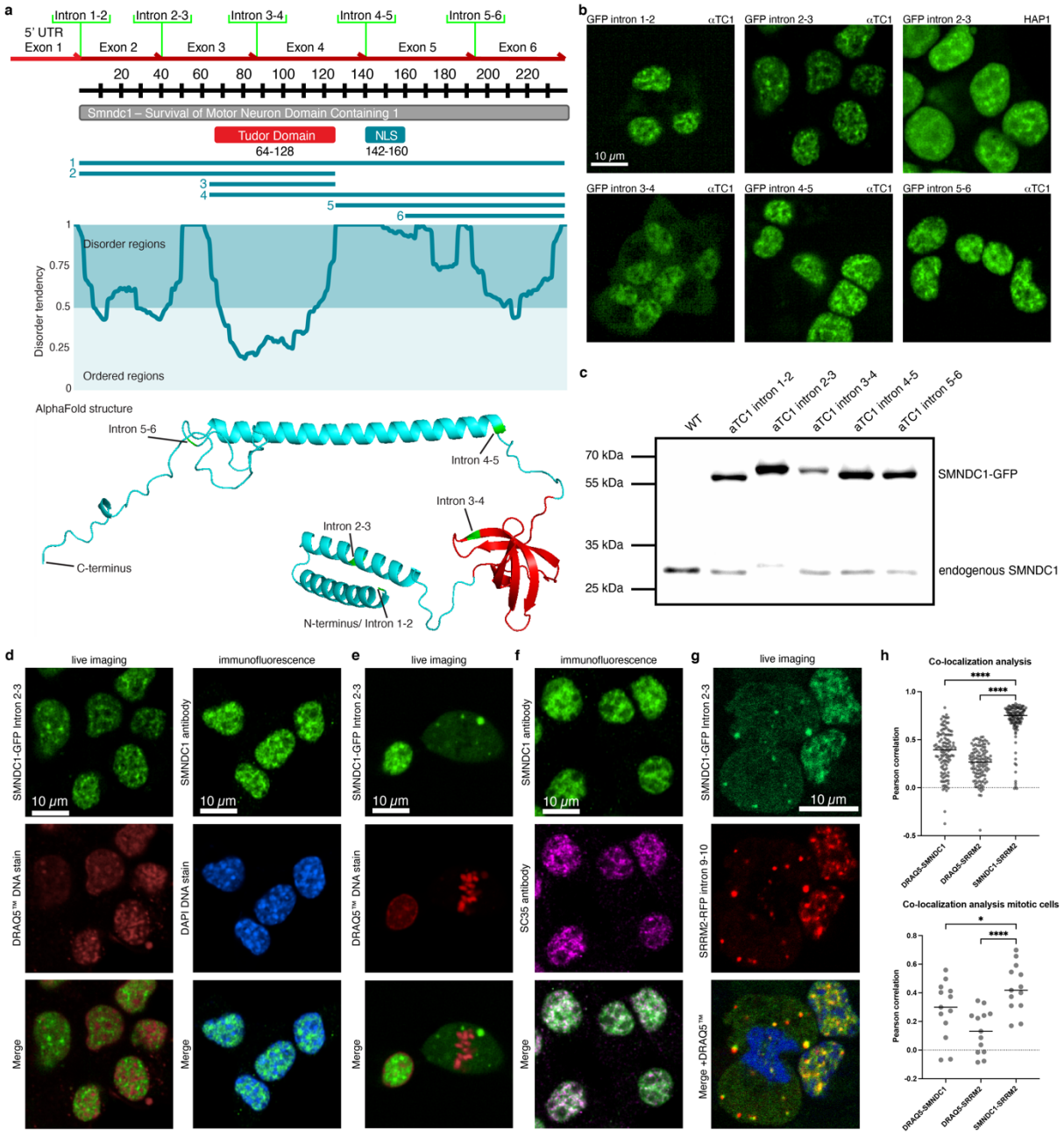


Fig. 1 | SMNDC1 co-localizes with nuclear speckle markers. a, Overview of SMNDC1's structure with numbered truncations (for Fig. 2f), intrinsic disorder prediction plot (MetaDisorder[36]), AlphaFold structure prediction with Tudor domain marked in red and positions of GFP intron-tags in green. **b**, Live images of clonal cell lines (αTC1 and HAP1) with the endogenous GFP-tag in green. **c**, Immunoblots showing expression of WT SMNDC1 and SMNDC1-GFP fusion proteins in clonal cell lines with GFP-tag in different introns. **d**, Live (SMNDC1-GFP intron 2-3, αTC1) and immunofluorescence images (αTC1 WT) with nuclear staining (DRAQ5™ in red/ DAPI in blue). **e**, Live imaging (SMNDC1-GFP intron 2-3, αTC1) with DRAQ5™ nuclear staining showing a cell during M-phase. **f**, Immunofluorescence images (αTC1 WT) with SMNDC1 (green) and SC35 antibody (magenta), overlap of green and magenta is white. DAPI nuclear staining blue. **g**, Live imaging (SMNDC1-GFP intron 2-3,

SRRM2-RFP intron 9-10, αTC1) with DRAQ5™ nuclear staining (blue) showing a cell during telophase. **h**, Co-localization analyses of interphase and mitotic cells, Pearson correlation between different channels of maximum intensity projections of z-stack images. Data shown as scatter plot + median, analyzed by unpaired t-test.

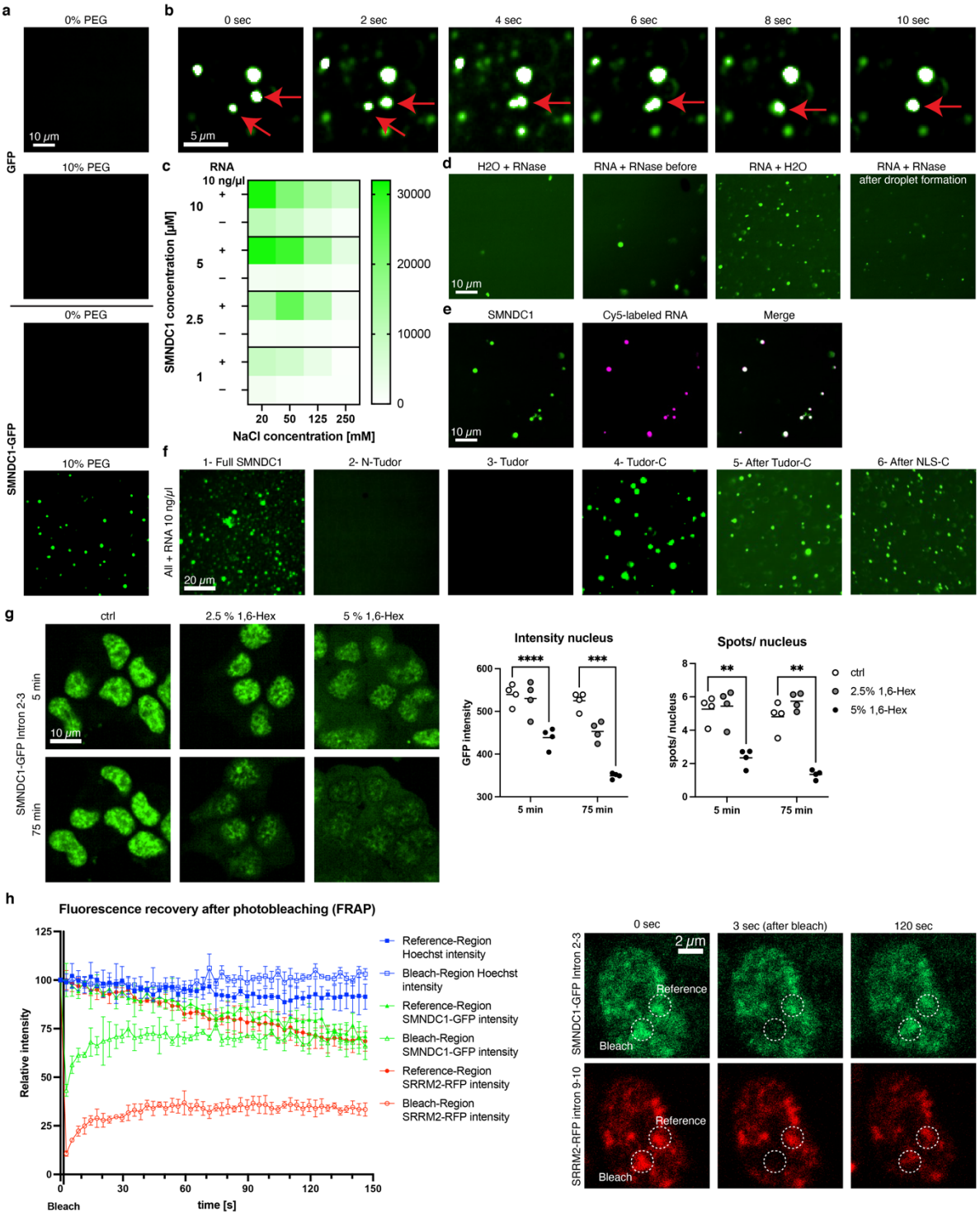


Fig. 2 | SMNDC1 shows biomolecular condensation *in vitro* and in cellular systems. **a,** *In vitro* droplet formation assay with 10 µM GFP or SMNDC1-GFP fusion protein +/- 10% PEG-8000. **b**, *In vitro* droplet formation assay of SMNDC1-GFP over time with droplet fusion event, marked by red arrows. **c**, *In vitro* droplet formation assay of SMNDC1-GFP with quantified number of droplets with different protein and NaCl concentrations, +/- 10 ng/µl RNA. **d**, *In vitro* droplet formation assay of SMNDC1-GFP with the addition of 10 ng/µl total cellular RNA and RNase. **e**, *In vitro* droplet formation assay of SMNDC1-GFP (green) with 100 ng/µl

Cy5-labeled RNA (magenta), overlap white. **f**, *In vitro* droplet formation assay of different truncations of SMNDC1-GFP + 10 ng/µl total cellular RNA. **g**, Live imaging (SMNDC1-GFP intron 2-3, αTC1), cells were treated with 2.5% or 5 % 1,6-hexanediol. Quantifications of GFP intensity and GFP spots/nucleus in 4 different clonal cell lines. Data are shown as mean with individual values, analyzed by multiple paired t-tests with False Discovery Rate q calculated by Two-stage step-up[73]. **h**, Fluorescence recovery after photobleaching (FRAP) experiment in SMNDC1-GFP intron 2-3, SRRM2-RFP intron 9-10, αTC1-cells. Left: Relative intensity of Hoechst (blue), SMNDC1-GFP (green), and SRRM2-RFP (red) in reference (filled symbols) and bleach region (empty symbols) over time. Data plotted as mean with standard deviation, n=3. Right: representative images of nucleus with marked reference and bleach region at 3 different time points, 0 sec (before bleaching), 3 sec (directly after bleaching), and 120 sec (after recovery).

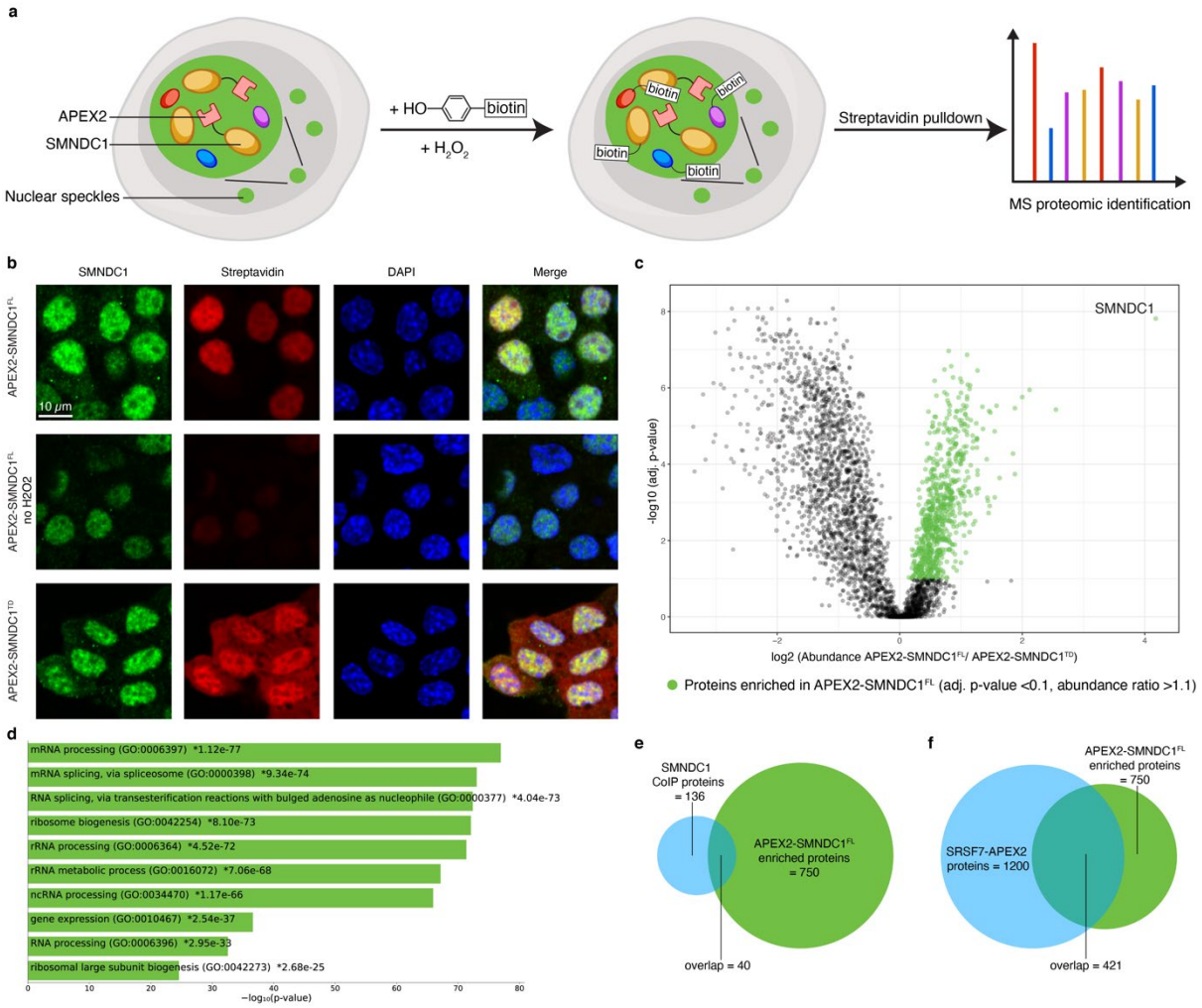


Fig. 3 | Characterization of SMNDC1's interactome by proximity labeling. a, Scheme of proximity labeling by APEX2 fusion proteins followed by mass spectrometry-based proteomics. **b**, Immunofluorescence images (αTC1 WT) with staining against SMNDC1 (green), Biotin via Streptavidin (red) and nuclear staining DAPI (blue). **c**, Volcano plot showing log2 abundance against -log10 adjusted p-value of APEX2-SMNDC1^{FL} versus APEX2-SMNDC1^{TD} biotinylated and enriched proteins. Highlighted dots in green indicate 750 enriched proteins (adjusted p-value <0.1, abundance ratio >1.1). **d**, Enrichr analysis of APEX2-SMNDC1^{FL} enriched proteins, top 10 terms sorted by p-value. **e**, Venn diagram showing the overlap of proteins identified by SMNDC1-CoIP (light blue), and APEX2-SMNDC1^{FL} enriched (green). **f**, Venn diagram showing the overlap of proteins identified by SRSF7-APEX2 (light blue), and APEX2-SMNDC1^{FL} enriched (green).

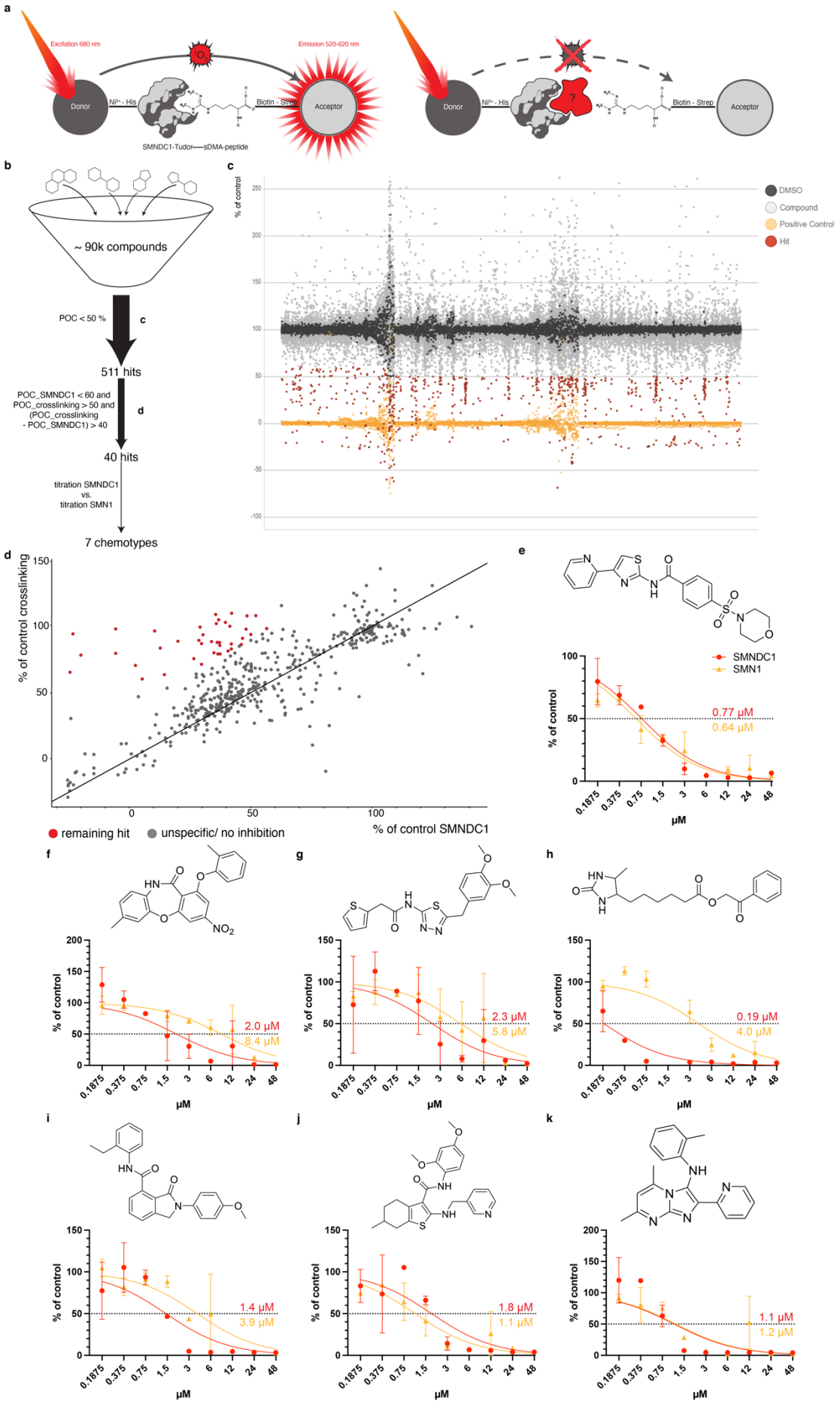

a
Excitation 680 nm
Emission 520-620 nm
Donor
Acceptor
Biotin - Strep
SMNDC1-Tudor—sDMA-peptide
b
~ 90k compounds
POC < 50 %
c
511 hits
POC_SMNDC1 < 60 and
POC_crosslinking > 50 and
(POC_crosslinking
- POC_SMNDC1) > 40
d
40 hits
titration SMNDC1
vs.
titration SMN1
7 chemotypes
c
% of control
DMSO
Compound
Positive Control
Hit
d
% of control crosslinking
% of control SMNDC1
remaining hit
unspecific/ no inhibition
e
SMNDC1
SMN1
0.77 µM
0.64 µM
µM
f
2.0 µM
8.4 µM
g
2.3 µM
5.8 µM
h
0.19 µM
4.0 µM
i
1.4 µM
3.9 µM
j
1.8 µM
1.1 µM
k
1.1 µM
1.2 µM

Fig. 4 | Identification of an inhibitor against SMNDC1's Tudor domain. a, Scheme of AlphaScreen set-up. **b**, Screening strategy starting with ~90,000 compound library. **c**, Overview of AlphaScreen of full ~90,000 compound library (light grey) with DMSO (dark grey), positive control (quencher, yellow), and compound hits (red). **d**, AlphaScreen percentage of DMSO control with SMNDC1/ sDMA-peptide vs crosslinking peptide. Remaining hits marked in red. **e-k**, Chemical structure and AlphaScreen 9-point compound titration with SMNDC1/ sDMA-peptide (red) vs. SMN/ sDMA-peptide (yellow).

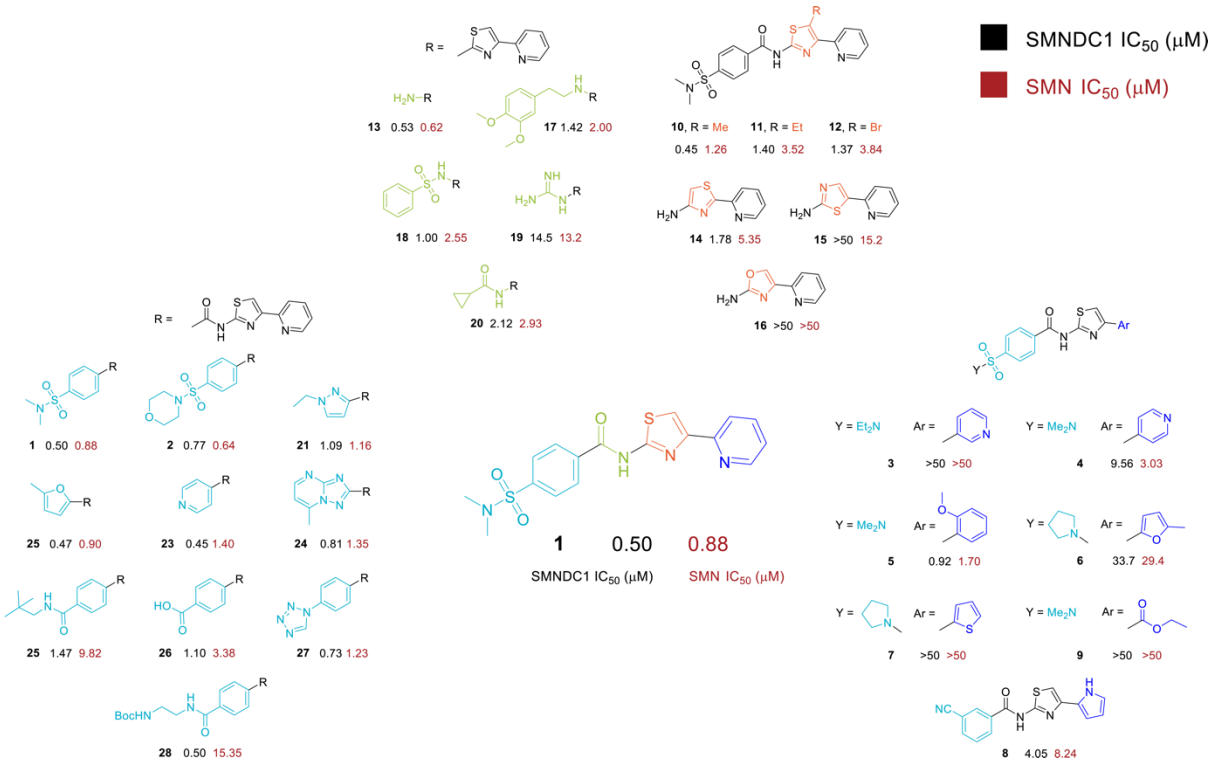


Fig. 5 | Structure-activity relationships of the 2-aryl-4-aminothiazole Tudor domain inhibitors. Color-coded chemical structures illustrating structure activity relationships for compound **1**. Modifications to the 2-aryl moiety are shown in dark blue, replacements of the thiazole group in orange, linker modifications in green and arylamide analogs in turquoise. Underneath are IC_{50} values (μM) for SMNDC1 in black and SMN in dark red.

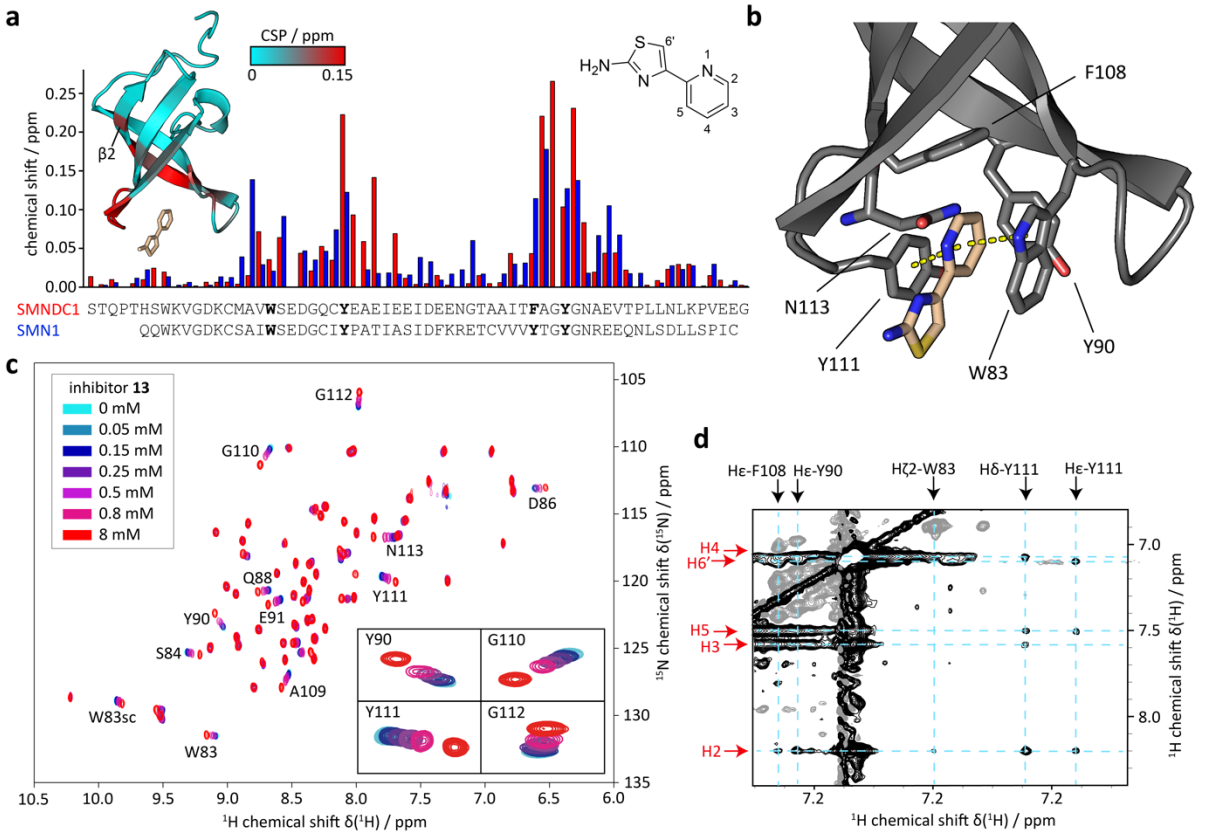


Fig. 6 | Inhibitor binds at the aromatic cage of the Tudor domain. a, Chemical shift perturbations (CSP) of Tudor domains SMNDC1 (red) and SMN (blue) in presence of 0.8 mM compound **13**. Residues forming the aromatic cage are highlighted in bold. Left inlet shows cartoon representation of SMNDC1 in complex with compound **13** (light-brown sticks) calculated using semi-rigid body docking of compound **13** based on 23 intermolecular NOE restraints (see d and Supplementary Table 3). CSP per residue in presence of 8 mM compound **13** are displayed in cyan to red shades. Right inlet shows compound **13** with numbers indicating the assignment used in d and Supplementary Table 3. **b**, Zoomed view of the binding site with residues forming the aromatic cage shown as sticks. Stacking contacts between the thiazole moiety of compound **13** and tryptophane 83 or tyrosine 111 indicated with dashed lines. **c**, Overlay of $^{1}H,^{15}N$-HSQC spectra of SMNDC1 in the presence of 0, 0.05, 0.15, 0.25, 0.5, 0.8, and 8 mM compound **13**. Zoomed view shows residues in and near the aromatic cages. **d**, Section of an in ω_1-^{13}C-filtered NOESY spectrum shows crosspeaks between inhibitor protons (red arrows) and aromatic cage protons (black) arrows. The crossing points of the dashed cyan lines indicate the locations of the intermolecular NOEs.

