Inaugural dissertation

for

obtaining the doctoral degree

of the

Combined Faculty of Mathematics, Engineering and Natural Sciences

of the

Ruprecht - Karls - University

Heidelberg

Presented by

M.Sc. Varshni Rajagopal

born in: Chennai (India)

Oral examination: 07.03.2025

RNA-dependent proteins in lung cancer and
their riboregulation in mitosis

Referees: Prof. Dr. Gislene Pereira

PD. Dr. Maïwen Caudron-Herger

Table of Contents

Declaration **1**

Abstract **3**

Zusammenfassung **5**

List of figures **9**

List of tables **11**

Abbreviations **13**

1 Introduction **17**

1.1 Coding and non-coding RNAs *17*

1.2 The RNA-binding proteins (RBPs) *19*

1.3 Conventional RNA-binding proteins *20*

1.4 Unconventional RNA-binding proteins *21*

1.5 RNA-binding proteins in human diseases *21*

1.6 Proteome-wide screens to identify RNA-binding proteins *22*

1.7 Proteome-wide screen to identify RNA-dependent proteins (R-DeeP) *24*

1.8 The different phases of cell cycle *28*

1.9 Cell division: a brief overview *30*

1.10 Cancer: a disease linked to cell cycle defects *35*

2 Aim of the thesis **39**

3 Materials and methods **41**

3.1 Gene ontology analysis *41*

3.2 Cell culture *41*

3.3 Cell synchronization *41*

3.4 R-DeeP screen *42*

3.5 SDS-PAGE and western blot analysis *42*

3.6 AURKA immunoprecipitation (IP) followed by LC-MS/MS-based protein analysis *43*

3.7 Protein digestion of AURKA IP samples for LC-MS/MS analysis 44

3.8 LC-MS/MS analysis of AURKA IP 45

3.9 Data analysis of LC-MS/MS data 45

3.10 AURKA immunoprecipitation (IP) 46

3.11 Proximity ligation assay (PLA) 47

3.12 PLA +/- RNase 47

3.13 Individual-nucleotide resolution UV cross-linking and immunoprecipitation (iCLIP2) 48

3.14 iCLIP2 library preparation and sequence analysis 50

3.15 RNA affinity purification (RAP) 54

3.16 Cloning 56

3.17 In vitro kinase assay 58

4 Results 65

4.1 Part I: R-DeeP screen in lung cancer cells 65

4.1.1 Analysis of the RNA-dependent shifts 66

4.1.2 Properties of shifting proteins 69

4.1.3 Western blot validation of the R-DeeP screen 71

4.1.4 Direct RNA binding of the RNA-dependent proteins 80

4.2 Part II: RNA-dependent functions of AURKA in mitosis 83

4.2.1 AURKA is an RNA-binding protein 86

4.2.2 RNA-dependent interactors of AURKA 89

4.2.3 RNA dependence of AURKA interactors KIFC1 and TPX2 92

4.2.4 Interaction of AURKA-KIFC1 and TPX2-KIFC1 across cell cycle 95

4.2.5 Validating RNA-dependent interactions of KIFC1 and TPX2 with AURKA 98

4.2.6 KIFC1 is an RNA-binding protein 109

4.2.7 KIFC1 - Yet another substrate of AURKA in mitosis 118

5 Discussion 125

6 Acknowledgements 133

7 References 135

Declaration

The analyses presented in the first part of this thesis have been included in a peer-reviewed publication:
"**Proteome-Wide Identification of RNA-Dependent Proteins in Lung Cancer Cells**" by Rajagopal et al., Cancers, 2022, 14(24):6109. DOI: https://doi.org/10.3390/cancers14246109

This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license. According to MDPI open access license, no special permission is required to reuse all or part of the article, including figures and tables. For article published under an open access Creative Common CCBY license, any part of the article may be reused without permission provided that the original article is clearly cited. In addition, the "author contributions" section mentions that V.R. (Varshni Rajagopal) wrote the original draft.

The analyses of the second part of this thesis have been included in a manuscript currently under peer review procedure. The current version of the manuscript is accessible on the bioRxiv server under a under a CC-BY-NC-ND 4.0 International license:
"**An atlas of RNA-dependent proteins in cell division reveals the riboregulation of mitotic protein-protein interactions**" by Rajagopal et al., bioRxiv, 2024.
DOI: https://doi.org/10.1101/2024.09.25.614981

Accordingly, I declare that I have included in the second part of this thesis modified parts of this manuscript including text and figures. In addition, the "author contributions" clearly mentions that V.R. (Varshni Rajagopal) wrote the initial version of the manuscript. The "declaration of interests" mentions that the study is part of the PhD thesis of V.R.

In the following, both manuscripts are thoroughly cited at the appropriated positions.

Abstract

From the onset of their transcription in the nucleus until their degradation in the cytoplasm, Ribonucleic acid (RNA) transcripts associate with RNA-binding proteins (RBPs) to form ribonucleoprotein (RNP) complexes. RNPs are dynamic macromolecular assemblies that regulate the fate of RNA molecules by coordinating all aspects of their post transcriptional maturation and regulation such as splicing, modification, transport, translation and decay. Conversely, recent studies pointed to a more RNA centric view of RNA-protein interactions, termed "riboregulation", where RNA modulates RBP localization, conformation, interactions and function. RBPs can bind their RNA targets with a wide range of affinities, modulated by *e.g.* post-translational modifications, interacting partners and structure. Due to their central role in various key cellular processes, dysregulation of RBP functions is implicated in the initiation and development of diseases such as neurological disorders, cancer and muscular atrophies. RBPs represent one of the largest protein groups and >1000 RBPs are associated with diseases. Therefore, RBPs have attracted increased interest in the past years, leading to the development of multiple strategies and tools to establish a comprehensive catalogue of RBPs in human and other various species.

Complementary methods such as the R-DeeP screen have been developed to identify RNA-dependent proteins (proteins that rely on RNA to form larger RNPs and directly or indirectly bind to RNA). R-DeeP is based on sucrose density gradient ultracentrifugation and fractionation, followed by mass spectrometry analysis. It provides quantitative information on the fraction of a protein being RNA dependent as well as allows the reconstruction of protein complexes based on co-segregation. Hence, to obtain a comprehensive view on RBPs in lung cancer, in the first part of my PhD project, I took advantage of the adaptability of the R-DeeP screen and identified RNA-dependent proteins in A549 lung adenocarcinoma cells. With this proteome-wide technique, I identified 1189 RNA-dependent proteins which includes 170 proteins which had never been linked to RNA before. Out of the 170 novel RNA-dependent proteins, I validated the RNA dependence of three newly identified shifting proteins: DOCK5, ELMO2, and ABRAXAS1 using western blot analysis. Further, the direct RNA interaction of the cell migration-related protein DOCK5 and BRCA1-associated protein ABRAXAS1 was verified using iCLIP2. The R-DeeP 2.0 database https://R-DeeP2.dkfz.de provides proteome-wide and cell line-specific information on proteins and their RNA dependence from A549 and HeLa S3 cell lines. This study contributes to increase our understanding of the functional role of RNA and RNA-binding proteins in cancer cells.

Given the role of RBPs in a wide range of cellular processes, and particularly in the context of cell division, several studies have reported the presence of RNA, *i.e*. protein-coding (mRNAs), non-coding RNAs (ncRNAs) and various RBPs within structures of the mitotic spindle apparatus such as the centrosomes and the spindle MTs. Collectively, these works pointed to the importance of RNA and RBPs for the structural and functional integrity of the mitotic spindle. However, the underlying mechanisms remain unknown. A recent R-DeeP screen in HeLa cells synchronized in prometaphase generated a huge resource on RNA-dependent proteins with cell cycle-specific information.

In the second part of my PhD project, I took advantage of this new resource and identified the RNA dependence of several key mitotic factors: AURKA, KIFC1 and TPX2. Further using immunoprecipitation coupled with mass spectrometry analysis, I uncovered new interacting partners of AURKA including KIFC1. KIFC1 interacted with AURKA and TPX2 in an RNA-dependent manner. Importantly, I discovered that TPX2 also interacted with AURKA in an RNA-dependent fashion. Though the interaction between AURKA and TPX2 has been well characterized, their RNA-mediated interaction was unknown. To uncover the functional significance of the interaction between AURKA and KIFC1, an *in vitro* kinase assay was performed and it revealed that AURKA phosphorylated KIFC1 in an RNA-dependent manner at S349 and T359 amino acid residues. With the aim to identify the RNA targets mediating their interactions, iCLIP2-sequencing of KIFC1-bound RNAs was performed in HeLa cells synchronized in prometaphase. The sequencing results demonstrated that, KIFC1 predominantly bound to ribosomal RNAs (rRNAs) and protein-coding RNAs (mRNAs) in prometaphase without sequence specificity. Though, KIFC1 lacked sequence specificity in binding to RNA, I identified that these RNA targets also bound AURKA and TPX2 in prometaphase cell lysates. These data suggest that these three key mitotic factors interact with each other and exist within the same complex. Notably, RNA played a crucial role in mediating their interaction, which is vital for spindle assembly, indicating the riboregulation of mitotic protein-protein interactions during cell division. This offers new perspectives on the control of cell division processes by RNA-protein complexes.

Zusammenfassung

Vom Beginn ihrer Transkription im Zellkern bis zu ihrem Abbau im Zytoplasma verbinden sich RNA-Transkripte mit RNA-bindenden Proteinen (RBPs) zu Ribonukleoprotein-Komplexen (RNP). RNPs sind dynamische makromolekulare Einheiten, die die Zukunft von RNA-Molekülen regulieren, indem sie alle Aspekte ihrer posttranskriptionellen Reifung und Regulierung wie Spleißen, Modifikation, Transport, Übersetzung und Abbau koordinieren. Umgekehrt weisen neuere Studien auf eine eher RNA-zentrierte Sichtweise der RNA-Protein-Interaktionen hin, die als „Riboregulation" bezeichnet wird und bei der RNA die Lokalisierung, Konformation, Interaktionen und Funktion von RBPs moduliert. RBPs können RNA-transkripte mit einer großen Bandbreite an Affinitäten binden, die z. B. durch posttranslationale Modifikationen, interagierende Partner und Struktur moduliert werden. Aufgrund ihrer zentralen Rolle in verschiedenen zellulären Schlüsselprozessen wird eine Dysregulation der RBP-Funktionen mit der Entstehung und Entwicklung von Krankheiten wie neurologischen Störungen, Krebs und Muskelatrophien in Verbindung gebracht. RBPs stellen eine der größten Proteingruppen dar, und mehr als 1000 RBPs werden mit Krankheiten in Verbindung gebracht. Daher sind RBPs in den letzten Jahren auf zunehmendes Interesse gestoßen, was zur Entwicklung zahlreicher Strategien und Instrumente geführt hat, um einen umfassenden Katalog von RBPs beim Menschen und bei anderen Tierarten zu erstellen.

Ergänzende Methoden wie das R-DeeP-Screen wurden entwickelt, um RNA-abhängige Proteine zu identifizieren (Proteine, die auf RNA angewiesen sind, um größere RNPs zu bilden und direkt oder indirekt an RNA binden). R-DeeP basiert auf Saccharose-Dichtegradienten-Ultrazentrifugation und Fraktionierung, gefolgt von einer massenspektrometrischen Analyse. Es liefert quantitative Informationen über den Anteil eines Proteins, der RNA-abhängig ist, und ermöglicht die Rekonstruktion von Proteinkomplexen auf der Grundlage der Ko-Segregation. Um einen umfassenden Überblick über RBPs bei Lungenkrebs zu erhalten, habe ich im ersten Teil meines Promotionsprojekts die Anpassungsfähigkeit des R-DeeP-Screens genutzt und RNA-abhängige Proteine in A549-Lungenadenokarzinomzellen identifiziert. Mit dieser proteomweiten Technik identifizierte ich 1189 RNA-abhängige Proteine, darunter 170 Proteine, die noch nie mit RNA in Verbindung gebracht worden waren. Von den 170 neuen RNA-abhängigen Proteinen habe ich die RNA-Abhängigkeit von drei neu identifizierten Shifting-Proteinen validiert: DOCK5, ELMO2 und ABRAXAS1 mittels Western-Blot-Analyse. Außerdem wurde die direkte RNA-Interaktion des migrationsbezogenen Proteins DOCK5 und des BRCA1-assoziierten Proteins ABRAXAS1 mit Hilfe von iCLIP2 verifiziert.

Die R-DeeP 2.0-Datenbank https://R-DeeP2.dkfz.de liefert proteomweite und zelllinienspezifische Informationen über Proteine und ihre RNA-Abhängigkeit aus den Zelllinien A549 und HeLa S3. Diese Studie trägt dazu bei, die funktionelle Rolle von RNA und RNA-bindenden Proteinen in Krebszellen zu verstehen.

In Anbetracht der Rolle von RBPs in zahlreichen zellulären Prozessen wurde im Zusammenhang mit der Zellteilung in mehreren Studien über das Vorhandensein von RNA, d. h. von proteincodierenden (mRNAs), nichtcodierenden RNAs (ncRNAs) und verschiedenen RBPs in Strukturen des mitotischen Spindelapparats wie den Zentrosomen und den Spindelmikrotubuli berichtet. Insgesamt weisen diese Arbeiten auf die Bedeutung von RNA und RBPs für die strukturelle und funktionelle Integrität der mitotischen Spindel hin. Die zugrunde liegenden Mechanismen sind jedoch nach wie vor unbekannt. Der jüngste R-DeeP Screen in HeLa Zellen, die in der Prometaphase synchronisiert wurden, lieferte eine große Menge an RNA-abhängigen Proteinen mit zyklusspezifischen Informationen.

Im zweiten Teil meines Promotionsprojekts nutzte ich diese neue Ressource und identifizierte die RNA-Abhängigkeit mehrerer wichtiger mitotischer Faktoren: AURKA, KIFC1 und TPX2. Durch Immunpräzipitation in Verbindung mit Massenspektrometrie-Analysen habe ich außerdem neue Interaktionspartner von AURKA wie z.B. KIFC1 aufgedeckt. KIFC1 interagierte mit AURKA und TPX2 auf RNA-abhängige Weise. Dabei entdeckte ich, dass auch TPX2 mit AURKA in einer RNA-abhängigen Weise interagierte. Obwohl die Interaktion zwischen AURKA und TPX2 gut charakterisiert ist, war ihre RNA-vermittelte Interaktion bisher unbekannt. Um die funktionelle Bedeutung der Interaktion zwischen AURKA und KIFC1 aufzudecken, wurde ein *in vitro* Kinase-Assay durchgeführt, wodurch ich fest stellte, dass AURKA KIFC1 in RNA-abhängiger Weise an den Aminosäureresten S349 und T359 phosphorylierte. Mit dem Ziel, die Ziel-RNAs zu identifizieren, die KIFC1 Protein-Protein Interaktionen vermitteln, wurde eine iCLIP2-Sequenzierung von KIFC1-gebundenen RNAs in HeLa Zellen durchgeführt, die in der Prometaphase synchronisiert wurden. Die Sequenzierungsergebnisse zeigten, dass KIFC1 in der Prometaphase vorwiegend an ribosomale RNAs (rRNAs) und proteinkodierende RNAs (mRNAs) ohne Sequenzspezifität gebunden war. Obwohl KIFC1 bei der Bindung an RNA keine Sequenzspezifität aufwies, konnte ich feststellen, dass diese KIFC1-gebundenen RNAs in Prometaphase-Zelllysaten auch AURKA und TPX2 gebunden haben.

Diese Daten deuten darauf hin, dass diese drei mitotischen Schlüsselfaktoren miteinander interagieren und im selben Komplex existieren. Vor allem spielte die RNA eine entscheidende Rolle bei der Vermittlung ihrer Interaktion, die für den Spindelaufbau und die korrekte

Zellteilung von entscheidender Bedeutung ist. Insgesamt deuten diese Daten auf eine Riboregulierung der mitotischen Protein-Protein-Interaktionen während des Spindelaufbaus hin und eröffnen neue Perspektiven für die Kontrolle von Zellteilungsprozessen durch RNA-Protein-Komplexe.

Übersetzt mit DeepL.com (kostenlose Version)

List of figures

Figure 1: Types of RNA *18*
Figure 2: mRNA life cycle *20*
Figure 3: Potential consequences of mutation of an RBP *22*
Figure 4: R-DeeP sucrose density gradient method *26*
Figure 5: Data integration and representation in R-DeeP 2.0 database *27*
Figure 6: Cell cycle stages *29*
Figure 7: Cell division cycle *34*
Figure 8: Hallmarks of cancer. *36*
Figure 9: Schematic describing the R-DeeP method in lung cancer cells *66*
Figure 10: Analysis of the RNA-dependent shifts *68*
Figure 11: Properties of shifting proteins *70*
Figure 12: Western blot validation of the R-DeeP screen *74*
Figure 13: Validating the RNA dependence of DOCK5, ELMO2 and ABRAXAS1 using western blot analysis *80*
Figure 14: Direct RNA binding of the DOCK5 and ABRAXAS1 ..... *81*
Figure 15: R-DeeP 2.0 Database *82*
Figure 16: RNA-dependent proteins are significantly enriched in mitotic factors *83*
Figure 17: AURKA is an RNA-dependent protein in mitosis *87*
Figure 18: AURKA is an RNA-binding protein *89*
Figure 19: RNA-mediated proteins interactors of AURKA in mitosis *92*
Figure 20: RNA dependence of KIFC1 and TPX2 in mitosis *94*
Figure 21: PLA assay representing the interaction of AURKA-KIFC1 and TPX2-KIFC1 across cell cycle *98*
Figure 22: AURKA immunoprecipitation validating the RNA-dependent interactions of KIFC1 and TPX2 with AURKA in mitosis *99*
Figure 23: PLA assay validating the RNA-dependent interactions of KIFC1 and TPX2 with AURKA in mitosis *107*
Figure 24: Interaction between α-tubulin and β-tubulin upon RNase treatment *108*
Figure 25: KIFC1 is an RNA-binding protein *110*
Figure 26: KIFC1 binds to ribosomal RNA and protein-coding RNA transcripts in mitosis *112*
Figure 27: RNA affinity purification of KIFC1-bound transcripts in HeLa prometaphase cells *114*
Figure 28: Kinase assay showing the phosphorylation of KIFC1 by AURKA *119*
Figure 29: AURKA phosphorylates KIFC1 at S349 and T359 in an RNA-dependent manner *121*

List of tables

Table 1: gBlock DNA sequences used for RNA affinity pulldown *61*
Table 2: Primers, adapters and barcode sequences used for iCLIP-Seq of KIFC1 ..... *62*
Table 3: Primer and other sequences used for cloning and in vitro kinase assay *63*
Table 4: Equipments used in the project *64*
Table 5: Software used in the project *64*
Table 6: Top 150 KIFC1-bound RNAs in prometaphase *114*

Abbreviations

Abbreviation	Full form
%	Percent
4SU	4-thiouridine
6SG	6-thioguanosine
aa	Amino acid
ABRAXAS1	BRCA1 A complex subunit
APC/C	Anaphase promoting complex/cyclosome
ASNS	Asparagine synthetase
ATM	Ataxia telangiectasia mutated protein
ATP	Adenosine triphosphate
ATR	Ataxia telangiectasia and Rad3 related protein
AURKA	Aurora kinase A
AURKB	Aurora kinase B
CARIC	Click chemistry-assisted RNA interactome capture
CDK**1**	Cyclin-dependent kinase 1
CEP192	Centrosomal protein 192
CHK1	Checkpoint kinase 1
CHK2	Checkpoint kinase 2
CLASP1	Cytoplasmic linker associated protein 1
CLIP	Crosslinking immunoprecipitation
CTCF	CCCTC-binding factor protein
DNA	Deoxyribonucleic acid
DOCK5	Dedicator of cytokinesis protein 5
dsDNA	Double-stranded DNA
dsRNA	Double-stranded RNA
dT	Deoxythymidines
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic acid
ELMO2	Engulfment and cell motility protein 2
EU	Ethynyl uridine
G0	Resting phase
G1	Gap 1
G2	Gap 2
GAPDH	Glyceraldehyde-3-phosphate dehydrogenase
GO	Gene ontology
GTP	Guanosine-5'-triphosphate
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HMGA1	High-mobility group protein
hnRNA	Heterogeneous nuclear RNA
hnRNP	Heterogeneous nuclear RNP
HNRNPU	Heterogeneous nuclear ribonucleoprotein U

HSP	Heat shock protein
HuR	Hu-antigen R
iCLIP2	Improved individual-nucleotide resolution UV crosslinking and immunoprecipitation
IDR	Intrinsically disordered region
IgG	Immunoglobulin G
INCENP	Inner centromere protein
IP	Immunoprecipitation
Kcl	Potassium Chloride
KH	K-homology domain
KIF14	Kinesin family member 14
KIFC1	Kinesin family member C, also termed HSET
KT	Kinetochores
LC-MS	Liquid chromatography–mass spectrometry
LDS	Lithium dodecyl sulphate
let-7	Lethal-7
lncRNA	Long non-coding RNA
M	Mitosis
MALAT1	Metastasis-associated lung adenocarcinoma transcript 1
MAPs	Microtubule-associated proteins
MCM2	Minichromosome maintenance complex component 2
MEG3	Maternally expressed gene 3
MgCl2	Magnesium Chloride
miRNA	MicroRNA
MKLP1	Mitotic kinesin-like protein 1
mRNA	Messenger RNA
mRNP	Messenger RNP
MS	Mass spectrometry
Ms	Mouse
MT	Microtubule
NaCl	Sodium Chloride
ncRNA	Non-coding RNA
NEB	Nuclear envelope breakdown
nm	Nanometer
nt	Nucleotide
NuMA	Nuclear mitotic apparatus protein
P32	Phosphorus-32
p53	Tumor protein 53
PAR-CLIP	Photoactivable ribonucleoside enhanced crosslinking immunoprecipitation
PCR	Polymerase chain reaction
piRNA	Piwi-associated RNAs
PLA	Proximity ligation assay
PLK1	Polo-like kinase 1

PNK	Polynucleotide kinase
Poly-A	Polyadenylation
PPI	Protein-protein interaction
PRC1	Protein regulator of cytokinesis 1
pre-mRNA	Pre-messenger RNA
PTex	Phenol toluol extraction
R-DeeP	RNA-dependent proteins
RACGAP1	Rac GTPase-activating protein 1
Rb	Rabbit
RBD	RNA-binding domain
RBDmap	RNA-binding domain map
RBP	RNA-binding protein
RICK	RNA interactome capture using click chemistry
RNA	Ribonucleic acid
RNP	Ribonucleoprotein
rpm	Revolutions per minute
RRM	RNA recognition motif
rRNA	Ribosomal RNA
RT	Room temperature
S	Synthesis
S349	Serine 349
SAC	Spindle assembly checkpoint
SDS-PAGE	Sodium dodecyl sulphate polyacrylamide gel electrophoresis
SEM	Standard error of the mean
serIC	Serial interactome capture
SNAIL1	Snail family transcriptional repressor 1
snRNA	Small nuclear RNA
T359	Threonine 359
TPX2	Targeting protein for Xklp2
tRNA	Transfer RNA
TUBA1C	Tubulin alpha 1c
UTP	Uridine-5′-triphosphate
UTR	Untranslated region
UV	Ultra-violet
WB	Western blot
ZF	Zinc finger protein 9
γ-TuRCs	γ-tubulin ring complexes

1 Introduction

RNA and RNA-binding proteins (RBPs) are gaining attention due to their crucial role in various cellular functions and disease-causing property upon deregulation. Towards the goal of understanding the molecular details of their functions and interactions, various proteome-wide screens were developed to identify RBPs. Using these proteome-wide screens, we observed an enrichment of mitotic factors within RBPs.

For the past decades, extensive research has been performed and we have an in-depth understanding of the mitotic-related pathways. Majority of the studies focused on deciphering the role of proteins and protein-protein interactions, with meagre focus on the role of RNA or RNA-binding proteins (RBPs) in such critical biological process with relevance to cancer.

Hence, in this thesis, I focused on identifying novel RBPs in cancer and investigated their role in mitosis. I will be introducing various topics related to my thesis in the following sections, starting from coding and non-coding RNAs to cancer – one of the major diseases caused due to the dysregulation of RNA, RBPs and mitotic factors.

1.1 Coding and non-coding RNAs

The discovery of RNA has a deep history that extends over a century. Starting with its initial identification, until the present, the field has evolved with a profound understanding of its biological significance. RNA is a regulatory molecule that is a part of Cricks "Central dogma of life", originally believed only to be the carrier of genetic information from DNA to the protein (1). Later, with the discovery of intronic and untranslated regions of messenger RNA (mRNA), along with the identification of a class of RNAs known as non-coding RNAs like small nuclear RNAs (snRNAs), heterogeneous nuclear RNAs (hnRNAs), micro RNAs (miRNAs) and long non-coding RNAs (lncRNAs), it became evidently clear that, RNA is a key regulatory molecule that modulates most cellular processes, largely independent of its ability to encode for proteins (1,2).

Aside from its role as a transitory template for protein synthesis, mRNA functions as a transcriptional and translational regulator. It mediates the localization and scaffolding of proteins through their UTRs, and is also responsible for their self-stability and translational efficiency (2,3). Whereas, in certain cases, like in P53 mRNA, the region responsible for the regulatory functions is not clearly separable from its coding region (3,4).

While 80% of the genome is transcribed, only a small fraction of the transcripts (less than 3%) is translated into proteins leaving the vast majority to be comprised of ncRNAs (Figure 1) (5,6). This major class of RNA - the ncRNAs are made of several different types that are classified based on their size and functions. The house keeping RNAs include rRNAs which are essential for protein translation, tRNAs which decode the mRNA into a protein by recruiting correct amino acids (aa) to the ribosomes and snRNAs which are essential for splicing. These housekeeping RNAs are essential for the normal cellular functions (2,5). Next, the small regulatory RNAs like small interfering RNAs (siRNAs), miRNAs and piwi-associated RNAs (piRNAs) that are roughly 25 nt in size and are key elements of gene regulation (2,5,7,8). Lastly, the lncRNAs that are about 200 nt long play crucial role in variety of biological processes like X-chromosome inactivation, gene regulation and maintenance of nuclear architecture (2,9,10). Additionally, these regulatory ncRNAs, especially, miRNAs and lncRNAs for example, let-7, miR-21, *MALAT1* and *MEG3*, received the most attention and are implicated in different cancers types and several other diseases (5,9-18). In addition to their role in diseases and other regulatory functions, RNAs mediate the formation of large protein complexes and regulate the fate of the protein in terms of localization and function. Importantly, all of these RNAs operate as RNA-protein complexes (2,19,20).

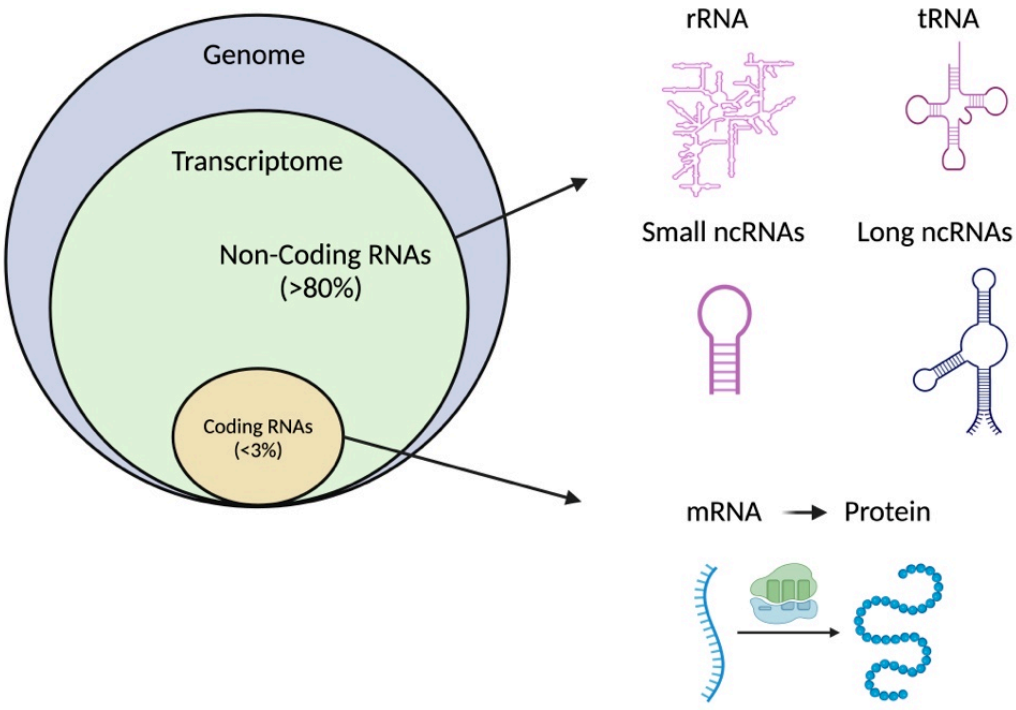


Figure 1: Types of RNA

Example of different types of coding and non-coding RNA in a cell (mRNA, rRNA, tRNA, small ncRNA and long ncRNA). This figure was created using BioRender.com.