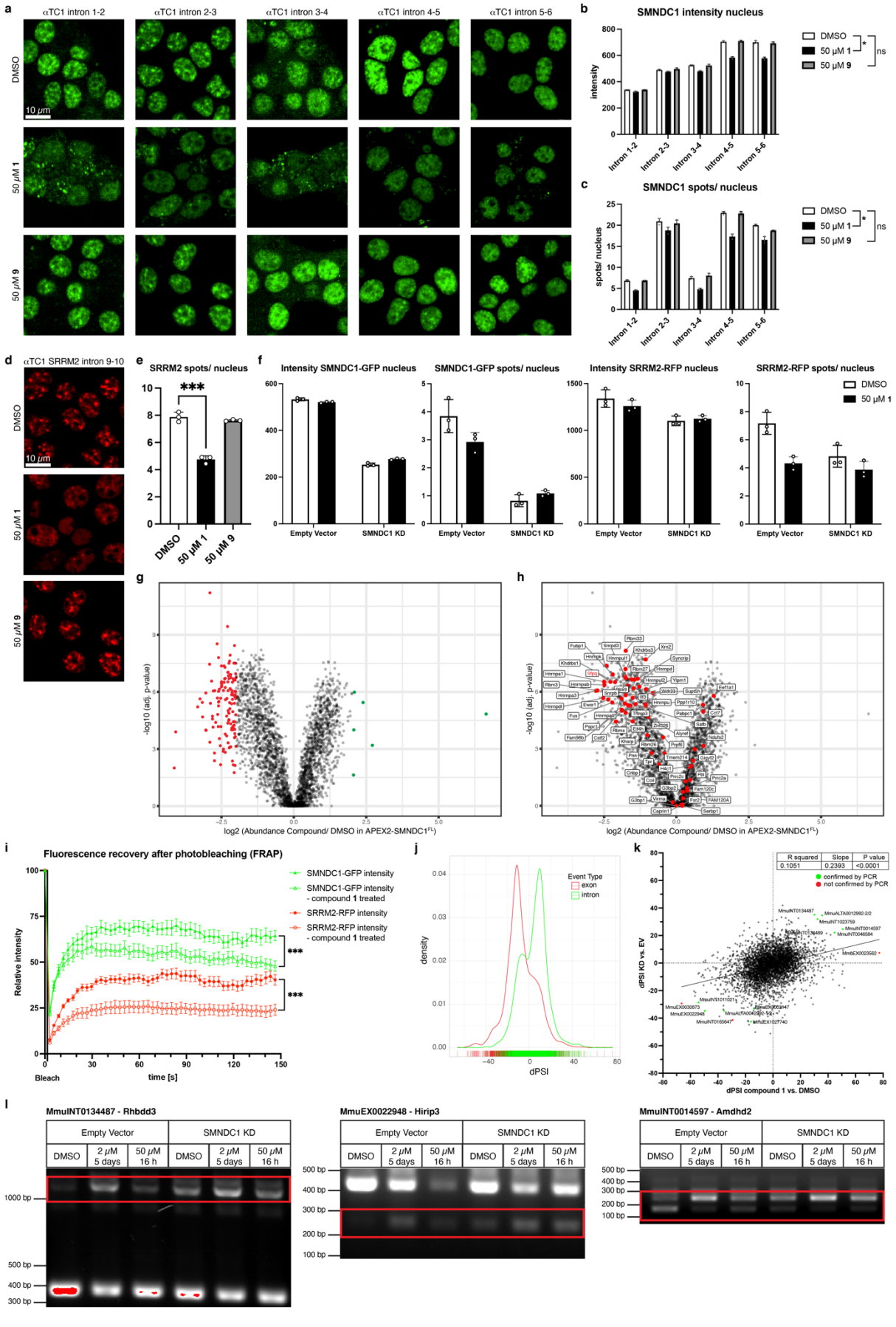
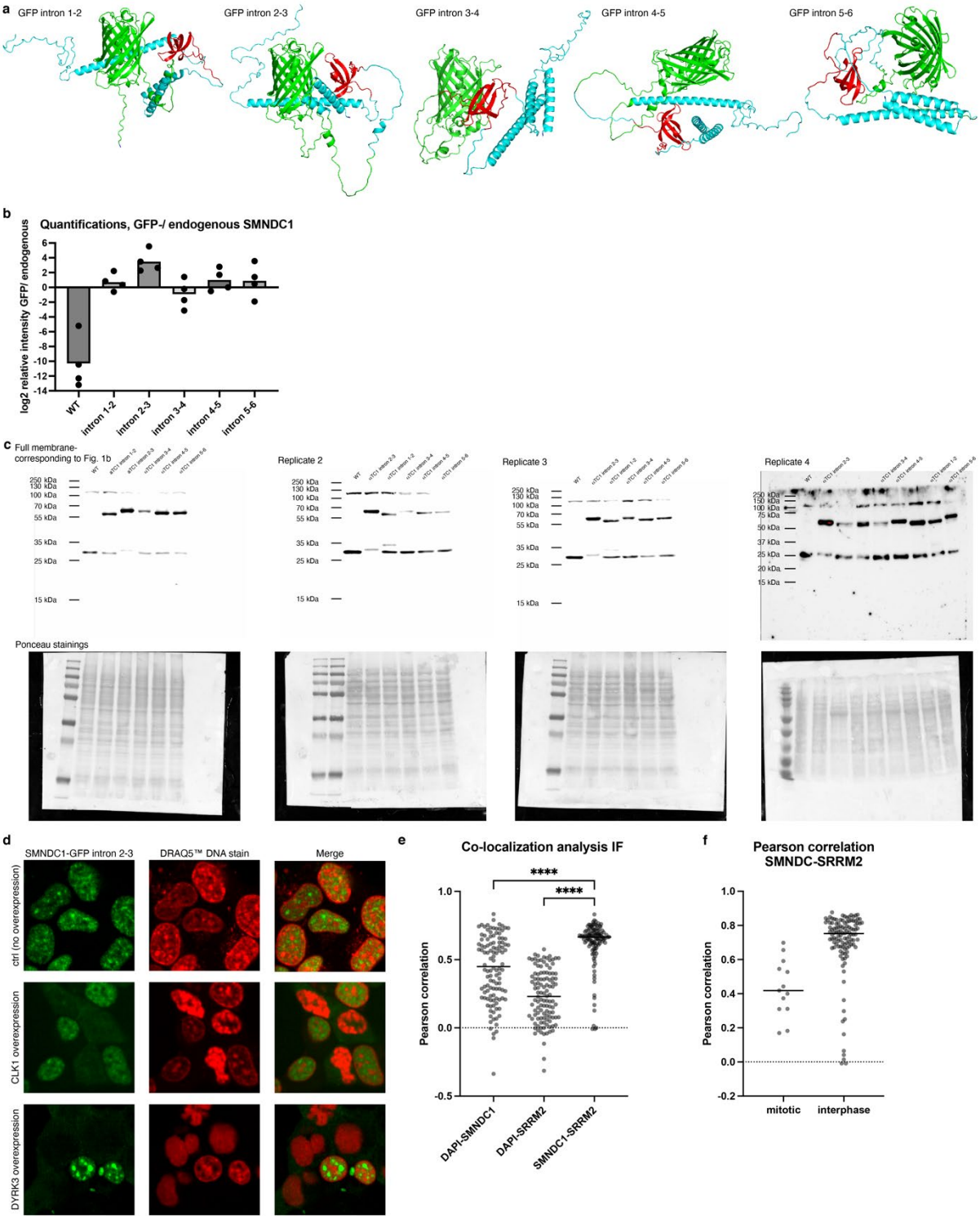
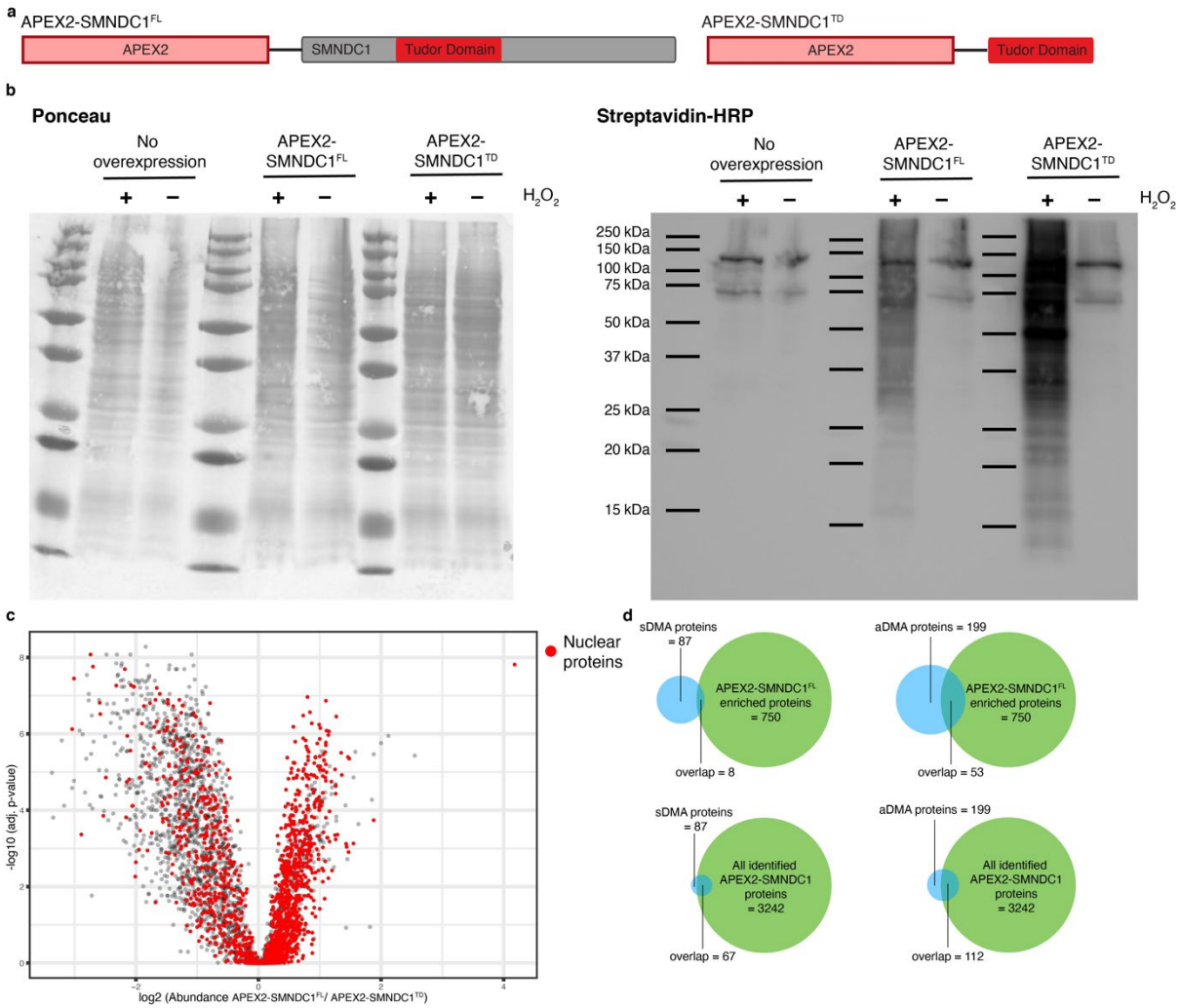

a
αTC1 intron 1-2
αTC1 intron 2-3
αTC1 intron 3-4
αTC1 intron 4-5
αTC1 intron 5-6
DMSO
50 μM 1
50 μM 9
10 μm
b
SMNDC1 intensity nucleus
intensity
DMSO
50 μM 1
50 μM 9
ns
Intron 1-2
Intron 2-3
Intron 3-4
Intron 4-5
Intron 5-6
c
SMNDC1 spots/ nucleus
spots/ nucleus
d
αTC1 SRRM2 intron 9-10
e
SRRM2 spots/ nucleus
***
f
Intensity SMNDC1-GFP nucleus
SMNDC1-GFP spots/ nucleus
Intensity SRRM2-RFP nucleus
SRRM2-RFP spots/ nucleus
Empty Vector
SMNDC1 KD
g
-log10 (adj. p-value)
log2 (Abundance Compound/ DMSO in APEX2-SMNDC1FL)
h
i
Fluorescence recovery after photobleaching (FRAP)
Relative intensity
SMNDC1-GFP intensity
SMNDC1-GFP intensity - compound 1 treated
SRRM2-RFP intensity
SRRM2-RFP intensity - compound 1 treated
Bleach
time [s]
j
Event Type
exon
intron
density
dPSI
k
R squared
Slope
P value
0.1051
0.2393
<0.0001
confirmed by PCR
not confirmed by PCR
dPSI KD vs. EV
dPSI compound 1 vs. DMSO
l
MmuINT0134487 - Rhbdd3
MmuEX0022948 - Hirip3
MmuINT0014597 - Amdhd2
Empty Vector
SMNDC1 KD
DMSO
2 μM 5 days
50 μM 16 h
1000 bp
500 bp
400 bp
300 bp
200 bp
100 bp

Fig. 7 | Cellular effects of SMNDC1 Tudor domain inhibition. a-c, Live imaging and quantifications of different SMNDC1-GFP cell lines (αTC1) treated with DMSO, 50 µM compound 1, or 50 µM compound. Data shown as mean + standard deviation, *n*=3, analyzed by ratio-paired t-test. **d**, Live imaging of cells (SRRM2-RFP intron 9-10, αTC1) treated with DMSO, 50 µM compound **1,** or 50 µM compound **9**. **e**, Quantification of SRRM2-RFP spots/ nucleus in live imaging data of SRRM2-RFP cell line (αTC1) treated with DMSO, 50 µM compound **1,** or 50 µM compound **9**. Data shown as mean + standard deviation, *n*=3, analyzed by unpaired t-test, n=3. **f**, Quantification of nuclear SMNDC1-GFP, SMNDC1-GFP spots/ nucleus, nuclear SRRM2-RFP, and SRRM2-RFP spots/ nucleus in live imaging data of double-tagged SMNDC1-GFP (intron 2-3) SRRM2-RFP (intron 9-10) cell line (αTC1) treated with DMSO or 50 µM compound **1** and transduced with Empty Vector or SMNDC1 knock-down (KD) plasmids. Data shown as mean + standard deviation, *n*=3. **g**, Volcano plot showing log2 protein abundance against -log10 adjusted p-value of compound **1** treated cells over DMSO control after APEX2-SMNDC1^{FL} proximity labeling and biotin enrichment. 126 proteins significantly depleted (red) vs. 6 proteins significantly enriched (green), adjusted p-value <0.05, |log2FC| ≥ 2. **h**, Same as in g. Proteins with known sDMA modification marked in red and named. **i,** Fluorescence recovery after photobleaching (FRAP) experiment in SMNDC1-GFP intron 2-3, SRRM2-RFP intron 9-10, αTC1-cells, treated with DMSO (filled symbols) or 50 µM compound **1** (empty symbols). Relative intensity of Hoechst SMNDC1-GFP (green), and SRRM2-RFP (red) in bleach region over time (Reference region: see Supplementary Fig. 6e). Data plotted as mean with standard error of the mean, n ≥ 11. Last time point analyzed by unpaired t-test. **j**, Splicing analysis of RNA-sequencing data, plotted are all alternative splicing events and their density with their differential percentage spliced-in (dPSI) value of compound **1** over DMSO treatment. Exon events are colored red, while intron events are colored green. **k**, Splicing analysis of RNA-sequencing data, plotted are all overlapping alternative splicing events between compound 1 over DMSO treatment (x-axis) and SMNDC1 knockdown (KD) over empty vector (EV) (y-axis) with their respective dPSI-values. Results from a simple linear regression analysis. Events confirmed via PCR in green, not confirmed events in red. **l**, DNA-bands on agarose gel after reverse transcription and PCR amplification of RNA to confirm alternative splicing events. RNA was isolated from αTC1 cells transfected with empty vector or SMNDC1 knock-down (KD) plasmid and treated with DMSO, 2 µM compound **1** for 5 days, or 50 µM compound **1** for 16 h.

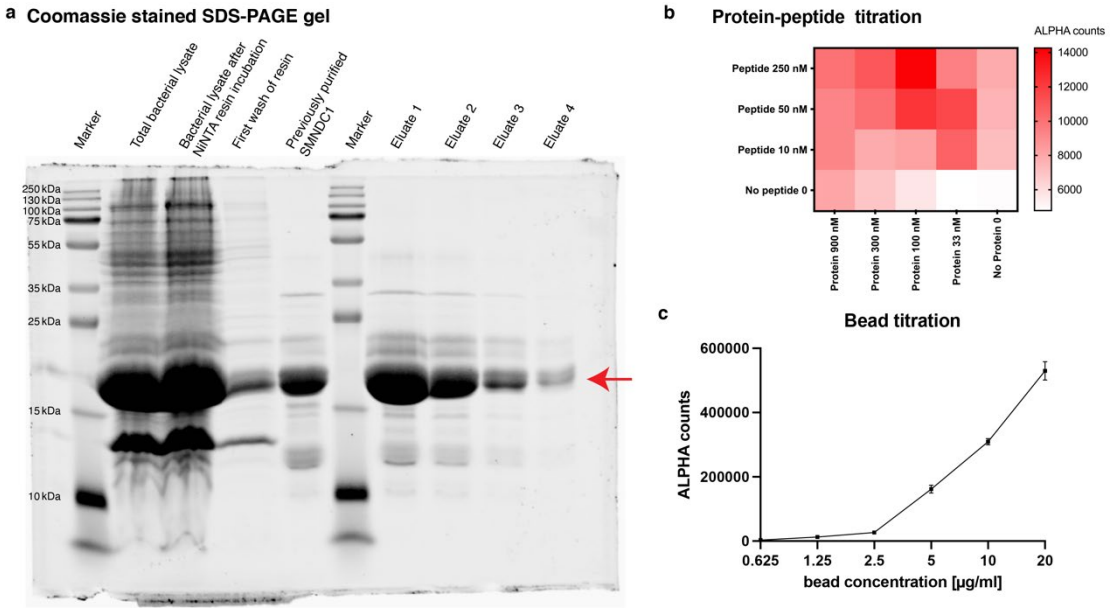


Supplementary Fig. 1 | SMNDC1 co-localizes with nuclear speckle markers. a, AlphaFold structure predictions of SMNDC1 with GFP (marked in green) integrated in different introns. **b**, Quantification of SMNDC1-GFP band divided by endogenous SMNDC1 band in different clonal intron-tag cell lines (αTC1), on a log2 scale. Data shown as mean, *n*=4. **c**, Immunoblots showing expression of WT SMNDC1 and SMNDC1-GFP fusion proteins in clonal cell lines with GFP-tag in different introns, 4 replicates. Ponceau staining to show comparable loading. **d**, Live imaging of SMNDC1-GFP intron 2-3 cell line with DRAQ5™ nuclear staining, control

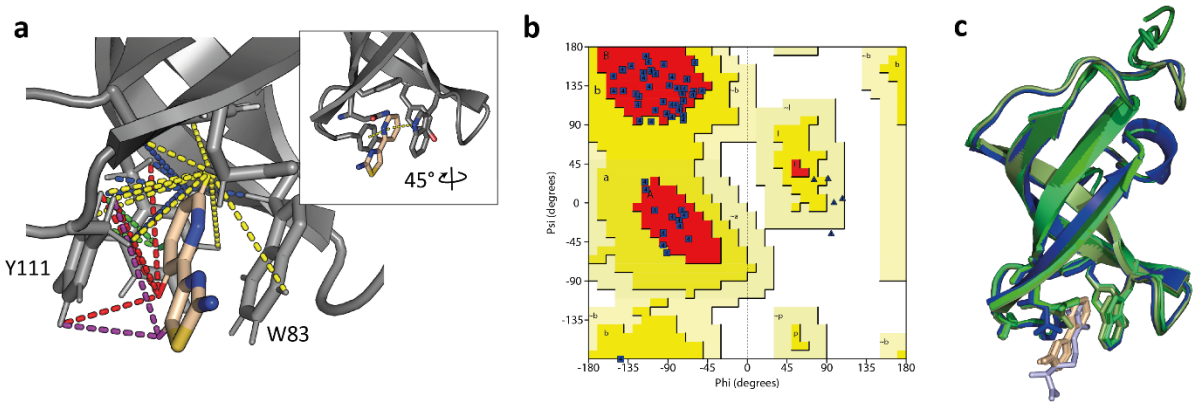
(no overexpression), overexpressing CLK1 or DYRK3. **e**, Co-localization analyses of IF images of αTC1 WT with SMNDC1-antibody, SC35-antibody, and DAPI nuclear staining. Pearson correlation between different channels of maximum intensity projections of z-stack images. Data shown as scatter plot + median, analyzed by unpaired t-test. **f**, Comparison of Pearson correlation values in co-localization analyses of mitotic and interphase cells, live imaging corresponding to Fig. 1g+h. Data shown as scatter plot + median.



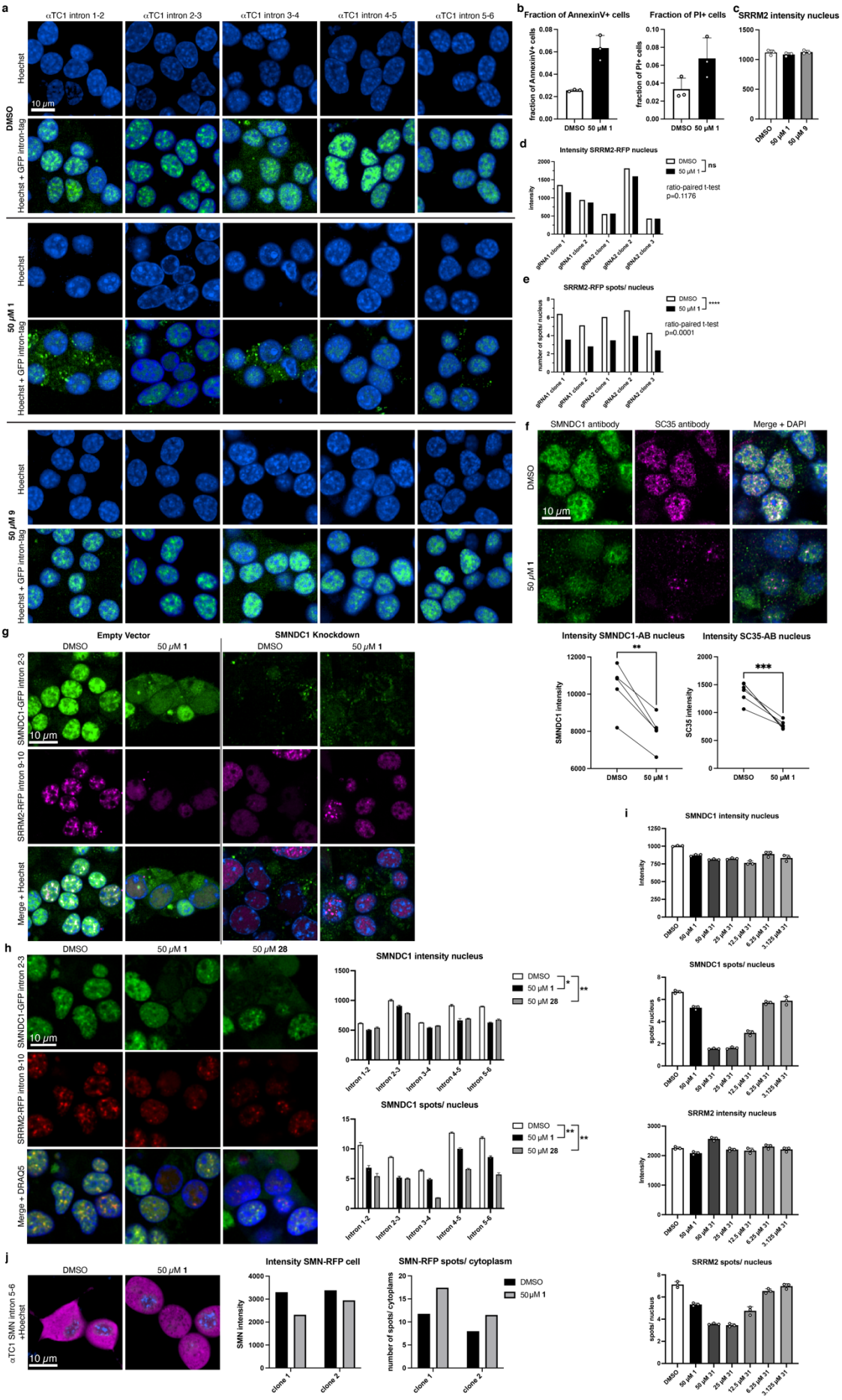
Supplementary Fig. 2 | Characterization of SMNDC1's interactome by proximity labeling. a, Depiction of APEX2-fusion constructs APEX2-SMNDC1^{FL} and APEX2-SMNDC1^{TD}. **b**, Ponceau S staining of all proteins and western blot with Streptavidin-HRP showing all biotinylated proteins. **c**, Volcano plot showing log2 abundance against -log10 adjusted p-value of APEX2-SMNDC1^{FL} versus APEX2-SMNDC1^{TD} biotinylated and enriched proteins. Nuclear proteins highlighted in red. **d**, Venn diagrams showing the overlap of all proteins with a known sDMA modification or aDMA modification (light blue), and APEX2-SMNDC1^{FL} enriched proteins or all proteins identified with SMNDC1 proximity labeling (green).



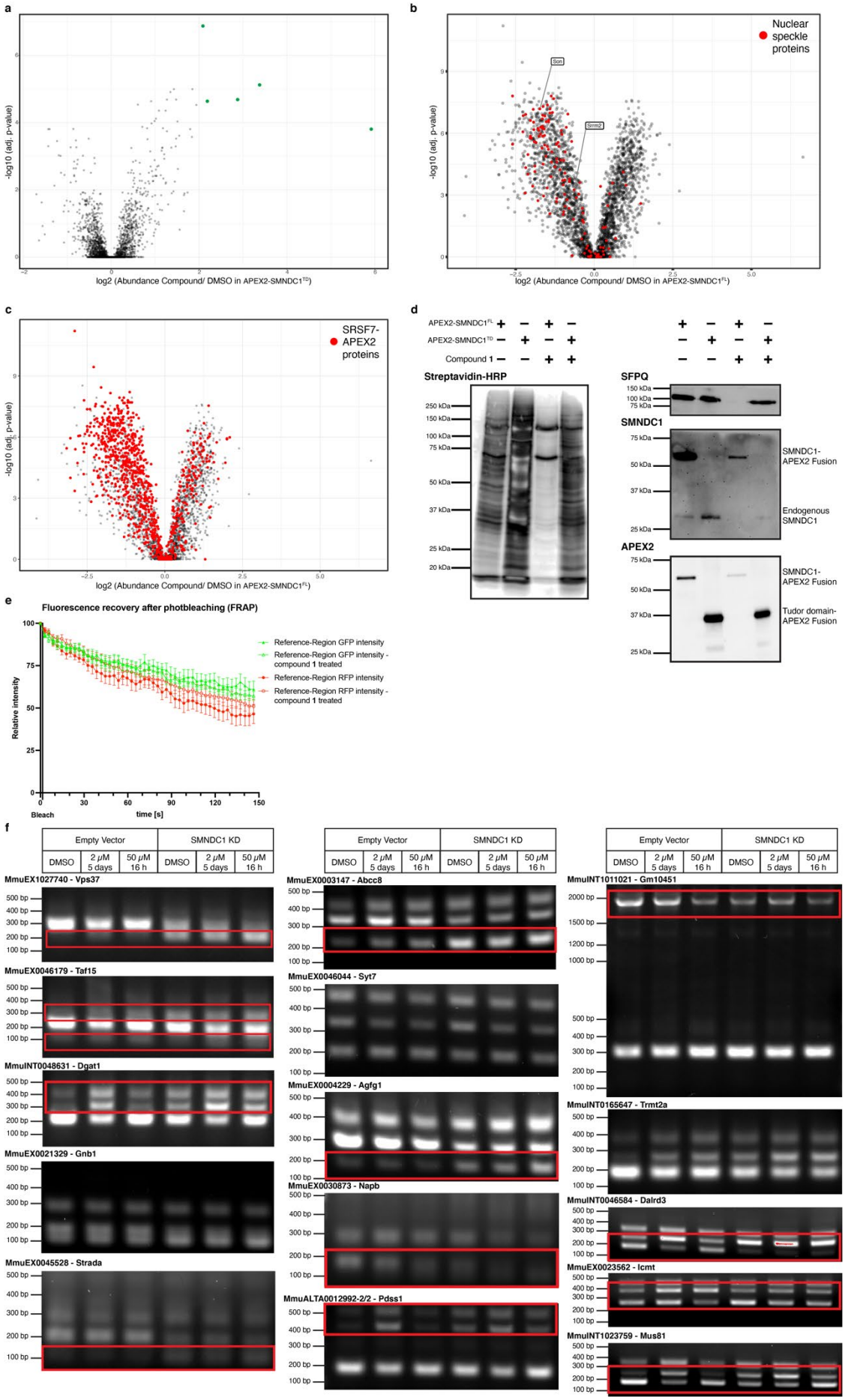
Supplementary Fig. 3 | Establishment of AlphaScreen. a, Coomassie staining of all proteins in different samples along the protein purification process. SMNDC1-Tudor domain marked by red arrow. **b**, Cross-titration of different protein and peptide concentrations in AlphaScreen. **c**, Titration of AlphaScreen acceptor and donor beads with 50 nM peptide, 75 nM protein. Data shown as mean + standard deviation, *n*=3.



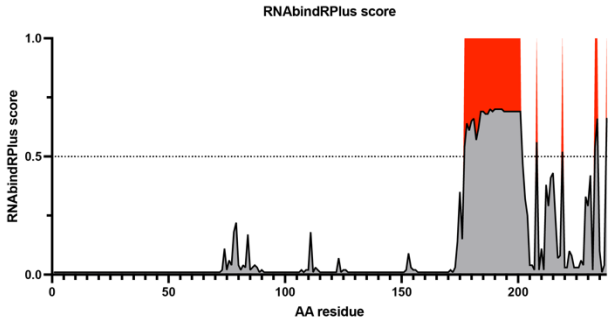
Supplementary Fig. 4 | SMNDC1/inhibitor complex. a, Inter-molecular NOEs (dashed line) used in the restraint-driven docking simulation. NOEs derived from different protons of the inhibitor are colored individually. Phe or Tyr protons HE and HD were not specifically assigned, and a dashed line is only shown for the proton with the closest distance to the NOE contact. The inlet shows the relative orientation of the structure compared to Figure 6. **b**, Ramachandran plot of the four lowest-energy structures, 88.9% and 11.1% of the backbone torsion angles are found in the favored regions and additional allowed regions, respectively. **c**, Overlay of SMNDC1/ compound **13** complex (four lowest-energy structures, shades of green; compound **13**, light brown) and SMNDC1/sDMA (blue/light blue, PDB: 4A4H) shows high structural similarity (backbone rmsd 0.295-0.375 Å).


a
αTC1 intron 1-2
αTC1 intron 2-3
αTC1 intron 3-4
αTC1 intron 4-5
αTC1 intron 5-6
DMSO
50 μM 1
50 μM 9
Hoechst
Hoechst + GFP intron-tag
10 μm
b
Fraction of AnnexinV+ cells
Fraction of PI+ cells
fraction of AnnexinV+ cells
fraction of PI+ cells
DMSO
50 μM 1
c
SRRM2 intensity nucleus
50 μM 9
d
Intensity SRRM2-RFP nucleus
Intensity
gRNA1 clone 1
gRNA1 clone 2
gRNA2 clone 1
gRNA2 clone 2
gRNA2 clone 3
ns
ratio-paired t-test
p=0.1176
e
SRRM2-RFP spots/ nucleus
number of spots/ nucleus
****
ratio-paired t-test
p=0.0001
f
SMNDC1 antibody
SC35 antibody
Merge + DAPI
g
Empty Vector
SMNDC1 Knockdown
SMNDC1-GFP intron 2-3
SRRM2-RFP intron 9-10
Merge + Hoechst
Intensity SMNDC1-AB nucleus
Intensity SC35-AB nucleus
SMNDC1 intensity
SC35 intensity
h
50 μM 28
Merge + DRAQ5
SMNDC1 intensity nucleus
SMNDC1 spots/ nucleus
Intron 1-2
Intron 2-3
Intron 3-4
Intron 4-5
Intron 5-6
i
SMNDC1 intensity nucleus
SMNDC1 spots/ nucleus
SRRM2 intensity nucleus
SRRM2 spots/ nucleus
25 μM 31
12.5 μM 31
6.25 μM 31
3.125 μM 31
j
αTC1 SMN intron 5-6 +Hoechst
Intensity SMN-RFP cell
SMN-RFP spots/ cytoplasm
SMN intensity
number of spots/ cytoplasm
clone 1
clone 2

Supplementary Fig. 5 | Effects of SMNDC1 Tudor domain inhibition on localization. a, Live imaging of cells (SMNDC1-GFP with different introns tagged, αTC1) treated with DMSO, 50 μM compound **1,** or 50 μM compound **9**. Nuclei stained with Hoechst, corresponding to images in Fig. 7a. **b**, Quantification of AnnexinV+ and PI+ cells (αTC1), treated with DMSO or 50 μM compound **1**. Data shown as mean + standard deviation, *n*=3. **c**, Quantification of nuclear SRRM2-RFP in live imaging data of SRRM2-RFP cell line (αTC1) treated with DMSO, 50 μM compound **1,** or 50 μM compound **9**. Data shown as mean + standard deviation, *n*=3. **d**, Quantification of nuclear SRRM2-RFP in live imaging data of different SRRM2-RFP cell lines (αTC1) treated with DMSO, or 50 μM compound **1.** Data analyzed by ratio-paired t-test. **e**, Quantification of SRRM2-RFP spots/ nucleus in live imaging data of different SRRM2-RFP cell lines (αTC1) treated with DMSO, or 50 μM compound **1**. Data analyzed by ratio-paired t-test. **f**, Upper panel: Immunofluorescence images of cells (αTC1 WT) stained with antibodies against SMNDC1 (green), SC25 (magenta), and nuclear marker DAPI (blue), treated with DMSO or 50 μM compound **1**. Lower panel: Quantification of SMNDC1-AB intensity and SC35-AB intensity in the nucleus in immunofluorescence imaging data (αTC1), cells treated with DMSO or 50 μM compound **1**. Data analyzed by ratio-paired t-test. **g**, Representative images of quantifications shown in Fig. 7f. Live imaging data of double-tagged SMNDC1-GFP (intron 2-3) SRRM2-RFP (intron 9-10) cell line (αTC1) treated with DMSO or 50 μM compound **1** and transduced with Empty Vector or SMNDC1 knock-down (KD) plasmids. **h**, Live imaging data of double-tagged SMNDC1-GFP (intron 2-3) SRRM2-RFP (intron 9-10) cell line (αTC1) and multiple SMNDC1-GFP clonal cell lines single-tagged in different introns, treated with DMSO, 50 μM compound **1,** or 50 μM compound **28.** Quantification of nuclear SMNDC1-GFP and SMNDC1-GFP spots/ nucleus. Data shown as mean + standard deviation, *n*=3, analyzed by ratio-paired t-test. **i**, Quantification of live imaging data of double-tagged SMNDC1-GFP (intron 2-3) SRRM2-RFP (intron 9-10) cell line (αTC1) treated with DMSO, 50 μM compound **1,** or 50/ 25/ 12.5/ 6.25/ 3.125 μM compound **28.** Nuclear SMNDC1-GFP, SMNDC1-GFP spots/ nucleus, nuclear SRRM2-RFP, and SRRM2-RFP spots/ nucleus. Data shown as mean + standard deviation, *n*=3. **j**, Left panel: Live imaging of cells (SMN-RFP intron 5-6, αTC1) treated with DMSO or 50 μM compound **1**. Quantification of whole cell SMN-RFP and SMN-RFP spots/ cytoplasm in live imaging data of different SMN-RFP cell lines (αTC1).


a
-log10 (adj. p-value)
log2 (Abundance Compound/ DMSO in APEX2-SMNDC1TD)
b
Nuclear speckle proteins
Son
Srrm2
-log10 (adj. p-value)
log2 (Abundance Compound/ DMSO in APEX2-SMNDC1FL)
c
SRSF7-APEX2 proteins
-log10 (adj. p-value)
log2 (Abundance Compound/ DMSO in APEX2-SMNDC1FL)
d
APEX2-SMNDC1FL
APEX2-SMNDC1TD
Compound 1
Streptavidin-HRP
250 kDa
150 kDa
100 kDa
75 kDa
50 kDa
37 kDa
25 kDa
20 kDa
SFPQ
150 kDa
100 kDa
75 kDa
SMNDC1
75 kDa
50 kDa
37 kDa
25 kDa
SMNDC1-APEX2 Fusion
Endogenous SMNDC1
APEX2
75 kDa
50 kDa
37 kDa
25 kDa
SMNDC1-APEX2 Fusion
Tudor domain-APEX2 Fusion
e
Fluorescence recovery after photbleaching (FRAP)
Relative intensity
time [s]
Bleach
Reference-Region GFP intensity
Reference-Region GFP intensity - compound 1 treated
Reference-Region RFP intensity
Reference-Region RFP intensity - compound 1 treated
f
Empty Vector
SMNDC1 KD
DMSO
2 μM 5 days
50 μM 16 h
MmuEX1027740 - Vps37
MmuEX0046179 - Taf15
MmuINT0048631 - Dgat1
MmuEX0021329 - Gnb1
MmuEX0045528 - Strada
MmuEX0003147 - Abcc8
MmuEX0046044 - Syt7
MmuEX0004229 - Agfg1
MmuEX0030873 - Napb
MmuALTA0012992-2/2 - Pdss1
MmuINT1011021 - Gm10451
MmuINT0165647 - Trmt2a
MmuINT0046584 - Dalrd3
MmuEX0023562 - Icmt
MmuINT1023759 - Mus81

Supplementary Fig. 6 | Effects of SMNDC1 Tudor domain inhibition on interactome and splicing. a, Volcano plot showing log2 protein abundance against -log10 adjusted p-value of compound **1** treated cells over DMSO control after APEX2-SMNDC1^{TD} proximity labeling and biotin enrichment. Significantly enriched proteins in green and significantly depleted proteins in red. 0 proteins significantly depleted vs. 5 proteins significantly enriched, adjusted p-value <0.05, |log2FC| ≥ 2. **b**, Volcano plot showing log2 protein abundance against -log10 adjusted p-value of compound **1** treated cells over DMSO control after APEX2-SMNDC1^{FL} proximity labeling and biotin enrichment. Nuclear speckle proteins marked in red. **c**, Volcano plot showing log2 protein abundance against -log10 adjusted p-value of compound **1** treated cells over DMSO control after APEX2-SMNDC1^{FL} proximity labeling and biotin enrichment. Proteins identified by SRSF7-APEX2 proximity labeling marked in red. **d**, Western blot with Streptavidin-HRP showing all biotinylated proteins, antibodies against SFPQ, APEX2, and SMNDC1. Cells overexpressing APEX2-SMNDC1^{FL} or APEX2-SMNDC1^{TD} were treated with DMSO or compound **1**, and proximity-labeled. Biotinylated proteins were enriched and separated in an SDS-PAGE. Representative images of *n*=2. **e**, Fluorescence recovery after photobleaching (FRAP) experiment in SMNDC1-GFP intron 2-3, SRRM2-RFP intron 9-10, αTC1-cells, treated with DMSO (filled symbols) or 50 µM compound **1** (empty symbols). Relative intensity of Hoechst SMNDC1-GFP (green), and SRRM2-RFP (red) only in reference region over time (bleach region: see Fig. 7i). Data plotted as mean with standard error of the mean, n ≥ 11. **f**, DNA-bands on agarose gel after reverse transcription and PCR amplification of RNA to confirm alternative splicing events. RNA was isolated from αTC1 cells transfected with empty vector or SMNDC1 knock-down (KD) plasmid and treated with DMSO, 2 µM compound **1** for 5 days, or 50 µM compound **1** for 16 h. Red boxes show relevant bands for confirmed events.



Supplementary Fig. 7 | RNAbindRplus of SMNDC1. RNAbindRplus score of amino acids (AA) over length of SMNDC1 protein. Areas over threshold of 0.5 marked in red.

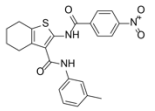
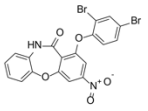
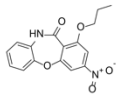
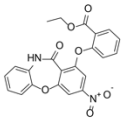
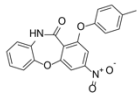
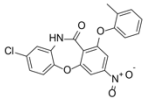
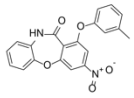
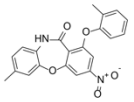
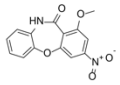
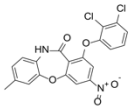
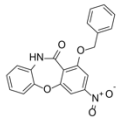
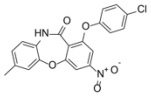
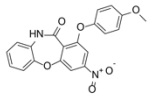
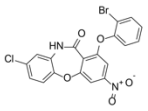
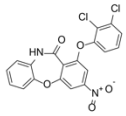
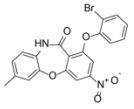
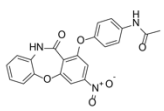
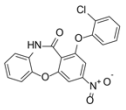
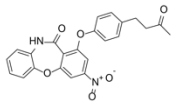
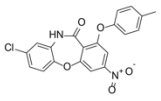
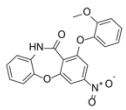
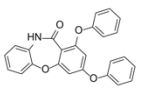
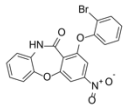
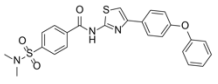
Supplementary Table 1. Small molecule screening data.

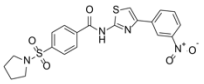
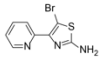
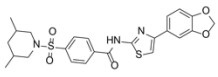
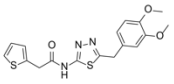
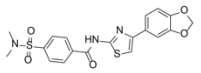
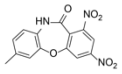
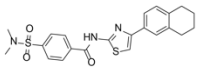
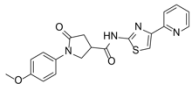
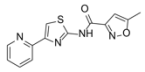
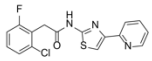
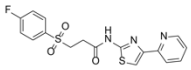
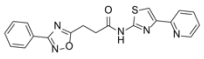
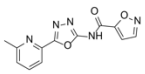
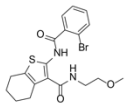
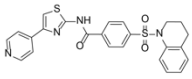
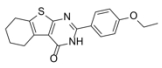
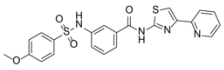
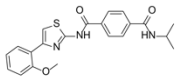
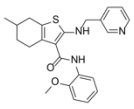
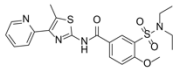
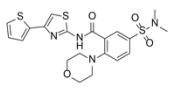
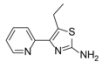
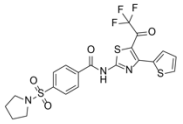
Category	Parameter	Description
Assay	Type of assay	*In vitro* AlphaScreen / luminescence proximity
	Target	Tudor domain of SMNDC1
	Primary measurement	Decrease in luminescence signal reflecting disruption of heterodimeric complex between SMNDC1-Tudor domain and a peptide corresponding to the C-terminus of Small nuclear ribonucleoprotein Sm D3 containing 4 sDMAs (Sequence: AAR*GR*GR*GMGR*GNIFQKRR, R*=sDMA)
	Key reagents	AlphaScreen no-wash assay kit containing Streptavidin Donor beads and nickel chelate (Ni-NTA) AlphaScreen Acceptor beads (PerkinElmer Part Number 6760619).
	Assay protocol	Described in Methods section
	Additional comments	-
Library	Library size	89,355 small molecules
	Library composition	Structural diversity, NIH clinical collection, natural products, approved drugs, known bioactives (e.g., kinase, epigenetic modifiers, ...), natural products, drug-like molecules
	Source	Cayman chemical, Enamine Ltd, LC Labs, MedChem Express, Selleck Chemicals, Sigma Aldrich, Tocris, Toronto Research Chemicals, Chemietek, Merck Millipore, Specs, ChemDiv, Zelinsky

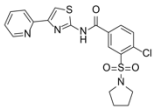
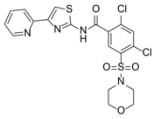
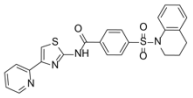
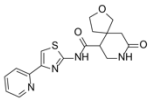
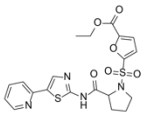
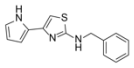
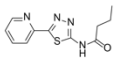
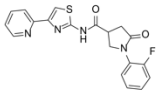
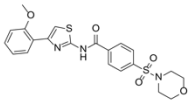
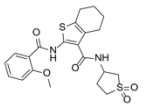
	Additional comments	
Screen	Format	PerkinElmer OptiPlate-384 well plate
	Concentration(s) tested	Typically 10 µM (0.1% DMSO)
	Plate controls	32 positive control wells (Mitoxantrone, quencher), 32 negative control wells (DMSO)
	Reagent/ compound dispensing system	Echo 520 Liquid Handler Multidrop™ Combi Reagent Dispenser
	Detection instrument and software	2104 EnVision Multilabel Plate Reader
	Assay validation/QC	Average Z′-score= 0.821
	Correction factors	-
	Normalization	Raw signal was normalized plate-specifically by correcting row and column-specific mean signals to the mean signal of the entire plate, each after removing the highest and lowest 25% of values. Raw signal was then converted to percent of control signal, but linear regression to plate-specific mean signal of DMSO wells (set to 100 percent of control) and positive control wells (set to 0% of control), after outlier removal using a Grubbs test.
	Additional comments	
Post-HTS analysis	Hit criteria	Percent of control ≤ 50 %
	Hit rate	Primary screen: 511 small molecules (including unspecific quenchers), Secondary screen (with crosslinking peptide): 40 small molecules, Tertiary screen (titration): 14 small molecules.
	Additional assay(s)	Secondary screen with crosslinking peptide, Tertiary screen as 4-point titration, also with related Tudor domain of SMN

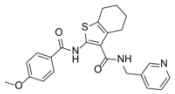
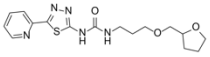
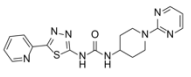
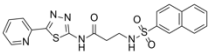
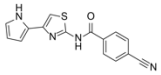
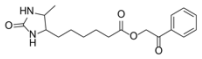
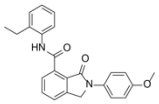
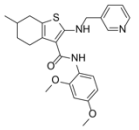
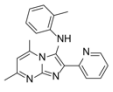
Confirmation of hit purity and structure	Re-ordered and tested selected hits (Fig. 3 f-l)
Additional comments	

Supplementary Table 2. Overview of all tested small molecules.

	structure	SMNDC1 IC_{50} (µM)	SMN1 IC_{50} (µM)		structure	SMNDC1 IC_{50} (µM)	SMN1 IC_{50} (µM)
29		29.16	>50	**41**		9.90	48.39
30		29.27	>50	**42**		22.57	>50
31		9.51	42.12	**43**		2.41	6.80
32		23.91	>50	**44**		2.02	8.41
33		>50	49.95	**45**		1.14	4.72
34		24.91	22.95	**46**		2.38	7.03
35		11.27	16.19	**47**		4.43	10.94
36		8.61	41.27	**48**		3.43	16.65
37		8.56	22.40	**49**		10.61	>50
38		17.84	38.72	**50**		2.17	15.91
39		12.53	15.61	**51**		38.37	>50
40		31.15	>50	**52**		>50	>50

	structure	SMNDC1 IC_{50} (μM)	SMN1 IC_{50} (μM)		structure	SMNDC1 IC_{50} (μM)	SMN1 IC_{50} (μM)
53		>50	>50	**65**		3.05	8.36
54		>50	>50	**66**		2.33	5.77
55		36.72	>50	**67**		4.13	12.25
56		26.52	>50	**68**		3.01	3.49
57		0.47	0.90	**69**		3.06	3.51
58		5.15	>50	**70**		16.83	15.43
59		>50	>50	**71**		>50	>50
60		>50	>50	**72**		>50	>50
61		2.82	4.49	**73**		>50	>50
62		1.53	2.88	**74**		7.35	42.86
63		5.89	21.68	**75**		26.05	34.22
64		1.18	3.14	**76**		49.92	>50

	structure	SMNDC1 IC_{50} (μM)	SMN1 IC_{50} (μM)
77		3.47	3.17
78		2.47	3.79
79		11.23	5.68
80		1.84	1.76
81		18.51	>50
82		9.60	22.12
83		>50	>50
84		2.43	3.58
85		>50	36.93
86		>50	>50

	structure	SMNDC1 IC_{50} (μM)	SMN1 IC_{50} (μM)
87		>50	>50
88		>50	>50
89		>50	>50
90		>50	>50
91		4.82	6.29
h		0.19	3.99
i		1.44	3.91
j		1.81	1.07
k		1.09	1.15

Supplementary Table 3. Assigned intermolecular NOEs between SMNDC1 and inhibitor **13** and associated restraint upper limits.

Protein / inhibitor contacts		HE-Y111 / H2-**13**	4.6	HE-Y111 / H4-**13**	4.9
HZ2-W83 / H2-**13**	4.7	HB2-W83 / H3-**13**	5.0	HB1-N113 / H4-**13**	5.0
HE-Y90 / H2-**13**	4.4	HB1-Y90 / H3-**13**	5.0	HB2-Y90 / H5-**13**	5.0
HB2-F108 / H2-**13**	5.0	HB2-Y90 / H3-**13**	5.0	HB2-Y111 / H5-**13**	5.0
HE-F108 / H2-**13**	4.7	HB2-F108 / H3-**13**	5.0	HD-Y111 / H5-**13**	4.7
HB1-Y111 / H2-**13**	5.0	HB1-Y111 / H3-**13**	5.0	HE-Y111 / H5-**13**	4.7
HB2-Y111 / H2-**13**	5.0	HD-Y111 / H3-**13**	4.7	HB2-Y111 / H6'-**13**	5.0
HD-Y111 / H2-**13**	4.3	HD-Y111 / H4-**13**	4.6	HE-Y111 / H6'-**13**	4.6

Supplementary Table 4. SMNDC1/ compound **13** molecular-docking results of structure ensemble generated using the HADDOCK webserver[52,53]. Statistics generated over all 200 analyzed structures.

HADDOCK score	-23.1 ± 0.0		
Cluster size	200		
Number of inter-molecular distance restraints	23[a]		
Number of violations (mean)	>0.5 Å: 1[b]	>0.2 Å: 1	>0.1 Å: 2.3
Violations (mean ± standard deviation) (Å)	0.39 ± 0.014		
RMSD from overall lowest-energy structure (Å)	0.6 ± 0.6		
Van der Waals energy	-17.4 ± 0.6		
Electrostatic energy	-15.2 ± 0.5		
Desolvation energy	-7.9 ± 0.2		
Restraint violation energy	37.5 ± 0.14		
Buried Surface Area	312.8 ± 1.4		
Average pairwise backbone RMSD (Å)	0.61 ± 0.79		

[a] Restraints listed in Supplementary Table 3.

[b] The violated distance involves proton **13**-H2 and Y90-HE.

Methods

Nomenclature

To reduce confusion due to the difference between gene and protein name we have decided to only use SMNDC1 for both.

AlphaFold

AlphaFold[37] predictions were run via ColabFold[74] with the AlphaFold2 algorithm and the following parameters:

msa_method=mmseqs2
homooligomer=1
pair_mode=unpaired
cov=0
qid=0
max_msa=512:1024
subsample_msa=True
num_relax=0
use_turbo=True
use_ptm=True
rank_by=pLDDT
num_models=5
num_samples=1
num_ensemble=1
max_recycles=3
tol=0
is_training=False
use_templates=False

Intron tagging and live imaging of cells

Cell lines with fluorescent tags in the endogenous intron loci of different genes were generated as described in Serebrenik et al., and Reicher et al.[39,75]. Cells were transiently transfected using Avalanche - Everyday Transfection Reagent with three plasmids in parallel: (1) the donor plasmid containing the artificial intron with splice acceptor and splice donor site, the fluorescent tag GFP or RFP, and a possible correction for the frame of the targeted intron, (2) a pX330 backbone containing Cas9 and the gRNA against the donor plasmid, and (3) a plasmid expressing the gRNA against the target intron (see table below). After 3-5 days, GFP- and/or RFP-positive cells were sorted on a SONY SH800 Cell Sorter to get fluorescent single cell clones. Clones were validated for the correct integration of the intron-tag via comparison

of live cell images to publicly available or in-house IF images, genomic DNA PCR amplification of the respective loci, and western blots with antibodies against the target protein and/or the fluorescent tag.

Cells were imaged on a PerkinElmer Opera Phenix automated microscope with 500 ms exposure time in either GFP or RFP channel, or on a Zeiss LSM 980 microscope. For condition-independent identification nuclear markers such as Hoechst or DRAQ5™ were used.

Gene	Species	Intron	gRNA Sequence
SMNDC1	mouse	1-2	GGACCCGTATGTTTGCCCCG
	mouse	2-3	AGACTTCCAGGCCAGCCAAG
	human	2-3	CTTGTGGAAATTGAACTATG
	mouse	3-4	TCACCTACACAGATCACGAT
	mouse	4-5	GCTAACCTGAGTTTAACCAT
	mouse	5-6	GTACCTAATGACTATTGACA
SRRM2	mouse	9-10	GATAGCTTAATGGGCCCATG
SMN1	mouse	5-6	TGAGCACTGGAGATACGGCG

Immunofluorescence

Cells were fixed in the 96-well imaging plates they were growing in before by adding 37% formaldehyde solution 1:10 to the culture medium for a final concentration of 3.7%. Cells were incubated with this for 15 min at room temperature (RT). Next, cells were washed once with PBS, followed by a 30 min permeabilization step with PBST (0.2% Tween). Afterwards, cells were blocked with a 3% BSA in PBST solution for 1h. Primary antibodies in 1.5% BSA in PBST were added in their individual concentrations and incubated overnight (o/n) at 4°C. On the next day, wells were washed 3x with PBST, before incubation with secondary antibodies (diluted 1:500) and DAPI (5 mg/ml, 1:2000) for 1-2h. After 3 washing steps with PBST, cells were ready to be imaged. Cells can be stored at 4°C before imaging on the PerkinElmer Opera Phenix automated microscope or on a Zeiss LSM 980 microscope.

Imaging quantifications

Images were analyzed using the high-content image acquisition and analysis software Harmony® 4.9 developed by PerkinElmer. First, nuclei were identified in the channel of the nuclear marker (DAPI/ Hoechst/ DRAQ5™) (with Method C, Common Threshold 0.75, Area > 10 μm^2). After the identification of nuclei, their corresponding cytoplasm was also identified using the respective nucleic acid marker (with Method A, Individual Threshold 0.15). Even

though the highest staining of these nuclear markers is obviously detected in the nucleus they still produce a significant staining of the cytoplasm above background. After defining the respective cell areas, mean intensity in the different channels was measured. Finally, spots were identified with the according “Spots” algorithm (with Method A, Relative Spot Intensity > 0.053, Splitting Sensitivity: 1.0).

Colocalization analysis

Images were preprocessed in Python version 3.7.9. Z-stacks in czi format were loaded with czifile library, version 2019.7.2, and reduced using maximum intensity Z projection. Segmentation of nuclei was carried out with Cellpose[76] (version 0.6.1) based on the DAPI channel. Additional segmentation masks (mitotic nuclei only) were created manually. Preprocessed images and segmentation masks were saved in PNG format. CellProfiler[77] (4.0.7) was used to extract fluorescence intensity measurements for non-mitotic and mitotic nuclei separately.

All preprocessing code and the CellProfiler pipeline are available at https://github.com/reinisj/colocalization_analysis.

In-vitro protein expression

Expression plasmids for SMNDC1’s and SMN1’s Tudor domain as used in Tripsianes et al., 2011[7] were a kind gift from Michael Sattler. Protein expression plasmids with a GFP-fusion for droplet assays were generated by ligation independent cloning[78,79] using pET His6 GFP TEV LIC cloning vector (1GFP) which was a gift from Scott Gradia (Addgene plasmid # 29663 ; http://n2t.net/addgene:29663 ; RRID:Addgene_29663) and amplification of the respective sequences from cDNA.

BL21(DE3) competent *E. coli* cells were transformed with the respective plasmids and liquid stocks frozen at -80°C. Volumes described here are for 450 ml total volume bacterial culture but were adjusted according to protein amounts needed. From frozen liquid stocks, 200 ml LB Kanamycin cultures were grown at 30°C overnight, diluted with 250 ml fresh LB and grown until OD_{600} reached 0.8-1. Protein expression was induced with 1mM IPTG and bacteria grown for another 24h at 20°C. Bacteria were harvested by centrifugation at 4000xg for 15min at 4°C. Pellets were washed in 35ml PBS and spun down again at 6000xg for 10 min at 4°C. After removal of supernatant PBS, pellets can be stored at -80°C.

For protein purification, pellets were resuspended in 13 ml Lysis buffer (50 mM TRIS pH 7.7, 500 mM NaCl, 1 % Igepal, 2.5 mg/ml Lysozyme, 0.1 mg/ml DNase I), incubated for at least 15min and sonicated to ensure cell lysis. Afterwards, lysates were spun down again for 20min at 8500xg and 4°C to remove debris pellet. In parallel, 1ml of Ni-NTA resin (Qiagen) were

added to a 15ml tube and centrifuged at 700xg for 2min. Supernatant was removed, and resin washed once by RIPA w/o EDTA (50 mM TRIS pH 7.7, 500 mM NaCl, 1 % Igepal). Lysate supernatant was then added to equilibrated resin and rotated at 4°C for 3h. Beads were then spun down again at 700xg for 2min, and washed rotating for 10min twice by adding 14 ml RIPA w/o EDTA. Eventually, bound protein was eluted 5x with 1 ml elution buffer (250 mM Imidazole in RIPA w/o EDTA) by rotating at room temperature for 20 min, 25 min, 30 min, 45 min, o/n.
Proteins for droplet assays were then purified further by size exclusion chromatography (SEC) on a Superdex increase 200 10/300 GL column with 50mM Tris pH 7.5, 125mM NaCl, 10% glycerol and 1mM DTT running buffer.

In-vitro droplet assays

In vitro droplet assays were performed as described in Klein, Boija et al.[80]. Recombinant GFP-fusion protein purified by SEC in 50mM Tris pH 7.5, 125mM NaCl, 10% glycerol and 1mM DTT running buffer was diluted to 10 µM with a concentrated PEG-8000 solution in the same buffer (and additional buffer according to protein concentration) to a final PEG-8000 concentration of 15%. In some of the experiments, total RNA isolated from αTC1 cells (10 ng/µl, Fig. 1i, k) or *in vitro*-transcribed RNA (100 ng/µl, Fig. 1j) was added. 10 µl of this solution were loaded onto PerkinElmer PhenoPlate™ 384-well microplates (formerly named CellCarrier Ultra microplates) and imaged immediately on the PerkinElmer Opera Phenix automated microscope with a 63x objective at the bottom of the well.

Fluorescence recovery after photobleaching (FRAP)

For FRAP experiments, cells harboring intron-tags in SMNDC1 and SRRM2 were seeded 24h before imaging on a Zeiss LSM 980 microscope. 15 min before imaging, medium was changed to medium without phenol red containing DRAQ5™ 1:1000 to reduce autofluorescence and to mark nuclei. If cells were treated with compounds, these were added in the same step. After identifying a suitable cell, bleach and reference regions were defined. After taking one reference image, the bleach region was bleached 15 times for 5 milliseconds with 100% laser power at 488 nm for GFP and with 20% laser power at 546 nm for RFP. After bleaching, a new image was taken approximately every 3 seconds until 150 seconds after bleaching. Fluorescence intensities were quantified in the bleach and reference regions for every image and normalized to the intensity before bleaching.

AlphaScreen

Compounds and controls were transferred on PerkinElmer OptiPlate-384 plates using an acoustic liquid handler (Echo, Labcyte). The AlphaScreen was conducted in 20 mM sodium

phosphate pH 6.5, 50 mM NaCl, Tween 0.01%, BSA 0.1%. Protein concentration was optimized for each batch of purified protein. Optimal concentration was chosen at the lowest concentration with ~80% of maximum signal. The biotinylated binding peptide (Sequence: AAR*GR*GR*GMGR*GNIFQKRR, R*=sDMA)[7] was used at 50 nM, and donor and acceptor beads at 5 µg/ml final concentration. In the first step, 10 µl of the protein solution containing either SMNDC1's or SMN's Tudor domain coupled to a 6xHis-Tag were distributed to 384-well plates pre-spotted with compounds and controls, shaken and incubated for 30 min at RT. Afterwards, 10 µl of peptide solution was added to each well and incubated for 1h at RT. Finally, 5 µl of a solution containing both Streptavidin Donor and nickel chelate (Ni-NTA) Acceptor beads was added and again incubated for 1h at RT. AlphaScreen signal was read out on a 2104 EnVision Multilabel Plate Reader with AlphaScreen settings, excitation time 180 ms, total measurement time 550 ms.

Proximity labeling with APEX2

Proximity labeling with APEX2 was done following the described protocol for imaging and proteomic analysis[47]. Cell lines with a stable expression of APEX2-fusion proteins were generated using lentiviral transduction of plasmids generated with Gateway cloning of the respective fusion protein into pLEX305. pLEX_305 was a gift from David Root (Addgene plasmid # 41390 ; http://n2t.net/addgene:41390 ; RRID:Addgene_41390). The APEX2 sequence was amplified from APEX2-csGBP which was a gift from Rob Parton (Addgene plasmid # 108874 ; http://n2t.net/addgene:108874 ; RRID:Addgene_108874). Briefly, cells were incubated with 0.5mM biotin-phenol for 30min, after which 1mM H_2O_2 was added for exactly 1min. Afterwards the labeling reaction was quenched by 3 quick washes with Quenching solution (10mM sodium ascorbate, 10mM sodium azide, 5mM Trolox in PBS). Cells were then fixed for IF analysis or detached from the plates with a cell scraper for WB or MS analysis.