1.2 The RNA-binding proteins (RBPs)

RNA transcripts associate with various RBPs to form ribonucleoprotein complexes (RNPs) that are dynamic in nature and fundamental for biological processes like splicing, translation, transcription and decay (20,21). The dynamic nature of RNPs is due to the frequent alterations in the RNP composition that depends on the cellular context and maturation state of the RNA transcript. RBPs regulate RNA transcripts throughout their life cycle (20-22). Here, processing of a freshly synthesized mRNA *i.e.* pre-mRNA or heterogenous nuclear RNAs (hnRNAs) would serve as a good example (Figure 2). As and when the pre-mRNAs are synthesized, the heterogenous ribonucleoprotein complexes (hnRNPs/mRNPs) associate with the formed pre-mRNAs co-transcriptionally. Many of these proteins remain bound to the mRNA until it is translated by the ribosome and shuttle between the nucleus and cytoplasm. Whereas, some proteins dissociate, while the others bind the mRNA as a result of splicing or maturation (23). mRNA is associated with various RBPs in every step of the process starting from the initial synthesis, 5' capping, splicing, 3' end cleavage, polyadenylation in the nucleus, followed by cytoplasmic export for protein synthesis by the ribosomes until its final decay in the cytoplasmic P-bodies, contributing to the dynamic state of RNPs (23,24). RBPs represent a large class of proteins (over 1000) (22), that are heterogenous and ubiquitous in nature. These RBPs are classified as conventional or unconventional RBPs based on the presence or absence of distinct structurally well-defined RNA-binding domains (RBDs) (21,22,25).

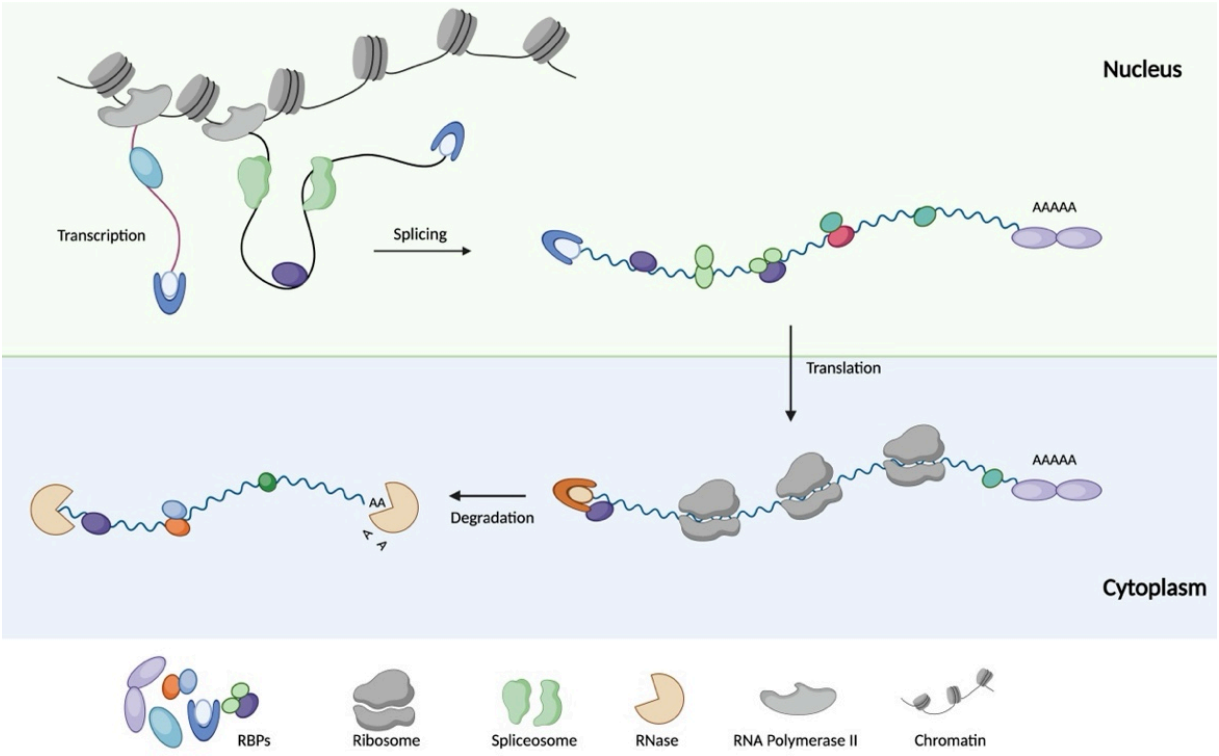


Figure 2: mRNA life cycle

Interaction of mRNA with various RBPs in different cellular compartments (nucleus and cytoplasm) and biological processes (transcription, splicing, translation and degradation). This figure was created using BioRender.com based on (20).

1.3 Conventional RNA-binding proteins

The diverse functions of the RBPs arise from the recognition of different RNA molecules mediated by the diverse combination of structurally well-defined RBDs. However, the number of these RNA-binding modules such as the RNA recognition motif (RRM), DEAD box helicase, Zinc finger domain (ZF) or K homology domain (KH) are quite limited (21,22,25). This structural diversity arises from a single or multiple copies of these RBDs distributed in a unique way that contributes to the specificity and functional repertoire of these RBPs. For example, an RBP might contain same or multiple modules of RBDs that are located near the enzymatic pockets that helps in the catalytic activity or define the specific targets, thereby leading to the formation of versatile macromolecular binding surfaces (21,26). Classic example of a canonical RBP is the CCCTC-binding factor (CTCF) protein. It is a transcriptional repressor and a key protein for chromatin organization. It is known to bind RNA through one of the typical RBDs, the ZF-binding domain (27-29). It contains 11 C2H2 type ZF domains that are distributed across the structure, with linker sequence between each domain. CTCF is known to bind RNA through its ZF1 (264–275 aa) and ZF10 (536–544 aa) that is distinct from

its DNA binding domain. Deletion of RNA-binding sites in CTCF leads to loss of ability to form chromatin loops and deletion of RNA-binding sites in ZF10 and ZF11 leads to loss of multimerization thereby affect higher-order chromatin structures (27,28). Though most of the biochemical and structural knowledge is based on decades of cumulative study on the limited pool of canonical RBPs, recent studies that determine the structure or large RNP complexes like spliceosome have reported the presence of complex RNA-protein interaction, including RBPs that do not contain conventional RBDs (21). These RBPs are termed unconventional RBPs (21).

1.4 Unconventional RNA-binding proteins

RBPs that lack discernible RBDs are classified as non-canonical or unconventional RBPs (20,21). They might contain intrinsically disordered regions (IDRs), which are the regions within the protein, that lack well-defined secondary or tertiary structure, with dynamic composition (20,22). They can switch into an ordered state upon RNA binding (22). Like the conventional RBDs, IDRs can be found multiple times in one RBP and can coordinate RNA binding with other globular RBDs. IDRs are known to have higher affinity to RNA and can solely orchestrate majority of protein-RNA interactions in a cell (22). In a proteome-wide study that mapped the RBDs in human cells, out of the 1,174 RNA-binding sites that were detected, nearly half of the RNA-binding sites mapped to the disordered regions (30). Though IDRs have very little specificity for RNA sequences, they represent malleable and multifunctional RNA-binding motifs and are known to drive liquid-liquid phase separation of their constituent RBPs (21,22). One example of non-canonical RBP is Enolase 1. It is a glycolytic enzyme that lacks classical RBDs and its RNA-binding activity is activated by acetylation. It is riboregulated by its interacting RNAs, mediates metabolic rewiring of mouse embryonic stem cells and alters stem cell differentiation (31).

1.5 RNA-binding proteins in human diseases

RBPs represent the largest protein group and are often implicated in pathological diseases like cancer, neurological diseases and muscular atrophies (20,32,33). There are distinct mechanisms by which a mutation can influence the function of RBPs and drive towards the onset of diseases (Figure 3). For instance, mutations in a gene encoding an RBP could lead to the synthesis of proteins with alternate isoforms or altered expression levels that affects the RBP function. It could modify the amino acid composition that could in turn alter RNA targets or binding

affinity. Also, it could lead to protein aggregation, mislocalization and affect the catalytic properties of enzymes (20,32,33). For example, HuR protein associates with metabolite UDP-glucose, that prevents the binding of HuR to *SNAIL1* mRNA under normal conditions. Upon mutation, interaction of HuR protein with UDP-glucose is abrogated that results in stabilization of *SNAIL1* mRNA. *SNAIL1* mRNA in turn encodes an epithelial to mesenchymal transition marker, resulting in invasive properties in various cancer cells (20,34). Given that RBPs regulate a vast network of RNA-protein and protein-protein interactions, interfering with their functions can impact several genes, pathways that could result in adverse phenotypes (32). Owing to the vast role of RBP in biological processes and diseases mechanisms, it is of utmost importance to identify, analyse and understand the RBP interactome and functions.

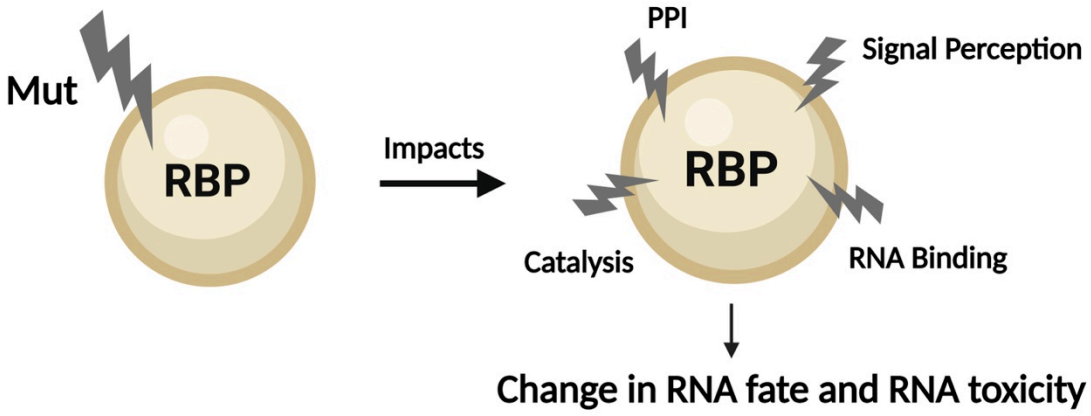


Figure 3: Potential consequences of mutation of an RBP

RBPs can influence protein-protein interaction (PPI), influence RNA binding or abolish the interaction with RNA, influence the enzymatic activity of the protein which in turn affects the RNA metabolism, signal perception and functions. This figure was created using BioRender.com.

1.6 Proteome-wide screens to identify RNA-binding proteins

Over the past decades various proteome-wide screens were developed to comprehensively and systematically identify RBPs, uncovering thousands of new ones. Initial attempts were established to identify the mRNA-bound proteome. First, the cells were cultured, and crosslinked using UV irradiation at 254 nm, that forms a covalent bond between proteins and RNA. Further, exploiting the presence of poly-A tail in the mRNA, DNA oligos with repeating deoxythymidine (oligo dT) were used to specifically capture the mRNAs. Using this concept, proteins interacting with the poly-A tail of mRNA were identified using oligo-dT capture and sepharose chromatography either using UV-crosslinked polysome fractions of purified mRNP complexes (35) or cells treated with (36) or without UV-irradiation (37). Similar approaches were implemented to identify and characterize mRNA poly-A-binding proteins (ploy-A) *i.e.*

the hnRNP particles (38,39). Later, 4-thiouridine (4SU) or 6-thioguanosine (6SG), were used that are readily taken up by cultured mammalian cells. Both photoreactive nucleoside analogs are metabolically incorporated into transcribed RNAs without detectable incorporation into DNA (40). The proteins were crosslinked to RNA using UV light at 365 nm, a technique which is also known as PAR-CLIP (photoactivatable ribonucleoside-enhanced crosslinking and immunoprecipitation) (40). The proteins bound to the labelled mRNA were then isolated using oligo-dT capture coupled with high-resolution quantitative mass spectrometry to identify mRNA-bound proteome and to globally map the RNA-binding sites on these proteins in mammalian cells (30,41). This concept was adapted and several mRNA-binding proteins were identified either using UV crosslinking at 254 nm or photoreactive nucleoside analogs with UV crosslinking at 365 nm or both in different mammalian cell compartments, cell types and also in different model organisms (42-48). One major limitation of these methods is the use of oligo-dT beads that limited the identification of RBPs binding to poly-A mRNAs and completely overlooked the proteins binding to non-polyadenylated RNAs. An additional limitation is the ineffective incorporation of 4SU and the inefficient crosslinking using UV light, though the method is highly sensitive.

Hence, to address these limitations, orthogonal strategies were employed to identify RBPs independently of poly-A binding property. That includes proteins binding to pre-mRNA, non-coding and long non-coding RNAs using techniques where, the RNA-protein complexes were isolated using their physicochemical properties. Different methods were adapted based the concept that, upon the lysis of UV irradiated cells with either acidic phenol-chloroform or phenol-toluol mixture and centrifugation, RNA molecules are separated into aqueous phase, protein and DNA molecules into organic phase whereas, the RNA-protein complexes are found in the interphasic layer mediating the efficient isolation of RBPs (49,50). While XRNAX (protein-crosslinked RNA extraction) technique implemented a single AGPC (guanidium-thiocyanate phenol-chloroform) extraction (50), the orthogonal organic phase separation (OOPS) method, utilised a double AGPC (guanidium-thiocyanate phenol-chloroform) extraction to purify crosslinked complexes from isolated RNA, protein and DNA and then to isolate the RNA and protein from the RNA-protein complexes followed by protease digestion (51). Additionally, the phenol toluol extraction (PTex) approach used double phenol toluol-based extraction to isolate RBPs. First phenol-toluol extraction was performed at a pH of 7, to retrieve RNA, proteins and crosslinked complexes in the aqueous phase, and to separate them from lipids and DNA. Then the second phenol-toluol extraction was performed at a pH below 5, to isolate the crosslinked complexes in the interphase (52). Although these methods isolated

RBPs without any bias towards RNA binding, they contain high background compared to other methods due to DNA, protein and lipid contamination, despite the washing conditions as described in the protocol.

Further, using modified nucleotides, RBPs binding to nascent RNAs were identified and various new RBPs including mitotic and metabolic factors were detected. Here, the cells were treated with 5-ethynyluridine (EU) which was then incorporated into newly synthesized RNA. The proteins and RNA were crosslinked using UV at 254 nm, lysed and biotin was attached to the EU-incorporated RNA, using click chemistry. Then, the RNA-protein complexes were isolated using pulldown with streptavidin beads (53). The same concept was employed in another method called CARIC (click chemistry-assisted RNA interactome capture), which combined the PAR-CLIP crosslinking step along with click chemistry (54). Although both methods yielded a comprehensive RBPome, devoid of any selection bias towards specific types of RNA, they may suffer from the disadvantage of identifying non-specific binding leading to false positives.

In addition to the experimental approaches, various RBPs were predicted based on machine learning algorithms for example, the SONAR (Support Vector Machine Obtained from Neighbourhood Associated RBPs) analysis, which depends on protein-protein interactions under assumptions that RBPs often interact with other RBPs (55). These massive efforts over the decades resulted in huge datasets in different species. Importantly, 43 different datasets of RBPs in human alone (25,56). Despite the existence of these huge datasets and databases (25,56-58), previous studies have showed meagre overlap between the datasets with several proteins being identified as an RBP only once (56). This highlights the need for complementary techniques to identify core set of novel RBPs, independent of physical or affinity-based purification.

1.7 Proteome-wide screen to identify RNA-dependent proteins (R-DeeP)

The R-DeeP screen is based on the concept of RNA dependence, according to which a protein is classified as RNA dependent if its interactome or eventually its functions are dependent on RNA. This includes proteins that directly and indirectly bind to RNA (59,60). The screen is devoid of any potential biases and is complementary to previously established techniques that were discussed above. The technique does not require crosslinking or additional experimental steps and thus, is independent of efficiency of RNA-protein crosslinking using UV light, incorporation of modified nucleotides, physiochemical properties or anti-sense oligos. This

aids the identification of conventional and non-conventional RBPs without any enrichment strategies. This screen was adapted based on the method developed for identifying the composition of argonaut mRNP complexes which also has similarities with polysome profiling. Both the methods were performed using sucrose density gradient and for example applied to identify the ribosomes present in the mRNA and in turn asses the translational activity (61,62). In summary, the underlying principle is that the migration pattern of a protein varies based on its interaction state. When a protein exists in a free, isolated state for *e.g*. as a monomer—meaning it is not interacting with other proteins or molecules—it is detected at a position in the sucrose density gradient that reflects its monomeric molecular weight. Conversely, if the protein is engaged in interactions with other proteins, particularly those mediated by RNA, it will exhibit a higher molecular mass as it becomes part of a complex. Consequently, it will be detected at a position that corresponds to its apparent molecular weight, in a fraction containing a higher sucrose density. This results in a distinct migration pattern for each protein, depending on the presence or absence of RNA, thereby highlighting the RNA dependence of the protein (59,60).

In summary, the cells were first cultured, lysed and the lysates were either treated with RNase to degrade the RNA transcripts or not treated with RNase *i.e.* the control lysates. After processing both the control and RNase-treated lysates, they were loaded on the sucrose density gradients which contains varying amounts of sucrose ranging from 5% to 50% and subjected to ultracentrifugation for 18 h at 110,000–115,000 g. Further, the gradients were fractionated and each fraction was analysed using quantitative mass spectrometry and followed by western blot analysis to establish migration profile for each protein in both the gradients (control and RNase-treated). If a protein requires RNA to form a complex or interact with other proteins, it is expected to be detected in the later fractions indicating its apparent higher molecular weight in the control gradients. In contrast, in RNase-treated gradients—where the absence of RNA disrupts complex formation or interactions with other proteins—the protein is expected to be detected in the earlier fractions (Figure 4) (59,60). This technique was successfully applied to identify over 700 novel RBPs in different cancer cell types like HeLa (cervical cancer) (59) and A549 (lung cancer) (63) and was also adapted to discover additional 545 RBPs in *Plasmodium falciparum (64)*. Importantly, the R-DeeP technique enables to conserve the native interaction between RNA-protein complexes in the cells allowing the comparison between different cellular phases, cell types and cellular compartments.

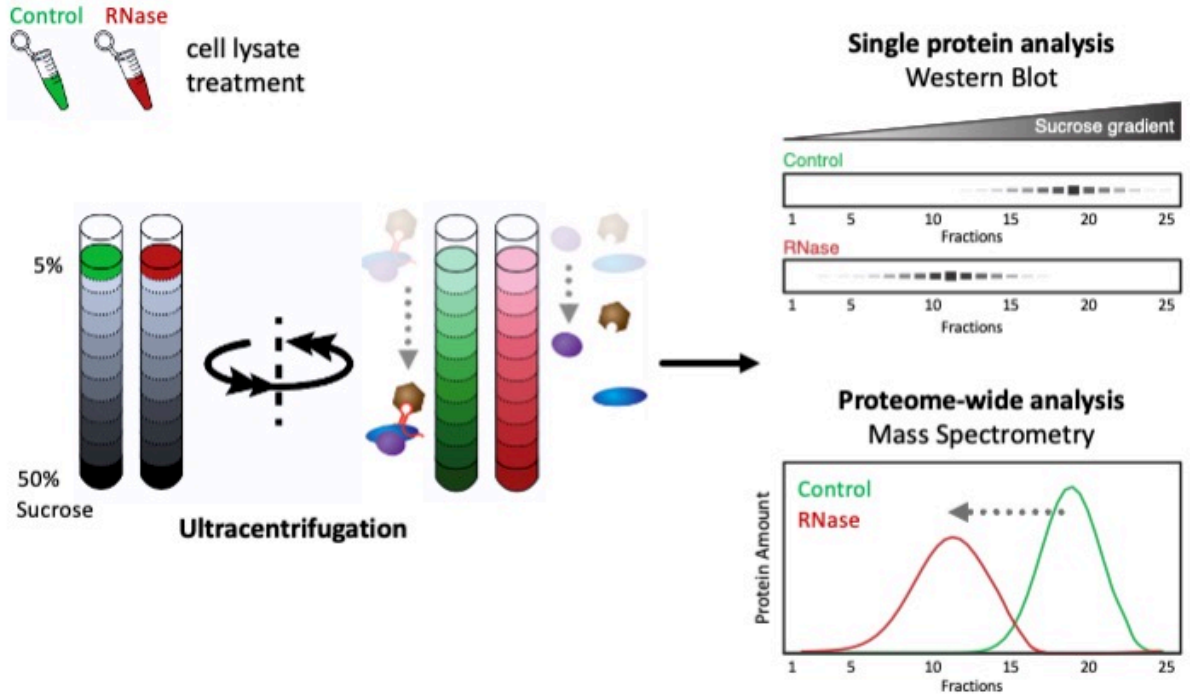


Figure 4: R-DeeP sucrose density gradient method

The cells were lysed, treated or not treated with RNase, loaded onto a sucrose density gradient containing 5-50% sucrose and finally ultracentrifuged. Further, the gradients were fractionated and analysed using mass spectrometry and western blot analysis. Figure adapted from (63).

Conveniently, the entire R-DeeP datasets on identified human RBPs are compiled and represented in R-DeeP 2.0 database available at https://R-DeeP2.dkfz.de (Figure 5). It contains various search options including single or advanced search that aids in visualizing the R-DeeP profile of a protein in one or both cell lines (HeLa or A549) simultaneously, with an option to download the results of the search (63). This helps in easy access and simple representation of the data. Also, no additional processing or analysis of the data by the user is required to explore the RNA dependence of a protein of interest.

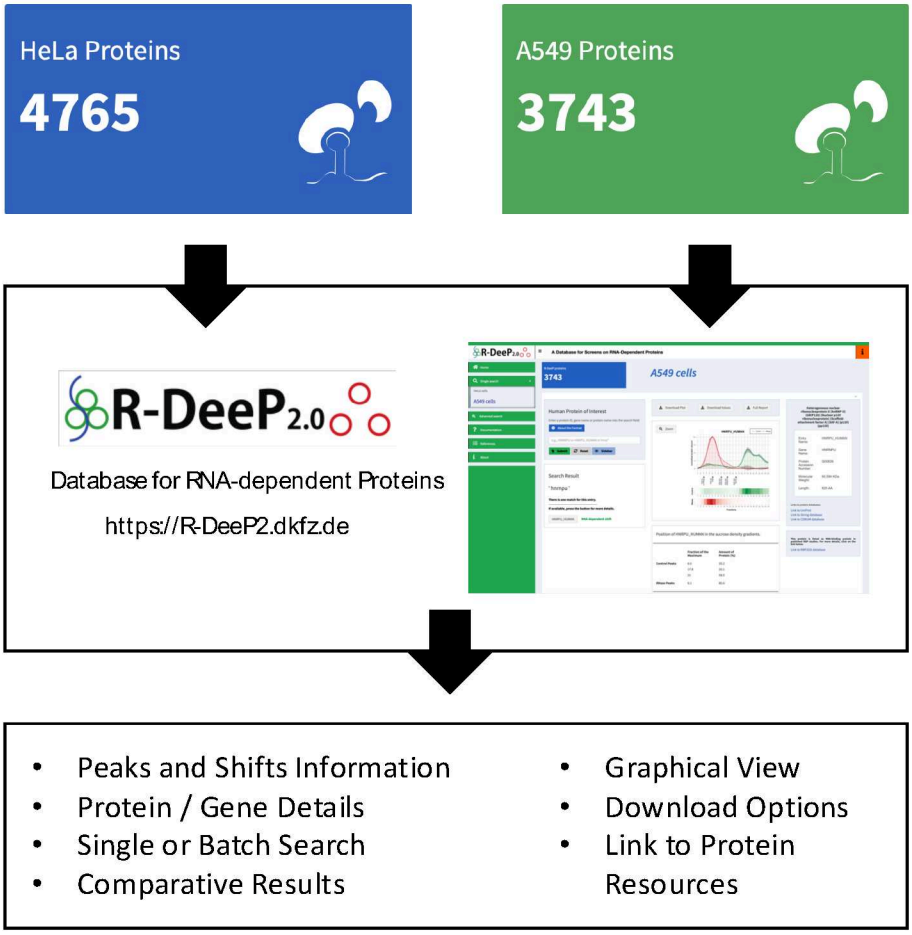


Figure 5: Data integration and representation in R-DeeP 2.0 database

The R-DeeP 2.0 database provides information on R-DeeP results on the proteins detected in HeLa S3 cells (4765) and the proteins detected in A549 cells (3743). The results are available at https://R-DeeP2.dkfz.de, in a downloadable format with multiple search options and link to external sources for detailed information on a protein. Figure adapted from (63).

Though it is possible to investigate the role and relevance of RBPs and RNA in various cellular processes with the currently available proteome-wide RBP screens, they do not reflect the RNA dependence of several important proteins in key cell cycle phases like mitosis. This is due to the differential protein expression levels throughout the cell cycle, low number of mitotic cells in non-synchronized cell populations, and the dynamics of post-translational modifications that may alter the affinity of proteins towards RNA. Therefore, by exploiting the adaptability of this method to different cellular phases, R-DeeP screen was performed in synchronised HeLa cells, resulting in the identification of several novel RBPs in a cell cycle specific manner. Notably, this approach uncovered the RNA dependence of several key mitotic factors, a previously unknow characteristic.

This thesis partly focuses on the RNA dependence of the mitotic factors identified in this new R-DeeP screen on HeLa cells synchronised in different cell cycle phases such as in interphase and mitosis. Hence, in the next sections, I will focus on introducing the cell cycle, more specifically mitosis that is central to my thesis.

1.8 The different phases of cell cycle

Cell cycle is a complex cellular process that is fundamental for all living organisms (65-67). It is a highly regulated event that demands high energy for multiple well-regulated molecular events including the synthesis of proteins that drive the replication of the parent cell genome, the precise partitioning of the cytoplasmic entities, the even distribution of the replicated chromosomes, and the generation of two individual daughter cells (65-67). Originally, the simplest form of cell cycle was assumed to be comprised of two stages; the M (mitotic) phase and interphase *i.e.* the interlude between two mitotic phases. Later, it became clear that interphase consists of 3 others stages called G1, S and G2 phase (Figure 6) (66,68,69). The G1 and G2 phases are the gaps in the cell cycle, which precede the two major stages, the synthesis (S) phase, where DNA replication takes place and M phase (69). During the G1 pre-replicative phase, the cell prepares for DNA replication. During this phase, before committing to the S phase, the cells can enter G0 phase–the resting or quiescent state which consists of non-proliferating, non-growing cells (65,66,69). One of the key molecular events is the origin licensing, during which the inactive helicases that are responsible for DNA replication such as MCM2 are loaded to the DNA. This step is essential for establishing conditions for efficient DNA replication that takes place in S phase (70). Upon entry into the S phase, the replication origins set in G1 phase are fired in a coordinated manner that initiates replication progression for timely, precise and complete chromosome replication as a result of which, the cells contain increased DNA content ranging from 2N to 4N (65,70). The MCM complexes are unloaded by replisome disassembly complex and the cells progress into G2 phase after complete replication of the genome at the end of S phase (70). During the second gap phase, the cells prepare for mitosis by synthesizing proteins that are essential for mitotic progression and determine the timing of mitotic entry (68,71). The G2/M transition is the rate limiting step that controls the progression of a cell through division, regulated by protein kinases (68). Finally, the cells enter mitosis, where the replicated genetic material is faithfully divided and segregated to the two daughter cells. The cell cycle restarts again from G1 phase where the decision is made to re-enter cell cycle or to enter G0 phase (65,66,68).

Multicellular organisms go through consecutive rounds of cell division to generate cells required for growth, development and homeostasis. Cyclin dependent kinase (CDK1) is the master regulator of the cell cycle (65,66,68,69,72). Accumulation of specific cyclins in different stages of cell cycle regulates transcription, protein stability or degradation and ensures sequential and unidirectional cell cycle progression (69,72). To prevent accumulation and propagation of genetic errors to daughter cells through cell division, cells depend on complex regulatory mechanisms and cell cycle checkpoints (69). These cell cycle checkpoints are evolutionarily conserved and monitor every stage of cell cycle such as during DNA damage in interphase, replication fork integrity in S phase and spindle assembly in M phase. These checkpoints are regulated by a huge network of proteins like p53, ATM, CHK2 in DNA damage control, ATR kinase and CHK1 in S phase to monitor DNA replication stress and multiprotein spindle assembly checkpoint (SAC) complex consisting of proteins like MAD, BUBR1 and CDC20 to ensure proper spindle assembly has occurred and equal partitioning of the replicated chromosomes to daughter cells (69,72). Any perturbations in these checkpoint mechanisms could lead to mis-regulated cell cycle progression, genomic instability leading to pathological diseases like cancer (69).

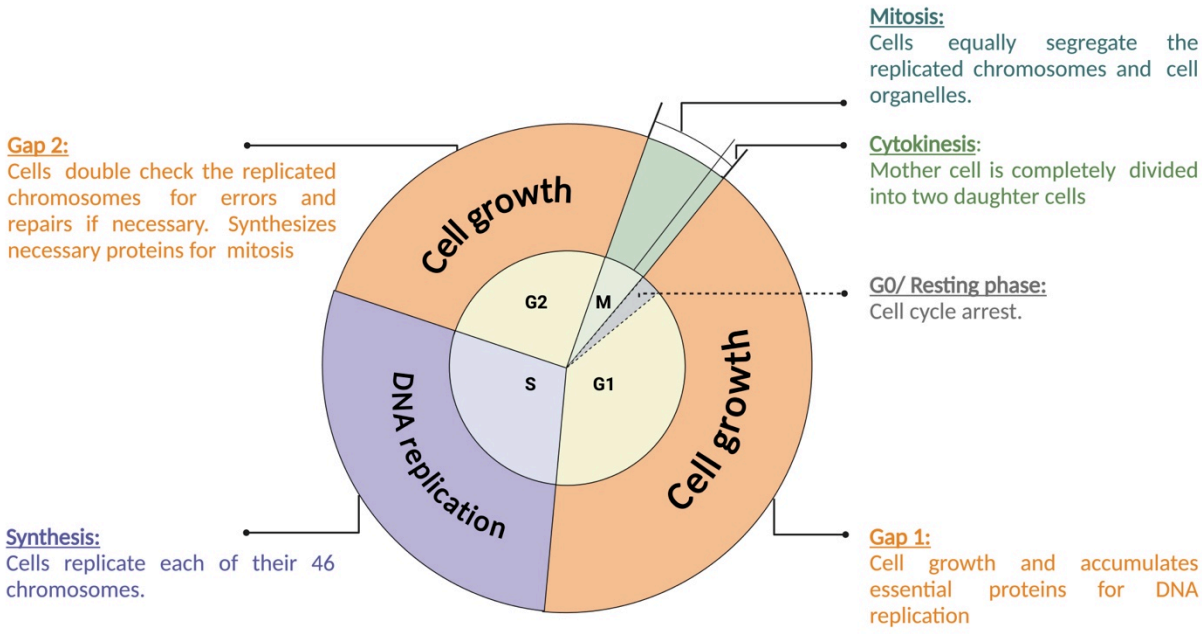


Figure 6: Cell cycle stages

Cell cycle consists of different phases such as gap 1 (G1), synthesis (S), gap 2 (G2), M-phase (M) and resting phase (G0). This figure was created using BioRender.com.