Biotin enrichment after proximity labeling

After proximity labeling, cells (approx. 10 Mio. cells, 15 cm dish) were harvested, washed 2x in PBS, snap frozen and stored at -80°C. Cell pellets were resuspended in 200 µL freshly prepared lysis buffer (1x PBS, 1% SDS, 2mM $MgCl_2$, Protease inhibitors, Benzonase), vortexed and incubated at 37°C for 30min. Samples were then centrifuged for 30min at 18,000xg and +4°C, supernatants transferred into fresh 1.5ml lo-bind tubes on ice. After quantification of protein amounts by Pierce™ 660nm Protein Assay, samples were normalized to 500 µg total protein input in a final volume of 300 µl lysis buffer. For reduction, 30 µl of 50 mM TCEP were added for a final concentration of 4.5 mM, vortexed and incubated on a shaking thermoblock at 56°C for 1h. After adjustment of pH by addition of 80µl 1M HEPES pH

7.5, 45µl of freshly prepared 200mM iodoacetamide were added for alkylation. Samples were vortexed and incubated on a shaking thermoblock at 25°C for 30min with light protection.
During reduction and alkylation 100 µl of streptavidin agarose beads (Pierce™ Streptavidin Agarose, Thermo Scientific, 20353) per sample were taken to 5 ml tubes in batches of 400 µl. To settle down the beads, tubes were centrifuged for 30sec in a table-top spin centrifuge and settled further on ice for 3min before taking off the supernatant. Beads were washed twice in 4ml PBS. After the last washing, beads were resuspended in PBS and combined to a final volume of 100 µl/ sample. After distribution of 100 µl/ sample, 1.35 ml of PBS were added, and beads stored at 4°C.
For enrichment, reduced and alkylated samples were then added to the prepared beads and rotated at 25°C for 1h. To settle down the beads, tubes were centrifuged for 30sec in a table-top spin centrifuge and settled further at RT before taking off the supernatant.
BioRad Minispin columns were equilibrated on vacuum manifold with 1ml Wash buffer 1 (0.2% SDS in PBS). Beads with enriched proteins were transferred from tubes to columns by resuspending in 2x 0.5ml Wash buffer 1. Afterwards, beads were washed 10x in 0.5ml Wash buffer 2 (8M Urea in PBS) and 4x in 0.5ml PBS. After closing of columns, beads were resuspended in 2x 0.5ml digestion buffer (H_2O (HPLC grade), 50mM Ammonium bicarbonate, 0.2M Guanidine hydrochloride, 1mM Calcium chloride) and transferred into fresh 1.5ml lo-bind tubes. To settle down the beads, tubes were centrifuged for 30sec in a table-top spin centrifuge and settled further on ice for 3min before taking off the supernatant. 250µl Digestion buffer were added to the beads, and beads were stored at 4°C before the overnight digest. 10µl trypsin (0.1µg/µl, total 1µg) were added to each tube at the end of the day, incubation at 37°C rotating inside the incubator overnight (~14h).
For solid phase extraction (SPE) stage tips were prepared as follows. 32x 1mm in diameter C18 material was punched out from Empore C18 disk using blunt syringe needle and plunged into filter-less P200 pipette tip, pushing towards narrow end of the tip. The metal piston was pressed down to fix the C18. 24µl oligo R3 solution (15mg/ml in 100% acetonitrile (ACN)) were applied to the C18 tip, centrifuged at 1,000xg for 1min inside of a collection tube. C18 was activated by washing 2x with 100µl 100% ACN, centrifugation at 1,000xg for 1min. Columns were equilibrated with 200µl 0.1% TFA, centrifuged at 1,000xg for 30sec, wrapped in parafilm and stored at 4°C overnight. Right before using them for clean-up of digests the next day, C18 columns were centrifuged at 1,000xg for 2min, equilibrated again with 200µl 0.1% TFA, and centrifuged at 1,000xg for 3min.
After overnight digest, beads were separated via centrifugation at 1,000xg for 30sec, and complete supernatants transferred into fresh 1.5ml lo-bind tubes. Beads were washed with 200µl H_2O for HPLC using wide pipette tips, centrifuged again for 30sec at 1,000xg and supernatant combined with digest. The peptide samples were then acidified with 16µl 30%

TFA (~1% final) and loaded to the C18 columns in fractions of max. 250 µl, and centrifuged at 1,000xg for 3min each. After loading the full volume, columns were washed with 200µl 0.1% TFA, and centrifuged at 1,000xg for 3min. Samples were eluted with 2x 50 µl elution buffer (90% ACN, 10% of 0.1% Trifluoroacetic acid (final 0.01%)) by centrifugation at 1,000xg for 3min. Eluates were dried in vacuum centrifuge at V-AQ, 45°C for 1.5h and stored at -20°C until TMT-labeling.

Dried pellets after SPE were reconstituted in 15µl of 100mM HEPES pH 8.5 in H_2O for HPLC (diluted from 1M HEPES pharmaceutical standard stock solution, pH adjusted using NaOH for HPLC). Aliquots of frozen TMT labels (Lotnr. WA314599) were equilibrated at RT for 5min, spun down in spin-centrifuge, vortexed and spun down again. 4µl of respective TMTpro label were added, vortexed, spun down in spin-centrifuge and incubated at 25°C and 300rpm for 1h. Reaction was stopped by adding 1.5µl of 5% hydroxylamine solution in H_2O for HPLC (prepared fresh from 50% hydroxylamine stock solution), vortexing, spinning down in spin-centrifuge and incubation at 25°C and 300rpm for 15min. Full volumes of respective TMTpro channels were then pooled into fresh 1.5ml lo-bind tube.

For a 2D analysis, samples were fractionated by on-tip high pH fractionation. Fresh ammonium formate (AF) buffer was prepared right before using, as it is volatile: 100mM ammonium formate in 2ml tube (6.3mg into 1ml H_2O for HPLC) mixed into 4ml H_2O for HPLC in 15ml tube, pH 10 adjusted with two drops of 25% ammonia solution (~35µl, final concentration 20mM). For 2D analysis, 1ml of 20mM freshly prepared AF was added to 320µl of pooled sample. C18 columns were prepared as described above. The eluate was loaded in fractions (max. capacity 200µl at once), centrifuged at 1,000xg for 3min each. The column was washed with 200µl 20mM AF, and centrifuged at 1,000xg for 3min. Each fraction was eluted in a fresh 1.5ml lo-bind tube. All fractionation buffers (100% ACN and 20mM AF mixed at different ratios) were prepared fresh:

Fraction 1: Elution with 50µl 16% ACN (24µl ACN +126µl 20mM AF), centrifuged at 1,000xg for 2min, washed with 20µl of same buffer, collected together in tube #1, centrifuged at 1,000xg for 1min.

Fraction 2: Elution with 50µl 20% ACN (30µl ACN +120µl 20mM AF), centrifuged at 1,000xg for 2min, washed with 20µl of same buffer, collected together in tube #2, centrifuged at 1,000xg for 1min.

Fraction 3: Elution with 50µl 24% ACN (36µl ACN 114µl 20mM AF), centrifuged at 1,000xg for 2min, washed with 20µl of same buffer, collected together in tube #3, centrifuged at 1,000xg for 1min.

Fraction 4: Elution with 50µl 28% ACN (42µl ACN +108µl 20mM AF), centrifuged at 1,000xg for 2min, washed with 20µl of same buffer, collected together in tube #4, centrifuged at 1,000xg for 1min.

Fraction 5: Elution with 50µl 80% ACN (120µl ACN +30µl 20mM AF), centrifuged at 1,000xg for 2min, washed with 20µl of same buffer, collected together in tube #5, centrifuged at 1,000xg for 1min.

All 5 eluates were dried in vacuum centrifuge at 45°C, V-AQ for at least 2h (until dry) and frozen at -20°C until analysis.

For a WB analysis instead of the described preparation of samples for MS, samples were not reduced and alkylated, but instead loaded on to streptavidin beads directly after lysis, quantification, and normalization. Instead of digesting proteins on the beads after enrichment, beads were transferred to lo-bind tubes with 2x 0.5mL PBS. After removal of supernatant, proteins were eluted from the beads in 3 rounds. First, 50 µl 4x LB was added, beads incubated at 95°C for 10 mins, spun down and supernatant transferred to a new tube. Second, 50 µl 1x LB was used and combined. Last, 50 µl PBS was used and combined. Typically, 30 µl of sample were loaded on an SDS-PAGE gel.

SDS-PAGE followed by Coomassie staining or western blotting

To separate proteins according to size, cell lysates/ protein solutions were loaded onto SDS-polyacrylamide gels (12%) with 4x Laemmli loading buffer (LB):

17.6ml 0.5M Tris pH 6.8
17.6ml Glycerol
8.8ml 20% SDS
2ml 1% bromophenol blue
2ml beta-mercaptoethanol

Afterwards, proteins were separated through application of an electric field (120V for 15 min, 160V for 90 min). For visualization of total protein, gels were stained with Coomassie Blue. To do so, the gel was fixed in fixing solution (50% methanol, 10% glacial acetic acid) for 1h with gentle agitation. The gel was then stained in staining solution (0.1% Coomassie Brilliant Blue R-250, 50% methanol and 10% glacial acetic acid) for 20 min, followed by several rounds of destaining with destaining solution (40% methanol, 10% glacial acetic acid).

For visualization of individual proteins, they were transferred to a nitrocellulose membrane (GE Healthcare Life Science) by electrophoresis. The membrane was blocked by 5% Milk solution in TBST for at least 1h at RT, followed by incubation in primary antibody solution (dilution 1:1000 in 5% Milk TBST) at 4°C o/n. Membranes were then washed 3 times in TBST, followed by incubation with HRP-coupled secondary antibody solution (dilution 1:20000 in 5% Milk TBST) for at least 1h at RT. After 3 more washing steps, signal was detected by application of Clarity ECL Western Blotting Substrate (Bio-Rad) to the membrane with a ChemiDoc MP Imaging System (Bio-Rad).

2D-RP/RP Liquid Chromatography – Tandem Mass Spectrometry analysis

Mass spectrometry analysis was performed on an Orbitrap Fusion Lumos Tribrid mass spectrometer (ThermoFisher Scientific, San Jose, CA) coupled to a Dionex Ultimate 3000 RSLCnano system (ThermoFisher Scientific, San Jose, CA) via a Nanospray Flex Ion Source (ThermoFisher Scientific, San Jose, CA) interface. Peptides were loaded onto a trap column (PepMap 100 C18, 5 µm, 5 × 0.3 mm, ThermoFisher Scientific, San Jose, CA) at a flow rate of 10 µL/min using 0.1% TFA as loading buffer. After loading, the trap column was switched in-line with an Acclaim PepMap nanoHPLC C18 analytical column (2.0 µm particle size, 75µm IDx500mm, catalog number 164942, ThermoFisher Scientific, San Jose, CA). Column temperature was maintained at 50 °C. Mobile-phase A consisted of 0.4% formic acid in water and mobile-phase B of 0.4% formic acid in a mix of 90% acetonitrile and 10% water. Separation was achieved by applying a four-step gradient over 151 min at the flow rate of 230 nL/min (initial gradient increase from 6% to 9% solvent B within 1 min, 9% to 30% solvent B within 146 min, 30% to 65% solvent B within 8 min and, 65% to 100% solvent B within 1 min, 100% solvent B for 6 min before equilibrating at 6% solvent B for 23 min prior to next injection). In a liquid-junction set-up, electrospray ionization was enabled by applying a voltage of 1.8 kV directly to the liquid to be sprayed, and non-coated silica emitters were used.

The mass spectrometer was operated in a data-dependent acquisition mode (DDA) and used a synchronous precursor selection (SPS) approach, which enables more accurate multiplexed quantification of peptides and proteins at the MS3 level. For both MS2 and MS3 level we collected a survey scan of 400–1600 m/z in the Orbitrap at a resolution of 120 000 (FTMS1), an AGC target was set to ‘standard’ and a maximum injection time (IT) of 50 ms was applied. Precursor ions were filtered according to charge state (2-6), dynamic exclusion (60 s with a ±10 ppm window), and monoisotopic precursor selection. Precursor ions for data-dependent MS^n ($ddMS^n$) analysis were selected using 10 dependent scans (TopN approach). Charge state filter was used to select precursors for data-dependent scans. In $ddMS^2$ analysis, spectra were acquired using a single charge state per branch (from z=2 to z=5) in a dual-pressure linear ion trap (ITMS2). Quadrupole isolation window was set to 0.7 Da and collision induced dissociation (CID) fragmentation technique was used at a normalized collision energy of 35%. Normalized AGC target value was set to 200% with a maximum IT of 35 ms. During the $ddMS^3$ analyses, precursors were isolated using SPS waveform and different MS1 isolation windows (1.3 m/z for z=2, 1.2 m/z for z=3, 0.8 m/z for z=4 and 0.7 m/z for z=5). Target MS2 fragment ions were further fragmented by high-energy collision induced dissociation (HCD) followed by Orbitrap analysis (FTMS3). The HCD normalized collision energy was set to 45% and normalized AGC target was set to 300% with a maximum IT of 100 ms. The resolution was set to 50 000 with defined scan range from 100 to 500 m/z. Xcalibur version 4.3.73.11 and Tune 3.4.3072.18 were used to operate the instrument.

Data processing and data analysis

Following data acquisition, acquired raw data files were processed using the Proteome Discoverer v.2.4.1.15 platform, choosing a TMT16plex quantification method. In the processing step we used Sequest HT database search engine and Percolator validation software node to remove false positives with a false discovery rate (FDR) of 1% on peptide and protein level under strict conditions. Searches were performed with full tryptic digestion against the mouse SwissProt database v.2017.10.25 (SwissProt TaxID=10090, 25 097 sequences and appended known contaminants and streptavidin) with a maximum of two allowed miscleavage sites. Oxidation (+15.994 Da) of methionine and acetylation of protein N termini (+42.011 Da), as well as methionine loss (-131.040 Da) and acetylation of protein N termini with methionine loss (-89.030Da) were set as variable modification, while carbamidomethylation (+57.021Da) of cysteine residues and tandem mass tag (TMT) 16-plex labeling of peptide N termini and lysine residues (+304.207Da) were set as fixed modifications. Data was searched with mass tolerances of ±10 ppm and ±0.6 Da on the precursor and fragment ions, respectively. Results were filtered to include peptide spectrum matches with Sequest HT cross-correlation factor (Xcorr) scores of ≥1 and high peptide confidence assigned by Percolator. MS3 signal-to-noise values (S/N) values of TMTpro reporter ions were used to calculate peptide/protein abundance values. Peptide spectrum matches with precursor isolation interference values of ≥70%, SPS mass matches ≤65% and average TMTpro reporter ion S/N≤10 were excluded from quantitation. Both unique and razor peptides were used for TMT quantitation. Isotopic impurity correction was applied. Data were normalized on total peptide amount for correction of experimental bias and scaled ‘on all average’. Protein ratios are directly calculated from the grouped protein abundances using an ANOVA hypothesis test. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD037092 and 10.6019/PXD037092.

Enrichr analysis

Gene symbols of proteins identified by mass-spectrometry or subsets thereof were analyzed for enrichment of gene ontology (GO) Biological Process 2021 terms with the online tool “Enrichr” (https://maayanlab.cloud/Enrichr/)[81]. Top 10 terms ranked by p-value were plotted.

NMR experiments

Isotope-enriched SMN_{84-147} and $SMNDC_{165-128}$ were expressed and purified as described in Tripsianes et al., 2011[7]. NMR experiments were performed on Bruker Avance III spectrometers operating at 600 MHz or 800 MHz ^{1}H frequencies using H/N/C triple-resonance

cryogenic probes. All NMR acquisition was performed in 3 mm tubes at 25°C. Spectra were processed using Topspin 3.5 (Bruker) and analyzed with Cara 1.9.17[82] or NMRglue[83]-based Python scripts. Chemical shift assignments were transferred from Selenko et al.[84] and Tripsianes et al., 2011[7] (BMRB: 4899 for SMN, 18006 for SMNDC1). All titration measurements were performed in 20 mM sodium phosphate buffer pH 6.5, 50 mM NaCl, 4 mM dithiothreitol and 10% (v/v) D_2O as deuterium lock. Aqueous, buffered inhibitor stock solutions of maximal 20 mM concentration were prepared and carefully adjusted to pH 6.5. Inhibitor concentration was measured by addition of 100 µM DSS, peak integration and calculating with $C = \frac{I \cdot N_{DSS} \cdot C_{DSS}}{I_{DSS} \cdot N}$ where C, I, N and C_{DSS}, I_{DSS}, N_{DSS} is the concentration, peak intensity, and number of protons of inhibitor and DSS, respectively. Titration experiments were performed with 50 µM ^{15}N-labeled SMN_{84-147} and $SMNDC_{165-128}$ and addition of 0, 0.05, 0.1, 0.25, 0.5 and 0.8 mM compound **13**. Since the titration did not reach saturation, an additional point was measured for SMNDC1 with 8 mM compound **13**. An apparent dissociation constants K_D of around 1 mM was calculated from CSP data using $CSP = \frac{CSP_{max} \cdot c}{c + K_D}$ with c being the concentration of compound **13** and CSP_{max} the CSP at saturation. However, this value is obstructed by solubility issues of **13** and therefore not comparable with the IC50 values from the AlphaScreen assay, where lower concentrations were used. To record intermolecular NOE data, a 1 mM sample of ^{13}C,^{15}N-labeled $SMNDC_{165-128}$ was prepared, lyophilized, and resolved in D_2O containing 20 mM buffered compound **13**. To confirm saturation of binding, a ^{1}H,^{15}N-HSQC spectrum was acquired and compared with the titration data. ω_1-^{13}C-filtered and ω_2-^{13}C-filtered two-dimensional NOESY and ω_2-^{13}C-filtered, ω_1-^{13}C-edited three-dimensional NOESY experiments[85] were recorded with 150 ms mixing times. Chemical shift assignments were transferred from Tripsianes et al., 2011[7].

Docking calculation

Docking calculations were performed using the HADDOCK webserver[52,53]. Structure and topology files for compound **13** were generated by prodrg2[86]. The SMNDC1/sDMA structure (PDB: 4A4H)[7] was used as a protein model with the sDMA removed beforehand. Instead of defining active/passive residues, intermolecular NOE contacts were introduced as ambiguous restraints. Visible and assigned NOE crosspeaks were defined as distance restraints with a lower limit of 0.5 Å and upper limit of 5 Å. Peak intensities of crosspeaks were measured and normalized to the strongest peak. According to their relative intensities, upper distance limits were gradually lowered to 3.5 Å for non-overlapping crosspeaks. 1000, 400 and 400 structures were calculated for the different stages of rigid body docking, semi-flexible refinement, and final refinement, respectively. Parameters were chosen as suggested by HADDOCK for small

molecule docking. 200 structures were analyzed and clustered by RMSD with a cutoff of 1.5 Å resulting in one cluster, which contains all analyzed structures.

SMNDC1 knock-down

SMNDC1 knock-down was performed as described in Casteels et al.[9]. Briefly, Smndc1 shRNA from the TRC shRNA library (https://portals.broadinstitute.org/gpp/public/) (TRCN0000123795) was cloned into pLKO.1 (Addgene plasmid #10878). This plasmid was packaged into lentivirus in Lenti-X™ 293 T cells with Lipofectamine™ 3000 (Thermo Fisher Scientific L3000008) and packaging plasmids psPAX2 (Addgene plasmid #12260) and pMD2.G (Addgene plasmid #12259). Target cells were transduced with viral supernatant after filtering and addition of 8 µg/ml Polybrene® (Santa Cruz Biotechnology sc-134220) 48h after transfection. Medium was changed 24h later.

Splicing PCRs

To perform splicing PCRs, RNA was isolated from pelleted cells using the RNeasy Mini Kit (Qiagen, #74106). RNA was then reverse transcribed with LunaScript RT SuperMix Kit (NEB #E3010). cDNA was PCR-amplified with OneTaq® Quick-Load® 2X Master Mix (NEB # M0486) for 35 cycles with the following primers as suggested on vastdb.crg.eu:

Gene	Event	Orientation	Sequence
Rhbdd3	MmuINT0134487	F	TTCCTGCACAACTCCACTGTG
Rhbdd3	MmuINT0134487	R	GCCAGAGACTTGCAAAGGACA
Hirip3	MmuEX0022948	F	AGGCAGCAGTAATGGTGACAG
Hirip3	MmuEX0022948	R	GCGACACTTCTCCAAGGAAGG
Amdhd2	MmuINT0014597	F	GCCTGGCTTTATCGATGTGCA
Amdhd2	MmuINT0014597	R	TGTGATAAACCTCTGGTGGGGA
Vps37a	MmuEX1027740	F	AGGCAAAAGGCAAACCGTTTT
Vps37a	MmuEX1027740	R	TGTTCTCTTTTCCTTGAAGCTATTGA
Taf15	MmuEX0046179	F	ATGACCGTCGTGATGTGAGTA
Taf15	MmuEX0046179	R	CAGCATCTGGTCTGGGTCCAT
Dgat1	MmuINT0048631	F	TGGGTTCCGTGTTTGCTCTG
Dgat1	MmuINT0048631	R	CGGTAGGTCAGGTTGTCTGGA
Gnb1	MmuEX0021329	F	GACCAGCCTCGCCGACTC
Gnb1	MmuEX0021329	R	GCAGCTGGTCAAGTTCACTCA
Strada	MmuEX0045528	F	TCTTGTAAGTAAACCAGAGCGCA
Strada	MmuEX0045528	R	GGTGAGCAGCTCATAACACCC
Abcc8	MmuEX0003147	F	TACGAGGCCCGGTTCCAG

Abcc8	MmuEX0003147	R	AGCCCCTCATAGCTCTCTGC
Syt7	MmuEX0046044	F	GGCAAACGCTACAAGAATTCCT
Syt7	MmuEX0046044	R	GCATCTCGCTGGTAAGGGA
Agfg1	MmuEX0004229	F	CTGCTCAGACACAACCTGCTT
Agfg1	MmuEX0004229	R	CTGTCTGCTGAGGGAAAGCTG
Napb	MmuEX0030873	F	AAACTCCACATGCAGCTCCAG
Napb	MmuEX0030873	R	TCGGCAATGGTGATATGGTGC
Pdss1	MmuALTA0012992-2/2	F	GAGCTGCACATCTCCACCAGA
Pdss1	MmuALTA0012992-2/2	R	TGGATCATTTCTGCAACTAAGGCT
Gm10451	MmuINT1011021	F	CACTGTGGCCAAACATCCCTG
Gm10451	MmuINT1011021	R	TGGATCAAGATGTTGCAATTTTTATC
Trmt2a	MmuINT0165647	F	AGGTGAAGAGAGTAGTGGGAA
Trmt2a	MmuINT0165647	R	CGTGGTGGGTCTAGAACAGCT
Dalrd3	MmuINT0046584	F	ATCCTCTCCGTGGCTACCATC
Dalrd3	MmuINT0046584	R	TACAGACCTTGCTCCGTTCCA
Icmt	MmuEX0023562	F	TCAGAGCTTGTTTCCTTGGCT
Icmt	MmuEX0023562	R	GGCCAGAAGATGTTCTCGAGC
Mus81	MmuINT1023759	F	AGCCTTCCACAAACCCTCTCT
Mus81	MmuINT1023759	R	GGTGCTGTATCGATCCACCAC

The PCR products were run on a 1% Agarose gel for 30 min at 100 Volt.

RNA sequencing and transcriptome analysis

RNA sequencing libraries were prepared from low-input samples using the Smart-seq2 protocol[87]. The subsequent library preparation from the amplified cDNA was performed using the Nextera XT DNA library prep kit (Illumina, San Diego, CA, USA). Library concentrations were quantified with the Qubit 2.0 Fluorometric Quantitation system (Life Technologies, Carlsbad, CA, USA) and the size distribution was assessed using the Experion Automated Electrophoresis System (Bio-Rad, Hercules, CA, USA). For sequencing, samples were diluted and pooled into NGS libraries in equimolar amounts.

Expression profiling libraries were sequenced on NovaSeq 6000 instrument (Illumina, San Diego, CA, USA) with a 100-base-pair, paired-end setup. Raw data acquisition and base calling was performed on-instrument. Subsequent raw data processing off the instruments involved two custom programs (https://github.com/epigen/picard/) based on Picard tools (2.19.2) (https://broadinstitute.github.io/picard/). In a first step, base calls were converted into lane-specific, multiplexed, unaligned BAM files suitable for long-term archival

(IlluminaBasecallsToMultiplexSam, 2.19.2-CeMM). In a second step, archive BAM files were demultiplexed into sample-specific, unaligned BAM files (IlluminaSamDemux, 2.19.2-CeMM). NGS reads were mapped to the Genome Reference Consortium GRCm38 assembly via "Spliced Transcripts Alignment to a Reference" (STAR)[88] utilising the "basic" Ensembl transcript annotation from version e100 (April 2020) as reference transcriptome. The mm10 assembly of the UCSC Genome Browser was used for downstream data processing, and the Ensembl transcript annotations were adjusted to UCSC Genome Browser sequence region names. STAR was run with options recommended by the ENCODE project. NGS read alignments overlapping Ensembl transcript features were counted with the Bioconductor (3.11) GenomicAlignments (1.24.0) package via the summarizeOverlaps function in Union mode, ignoring secondary alignments and alignments not passing vendor quality filtering. Since the Smart-seq2 protocol is not strand specific, all alignments irrespective of the gene or transcript orientation were counted. Transcript-level counts were aggregated to gene-level counts and the Bioconductor DESeq2[89] (1.28.1) package was used to test for differential expression based on a model using the negative binomial distribution.

Splicing analysis

Alternative splicing events were characterised and quantified using VAST-TOOLS[54] (2.5.1) in conjunction with the *Mus musculus* database (vastdb.mm2.23.06.20), based on the Genome Reference Consortium assembly GRCm38.p5 and Ensembl transcript annotation 88 (March 2017). Briefly, NGS reads were aligned for each read group independently, read groups were merged into samples and samples were combined into a summary table. The differential splicing events were called via the VAST-TOOLS "compare" algorithm (min_dPSI > 15, min_range > 5) and further filtered for genes showing statistical significance (adjusted P-value <= 0.1) and a sizable effect (absolute log2-fold change >= 1.0) in the differential expression analysis.

Data availability statement

The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD037092 and 10.6019/PXD037092.

NMR structures will be deposited to PDB.

RNA-seq data will be deposited to GEO.

Code availability

All preprocessing code and the CellProfiler pipeline for the colocalization analysis are available at https://github.com/reinisj/colocalization_analysis.

Acknowledgements

We thank CeMM's Molecular Discovery Platform for help with small molecule screening and generation and processing of proteomics data and CeMM's Biomedical Sequencing Facility for generation and processing of our RNA-sequencing data. Yuchen Wu assisted with clone-picking and computational docking. We acknowledge NMR measurements at the Bavarian NMR Center, TUM, Garching. Lennart Enders gratefully acknowledges support by a DOC PhD Fellowship of the Austrian Academy of Sciences. This work was supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – SFB 1309 – 325871075 (M.S). Research in the Kubicek lab is supported by the Austrian Federal Ministry for Digital and Economic Affairs and the National Foundation for Research, Technology, and Development; the Austrian Science Fund (FWF) F4701; and the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (ERC-CoG-772437).

Author contributions

L.E. and S.K. planned the study and designed the experiments; L.E. performed most of the experiments with help from A.K., M.M., T.T., A.Re. and T.C.; M.Si. performed chemical syntheses and SAR studies; J.B. and S.G. performed NMR experiments; E.H. and A.Ru. performed biotin-purification and MS sample preparation; L.E., J.R., M.S., A.K., M.M., and A. Ru. analyzed the data; L.E. and S.K. wrote the manuscript with input from all co-authors; C.B., G.E.W., J.T.H., M.Sa., and S.K. supervised the work.

Declaration of interests

Lennart Enders, Marton Siklos and Stefan Kubicek are filing a patent based on the findings in this manuscript. The other authors declare no competing interests.

References

1. Meister, G. *et al.* SMNrp is an essential pre-mRNA splicing factor required for the formation of the mature spliceosome. *The EMBO Journal* **20**, 2304–2314 (2001).

2. Rappsilber, J., Ajuh, P., Lamond, A. I. & Mann, M. SPF30 Is an Essential Human Splicing Factor Required for Assembly of the U4/U5/U6 Tri-small Nuclear Ribonucleoprotein into the Spliceosome. *J. Biol. Chem.* **276**, 31142–31150 (2001).

3. Côté, J. & Richard, S. Tudor Domains Bind Symmetrical Dimethylated Arginines *. *Journal of Biological Chemistry* **280**, 28476–28483 (2005).

4. Cheng, D., Côté, J., Shaaban, S. & Bedford, M. T. The Arginine Methyltransferase CARM1 Regulates the Coupling of Transcription and mRNA Processing. *Molecular Cell* **25**, 71–83 (2007).

5. Pellizzoni, L., Kataoka, N., Charroux, B. & Dreyfuss, G. A Novel Function for SMN, the Spinal Muscular Atrophy Disease Gene Product, in Pre-mRNA Splicing. *Cell* **95**, 615–624 (1998).

6. Fischer, U., Liu, Q. & Dreyfuss, G. The SMN–SIP1 Complex Has an Essential Role in Spliceosomal snRNP Biogenesis. *Cell* **90**, 1023–1029 (1997).

7. Tripsianes, K. *et al.* Structural basis for dimethylarginine recognition by the Tudor domains of human SMN and SPF30 proteins. *Nature Structural & Molecular Biology* **18**, 1414–1420 (2011).

8. Bowerman, M. *et al.* Glucose metabolism and pancreatic defects in spinal muscular atrophy. *Annals of Neurology* **72**, 256–268 (2012).

9. Casteels, T. *et al.* SMNDC1 links chromatin remodeling and splicing to regulate pancreatic hormone expression. *Cell Reports* **40**, (2022).

10. Zhu, R., Wang, X., Yu, Q., Guo, W. & Zhu, L. A systems biology-based approach to screen key splicing factors in hepatocellular carcinoma. *Molecular Carcinogenesis* **n/a**, (2023).

11. Courchaine, E. M. *et al.* DMA-tudor interaction modules control the specificity of in vivo condensates. *Cell* **184**, 3612-3625.e17 (2021).

12. Chitiprolu, M. *et al.* A complex of C9ORF72 and p62 uses arginine methylation to eliminate stress granules by autophagy. *Nature Communications* **9**, 2794 (2018).

13. Hofweber, M. *et al.* Phase Separation of FUS Is Suppressed by Its Nuclear Import Receptor and Arginine Methylation. *Cell* **173**, 706-719.e13 (2018).

14. Qamar, S. *et al.* FUS Phase Separation Is Modulated by a Molecular Chaperone and Methylation of Arginine Cation-π Interactions. *Cell* **173**, 720-734.e15 (2018).

15. Chong, P. A., Vernon, R. M. & Forman-Kay, J. D. RGG/RG Motif Regions in RNA Binding and Phase Separation. *Journal of Molecular Biology* **430**, 4650–4665 (2018).

16. Gao, X. *et al.* Tudor-SN interacts with and co-localizes with G3BP in stress granules under stress conditions. *FEBS Lett.* **584**, 3525–3532 (2010).

17. Su, C. *et al.* Phosphorylation of Tudor-SN, a novel substrate of JNK, is involved in the efficient recruitment of Tudor-SN into stress granules. *Biochim Biophys Acta Mol Cell Res* **1864**, 562–571 (2017).

18. Zhu, L. & Brangwynne, C. P. Nuclear bodies: the emerging biophysics of nucleoplasmic phases. *Curr. Opin. Cell Biol.* **34**, 23–30 (2015).

19. Boeynaems, S. *et al.* Protein Phase Separation: A New Phase in Cell Biology. *Trends in Cell Biology* **28**, 420–435 (2018).

20. Strom, A. R. & Brangwynne, C. P. The liquid nucleome - phase transitions in the nucleus at a glance. *J. Cell. Sci.* **132**, (2019).

21. Wright, P. E. & Dyson, H. J. Intrinsically disordered proteins in cellular signalling and regulation. *Nat Rev Mol Cell Biol* **16**, 18–29 (2015).

22. Shao, W. *et al.* Phase separation of RNA-binding protein promotes polymerase binding and transcription. *Nat Chem Biol* (2021) doi:10.1038/s41589-021-00904-5.

23. Garcia-Jove Navarro, M. *et al.* RNA is a critical element for the sizing and the composition of phase-separated RNA-protein condensates. *Nat Commun* **10**, 3230 (2019).

24. Banerjee, P. R., Milin, A. N., Moosa, M. M., Onuchic, P. L. & Deniz, A. A. Reentrant Phase Transition Drives Dynamic Substructure Formation in Ribonucleoprotein Droplets. *Angewandte Chemie International Edition* **56**, 11354–11359 (2017).

25. Langdon, E. M. *et al.* mRNA structure determines specificity of a polyQ-driven phase separation. *Science* **360**, 922–927 (2018).

26. Maharana, S. *et al.* RNA buffers the phase separation behavior of prion-like RNA binding proteins. *Science* **360**, 918–921 (2018).

27. Frege, T. & Uversky, V. N. Intrinsically disordered proteins in the nucleus of human cells. *Biochem Biophys Rep* **1**, 33–51 (2015).

28. Aumiller, W. M. & Keating, C. D. Phosphorylation-mediated RNA/peptide complex coacervation as a model for intracellular liquid organelles. *Nat Chem* **8**, 129–137 (2016).

29. Bhat, P., Honson, D. & Guttman, M. Nuclear compartmentalization as a mechanism of quantitative control of gene expression. *Nat Rev Mol Cell Biol* **22**, 653–670 (2021).

30. Strom, A. R. *et al.* Phase separation drives heterochromatin domain formation. *Nature* **547**, 241–245 (2017).

31. Larson, A. G. *et al.* Liquid droplet formation by HP1α suggests a role for phase separation in heterochromatin. *Nature* **547**, 236–240 (2017).

32. Cho, W.-K. *et al.* Mediator and RNA polymerase II clusters associate in transcription-dependent condensates. *Science* **361**, 412–415 (2018).

33. Gueroussov, S. *et al.* Regulatory Expansion in Mammals of Multivalent hnRNP Assemblies that Globally Control Alternative Splicing. *Cell* **170**, 324-339.e23 (2017).

34. Liu, Y. *et al.* A small molecule antagonist of SMN disrupts the interaction between SMN and RNAP II. *Nat Commun* **13**, 5453 (2022).

35. Kroschwald, S., Maharana, S. & Simon, A. Hexanediol: a chemical probe to investigate the material properties of membrane-less compartments. *Matters* **3**, e201702000010 (2017).

36. Kozlowski, L. P. & Bujnicki, J. M. MetaDisorder: a meta-server for the prediction of intrinsic disorder in proteins. *BMC Bioinformatics* **13**, 111 (2012).

37. Jumper, J. *et al.* Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).

38. Varadi, M. *et al.* AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. *Nucleic Acids Research* **50**, D439–D444 (2022).

39. Serebrenik, Y. V., Sansbury, S. E., Kumar, S. S., Henao-Mejia, J. & Shalem, O. Efficient and flexible tagging of endogenous genes by homology-independent intron targeting. *Genome Res* **29**, 1322–1328 (2019).

40. Ferreira, J. A., Carmo-Fonseca, M. & Lamond, A. I. Differential interaction of splicing snRNPs with coiled bodies and interchromatin granules during mitosis and assembly of daughter cell nuclei. *Journal of Cell Biology* **126**, 11–23 (1994).

41. Tripathi, K. & Parnaik, V. K. Differential dynamics of splicing factor SC35 during the cell cycle. *J Biosci* **33**, 345–354 (2008).

42. Avşar Ilık, İ. & Aktaş, T. Nuclear speckles: dynamic hubs of gene expression regulation. *FEBS J* (2021) doi:10.1111/febs.16117.

43. Rai, A. K., Chen, J.-X., Selbach, M. & Pelkmans, L. Kinase-controlled phase transition of membraneless organelles in mitosis. *Nature* **559**, 211–216 (2018).

44. Sacco-Bubulya, P. & Spector, D. L. Disassembly of interchromatin granule clusters alters the coordination of transcription and pre-mRNA splicing. *Journal of Cell Biology* **156**, 425–436 (2002).

45. Ilik, İ. A. *et al.* SON and SRRM2 are essential for nuclear speckle formation. *eLife* **9**, e60579 (2020).

46. Xu, S. *et al.* SRRM2 organizes splicing condensates to regulate alternative splicing. *Nucleic Acids Research* gkac669 (2022) doi:10.1093/nar/gkac669.

47. Hung, V. *et al.* Spatially resolved proteomic mapping in living cells with the engineered peroxidase APEX2. *Nat Protoc* **11**, 456–475 (2016).

48. Barutcu, A. R. *et al.* Systematic mapping of nuclear domain-associated transcripts reveals speckles and lamina as hubs of functionally distinct retained introns. *Mol Cell* **82**, 1035-1052.e9 (2022).

49. Hartel, N. G., Chew, B., Qin, J., Xu, J. & Graham, N. A. Deep Protein Methylation Profiling by Combined Chemical and Immunoaffinity Approaches Reveals Novel PRMT1 Targets*. *Molecular & Cellular Proteomics* **18**, 2149–2164 (2019).

50. Bosse, R., Illy, C., Elands, J. & Chelsky, D. Miniaturizing screening: how low can we go today? *Drug Discovery Today* **5**, 42–47 (2000).

51. Göbl, C., Madl, T., Simon, B. & Sattler, M. NMR approaches for structural analysis of multidomain proteins and complexes in solution. *Progress in Nuclear Magnetic Resonance Spectroscopy* **80**, 26–63 (2014).

52. van Zundert, G. C. P. *et al.* The HADDOCK2.2 Web Server: User-Friendly Integrative Modeling of Biomolecular Complexes. *Journal of Molecular Biology* **428**, 720–725 (2016).

53. Honorato, R. V. *et al.* Structural Biology in the Clouds: The WeNMR-EOSC Ecosystem. *Frontiers in Molecular Biosciences* **8**, (2021).

54. Tapial, J. *et al.* An atlas of alternative splicing profiles and functional associations reveals new regulatory programs and genes that simultaneously express multiple major isoforms. *Genome Res.* **27**, 1759–1768 (2017).

55. Walia, R. R. *et al.* RNABindRPlus: A Predictor that Combines Machine Learning and Sequence Homology-Based Methods to Improve the Reliability of Predicted RNA-Binding Residues in Proteins. *PLOS ONE* **9**, e97725 (2014).

56. Arrowsmith, C. H. & Schapira, M. Targeting non-bromodomain chromatin readers. *Nat Struct Mol Biol* **26**, 863–869 (2019).

57. Zhu, H., Wei, T., Cai, Y. & Jin, J. Small Molecules Targeting the Specific Domains of Histone-Mark Readers in Cancer Therapy. *Molecules* **25**, 578 (2020).

58. Sun, Y. *et al.* Discovery of a novel 53BP1 inhibitor through AlphaScreen-based high-throughput screening. *Bioorganic & Medicinal Chemistry* **34**, 116054 (2021).

59. Perfetti, M. T. *et al.* Identification of a fragment-like small molecule ligand for the methyl-lysine binding protein, 53BP1. *ACS Chem. Biol.* **10**, 1072–1081 (2015).

60. Wagner, T. *et al.* Identification of a small-molecule ligand of the epigenetic reader protein Spindlin1 via a versatile screening platform. *Nucleic Acids Res.* **44**, e88 (2016).

61. Bae, N. *et al.* Developing Spindlin1 Small Molecule Inhibitors Using Protein Microarrays. *Nat Chem Biol* **13**, 750–756 (2017).

62. Fagan, V. *et al.* A Chemical Probe for Tudor Domain Protein Spindlin1 to Investigate Chromatin Function. *J. Med. Chem.* **62**, 9008–9025 (2019).

63. Xiong, Y. *et al.* Discovery of a Potent and Selective Fragment-like Inhibitor of Methyllysine Reader Protein Spindlin 1 (SPIN1). *J. Med. Chem.* (2019) doi:10.1021/acs.jmedchem.9b00522.

64. Senisterra, G. *et al.* Discovery of Small-Molecule Antagonists of the H3K9me3 Binding to UHRF1 Tandem Tudor Domain. *SLAS DISCOVERY: Advancing the Science of Drug Discovery* **23**, 930–940 (2018).

65. Chang, L. *et al.* Discovery of small molecules targeting the tandem tudor domain of the epigenetic factor UHRF1 using fragment-based ligand discovery. *Sci Rep* **11**, (2021).

66. Kori, S. *et al.* Structure-based screening combined with computational and biochemical analyses identified the inhibitor targeting the binding of DNA Ligase 1 to UHRF1. *Bioorganic & Medicinal Chemistry* **52**, 116500 (2021).

67. Liu, J. *et al.* Structural plasticity of the TDRD3 Tudor domain probed by a fragment screening hit. *FEBS J.* **285**, 2091–2103 (2018).

68. Upadhyay, A. K. *et al.* Targeting lysine specific demethylase 4A (KDM4A) tandem TUDOR domain - A fragment based approach. *Bioorg. Med. Chem. Lett.* **28**, 1708–1713 (2018).

69. Mader, P. *et al.* Identification and characterization of the first fragment hits for SETDB1 Tudor domain. *Bioorganic & Medicinal Chemistry* **27**, 3866–3878 (2019).

70. Engelberg, I. A. *et al.* Discovery of an H3K36me3-Derived Peptidomimetic Ligand with Enhanced Affinity for Plant Homeodomain Finger Protein 1 (PHF1). *J Med Chem* (2021) doi:10.1021/acs.jmedchem.1c00430.

71. Blanc, R. S. & Richard, S. Arginine Methylation: The Coming of Age. *Mol. Cell* **65**, 8–24 (2017).

72. Gayatri, S. & Bedford, M. T. Readers of histone methylarginine marks. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms* **1839**, 702–710 (2014).

73. Benjamini, Y., Krieger, A. M. & Yekutieli, D. Adaptive linear step-up procedures that control the false discovery rate. *Biometrika* **93**, 491–507 (2006).

74. Mirdita, M. *et al.* ColabFold: making protein folding accessible to all. *Nat Methods* **19**, 679–682 (2022).

75. Reicher, A., Koren, A. & Kubicek, S. Pooled protein tagging, cellular imaging, and in situ sequencing for monitoring drug action in real time. *Genome Res.* (2020) doi:10.1101/gr.261503.120.

76. Stringer, C., Wang, T., Michaelos, M. & Pachitariu, M. Cellpose: a generalist algorithm for cellular segmentation. *Nat Methods* **18**, 100–106 (2021).

77. Stirling, D. R. *et al.* CellProfiler 4: improvements in speed, utility and usability. *BMC Bioinformatics* **22**, 433 (2021).

78. Aslanidis, C. & de Jong, P. J. Ligation-independent cloning of PCR products (LIC-PCR). *Nucleic Acids Res* **18**, 6069–6074 (1990).

79. Stols, L. *et al.* A new vector for high-throughput, ligation-independent cloning encoding a tobacco etch virus protease cleavage site. *Protein Expr Purif* **25**, 8–15 (2002).

80. Klein, I. A. *et al.* Partitioning of cancer therapeutics in nuclear condensates. *Science* **368**, 1386–1392 (2020).

81. Chen, E. Y. *et al.* Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. *BMC Bioinformatics* **14**, 128 (2013).

82. Keller, R. L. J. Optimizing the process of nuclear magnetic resonance spectrum analysis and computer aided resonance assignment. (ETH Zurich, 2005). doi:10.3929/ethz-a-005068942.

83. Helmus, J. J. & Jaroniec, C. P. Nmrglue: an open source Python package for the analysis of multidimensional NMR data. *J Biomol NMR* **55**, 355–367 (2013).

84. Selenko, P. *et al.* SMN Tudor domain structure and its interaction with the Sm proteins. *Nat Struct Mol Biol* **8**, 27–31 (2001).

85. Sattler, M., Schleucher, J. & Griesinger, C. Heteronuclear multidimensional NMR experiments for the structure determination of proteins in solution employing pulsed field gradients. *Progress in Nuclear Magnetic Resonance Spectroscopy* **34**, 93–158 (1999).

86. Schüttelkopf, A. W. & van Aalten, D. M. F. PRODRG: a tool for high-throughput crystallography of protein–ligand complexes. *Acta Cryst D* **60**, 1355–1363 (2004).

87. Picelli, S. *et al.* Full-length RNA-seq from single cells using Smart-seq2. *Nat Protoc* **9**, 171–181 (2014).

88. Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).

89. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology* **15**, 550 (2014).

3. DISCUSSION

3.1 General Discussion

3.1.1 SMNDC1's phase separation behavior

In the paper detailed in section 2 we showed that SMNDC1 localizes to nuclear speckles and can undergo LLPS *in vitro*. We confirmed the observation by Courchaine *et al*, 2021 that the Tudor domain is not sufficient for phase separation. Instead, the disordered C-terminal regions and the full-length protein could form droplets *in vitro*.

We also provided evidence that SMNDC1 phase separation is happening in cells. For example, we were able to change the spotted localization of SMNDC1 in the nucleus to a uniform distribution throughout the cell using 1,6-hexanediol. While 1,6-hexanediol has been used to disrupt separated phases (Kroschwald *et al*, 2017), its unspecific side-effects like inhibition of kinases and phosphatases have been debated as well (Düster *et al*, 2021). Fluorescence recovery after photobleaching (FRAP) measurements showed that SMNDC1 is highly mobile, another typical feature of proteins in liquid phases which discriminates them from insoluble protein aggregates.

Nevertheless, the cell is an extremely complex system and until now our understanding on what factors are necessary and can influence SMNDC1's phase separation behavior is limited. While the C-terminal region is sufficient for *in vitro* phase separation together with RNA, we have not tested the influence of the different parts of SMNDC1 for phase separation in cells. This would be particularly demanding as the nuclear localization signal is also located in the C-terminal region, and the regions lateral from the NLS are particularly disordered. Unraveling the contributions of certain parts of the protein without affecting its general localization within the cell might therefore be impossible.

Due to the high score of disorder prediction of the C-terminal region and a high RNA binding score (RNAbindRplus) (Walia *et al*, 2014) we hypothesize that this part of the protein could be responsible for the interactions of SMNDC1 with RNA. However, this still requires experimental evidence. In a classical immunoprecipitation (IP) experiment, we could show that a lot of interactions were dependent on the presence of RNA as shown in RNase treated samples (Casteels *et al*, 2022). We therefore suspect that RNA is very important as a mediator of SMNDC1's interactions in nuclear speckles.

Furthermore, we can only speculate how important SMNDC1 is for the formation of nuclear speckles in general, i.e., whether it can act as a "scaffold" for many other proteins and nucleic acids or whether it is a "client" localizing to the phase separated compartment through the interaction with other biomolecules (Banani *et al*, 2016). Inhibition of SMNDC1's Tudor domain

leads to a dissolution of SMNDC1 and SRRM2 and therefore of at least one other important scaffold of nuclear speckles (Ilik *et al*, 2020). Knock-down of SMNDC1 has a very similar effect on SRRM2. Nevertheless, we don't know how many other nuclear speckle proteins depend on the presence of SMNDC1 or SMNDC1's valency, so we can't definitively call it a scaffold for nuclear speckles.

Overall, we hypothesize that only SMNDC1's unstructured C-terminal region in combination with the Tudor domain can provide sufficient valency for the formation of a biomolecular condensate in cells (see also Figure 14).

3.1.2 Pharmacological perturbation of SMNDC1

Since SMNDC1 is an essential splicing factor (Rappsilber *et al*, 2001; Meister *et al*, 2001), a knock-out of the gene quickly leads to cell death. Knock-down (and overexpression) can be used to study the effects of SMNDC1 levels, but these tools are not well suited to study immediate effects. Very often, cells are able to react to such a perturbation with a plethora of counter-balancing measuring and return to a steady-state. This obstructs the dissection of direct from secondary effects.

To overcome these limitations and investigate the function of SMNDC1 in a dose- and time-dependent manner, we turned our attention to the protein's Tudor domain, which is well-structured and features a characteristic aromatic cage that allows for specific binding of dimethylarginine ligands. This characteristic presents a potential opportunity also for the recognition of inhibitory small molecules, also known as druggability, unlike the C-terminal region.

To identify small molecule inhibitors that bind specifically to SMNDC1, we set up an AlphaScreen assay combining its Tudor domain with a peptide containing four sDMAs. After validating this set-up, we screened a library of 90,000 compounds. Some of the hit structures displayed selectivity by binding only to SMNDC1 and not the SMN Tudor domain. Multiple rounds of testing analog molecules enabled us to establish the structure-activity relationships of these compounds, discussed in more detail below. It is worth noting that these compounds showed activity in cellular assays; however, high concentrations were required.

3.1.3 SMNDC1 inhibition – immediate and later effects

We then focused on the most promising inhibitor, which caused SMNDC1 to exit the nucleus and nuclear speckles. This effect could not be observed immediately after treatment, though. Instead, it takes about 12-16 hours until a clear difference in SMNDC1's localization is visible. To look for more prompt effects we turned to fluorescence recovery after photobleaching

(FRAP). When measuring SMNDC1 FRAP immediately (15-45 min) after inhibitor treatment we observed a reduced mobility compared to solvent. These observations suggest that inhibiting SMNDC1's Tudor domain impacts the architecture of nuclear speckles over time, albeit without immediately disrupting the existing speckles. This is because mobility and the incorporation of new proteins into the phase-separated compartment might be affected, whereas SMNDC1 protein already in nuclear speckles is not affected.

Furthermore, inhibitor treatment decreased SMNDC1's interaction with its binding partners identified by proximity labeling. We did not study SMNDC1's interaction with binding partners in a time-resolved manner, but we hypothesize that the interactions dependent on the Tudor domain-sDMA interaction would be the first ones lost upon inhibitor treatment. Secondary interactions occurring within nuclear speckles would be lost only at later time points when the overall architecture of nuclear speckles is disrupted.

Finally, we also observe effects of inhibitor treatment on splicing. While we only checked global splicing effects by RNA sequencing after 16 hour treatment, we confirmed most of the splicing changes with the biggest dPSI values also by PCR with a low dose – long term treatment (5 days at 2 µM vs. 16 hours at 50 µM). Some of the splicing PCRs showed a clearer difference upon low dose – long term compared to high dose – short term hinting that the effects on splicing might be the ones taking the longest time. This is easily imaginable since the half-life of certain mRNAs can extend to a time-scale of days (Sharova *et al*, 2009).

Overall, while we could obtain some first experimental data on the temporal order of events, further experiments at additional time-points could give interesting insights into the chronological and potentially causative order of events happening upon SMNDC1 small molecule inhibition.

3.1.4 Comparison of SMNDC1 inhibitors to other Tudor domain inhibitors

As described extensively in the introduction, a number of Tudor domains have been targeted by small molecule inhibitors (see 1.2.5 for a detailed description). In this section, I will compare the chemical structures and their binding mode to the inhibitors that we have found.

Starting with Spindlin 1, the KMT G9a inhibitor A366 was found to also inhibit Spindlin1 in an AlphaScreen, but a binding mode could not be established yet (Wagner *et al*, 2016). This changed for EML631 (Bae *et al*, 2017) binding Tudor domain I and partly II of Spindlin1 which was further developed into VinSpinIn (Fagan *et al*, 2019), which used a bidentate binding mode to also bind Tudor domain I and II simultaneously. This is a useful strategy for proteins with more than one binding pocket. Unfortunately, SMNDC1 only has one deep "druggable" binding pocket. MS31 uses similar interactions to the aromatic cage of Tudor domain II of Spindlin1 as our inhibitor fragment **13** to SMNDC1's Tudor domain, namely cation−π and π−π-

stacking (Xiong *et al*, 2019). Another compound employing π-stacking is NV01 which binds to UHRF1's TTD aromatic cage with its oxothieno-pyrrolo-triazine moiety (Senisterra *et al*, 2018). The last known compound engaging a Tudor domain through π-π interaction is structure "1" for TDRD3 (Liu *et al*, 2018). While different binding modes have been reported for Tudor domain inhibitors, π-π interactions of ring-systems in small molecule binders with aromatic residues in the Tudor domain pockets are a common pattern which is also shared by our inhibitor. Further interactions often occur via nitrogen atoms, mimicking the natural ligand arginine.

3.1.5 Identification of off-targets

Given the structural similarity of the different Tudor domains, especially between SMNDC1 and SMN, identifying selective compounds is crucial. We therefore established the AlphaScreen for SMN as well to be able to measure IC_{50} values for these two proteins in parallel. Due to the high conservation between these two Tudor domains (see also Figure 5A) we expect that compounds which achieve selectivity for SMNDC1 over SMN are also selective for SMNDC1 over other Tudor domains which are less conserved. Nevertheless, it is still important to check all other Tudor domain proteins and ideally the whole proteome for binding of the SMNDC1 inhibitors. This could be done with a panel of Tudor domain proteins as shown for SMN (Liu *et al*, 2022) or Spindlin1 (Bae *et al*, 2017). A proteome-wide approach could employ the derivatization of an identified inhibitor with a diazirine-alkyne handle (Thomas *et al*, 2017). Proteins bound by the derivatized inhibitor in cells or lysates are covalently linked via photo-crosslinking and extracted. The alkyne handle of the fragment is then conjugated to biotin-azide by means of click chemistry, which allows the ensuing enrichment of target proteins via streptavidin-pulldown. Afterwards, all binding proteins can be identified by mass spectrometry. An attachment of a diazirine-alkyne handle is not the only option for derivatization. Functional derivatizations will be discussed in the next section.

3.1.6 Further development and derivatization of the inhibitor

The inhibitors presented in the publication in section 2 are the first targeted towards SMNDC1 with IC_{50} values below 1 µM. Of course, a more potent inhibitor in the nanomolar range would be desirable. Still, the natural ligands aDMA and sDMA bind the SMNDC1's Tudor domain with K_d values of 1.706 mM and 1.317 mM, respectively (Tripsianes *et al*, 2011). Only a peptide containing four sDMA residues (the same that was also used for the AlphaScreen to identify inhibitors) showed a lower binding affinity of 0.024 mM/ 24 µM. A sub-micromolar inhibitor could therefore already lead to substantial inhibition of binding to the natural ligands.

Nevertheless, relatively high concentrations of up to 50 µM were necessary to observe effects *in cellulo*, immediately for SMNDC1 mobility and for 12 hours for SMNDC1 localization. A low dose of 2 µM elicited a difference in splicing only after 5 days. This could also be due to low permeability in cells which we did not check for. Overall, more analogues should be tested to further improve IC_{50} values and potentially cellular uptake.

Then again, we have identified the first compounds with selectivity for SMNDC1 over SMN. Compound **28** with a larger substituent instead of the arylsulfonamide showed a ~30-fold selectivity (IC_{50} SMNDC1: 0.50 µM, IC_{50} SMN: 15.35 µM). More and different substituent and further extensions here should be tested to potentially further improve selectivity for SMNDC1 over SMN.

Another approach towards improved or different cellular and subsequently *in vivo* activity is the development of heterobifunctional molecules. Best-studied are Proteolysis-Targeting Chimeric Molecules (PROTACs) for proteasomal degradation of the target protein (Békés *et al*, 2022), but a plethora of other functions can in principle be utilized by recruiting the respective cellular machinery to the target protein (Hua *et al*, 2022).

Since we have only limited knowledge of the contribution of the Tudor domain alone (and its inhibition) to SMNDC1's function, proteasomal degradation to lower the levels of the full SMNDC1 protein could have different consequences. Lower SMNDC1 protein levels would potentially mimic an SMNDC1 knock-down as used and studied extensively in Casteels *et al*, 2022. This could also render possible the induction of insulin expression in α-cells which we have not observed upon SMNDC1 small molecule inhibition yet and which leads to the next section – potential therapeutic applications of the inhibitors.

3.1.7 Potential therapeutic applications

Eventually, an inhibitor for SMNDC1 could not only be used to study SMNDC1's function in cellular or – potentially – animal models, but also for therapeutic applications. The most obvious application arises from our observation that SMNDC1 knock-down leads to insulin expression in pancreatic α-cells (Casteels *et al*, 2022). We could show that lower SMNDC1 levels lead to global changes in alternative splicing with more retained introns and skipped exons. One of the affected transcripts is the chromatin remodeler ATRX. Its mRNA and protein levels are reduced upon SMNDC1 knock-down due to the introduction of a poison exon. Consequently, we observed an upregulation of pluripotency and β-cell genes like the transcription factor Pdx1. This transdifferentiation of α-cells towards a β-cell like identity eventually leads to an upregulation of insulin expression (Figure 5 C+D).

This effect could be used for the therapy of diabetes where β-cells are lost or lose their function. Targeting SMNDC1 could replenish the pool of functional β-cells. While genetic

therapies are potentially irreversible and still very difficult to implement a small molecule inhibitor would be much easier to deliver to patients.

Until now, we could not detect an upregulation of insulin when treating α-cells with the SMNDC1 inhibitor. This might be explained by a contrary role of SMN in the regulation of insulin expression. In a murine model of SMA, mutations in SMN led to a shift in the composition of pancreatic islets towards more glucagon- and less insulin-expressing cells (Bowerman *et al*, 2012, 2014)(Figure 12). SMN could therefore have the opposite effect on cell identity and insulin expression compared to SMNDC1. If our SMNDC1 inhibitors still inhibit SMN as well, this could explain why we cannot replicate the effects of SMNDC1 knock-down with the inhibitors.

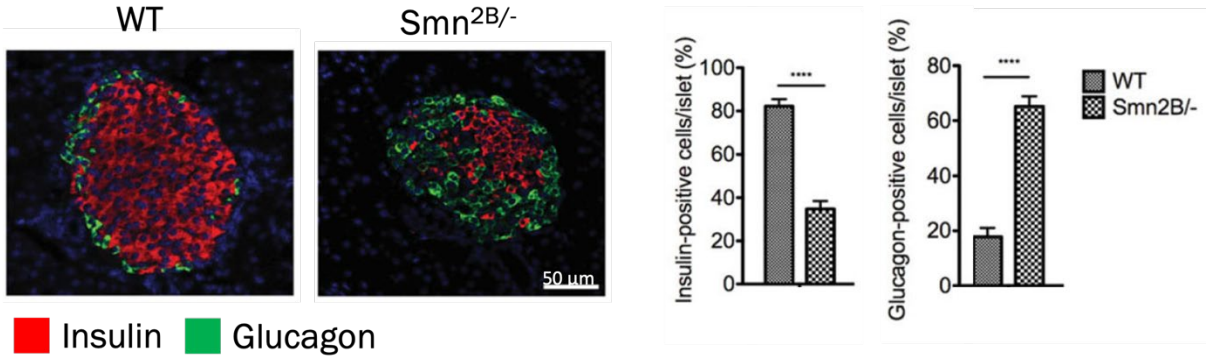


Figure 12: A mouse model of SMA shows changes in the cell type composition of pancreatic islets.

Adapted from Bowerman *et al*, 2012. Pancreatic islet cross-sections of wild-type (WT) and *Smn2B/-* mice. *Smn2B/-* is an intermediate SMA mouse model which has a recombined allele harboring mutations within exon 7 of the endogenous Smn gene (Bowerman *et al*, 2009). Sections were taken at postnatal (P) day 21. To the right percentages of insulin/ glucagon positive cells. Red: Insulin. Green: Glucagon. Blue: DAPI. *Smn2B/-* mice have more glucagon- and less insulin-positive cells.

Another application could be the treatment of hepatocellular carcinoma. A recent publication could show that high SMNDC1 expression is associated with worse survival in patients suffering from hepatocellular carcinoma (Zhu *et al*, 2023). In addition, the authors could show that in hepatocellular carcinoma cell lines, a knock-down of SMNDC1 reduces proliferation and migration.

Nevertheless, SMNDC1's essentiality might pose a challenge for a systemic administration of an inhibitor. SMNDC1 knock-out, long-term knock-down and long-term small molecule inhibition have proven to be toxic to cells in culture. Finding a therapeutic window could

therefore be challenging. Furthermore, SMNDC1 is ubiquitously expressed throughout the human body as shown by a tissue-wide comparison of expression levels from the Human Cell Atlas (Uhlén *et al*, 2015) (Figure 13). Its mRNA could be identified in all analyzed tissues. Pancreas and liver as the two tissues which are the most interesting for a therapeutic application do not appear as tissues with a particularly high expression.

Still, there are possibilities of specifically targeting tissues and cell types by employing the enrichment of certain elements for example, as shown for Zinc-containing molecules and β-cells (Horton *et al*, 2018). Another alternative for cell-type specific targeting are drug-antibody conjugates in which a monoclonal antibody is connected to a small molecule payload with a linker (Pettinato, 2021). Cell-type specific cell surface markers like TM4SF4 for α-cells (Muraro *et al*, 2016) could be exploited to deliver an inhibitor only to the target cells.

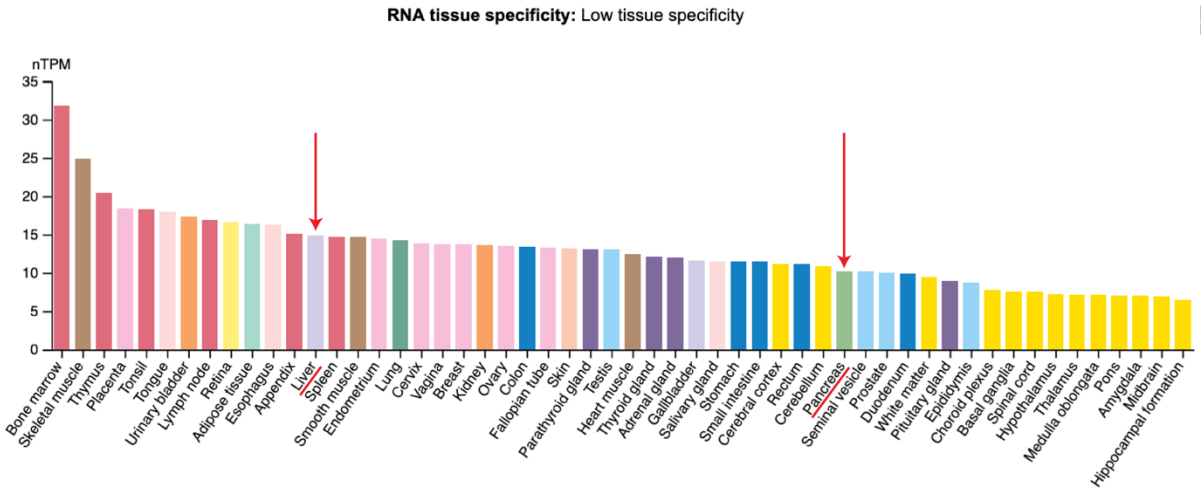


Figure 13: SMNDC1 expression compared across human tissues.
Analyzed on the level of RNA, sorted by expression level. Liver and pancreas marked with a red arrow. Modified from https://www.proteinatlas.org/ENSG00000119953-SMNDC1/tissue (Uhlén *et al*, 2015).

3.2 Conclusion and Future Prospects

To visualize the findings described in this thesis and in the publication in section 2, I have created a graphical model (Figure 14).

In summary, I have shown that SMNDC1 localizes to nuclear speckles and interacts with other known nuclear speckle proteins. It can undergo phase separation both *in vitro* and *in cellulo*. After screening for a small molecule inhibitor directed against SMNDC1's Tudor domain we conducted extensive SAR studies with the most promising hits. Furthermore, we elucidated the binding mode of a soluble fragment (**13**) of the best inhibitor (**1**) through NMR experiments

in collaboration with Jan Borggräfe, Stefan Gaussmann and Michael Sattler from the Technical University of Munich.

We then went on to test the effects of inhibitor **1** in cells. The first observation upon inhibitor treatment was a dilution of SMNDC1 away from nuclear speckles into nucleoplasm and cytoplasm. A similar behavior was detected for the nuclear speckle marker SRRM2. Furthermore, SMNDC1's interactions with other proteins were lost, especially to proteins known to carry the sDMA post-translational modification and to nuclear speckle proteins. Finally, inhibiting SMNDC1's Tudor domain led to global changes in alternative splicing patterns.