1.9 Cell division: a brief overview

Mitosis is a process by which a mother cell equally divides its nuclear and cytoplasmic entities into two daughter cells (73,74). Upon mitotic entry, the structure of the replicated chromosomes is reorganised by a series of molecular events, resulting in highly condensed, cylindrical chromosomes. Simultaneously, the cytoskeleton is rearranged to assemble the mitotic spindle around the chromosomes. Further, the condensed chromosomes are eventually segregated into daughter cells in the consecutive steps mediated through the coordinated actions of *e.g.* microtubules (MTs), mitotic motors and regulatory proteins including kinases (75). It induces enormous cytoskeletal changes impacting cell shape and is one of the largest and crucial processes performed by a dividing cell. The interphasic MTs are destabilized by inhibitory phosphorylation of the interphasic MT-associated protein MAP7 by cyclin B/CDK1 complex, resulting in increased levels of soluble α/β tubulin in the cytoplasm. The α/β tubulin heterodimers later support mitotic MTs nucleation from the centrosomes, that is driven by the γ-tubulin ring complexes (γ-TuRCs), which further plays a crucial role in mitosis along with mitotic factors such as Targeting protein for Xklp2 (TPX2) and Aurora kinase A (AURKA) (76). Interestingly, budding yeasts make one mistake in over 100,000 division whereas, accuracy of cell division is considerably less in mammalian cells (74). Any dysfunction or mis-regulation can lead to growth, ageing related and proliferative diseases (73,77). Cell division is subdivided into series of stages: prophase, prometaphase, metaphase, anaphase and telophase, which altogether constitute mitosis, followed by cytokinesis (Figure 7) (73,74,77).

Prophase is defined by the separation of centrosomes and condensation of the chromosomes, where they individually become visible (74). The chromosome condensation occurs through the wrapping of DNA fibres around the histone octamers mediated by the cohesins–a family of proteins, that aids in intensive scaffolding by encircling one or more strands of the double stranded DNA. The chromosomal condensation is also partly mediated by post-translational modifications of histone proteins, such as phosphorylation of histone H3 at serine 10 (74). As and when the chromosomes condense serves as a signalling mechanism for the nuclear envelope breakdown (NEB), thereby dissolving the barrier between nuclear and cytoplasmic entities (78). Further, the cell reorganizes its interphasic cytoskeleton which leads to a round symmetry of a cell, thereby boosting equal distribution of its cytoplasmic entities and equipartition of its organelles during cytokinesis (74). During prophase, transcription and

translation are globally repressed due to the highly condensed state of the chromosomes and slower transit of mRNA by the ribosomes (74). The condensation process transforms the chromosomes which become significantly shorter and aids in error-free and even segregation of the genetic material into daughter cells by the mitotic spindle and motors proteins that are required for chromosome movement and spindle orientations (75).

Prometaphase is considered as the longest phase of mitosis ranging from the end of prophase, *i.e.* NEB and onset of metaphase *i.e.* characterised by the chromosome alignment at the metaphasic plate (76). Thereby, kinetochores are the most important parts of the mitotic chromosomes, as they are essential for MT attachment. One kinetochore structure is present in each chromatid and is made up of a multiprotein complex consisting of Ncd80 complex, KNL1 and CLASP1 (74,79). The MTs, which are dynamic and instable in nature continuously shrink and grow and are stabilized upon binding to the chromosomes. MT end facing the chromosomes are termed as the plus end and the side facing far away from the chromosomes and towards the spindle poles is termed as the minus end. This dynamic instability of the MTs supports the search and capture model *i.e.* helps to probe the entire cytoplasm in various trajectories until MTs capture a chromosome at their kinetochores (74,79). These particular MTs, which are stabilized upon attachment to the kinetochores are termed as KT-fibres.

There are also other models suggesting chromatin-dependent MT assembly from the chromosomes which is mediated by the small GTPase Ran. Such a process enables acentrosomal cell divisions like in *Xenopus laevis* egg extract (76). The first evidence for MT organization from kinetochores was published in 1975, in *in vitro* purified chromosomes where the MT outgrowth was observed from the kinetochore (80). In the context of the RanGTP gradient emerging from the chromosomes, RanGTP binding to importins triggers the release of MT regulatory proteins, including stabilizing proteins like Augmin, TPX2 and CLASP1 that support spindle formation (76,81,82). Further with the help of plus and minus end motor proteins like Eg5, dynein and HSET/KIFC1, the MTs are clustered at the minus end to form the two spindle poles (74,83).

Though proper attachment of chromosomes to the MTs in essential for faithful cell division, it is prone to errors during early prometaphase mediated by inaccurate attachments (74,79). For example, a single kinetochore could be connected to both the poles, a condition termed as merolytic attachment. If uncorrected, this condition leads to the formation of lagging chromosomes and serves as a major cause for aneuploidy and chromosomal instability (74,79). Prometaphase is prolonged when the kinetochore-microtubule attachment is improper. This

delay is caused by the SAC complex which consists of a huge protein network including the kinase Aurora kinase B (AURKB) (74,79). A biochemical cascade is initiated upon unattached kinetochore and terminates when the last chromosome is properly attached. When all the kinetochores from the pairs are attached appropriately to MTs emerging from the two opposite poles, it creates a chromosomal bi-orientation leading to metaphase (74,79,84).

Metaphase is denoted by the bi-oriented and aligned chromosomes (74,84). It is characterized by a bipolar spindle structure, that consists of KT-fibres, interpolar MTs that form an anti-parallel array by overlapping in the centre and astral MTs that interact with the cell cortex (74,84). Although this portrays a good overview of metaphasic organization of the spindles and the chromosomes, in reality, most chromosomes do not stay at the equator, but oscillate along the spindle axis which varies between the chromosomes and the cell (74,84). This could be due to constant MT flux *i.e.* a constant net addition of α/β tubulin units to the plus end, where they are attached to the kinetochores and depolymerization at the minus end of the MTs *i.e.* at the poles, mediated by the kinesin-13 family proteins (74,84). Nevertheless, the average length of the MTs is roughly constant during metaphase. When all chromosomes are properly oriented and aligned, cellular checkpoints are fulfilled and silenced. This triggers the ubiquitination and phosphorylation of various regulatory proteins such as securin and cyclin B (74,84). Further, it leads to the proteasomal degradation and loss of securin, which in turn initiates the separation of sister chromatids thereby marking the onset of anaphase (74,84).

Anaphase is the phase where the sister chromatids migrate away from each other towards the opposite ends of the cell (74,85,86). During the transition to anaphase, the E3 ubiquitin ligase anaphase-promoting complex (APC/C) is activated, due to which the global protein phosphorylation state is reversed along with downregulation of CDK1 activity and increase in PP1/PP2A phosphatase activity (85). Not only these phosphatases are essential for metaphase-anaphase transition, they are also the major regulators driving mitotic exit, as they help in restoring the phosphorylation of the proteins to their interphasic level–hypo phosphorylation (87). The switch in protein phosphorylation state influences the protein localization pattern and induces profound changes in spindle behaviour (85). Anaphase is divided into two distinct phases: anaphase A and anaphase B (74,85,86). Anaphase A is defined by the movement of chromosomes towards the poles. Here, kinetochores tend to move in straight paths towards the spindle poles, followed by chromosomal arms that follows more complex movement patterns (86). The kinetochores mostly assume a unidirectional movement with some minor changes in

the directionality due to the dynamic instability of the MTs (88). The minus end motors like dynein and kinesin-14 create a poleward flux, which pulls the chromosomes towards the spindle poles, thereby, shortening the length of KT-fibres (74,86). Whereas anaphase B is denoted by the elongation of the interpolar MTs, that further separates the disjointed sister chromatids (74,85,86). Anaphase B usually starts with a 30-50s delay compared to anaphase A in human mitosis and the spindle elongates by ~8 μm (85,89). The spindle elongation is accompanied by plasma membrane elongation, creating space for the separated chromosomes. Many new MTs are formed whose plus end points towards the cell equator (74,85). These MTs recruit kinesin-6, which is a plus end directed motor protein, that conveys the central spindlin complex (MKLP1 and RACGAP1) to the MT tips and thereby to the cell cortex (74,85). This activates the Rho-GTPase which regulates the formation of a cleavage furrow that bundles the MTs to form the midbody. The midbody consists of proteins like MKLP1, PRC1, RACGAP1 and PLK1 (74). Finally, AURKB activity coordinates anaphase exit by ensuring proper separation of the sister chromatids before the reformation of nuclear envelope and chromatin decondensation, thus defining the checkpoint in late anaphase (85,90).

Telophase which follows anaphase is characterized by the onset of nuclear envelope reassembly and chromosome decondensation (74,85). Along with decrease in CDK1 activity, the activity of other mitotic kinases is altered (74,85). The chromatin arms contract which separate the chromosome, pulls them closer to the spindle poles and the nuclear membrane is reformed (74,91). Further, the nuclear pore complexes reassemble, forming a new nuclear envelop, reassuming the interphasic conformation, that signals the decondensation of the chromosomes (74).

Cytokinesis starts once the mitotic spindle segregates the separated chromosomes into individual nuclei, and the cleavage furrow is formed (74,92). The major mitotic kinases like CDK1 and PLK1 regulate the central spindlin complex (MKLP1 and RACGAP1) and chromosome passenger complex (CPC) (AURKB, INCENP, Borealin and Survivin), which specify the location for the formation of the contractile ring (92). Additionally, the central spindlin complex and CPC activates the downstream signalling proteins like Rho-GEF and ECT2 to assemble the contractile ring (92,93). The contractile ring consists of actin filaments and myosin-II that makes up the cleavage furrow. It is positioned at the equator of the cell as demonstrated by various studies over the past decades. Once the contractile ring is formed, the

actin, myosin and other associated proteins like arp2/3 complex constrict and generate an inwards contractile force, that drives furrow ingression and finally separation of the daughter cells, a step which completes cell division (74,84,92).

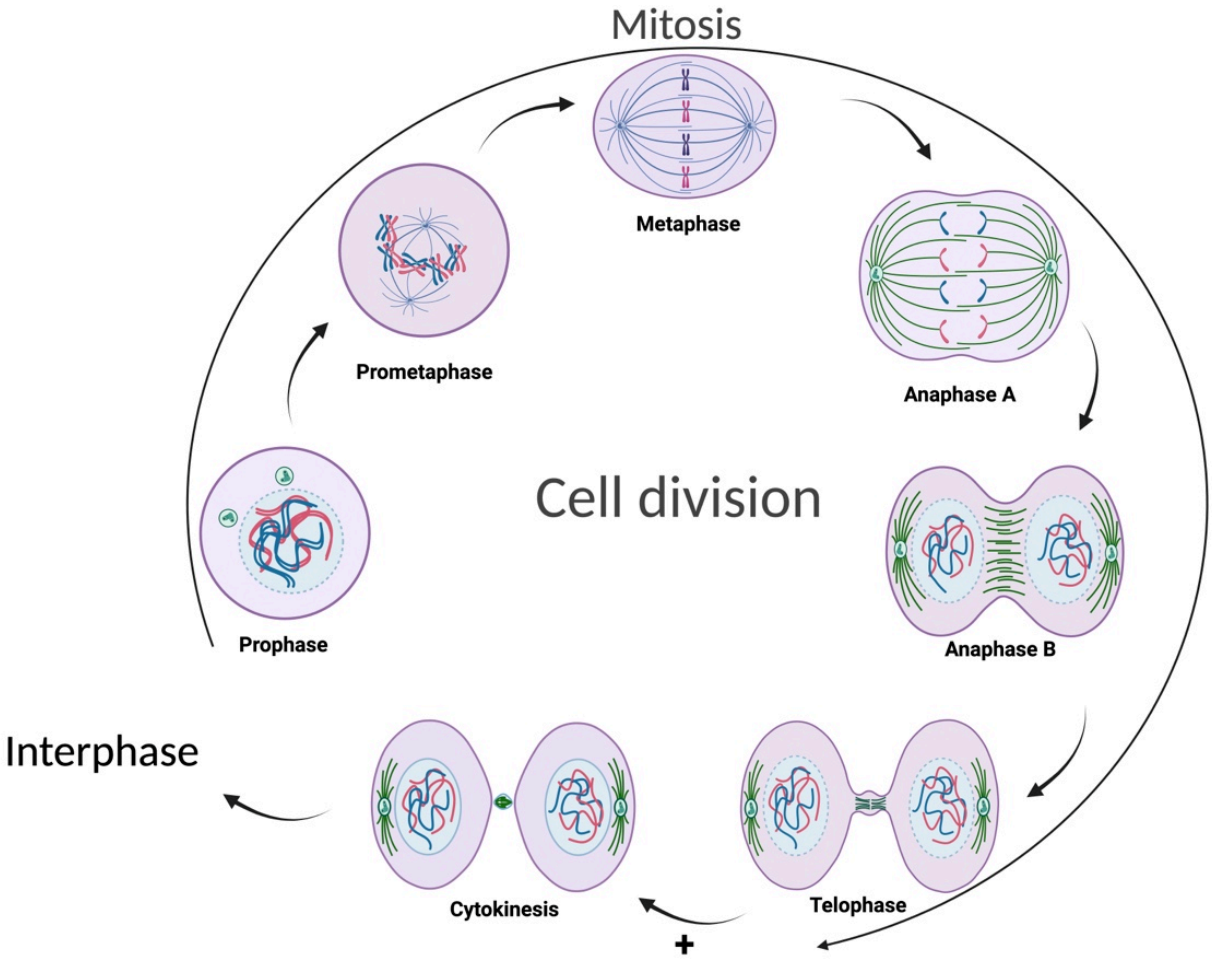


Figure 7: Cell division cycle

Mitosis is divided into prophase, prometaphase, metaphase, anaphase (A & B) and telophase. Finally, a mother cells is divided into two daughter cells during cytokinesis and these new cells further proceed into interphase, where the cell cycle starts again. This figure was created using BioRender.com.

The transition from mitosis to interphase is termed as mitotic exit. It is a critical process mediated by several regulatory proteins and monitored by the mitotic checkpoint (87,94). The regulatory functions of APC/C complex, cyclin B, kinases and phosphatases play critical role for proper mitotic exit. For instance, upon activation of the APC/C complex, it targets its substrates including cyclin B to the 26S proteasomal degradation pathway. This in turn downregulates CDK1 activity and increases PP1/PP2A phosphatase activity (87,94). This increase in phosphatase activity helps in the restoration of the global phosphorylation levels of the proteins down to the basal interphasic levels, which also ensures proper exit and no abrupt

reinitiation of mitosis (87,94). Any loss or mutation of these checkpoint proteins could lead to mitotic dysregulation, aberrant cell proliferation and genomic instability that ultimately leads to pathological diseases like cancer (74,87,94,95).

1.10 Cancer: a disease linked to cell cycle defects

Cancer is a pathological disease, where the cells entail defects in regulatory pathways that govern normal cell division leading to uncontrolled proliferation (96,97). Tumorigenesis is a multi-step process, that drives the transformation of normal cells into malignant cancer cells which contains genetic aberrations. These genetic mutations produce oncogenes with dominant gain of functions like KRAS and tumour suppressor genes like P53 with recessive loss of function (96,97). During this multi-step process, the genome of the cancer cells has altered invariably at multiples sites, through lesions as subtle as point mutations and as enormous as changes in the number of chromosomes (96,97).

In 2000, six hallmarks of cancer were proposed that were common to different cancer types, which were believed to collectively drive the malignant transformation of cells: (i) self-sufficiency in growth signals, (ii) resistance to growth inhibitory signals (iii) evading apoptosis, (iv) limitless replicative potential, (v) sustained angiogenesis and (vi) tissue invasion and metastasis (96,97). Briefly, the fundamental trait of a cancer cell is to survive beyond their life span and sustain uncontrolled proliferation. In order to grow exponentially, cancer cells circumvent growth inhibitory signals and evade programmed cell death (apoptosis) (96,98,99). Additionally, tumour cells acquire angiogenetic phenotype, that is essential for their growth i.e. new blood vessels are formed that supply oxygen and nutrients to support tumour growth and promote metastasis (96,98,99). Metastasis is another critical stage of tumorigenesis, that is divided into a number of steps consisting of detachment of tumour cells from their primary environment, invasion, migration, vascularization and colony formation in a secondary site. Importantly, metastasis is the primary reason for the high mortality rate and the majority of cancer related deaths (96,98,99). As our understanding of cancer improved, the hallmarks of cancer evolved from six to ten, with the addition of four new characteristics: deregulated energy metabolism, evading immune response, genomic instability and mutation, and tumour promoting inflammation (96,98). In 2022, these 10 hallmarks of cancer further expanded to 14, which includes emerging characteristics that are based on extensive studies from the past decades. These novel hallmarks include: unlocking phenotypic plasticity, senescent cells, polymorphic microbiome and non-mutational epigenetic reprogramming (Figure 8) (99). The

hallmarks are updated with time and gain of evidence over the years, adding more layers to the existing complexity (96,99).

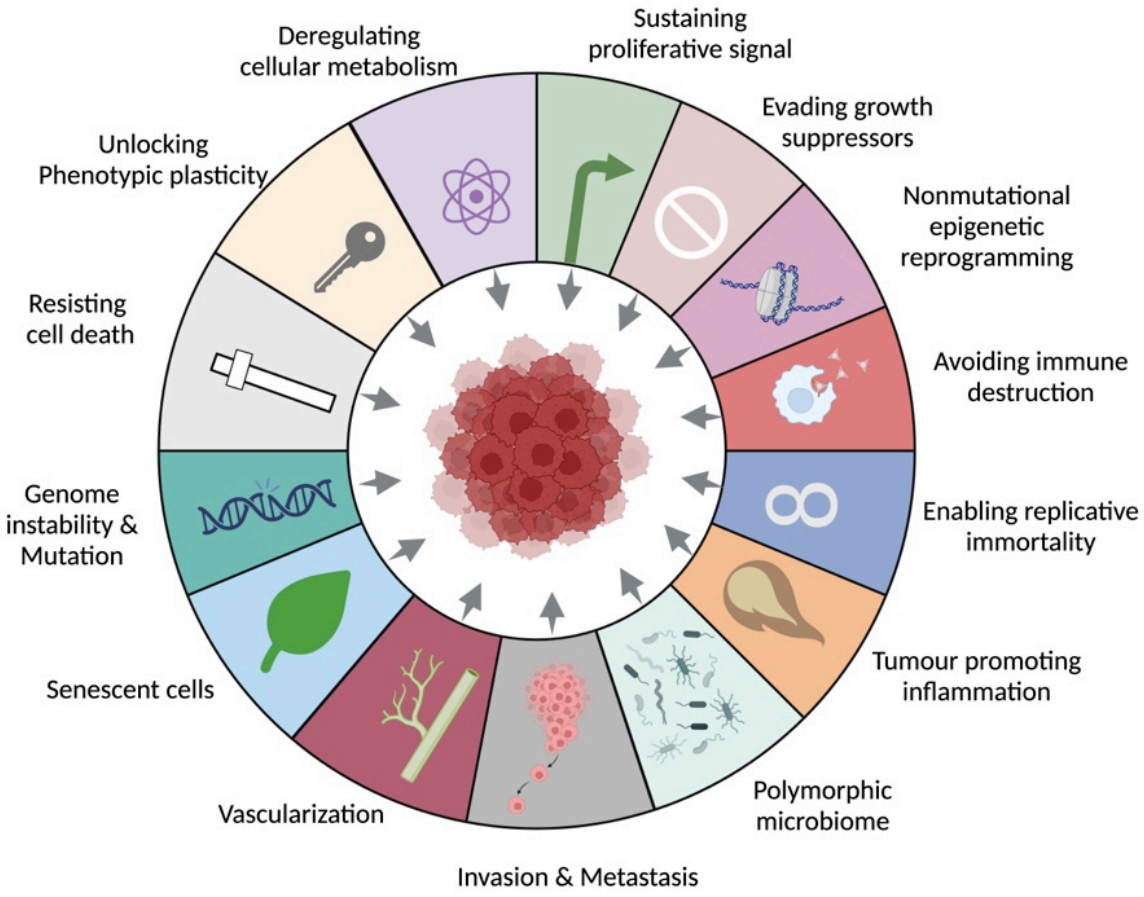


Figure 8: Hallmarks of cancer.
The 14 hallmarks of cancer based on the review article Hallmarks of cancer: New Dimensions (99). This figure was created using BioRender.com.

There are more than 100 cancer types and subtypes of tumours found within each organ (97). According to the World health organization (WHO) report, cancer is the leading cause of death accounting to roughly 10 million worldwide, with lung cancer being at the top with roughly 1.80 million deaths, followed by colon cancer (916,000 deaths) and liver cancer (830,000 deaths) in the year 2020 (100). Hence, the ultimate aim of current cancer research is to develop a comprehensive strategy for effective treatment of cancer and to improve the quality of life of the patients (101,102). Remarkable progress in understanding cancer as a disease has led to major advances in the treatment options and evolution of therapies over the past 170 years. Initially, in the mid 1800s, surgical resections were performed with the help of anaesthesia. The invention of X-rays (103) paved the way for combination treatment of cancer with radiation

and surgery, resulting in the breakthrough in chemotherapy during World War II (101,102). Further, in 1990, the use of BCG (vaccine for tuberculosis) was approved for the treatment of superficial bladder cancer, followed by rituximab–the first monoclonal antibody that targets CD20 antigen to treat B-lymphomas–marking the new era of targeted cancer therapy (101,104). Further, the approval of imatinib in 2001–a small molecule inhibitor which targets tyrosine kinases–revolutionized the treatment of chronic myeloid leukaemia and gastrointestinal tumours (101,102,105). Later in 2011, the development of engineered immune cells (CAR-T cells) to treat leukaemia marked a significant step towards cancer immunotherapy (101,106). Several other small molecule inhibitors, monoclonal antibodies, PROTACs and peptide-based drugs like gefitinib, bevacizumab and zoladex were developed targeting multiple pathways, regulatory proteins and kinases like VEGF, EGFR, AURKA and IDO1 (101,107,108). The trend in cancer treatment shifted from monotherapy to combination therapy to inhibit more proteins and pathways with re-use or re-purposing of commercially available drugs (encompassing a variety of immunomodulators, anti-angiogenic drugs and chemotherapies), to targeted therapies and finally to personalised therapy (101,102). With such a progress, the field of oncology is currently facing a tremendous proliferation of anti-cancer drugs in the market (101), yet with highest mortality rate and poor clinical outcome.

2 Aim of the thesis

Lung cancer is the leading cause of cancer related deaths worldwide, with a low 5-year survival rate of 26.7% (109). It is mostly diagnosed in advanced stage of the disease, resulting in limited treatment options (100,110). Though intensive research over the past decades have developed better treatment strategies for lung cancer, the treatments are often expensive and the success rate of such treatments is still low with the risk of tumour recurrence and therapy resistance (110). To device an effective therapeutic strategy, it is crucial to understand the molecular details of the disease mechanism which includes proteins and nucleic acids. The crucial role of RBPs and RNAs in several pathological diseases like neurological diseases and cancer has been emphasized in numerous studies (9,12,20,111-114). Still, many questions remain unanswered:

- Why does majority of the treatments focus on targeting a single protein and not the interacting RNAs, or the interactions between the RNAs and the RBPs?
- Are the treatment strategies heading in the right direction? Could it be that the well-known lungs cancer targets are RBPs?
- Could the lungs cancer targets be deregulated due to the changes in their interaction networks, e.g. due to interacting RNAs? If so, could these proteins also be relevant in other cancer types?

To answer these questions and to have an overview of RBPs in lung cancer, I focussed on identifying novel RNA-dependent proteins in lung cancer in the first part of my PhD thesis. Towards this aim, I exploited the R-DeeP methodology and identified 170 novel RNA-dependent proteins in A549 cells, which served as a novel lung cancer cell line model (63).

Loss of normal cell cycle control is one of the hallmarks of cancer (115). Various check point mechanisms have been established by the cell for proper cell cycle control. Yet, cancer cells accumulate several mutations that results in genomic instability and uncontrolled proliferation (115). Importantly, mitotic kinases, SAC components and other mitotic proteins responsible for faithful cell division such as Aurora kinases, BUBR1 and kinesins are deregulated in cancer and are one of the major cancer targets (115). Though several proteome-wide screens were developed, some studies report underrepresentation of mitotic factors due to differential protein expression (46). However, the R-DeeP screen in unsynchronized HeLa cells indicated an enrichment of mitotic factors within the detected RNA-dependent proteins (59,116).

- This raised the question whether these mitotic cancer targets could be RBPs, *i.e.*, directly binding to RNA.

- Does RNA play a crucial role in mitosis to regulate these proteins, irrespective of global transcriptional and translational repression?

Though mitotic pathways have been intensively studied for several years, the role of RNA and the involvement of RBPs in mitosis is seldom investigated (117,118). Hence, to address these questions and to obtain a complete landscape of RNA-dependent proteins and their role in mitosis, the R-DeeP screen was performed in HeLa cells synchronised in mitosis and interphase. This led to the identification of more than 700 novel RNA-dependent proteins in both cell cycle phases together. This analysis led to the discovery of the RNA dependence of one of the major mitotic kinases: Aurora kinase A (AURKA).

This leads to the second part of my PhD project, where I focussed on characterizing the RNA-dependent functions of AURKA in mitosis. In this second part, I characterized the RNA dependence of AURKA and identified new interaction partners of AURKA in mitosis that were also RNA dependent and more specifically directly interacting with RNA. Additionally, the RNA transcripts mediating the interaction of these proteins were uncovered. Finally, I identified the functional relevance of these RNA-protein interactions and RNA-mediated protein-protein interactions in mitosis, leading to new cues and perspectives in the regulation of mitotic spindle assembly and cell division.

3 Materials and methods

Most of the materials and methods section is adapted from the two existing manuscripts (63,116).

3.1 Gene ontology analysis

The shifting proteins (1751 proteins in total), identified in the R-DeeP screen performed in unsynchronized HeLa cells (59) were selected for gene ontology (GO) enrichment analysis. The analysis was performed using the GO enrichment analysis tool from the Gene Ontology Resource (119). Fold enrichment analysis was performed between the list of proteins provided as an input and the whole human proteome. Further, the GO analysis provided information on the number of proteins enriched in different biological process and cellular compartments from the given list. The statistical significance for the enrichment of proteins was performed using Fishers exact test with correction for multiple testing based on false discovery rate (FDR). Adjusted p-values < 0.05 were considered significant.

3.2 Cell culture

Hela wild type (WT) (ATCC, cat no. CCL-2) cervical cancer cells were grown in DMEM high glucose medium (Sigma Aldrich, cat no. D65796) supplemented with 10% FBS and A549 WT (ATCC, cat no. CCL-185) lung cancer cells were grown in RPMI 1640 medium (Gibco, Thermo Fisher Scientific, cat no. 11875093) supplemented with 10% FBS. Both the cell lines were incubated in a humidified incubator at 37 °C with 5% CO_2. The cell lines were regularly authenticated and tested for the absence of mycoplasma contamination.

3.3 Cell synchronization

2.5-3 million cells were seeded for synchronization in prometaphase and 2 million cells were seeded for synchronization at interphase in 15 cm dishes (Techno plastic products, cat no. 93150) on the first day. On the second day, 2mM (final concentration) thymidine (Sigma Aldrich, cat no. T1895-1G) was added to the cells and incubated for 16 h to arrest the cells in S-phase. The cells were washed once with warm PBS (Sigma Aldrich, cat no. D8537) and released for 9 h in fresh media. On day 3, 2mM (final concentration) thymidine was added again to the cells (incubation for 16 h) to arrest the cells in S-phase (double thymidine block). To further synchronize the cells in prometaphase, the cells were washed once with PBS and

fresh media was added, that contained 100 ng/ml nocodazole (Sigma Aldrich, cat no. M1404-2MG). Finally, after 12 h, the cells were harvested, flash frozen and stored at -80°C until lysate preparation.

To synchronize cells at metaphase, 5 million cells were seeded on a 15 cm plate on the first day. On the following day, 2mM (final concentration) thymidine was added to the cells and incubated for 24 h to arrest the cells in S-phase. On the third day, cells were washed once with warm PBS and were incubated with fresh medium containing 40 ng/ml nocodazole for 12 h to arrest the cells in prometaphase. After the incubation, the cells were washed three times with warm PBS and released with fresh media for 40 min to capture them in metaphase. Finally, the cells were harvested, flash frozen and stored at -80°C until lysate preparation.