The next steps in developing the inhibitor further will include testing new analogues to improve selectivity and activity, especially in cells. Derivatizations like PROTACs might help in this regard as well. To rule out off-targets, global chemoproteomics might be necessary.

Eventually, these efforts might lead to an inhibitor that increases the insulin expression of treated α-cells. In case this goal was achieved a logical next step would be trials in animal models of diabetes.

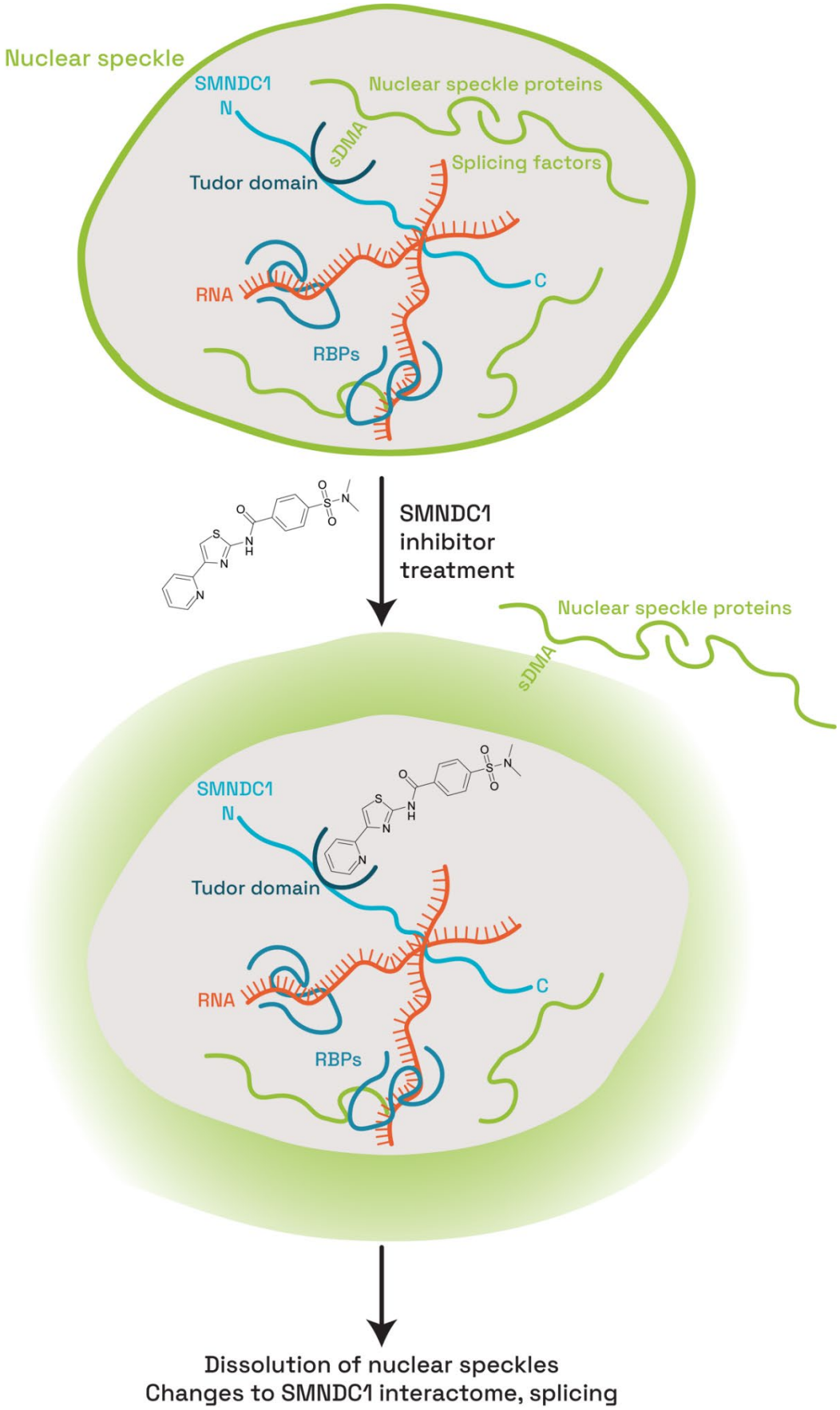


Figure 14: A model for the mechanisms behind SMNDC1's phase separation and its pharmacological perturbation

4. MATERIALS AND METHODS

Corresponding materials and methods are described in detail within the published manuscript in Section 2.

5. REFERENCES

Acsadi G, Lee I, Li X, Khaidakov M, Pecinova A, Parker GC & Hüttemann M (2009) Mitochondrial dysfunction in a neural cell model of spinal muscular atrophy. *J Neurosci Res* 87: 2748–2756

Afroz T, Pérez-Berlanga M & Polymenidou M (2019) Structural Transition, Function and Dysfunction of TDP-43 in Neurodegenerative Diseases. *CHIMIA* 73: 380–380

Ahn JH, Davis ES, Daugird TA, Zhao S, Quiroga IY, Uryu H, Li J, Storey AJ, Tsai Y-H, Keeley DP, *et al* (2021) Phase separation drives aberrant chromatin looping and cancer development. *Nature*

Aho ER, Wang J, Gogliotti RD, Howard GC, Phan J, Acharya P, Macdonald JD, Cheng K, Lorey SL, Lu B, *et al* (2019) Displacement of WDR5 from Chromatin by a WIN Site Inhibitor with Picomolar Affinity. *Cell Rep* 26: 2916-2928.e13

Ainani H, Bouchmaa N, Ben Mrid R & El Fatimy R (2023) Liquid-liquid phase separation of protein tau: An emerging process in Alzheimer's disease pathogenesis. *Neurobiology of Disease* 178: 106011

Akiba K, Katoh-Fukui Y, Yoshida K, Narumi S, Miyado M, Hasegawa Y & Fukami M (2021) Role of Liquid-Liquid Separation in Endocrine and Living Cells. *J Endocr Soc* 5: bvab126

Alberti S & Dormann D (2019) Liquid-Liquid Phase Separation in Disease. *Annu Rev Genet* 53: 171–194

Alberti S & Hyman AA (2016) Are aberrant phase transitions a driver of cellular aging? *BioEssays* 38: 959–968

Amato A, Lucas X, Bortoluzzi A, Wright D & Ciulli A (2018) Targeting Ligandable Pockets on Plant Homeodomain (PHD) Zinc Finger Domains by a Fragment-Based Approach. *ACS Chem Biol* 13: 915–921

Ambler RP & Rees MW (1959) Epsilon-N-Methyl-lysine in bacterial flagellar protein. *Nature* 184: 56–57

Anbo H, Sato M, Okoshi A & Fukuchi S (2019) Functional Segments on Intrinsically Disordered Regions in Disease-Related Proteins. *Biomolecules* 9

Arribas-Layton M, Dennis J, Bennett EJ, Damgaard CK & Lykke-Andersen J (2016) The C-Terminal RGG Domain of Human Lsm4 Promotes Processing Body Formation Stimulated by Arginine Dimethylation. *Molecular and Cellular Biology* 36: 2226–2235

Arrowsmith CH, Bountra C, Fish PV, Lee K & Schapira M (2012) Epigenetic protein families: a new frontier for drug discovery. *Nat Rev Drug Discov* 11: 384–400

Arrowsmith CH & Schapira M (2019) Targeting non-bromodomain chromatin readers. *Nat Struct Mol Biol* 26: 863–869

Aumiller WM & Keating CD (2016) Phosphorylation-mediated RNA/peptide complex coacervation as a model for intracellular liquid organelles. *Nat Chem* 8: 129–137

Avşar Ilık İ & Aktaş T (2021) Nuclear speckles: dynamic hubs of gene expression regulation. *FEBS J*

Bae N, Viviano M, Su X, Lv J, Cheng D, Sagum C, Castellano S, Bai X, Johnson C, Khalil MI, *et al* (2017) Developing Spindlin1 Small Molecule Inhibitors Using Protein Microarrays. *Nat Chem Biol* 13: 750–756

Baertschi S, Baur N, Lueders-Lefevre V, Voshol J & Keller H (2014) Class I and IIa histone deacetylases have opposite effects on sclerostin gene regulation. *J Biol Chem* 289: 24995–25009

Ball LJ, Murzina NV, Broadhurst RW, Raine ARC, Archer SJ, Stott FJ, Murzin AG, Singh PB, Domaille PJ & Laue ED (1997) Structure of the chromatin binding (chromo) domain from mouse modifier protein 1. *The EMBO Journal* 16: 2473–2481

Banani SF, Rice AM, Peeples WB, Lin Y, Jain S, Parker R & Rosen MK (2016) Compositional Control of Phase-Separated Cellular Bodies. *Cell* 166: 651–663

Banerjee PR, Milin AN, Moosa MM, Onuchic PL & Deniz AA (2017) Reentrant Phase Transition Drives Dynamic Substructure Formation in Ribonucleoprotein Droplets. *Angewandte Chemie International Edition* 56: 11354–11359

Barski A, Cuddapah S, Cui K, Roh T-Y, Schones DE, Wang Z, Wei G, Chepelev I & Zhao K (2007) High-Resolution Profiling of Histone Methylations in the Human Genome. *Cell* 129: 823–837

Beck JS (1961) VARIATIONS IN THE MORPHOLOGICAL PATTERNS OF 'AUTOIMMUNE' NUCLEAR FLUORESCENCE. *The Lancet* 277: 1203–1205

Békés M, Langley DR & Crews CM (2022) PROTAC targeted protein degraders: the past is prologue. *Nat Rev Drug Discov* 21: 181–200

Bernard D, Martinez-Leal JF, Rizzo S, Martinez D, Hudson D, Visakorpi T, Peters G, Carnero A, Beach D & Gil J (2005) CBX7 controls the growth of normal and tumor-derived prostate cells by repressing the Ink4a/Arf locus. *Oncogene* 24: 5543–5551

Bernard P, Drogat J, Dheur S, Genier S & Javerzat J-P (2010) Splicing Factor Spf30 Assists Exosome-Mediated Gene Silencing in Fission Yeast. *Mol Cell Biol* 30: 1145–1157

Biggar KK & Li SS-C (2015) Non-histone protein methylation as a regulator of cellular signalling and function. *Nat Rev Mol Cell Biol* 16: 5–17

Binda O, Ishimwe ABK, Galloy M, Jacquet K, Corpet A, Fradet-Turcotte A, Côté J & Lomonte P (2023a) The TUDOR domain of SMN is an H3K79me1 histone mark reader. *Life Science Alliance* 6

Binda O, Juillard F, Ducassou JN, Kleijwegt C, Paris G, Didillon A, Baklouti F, Corpet A, Couté Y, Côté J, *et al* (2023b) SMA-linked SMN mutants prevent phase separation properties and SMN interactions with FMRP family members. *Life Science Alliance* 6

Blanc RS & Richard S (2017) Arginine Methylation: The Coming of Age. *Mol Cell* 65: 8–24

Boehning M, Dugast-Darzacq C, Rankovic M, Hansen AS, Yu T, Marie-Nelly H, McSwiggen DT, Kokic G, Dailey GM, Cramer P, *et al* (2018) RNA polymerase II clustering through carboxy-terminal domain phase separation. *Nat Struct Mol Biol* 25: 833–840

Boeynaems S, Alberti S, Fawzi NL, Mittag T, Polymenidou M, Rousseau F, Schymkowitz J, Shorter J, Wolozin B, Bosch LVD, *et al* (2018) Protein Phase Separation: A New Phase in Cell Biology. *Trends in Cell Biology* 28: 420–435

Boija A, Klein IA, Sabari BR, Dall'Agnese A, Coffey EL, Zamudio AV, Li CH, Shrinivas K, Manteiga JC, Hannett NM, *et al* (2018) Transcription Factors Activate Genes through the Phase-Separation Capacity of Their Activation Domains. *Cell* 175: 1842-1855.e16

Bokar JA, Shambaugh ME, Polayes D, Matera AG & Rottman FM (1997) Purification and cDNA cloning of the AdoMet-binding subunit of the human mRNA (N6-adenosine)-methyltransferase. *RNA* 3: 1233–1247

Bolshan Y, Getlik M, Kuznetsova E, Wasney GA, Hajian T, Poda G, Nguyen KT, Wu H, Dombrovski L, Dong A, *et al* (2013) Synthesis, Optimization, and Evaluation of Novel Small Molecules as Antagonists of WDR5-MLL Interaction. *ACS Med Chem Lett* 4: 353–357

Bonasio R, Lecona E & Reinberg D (2010) MBT domain proteins in development and disease. *Seminars in Cell & Developmental Biology* 21: 221–230

Bostick M, Kim JK, Estève P-O, Clark A, Pradhan S & Jacobsen SE (2007) UHRF1 Plays a Role in Maintaining DNA Methylation in Mammalian Cells. *Science* 317: 1760–1764

Boswell RE & Mahowald AP (1985) tudor, a gene required for assembly of the germ plasm in Drosophila melanogaster. *Cell* 43: 97–104

Böttcher J, Dilworth D, Reiser U, Neumüller RA, Schleicher M, Petronczki M, Zeeb M, Mischerikow N, Allali-Hassani A, Szewczyk MM, *et al* (2019) Fragment-based discovery of a chemical probe for the PWWP1 domain of NSD3. *Nat Chem Biol* 15: 822–829

Botuyan MV, Lee J, Ward IM, Kim J-E, Thompson JR, Chen J & Mer G (2006) Structural Basis for the Methylation State-Specific Recognition of Histone H4-K20 by 53BP1 and Crb2 in DNA Repair. *Cell* 127: 1361–1373

Bouchard JJ, Otero JH, Scott DC, Szulc E, Martin EW, Sabri N, Granata D, Marzahn MR, Lindorff-Larsen K, Salvatella X, *et al* (2018) Cancer Mutations of the Tumor Suppressor SPOP Disrupt the Formation of Active, Phase-Separated Compartments. *Mol Cell* 72: 19-36.e8

Boulay G, Sandoval GJ, Riggi N, Iyer S, Buisson R, Naigles B, Awad ME, Rengarajan S, Volorio A, McBride MJ, *et al* (2017) Cancer-Specific Retargeting of BAF Complexes by a Prion-like Domain. *Cell* 171: 163-178.e19

Bowerman M, Anderson CL, Beauvais A, Boyl PP, Witke W & Kothary R (2009) SMN, profilin IIa and plastin 3: a link between the deregulation of actin dynamics and SMA pathogenesis. *Mol Cell Neurosci* 42: 66–74

Bowerman M, Michalski J-P, Beauvais A, Murray LM, DeRepentigny Y & Kothary R (2014) Defects in pancreatic development and glucose metabolism in SMN-depleted mice independent of canonical spinal muscular atrophy neuromuscular pathology. *Hum Mol Genet* 23: 3432–3444

Bowerman M, Swoboda KJ, Michalski J-P, Wang G-S, Reeks C, Beauvais A, Murphy K, Woulfe J, Screaton RA, Scott FW, *et al* (2012) Glucose metabolism and pancreatic defects in spinal muscular atrophy. *Annals of Neurology* 72: 256–268

Boyko S, Qi X, Chen T-H, Surewicz K & Surewicz WK (2019) Liquid–liquid phase separation of tau protein: The crucial role of electrostatic interactions. *Journal of Biological Chemistry* 294: 11054–11059

Brahms H, Meheus L, Brabandere VD, Fischer U & Lührmann R (2001) Symmetrical dimethylation of arginine residues in spliceosomal Sm protein B/B′ and the Sm-like protein LSm4, and their interaction with the SMN protein. *RNA* 7: 1531–1542

Brangwynne CP, Eckmann CR, Courson DS, Rybarska A, Hoege C, Gharakhani J, Jülicher F & Hyman AA (2009) Germline P Granules Are Liquid Droplets That Localize by Controlled Dissolution/Condensation. *Science* 324: 1729–1732

Brangwynne CP, Mitchison TJ & Hyman AA (2011) Active liquid-like behavior of nucleoli determines their size and shape in Xenopus laevis oocytes. *Proc Natl Acad Sci USA* 108: 4334–4339

Branscombe TL, Frankel A, Lee JH, Cook JR, Yang Z, Pestka S & Clarke S (2001) PRMT5 (Janus kinase-binding protein 1) catalyzes the formation of symmetric dimethylarginine residues in proteins. *J Biol Chem* 276: 32971–32976

Braunschweig U, Gueroussov S, Plocik AM, Graveley BR & Blencowe BJ (2013) Dynamic Integration of Splicing within Gene Regulatory Pathways. *Cell* 152: 1252–1269

Brown JM, Green J, das Neves RP, Wallace HAC, Smith AJH, Hughes J, Gray N, Taylor S, Wood WG, Higgs DR, *et al* (2008) Association between active genes occurs at nuclear speckles and is modulated by chromatin environment. *J Cell Biol* 182: 1083–1097

Callebaut I & Mornon JP (1997) The human EBNA-2 coactivator p100: multidomain organization and relationship to the staphylococcal nuclease fold and to the tudor protein involved in Drosophila melanogaster development. *Biochem J* 321: 125–132

Camerino MA, Zhong N, Dong A, Dickson BM, James LI, Baughman BM, Norris JL, Kireev DB, Janzen WP, Arrowsmith CH, *et al* (2013) The structure-activity relationships of L3MBTL3 inhibitors: flexibility of the dimer interface. *Medchemcomm* 4: 1501–1507

Casanova M, Preissner T, Cerase A, Poot R, Yamada D, Li X, Appanah R, Bezstarosti K, Demmers J, Koseki H, *et al* (2011) Polycomblike 2 facilitates the recruitment of PRC2 Polycomb group complexes to the inactive X chromosome and to target loci in embryonic stem cells. *Development* 138: 1471–1482

Casteels T, Bajew S, Reiniš J, Enders L, Schuster M, Fontaine F, Müller AC, Wagner BK, Bock C & Kubicek S (2022) SMNDC1 links chromatin remodeling and splicing to regulate pancreatic hormone expression. *Cell Reports* 40: 111288

Casteels T, Zhang Y, Frogne T, Sturtzel C, Lardeau C-H, Sen I, Liu X, Hong S, Pauler FM, Penz T, *et al* (2021) An inhibitor-mediated beta cell dedifferentiation model reveals distinct roles for FoxO1 in glucagon repression and insulin maturation. *Molecular Metabolism*: 101329

Caudy AA, Ketting RF, Hammond SM, Denli AM, Bathoorn AMP, Tops BBJ, Silva JM, Myers MM, Hannon GJ & Plasterk RHA (2003) A micrococcal nuclease homologue in RNAi effector complexes. *Nature* 425: 411–414

Ceol CJ, Houvras Y, Jane-Valbuena J, Bilodeau S, Orlando DA, Battisti V, Fritsch L, Lin WM, Hollmann TJ, Ferré F, *et al* (2011) The histone methyltransferase SETDB1 is recurrently amplified in melanoma and accelerates its onset. *Nature* 471: 513–517

Chakraborty S, Senyuk V, Sitailo S, Chi Y & Nucifora G (2001) Interaction of EVI1 with cAMP-responsive element-binding protein-binding protein (CBP) and p300/CBP-associated factor (P/CAF) results in reversible acetylation of EVI1 and in co-localization in nuclear speckles. *J Biol Chem* 276: 44936–44943

Chang L, Campbell J, Raji IO, Guduru SKR, Kandel P, Nguyen M, Liu S, Tran K, Venugopal NK, Taylor BC, *et al* (2021) Discovery of small molecules targeting the tandem tudor domain of the epigenetic factor UHRF1 using fragment-based ligand discovery. *Sci Rep* 11

Chaytow H, Huang Y-T, Gillingwater TH & Faller KME (2018) The role of survival motor neuron protein (SMN) in protein homeostasis. *Cell Mol Life Sci* 75: 3877–3894

Chen G, Wang D, Wu B, Yan F, Xue H, Wang Q, Quan S & Chen Y (2020a) Taf14 recognizes a common motif in transcriptional machineries and facilitates their clustering by phase separation. *Nat Commun* 11: 4206

Chen X, Wu X, Wu H & Zhang M (2020b) Phase separation at the synapse. *Nat Neurosci* 23: 301–310

Chen Y, Zhang Y, Wang Y, Zhang L, Brinkman EK, Adam SA, Goldman R, van Steensel B, Ma J & Belmont AS (2018) Mapping 3D genome organization relative to nuclear compartments using TSA-Seq as a cytological ruler. *Journal of Cell Biology* 217: 4025–4048

Chen Z, Huai Y, Mao W, Wang X, Ru K, Qian A & Yang H (2022) Liquid-Liquid Phase Separation of Biomacromolecules and Its Roles in Metabolic Diseases. *Cells* 11: 3023

Chitiprolu M, Jagow C, Tremblay V, Bondy-Chorney E, Paris G, Savard A, Palidwor G, Barry FA, Zinman L, Keith J, *et al* (2018) A complex of C9ORF72 and p62 uses arginine methylation to eliminate stress granules by autophagy. *Nature Communications* 9: 2794

Cho W-K, Spille J-H, Hecht M, Lee C, Li C, Grube V & Cisse II (2018) Mediator and RNA polymerase II clusters associate in transcription-dependent condensates. *Science* 361: 412–415

Chong PA, Vernon RM & Forman-Kay JD (2018a) RGG/RG Motif Regions in RNA Binding and Phase Separation. *Journal of Molecular Biology* 430: 4650–4665

Chong S, Dugast-Darzacq C, Liu Z, Dong P, Dailey GM, Cattoglio C, Heckert A, Banala S, Lavis L, Darzacq X, *et al* (2018b) Imaging dynamic and selective low-complexity domain interactions that control gene transcription. *Science* 361

Cloos PAC, Christensen J, Agger K, Maiolica A, Rappsilber J, Antal T, Hansen KH & Helin K (2006) The putative oncogene GASC1 demethylates tri- and dimethylated lysine 9 on histone H3. *Nature* 442: 307–311

Conti BA & Oppikofer M (2022) Biomolecular condensates: new opportunities for drug discovery and RNA therapeutics. *Trends in Pharmacological Sciences*

Copeland AM, Altamura LA, Van Deusen NM & Schmaljohn CS (2013) Nuclear relocalization of polyadenylate binding protein during rift valley fever virus infection involves expression of the NSs gene. *J Virol* 87: 11659–11669

Courchaine EM, Barentine AES, Straube K, Lee D-R, Bewersdorf J & Neugebauer KM (2021) DMA-tudor interaction modules control the specificity of in vivo condensates. *Cell* 184: 3612-3625.e17

Cuinat S, Nizon M, Isidor B, Stegmann A, van Jaarsveld RH, van Gassen KL, van der Smagt JJ, Volker-Touw CML, Holwerda SJB, Terhal PA, *et al* (2022) Loss-of-function variants in SRRM2 cause a neurodevelopmental disorder. *Genet Med* 24: 1774–1780

Cuthbert GL, Daujat S, Snowden AW, Erdjument-Bromage H, Hagiwara T, Yamada M, Schneider R, Gregory PD, Tempst P, Bannister AJ, *et al* (2004) Histone Deimination Antagonizes Arginine Methylation. *Cell* 118: 545–553

Dale B, Cheng M, Park K-S, Kaniskan HÜ, Xiong Y & Jin J (2021) Advancing targeted protein degradation for cancer therapy. *Nat Rev Cancer* 21: 638–654

Deng L-W, Chiu I & Strominger JL (2004) MLL 5 protein forms intranuclear foci, and overexpression inhibits cell cycle progression. *Proc Natl Acad Sci U S A* 101: 757–762

Dilworth D, Hanley RP, Ferreira de Freitas R, Allali-Hassani A, Zhou M, Mehta N, Marunde MR, Ackloo S, Carvalho Machado RA, Khalili Yazdi A, *et al* (2021) A chemical probe targeting the PWWP domain alters NSD2 nucleolar localization. *Nat Chem Biol*

Dimitriadi M, Derdowski A, Kalloo G, Maginnis MS, O'Hern P, Bliska B, Sorkaç A, Nguyen KCQ, Cook SJ, Poulogiannis G, *et al* (2016) Decreased function of survival motor neuron protein impairs endocytic pathways. *Proc Natl Acad Sci U S A* 113: E4377-4386

Ditlev JA, Vega AR, Köster DV, Su X, Tani T, Lakoduk AM, Vale RD, Mayor S, Jaqaman K & Rosen MK (2019) A composition-dependent molecular clutch between T cell signaling condensates and actin. *Elife* 8: e42695

Dominissini D, Moshitch-Moshkovitz S, Schwartz S, Salmon-Divon M, Ungar L, Osenberg S, Cesarkas K, Jacob-Hirsch J, Amariglio N, Kupiec M, *et al* (2012) Topology of the human and mouse m6A RNA methylomes revealed by m6A-seq. *Nature* 485: 201–206

Düster R, Kaltheuner IH, Schmitz M & Geyer M (2021) 1,6-Hexanediol, commonly used to dissolve liquid-liquid phase separated condensates, directly impairs kinase and phosphatase activities. *J Biol Chem* 296: 100260

Eissenberg JC (2012) Structural biology of the chromodomain: Form and function. *Gene* 496: 69–78

Engelberg IA, Liu J, Norris-Drouin JL, Cholensky SH, Ottavi SA, Frye SV, Kutateladze TG & James LI (2021) Discovery of an H3K36me3-Derived Peptidomimetic Ligand with Enhanced Affinity for Plant Homeodomain Finger Protein 1 (PHF1). *J Med Chem*

Erdel F & Rippe K (2018) Formation of Chromatin Subcompartments by Phase Separation. *Biophys J* 114: 2262–2270

Fagan V, Johansson C, Gileadi C, Monteiro O, Dunford JE, Nibhani R, Philpott M, Malzahn J, Wells G, Faram R, *et al* (2019) A Chemical Probe for Tudor Domain Protein Spindlin1 to Investigate Chromatin Function. *J Med Chem* 62: 9008–9025

Fallini C, Donlin-Asp PG, Rouanet JP, Bassell GJ & Rossoll W (2016) Deficiency of the Survival of Motor Neuron Protein Impairs mRNA Localization and Local Translation in the Growth Cone of Motor Neurons. *J Neurosci* 36: 3811–3820

Fang X, Wang L, Ishikawa R, Li Y, Fiedler M, Liu F, Calder G, Rowan B, Weigel D, Li P, *et al* (2019) Arabidopsis FLL2 promotes liquid-liquid phase separation of polyadenylation complexes. *Nature* 569: 265–269

Feng Y, Maity R, Whitelegge JP, Hadjikyriacou A, Li Z, Zurita-Lopez C, Al-Hadid Q, Clark AT, Bedford MT, Masson J-Y, *et al* (2013) Mammalian protein arginine methyltransferase 7 (PRMT7) specifically targets RXR sites in lysine- and arginine-rich regions. *J Biol Chem* 288: 37010–37025

Feric M, Vaidya N, Harmon TS, Mitrea DM, Zhu L, Richardson TM, Kriwacki RW, Pappu RV & Brangwynne CP (2016) Coexisting Liquid Phases Underlie Nucleolar Subcompartments. *Cell* 165: 1686–1697

Ferreira de Freitas R, Liu Y, Szewczyk MM, Mehta N, Li F, McLeod D, Zepeda-Velázquez C, Dilworth D, Hanley RP, Gibson E, *et al* (2021) Discovery of Small-Molecule Antagonists of the PWWP Domain of NSD2. *J Med Chem*

Fischer U, Liu Q & Dreyfuss G (1997) The SMN–SIP1 Complex Has an Essential Role in Spliceosomal snRNP Biogenesis. *Cell* 90: 1023–1029

FitzGerald JE, Grenon M & Lowndes NF (2009) 53BP1: function and mechanisms of focal recruitment. *Biochem Soc Trans* 37: 897–904

Fodor BD, Kubicek S, Yonezawa M, O'Sullivan RJ, Sengupta R, Perez-Burgos L, Opravil S, Mechtler K, Schotta G & Jenuwein T (2006) Jmjd2b antagonizes H3K9 trimethylation at pericentric heterochromatin in mammalian cells. *Genes Dev* 20: 1557–1562

Fox AH, Nakagawa S, Hirose T & Bond CS (2018) Paraspeckles: Where Long Noncoding RNA Meets Phase Separation. *Trends in Biochemical Sciences* 43: 124–135

Frazer C, Staples MI, Kim Y, Hirakawa M, Dowell MA, Johnson NV, Hernday AD, Ryan VH, Fawzi NL, Finkelstein IJ, *et al* (2020) Epigenetic cell fate in Candida albicans is controlled by transcription factor condensates acting at super-enhancer-like elements. *Nat Microbiol* 5: 1374–1389

Frege T & Uversky VN (2015) Intrinsically disordered proteins in the nucleus of human cells. *Biochem Biophys Rep* 1: 33–51

Friberg A, Corsini L, Mourão A & Sattler M (2009) Structure and Ligand Binding of the Extended Tudor Domain of D. melanogaster Tudor-SN. *Journal of Molecular Biology* 387: 921–934

Friesen WJ, Massenet S, Paushkin S, Wyce A & Dreyfuss G (2001) SMN, the Product of the Spinal Muscular Atrophy Gene, Binds Preferentially to Dimethylarginine-Containing Protein Targets. *Molecular Cell* 7: 1111–1117

Frottin F, Schueder F, Tiwary S, Gupta R, Körner R, Schlichthaerle T, Cox J, Jungmann R, Hartl FU & Hipp MS (2019) The nucleolus functions as a phase-separated protein quality control compartment. *Science* 365: 342–347

Fu Y & Zhuang X (2020) m6A-binding YTHDF proteins promote stress granule formation - Nature Chemical Biology. *Nature Chemical Biology* 16: 955–963

Galganski L, Urbanek MO & Krzyzosiak WJ (2017) Nuclear speckles: molecular organization, biological function and role in disease. *Nucleic Acids Res* 45: 10350–10368

Gallego LD, Schneider M, Mittal C, Romanauska A, Carrillo RMG, Schubert T, Pugh BF & Köhler A (2020) Phase separation directs ubiquitination of gene-body nucleosomes. *Nature* 579: 592–597

Gao X, Ge L, Shao J, Su C, Zhao H, Saarikettu J, Yao X, Yao Z, Silvennoinen O & Yang J (2010) Tudor-SN interacts with and co-localizes with G3BP in stress granules under stress conditions. *FEBS Lett* 584: 3525–3532

Gao XK, Rao XS, Cong XX, Sheng ZK, Sun YT, Xu SB, Wang JF, Liang YH, Lu LR, Ouyang H, *et al* (2022) Phase separation of insulin receptor substrate 1 drives the formation of insulin/IGF-1 signalosomes. *Cell Discov* 8: 60

Gao Y, Pei G, Li D, Li R, Shao Y, Zhang QC & Li P (2019) Multivalent m6A motifs promote phase separation of YTHDF proteins. *Cell Res* 29: 767–769

Garcia-Jove Navarro M, Kashida S, Chouaib R, Souquere S, Pierron G, Weil D & Gueroui Z (2019) RNA is a critical element for the sizing and the composition of phase-separated RNA-protein condensates. *Nat Commun* 10: 3230