3.4 R-DeeP screen

Sucrose density gradients: the unsynchronized A549 cells and HeLa cells synchronized in prometaphase and interphase as stated above were used. The gradients, cell lysate preparation, RNase treatment, ultracentrifugation and fractionation were performed as previously published (59,60). Briefly, the cell lysates were either untreated (control gradients) or treated with an RNase cocktail (RNase-treated gradients), loaded on the sucrose gradients and subjected to ultracentrifugation. The gradients were fractionated into 25 different fractions, and further analysed either using mass spectrometry or western blot analysis.
Bioinformatic analysis: mass spectrometry datasets were analysed based on the pipeline as described in the published protocol (60).

3.5 SDS-PAGE and western blot analysis

Proteins were separated based on their molecular weight using sodium dodecyl sulphate–polyacrylamide gel electrophoresis (SDS-PAGE). The samples in 1× SDS sample buffer (30% (v/v) glycerol, 12% (w/v) SDS, 3.6 M DTT, 0.012% (w/v) bromophenol blue, and 500 mM Tris-HCl (pH 6.8)) or 1x Lithium dodecyl sulfate (LDS) buffer (Thermo fisher scientific, cat no. NP0007) were boiled at 95 °C for 5 min or 70°C for 10 min respectively, and briefly spun down to collect the samples at the bottom of the tube. Next, the samples were loaded onto appropriate Bio-Rad pre-cast protein gels (for gradients: 4–20% Criterion™ TGX Stain-Free™ Protein Gel, 26 well, 15 µl, BioRad, Cat no. 5678095. For other samples: 7.5% precast Mini-Protean-TGX gel, BioRad, cat no. 456-1024) and run at 120 V in the electrophoresis chamber containing 1x SDS Running buffer (25 mM Tris base, 192 mM glycine, 0.1% (w/v) SDS). Western blot analysis was performed on nitrocellulose membrane (Amersham™ Protran®

western blotting membranes, nitrocellulose, Sigma Aldrich, cat no. 10600002) in a Trans blot turbo wet transfer system using 1x Trans-Blot Turbo Transfer Buffer (BioRad, cat no. 10026938) containing 20% ethanol (mixed molecular weight program). The membrane was blocked for 1 h at room temperature (RT) with 5% milk in Tris-buffered saline (blocking solution: 24.7 mM Tris-HCl (pH 7.4), 137 mM NaCl, 2.7 mM KCl) containing 0.05% Tween-20 (TBST). Furthermore, the membrane was incubated at 4°C overnight with the respective antibodies (AURKA: (D3E4Q) rabbit mAb, 1:1000 dilution, Cell Signaling, cat no. 14475, TPX2: 1:100 dilution, Biolegend, cat no. 628002/ Sigma Aldrich, cat no. SAB4701065, KIFC1: 1:1000 dilution, Abcam, cat no. 172620). The following day, the membrane was washed three times with TBST for 5 min with at room temperature (RT) and incubated with the appropriate HRP-conjugated secondary antibody (Goat anti-mouse, Dianova, cat no. 115-035-003, Goat anti-rabbit, Dianova, cat no. 111-035-144) at 1:5000 dilution in blocking solution for 1 h at RT. The membrane was washed three times with TBST for 5 min at RT. Finally, the membrane was incubated with ECL reagent (ECL prime western blotting system, Cytiva, cat no. RPN2232) for 5 min and the blots were imaged using an INTAS ECL Chemocam imager (Imager ECL Chemo Cam CC5569). Quantitative analysis of western blot images was performed using Fiji (Image J) software.

3.6 AURKA immunoprecipitation (IP) followed by LC-MS/MS-based protein analysis

AURKA immunoprecipitation: Cells synchronized in prometaphase were used for immunoprecipitation. The cell pellets were lysed in three pellet volume of lysis buffer (50 mM HEPES-KOH, pH-7.5, 150 mM KCl, 0.5% NP-40, 2 mM EDTA, 1 mM NaF, 0.5 mM DTT and 1x complete EDTA free protease inhibitor cocktail (Sigma Aldrich, cat no. 4693132)), incubated on ice for 30 min and centrifuged at 17,000 g for 20 min at 4°C. The supernatant was transferred into fresh tubes followed by second centrifugation step at 17,000 g for 20 min at 4°C. The supernatant was stored in a fresh tube on ice until the beads were prepared for the pre-clearing step (see below).

30 µl of pierce ChIP-grade protein A/G magnetic beads (Thermo Fisher Scientific, cat no. 26162) were used per sample in the IP. The beads were aliquoted to a fresh 1.5 ml tube and washed three times with 1 ml lysis buffer. Unspecific interactions were removed by incubating the lysate and beads and for 1 h at 4 °C on a rotator. After 1 h, the beads were removed from the lysate using a magnetic stand and the pre-cleared lysate was transferred into fresh 1.5 ml tubes. BCA assay (BSA: Sigma Aldrich, cat no. A1470, bicinchoninic acid solution: Sigma

Aldrich, cat no. B9643-1L, Pierce™ BCA Protein Assay Reagent B, Thermo Fisher Scientific, cat no. 23224) was performed to measure the protein concentration.

The lysate was split into 2 different samples (containing 4 mg total lysate each) and incubated overnight for protein-antibody complex formation at 4°C on a rotator. Here, 0.8 µg AURKA antibody was used for AURKA IP and rabbit IgG was used as a negative control. Here, rabbit IgG was used as the AURKA antibody was raised from rabbit.

On the next day, beads were prepared by washing three times in 1 ml lysis buffer. The washed beads were split into 2 tubes, each lysate-antibody mix was added to the beads and was incubated for 2 h at 4°C on a rotator to capture the complex. After incubation, the beads were removed from the lysate using a magnetic stand and the flow through (FT) was discarded. The beads-antibody-protein complexes were washed three times with 1 ml wash buffer I (50 mM HEPES-KOH, pH-7.5, 150 mM KCl, 0.5% NP-40, 0.5 mM DTT and 1x complete EDTA free protease inhibitor cocktail). During the last wash step, each tube was split into two tubes for control and RNase treatment (total 4 tubes). Using a magnetic stand, the supernatant was discarded and the beads were resuspended with 100 µl wash buffer I.

10 µl wash buffer I was added to the sample in the control tubes. For the RNase-treated samples, 10 µl RNase cocktail was added (RNase cocktail: equal volume of RNase A, RNase I, RNase III, RNase H and RNase T1) (RNase A: Sigma Aldrich, cat no. 4875, RNase I: Thermo Fisher Scientific, cat no. 1056893, RNase III: Thermo Fisher Scientific, cat no. AM2290, RNase H: New England Biolabs, cat no. M0297S, RNase T1: Thermo Fisher Scientific, cat no. 10621621) and incubated for 1 h at 4°C on a rotator. After incubation, FTs were collected for LC-MS/MS-based protein analysis by capturing the beads using a magnetic stand. The beads were washed three times with wash buffer II (50 mM HEPES-KOH, pH-7.5, 300 mM KCl, 0.5% NP-40, 0.5 mM DTT and 1x complete EDTA free protease inhibitor cocktail). Finally, 30 µl 1x LDS containing 100 mM DTT was used for eluting the protein complexes and the samples were analysed using SDS-PAGE/western blot and LC-MS/MS-based protein analysis at Proteomics Core Facility (mass spectrometry-based protein analysis unit).

3.7 Protein digestion of AURKA IP samples for LC-MS/MS analysis

IP eluates were run for 0.5 cm into an SDS-PAGE and the entire gel piece was cut out and digested using trypsin according to Shevchenko et al. (120) adapted on a DigestPro MSi robotic system (INTAVIS Bioanalytical Instruments AG).

3.8 LC-MS/MS analysis of AURKA IP

The LC-MS/MS analysis was carried out on an Ultimate 3000 UPLC system (Thermo Fisher Scientific) directly connected to an Orbitrap Exploris 480 mass spectrometer for a total of 120 min. Peptides were desalted on a trapping cartridge (Acclaim PepMap300 C18, 5 µm, 300 Å wide pore; Thermo Fisher Scientific) for 3 min using 30 µl/min flow of 0.05% TFA in water. The analytical multistep gradient (300 nl/min) was performed using a nanoEase MZ Peptide analytical column (300 Å, 1.7 µm, 75 µm x 200 mm, Waters) using solvent A (0.1% formic acid in water) and solvent B (0.1% formic acid in acetonitrile). For 102 min the concentration of B was linearly ramped from 4% to 30%, followed by a quick ramp to 78%, after two minutes the concentration of B was lowered to 2% and a 10 min equilibration step appended. Eluting peptides were analysed in the mass spectrometer using data dependent acquisition (DDA) mode. A full scan at 120k resolution (380-1400 m/z, 300% AGC target, 45 ms maxIT) was followed by up to 2 seconds of MS/MS scans. Peptide features were isolated with a window of 1.4 m/z, fragmented using 26% NCE. Fragment spectra were recorded at 15k resolution (100% AGC target, 54 ms maxIT). Unassigned and singly charged eluting features were excluded from fragmentation and dynamic exclusion was set to 30 s. Each sample was followed by a wash run (40 min) to minimize carry-over between samples. Instrument performance throughout the course of the measurement was monitored by regular (approx. one per 48 h) injections of a standard sample and an in-house shiny application.

3.9 Data analysis of LC-MS/MS data

Data analysis was carried out by MaxQuant (121) using an organism specific database extracted from Uniprot.org (human reference database, containing 74,811 unique entries from 27th February 2020). Settings were left at default with the following adaptions. Match between runs (MBR) was enabled to transfer peptide identifications across RAW files based on accurate retention time and m/z. Fractions were set in a way that MBR was only performed within replicates. Label free quantification (LFQ) was enabled with default settings. The iBAQ-value (122) generation was enabled. Peptides from AURKA interactors were analysed using the log2(iBAQ) values from four replicates. Replicate four was excluded from the analysis due to unproper clustering pattern as compared to the other three replicates. Three conditions were analysed (IgG IP, AURKA IP, AURKA IP RNase treatment in three replicates, 9 samples in total). Interactors were filtered and only further analysed if they were detected in at least 70% of the samples. Missing values were imputed using random values based on a gaussian

distribution centered around the median of the sample and outliers were adjusted based on the mean method. Ratios between the (AURKA IP)/(IgG) samples and (AURKA IP)/(AURKA IP RNase treatment), *i.e.* differences of the log2(iBAQ) values were calculated for each replicate and adjusted p-values were computed by applying a t-test, corrected for multiple testing (FDR method). AURKA interactors were selected based on an at least two-fold increased (AURKA IP)/(IgG) ratio (adjusted p-values < 0.05). RNase sensitive AURKA interactors were identified based on an at least two-fold increased (AURKA IP)/(AURKA IP RNase treatment) ratio (adjusted p-values < 0.05).

3.10 AURKA immunoprecipitation (IP)

HeLa and A549 cells synchronized in prometaphase or metaphase were used for immunoprecipitation followed by western blot analysis. Cell pellets were lysed in 2 ml lysis buffer (50 mM HEPES-KOH, pH-7.5, 150 mM KCl, 0.5% NP-40, 2 mM EDTA, 1 mM NaF, 0.5 mM DTT and 1x complete EDTA free protease inhibitor cocktail) as explained in the previous section. Protein concentration in the lysate was measured using BCA assay and the lysate was diluted to 2 mg/ml with lysis buffer. The lysates were then split into 4 tubes containing 2 mg each.

2 µl of turbo DNase (Thermo Fisher Scientific, cat no. AM2238) were added to each tube. 10 µl of lysis buffer was added to control samples while, 10 µl RNase I (Thermo Fisher Scientific, cat no. AM2295) was added to RNase-treatment samples and were incubated at 37°C for 3 min at 1100 rpm in a thermomixer and cooled down by incubating on ice for 3 min. Later, the samples were centrifuged at 17,000 g for 20 min at 4°C. The supernatants were transferred to fresh tubes and were filtered through a proteus clarification mini spin column (Serva, cat no. 42225.01) by centrifuging at 16,000 g for 2 min at 4°C. The filtered lysates were transferred to fresh 2 ml tubes and kept on ice until the beads were ready for the pre-clearing step (see previous section).

The lysates were incubated with the respective antibodies overnight at 4°C on a rotator to form protein-antibody complexes. For AURKA IP: 0.4 µg (AURKA antibody (D3E4Q) rabbit mAb, 1:1000 dilution, Cell Signaling, cat no. 14475) per IP and rabbit IgG (Normal rabbit IgG, Millipore, cat no. 12-370) was used as a negative control.

On the following day, the beads were prepared by washing three times in 1 ml lysis buffer. The washed beads were added to the lysate-antibody mix and incubated for 2 h at 4°C on a rotator to capture the protein-antibody complex. After incubation, FT was discarded and the beads-

antibody-protein complexes were washed three times with 1 ml wash buffer (HeLa: 50 mM HEPES-KOH, pH-7.5, 150 mM KCl, 0.5% NP-40, 0.5 mM DTT and 1x complete EDTA free protease inhibitor cocktail, A549: 50 mM HEPES-KOH, pH-7.5, 15 mM KCl, 0.5% NP-40, 0.5 mM DTT and 1x complete EDTA free protease inhibitor cocktail). After the last wash, the beads were resuspended in 20 µl lysis buffer, 2 µl RNase I and incubated at 37°C for 3 min at 1100 rpm in a thermomixer. Finally, the samples were eluted using 7.5 µl of 4x LDS (with 200 mM DTT) and boiling at 70°C for 10 min. The samples were stored in fresh tubes at -20°C until SDS-PAGE/western blot analysis.

3.11 Proximity ligation assay (PLA)

All the buffers and solutions were provided in the Duolink® in-situ PLA kit.
PLA on cell cycle phases: 120,000 cells/well were seeded on a coverslip (Microscope cover glasses, 12mm, Nr. 1.5, Neolab, cat no. 0112520) in a 12-well plate (Techno Plastic Products, cat no. 92412) and were allowed to grow overnight. On the next day, cell medium was discarded, cells were washed once with warm PBS and fixed with either methanol or 4% PFA (depending on the antibody) for 10 min or 15 min respectively at RT. After fixation, the cells were permeabilized in 0.25% Triton in PBS for 10 min at RT. Further, cells were washed two times with warm PBS and blocked with 40 µl Duolink® blocking solution for 1 h at 37°C in a heated humidity chamber. In the meantime, primary antibodies were diluted to appropriate concentrations in the Duolink® antibody diluent. After blocking, the cells were incubated with primary antibody overnight at 4°C in a humidity chamber. All subsequent steps were performed following the instructions as provided with the Duolink® in-situ PLA kit (Duolink® In Situ PLA® Probe Anti-Mouse MINUS, Affinity purified Donkey anti-Mouse IgG (H+L), cat no. DUO92004, Duolink® In Situ PLA® Probe Anti-Rabbit PLUS, Affinity purified Donkey anti-Rabbit IgG (H+L), cat no. DUO92002, Duolink® In Situ Wash Buffers, Fluorescence, cat no. DUO82049, Duolink® In Situ Detection Reagents Red, cat no. DUO92008, Duolink In Situ Mounting medium, cat no. DUO82040, Sigma Aldrich). Mouse or rabbit secondary antibody Alexa Fluor 488 (Alexa Fluor 488 goat anti-rabbit/ anti-mouse IgG, Thermo Fisher Scientific, cat no. A11034/ A32723) was added to the amplification mix in a 1:500 dilution.

3.12 PLA +/- RNase

200,000 cells/well were seeded on a coverslip in a 12-well plate and the cells were synchronized in metaphase. Following the synchronization, the media was aspirated, and the

cells were first treated for 30 s at RT with 0.1% Triton in PBS with RNase and without RNase for control slides. Next, the cells were fixed with methanol at RT for 10 min, washed once with warm PBS and permeabilized with 0.25% Triton for 10 min at RT. Further, the cells were washed two times with warm PBS and the rest of the PLA protocol from blocking step was followed as described above.

Imaging: Images were acquired on a Zeiss LSM 980 Airyscan NIR, in confocal acquisition mode (best signal setting) or on a Zeiss LSM 710 ConfoCor 3. Both the microscopes are equipped with diodes for the excitation of DAPI (405 nm), Alexa 488 (488 nm) and Alexa 594 (561 nm/555nm) fluorophores. Samples within one replicate were all acquired with the same settings. Z-stacks were acquired with 0.5 µm or 1 µm interval between each slice and maximal projection images (512 x 512 pixels, 8 bits) were analysed using Fiji (ImageJ) software (123). Background pixel values up to 10 were removed for intensity calculation.

3.13 Individual-nucleotide resolution UV cross-linking and immunoprecipitation (iCLIP2)

Frozen pellets from HeLa and A549 cells synchronized in prometaphase (for AURKA and KIFC1) were used for the iCLIP2 assay.

Beads-antibody preparation: the antibodies were conjugated to beads first. Dynabeads Protein A (100 µl per IP) were washed three times with 1 ml lysis buffer. After the last wash the beads were re-suspended in 500 µl lysis buffer and split into two tubes, one for the IgG control (100 µl) and one for the protein of interest AURKA/KIFC1/DOCK5/ABRAXAS1 (400 µl), and further incubated with the antibodies (IgG/DOCK5 12 µg/IP) or for ABRAXAS1 iCLIP2 (IgG/ABRAXAS1: 1 µg/IP) or (IgG/AURKA/KIFC1: 2 µg per IP) for 1 h at RT on a rotator (10 rpm). Normal Rabbit IgG, (Millipore, cat no. 12-370), DOCK5: Anti-DOCK5 (Biomol, cat no. A3049887), ABRAXAS1/CCDC98: Anti-CCDC98 (Abcam, cat no. EPR6310(2)). The bead–antibody complexes were captured on a magnetic rack and washed once with 1 ml high-salt wash buffer (50 mM Tris-HCl pH 7.4, 1.5 M NaCl, 1 mM EDTA pH 8.0, 1% Igepal CA-630, 0.1% SDS, 0.5% sodium deoxycholate) and twice with 1 ml lysis buffer. The beads were resuspended in 100 µl lysis buffer for IgG or 400 µl lysis buffer for pulldown.

Cell lysis: UV cross-linked (254 nm, 200 mJ/cm2) and non-crosslinked cells were lysed in 2 ml lysis buffer per cell pellet (50 mM Tris-HCl pH 7.4, 100 mM NaCl, 1% Igepal (CA-630), 0.1% SDS, 0.5% Sodium deoxycholate, 1x protease inhibitor cocktail). Protein concentration was measure using BCA assay, diluted to 2 mg/ml and distributed into different 1.5 ml low-

bind tubes containing 1 ml total lysate. Then, the lysates were treated with 4 µl turbo DNase and different RNase I dilutions ranging from 1:5 to 1:1000 for AURKA and 1:5 to 1:500 for DOCK5, ABRAXAS1 and KIFC1. The RNase and DNase treatments were performed at 37°C at 1100 rpm on a thermomixer for 3 min, immediately incubated on ice for additional 3 min and centrifuged at 17,000 g for 20 min at 4°C. The supernatant was collected into a fresh 2 ml low-bind tube and filtered through proteus clarification mini spin column by centrifuging at 16,000 g for 2 min at 4°C. Further, the filtered lysates were transferred to a fresh 2 ml low-bind tubes and kept on ice.

Pulldown/Immunoprecipitation: Immunoprecipitation was performed by adding 100 µl of the resuspended beads to the respective tubes containing cleared lysate and incubated for 2 h rotating at 4°C. After 2 h, the complex was captured on a magnetic rack, FT was removed and the beads were washed twice with 1 ml high-salt wash buffer with rotation at 10 rpm at 4°C for 1 min and then washed twice with 1 ml PNK wash buffer (KIFC1/DOCK5/ABRAXAS1: 20 mM Tris-HCl pH 7.4, 10 mM MgCl2, 0.2% Tween-20, AURKA: 20 mM Tris-HCl pH 7.4, 10 mM MgCl2, 0.2% Tween-20, PhosStop). During the last wash, the beads were transferred to fresh 1.5 ml low-bind tubes and stored at 4°C.

On the following day, for AURKA iCLIP2: the samples were placed on magnetic stand, beads in PNK buffer containing PhosStop was removed and resuspended in 1 ml PNK buffer (20 mM Tris-HCl pH 7.4, 10 mM MgCl2, 0.2% Tween-20) without PhosStop.

Further, 100 µl of beads were used for western blot and the remaining 900 µl were used for labelling the RNA. For western blot, the supernatant was removed, and the protein complexes bound to the beads were eluted using 1× LDS buffer containing 50 mM DTT at 70°C for 10 min. The eluate was collected and stored at −20°C to check for immunoprecipitation efficiency using western blot analysis.

RNA labelling with radioactive 32P: RNA labelling was performed using the remaining 900 µl sample. The radioactive labelling of RNA using 32P was performed using a master mix containing 11.85 µl nuclease free water, 0.75 µl T4 PNK enzyme (NEB, cat no. M0201), 1.5 µl 10× PNK buffer, and 0.9 µl 32P-γ-ATP ([gamma-P32] Adenosine 5'-triphosphate (ATP) /9,25MBq, Hartmann Analytic, SRP-501) per sample. The supernatant was removed and the beads were resuspended in 15 µl PNK mix. The samples were incubated on a thermomixer at 37°C for 5 min at 1100 rpm for labelling the RNA. Later, to remove excess radioactivity, the samples were washed twice with 1 ml PNK wash buffer and eluted in 25 µl 1× LDS buffer containing 50 mM DTT on a thermomixer at 70°C for 10 min at 1100 rpm.

To visualize the RNA-labelling, SDS-PAGE and western blot analysis were performed. The samples were loaded in a 7.5% Mini-PROTEAN® TGX™ precast protein gel (Bio-Rad, cat no. 456-1024) and run 120 V in a vertical electrophoresis chamber filled with 1× SDS running buffer (25 mM Tris base, 192 mM glycine, 0.1% (w/v) SDS). western blot was performed using 0.45 µm nitrocellulose membrane and the proteins were transferred to the membrane with a wet transfer system with transfer buffer (25 mM Tris base, 192 mM glycine) containing 20% methanol for 1.5 h at 120 V in an ice bath. Finally, the membrane was washed once in nuclease-free water, covered with plastic wrap and exposed to a phosphor imager screen. After exposing the screen to the membrane for an appropriate amount of time, the screen was imaged using a Typhoon laser scanner phosphor imager at 200 µm, high speed and intensity 3.

3.14 iCLIP2 library preparation and sequence analysis

The iCLIP2 library preparation was performed based on the publication "Improved library preparation with the new iCLIP2 protocol" (124). For adapter, barcodes or primer sequences refer to Table 2.

For KIFC1 iCLIP2 library preparation, UV-crosslinked prometaphase cells synchronized and harvested on four different dates were used. Rabbit IgG was used as a negative control. Here, 1:10 dilution RNase treatment was performed on the lysates. All the steps from lysate preparation, beads preparation, and pulldown were performed as described in the above section: Individual Nucleotide Resolution and UV Cross-linked Immunoprecipitation (iCLIP2).

Dephosphorylation: The master mix containing 1x PNK buffer pH-6.5 (350 mM Tris-HCl, pH 6.5, 50 mM MgCl2, 5 mM DTT, 0.5 µl SUPERase-In, 0.5 µl of T4 PNK enzyme) in a total volume of 15 µl per sample was used to perform 3' dephosphorylation of KIFC1 bound RNA. Next, the PNK buffer was removed, the beads were resuspended in 20 µl of the dephosphorylation master mix and incubated for 20 min at 37°C at 1100 rpm on a thermomixer. After dephosphorylation, the beads were washed once with 1 ml PNK wash buffer (20 mM Tris-HCl pH 7.4, 10 mM MgCl2, 0.2% Tween-20), two times with 1 ml high-salt wash buffer (50 mM Tris-HCl pH 7.4, 1.5 M NaCl, 1 mM EDTA pH 8.0, 1% Igepal CA-630, 0.1% SDS, 0.5% Sodium deoxycholate) for 2 min at 4°C on a rotator and twice again with 1 ml PNK wash buffer (20 mM Tris-HCl pH 7.4, 10 mM MgCl2, 0.2% Tween-20).

3' adapter ligation: The following ligation mix containing (4x ligation buffer (200 mM Tris-HCl, pH 7.8, 40 mM MgCl2, 4 mM DTT, PEG 400, 3 µl of L3-App-Fluo adapter (10 µM)

(/rApp/AGATCGGAAGAGCGGTTCAG/ddC/), 0.5 µl SUPERase-In, 1 µl of T4 ligase in a total volume of 6.5 µl per sample) was used for adapter ligation at the 3' end of the dephosphorylated RNA. Now, the PNK buffer was removed, the beads were resuspended in 20 µl ligation mix and incubated overnight at 16°C at 1100 rpm on a thermomixer. On the following day, the beads were washed once with 0.5 ml PNK wash buffer (20 mM Tris-HCl pH 7.4, 10 mM MgCl2, 0.2% Tween-20), twice with 1 ml high-salt wash buffer (50 mM Tris-HCl pH 7.4, 1.5 M NaCl, 1 mM EDTA pH 8.0, 1% Igepal CA-630, 0.1% SDS, 0.5% sodium deoxycholate) for 2 min at 4°C on a rotator and once again with 0.5 ml PNK wash buffer. During the last wash, the samples were transferred to a fresh 1.5 ml low-bind tubes and re-suspended in 1ml PNK wash buffer. For western blot analysis, 100 µl of the samples were stored in a separate tube and the remaining 900 µl of the samples were used for RNA extraction. The samples were placed on a magnetic stand, buffer was removed and the samples were eluted using 35 µl 1x LDS buffer, boiled at 70°C for 10 min. Further, the eluted samples were loaded 7.5% Mini-PROTEAN® TGX™ precast protein gel and run at 120 V in a vertical electrophoresis chamber filled with 1× SDS running buffer (25 mM Tris base, 192 mM glycine, 0.1% (w/v) SDS). Western blot was performed using 0.45 µm nitrocellulose membrane using a wet transfer system with transfer buffer (25 mM Tris base, 192 mM glycine) containing 20% methanol for 1.5 h at 120 V in an ice bath.

Proteinase K digestion: Following the western blot, for extracting the RNA without the bound protein, the membrane was cut at 90 kDa-150 kDa to 4-5 small pieces and transferred into a new 2 ml low-bind tube. Further, the master mix (2x proteinase K buffer, 1 mg proteinase K, Ambion, cat no. AM2546) for proteinase K digestion was prepared. 400 µl of the mix was added to each tube containing the cut membrane pieces, vortexed for 20 seconds and incubated for 1 h 30 min at 55°C at 1000 rpm on a thermomixer for protein digestion.

RNA extraction: 2 volumes of acidic phenol-chloroform-IAA (pH 6.5-6.9) (Phenol-chloroform-Isoamylalkohol pH 6.5-6.9, Sigma Aldrich, cat no. P3803) was added directly to the proteinase K digested samples, mixed by inverting for 15 s and incubated at RT for 5 min for RNA extraction. In the meantime, phase lock gel heavy tubes (Serva, cat o. 733-2478) were prepared by spinning them at 12,000 g for 30 s to level the sample loading surface. The supernatant without the membrane pieces were transferred to the prepared phaselock gel heavy tubes, incubated at RT for 5 min at 1200 rpm on a thermomixer and centrifuged at 17,000 g for 15 min at RT. The aqueous layer was transferred to new 2 ml low-bind tubes. Further, all the steps until elution was performed based on the protocol provided in the RNA extraction kit (RNA Clean & Concentrator™-5 (50 Preps) w/ Zymo-Spin™ IC Columns (Capped), Zymo

research, cat no. R1015). Finally, the RNA was eluted into a fresh low-bind tube with nuclease free water by centrifuging at 15,000 g for 1 min at RT. The RNA was stored at -80°C until reverse transcription.

Reverse transcription: dNTPs and RT oligo mix containing (2 µl of RT oligo (1 µM), 1 µl dNTPs (10 mM each), and nuclease free water in a total volume of 5 µl per sample) were prepared and added to each RNA extracted from the membrane. The samples were mixed, briefly centrifuged and incubated in a thermomixer for 5 min at 65°C and then on ice for at least 1 min. 5x superscript IV buffer (SSIV) was vortexed, briefly centrifuge and the reverse transcription (RT) reaction (Superscript IV reverse transcriptase, Thermo Fisher Scientific, cat no. 18090050) mix was prepared (4 µl of 5x SSIV buffer, 1 µl 0.1 M DTT, 1 µl RNase OUT, 1 µl superscript IV in a total volume of 7 µl per sample). 7 µl of the RT reaction mix was added to each tube containing the RNA and incubated in a thermomixer at 25°C for 5 min, 42°C for 20 min, 50°C for 10 min, 80°C for 5 min and hold at 4°C. Later, 1 µl RNase H was added to each tube and incubated at 37°C for 20 min.

Cleanup I: cDNA cleanup was performed using MyONE silane beads (Thermo fisher Scientific, cat no. 37002D). The beads were briefly vortexed 10 µl beads were used per sample. The beads were washed once with 500 µl RLT buffer (Qiagen, cat no. 79216). After the wash step, the beads were resuspended in 125 µl RLT buffer and added to each sample, transferred to new 1.5 ml low-bind tube and mixed well by pipetting. 150 µl of 100% ice cold ethanol was added to the cDNA-beads complex, mixed well and incubated at RT for 5 min. The samples were mixed once again and incubated for additional 5 min at RT. Next, the beads were captured using a magnetic stand, the supernatant was discarded, and the beads were re-suspended in 900 µl of freshly prepared 80% ice cold ethanol and mixed by pipetting and transferred to a new low-bind tube. Further, the supernatant was discarded and the above step was repeated twice. The tubes were spun briefly to remove as much ethanol as possible, air dried at RT and finally, the beads were re-suspended in 5 µl nuclease free water.

Adapter ligation: 2 µl adapters L02clip2.0 (IgG), L02clip2.0 (KIFC1 replicate 1), L05clip2.0 (KIFC1 replicate 2), L10clip2.0 (KIFC1 replicate 3), L19clip2.0 (KIFC1 replicate 4), and L02clip2.0 (nuclease free water) from 10 µm stock were used from the publication Buchbender et al. (63). 1 µl of 100% DMSO was added to each tube, mixed well and heated on a thermomixer for 2 min at 75°C and the samples were immediately kept on ice for less than 1 min. The ligase mix (2 µl 10x NEB RNA ligase buffer with 10 mM DTT, 0.2 µl 100 mM ATP, 9 µl 50% PEG 8000, 0.5 µl high conc. RNA ligase in a total volume of 12 µl per sample) (T4 RNA Ligase 1 (ssRNA Ligase), High Concentration, NEB, cat no. M0437M) were prepared

and added to the tubes containing the beads and the adapters samples and mixed well. Additionally, 1 µl of high conc. RNA ligase was added to each sample, mixed well and the samples were incubated overnight at RT (20°C) at 1100 rpm on a thermomixer.

Cleanup II: MyONE beads were used for the second cleanup procedure and steps were followed as mentioned in the cleanup I section (see above). During the last step, the beads were resuspended in 23 µl nuclease free water, incubated at RT for 5 min, the beads were captured and the supernatant without the beads were transferred to new PCR tubes.

cDNA pre-amplification: Phusion master mix (2.5 µl P3Solexa_s and P5Solexa_s mix, 10 µM each, 25 µl 2x Phusion HF PCR master mix to a total volume of 27.5 µl per sample) was added to 22.5 µl cDNA and PCR amplification (98°C for 30 s, 98°C for 10 s, 65°C for 30 s, 72°C for 15 s, 72°C for 3 min) (2x Phusion High-Fidelity PCR master mix, NEB, cat no. M0531S) was performed for 6 cycles. The amplified cDNA was then size selected using the ProNex beads (ProNex® Size-Selective Purification System, Promega, cat no. NG2001) to reduce the primer-dimers formed during the PCR reaction.

ProNex size selection I: To discard fragments less than 55 nt and to retain fragments longer than 75 nt, size selection using ProNex beads was performed. Ultra-low range (ULR) ladder was used for reference (1 µl ULR ladder, 49 µl water) and for size selection (1 µl ULR ladder, 25 µl 2x Phusion HF PCR master mix in a total volume of 50 µl). First, the beads were incubated at RT for 30 min on a rotator for equilibrating. Next, 145 µl ProNex beads were added per sample (beads to sample ratio: 1:2.9), mixed well by pipetting and incubated at RT for 10 min. The beads were captured, supernatant was discarded, 300 µl ProNex buffer was added to the beads, incubated for 30-60 s and the supernatant was discarded. This step was repeated once more, the beads were air-dried. Later, the beads were resuspended in 23 µl nuclease free water. For size selection, the ULR ladder was resuspended in 50 µl nuclease free water and incubated at RT for 5 min. The samples were placed on a magnetic stand to capture the beads and the eluted cDNA was carefully transferred to a fresh PCR tube. The amplicon size of the cDNA was checked using ULR reference ladder for size selection using high sensitivity D1000 tape station kit (Agilent technologies, cat no. 5067-5584).

PCR cycle optimization: 1 µl of the size selected cDNA was used for PCR cycle optimization. Phusion master mix (0.5 µl P3 solexa PCR primer and P5 solexa PCR primer mix, 10 µM each, 5 µl 2x Phusion HF PCR master mix to a total volume of 9 µl per sample) were added to 1 µl cDNA and PCR amplification (98°C for 30 s, 98°C for 10 s, 65°C for 30 s, 72°C for 30 s, 72°C for 3 min) was initially performed for 7 and 10 cycles. Depending on the amount of cDNA obtained during the cycles and to limit the amplification within 10 cycles to minimize the PCR

duplicates, for this library I accordingly decided to continue with 8 and 9 cycles for preparative PCR.

Preparative PCR: Phusion master mix (2 µl P3Solexa_s and P5Solexa_s mix, 10 µM each, 20 µl 2x Phusion HF PCR master mix to a total volume of 30 µl per sample) was added to 10 µl cDNA and amplification was performed. 2 µl of the amplified library was used for run with a High Sensitivity D1000 tape station kit. The library obtained from 8 and 9 cycles were combined and all four replicates (KIFC1 replicates 1-4) were multiplexed. Further, residual primer-dimers from the PCR amplification were removed through second size selection using ProNex beads.

ProNex size selection II: ProNex selection was performed as described in the above section with ProNex size selection II with samples to beads ratio of 1:2.2. Finally, the library samples were eluted in 63 µl nuclease free water and the concentration was measured using a Qubit device. The samples were sequenced using Illumina Inc., NextSeq 550 high output v2.5, 150 cycles, 320 million reads platform at the DKFZ high-throughput core facility.

Sequence analysis: The 150 nt long reads were mapped to hg38 human genome (GENCODE v39) using STAR 2.5.3a (125) and uniquely mapped reads were used to perform further analysis, after the evaluating of the amount of rRNA sequences as previously described (126,127) BindingSiteFinder v2.0.0 was used to define binding sites as described in Busch et al. (128) and the binding site width was set to 7 nt. Maximum PureCLIP score from the replicates was assigned to each binding site and the reads were overlapped with gene annotations from GENCODE on human genome (GENCODE v43). The binding sites with the 5% lowest and highest scores were removed and the target spectrum of KIFC1 was generated after assigning the binding sites to distinct genes and gene regions. The pentamer frequencies were evaluated in the top 20% binding site as compared to the bottom 20% binding sites.

3.15 RNA affinity purification (RAP)

HeLa were cultured in a 15 cm, synchronized in prometaphase as described earlier and the frozen cell pellets (6 pellets per experiment) were used for the RAP assay.

Lysate preparation: The cell pellets were thawed on ice, lysed in 4 ml lysis buffer (50 mM HEPE-KOH pH7.5, 150 mM KCl, 2 mM EDTA, 1 mM NaF, 0.5% NP40, 0.5 mM DTT, 1X complete EDTA free protease inhibitor cocktail, 100 U/ml Ribolock) and incubated on ice for 30 min. The cell lysate was cleared by centrifugation at 17,000 g for 20 min at 4 °C. The supernatant was transferred to a new 2 ml tube, filtered through proteus clarification mini spin

column by centrifuging at 16,000 g for 2min at 4 °C and collected in a fresh 2ml tubes. BCA assay was performed to measure the protein concentration and the lysate was diluted to 10 mg/ml using lysis buffer (2 mg/sample) and kept on ice until pre-clearing.

Beads preparation: Dynabeads™ M-280 Streptavidin magnetic beads (Invitrogen, cat no. 11205D) were used. They were resuspended by vortexing for 30 s and desired volume (25 µl for 2 mg lysate) were transferred to a fresh 2 ml tube and washed using at least 1 ml of washing buffer (5 mM Tris-HCl pH 7.5, 1 M NaCl, 0.5 mM EDTA) and vortexing for 5 s. The supernatant was discarded and the wash step was repeated once more. The beads were washed twice with same as initial or larger volume with solution A (0.1 M NaOH, 50 mM NaCl) for 2 min and twice with solution B (0.1 M NaCl).

Pre-clearing the lysate: Supernatant was discarded from the beads, resuspended in 1 ml lysate and they incubated for 1-2 h at 4 °C on a rotator. Further, the pre-cleared lysate was transferred to a fresh tube and flash frozen using liquid nitrogen and stored at -80 °C until further use.

In vitro transcription: For the gBlock sequences, refer to the Key Resource Table 1.

The RNA candidates were ordered as gblocks from IDT at a length of 125 nt around the binding site, including the minimal T7 promoter sequence for in-vitro transcription of the gblocks to RNA. Further, the RNA was biotinylated and the non-biotinylated RNA was used as a negative control for RAP. Biotin (Biotin-16-UTP, BIOZYM SCIENTIFIC GM / Epicenter, cat no. BU6105H), buffers and NTP stock solutions were thawed at RT and biotin was diluted to 25mM with nuclease free water. For all the 5 RNA candidates, the number of U-residues were calculated and the *in vitro* transcription reaction for biotinylated/non-biotinylated candidates were performed in a total volume of 20 µl per sample TUBA1C (2 µl ATP 75 mM, 2 µl GTP 75 mM, 2 µl CTP 75 mM, 2 µl UTP 75 mM, 7.1 µl gblock 10 ng/µl, 2 µl 10X reaction buffer, 2 µl enzyme mix additionally for biotinylated RNA, 1.5 µl UTP 75 mM and 1.4 µl Biotin-UTP 25 mM were added), KIF14 (2 µl ATP 75 mM, 2 µl GTP 75 mM, 2 µl CTP 75 mM, 2 µl UTP 75 mM, 7.5 µl gblock 10 ng/µl, 2 µl 10X reaction buffer, 2 µl enzyme mix additionally for biotinylated RNA, 1.7 µl UTP 75 mM and 0.8 µl biotin-UTP 25 mM were added), HMGA1 (2 µl ATP 75 mM, 2 µl GTP 75 mM, 2 µl CTP 75 mM, 2 µl UTP 75 mM, 7 µl gblock 10 ng/µl, 2 µl 10X reaction buffer, 2 µl enzyme mix additionally for biotinylated RNA, 1.5 µl UTP 75 mM and 1.5 µl biotin-UTP 25 mM were added), HNRNPU (2 µl ATP 75 mM, 2 µl GTP 75 mM, 2 µl CTP 75 mM, 2 µl UTP 75 mM, 7.3 µl gblock 10 ng/µl, 2 µl 10X reaction buffer, 2 µl enzyme mix additionally for biotinylated RNA, 1.7 µl UTP 75 mM and 1 µl biotin-UTP 25 mM were added), LMO4 (2 µl ATP 75 mM, 2 µl GTP 75 mM, 2 µl CTP 75 mM, 2 µl UTP 75

mM, 7 µl gblock 10 ng/µl, 2 µl 10X reaction buffer, 2 µl enzyme mix additionally for biotinylated RNA, 1.5 µl UTP 75 mM and 1.5 µl biotin-UTP 25 mM were added). All the reagents were pipetted in a PCR tube, mixed, spun down and were incubated overnight at 37 °C for up to 16 h on a PCR cycler. Further, the following day, 1µl DNase I was added and incubated at 37 °C for 15 min to digest any residual DNA. The RNA was purified using RNA clean and concentrator kit, eluted in 20 µl nuclease free water and the RNA concentration was measured using nano-drop.

RNA pulldown: 50 pmol RNA was used per sample. RNA was pipetted into a fresh 1.5 ml low-bind tube, denatured at 65 °C for 10 min and cooled down to RT for over 20 min on a thermomixer. Meanwhile, the pre-cleared lysates were thawed on ice and few microlitres were used for preparing the input control with 1X LDS and DTT. The remaining lysate was supplemented with 100 U/ml ribolock. 2 mg total lysate (per sample) was added to the denatured RNA, incubated at 4 °C for 1 h on a rotator.

Meanwhile, magnetic Dynabeads™ M-280 streptavidin beads (50 µl per sample) were prepared by washing as mentioned in the earlier section (Beads preparation). After the incubation, 50 µl beads were added to the RNA-protein complex and incubated for 1.5 h at 4 °C on a rotator for capture. Further, the supernatant was removed and the beads were washed thrice with 1 ml wash buffer (5 mM Tris-HCl pH 7.5, 150 mM NaCl, 0.1 % tween-20, 0.5 mM DTT, 1X complete EDTA dree protease inhibitor cocktail). During the last wash, the beads were transferred to a fresh 1.5 ml low-bind tube and the RNA-protein complexes were eluted using 30 µl 1X LDS supplemented with 50 mM DTT, boiled at 70 °C for 10 min. The eluate was stored in a fresh 1.5 ml low-bind tube and were analysed using SDS-PAGE and western blot.

3.16 Cloning

For primer sequences, please refer to Table 3.

KIFC1 cDNA ORF clone in cloning vector was purchased (Sino biologicals, cat no. HG15958-G). KIFC1 WT pENTRY clone was cloned into pDONOR221 plasmid by PCR (25 µl 2x Phusion high-fidelity PCR master mix, 1 µl DNA (10 ng in total), 2.5 µl forward and reverse primer (10 µM), 1.5 µl DMSO in a total volume 50 µl per reaction) amplification at 98°C for 1 min, 98°C for 10 s, 55°C for 30 s, 72°C for 1 min, 72°C for 10 min and hold at 4°C) for 40 cycles. The amplicon was run on 0.8% agarose gel, the band was cut out, purified using the GeneJET gel extraction kit (Thermo Fisher Scientific, cat no. K0692) according to the

manufacturer's instructions. The plasmid was eluted in 50 µl nuclease free water. Further, BP reaction (1 µl PCR product (~50 ng/µl), 1 µl pDONOR221 (150 ng/µl), 6 µl TE buffer pH 8, 2 µl 5x BP clonase enzyme mix in a total volume of 10 µl) (Gateway™ BP Clonase™ II Enzyme-Mix, Thermo Fisher Scientific, cat no. 11789100) was carried out at 25 °C overnight in a PCR machine. Finally, 1 µl proteinase K was added, incubated at 37 °C for 10 min in a PCR machine to terminate the reaction.

Transformation: One shot top 10 chemically competent E. coli cells (Thermo Fisher <scientific, cat no. C404003) were transformed with 1.5 µl BP reaction mix. The cells were incubated on ice for 30 min, followed by heat shock at 42 °C for 2 min, and again incubated on ice again for 2 min. Further, 250 µl LB medium was added to the cells and incubated for 1 h at 37 °C with mild shaking. Later, 100 µl cells were plated out on LB agarose plates containing 10 µg/ml kanamycin and incubated overnight at 37 °C for the bacteria to grow.

The following day, 2 colonies were picked per plate, plasmid isolation was performed using the Nucleospin plasmid isolation kit (Macherey Nagel, cat no. 740588.250) based on manufacturers protocol and were sequenced prior to use.

The KIFC1 WT sequences were shuttled into pFRT-flag-HA-ΔCmR-ΔccdB vector using gateway cloning system for further use.

Serine (S) and threonine (T) residues were mutated to alanine, an amino acid residue which converts the protein into non-phosphorylatable form at these residues (KIFC1-S6A, KIFC1-S26A, KIFC1-S31A, KIFC1-S96A, KIFC1-T187A, KIFC1-S221A, KIFC1-S349A, KIFC1-T359A, KIFC1-S349A/T359A). The mutation was performed using the Q5 Site-Directed Mutagenesis Kit (NEB, cat no. E0554S). Single or double base substitutions were performed using respective primers and PCR amplification(1 µg DNA template, 12.5 µl Q5 Hot Start High-Fidelity 2x Master Mix, 1.25 µl 10 µM forward and reverse primer each, in a total volume of 25 µl per sample) at (Initial denaturation 98 °C for 30 s, denaturation at 98 °C for 10 s, annealing at 61-72°C for 30 s depending on the individual primers, extension at 72 °C for 3 min, final extension at 72 °C for 2 min and hold at 4 °C for 25 cycles). The mutations were performed on the KIFC1 WT sequence on pFRT-flag-HA-ΔCmR-ΔccdB plasmid backbone. The amplicon size and integrity were verified using 10 % of the PCR amplified product run on 0.8 % agarose gel.

Transformation: Kinase, ligase and DpnI (KLD) treatment (PCR product 1 µl, 5 µl 2x KLD reaction buffer, 1 µl 10x KLD enzyme were added to a total volume of 10 µl per reaction) was performed for 5 min at RT prior to transformation. Following the treatment, 5 µl of the KLD reaction mixture was added to 50 µl NEB® 5-alpha Competent E. coli (High Efficiency) cells

(NEB, cat no. C2987H). Further, the cells were incubated for 30 min on ice, followed by heat shock at 42 °C for 30 s, and again incubation on ice for 5 min. Further, the cells and incubated at 37 °C for 1 h with 950 µl SOC media and gentle shaking. Later, 50-100 µl of the cells were plated on LB agar plate with appropriate selection antibiotics and incubated overnight at 37 °C for the bacteria to grow.

The following day, 2 colonies were picked per plate and per mutant, plasmid isolation was performed using the Nucleospin plasmid isolation kit (Macherey Nagel) based on manufacturers protocol and sequences were verified to use.

3.17 *In vitro* kinase assay

Transfection: On the first day, 5 million HeLa cells were plated out on a 15 cm dish. On the following day, the cells were transfected with plasmids (pFRT-flag-HA-ΔCmR-ΔccdB as empty vector control, flagHA-KIFC1_WT, flagHA-KIFC1_S6A, flagHA-KIFC1_S26A, flagHA-KIFC1_S31A, flagHA-KIFC1_S96A, flagHA-KIFC1_T187A, flagHA-KIFC1_S221A, flagHA-KIFC1_S349A, flagHA-KIFC1_T359A) using lipofectamine 2000 (Thermo Fisher Scientific, cat no. 11668500) at a DNA:lipofectamine ratio of 1:2.5, based on the manufacturer's instructions. After 6 h, the plates were washed once with warm PBS and fresh media was added to the cells. After 24 h, the cells were harvested, the pellets were flash frozen and stored at -80 °C until further use.

Cell lysis: Each cell pellet was lysed in 1 ml lysis buffer (50 mM Tris-HCl pH 7.4, 150 mM NaCl, 1 mM EDTA, 1 mM NaF, 1% triton X-100, 1x complete EDTA free protease inhibitor cocktail). The cells were incubated on ice for 30 min for lysis and centrifuged at 17,000 g for 20 min at 4 °C and the supernatant was collected into a fresh 2 ml low-bind tube. Further, the supernatants were filtered using proteus clarification mini spin column and centrifuging at 16,000 g for 2 min at 4°C. Protein concentration was determined using BCA assay and the lysates were diluted to 1.5 mg/ml using lysis buffer.

Beads preparation and immunoprecipitation: Anti-DYKDDDDK magnetic agarose beads (Pierce™ Anti-DYKDDDDK Magnetic Agarose, Thermo Scientific, cat no. YD368599) were mixed by gentle vortexing and 37.5 µl beads were used per sample containing 1.5 mg lysate. The beads were washed three times with 1 ml lysis buffer. During the last wash, the beads were transferred to a fresh 1.5 ml tube, the supernatant was discarded and lysate was added to the beads. The lysate-beads mix were incubated for 2 h at 4 °C on a rotator. Further, the supernatant was removed and the beads were washed 4 times with 1 ml wash buffer (50 mM Tris-HCl pH

7.4, 350 mM NaCl, 1% triton X-100, 1x complete EDTA free protease inhibitor cocktail) for 5 min at 4 °C on a rotating wheel. During the last wash, the beads were transferred to a fresh 1.5 ml tube, the supernatant was discarded and the beads were resuspended in 1 ml PNK buffer (20 mM Tris-HCl pH 7.4, 10 mM MgCl2, 0.2% tween-20, 1x complete EDTA free protease inhibitor cocktail) and stored at 4 °C overnight. The following day, 100 µl of the samples were used for western blot analysis to check for IP efficiency.

In vitro kinase assay for single mutants: 1x kinase assay buffer was prepared from 10 x kinase assay buffer (500 mM Tris-HCl pH 7.5, 100 mM NaCl, 100 mM MgCl2, 10 mM DTT, 1x complete EDTA free protease inhibitor cocktail). The insect cell purified AURKA protein (obtained from EMBL Protein Expression and Purification Core Facility) was diluted to 0.5 mg/ml in 1x kinase assay buffer. Next, the kinase assay master mix was prepared (negative control without purified AURKA protein: 2 µl 1x kinase assay buffer, 2 µl 10x kinase assay buffer, 2 µl 1 mM cold ATP and 0.5 µl 32P-γ-ATP; kinase assay with purified AURKA protein: 2 µl purified AURKA protein 0.5 mg/ml, 2 µl 10x kinase assay buffer, 2 µl 1mM cold ATP and 0.5 µl 32P-γ-ATP in a total volume of 6.5 µl per sample). Further, supernatant was removed, and the beads were resuspended in 6.5 µl of the master mix and incubated for 30 min at 30°C in a thermomixer at 800 rpm. After the incubation, the samples were eluted using 5 µl 4x LDS supplemented with 200 mM DTT, boiled at 70 °C for 10 min.

To visualize the phosphorylation signal, SDS-PAGE analysis was performed. The samples were loaded in a 7.5% Mini-PROTEAN® TGX™ precast protein gel and run at 120 V in a vertical electrophoresis chamber filled with 1× SDS running buffer (25 mM Tris base, 192 mM glycine, 0.1 % (w/v) SDS). Later, the gel was fixed for 1 h at RT with slow rocking using 15 ml fixation solution (50 % methanol, 10 % acetic acid), washed three times with nuclease free water for 5 min and dried for 1-1.5 h at 80 °C. The dried gel was then exposed to phosphor imager screen and was scanned after appropriate amount of time using Typhoon laser scanner phosphor imager at 200 µm, high speed and intensity 3.

In-vitro kinase assay for double mutant: 5 million HeLa cells were seeded on a 15 cm dish. On the next day, the cells were transfected with plasmids (pFRT-flag-HA-ΔCmR-ΔccdB (empty vector control), flagHA-KIFC1_WT, flagHA-KIFC1_S349A, flagHA-KIFC1_T359A, flagHA-KIFC1_S349A/T359A) using lipofectamine 2000 at DNA:lipofectamine ratio of 1:2.5, based on the manufacturer's instructions. After 6 h, the plates were washed once with warm PBS and fresh media was added to the cells. 24 h post transfection, the cells were harvested, the pellets were flash frozen and stored at -80 °C until further use.

The kinase assay was performed just as described above in this section. During the kinase reaction, insect cell purified AURKA kinase dead mutant D274A (obtained from EMBL Protein Expression and Purification Core Facility) was used as a negative control and AURKA WT protein was used for the kinase activity to demonstrate the phosphorylation of KIFC1 WT vs the phosphorylation mutants. Further, after fixing the gel, the gel was first stained with Coomassie stain (to visualize protein bands) overnight at RT on a shaker. Further, the gel was washed with nuclease free water for 1 h at RT, imaged and then dried and exposed to phosphor imager screen.

In vitro kinase assay +/- RNase: 5 million HeLa cells were seeded on a 15 cm dish and transfected with plasmids (Empty vector control: pFRT-flag-HA-ΔCmR-ΔccdB, flagHA-KIFC1_WT) the following day, using lipofectamine 2000 at DNA: lipofectamine ratio of 1:2.5, based on the manufacturer's instructions. After 6 h, the plates were washed once with warm PBS and fresh media was added to the cells. 24 h post transfection, the cells were harvested, flash frozen and stored at -80 °C until further use.

Cell lysis: Each cell pellet was lysed in 1 ml lysis buffer (50 mM Tris-HCl pH 7.4, 150 mM NaCl, 1 mM EDTA, 1 mM NaF, 1% triton X-100, 1X complete EDTA free protease inhibitor cocktail), incubated on ice for 30 min and BCA assay was performed to measure the protein concentration. The lysates were diluted to 4 mg/ml and the lysates were aliquoted to 4 tubes per construct. The lysates were treated with 4 µl turbo DNase, 10 µl 1:5 diluted RNase 1 for +RNase samples and 10 µl lysis buffer of control samples at 37 °C for 3 min at 1100 rpm on a thermomixer. After the nuclease treatment, the control and RNase-treated lysates were cooled down on ice for 3 min and centrifuged at 17,000 g for 20 min at 4 °C. The supernatant was collected into a fresh 2 ml low-bind tubes and filtered through proteus clarification mini spin column by centrifuging at 16,000 g for 2 min at 4 °C. BCA assay was performed again to determine the protein concentration and the lysates were diluted to 2 mg/ml using lysis buffer. 50 µl beads were used per sample.

Further the beads preparation, immunoprecipitation and the kinase assay were performed as described in the above section (In-vitro kinase assay for single mutants). During the kinase reaction, insect cell purified AURKA kinase dead mutant D274A (obtained from EMBL Protein Expression and Purification Core Facility) was used as a negative control and AURKA WT protein was used for the kinase activity to demonstrate the phosphorylation of KIFC1 WT vs the phosphorylation mutants. Further, after fixing the gel, the gel was first stained with Coomassie stain (to visualize protein bands) overnight at RT on a shaker. Further, the gel was

washed with nuclease free water for 1 h at RT, imaged and then dried and exposed to phosphor imager screen.

Table 1: gBlock DNA sequences used for RNA affinity pulldown

gBlocks	**Sequence**	**Company**
TUBA1C	TAATACGACTCACTATAGGGTGCCTGGCGGAGACCT GGCCAAGGTACAGAGAGCTGTGTGCATGCTGAGCAA TACCACAGCTGTTGCCGAGGCCTGGGCTCGCCTGGA CCACAAGTTTGACCTGA	IDT
KIF14	TAATACGACTCACTATAGGGTGTTTTGTTAAAGGTAT AATGAAATAATTTGTATATGATTTGATGAAGATTAA AGACCCTTATTTTCCACAGCTTTAAAAAAAAACCTTT ATTTATGATCAAGTA	IDT
HMGA1	TAATACGACTCACTATAGGGCCGCCCACCCACGCAT ACACACATGCCCTCCTGGACAAGGCTAACATCCCAC TTAGCCGCACCCTGCACCTGCTGCGTCCCCACTCCCT TGGTGGTGGGGACATT	IDT
HNRNPU	TAATACGACTCACTATAGGGTCGTATTGGCTGGTCAC TAACTACAAGTGGAATGTTACTTGGTGAAGAAGAAT TTTCTTATGGGTATTCTCTAAAAGGAATAAAAACATG CAACTGTGAGACTGA	IDT
LMO4	TAATACGACTCACTATAGGGAAAAAGATCCACCAGA GGACATCTTGGGGAGGGGGAGGGAGCTGGGGGGGA GGGAAATGACTAATGAAGCTAATTAAAAGAAGCATT CAAATCTGCTTTCTACCC	IDT

Table 2: Primers, adapters and barcode sequences used for iCLIP-Seq of KIFC1

Adapters or barcodes or primers	**Sequence**	**Company**
L3-App	/5rApp/AG ATC GGA AGA GCG GTT CAG /3ddC/	IDT
RT oligo	GGATCCTGAACCGCT	Sigma-Aldrich
P3Solexa PCR primer	CACGACGCTCTTCCGATCT	Sigma-Aldrich
P5Solexa PCR primer	CTGAACCGCTCTTCCGATCT	Sigma-Aldrich
P3Solexa adapter	AATGATACGGCGACCACCGAGATCTACACTCTTTCCC TACACGACGCTCTTCCGATCT	Sigma-Aldrich
P5Solexa adapter	CAAGCAGAAGACGGCATACGAGATCGGTCTCGGCAT TCCTGCTGAACCGCTCTTCCGATCT	Sigma-Aldrich
L02clip2.0	/5Phos/NN NNC GAT GTN NNN NAG ATC GGA AGA GCG TCG TG/3ddC/	IDT
L05clip2.0	/5Phos/NN NNA CAG TGN NNN NAG ATC GGA AGA GCG TCG TG/3ddC/	IDT
L10clip2.0	/5Phos/NN NNT AGC TTN NNN NAG ATC GGA AGA GCG TCG TG/3ddC/	IDT
L19clip2.0	/5Phos/NN NNG TGA AAN NNN NAG ATC GGA AGA GCG TCG TG/3ddC/	IDT

Table 3: Primer and other sequences used for cloning and in vitro kinase assay

Cloning primers	**Sequence**		**Company**
KIFC1 Forward	GGGGACAAGTTTGTACAAAAAAG CAGGCTTCATGGATCCGCAGAGGT CCCCCCTATTG	Primers to generate Entry clone for Gateway cloning.	Sigma-Aldrich
KIFC1 Reverse	GGGGACCACTTTGTACAAGAAAGC TGGGTTCTATCACTTCCTGTTGGCC TGAGCAGTACC		Sigma-Aldrich
S6A_FP	TCCGCAGAGGGCCCCCCTATTGG		Sigma-Aldrich
S6A_RP	TCCATTCACTTCCTGTTGGCCTGAG C		Sigma-Aldrich
S26A_FP	TAAGGCCCCTGCCCAGCTGCCTC		Sigma-Aldrich
S26A_RP	ATCAGAGGTCTCTTCAGTTCTATG TTCCCCTTTAC		Sigma-Aldrich
S31A_FP	GCTGCCTCTCGCAGGAAGCAGAC		Sigma-Aldrich
S31A_RP	TGGGAAGGGGCCTTAATCAG		Sigma-Aldrich
S349A_FP	AACCCGCCTTGCGCTCTCCCGGTC TGAC		Sigma-Aldrich
S349A_RP	GGAGGATCAGAGGGCCCA		Sigma-Aldrich
T359A_FP	GCGGCGTGGGGCCCTGAGTGGGG		Sigma-Aldrich
T359A_RP	TCGTCAGACCGGGAGAGGCTAAG GCG		Sigma-Aldrich
pFastBac_KIFC1-FP	ATGGATCCGGAATTCAAAG	Primer for NEBuilder Cloning to generate KIFC1 for expression in insect cells	Sigma-Aldrich
pFastBac_KIFC1-RP	GGCGCCCTGAAAATACAG		Sigma-Aldrich
KIFC1-FP	ACCTGTATTTTCAGGGCGCCATGG ATCCGCAGAGGTCC		Sigma-Aldrich
KIFC1-RP	CCTTTGAATTCCGGATCCATTCACT TCCTGTTGGCCTG		Sigma-Aldrich
pFastBac_AURKA-FP	ACAGTCTTAGTAATCAGCCATACC ACATTTG	Primer for NEBuilder Cloning to generate AURKA for expression in insect cells	Sigma-Aldrich
pFastBac_AURKA-RP	ATCGGTCCATGGCGCCCTGAAAAT ACAG		Sigma-Aldrich
AURKA-FP	TCAGGGCGCCATGGACCGATCTAA AGAAAAC		Sigma-Aldrich
AURKA-RP	TGGCTGATTACTAAGACTGTTTGC TAGC		Sigma-Aldrich
D274A_FP	TAAAATTGCAGCTTTTGGGTGGTC		Sigma-Aldrich
D274A_RP	AGCTCTCCAGCTGATCCA		Sigma-Aldrich

Table 4: Equipments used in the project

Equipments	**Catalogue Number**	**Company**
Imager ECL Chemo Cam CC5569	N/A	INTAS
Stratalinker 2400	N/A	Stratagene
SW 40 Ti Swinging-Bucket Rotor	331302	Beckman Coulter
Ultra-Clear Tube	344060	Beckman Coulter
Orbitrap Fusion LC-MS/MS platform	N/A	Thermo Fisher Scientific
Discovery 90SE Ultracentrifuge	N/A	Sorvall
Sequencing Grade Modified Trypsin	V5113	Promega
TMTsimplex™ Isobaric Label Reagent Set	90066	Thermo Fisher Scientific

Table 5: Software used in the project

Bioconductor	Gentleman et al., 2004	www.bioconductor.org
Comet	Eng et al., 2013	http://comet-ms.sourceforge.net
ImageJ	Schneider et al., 2012	https://imagej.net/Downloads
LabImage 1D 2006	Kapelan Bio-Imaging GmbH	www.labimage.com
Microsoft Excel	Microsoft	www.microsoft.com
Primer Blast	NCBI	https://www.ncbi.nlm.nih.gov/tools/primer-blast/
R programming	The R project	www.r-project.org
Shiny	R Studio	https://shiny.rstudio.com
GO Enrichment Analysis	Thomas PD et al., 2022	https://geneontology.org/
Serial cloner	Serial Basics	http://serialbasics.free.fr/Serial_Cloner.html
GIMP	The GNU Project	https://www.gimp.org
Ensemble	Harrison et al., 2024	https://www.ensembl.org/index.html

4 Results

4.1 Part I: R-DeeP screen in lung cancer cells

Cancer is a complex disease with multiple mutations and dysregulated factors across different cancer diseases. Lung cancer is the leading cause of cancer related deaths, contributing to 18% of the mortality worldwide in 2020 (129,130). More than 80 % of the lung cancer cases are classified as non-small cell lung cancer (NSCLC) that encompasses adenocarcinoma, squamous cell carcinoma and large cell carcinoma with a poor 5-year survival rate (63,130). To device an effective therapeutic strategy to treat such complex disease, it is crucial to understand and characterize the molecular details of disease mechanism at RNA and protein level. Additionally, to obtain an overview on RBPs in lung cancer, in the first part of the project, I focussed on identifying RNA-dependent proteins in lung cancer by adapting R-DeeP screen in A549 cell line (63). A549 cell lines are epithelial cells derived from lung adenocarcinoma and have been widely used as a model for characterizing lung cancer (131). First, the cells were cultured, lysed and the lysates were either treated with RNase (RNase sample) or not treated with RNase (control sample). After treatment, the lysates were loaded onto the sucrose density gradients and subjected to ultracentrifugation (see Method section). Later, the gradients were fractionated into 25 different fractions which were quantitatively analysed using mass spectrometry. Further, the results were validated using western blot analysis for individual proteins of interest (Figure 9). The gradients and mass spectrometry analysis were performed by my colleague Dr. Astrid-Solveig Loubal, who was a postdoctoral researcher in the lab before I joined. The gradients were prepared in triplicates, amounting to 150 samples in total. The mass spectrometry analysis detected 3743 proteins across the samples. The established bioinformatic pipeline to identify RNA-dependent proteins (59,60) was used to calculate Gaussian fitted distribution profile for each protein (Figure 9).