Gayatri S & Bedford MT (2014) Readers of histone methylarginine marks. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms* 1839: 702–710

Getlik M, Smil D, Zepeda-Velázquez C, Bolshan Y, Poda G, Wu H, Dong A, Kuznetsova E, Marcellus R, Senisterra G, *et al* (2016) Structure-Based Optimization of a Small Molecule Antagonist of the Interaction Between WD Repeat-Containing Protein 5 (WDR5) and Mixed-Lineage Leukemia 1 (MLL1). *J Med Chem* 59: 2478–2496

Giavazzi A, Setola V, Simonati A & Battaglia G (2006) Neuronal-specific roles of the survival motor neuron protein: evidence from survival motor neuron expression patterns in the developing human central nervous system. *J Neuropathol Exp Neurol* 65: 267–277

Goulet I, Boisvenue S, Mokas S, Mazroui R & Côté J (2008) TDRD3, a novel Tudor domain-containing protein, localizes to cytoplasmic stress granules. *Hum Mol Genet* 17: 3055–3074

Grebien F, Vedadi M, Getlik M, Giambruno R, Grover A, Avellino R, Skucha A, Vittori S, Kuznetsova E, Smil D, *et al* (2015) Pharmacological targeting of the Wdr5-MLL interaction in C/EBPα N-terminal leukemia. *Nat Chem Biol* 11: 571–578

Green MR (2005) Eukaryotic transcription activation: right on target. *Mol Cell* 18: 399–402

Gu C, Yin Z, Nie H, Liu Y, Yang J, Huang G, Shen J, Chen L & Fei J (2020) Identification of berberine as a novel drug for the treatment of multiple myeloma via targeting UHRF1. *BMC Biol* 18: 33

Gueroussov S, Weatheritt RJ, O'Hanlon D, Lin Z-Y, Narula A, Gingras A-C & Blencowe BJ (2017) Regulatory Expansion in Mammals of Multivalent hnRNP Assemblies that Globally Control Alternative Splicing. *Cell* 170: 324-339.e23

Guilhas B, Walter J-C, Rech J, David G, Walliser NO, Palmeri J, Mathieu-Demaziere C, Parmeggiani A, Bouet J-Y, Le Gall A, *et al* (2020) ATP-Driven Separation of Liquid Phase Condensates in Bacteria. *Mol Cell* 79: 293-303.e4

Guillén-Boixet J, Kopach A, Holehouse AS, Wittmann S, Jahnel M, Schlüßler R, Kim K, Trussina IREA, Wang J, Mateju D, *et al* (2020) RNA-Induced Conformational Switching and Clustering of G3BP Drive Stress Granule Assembly by Condensation. *Cell* 181: 346-361.e17

Guo YE, Manteiga JC, Henninger JE, Sabari BR, Dall'Agnese A, Hannett NM, Spille J-H, Afeyan LK, Zamudio AV, Shrinivas K, *et al* (2019) Pol II phosphorylation regulates a switch between transcriptional and splicing condensates. *Nature* 572: 543–548

Guthmann M, Burton A & Torres-Padilla M (2019) Expression and phase separation potential of heterochromatin proteins during early mouse development. *EMBO Rep* 20

Hanahan D (2022) Hallmarks of Cancer: New Dimensions. *Cancer Discov* 12: 31–46

Hanahan D & Weinberg RA (2000) The Hallmarks of Cancer. *Cell* 100: 57–70

Hanahan D & Weinberg RA (2011) Hallmarks of cancer: the next generation. *Cell* 144: 646–674

Handler D, Olivieri D, Novatchkova M, Gruber FS, Meixner K, Mechtler K, Stark A, Sachidanandam R & Brennecke J (2011) A systematic analysis of Drosophila TUDOR domain-containing proteins identifies Vreteno and the Tdrd12 family as essential primary piRNA pathway factors. *The EMBO Journal* 30: 3977–3993

Hayashi Y, Ford LK, Fioriti L, McGurk L & Zhang M (2021) Liquid-Liquid Phase Separation in Physiology and Pathophysiology of the Nervous System. *J Neurosci* 41: 834–844

Hayes MH, Peuchen EH, Dovichi NJ & Weeks DL (2018) Dual roles for ATP in the regulation of phase separated protein aggregates in Xenopus oocyte nucleoli. *eLife* 7: e35224

He Y, Selvaraju S, Curtin ML, Jakob CG, Zhu H, Comess KM, Shaw B, The J, Lima-Fernandes E, Szewczyk MM, *et al* (2017) The EED protein-protein interaction inhibitor A-395 inactivates the PRC2 complex. *Nat Chem Biol* 13: 389–395

Herold JM, Ingerman James L, K. Korboukh V, Gao C, E. Coil K, J. Bua D, L. Norris J, B. Kireev D, J. Brown P, Jin J, *et al* (2012) Structure–activity relationships of methyl-lysine reader antagonists. *MedChemComm* 3: 45–51

Herold JM, Wigle TJ, Norris JL, Lam R, Korboukh VK, Gao C, Ingerman LA, Kireev DB, Senisterra G, Vedadi M, *et al* (2011) Small-molecule ligands of methyl-lysine binding proteins. *J Med Chem* 54: 2504–2511

Hirakata S, Ishizu H, Fujita A, Tomoe Y & Siomi MC (2019) Requirements for multivalent Yb body assembly in transposon silencing in Drosophila. *EMBO reports* 20: e47708

Hnisz D, Shrinivas K, Young RA, Chakraborty AK & Sharp PA (2017) A Phase Separation Model for Transcriptional Control. *Cell* 169: 13–23

Hofweber M & Dormann D (2019) Friend or foe-Post-translational modifications as regulators of phase separation and RNP granule dynamics. *J Biol Chem* 294: 7137–7150

Hofweber M, Hutten S, Bourgeois B, Spreitzer E, Niedner-Boblenz A, Schifferer M, Ruepp M-D, Simons M, Niessing D, Madl T, *et al* (2018) Phase Separation of FUS Is Suppressed by Its Nuclear Import Receptor and Arginine Methylation. *Cell* 173: 706-719.e13

Holt CE & Schuman EM (2013) The Central Dogma Decentralized: New Perspectives on RNA Function and Local Translation in Neurons. *Neuron* 80: 648–657

Horton TM, Allegretti PA, Lee S, Moeller HP, Smith M & Annes JP (2018) Zinc-Chelating Small Molecules Preferentially Accumulate and Function within Pancreatic β Cells. *Cell Chemical Biology* 0

Hsin J-P & Manley JL (2012) The RNA polymerase II CTD coordinates transcription and RNA processing. *Genes Dev* 26: 2119–2137

Hu S, Lv P, Yan Z & Wen B (2019) Disruption of nuclear speckles reduces chromatin interactions in active compartments. *Epigenetics Chromatin* 12: 43

Hua L, Zhang Q, Zhu X, Wang R, You Q & Wang L (2022) Beyond Proteolysis-Targeting Chimeric Molecules: Designing Heterobifunctional Molecules Based on Functional Effectors. *J Med Chem* 65: 8091–8112

Huang D, Tian S, Qi Y & Zhang JZH (2020) Binding Modes of Small-Molecule Inhibitors to the EED Pocket of PRC2. *Chemphyschem* 21: 263–271

Huang Y, Fang J, Bedford MT, Zhang Y & Xu R-M (2006) Recognition of Histone H3 Lysine-4 Methylation by the Double Tudor Domain of JMJD2A. *Science* 312: 748–751

Huang Y, Zhang J, Yu Z, Zhang H, Wang Y, Lingel A, Qi W, Gu J, Zhao K, Shultz MD, *et al* (2017) Discovery of First-in-Class, Potent, and Orally Bioavailable Embryonic Ectoderm Development (EED) Inhibitor with Robust Anticancer Efficacy. *J Med Chem* 60: 2215–2226

Hult C, Adalsteinsson D, Vasquez PA, Lawrimore J, Bennett M, York A, Cook D, Yeh E, Forest MG & Bloom K (2017) Enrichment of dynamic chromosomal crosslinks drive phase separation of the nucleolus. *Nucleic Acids Res* 45: 11159–11173

Hyman AA, Weber CA & Jülicher F (2014) Liquid-Liquid Phase Separation in Biology. *Annual Review of Cell and Developmental Biology* 30: 39–58

Ilik İA, Malszycki M, Lübke AK, Schade C, Meierhofer D & Aktaş T (2020) SON and SRRM2 are essential for nuclear speckle formation. *eLife* 9: e60579

Iserman C, Altamirano CD, Jegers C, Friedrich U, Zarin T, Fritsch AW, Mittasch M, Domingues A, Hersemann L, Jahnel M, *et al* (2020) Condensation of Ded1p Promotes a Translational Switch from Housekeeping to Stress Protein Production. *Cell* 181: 818-831.e19

Ishida Y-I, Miyao S, Saito M, Hiraishi N & Nagahama M (2021) Interactome analysis of the Tudor domain-containing protein SPF30 which associates with the MTR4-exosome RNA-decay machinery under the regulation of AAA-ATPase NVL2. *Int J Biochem Cell Biol*: 105919

James LI, Barsyte-Lovejoy D, Zhong N, Krichevsky L, Korboukh VK, Herold MJ, MacNevin CJ, Norris JL, Sagum CA, Tempel W, *et al* (2013a) Discovery of a chemical probe for the L3MBTL3 methyl-lysine reader domain. *Nat Chem Biol* 9: 184–191

James LI, Korboukh VK, Krichevsky L, Baughman BM, Herold JM, Norris JL, Jin J, Kireev DB, Janzen WP, Arrowsmith CH, *et al* (2013b) Small-molecule ligands of methyl-lysine binding proteins: optimization of selectivity for L3MBTL3. *J Med Chem* 56: 7358–7371

Jonas S & Izaurralde E (2013) The role of disordered protein regions in the assembly of decapping complexes and RNP granules. *Genes Dev* 27: 2628–2641

Kaniskan HÜ, Konze KD & Jin J (2015) Selective inhibitors of protein methyltransferases. *J Med Chem* 58: 1596–1629

Karatas H, Li Y, Liu L, Ji J, Lee S, Chen Y, Yang J, Huang L, Bernard D, Xu J, *et al* (2017) Discovery of a Highly Potent, Cell-Permeable Macrocyclic Peptidomimetic (MM-589) Targeting the WD Repeat Domain 5 Protein (WDR5)-Mixed Lineage Leukemia (MLL) Protein-Protein Interaction. *J Med Chem* 60: 4818–4839

Karatas H, Townsend EC, Cao F, Chen Y, Bernard D, Liu L, Lei M, Dou Y & Wang S (2013) High-affinity, small-molecule peptidomimetic inhibitors of MLL1/WDR5 protein-protein interaction. *J Am Chem Soc* 135: 669–682

Kent S, Brown K, Yang C-H, Alsaihati N, Tian C, Wang H & Ren X (2020) Phase-Separated Transcriptional Condensates Accelerate Target-Search Process Revealed by Live-Cell Single-Molecule Imaging. *Cell Rep* 33: 108248

Khanna N, Hu Y & Belmont AS (2014) HSP70 Transgene Directed Motion to Nuclear Speckles Facilitates Heat Shock Activation. *Current Biology* 24: 1138–1144

Khurana JS & Theurkauf W (2010) piRNAs, transposon silencing, and Drosophila germline development. *J Cell Biol* 191: 905–913

Kiebler M, Scheiffele P & Ule J (2013) What, where, and when: the importance of post-transcriptional regulation in the brain. *Frontiers in Neuroscience* 7

Kilgore HR & Young RA (2022) Learning the chemical grammar of biomolecular condensates. *Nat Chem Biol* 18: 1298–1306

Kim GH & Kwon I (2021) Distinct roles of hnRNPH1 low-complexity domains in splicing and transcription. *Proc Natl Acad Sci U S A* 118: e2109668118

Kim J, Daniel J, Espejo A, Lake A, Krishna M, Xia L, Zhang Y & Bedford MT (2006) Tudor, MBT and chromo domains gauge the degree of lysine methylation. *EMBO Rep* 7: 397–403

Kim J, Venkata NC, Hernandez Gonzalez GA, Khanna N & Belmont AS (2019) Gene expression amplification by nuclear speckle association. *Journal of Cell Biology* 219

Kim J-H, Shinde DN, Reijnders MRF, Hauser NS, Belmonte RL, Wilson GR, Bosch DGM, Bubulya PA, Shashi V, Petrovski S, *et al* (2016) De Novo Mutations in SON Disrupt RNA Splicing of Genes Essential for Brain Development and Metabolism, Causing an Intellectual-Disability Syndrome. *Am J Hum Genet* 99: 711–719

Klein IA, Boija A, Afeyan LK, Hawken SW, Fan M, Dall'Agnese A, Oksuz O, Henninger JE, Shrinivas K, Sabari BR, *et al* (2020) Partitioning of cancer therapeutics in nuclear condensates. *Science* 368: 1386–1392

Klose RJ, Yamane K, Bae Y, Zhang D, Erdjument-Bromage H, Tempst P, Wong J & Zhang Y (2006) The transcriptional repressor JHDM3A demethylates trimethyl histone H3 lysine 9 and lysine 36. *Nature* 442: 312–316

Konopka S, Trievel R & Albaugh B (2022) Investigating the structure of UHRF2 TTD-PHD with H3K9me3. *FASEB J* 36 Suppl 1

Kori S, Shibahashi Y, Ekimoto T, Nishiyama A, Yoshimi S, Yamaguchi K, Nagatoishi S, Ohta M, Tsumoto K, Nakanishi M, *et al* (2021) Structure-based screening combined with computational and biochemical analyses identified the inhibitor targeting the binding of DNA Ligase 1 to UHRF1. *Bioorganic & Medicinal Chemistry* 52: 116500

Kroschwald S, Maharana S & Simon A (2017) Hexanediol: a chemical probe to investigate the material properties of membrane-less compartments. *Matters* 3: e201702000010

Kwon I, Kato M, Xiang S, Wu L, Theodoropoulos P, Mirzaei H, Han T, Xie S, Corden JL & McKnight SL (2013) Phosphorylation-Regulated Binding of RNA Polymerase II to Fibrous Polymers of Low-Complexity Domains. *Cell* 155: 1049–1060

Kye MJ, Niederst ED, Wertz MH, Gonçalves I do CG, Akten B, Dover KZ, Peters M, Riessland M, Neveu P, Wirth B, *et al* (2014) SMN regulates axonal local translation via miR-183/mTOR pathway. *Hum Mol Genet* 23: 6318–6331

Langdon EM, Qiu Y, Niaki AG, McLaughlin GA, Weidmann CA, Gerbich TM, Smith JA, Crutchley JM, Termini CM, Weeks KM, *et al* (2018) mRNA structure determines specificity of a polyQ-driven phase separation. *Science* 360: 922–927

Larsen SC, Sylvestersen KB, Mund A, Lyon D, Mullari M, Madsen MV, Daniel JA, Jensen LJ & Nielsen ML (2016) Proteome-wide analysis of arginine monomethylation reveals widespread occurrence in human cells. *Sci Signal* 9: rs9–rs9

Larson AG, Elnatan D, Keenen MM, Trnka MJ, Johnston JB, Burlingame AL, Agard DA, Redding S & Narlikar GJ (2017) Liquid droplet formation by HP1α suggests a role for phase separation in heterochromatin. *Nature* 547: 236–240

Larson AG & Narlikar GJ (2018) The Role of Phase Separation in Heterochromatin Formation, Function, and Regulation. *Biochemistry* 57: 2540–2548

Le Romancer M, Treilleux I, Leconte N, Robin-Lespinasse Y, Sentis S, Bouchekioua-Bouzaghou K, Goddard S, Gobert-Gosse S & Corbo L (2008) Regulation of estrogen rapid signaling through arginine methylation by PRMT1. *Mol Cell* 31: 212–221

Lee J-H, Wang R, Xiong F, Krakowiak J, Liao Z, Nguyen PT, Moroz-Omori EV, Shao J, Zhu X, Bolt MJ, *et al* (2021) Enhancer RNA m6A methylation facilitates transcriptional condensate formation and gene activation. *Mol Cell* 81: 3368-3385.e9

Li L, Zhang H, Zhang M, Zhao M, Feng L, Luo X, Gao Z, Huang Y, Ardayfio O, Zhang J-H, *et al* (2017) Discovery and Molecular Basis of a Diverse Set of Polycomb Repressive Complex 2 Inhibitors Recognition by EED. *PLoS ONE* 12: e0169855

Li P, Banjade S, Cheng H-C, Kim S, Chen B, Guo L, Llaguno M, Hollingsworth JV, King DS, Banani SF, *et al* (2012) Phase transitions in the assembly of multivalent signalling proteins. *Nature* 483: 336–340

Liao SE & Regev O (2020) Splicing at the phase-separated nuclear speckle interface: a model. *Nucleic Acids Research*

Lim J & Yue Z (2015) Neuronal aggregates: formation, clearance and spreading. *Dev Cell* 32: 491–501

Linder B, Plöttner O, Kroiss M, Hartmann E, Laggerbauer B, Meister G, Keidel E & Fischer U (2008) Tdrd3 is a novel stress granule-associated protein interacting with the Fragile-X syndrome protein FMRP. *Hum Mol Genet* 17: 3236–3246

Liu H, Wang J-YS, Huang Y, Li Z, Gong W, Lehmann R & Xu R-M (2010a) Structural basis for methylarginine-dependent recognition of Aubergine by Tudor. *Genes Dev* 24: 1876–1881

Liu J, Zhang S, Liu M, Liu Y, Nshogoza G, Gao J, Ma R, Yang Y, Wu J, Zhang J, *et al* (2018) Structural plasticity of the TDRD3 Tudor domain probed by a fragment screening hit. *FEBS J* 285: 2091–2103

Liu K, Chen C, Guo Y, Lam R, Bian C, Xu C, Zhao DY, Jin J, MacKenzie F & Pawson T (2010b) Structural basis for recognition of arginine methylated Piwi proteins by the extended Tudor domain. *Proceedings of the National Academy of Sciences* 107: 18398–18403

Liu L, Qi H, Wang J & Lin H (2011) PAPI, a novel TUDOR-domain protein, complexes with AGO3, ME31B and TRAL in the nuage to silence transposition. *Development* 138: 1863–1873

Liu Q & Dreyfuss G (1996) A novel nuclear structure containing the survival of motor neurons protein. *EMBO J* 15: 3555–3565

Liu Q, Fischer U, Wang F & Dreyfuss G (1997) The spinal muscular atrophy disease gene product, SMN, and its associated protein SIP1 are in a complex with spliceosomal snRNP proteins. *Cell* 90: 1013–1021

Liu S, Wang T, Shi Y, Bai L, Wang S, Guo D, Zhang Y, Qi Y, Chen C, Zhang J, *et al* (2021) USP42 drives nuclear speckle mRNA splicing via directing dynamic phase separation to promote tumorigenesis. *Cell Death Differ* 28: 2482–2498

Liu Y, Iqbal A, Li W, Ni Z, Wang Y, Ramprasad J, Abraham KJ, Zhang M, Zhao DY, Qin S, *et al* (2022) A small molecule antagonist of SMN disrupts the interaction between SMN and RNAP II. *Nat Commun* 13: 5453

Lu H, Yu D, Hansen AS, Ganguly S, Liu R, Heckert A, Darzacq X & Zhou Q (2018) Phase-separation mechanism for C-terminal hyperphosphorylation of RNA polymerase II. *Nature* 558: 318–323

Lu Y, Wu T, Gutman O, Lu H, Zhou Q, Henis YI & Luo K (2020) Phase separation of TAZ compartmentalizes the transcription machinery to promote gene expression. *Nature Cell Biology* 22: 453–464

Lyons H, Veettil RT, Pradhan P, Fornero C, De La Cruz N, Ito K, Eppert M, Roeder RG & Sabari BR (2023) Functional partitioning of transcriptional regulators by patterned charge blocks. *Cell* 186: 327-345.e28

Mackenzie IR, Rademakers R & Neumann M (2010) TDP-43 and FUS in amyotrophic lateral sclerosis and frontotemporal dementia. *The Lancet Neurology* 9: 995–1007

Mader P, Mendoza-Sanchez R, Iqbal A, Dong A, Dobrovetsky E, Corless VB, Liew SK, Houliston SR, De Freitas RF, Smil D, *et al* (2019) Identification and characterization of the first fragment hits for SETDB1 Tudor domain. *Bioorganic & Medicinal Chemistry* 27: 3866–3878

Maharana S, Wang J, Papadopoulos DK, Richter D, Pozniakovsky A, Poser I, Bickle M, Rizk S, Guillén-Boixet J, Franzmann TM, *et al* (2018) RNA buffers the phase separation behavior of prion-like RNA binding proteins. *Science* 360: 918–921

Malinovska L, Kroschwald S & Alberti S (2013) Protein disorder, prion propensities, and self-organizing macromolecular collectives. *Biochim Biophys Acta* 1834: 918–931

Marzahn MR, Marada S, Lee J, Nourse A, Kenrick S, Zhao H, Ben-Nissan G, Kolaitis R-M, Peters JL, Pounds S, *et al* (2016) Higher-order oligomerization promotes localization of SPOP to liquid nuclear speckles. *EMBO J* 35: 1254–1275

Mateju D, Eichenberger B, Voigt F, Eglinger J, Roth G & Chao JA (2020) Single-Molecule Imaging Reveals Translation of mRNAs Localized to Stress Granules. *Cell* 183: 1801-1812.e13

Maurer-Stroh S, Dickens NJ, Hughes-Davies L, Kouzarides T, Eisenhaber F & Ponting CP (2003) The Tudor domain 'Royal Family': Tudor, plant Agenet, Chromo, PWWP and MBT domains. *Trends in Biochemical Sciences* 28: 69–74

Meister G, Hannus S, Plöttner O, Baars T, Hartmann E, Fakan S, Laggerbauer B & Fischer U (2001) SMNrp is an essential pre-mRNA splicing factor required for the formation of the mature spliceosome. *The EMBO Journal* 20: 2304–2314

Métivier R, Penot G, Hübner MR, Reid G, Brand H, Kos M & Gannon F (2003) Estrogen receptor-alpha directs ordered, cyclical, and combinatorial recruitment of cofactors on a natural target promoter. *Cell* 115: 751–763

Meyer KD, Saletore Y, Zumbo P, Elemento O, Mason CE & Jaffrey SR (2012) Comprehensive Analysis of mRNA Methylation Reveals Enrichment in 3′ UTRs and near Stop Codons. *Cell* 149: 1635–1646

Miller TCR, Rutherford TJ, Birchall K, Chugh J, Fiedler M & Bienz M (2014) Competitive Binding of a Benzimidazole to the Histone-Binding Pocket of the Pygo PHD Finger. *ACS Chem Biol* 9: 2864–2874

Milosevich N, Gignac MC, McFarlane J, Simhadri C, Horvath S, Daze KD, Croft CS, Dheri A, Quon TTH, Douglas SF, *et al* (2016a) Selective Inhibition of CBX6: A Methyllysine Reader Protein in the Polycomb Family. *ACS Med Chem Lett* 7: 139–144

Milosevich N, Warmerdam Z & Hof F (2016b) Structural aspects of small-molecule inhibition of methyllysine reader proteins. *Future Med Chem* 8: 1681–1702

Mitrea DM, Mittasch M, Gomes BF, Klein IA & Murcko MA (2022) Modulating biomolecular condensates: a novel approach to drug discovery. *Nat Rev Drug Discov* 21: 841–862

Morettin A, Paris G, Bouzid Y, Baldwin RM, Falls TJ, Bell JC & Côté J (2017) Tudor Domain Containing Protein 3 Promotes Tumorigenesis and Invasive Capacity of Breast Cancer Cells. *Sci Rep* 7: 5153

Mukherjee H, Chan K-P, Andresen V, Hanley ML, Gjertsen BT & Myers AG (2015) Interactions of the Natural Product (+)-Avrainvillamide with Nucleophosmin and Exportin-1 Mediate the Cellular Localization of Nucleophosmin and its AML-Associated Mutants. *ACS Chem Biol* 10: 855–863

Muraro MJ, Dharmadhikari G, Grün D, Groen N, Dielen T, Jansen E, van Gurp L, Engelse MA, Carlotti F, de Koning EJP, *et al* (2016) A Single-Cell Transcriptome Atlas of the Human Pancreas. *Cell Systems* 3: 385-394.e3

Nady N, Lemak A, Walker JR, Avvakumov GV, Kareta MS, Achour M, Xue S, Duan S, Allali-Hassani A, Zuo X, *et al* (2011) Recognition of Multivalent Histone States Associated with Heterochromatin by UHRF1 Protein. *J Biol Chem* 286: 24300–24311

Neubauer G, King A, Rappsilber J, Calvio C, Watson M, Ajuh P, Sleeman J, Lamond A & Mann M (1998) Mass spectrometry and EST-database searching allows characterization of the multi-protein spliceosome complex. *Nat Genet* 20: 46–50

Noordstra I, van den Berg CM, Boot FWJ, Katrukha EA, Yu KL, Tas RP, Portegies S, Viergever BJ, de Graaff E, Hoogenraad CC, *et al* (2022) Organization and dynamics of the cortical complexes controlling insulin secretion in β-cells. *J Cell Sci*: jcs.259430

Nott TJ, Petsalaki E, Farber P, Jervis D, Fussner E, Plochowietz A, Craggs TD, Bazett-Jones DP, Pawson T, Forman-Kay JD, *et al* (2015) Phase Transition of a Disordered Nuage Protein Generates Environmentally Responsive Membraneless Organelles. *Molecular Cell* 57: 936–947

Paik WK & Kim S (1975) Protein Methylation: Chemical, Enzymological, and Biological Significance. In *Advances in Enzymology and Related Areas of Molecular Biology* pp 227–286. John Wiley & Sons, Ltd

Palikyras S & Papantonis A (2019) Modes of phase separation affecting chromatin regulation. *Open Biol* 9

Papoutsopoulou S & Janknecht R (2000) Phosphorylation of ETS transcription factor ER81 in a complex with its coactivators CREB-binding protein and p300. *Mol Cell Biol* 20: 7300–7310

Parchure A, Tian M, Stalder D, Boyer CK, Bearrows SC, Rohli KE, Zhang J, Rivera-Molina F, Ramazanov BR, Mahata SK, *et al* (2022) Liquid-liquid phase separation facilitates the biogenesis of secretory storage granules. *J Cell Biol* 221: e202206132

Park I, Hwang YJ, Kim T, Viswanath ANI, Londhe AM, Jung SY, Sim KM, Min S-J, Lee JE, Seong J, *et al* (2017) In silico probing and biological evaluation of SETDB1/ESET-targeted novel compounds that reduce tri-methylated histone H3K9 (H3K9me3) level. *J Comput Aided Mol Des* 31: 877–889

Park I-G, Jeon M, Kim H & Lee JM (2022) Coordinated methyl readers: Functional communications in cancer. *Seminars in Cancer Biology* 83: 88–99

Patel A, Malinovska L, Saha S, Wang J, Alberti S, Krishnan Y & Hyman AA (2017) ATP as a biological hydrotrope. *Science* 356: 753–756

Paul S, Arias MA, Wen L, Liao SE, Zhang J, Wang X, Regev O & Fei J (2022) Nuclear speckle-localized RNAs exhibit preferential positioning and orientation. 2022.10.17.512423 doi:10.1101/2022.10.17.512423 [PREPRINT]

Pek JW, Anand A & Kai T (2012) Tudor domain proteins in development. *Development* 139: 2255–2266

Perfetti MT, Baughman BM, Dickson BM, Mu Y, Cui G, Mader P, Dong A, Norris JL, Rothbart SB, Strahl BD, *et al* (2015) Identification of a fragment-like small molecule ligand for the methyl-lysine binding protein, 53BP1. *ACS Chem Biol* 10: 1072–1081

Pettinato MC (2021) Introduction to Antibody-Drug Conjugates. *Antibodies (Basel)* 10: 42

Ping X-L, Sun B-F, Wang L, Xiao W, Yang X, Wang W-J, Adhikari S, Shi Y, Lv Y, Chen Y-S, *et al* (2014) Mammalian WTAP is a regulatory subunit of the RNA N6-methyladenosine methyltransferase. *Cell Res* 24: 177–189

Ponting CP (1997) Tudor domains in proteins that interact with RNA. *Trends Biochem Sci* 22: 51–52

Pytowski L, Lee CF, Foley AC, Vaux DJ & Jean L (2020) Liquid-liquid phase separation of type II diabetes-associated IAPP initiates hydrogelation and aggregation. *Proc Natl Acad Sci USA* 117: 12050–12061

Qamar S, Wang G, Randle SJ, Ruggeri FS, Varela JA, Lin JQ, Phillips EC, Miyashita A, Williams D, Ströhl F, *et al* (2018) FUS Phase Separation Is Modulated by a Molecular Chaperone and Methylation of Arginine Cation-π Interactions. *Cell* 173: 720-734.e15

Qi W, Zhao K, Gu J, Huang Y, Wang Y, Zhang H, Zhang M, Zhang J, Yu Z, Li L, *et al* (2017) An allosteric PRC2 inhibitor targeting the H3K27me3 binding pocket of EED. *Nat Chem Biol* 13: 381–388

Qiu C, Sawada K, Zhang X & Cheng X (2002) The PWWP domain of mammalian DNA methyltransferase Dnmt3b defines a new family of DNA-binding folds. *Nat Struct Mol Biol* 9: 217–224

Quinodoz SA, Ollikainen N, Tabak B, Palla A, Schmidt JM, Detmar E, Lai MM, Shishkin AA, Bhat P, Takei Y, *et al* (2018) Higher-Order Inter-chromosomal Hubs Shape 3D Genome Organization in the Nucleus. *Cell* 174: 744-757.e24

Raichle ME & Gusnard DA (2002) Appraising the brain's energy budget. *Proceedings of the National Academy of Sciences* 99: 10237–10239

Ramón y Cajal S (1903) Un sencillo método de coloración selectiva del retículo protoplásmico y sus efectos en los diversos órganos nerviosos. *Trab Lab Invest Biol(Madrid)* 2: 129–221

Ramón y Cajal S (1910) El núcleo de las células piramidales del cerebro humano y de algunos mamíferos

Rappsilber J, Ajuh P, Lamond AI & Mann M (2001) SPF30 Is an Essential Human Splicing Factor Required for Assembly of the U4/U5/U6 Tri-small Nuclear Ribonucleoprotein into the Spliceosome. *J Biol Chem* 276: 31142–31150

Ray S, Singh N, Kumar R, Patel K, Pandey S, Datta D, Mahato J, Panigrahi R, Navalkar A, Mehra S, *et al* (2020) α-Synuclein aggregation nucleates through liquid–liquid phase separation. *Nat Chem* 12: 705–716

Regan-Fendt KE & Izumi K (2023) Nuclear speckleopathies: developmental disorders caused by variants in genes encoding nuclear speckle proteins. *Hum Genet*