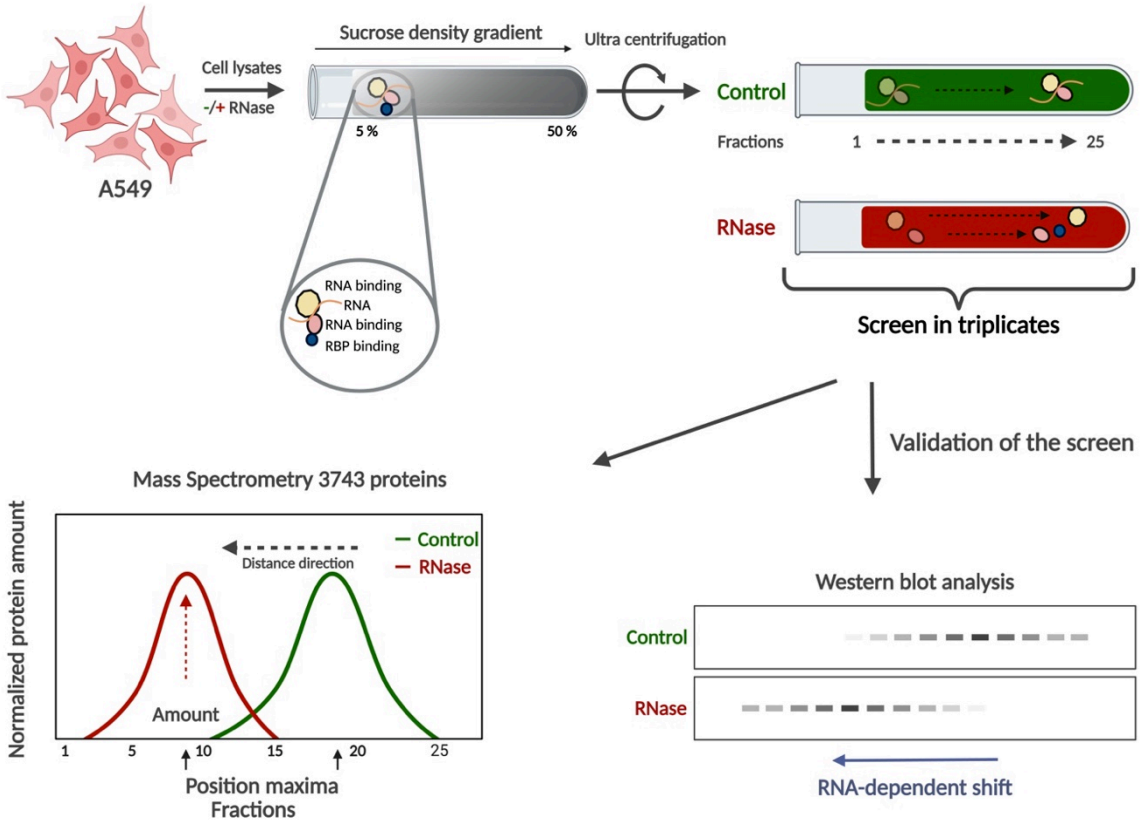


Figure 9: Schematic describing the R-DeeP method in lung cancer cells

The cells were cultured, lysed and the lysates were treated with (+RNase) or without (control) RNase. The lysates were subjected to ultracentrifugation for 18 h at high speed (see Materials and Methods section). Later, the control and RNase-treated gradients were fractionated into 25 different fractions and were analysed using mass spectrometry analysis followed by western blot for the validation of the screen. The gradients were prepared in triplicates and the raw mass spectrometry data were fitted using Gaussian curves to analyse various parameters such as position of the maxima, amplitude difference, shift distance, and amount of protein as given by the area under the curve at each peak. The screen is based on the concept of RNA dependence. Upon RNase treatment, the protein complexes that are formed by interacting with RNA are released into smaller complexes or individual proteins migrate to a position corresponding to their monomeric size causing a decrease in the apparent molecular weight of a protein between control and RNase-treated gradients. This results in the detection of the protein peak amount in earlier fractions in RNase-treated gradients or a "left shift" compared to the protein peak in the control gradients. This indicates the RNA dependence of a protein. This figure was generated using Biorender based on (59,63).

4.1.1 Analysis of the RNA-dependent shifts

Each protein displayed a specific migration pattern throughout the gradient depending on the presence or absence of RNA. The shift in the distribution of a protein between control and RNase-treated gradients indicated the RNA dependence of a protein. The shifts were characterized based on multiple criteria: (i) amount of protein shifting in RNase-treated and control gradients indicated by area under the curve. (ii) Position of the maxima in the curve

reflecting the apparent molecular weight of the protein. (iii) distance and the direction of shift (iv) the difference in amplitude between the maxima in control vs RNase-treated curves. (v) statistical significance in the difference between the amplitude maxima.

A protein was classified as RNA dependent if the distance between the maxima in control vs RNase-treated samples was strictly greater than 1 fraction with a significant difference (adjusted p-value < 0.05) in at least one of the maxima of either the control or the RNase profile. Upon fraction-wise normalization, the amount of protein present in each fraction correlated well between the three replicates for all the proteins detected in the screen, reflecting the reproducibility of the method (Figure 10A). Proteins exhibiting significant difference in their distribution pattern between the control and RNase-treated gradients were further grouped into four groups: (i) left-shifting proteins, *i.e.* proteins shifting to lower fractions upon RNase treatment, (ii) right-shifting proteins, *i.e.* proteins shifting to higher fractions upon RNase treatment (iii) non-shifting proteins, *i.e.* the distribution of a proteins is unchanged between control and RNase-treated samples and finally (v) precipitated proteins that are accumulated in the last fraction. Out of the 3743 total number of proteins detected in the screen, we observed 1525 left shifts (blue), 241 right shifts (red) and 260 precipitated proteins after RNase treatment (orange), whereas no significant shift was observed for 2554 proteins (grey) (Figure 10B, C). Here, it is important to consider that certain proteins might contain multiple peaks and such proteins with multiple peaks could also show multiple shifts. In comparison with previously published data on 43 RBP proteome-wide studies in human, we detected 1189 proteins with at least one significant shift out of which, 170 novel RNA-dependent proteins were identified, which had not been previously categorized as RBP (Figure 10D) (56). Further, exploiting the quantitative nature of the R-DeeP screen, a shifting coefficient was calculated for each protein based on the amount of protein present in a peak and its change upon RNase treatment. Based on the calculated shifting coefficient, proteins were further classified into (i) completely RNA dependent: complete shift in the protein amount upon RNase treatment, (ii) partially RNA dependent: only a fraction of the protein amount shifted after RNase treatment and (iii) RNA independent: the protein amount was not disturbed upon RNase treatment (Figure 10E).

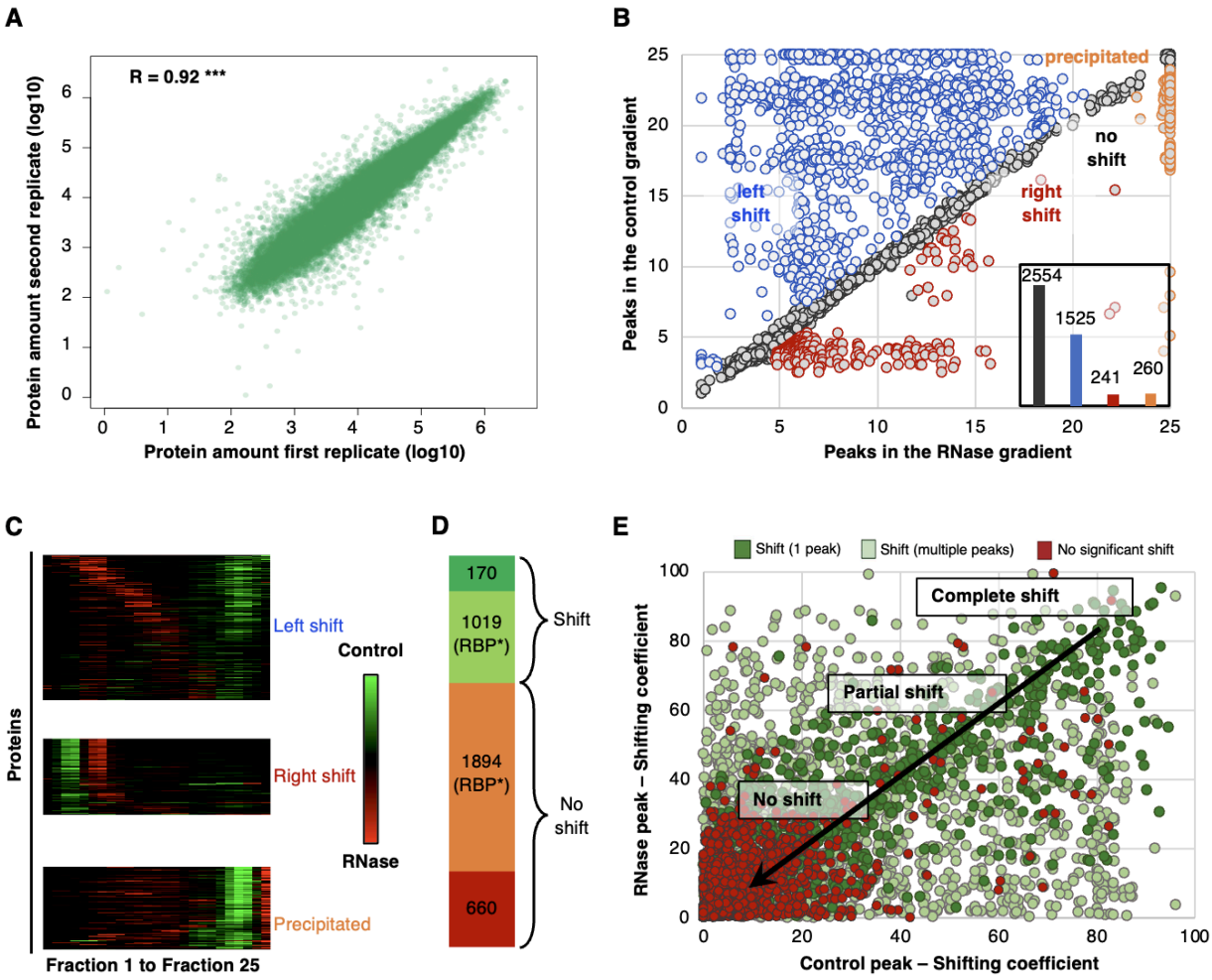


Figure 10: Analysis of the RNA-dependent shifts

*(A) Correlation plot representing the reproducibility of the method after fraction wise normalization of the amount of individual protein present in each fraction in all the three replicates. Pearson's correlation coefficient R = 0.92, *** p-value < 0.001.*

(B) The graph indicates the position of the peaks in control and RNase-treated fractions for each shift based on the mean of three replicates. The inset bar graph shows the number of proteins shifting towards left (blue), right (red), not shifting (grey) or precipitating (orange).

(C) Heatmap representing significant shifts based on their direction (left shifts, right shifts and precipitation). Green represents the proteins in control fractions and red represents the proteins in RNase-treated fractions.

(D) Shows the proteins detected in the R-DeeP screen in A549 cells, categorized as shifting and non-shifting proteins. Red box represents non-shifting proteins that were not identified as RBPs before (660). Orange box represents non-shifting proteins that were identified as RBPs at least once in the 43 previous proteome-wide studies (1894). Light green box represents shifting proteins that were also identified as RBPs at least once in other studies (1019). Dark green box represents the shifting proteins that were not identified as RBPs in the 43 previous proteome-wide studies (170). RBP indicates an RBP or an RBP candidate.*

(E) The graph represents the shifting coefficient (protein amount at maxima × loss or gain after the shift) for each pair of control and RNase-treated peak. Proteins with no significant shifts are indicated by red dots, proteins with significant shifts between one control and one RNase-treated peaks are indicated with green dots and light green represents proteins with multiple peaks. The top right region of the graph represents proteins with a "complete shift" (almost the entire protein amount is shifting), bottom left indicates proteins with no shift and the candidates in the middle are the proteins with a partial shift (a fraction of the protein is shifting, while another fraction does not show any shift or no RNA dependence). The figure was adapted from (63).

4.1.2 Properties of shifting proteins

To analyse the properties of the shifting proteins and to discard the false positives, multiple characteristics of shifting proteins were analysed that were related to RBPs. First, the RBP2GO score was calculated for shifting and non-shifting proteins. The RBP2GO score is calculated based on the frequency of a protein in being listed as an RBP in the 43 proteome-wide human studies. This reflects the probability of a protein to be an RBP *i.e.* provides an estimation of the likelihood for each protein for being a true RBP (56). The RBP2GO score for the shifting proteins were significantly higher than for the non-shifting group (Figure 11A). Second, the number of RBDs in the protein sequence was taken into account. The canonical RBPs usually bind RNA through one or more of their RBDs. Hence, the number of RBDs in the shifting and the non-shifting groups were compared. As expected, the number of RBDs were significantly enriched in the shifting groups as compared to the non-shifting proteins (Figure 11B). Also, the shifting proteins presented a higher fraction of RBDs relative to their length, when compared to the non-shifting proteins (Figure 11C). As canonical RBPs usually bind RNA through their well-defined RBDs, one other way through which RBPs could bind RNA is through their IDRs. The IDRs contain low complexity amino acid sequences and lack a well-defined structure. These regions are stabilized upon the interaction with RNA. Appropriately, the shifting proteins contained significantly higher disordered fractions relative to their length compared to the non-shifting proteins (Figure 11D).

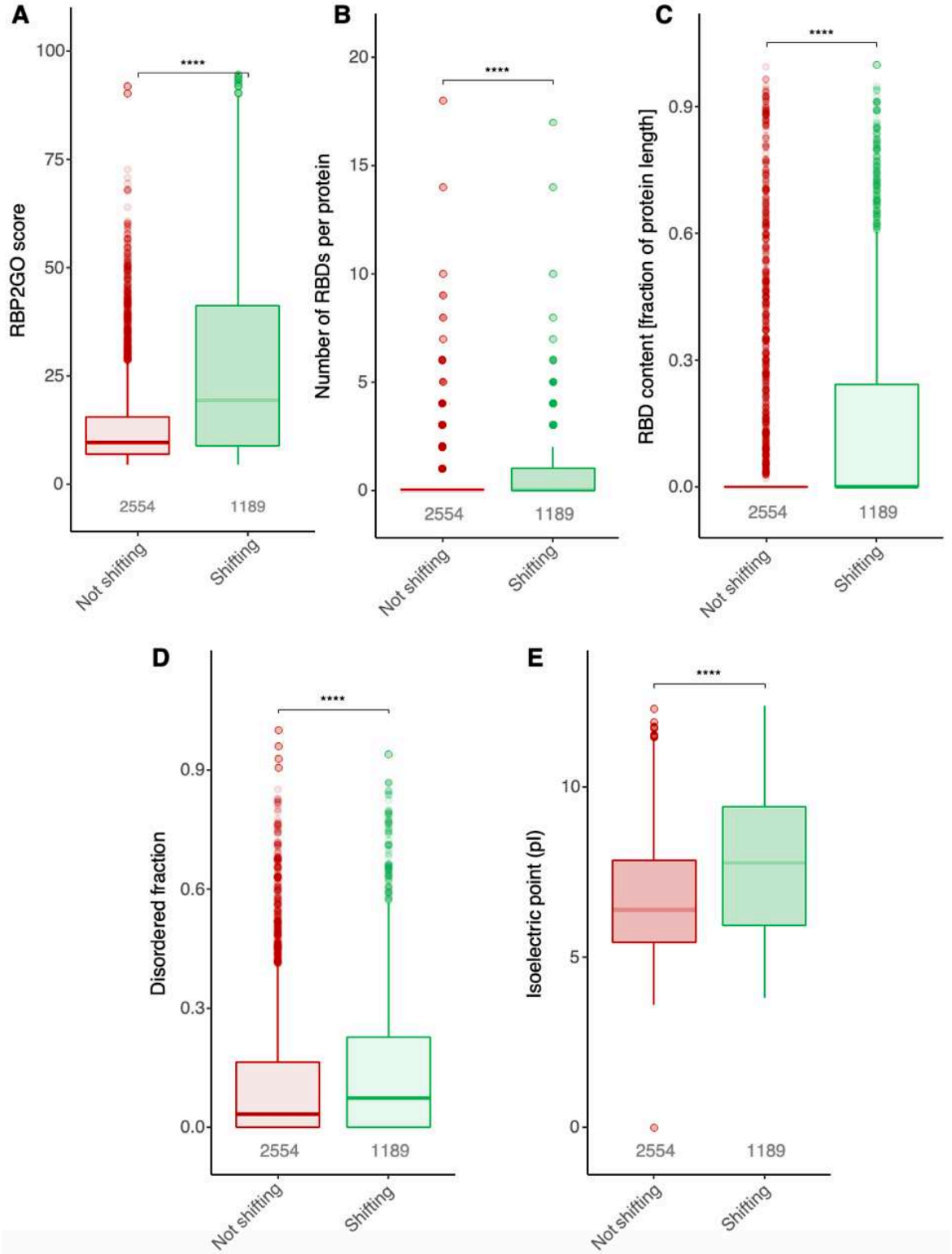


Figure 11: Properties of shifting proteins

(A) RBP2GO score in shifting vs non-shifting proteins. The RBP2GO score is calculated based on the frequency of a protein listed as an RBP in the 43 proteome-wide human studies.
(B) Graph indicating higher number of RBDs in shifting proteins compared to non-shifting proteins in the R-DeeP screen.

(C) Box plot showing significantly higher RBD content fraction relative to the length of the protein in shifting group of proteins compared to the non-shifting proteins.
(D) Shifting protein display a higher disordered region fraction non-shifting proteins.
*(E) Boxplot depicting higher isoelectric point for the shifting proteins when compared to the non-shifting proteins. All data are presented as boxplots with the median indicated by the bar and the box indicating the lower and upper quartiles. The outliers are represented by dots. The upper whisker extends from the hinge to the largest value not more than 1.5× interquartile range from the hinge and the lower whisker extends from the hinge to the smallest value at most 1.5× interquartile range of the hinge. Wilcoxon test was used to calculate the p-value, (*** $p < 0.001$, **** $p \leq 0.0001$). The figure was adapted from (63).*

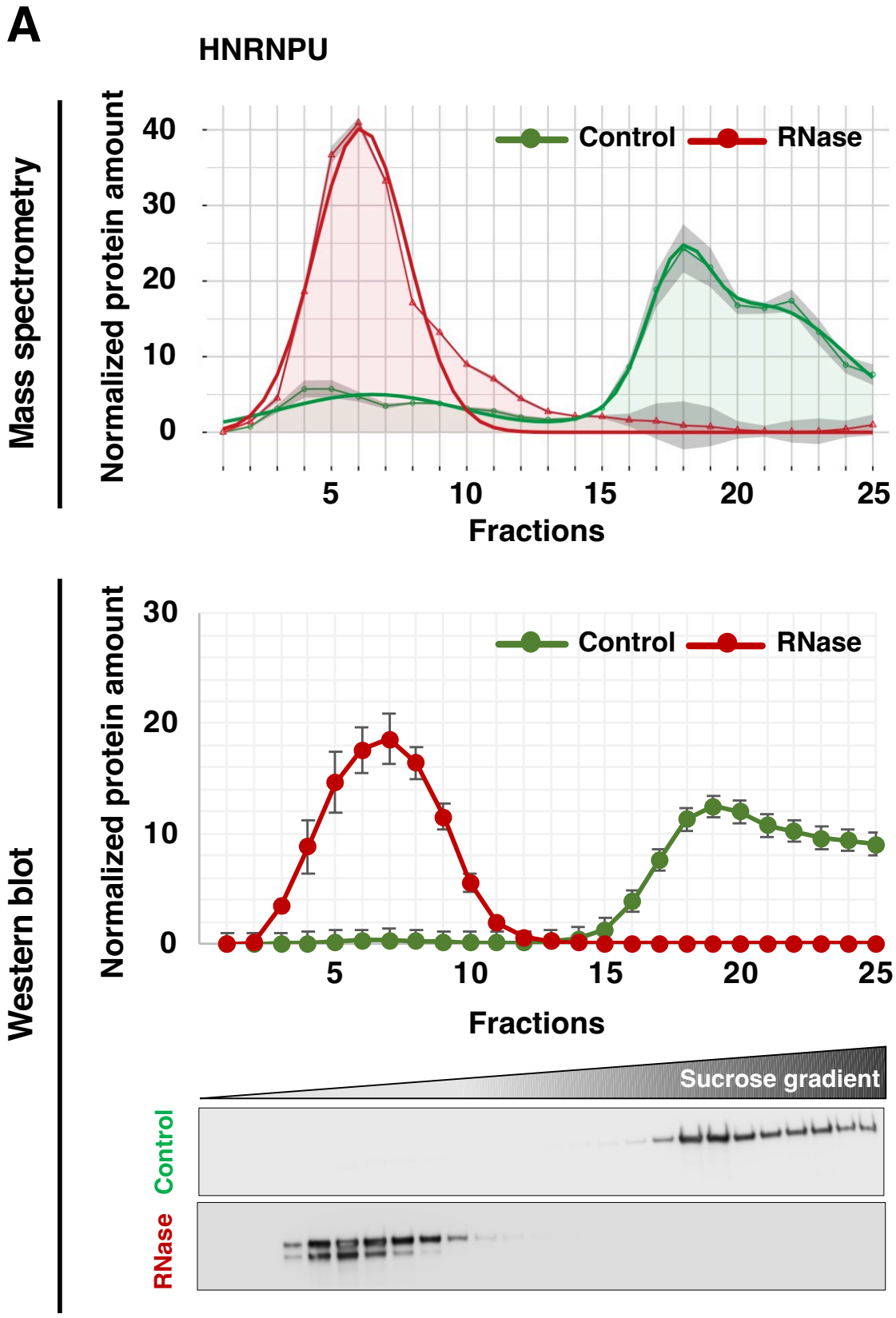
Finally, as the last parameter, the isoelectric points (pI) between the two groups of proteins were compared. RBPs are known to have relatively high positively charged amino acids in their sequence which aids in the RNA binding. Aptly, the shifting proteins contained significantly high isoelectric point (pI) compared to non-shifting proteins (Figure 11E). These parameters confirmed the high probability of shifting proteins for being a true RNA-binding protein.

4.1.3 Western blot validation of the R-DeeP screen

To validate the R-DeeP screen, all the 25 fractions from the control and RNase-treated gradients were loaded on the gel and SDS-PAGE, western blot was performed, followed by the quantification of individual bands using ImageJ software. After the quantification, the values were plotted as a graph which displayed the migration profile of the protein in comparison with mass spectrometry results (Figure 12).

First, the positive and negative controls were chosen: the well-known RBP-HNRNPU (heterogeneous nuclear ribonucleoprotein U), served as a positive control and ASNS (asparagine synthetase), served as a negative control. Further, the western blot was performed and the bands were quantified for the validation. The migration profile generated after the quantification of the western blot bands were comparable to the mass spectrometry dataset for these two proteins (Figure 12). The HNRNPU protein displayed a complete RNA-dependent shift towards the earlier fractions in the RNase-treated gradient compared to the control fractions as clearly indicated by the shift in peak protein amount (control: fractions 16-25 and RNase-treated: 3-9) (Figure 12A), whereas the negative control protein ASNS displayed an unaltered interactome profile in the control and the RNase-treated gradients (protein peak around the fractions: 4-9) as expected (Figure 12B). Additionally, previous Crosslinking and Immunoprecipitation (CLIP) experiments followed by radioactive labelling of RNA confirmed the RNA-binding property of the protein HNRNPU, and the lack of RNA binding by the protein

ASNS. These results were in line with the previously published CLIP data, further strengthening the robustness of this R-DeeP screen.



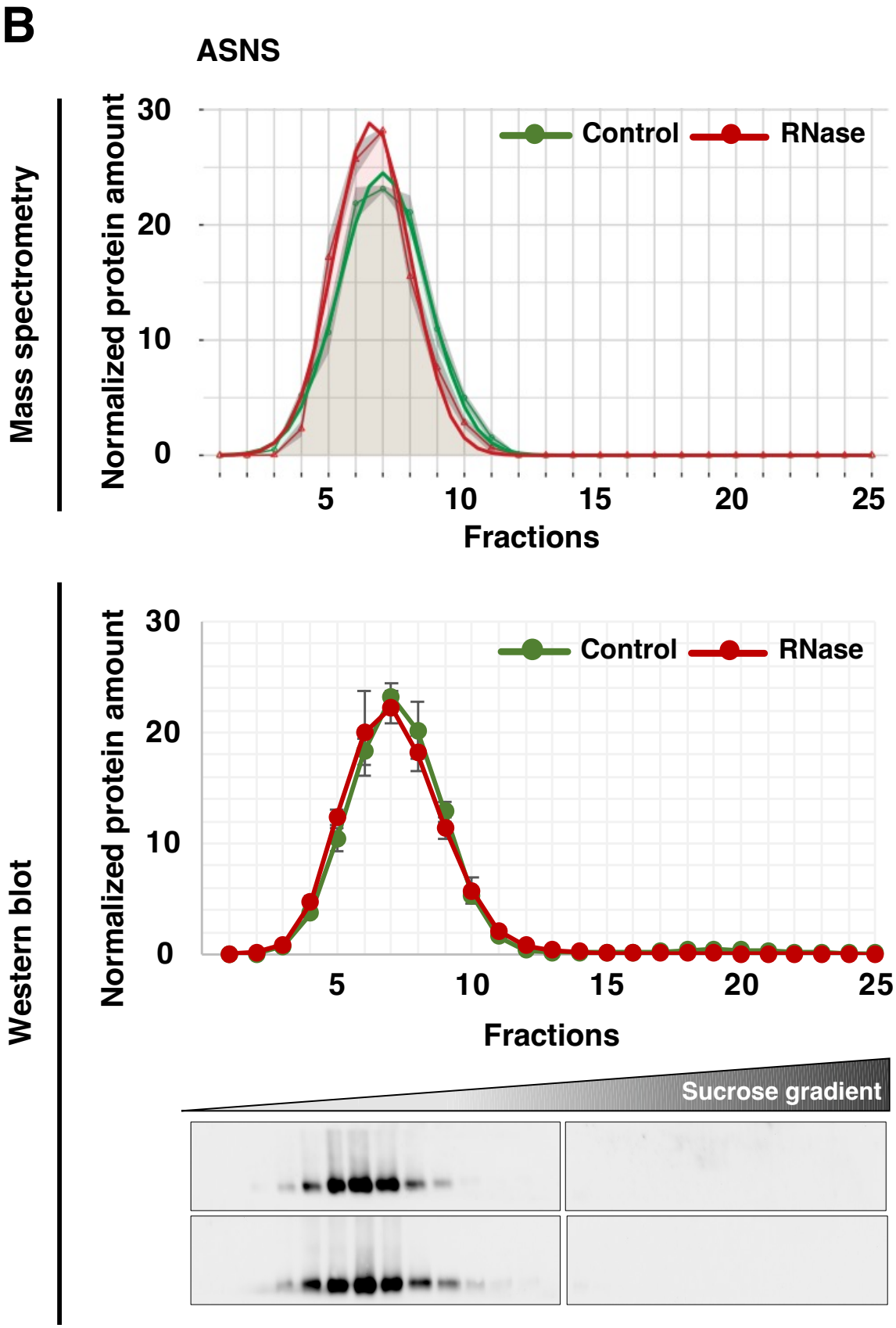

B
ASNS
Mass spectrometry
Normalized protein amount
30
20
10
0
Control
RNase
5
10
15
20
25
Fractions
Western blot
Normalized protein amount
30
20
10
0
Control
RNase
5
10
15
20
25
Fractions
Sucrose gradient

Figure 12: Western blot validation of the R-DeeP screen

(A) The top panel represents the mass spectrometry analysis of the known RBP HNRNPU (positive control) displaying the distribution of protein amount across the 25 fractions in the control and RNase-treated gradients. Raw data represented by the mean of three replicates is indicated by the curve with points and the Gaussian fit is represented by a smooth curve (green: Control, red: RNase-treated). The overall protein amount was normalized to 100. The middle panel depicts the quantification graph of the western blot bands represented by the mean and SEM for comparison to the mass spectrometry data. The bottom panel shows a representative example of the western blot analysis, carried out in triplicates. Displays the distribution of the protein HNRNPU across 25 fractions in control and RNase-treated gradients.
(B) Same as in 12A for the negative control protein ASNS.
The figure was adapted from(63).

In an effort to further validate and identify novel RNA-dependent proteins from the R-DeeP screen in A549 cells, proteins were additionally selected based on the following criteria, for the validation of the R-DeeP screen in A549 cells. (i) Novel RNA-dependent proteins *i.e.* the proteins that has not been identified as an RBP in the previous 43 proteome-wide studies in human, accessible from the RBP2GO database (56), (ii) strength of the shift (distance between the peaks in control and RNase-treated fractions), finally (iii) availability of the antibodies to perform western blot. Based on these three criteria, three proteins were selected for further validation using western blot analysis namely, Dedicator of cytokinesis protein 5(DOCK5), Engulfment and cell motility protein 2 (ELMO2) and BRCA1-A complex subunit protein ABRAXAS1 (CCDC98).

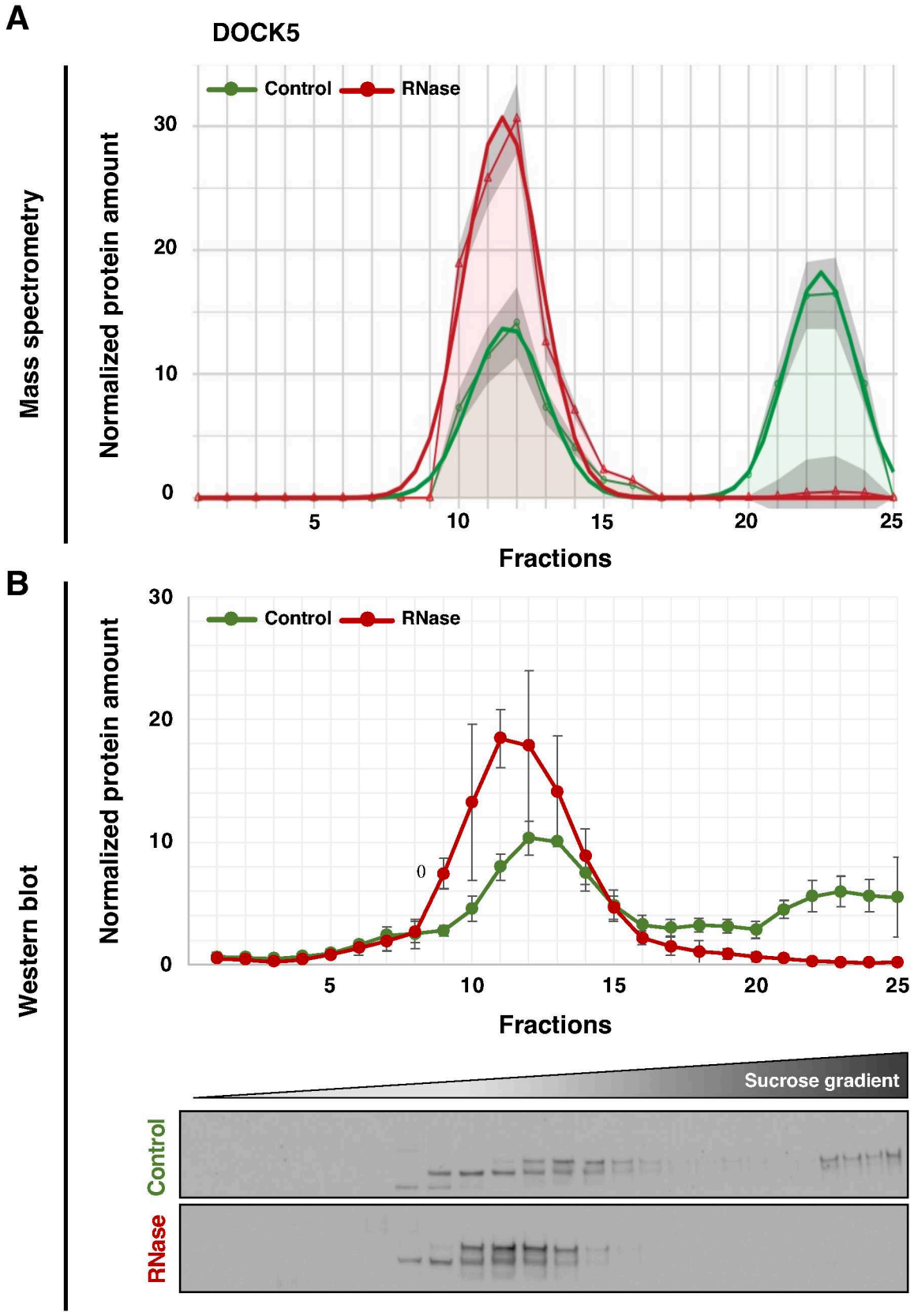
The first protein to be validated was the Dedicator of Cytokinesis 5 (DOCK5). It belongs to the DOCK family of proteins that are guanine nucleotide exchange factors (GEF) for proteins like Rac and Cdc42 GTPases (132,133). DOCK proteins are classified into 4 subgroups DOCK A-D that contain two evolutionarily conserved DOCK homology region (DHR): the lipid-binding DHR-1 and the catalytic DHR-2 domains. The DOCK family of proteins are comprised of 11 GEFs, out of which DOCK1, DOCK2 and DOCK5 belong to the DOCK A subfamily (132,133). DOCK5 is the least studied member of the DOCK family and the closest homologue of DOCK1. The interaction between these two proteins mediates cell migration and spreading (132,133). DOCK5 is known for its role in osteoclasts, cell migration, motility, invasion and murine embryonic development. Importantly, studies have shown that, inhibition of DOCK5 has reduced the invasiveness and tumour burden in mice injected with MDA-MB-231 breast cancer cells (132,133). Additionally, DOCK5 expression regulated by PHF5A has also been shown to promote head and neck squamous cell carcinoma (HNSCC) progression (134). Given the crucial role of DOCK5 in osteoclasts and other diseases like cancer, we analysed the RNA dependence of DOCK5 in A549 cells. The migration profile of DOCK5 clearly indicated its partial RNA dependence, as evident from the left shift of the protein amount. DOCK5 depicted

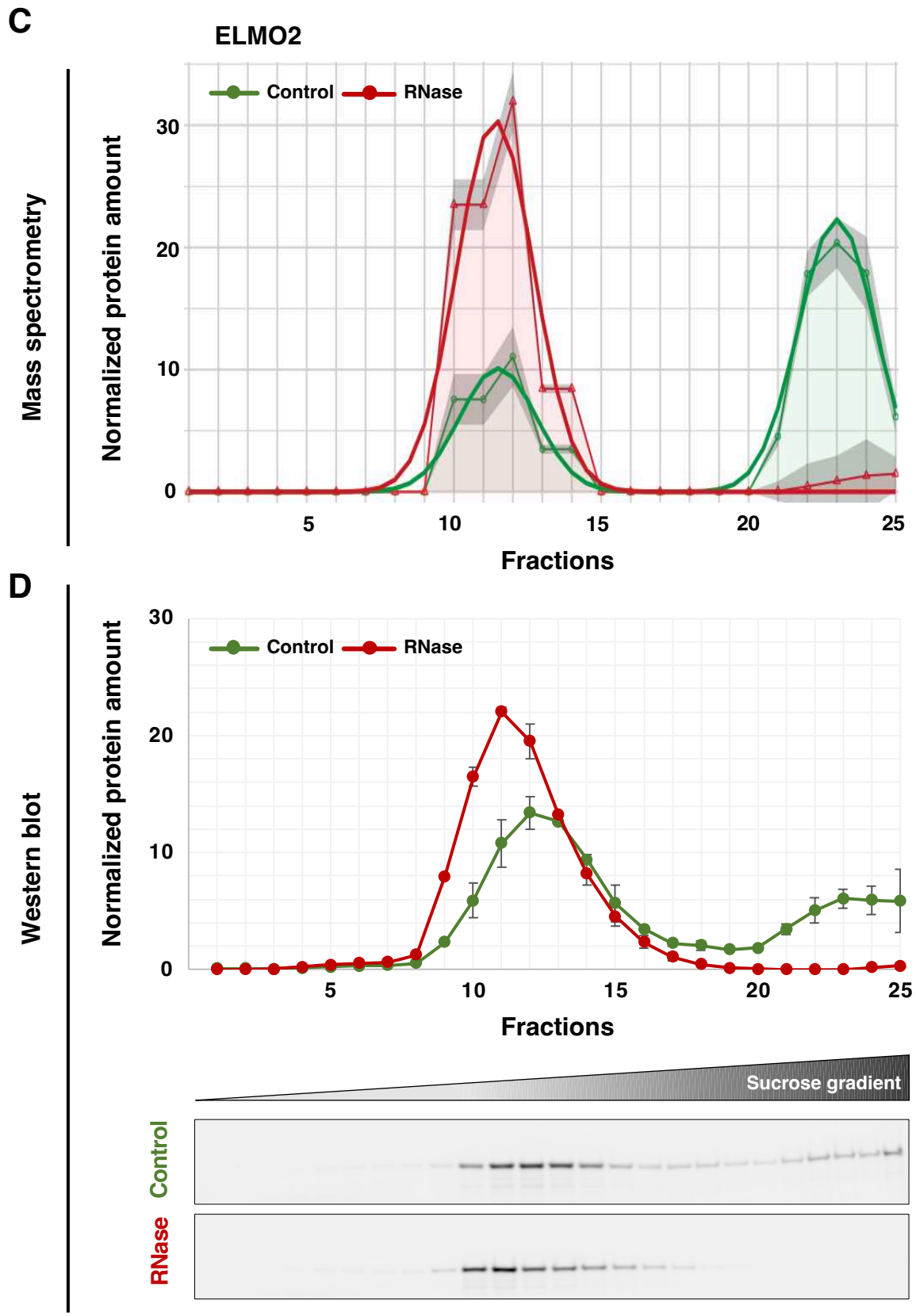
two populations. The first peak was observed around the fractions 9–14 and the second peak around the fractions 21–25. Upon RNase treatment, the protein was completely enriched in the earlier fractions (9–14) with no detectable signal in the later fractions (Figure 13A). The western blot quantification graph (Figure 13B) and the mass spectrometry analysis correlated very well with each other, validating that DOCK5 was RNA dependent.

The second validated protein was the Engulfment and Cell Motility protein 2 (ELMO2). It belongs to the evolutionarily conserved ELMO family of proteins consisting of three members ELMO1, ELMO2 and ELMO3 (135,136). These proteins are essential for the engulfment and motility of the cell and a well-known interactors of DOCK proteins. It interacts with DOCK proteins to activate Rac signalling pathways that initiates cytoskeleton remodelling which is in turn crucial for cellular processes such as cell migration, myoblast fusion or phagocytosis (135,136). Additionally, ELMO2 has been reported to be interacting with other proteins like Gαi2, Gβγ and Nck-1 that are involved in similar tasks (63,137). Importantly, ELMO2 has been reported to play an important role in chemotaxis, invasion and migration in pancreatic cancer, mediated by CXCL-2 (63,137). Although ELMO2 has not been extensively studied in cancer, in this project, I investigated the RNA dependence of ELMO2. As represented by the graphs and the western blot, ELMO2 shows a substantial left shift from the fractions 21–25 to the fractions 9–14 upon RNase treatment when compared to the control gradients (Figure 13C, D). This clearly proves that ELMO2 is partially RNA dependent (Figure 13C, D). Interestingly, both proteins (DOCK5 and ELMO2) were detected in similar fractions with similar migration profile in control and RNase-treated gradients indicating a potential interaction between these proteins.

Finally, ABRAXAS1/CCDC98 was validated. It is a part of the BRCA1-A complex involved in the DNA damage repair pathway (138,139). ABRAXAS1 acts as a scaffold protein for the BRCA1-A complex that consists of other proteins like MERIT40, RAP80, BRCC45 and BRCC36 (138). It contains a coiled-coil domain, regulates DNA damage checkpoint and DNA end resurrection in homologous recombination repair together with other proteins such as RAP80, BRE (138,139). It is a tumour suppressor gene in the BRCA pathway, which displays copy number loss or mutations in several types of cancer (138). Additionally, ABRAXAS1 depletion causes genome instability by impeding the recruitment of BRCA1 to DNA damage sites (138). Importantly, reduced ABRAXAS1 expression was implicated in many cancers and knockout mice exhibited decreased survival (138). Since, ABRAXAS1 is shown to be crucial for tumour suppression and better survival, I further investigated the RNA dependence of ABRAXAS in A549 cells. Similar to DOCK5 and ELMO2, ABRAXAS1 depicted two

populations with peaks around the fractions 8-13 and fractions 20-5 in the control gradient (Figure 13E, F). Upon RNase treatment, the entire protein amount shifted to the earlier fractions 8-13 (Figure 13E, F). This displayed a clear left shift and validated the mass spectrometry analysis. Thus, ABRAXAS1 is a partially RNA-dependent protein. Further, I compared the position of maxima for the reference proteins of known molecular weights (RNase A (14 kDa), BSA (60 kDa), Aldolase (160 kDa), Catalase (240 kDa), and Ferritin (480 kDa)) in the gradients from the R-DeeP screen in unsynchronized HeLa cells (59). Here, I observed that DOCK5 shifted to its monomeric size after RNase treatment, whereas ABRAXAS1 and ELMO2 persisted in a complex larger than their respective monomeric sizes.


A
DOCK5
Mass spectrometry
Control
RNase
Normalized protein amount
30
20
10
0
5
10
15
20
25
Fractions
B
Western blot
Normalized protein amount
Control
RNase
30
20
10
0
5
10
15
20
25
Fractions
Sucrose gradient
Control
RNase


C
ELMO2
Mass spectrometry
Control
RNase
Normalized protein amount
30
20
10
0
5
10
15
20
25
Fractions
D
Western blot
Normalized protein amount
Control
RNase
30
20
10
0
5
10
15
20
25
Fractions
Sucrose gradient
Control
RNase

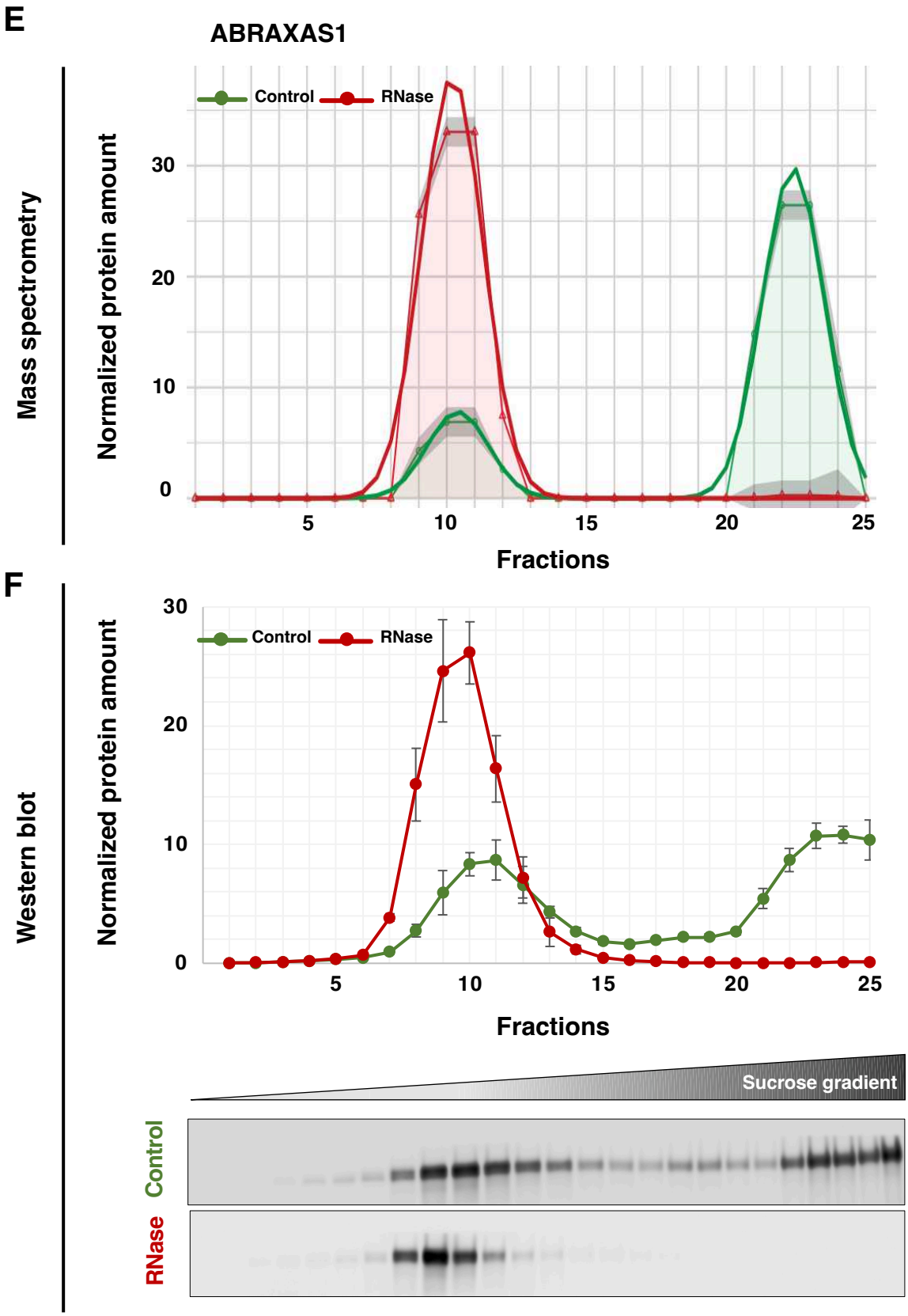

E
ABRAXAS1
Mass spectrometry
Normalized protein amount
Control
RNase
30
20
10
0
5
10
15
20
25
Fractions
F
Western blot
Normalized protein amount
Control
RNase
30
20
10
0
5
10
15
20
25
Fractions
Sucrose gradient
Control
RNase

Figure 13: Validating the RNA dependence of DOCK5, ELMO2 and ABRAXAS1 using western blot analysis

(A) Graph representing the protein amounts of DOCK5 distributed in control and RNase-treated gradients, as determined by mass spectrometry analysis in all 25 fractions. The experiments were performed in three replicates (n=3). The pointed curve represents the mean of the raw data (total protein amount normalized to 100), the grey shadows represent the standard deviation and the Gaussian fit is represented by the smooth curve.
(B) Top panel: Quantitative analysis of the western blot, represented by the mean of three replicates and SEM across 25 fractions. The green line represents the control gradients and the red line represents the RNase-treated gradients. The overall protein amount was normalized to 100 to simplify the comparison with the mass spectrometry analysis. Bottom panel: Representative image of DOCK5 (215 kDa) western blot in control and RNase-treated A549 gradients. The western blot validation was performed for all three replicates of the gradient (n=3).
(C-F) Similar to (A) top panel and (B) bottom panel for ELMO2 (83 kDa) and ABRAXAS1 (47 kDa) (n=3).
The figure was adapted from (63).

4.1.4 Direct RNA binding of the RNA-dependent proteins

In the previous section, it has been clearly demonstrated that the three shifting proteins DOCK5, ELMO2 and ABRAXAS1 were RNA dependent. This led me to the next question, are these RNA-dependent proteins RBPs? Analysis on the properties of the shifting proteins in terms of RBP2GO score, number of RBDs, IDR content and isoelectric point, correlated well with the properties of RBPs. Hence, I validated the ability of DOCK5 and ABRAXAS1 to directly bind to RNA using iCLIP2 assay (124).I was unable to verify the direct RNA binding of ELMO2 due to lack of appropriate antibody to immunoprecipitate ELMO2 from the A549 cell lysate. First, the UV-crosslinked and the non-crosslinked A549 cells were lysed and the lysates were treated with varying amounts of RNase 1 dilution ranging from 1:5 to 1:500. Later, the proteins of interest DOCK5 and ABRAXAS1 were immunoprecipitated and the co-purified RNA was labelled with [γ-32P] ATP. Finally, autoradiography revealed the direct binding of the immunoprecipitated protein to RNA. Here, IgG and non-crosslinked (NCL) samples were used as negative controls. Since the lysates were treated with varying concentrations of RNase, the length of the RNA bound to the protein increased with decrease in RNase 1 concentration. Consequently, the size of the complex shifted upwards on the blot with decreasing amounts of RNase due to decreasing partial degradation of the RNA. This method is a standard to verify the identity of the signal as RNA (Figure 14A, C). At the highest RNase 1 concentration (1:5 dilution), the RNA signal, was most intense at the height of the protein (DOCK5 or ABRAXAS1 respectively, (Figure 14), while a dispersed smear towards higher molecular weights was observed at lower RNase concentrations (for increasing dilutions: 1:50 and 1:500), ascertaining that both DOCK5 and ABRAXAS1 are RBPs (Figure 14A, C). The proper immunoprecipitation of DOCK5 and ABRAXAS1 was validated by western blotting as a control to demonstrate equal protein amounts in all the samples (Figure 14B, D).

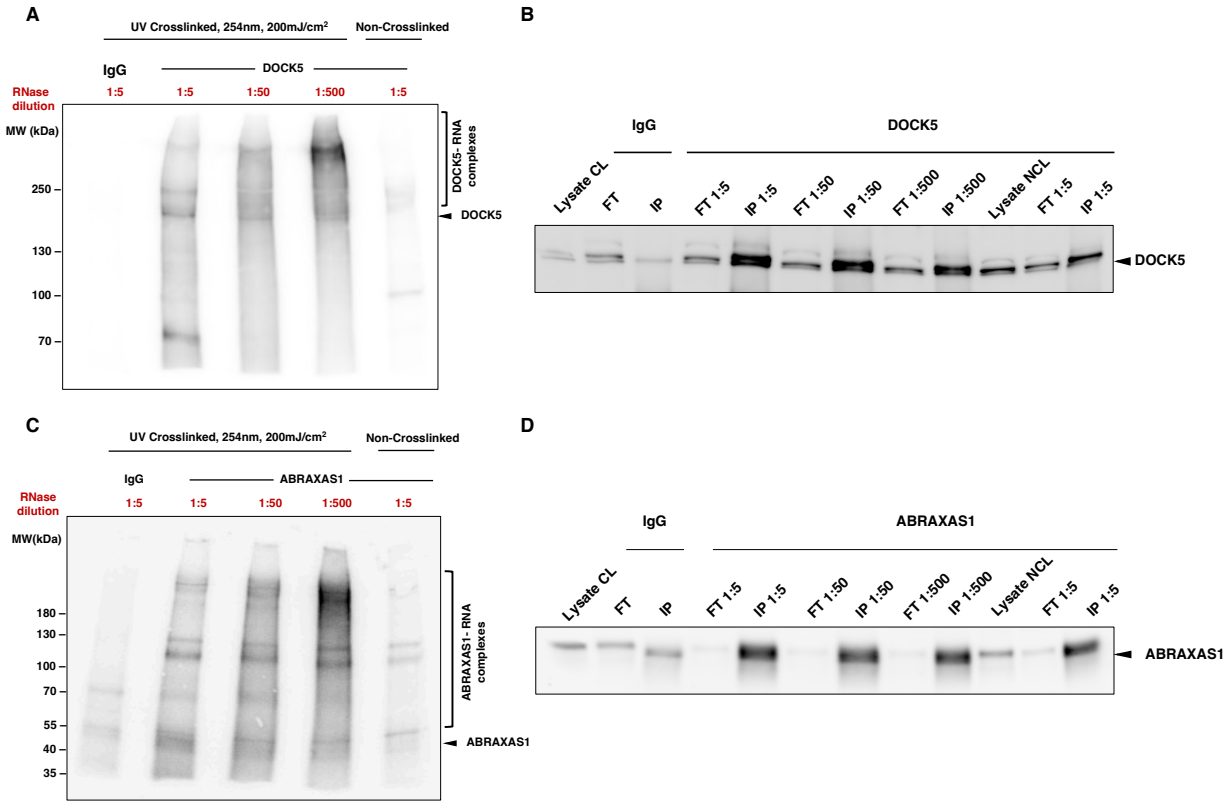


Figure 14: Direct RNA binding of the DOCK5 and ABRAXAS1

(A) Autoradiography presentation the direct binding of DOCK5 to RNA by iCLIP2 assay, as indicated by the shift in radioactive signal toward higher molecular weights with decreasing RNase I concentrations (n = 3).
(B) Western blot validating the equal amount of DOCK5 in samples treated with different RNase dilutions from 1:5 to 1:500. 2.5% of the total lysate from the crosslinked and non-crosslinked samples and 2.5% of the flowthrough (FT) were loaded on the gel to compare the IP efficiency and size of the pulled down protein (n = 3).
(C, D) Autoradiography and western blot images of ABRAXAS1 iCLIP2 assay (n=3) proving the direct RNA binding of the protein. Same as (A, B). The figure was adapted from (63).

Finally, the entire dataset from the A549 R-DeeP screen together with the data from the previous R-DeeP screen in unsynchronized HeLa S3 cells were compiled into the R-DeeP 2.0 database https://R-DeeP2.dkfz.de that can be freely accessed (Figure 15). The database contains the analysed proteome-wide mass spectrometry data of proteins after sucrose density gradient fractionation in presence of RNA molecules and after RNase treatment in both cell lines. It offers various search options and detailed information on the protein including the graphical representation of the migration profiles in a downloadable format. The user can make use of the single or advanced search options to look into proteins in a single cell line or to compare a protein directly between the two cell lines A549 and HeLa S3. The R-DeeP 2.0 analysis results are available for download which includes the results of statistical quantitative analysis, the maxima, their position and the amount of protein for the control and RNase gradients, parameters of the shifts and additional information about the protein. The database

work was completely performed by my colleague Niklas Engel, who was a master student in our lab (Figure 15).