Ren C, Morohashi K, Plotnikov AN, Jakoncic J, Smith SG, Li J, Zeng L, Rodriguez Y, Stojanoff V, Walsh M, *et al* (2015) Small-molecule modulators of methyl-lysine binding for the CBX7 chromodomain. *Chem Biol* 22: 161–168

Ren C, Smith SG, Yap K, Li S, Li J, Mezei M, Rodriguez Y, Vincek A, Aguilo F, Walsh MJ, *et al* (2016) Structure-Guided Discovery of Selective Antagonists for the Chromodomain of Polycomb Repressive Protein CBX7. *ACS Med Chem Lett* 7: 601–605

Ries RJ, Zaccara S, Klein P, Olarerin-George A, Namkoong S, Pickering BF, Patil DP, Kwak H, Lee JH & Jaffrey SR (2019) m6A enhances the phase separation potential of mRNA. *Nature* 571: 424–428

Ripolone M, Ronchi D, Violano R, Vallejo D, Fagiolari G, Barca E, Lucchini V, Colombo I, Villa L, Berardinelli A, *et al* (2015) Impaired Muscle Mitochondrial Biogenesis and Myogenesis in Spinal Muscular Atrophy. *JAMA Neurol* 72: 666–675

Rodriguez-Paredes M, Paz AM de, Simó-Riudalbas L, Sayols S, Moutinho C, Moran S, Villanueva A, Vázquez-Cedeira M, Lazo PA, Carneiro F, *et al* (2014) Gene amplification of the histone methyltransferase SETDB1 contributes to human lung tumorigenesis. *Oncogene* 33: 2807–2813

Roovers EF, Kaaij LJT, Redl S, Bronkhorst AW, Wiebrands K, de Jesus Domingues AM, Huang H-Y, Han C-T, Riemer S, Dosch R, *et al* (2018) Tdrd6a Regulates the Aggregation of Buc into Functional Subcellular Compartments that Drive Germ Cell Specification. *Dev Cell* 46: 285-301.e9

Rossoll W, Jablonka S, Andreassi C, Kröning A-K, Karle K, Monani UR & Sendtner M (2003) Smn, the spinal muscular atrophy-determining gene product, modulates axon growth and localization of beta-actin mRNA in growth cones of motoneurons. *J Cell Biol* 163: 801–812

Rottach A, Frauer C, Pichler G, Bonapace IM, Spada F & Leonhardt H (2010) The multi-domain protein Np95 connects DNA methylation and histone modification. *Nucleic Acids Res* 38: 1796–1804

Ryan VH, Dignon GL, Zerze GH, Chabata CV, Silva R, Conicella AE, Amaya J, Burke KA, Mittal J & Fawzi NL (2018) Mechanistic View of hnRNPA2 Low-Complexity Domain Structure, Interactions, and Phase Separation Altered by Mutation and Arginine Methylation. *Molecular Cell* 69: 465-479.e7

Sabari BR, Dall'Agnese A, Boija A, Klein IA, Coffey EL, Shrinivas K, Abraham BJ, Hannett NM, Zamudio AV, Manteiga JC, *et al* (2018) Coactivator condensation at super-enhancers links phase separation and gene control. *Science* 361: eaar3958

Saito K & Siomi MC (2010) Small RNA-Mediated Quiescence of Transposable Elements in Animals. *Developmental Cell* 19: 687–697

Saitoh N, Spahr CS, Patterson SD, Bubulya P, Neuwald AF & Spector DL (2004) Proteomic Analysis of Interchromatin Granule Clusters. *Mol Biol Cell* 15: 3876–3890

Sakabe K & Hart GW (2010) O-GlcNAc transferase regulates mitotic chromatin dynamics. *J Biol Chem* 285: 34460–34468

Sanchez G, Dury AY, Murray LM, Biondi O, Tadesse H, El Fatimy R, Kothary R, Charbonnier F, Khandjian EW & Côté J (2013) A novel function for the survival motoneuron protein as a translational regulator. *Hum Mol Genet* 22: 668–684

Sanchez R & Zhou M-M (2011) The PHD finger: a versatile epigenome reader. *Trends Biochem Sci* 36: 364–372

Sanders DW, Kedersha N, Lee DSW, Strom AR, Drake V, Riback JA, Bracha D, Eeftens JM, Iwanicki A, Wang A, *et al* (2020) Competing Protein-RNA Interaction Networks Control Multiphase Intracellular Organization. *Cell* 181: 306-324.e28

Santiago C, Nguyen K & Schapira M (2011) Druggability of methyl-lysine binding sites. *J Comput Aided Mol Des* 25: 1171–1178

Schapira M, Tyers M, Torrent M & Arrowsmith CH (2017) WD40 repeat domain proteins: a novel target class? *Nat Rev Drug Discov* 16: 773–786

Selenko P, Sprangers R, Stier G, Bühler D, Fischer U & Sattler M (2001) SMN Tudor domain structure and its interaction with the Sm proteins. *Nat Struct Mol Biol* 8: 27–31

Senisterra G, Wu H, Allali-Hassani A, Wasney GA, Barsyte-Lovejoy D, Dombrovski L, Dong A, Nguyen KT, Smil D, Bolshan Y, *et al* (2013) Small-molecule inhibition of MLL activity by disruption of its interaction with WDR5. *Biochem J* 449: 151–159

Senisterra G, Zhu HY, Luo X, Zhang H, Xun G, Lu C, Xiao W, Hajian T, Loppnau P, Chau I, *et al* (2018) Discovery of Small-Molecule Antagonists of the H3K9me3 Binding to UHRF1 Tandem Tudor Domain. *SLAS DISCOVERY: Advancing the Science of Drug Discovery* 23: 930–940

Senti K-A & Brennecke J (2010) The piRNA pathway: a fly's perspective on the guardian of the genome. *Trends in Genetics* 26: 499–509

Sevcuka A, White K & Terry C (2022) Factors That Contribute to hIAPP Amyloidosis in Type 2 Diabetes Mellitus. *Life* 12: 583

Sharova LV, Sharov AA, Nedorezov T, Piao Y, Shaik N & Ko MSH (2009) Database for mRNA half-life of 19 977 genes obtained by DNA microarray analysis of pluripotent and differentiating mouse embryonic stem cells. *DNA Res* 16: 45–58

Shen EC, Henry MF, Weiss VH, Valentini SR, Silver PA & Lee MS (1998) Arginine methylation facilitates the nuclear export of hnRNP proteins. *Genes Dev* 12: 679–691

Sikorsky T, Hobor F, Krizanova E, Pasulka J, Kubicek K & Stefl R (2012) Recognition of asymmetrically dimethylated arginine by TDRD3. *Nucleic Acids Res* 40: 11748–11755

Šimčíková D, Gelles-Watnick S & Neugebauer KM (2023) Tudor–dimethylarginine interactions: the condensed version. *Trends in Biochemical Sciences*

Simhadri C, Daze KD, Douglas SF, Quon TTH, Dev A, Gignac MC, Peng F, Heller M, Boulanger MJ, Wulff JE, *et al* (2014) Chromodomain antagonists that target the polycomb-group methyllysine reader protein chromobox homolog 7 (CBX7). *J Med Chem* 57: 2874–2883

Sinclair GD & Brasch K (1978) The reversible action of α-amanitin on nuclear structure and molecular composition. *Experimental Cell Research* 111: 1–14

Singh KK, Erkelenz S, Rattay S, Dehof AK, Hildebrandt A, Schulze-Osthoff K, Schaal H & Schwerk C (2010) Human SAP18 mediates assembly of a splicing regulatory multiprotein complex via its ubiquitin-like fold. *RNA* 16: 2442–2454

Sleeman JE & Trinkle-Mulcahy L (2014) Nuclear bodies: new insights into assembly/dynamics and disease relevance. *Current Opinion in Cell Biology* 28: 76–83

Snead WT & Gladfelter AS (2019) The Control Centers of Biomolecular Phase Separation: How Membrane Surfaces, PTMs, and Active Processes Regulate Condensation. *Molecular Cell* 76: 295–305

Soboleva TA, Parker BJ, Nekrasov M, Hart-Smith G, Tay YJ, Tng W-Q, Wilkins M, Ryan D & Tremethick DJ (2017) A new link between transcriptional initiation and pre-mRNA splicing: The RNA binding histone variant H2A.B. *PLOS Genetics* 13: e1006633

Spector DL & Lamond AI (2011) Nuclear Speckles. *Cold Spring Harb Perspect Biol* 3

Strom AR & Brangwynne CP (2019) The liquid nucleome - phase transitions in the nucleus at a glance. *J Cell Sci* 132

Strom AR, Emelyanov AV, Mir M, Fyodorov DV, Darzacq X & Karpen GH (2017) Phase separation drives heterochromatin domain formation. *Nature* 547: 241–245

Stubbe J, Tian J, He A, Sinskey AJ, Lawrence AG & Liu P (2005) Nontemplate-dependent polymerization processes: polyhydroxyalkanoate synthases as a paradigm. *Annu Rev Biochem* 74: 433–480

Stuckey JI, Dickson BM, Cheng N, Liu Y, Norris JL, Cholensky SH, Tempel W, Qin S, Huber KG, Sagum C, *et al* (2016) A cellular chemical probe targeting the chromodomains of Polycomb repressive complex 1. *Nat Chem Biol* 12: 180–187

Su C, Gao X, Yang W, Zhao Y, Fu X, Cui X, Zhang C, Xin L, Ren Y, Li L, *et al* (2017) Phosphorylation of Tudor-SN, a novel substrate of JNK, is involved in the efficient recruitment of Tudor-SN into stress granules. *Biochim Biophys Acta Mol Cell Res* 1864: 562–571

Su X, Ditlev JA, Hui E, Xing W, Banjade S, Okrut J, King DS, Taunton J, Rosen MK & Vale RD (2016) Phase separation of signaling molecules promotes T cell receptor signal transduction. *Science* 352: 595–599

Su Y, Maimaitiyiming Y, Wang L, Cheng X & Hsu C-H (2021) Modulation of Phase Separation by RNA: A Glimpse on N6-Methyladenosine Modification. *Front Cell Dev Biol* 9: 786454

Suh JL, Barnash KD, Abramyan TM, Li F, The J, Engelberg IA, Vedadi M, Brown PJ, Kireev DB, Arrowsmith CH, *et al* (2019) Discovery of selective activators of PRC2 mutant EED-I363M. *Sci Rep* 9: 6524

Sun Q-Y, Ding L-W, Xiao J-F, Chien W, Lim S-L, Hattori N, Goodglick L, Chia D, Mah V, Alavi M, *et al* (2015) SETDB1 accelerates tumourigenesis by regulating the WNT signalling pathway. *The Journal of Pathology* 235: 559–570

Sun Y, Lu H, Fang X, Xiao S, Yang F, Chen Y, Wang H, Li X, Lu J, Lin H, *et al* (2021) Discovery of a novel 53BP1 inhibitor through AlphaScreen-based high-throughput screening. *Bioorganic & Medicinal Chemistry* 34: 116054

Sweis RF, Pliushchev M, Brown PJ, Guo J, Li F, Maag D, Petros AM, Soni NB, Tse C, Vedadi M, *et al* (2014) Discovery and Development of Potent and Selective Inhibitors of Histone Methyltransferase G9a. *ACS Med Chem Lett* 5: 205–209

Swift H (1959) Studies on nuclear fine structure. *Brookhaven Symp Biol* 12: 134–152

Szakmary A, Reedy M, Qi H & Lin H (2009) The Yb protein defines a novel organelle and regulates male germline stem cell self-renewal in Drosophila melanogaster. *J Cell Biol* 185: 613–627

Takenouchi T, Miura K, Uehara T, Mizuno S & Kosaki K (2016) Establishing SON in 21q22.11 as a cause a new syndromic form of intellectual disability: Possible contribution to Braddock-Carey syndrome phenotype. *Am J Med Genet A* 170: 2587–2590

Talbot K, Miguel-Aliaga I, Mohaghegh P, Ponting CP & Davies KE (1998) Characterization of a Gene Encoding Survival Motor Neuron (Smn)-Related Protein, a Constituent of the Spliceosome Complex. *Human Molecular Genetics* 7: 2149–2156

Tantos A, Han K-H & Tompa P (2012) Intrinsic disorder in cell signaling and gene transcription. *Mol Cell Endocrinol* 348: 457–465

Taylor JP, Hardy J & Fischbeck KH (2002) Toxic Proteins in Neurodegenerative Disease. *Science* 296: 1991–1995

Teng I-F & Wilson SA (2013) Mapping Interactions between mRNA Export Factors in Living Cells. *PLOS ONE* 8: e67676

Thinnes CC, England KS, Kawamura A, Chowdhury R, Schofield CJ & Hopkinson RJ (2014) Targeting histone lysine demethylases - progress, challenges, and the future. *Biochim Biophys Acta* 1839: 1416–1432

Thomas JR, Brittain SM, Lipps J, Llamas L, Jain RK & Schirle M (2017) A Photoaffinity Labeling-Based Chemoproteomics Strategy for Unbiased Target Deconvolution of Small Molecule Drug Candidates. *Methods Mol Biol* 1647: 1–18

Thomas MG, Loschi M, Desbats MA & Boccaccio GL (2011) RNA granules: The good, the bad and the ugly. *Cellular Signalling* 23: 324–334

Tian J, Teuscher KB, Aho ER, Alvarado JR, Mills JJ, Meyers KM, Gogliotti RD, Han C, Macdonald JD, Sai J, *et al* (2020) Discovery and Structure-Based Optimization of Potent and Selective WD Repeat Domain 5 (WDR5) Inhibitors Containing a Dihydroisoquinolinone Bicyclic Core. *J Med Chem* 63: 656–675

Tokita MJ, Braxton AA, Shao Y, Lewis AM, Vincent M, Küry S, Besnard T, Isidor B, Latypova X, Bézieau S, *et al* (2016) De Novo Truncating Variants in SON Cause Intellectual Disability, Congenital Malformations, and Failure to Thrive. *Am J Hum Genet* 99: 720–727

Toledo PL, Torkko JM, Müller A, Wegbrod C, Sönmez A, Solimena M & Ermácora MR (2019) ICA512 RESP18 homology domain is a protein-condensing factor and insulin fibrillation inhibitor. *J Biol Chem* 294: 8564–8576

Toledo PL, Vazquez DS, Gianotti AR, Abate MB, Wegbrod C, Torkko JM, Solimena M & Ermácora MR (2023) Condensation of the β-cell secretory granule luminal cargoes pro/insulin and ICA512 RESP18 homology domain. *Protein Sci*

Tripsianes K, Madl T, Machyna M, Fessas D, Englbrecht C, Fischer U, Neugebauer KM & Sattler M (2011) Structural basis for dimethylarginine recognition by the Tudor domains of human SMN and SPF30 proteins. *Nature Structural & Molecular Biology* 18: 1414–1420

Trojanowski J, Frank L, Rademacher A, Mücke N, Grigaitis P & Rippe K (2022) Transcription activation is enhanced by multivalent interactions independent of phase separation. *Mol Cell* 82: 1878-1893.e10

Uhlén M, Fagerberg L, Hallström BM, Lindskog C, Oksvold P, Mardinoglu A, Sivertsson Å, Kampf C, Sjöstedt E, Asplund A, *et al* (2015) Tissue-based map of the human proteome. *Science* 347: 1260419

Upadhyay AK, Judge RA, Li L, Pithawalla R, Simanis J, Bodelle PM, Marin VL, Henry RF, Petros AM & Sun C (2018) Targeting lysine specific demethylase 4A (KDM4A) tandem TUDOR domain - A fragment based approach. *Bioorg Med Chem Lett* 28: 1708–1713

Usher ET, Sabri N, Rohac R, Boal AK, Mittag T & Showalter SA (2021) Intrinsically disordered substrates dictate SPOP subnuclear localization and ubiquitination activity. *J Biol Chem* 296: 100693

Wagner EK, Nath N, Flemming R, Feltenberger JB & Denu JM (2012) Identification and characterization of small molecule inhibitors of a plant homeodomain finger. *Biochemistry* 51: 8293–8306

Wagner T, Greschik H, Burgahn T, Schmidtkunz K, Schott A-K, McMillan J, Baranauskienė L, Xiong Y, Fedorov O, Jin J, *et al* (2016) Identification of a small-molecule ligand of the epigenetic reader protein Spindlin1 via a versatile screening platform. *Nucleic Acids Res* 44: e88

Wahl MC, Will CL & Lührmann R (2009) The spliceosome: design principles of a dynamic RNP machine. *Cell* 136: 701–718

Walia RR, Xue LC, Wilkins K, El-Manzalawy Y, Dobbs D & Honavar V (2014) RNABindRPlus: A Predictor that Combines Machine Learning and Sequence Homology-Based Methods to Improve the Reliability of Predicted RNA-Binding Residues in Proteins. *PLOS ONE* 9: e97725

Walport LJ, Hopkinson RJ, Chowdhury R, Schiller R, Ge W, Kawamura A & Schofield CJ (2016) Arginine demethylation is catalysed by a subset of JmjC histone lysine demethylases. *Nat Commun* 7: 1–12

Wang J, Wang L, Diao J, Shi YG, Shi Y, Ma H & Shen H (2020a) Binding to m6A RNA promotes YTHDF2-mediated phase separation. *Protein Cell* 11: 304–307

Wang S, Denton KE, Hobbs KF, Weaver T, McFarlane JMB, Connelly KE, Gignac MC, Milosevich N, Hof F, Paci I, *et al* (2020b) Optimization of Ligands Using Focused DNA-Encoded Libraries To Develop a Selective, Cell-Permeable CBX8 Chromodomain Inhibitor. *ACS Chem Biol* 15: 112–131

Wegmann S, Eftekharzadeh B, Tepper K, Zoltowska KM, Bennett RE, Dujardin S, Laskowski PR, MacKenzie D, Kamath T, Commins C, *et al* (2018) Tau protein liquid–liquid phase separation can initiate tau aggregation. *The EMBO Journal* 37: e98049

Wei M-T, Chang Y-C, Shimobayashi SF, Shin Y, Strom AR & Brangwynne CP (2020) Nucleated transcriptional condensates amplify gene expression. *Nat Cell Biol* 22: 1187–1196

Whetstine JR, Nottke A, Lan F, Huarte M, Smolikov S, Chen Z, Spooner E, Li E, Zhang G, Colaiacovo M, *et al* (2006) Reversal of Histone Lysine Trimethylation by the JMJD2 Family of Histone Demethylases. *Cell* 125: 467–481

Wigle TJ, Herold JM, Senisterra GA, Vedadi M, Kireev DB, Arrowsmith CH, Frye SV & Janzen WP (2010) Screening for inhibitors of low-affinity epigenetic peptide-protein interactions: an AlphaScreen-based assay for antagonists of methyl-lysine binding proteins. *J Biomol Screen* 15: 62–71

Wu H, Zeng H, Lam R, Tempel W, Amaya MF, Xu C, Dombrovski L, Qiu W, Wang Y & Min J (2011) Structural and Histone Binding Ability Characterizations of Human PWWP Domains. *PLOS ONE* 6: e18919

Xiao W, Adhikari S, Dahal U, Chen Y-S, Hao Y-J, Sun B-F, Sun H-Y, Li A, Ping X-L, Lai W-Y, *et al* (2016) Nuclear m(6)A Reader YTHDC1 Regulates mRNA Splicing. *Mol Cell* 61: 507–519

Xie S, Jakoncic J & Qian C (2012) UHRF1 Double Tudor Domain and the Adjacent PHD Finger Act Together to Recognize K9me3-Containing Histone H3 Tail. *Journal of Molecular Biology* 415: 318–328

Xing Y, Nandakumar A, Kakinen A, Sun Y, Davis TP, Ke PC & Ding F (2021) Amyloid Aggregation under the Lens of Liquid–Liquid Phase Separation. *J Phys Chem Lett* 12: 368–378

Xiong Y, Greschik H, Johansson C, Seifert L, Bacher J, Park K, Babault N, Martini M, Fagan V, Li F, *et al* (2019) Discovery of a Potent and Selective Fragment-like Inhibitor of Methyllysine Reader Protein Spindlin 1 (SPIN1). *J Med Chem*

Yang P, Mathieu C, Kolaitis R-M, Zhang P, Messing J, Yurtsever U, Yang Z, Wu J, Li Y, Pan Q, *et al* (2020) G3BP1 Is a Tunable Switch that Triggers Phase Separation to Assemble Stress Granules. *Cell* 181: 325-345.e28

Yang Y & Bedford MT (2013) Protein arginine methyltransferases and cancer. *Nat Rev Cancer* 13: 37–50

Yang Y, Hadjikyriacou A, Xia Z, Gayatri S, Kim D, Zurita-Lopez C, Kelly R, Guo A, Li W, Clarke SG, *et al* (2015) PRMT9 is a Type II methyltransferase that methylates the splicing factor SAP145. *Nat Commun* 6: 1–12

Yang Y, Lu Y, Espejo A, Wu J, Xu W, Liang S & Bedford MT (2010) TDRD3 is an effector molecule for arginine-methylated histone marks. *Mol Cell* 40: 1016–1023

Yano T, Nakamura T, Blechman J, Sorio C, Dang CV, Geiger B & Canaani E (1997) Nuclear punctate distribution of ALL-1 is conferred by distinct elements at the N terminus of the protein. *Proceedings of the National Academy of Sciences* 94: 7286–7291

Ye X, Chen G, Jin J, Zhang B, Wang Y & Ye F (2019) The development of inhibitors targeting the Mixed lineage leukemia 1 (MLL1)-WD repeat domain 5 protein (WDR5) Protein-Protein Interaction. *Curr Med Chem*

Yu J, Chen M, Huang H, Zhu J, Song H, Zhu J, Park J & Ji S-J (2018) Dynamic m6A modification regulates local translation of mRNA in axons. *Nucleic Acids Research* 46: 1412–1423

Zamparini AL, Davis MY, Malone CD, Vieira E, Zavadil J, Sachidanandam R, Hannon GJ & Lehmann R (2011) Vreteno, a gonad-specific protein, is essential for germline development and primary piRNA biogenesis in Drosophila. *Development* 138: 4039–4050

Zaware N & Zhou M-M (2019) Bromodomain biology and drug discovery. *Nat Struct Mol Biol* 26: 870–879

Zbinden A, Pérez-Berlanga M, De Rossi P & Polymenidou M (2020) Phase Separation and Neurodegenerative Diseases: A Disturbance in the Force. *Dev Cell* 55: 45–68

Zhang D, Yang J-F, Gao B, Liu T-Y, Hao G-F, Yang G-F, Fu L-J, Chen M-X & Zhang J (2019) Identification, evolution and alternative splicing profile analysis of the splicing factor 30 (SPF30) in plant species. *Planta* 249: 1997–2014

Zhang H, Liu K, Izumi N, Huang H, Ding D, Ni Z, Sidhu SS, Chen C, Tomari Y & Min J (2017) Structural basis for arginine methylation-independent recognition of PIWIL1 by TDRD2. *PNAS* 114: 12483–12488

Zhang L, Zhang Y, Chen Y, Gholamalamdari O, Wang Y, Ma J & Belmont AS (2020) TSA-seq reveals a largely conserved genome organization relative to nuclear speckles with small position changes tightly correlated with gene expression changes. *Genome Res*: gr.266239.120

Zhang M, Yang C, Zhu M, Qian L, Luo Y, Cheng H, Geng R, Xu X, Qian C & Liu Y (2021) Saturated fatty acids entrap PDX1 in stress granules and impede islet beta cell function. *Diabetologia* 64: 1144–1157

Zhang Q, Kota KP, Alam SG, Nickerson JA, Dickinson RB & Lele TP (2016) Coordinated Dynamics of RNA Splicing Speckles in the Nucleus. *J Cell Physiol* 231: 1269–1275

Zhou K, Chen Q, Chen J, Liang D, Feng W, Liu M, Wang Q, Wang R, Ouyang Q, Quan C, *et al* (2022) Spatiotemporal regulation of insulin signaling by liquid-liquid phase separation. *Cell Discov* 8: 64

Zhu H, Wei T, Cai Y & Jin J (2020) Small Molecules Targeting the Specific Domains of Histone-Mark Readers in Cancer Therapy. *Molecules* 25: 578

Zhu L & Brangwynne CP (2015) Nuclear bodies: the emerging biophysics of nucleoplasmic phases. *Curr Opin Cell Biol* 34: 23–30

Zhu R, Wang X, Yu Q, Guo W & Zhu L (2023) A systems biology-based approach to screen key splicing factors in hepatocellular carcinoma. *Molecular Carcinogenesis* n/a

Zhu X, Petrovski S, Xie P, Ruzzo EK, Lu Y-F, McSweeney KM, Ben-Zeev B, Nissenkorn A, Anikster Y, Oz-Levi D, *et al* (2015) Whole-exome sequencing in undiagnosed genetic diseases: interpreting 119 trios. *Genet Med* 17: 774–781

6. CURRICULUM VITAE

Personal Data

Address	Staudgasse 11/25 1180 Vienna, Austria
Mobile	+43 670 6046069
E-Mail	lennart.enders@gmx.de
Date, place of birth	March 8, 1993; Heidelberg, Germany
Marital status	Unmarried

Professional experience/ internships

October 2017- today	PhD student at CeMM, Research Center for Molecular Medicine of the Austrian Academy of Sciences, Vienna Laboratory Stefan Kubicek, Chemical Epigenetics, Control of cellular identity
November 2016- June 2017	Master thesis “The role of the piRNA machinery in intergenerational metabolic inheritance” at the Max Planck Institute of Immunobiology and Epigenetics, Freiburg Laboratory J. Andrew Pospisilik, Epigenetic control of complex disease (Diabetes and obesity)
April 2015- June 2017	Internship and following employment as research assistant at the Max Planck Institute of Immunobiology and Epigenetics, Freiburg Laboratory J. Andrew Pospisilik, Epigenetic control of complex disease (Diabetes and obesity)
April- July 2014	Bachelor thesis “Establishment of a robust workflow for cell surface proteomics” at the Institute of Molecular Medicine and Cell Research, University of Freiburg Laboratory Dr. Oliver Schilling, Proteases and Degradomics
May 2013- July 2014	Employment as research assistant at the Institute of Molecular Medicine and Cell Research, University of Freiburg Laboratory Dr. Oliver Schilling, Proteases and Degradomics
July 2012	Internship at the Institute of Anatomy and Cell Biology, Department of Molecular Embryology, University of Freiburg Laboratory Dr. Björn Spittau, Parkinson’s disease
November 2010- April 2011	Civil service in the nephrologic/ hepatologic ward of the pediatric clinic Heidelberg

Education

April 2015- June 2017	University of Freiburg Master of Science Molecular Medicine Final grade: very good / 1.2 (in a scale from 1 to 6; 1.0 is best)
October 2011- August 2014	University of Freiburg Bachelor of Science Molecular Medicine Final grade: very good / 1.5 (in a scale from 1 to 6; 1.0 is best)
2004-2010	Secondary school St. Raphael Gymnasium Heidelberg General qualification for university entrance after 8 years in a class for intellectually gifted students Final grade: 1.3 (in a scale from 1 to 6; 1.0 is best)
2002-2003	Secondary school Ottheinrich-Gymnasium Wiesloch

Additional Skills/ Achievements

Fellowships	**2019: DOC-Fellowship** of the Austrian Academy of Sciences, 38,000 Euros per year
Leadership/ Mentoring	**2019:** Supervised a rotating PhD student **2021:** Supervised a summer intern student **2020-2021:** PhD Student Representative
Poster presentations	**2018:** Islet Study Group of the European Association for the Study of Diabetes, Potsdam, Germany **2019:** Scientific meeting of the EU-LIFE institutes, Signalling & Gene Regulation in Health & Disease, Babraham Institute, Cambridge, UK **2020:** Cell Symposia: Biological Assemblies: Phase Transitions and More, online meeting **2020:** Annual Meeting of the European Association for the Study of Diabetes, online meeting **2022:** Gordon Research Conference on Post-Transcriptional Gene Regulation, Newry, ME, USA
Science communication	Produced a video for KinderUni/ Children's University
Certificates	Lab Animal Training FELASA B Sicherheit in der Gentechnik nach § 15 GenTSV
Software	MS Office, ImageJ, Cell Profiler, Zeiss ZEN, PerkinElmer Harmony, Prism, Spotfire, RStudio, Basic data analysis with Python, FlowJo, SnapGene, Benchling

Languages	**German:** native language
	English: full professional proficiency
	Spanish: limited working proficiency
	French: elementary proficiency

Publications

Pérez-Mojica, J.E., **Enders, L.,** Walsh, J., Lau, K.H., Lempradl, A., 2023. Continuous transcriptome analysis reveals novel patterns of early gene expression in Drosophila embryos. **Cell Genomics** 3, 100265. https://doi.org/10.1016/j.xgen.2023.100265

Casteels, T., Bajew, S., Reiniš, J., **Enders, L.**, Schuster, M., Fontaine, F., Müller, A.C., Wagner, B.K., Bock, C., Kubicek, S., **2022**. SMNDC1 links chromatin remodeling and splicing to regulate pancreatic hormone expression. **Cell Reports** 40, 111288. https://doi.org/10.1016/j.celrep.2022.111288

Tsiantoulas, D., Eslami, M., Obermayer, G., Clement, M., Smeets, D., Mayer, F.J., Kiss, M.G., **Enders, L.**, Weißer, J., Göderle, L., Lambert, J., Frommlet, F., Mueller, A., Hendrikx, T., Ozsvar-Kozma, M., Porsch, F., Willen, L., Afonyushkin, T., Murphy, J.E., Fogelstrand, P., Donzé, O., Pasterkamp, G., Hoke, M., Kubicek, S., Jørgensen, H.F., Danchin, N., Simon, T., Scharnagl, H., März, W., Borén, J., Hess, H., Mallat, Z., Schneider, P., Binder, C.J., **2021**. APRIL limits atherosclerosis by binding to heparan sulfate proteoglycans. **Nature** 597, 92–96. https://doi.org/10.1038/s41586-021-03818-3

Lempradl, A., Kugelberg, U., Iconomou, M., Beddows, I., Nätt, D., Casas, E., Örkenby, L., **Enders, L.**, Martinez, A.G., Lushchak, O., Boonen, E., Rückert, T., Sabev, M., Roth, M.G.L., Pettinga, D., Vavouri, T., Öst, A., Pospisilik, J.A., **2021**. Intergenerational metabolic priming by sperm piRNAs. **bioRxiv.** https://doi.org/10.1101/2021.03.29.436592

Lu, T.T.-H., Heyne, S., Dror, E., Casas, E., Leonhardt, L., Boenke, T., Yang, C.-H., Sagar, Arrigoni, L., Dalgaard, K., Teperino, R., **Enders, L.**, Selvaraj, M., Ruf, M., Raja, S.J., Xie, H., Boenisch, U., Orkin, S.H., Lynn, F.C., Hoffman, B.G., Grün, D., Vavouri, T., Lempradl, A.M., Pospisilik, J.A., **2018**. The Polycomb-Dependent Epigenome Controls β Cell Dysfunction, Dedifferentiation, and Diabetes. **Cell Metab.** 27, 1294-1308.e7. https://doi.org/10.1016/j.cmet.2018.04.013