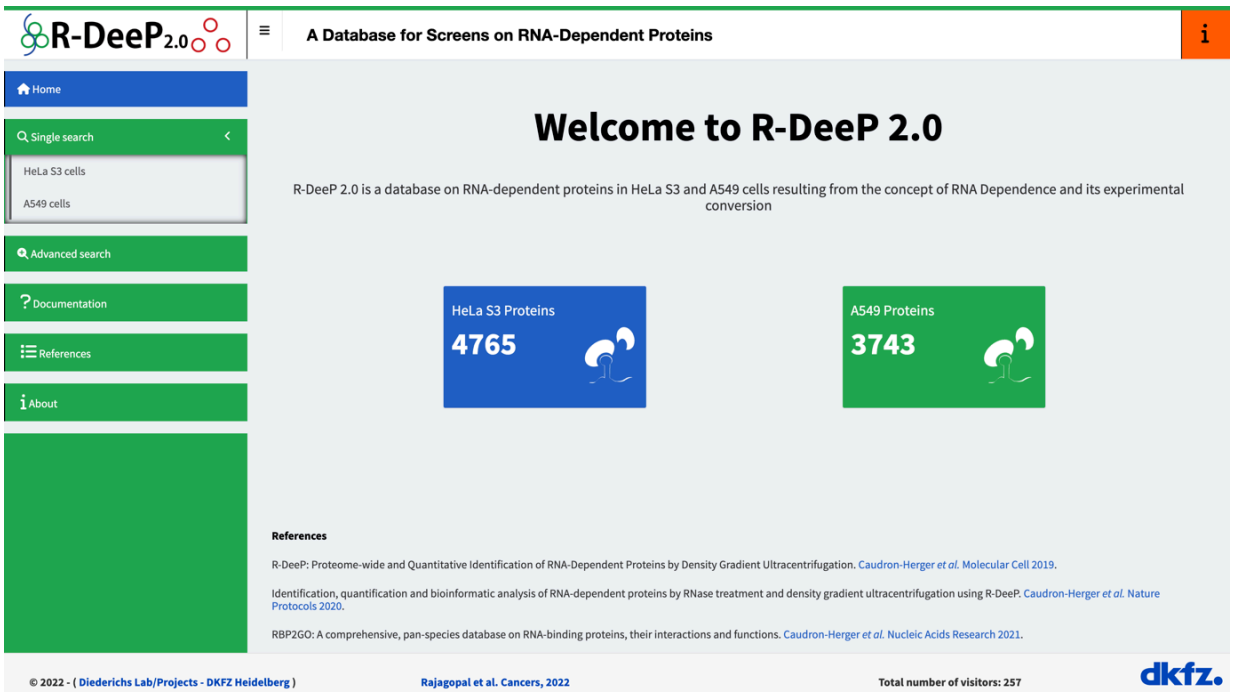


Figure 15: R-DeeP 2.0 Database

The R-DeeP 2.0 database provides information on the R-DeeP results on 4765 detected proteins in HeLa S3 cells and 3743 proteins in A549 cells, with multiple search and download options and provides information about each protein in terms of peaks and shifts in one or both cell lines. Further external resources which give information about the interactors and protein complexes are linked to the database for more detailed information on the protein. Here, a screenshot of the homepage of the database is shown. Adapted from (63)

Altogether, in the first part of my PhD project, I identified 170 novel RNA-dependent proteins in A549 cells using a proteome-wide R-DeeP screen, which provides an overview of the ribonucleoprotein complexes present in A549 lung cancer cells. Additionally, I confirmed that DOCK5 and ABRAXAS1 directly bind to RNA in these cells, demonstrating that, these proteins function not only in an RNA-dependent manner but also as RBPs. The entire dataset is presented in a user-friendly R-DeeP 2.0 database, that provides detailed information about RBPs in A549 along with data from HeLa cells to the scientific community.

This dataset will further help to understand and dissect the role of RNA and RNA-dependent proteins in lung cancer. In future, this data could be further utilised to investigate the difference between the RBPome of healthy cells and lung cancer cells, which would shed light on the changes in the interactome of proteins leading to cancer progression. The entire data was published in the peer-reviewed journal *Cancers* in 2022 (63).

4.2 Part II: RNA-dependent functions of AURKA in mitosis

Gene Ontology analysis (GO) of the shifting proteins from the first R-DeeP screen in HeLa unsynchronised cells revealed the significant enrichment of proteins with GO terms related to mitosis (Figure 16). The GO analysis compared the fold enrichment of shifting proteins to the human proteome and provided a list of GO terms with significant adjusted p-value (< 0.05) using Fisher's exact test. Due to variations in the expression level of mitotic factors in an unsynchronized cell population, the R-DeeP approach was established in HeLa cells synchronized in mitosis and in interphase, to obtain a more comprehensive landscape of the RNA-dependent proteins in a cell cycle-dependent manner. The screen detected 1751 shifting proteins in mitosis and 1912 shifting proteins in interphase. Out of the identified shifting proteins, 389 novel RNA-dependent proteins were detected in mitosis and 426 novel RNA-dependent proteins were detected in interphase (116). These novel RNA-dependent proteins were not identified as RBPs in the previous proteome-wide studies in human (25,56,116).

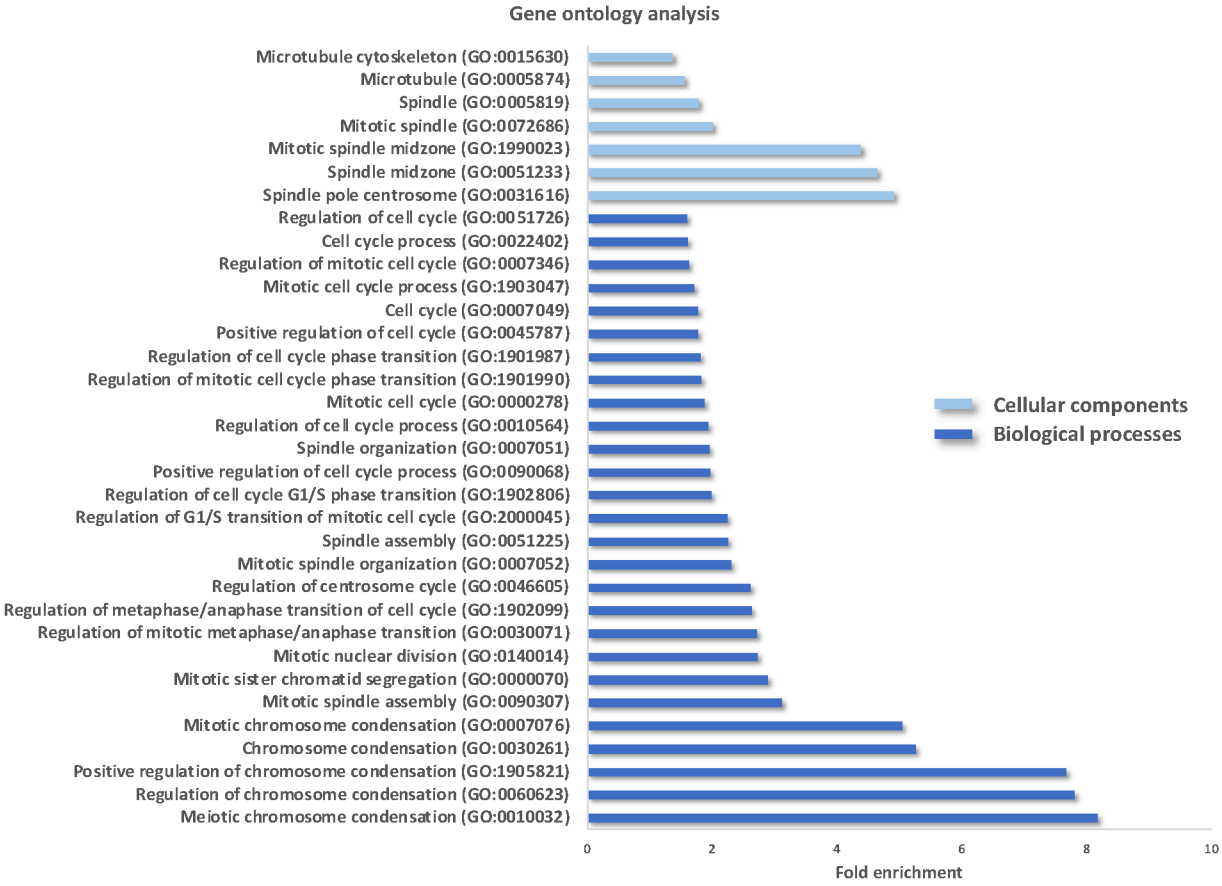


Figure 16: RNA-dependent proteins are significantly enriched in mitotic factors

The RNA-dependent protein list from the R-DeeP screen in unsynchronised HeLa cells (59) was used to perform a GO analysis using the PANTHER classification tool available online (119). 2-fold enrichment was observed for the mitosis-related terms like cell cycle regulation, spindle, or microtubule. Bars in light blue indicate GO analysis

under "cellular components" terms and bars in dark blue represent GO analysis under "biological processes" terms (116). Fisher's exact test was used to calculate the p-values, with further correction for multiple testing based on false discovery rate (FDR). Adjusted p-values (< 0.05) were considered significant. The figure was adapted from (116).

Upon further analysis, AURKA emerged as the top candidates that showed reproducible R-DeeP profiles in both R-DeeP screens (HeLa unsynchronised and HeLa mitosis vs interphase) with clear and strong left shift in RNase-treated gradients (116). In addition, AURKA belongs to the family of serine/threonine kinases that consists for two other kinases AURKB and Aurora kinase C, that are crucial for cell cycle control (140,141), making it a very interesting candidate. AURKA was first identified in *Drosophila*, in a screen to identify gene involved in mitotic spindle functions. Mutations in these genes resulted in severe mitotic abnormalities like defects in centrosome separation, spindle organization failure and formation of monopolar spindles. Hence, they were termed as Aurora kinases, reminiscent of North pole (141-143).

AURKA contains a regulatory domain in the NH2 terminal and catalytic domain in its COOH terminal part. The kinase domain responsible for the functions of AURKA is situated close to the COOH terminus around the amino acid residues 133-383, with its activation segment in amino acids residues 280-293 (141). Importantly, auto-phosphorylation at T288 is crucial for the activation of its kinase activity, that is required for mitotic progression. AURKA is mainly enriched at the centrosomes and the spindle, suggesting its dual role in mitotic spindle assembly (141,142).

The first and well-studied cofactor of AURKA is Targeting protein for Xklp2 (TPX2), which is a key player in spindle assembly. In the mitotic cytoplasm, TPX2 and several other mitotic factors are inhibited through their interaction with the importin α/β heterodimer (144-146). Upon the establishment of the RanGTP gradient around the mitotic chromosomes, TPX2 is released from the importins, promoting the binding of TPX2 to AURKA as well as releasing TPX2 functions in MT nucleation (144-147). The interaction between AURKA and TPX2 is mediated through the conserved NH2-terminal domain of AURKA, which induces a conformational change, moving the activation segment of AURKA inside the catalytic pocket of the kinase (144-147). Thereby, the crucial phospho-threonine (T288) in the activation segment is protected by TPX2 against the activity of phosphatase 1 (PP1)-dependent dephosphorylation and inactivation of AURKA. Further, TPX2-AURKA oligomerize and bind to other interaction partners such as hyaluronan mediated motility receptor (RHAMM), gamma-tubulin ring complex (γ-TuRC) and neural precursor cell expressed developmentally down-regulated protein 1 (NEDD1) (144,145). Later, AURKA phosphorylates several

substrates for their activation e.g. the adaptor subunit of γ-TuRC at its conserved serine residues, which then initiates MT nucleation, stabilization and bundling near the mitotic chromosomes (144,145). In addition to TPX2, NEDD1 and RHAMM, AURKA is also known to interact with several other spindle assembly factors like the nuclear mitotic apparatus protein (NuMA), aurora Borealis (BORA) and transforming acidic coiled-coil containing protein 3 (TACC3) for their activation and the progression through mitosis. Importantly, TPX2 depletion results in the abrogation of AURKA localization at the spindle, highlighting the importance of their interaction (144). Interestingly, depletion of TPX2 does not affect the localization of AURKA to the centrosomes suggesting different mechanism of action and interaction partners of AURKA at these two distinct structures (144,145).

A study in *Xenopus* egg extracts identified a new protein, centrosomal Protein 192 (CEP192), as a major cofactor of AURKA in centrosomes, which helped in deciphering the role of AURKA at centrosomes (144,148). Subsequent studies demonstrated that CEP192 recruits AURKA to the centrosomes leading to its activation (144,149). CEP192 is constitutively bound to a fraction of AURKA, and recruits it to the centrosomes during late G2 and mitosis. CEP192 also recruits polo-like kinase 1 (PLK1) to the centrosome. AURKA is first activated by autophosphorylation at T288 followed by phosphorylation of PLK1 at the conserved T210 amino acid residue by AURKA, which mediates the docking of PLK1 onto CEP192 at T44. This initiates a multistep signalling cascade that drives centrosome maturation, separation and aster formation (144,148,150).

Apart from the role of AURKA as a mitotic regulator, many studies have highted its role as an oncogene in tumorigenesis in multiple types of cancer including haematological malignancies and solid tumours when overexpressed. It leads to cancer cell proliferation, epithelial-mesenchymal transition (EMT), metastasis, apoptosis, and self-renewal of cancer stem cells (141,142,151,152). Also, substrates of AURKA such as yes-associated protein (YAP) and Twist are involved in crucial oncogenic signaling for tumour progression (152-154). Given the oncogenic role of AURKA, many small molecule inhibitors were produced over the past decades, with nearly fifty clinical trials testing the specific AURKA inhibitors (152). Also, activation of AURKA has proven to induce resistance towards third generation EGFR inhibitors in lung cancer, which may lead to tumour heterogeneity and the generation of distinct clones that serves as a driving force for drug resistance (152,155). Hence, to obtain the effective therapeutic response, it is essential to know the pathways, proteins and molecular interactors involved in AURKA-mediated oncogenic function (152).

Though the R-DeeP screen in unsynchronized HeLa cells demonstrated a clear RNA dependence of AURKA, none of the other studies have investigated its RNA-binding role. Additionally, due to the absence of canonical RBDs, AURKA has become an interesting candidate of choice to understand the role of unconventional RBPs in biological pathways such as mitosis and its role in tumorigenesis. Therefore, in this project, I focused on AURKA and its RNA-dependent functions in mitosis.

4.2.1 AURKA is an RNA-binding protein

First, the RNA dependence of AURKA was analysed using the R-DeeP gradients from HeLa cells synchronised in mitosis using double thymidine block and nocodazole treatment. The control and RNase-treated gradient fractions were loaded on a SDS-PAGE and western blot analysis was performed. Upon quantification of the western blot band intensities, good agreement between the mass spectrometry analysis and western blot validation were observed. In the control gradients, the peak of AURKA protein amount was observed around fractions 3-7 and 18-23. Upon RNase treatment, the entire protein amount shifted (left shift) to the earlier fractions 3-7, demonstrating a strong and significant partial RNA dependence. This proved the RNA dependence of AURKA in mitosis (Figure 17A-C).

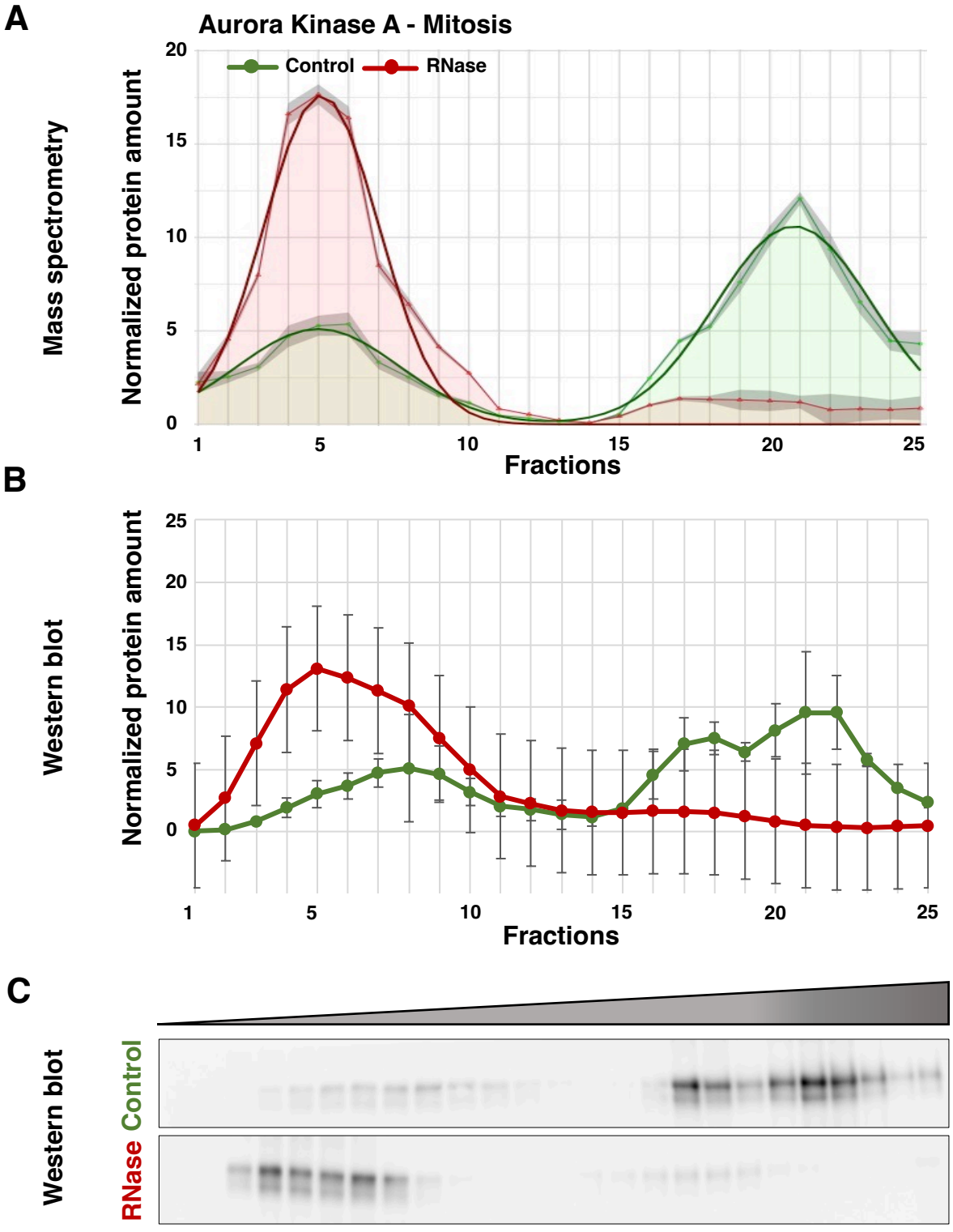


Figure 17: AURKA is an RNA-dependent protein in mitosis

(A) Graphical representation of the mass spectrometry analysis showcasing the protein amount in 25 different fractions of control (green) and RNase-treated (red) sucrose density gradients for AURKA. Lines with markers are depicted by raw data (mean of three replicates). Smooth lines represent the respective Gaussian fit. The overall protein amount of the raw data was normalized to 100.

(B) Graph showing the quantitative analysis of western blot replicates depicted by the mean of three replicates with SEM (n=3).
(C) Representative western blot depicting the distribution of AURKA in 25 different fractions in control and RNase-treated mitotic gradients.
The figure was adapted from (116).

Since the RNA dependence of AURKA was validated, the next question that raised was if AURKA could directly bind to RNA. In other words, is AURKA an RBP? To address this question, an iCLIP2 assay was performed in HeLa and A549 cells synchronised in prometaphase. The cells were lysed, treated with varying concentrations of RNase I and the lysate was cleared by centrifugation. Further, AURKA was pulled down using anti-AURKA antibody, and the bound RNA was labelled using [γ-32P] ATP. Finally, autoradiography revealed the RNA binding of AURKA (Figure 18A, C). Here, IgG and non-crosslinked (NCL) samples were used as negative controls. Similar to Figure 14, at highest RNase concentration, the signal was observed at the height of the protein (46 kDa). With decrease in the RNase concentration, an increasingly dispersed smear was observed towards higher molecular weight further substantiating that AURKA is a true RBP (Figure 18A, C). The immunoprecipitation of AURKA was verified by western blotting as a control to demonstrate equal protein amounts in all the samples (Figure 18B, D).

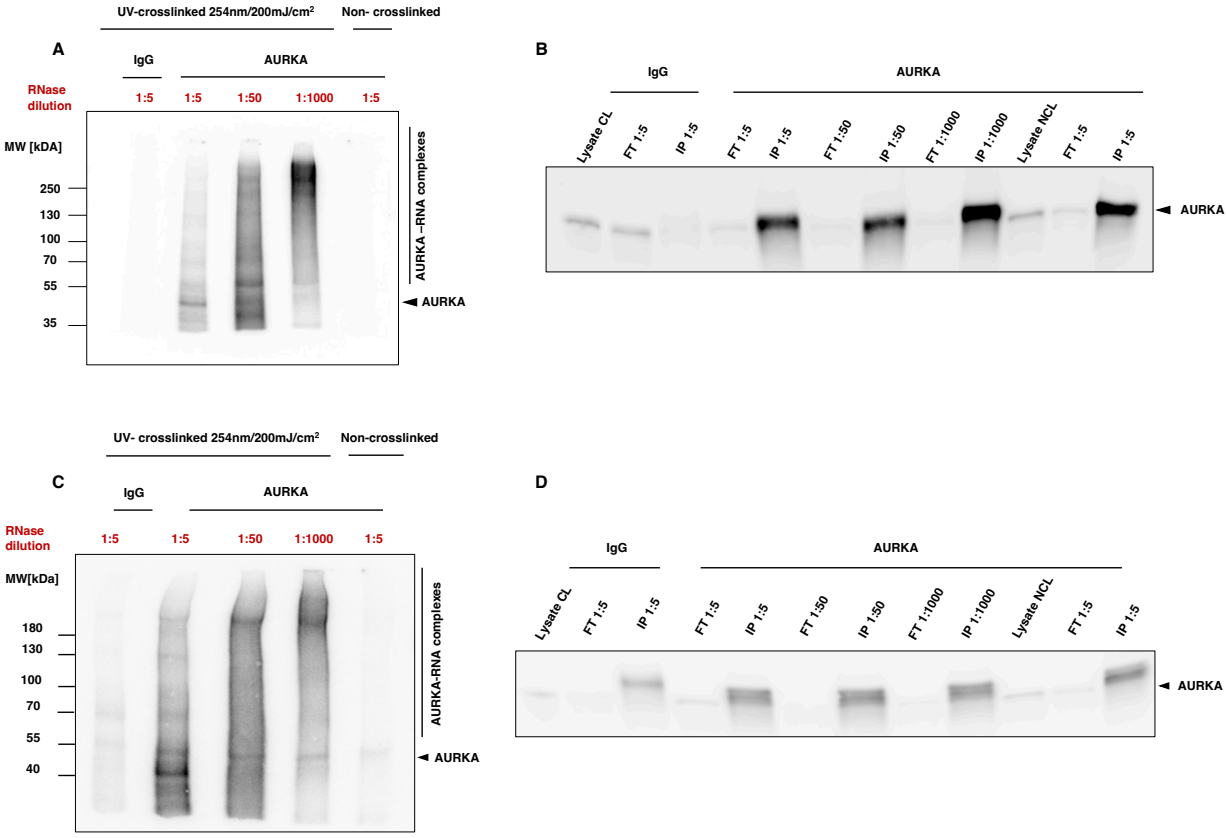


Figure 18: AURKA is an RNA-binding protein

(A) Autoradiography presentation the direct binding of AURKA to RNA (46 kDa) by iCLIP2 assay in HeLa prometaphase cells, as indicated by the shift in radioactive signal toward higher molecular weights with decreasing RNase I concentrations (n = 3).
(B) Western blot validating the equal amount of AURKA in all the samples. 2.5% of FT and total lysate from the CL and NCL samples were loaded on the gel for validating the immunoprecipitation efficiency and size confirmation (n = 3).
(C, D) Autoradiography and western blot images of AURKA iCLIP2 assay (n=3) in A549 prometaphase cells proving the direct binding of AURKA to RNA. Same as (A, B).
The figure was partially adapted from (116).

4.2.2 RNA-dependent interactors of AURKA

AURKA is a well-known mitotic regulator which has been thoroughly characterized and studied for decades. Likewise, most of the protein-protein interactors of AURKA is known. However, the RNA-binding property of AURKA is a novel discovery made during my PhD. Hence, as a next step, I focussed on identifying proteins interactors of AURKA, that were RNA mediated *i.e.* the protein-protein interactions mediated by RNA and the interactions with AURKA which are lost upon RNase digestion in mitosis. To this aim, cell lysates prepared from HeLa cells synchronised in prometaphase were used. These lysates were centrifuged, precleared and used for immunoprecipitation (see methods section for further details). Next, the lysates were incubated with AURKA-specific antibodies for the pulldown of AURKA-

protein complexes using magnetic beads. After washes, the isolated complexes were either treated with or without RNase, and finally the samples were eluted from the beads and analysed using LC-MS/MS-based analysis. Here, IgG was used as a negative control. From the mass spectrometry analysis, I observed that 86% of the protein interactors of AURKA were RNA dependent (Figure 19A). After the GO analysis using RBP2GO database, on mitosis related termed, we again observed a high enrichment (87%) for RNase-sensitive interactors in mitosis (Figure 19B). Furthermore, several new uncharacterized protein interactors of AURKA such as the major mitotic factors KIFC1 and INCENP were identified (Figure 19C). Most interestingly, we observed that AURKA interacted with TPX2, HMMR and CLASP1 in an RNA-dependent manner – few of the well-characterized AURKA interactors in mitosis (Figure 19C). TPX2 is a MT-associated protein that is crucial for AURKA activation and spindle assembly as described in the previous section (147). KIFC1 (alternatively termed HSET and also XCTK1 in *Xenopus*) is a minus-end directed motor protein. Both TPX2 and KIFC1 are localised to the nucleus during interphase (156-160). Upon nuclear membrane breakdown and the establishment of a RanGTP gradient around the chromosomes, KIFC1 is released from its inhibitory interaction with importins, as TPX2 is (158,161,162). Furthermore, KIFC1 is involved in spindle organization, MT focusing at the spindle poles and more generally MT crosslinking. KIFC1 recognises and crosslinks antiparallel MTs and slides the parallel MTs arising from the same centrosomes (163). Also, it is essential for cells that go through acentrosomal cell division. Centrosome serves as a major MT organizing centres (MTOCs) which mediate bipolar cell division. In cells that lack centrosomes such as the meiotic cells, KIFC1 compensates for the lack of centrosomes by clustering and focussing the minus ends of the spindles into bipolar structures (163,164). Similarly, it mediates bipolar spindle division by clustering supernumerary centrosomes in several cancer types with high centrosome amplification (163-167). At the same time, it transports cargo proteins such as NuMa across the spindle arm to the spindle poles (168). A recent study demonstrated that in cancer cells with high centrosome amplification, the interaction of KIFC1 and CEP215 was crucial for mediating pseudo bipolar spindle for viable cell division (166). In addition to the role of AURKA, TPX2 and KIFC1 in mitosis, all the three proteins are frequently over expressed in cancer and the co-overexpression of these proteins is strongly linked to genome instability and cancer progression (164,167,169-171). Given the important roles of these mitotic factors and their high relevance in cancer, I focussed on characterizing the RNA-dependent interactions between AURKA and

KIFC1 (a new uncharacterized interaction) and between AURKA and TPX2 (which RNA-dependent interaction had never been characterized before).

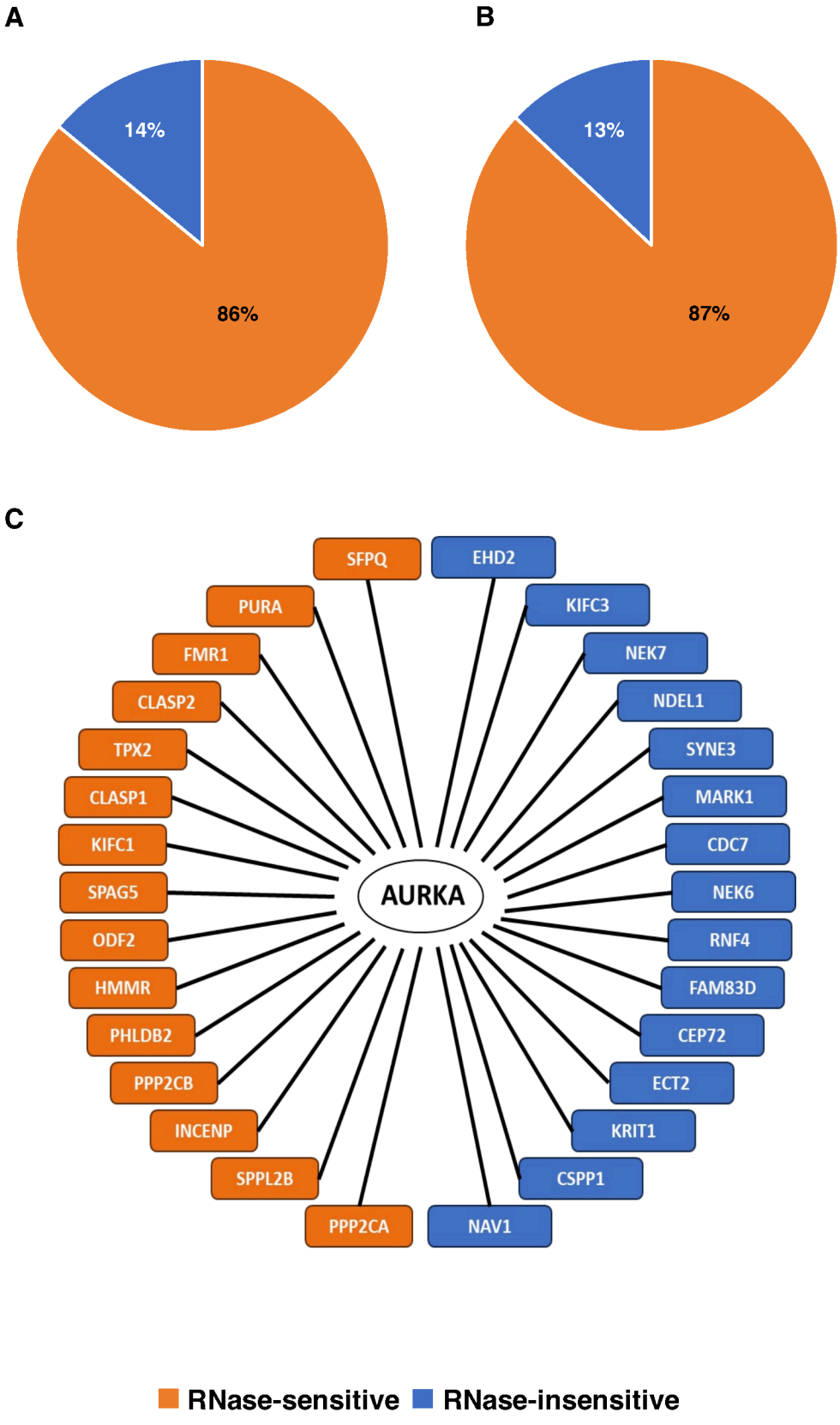
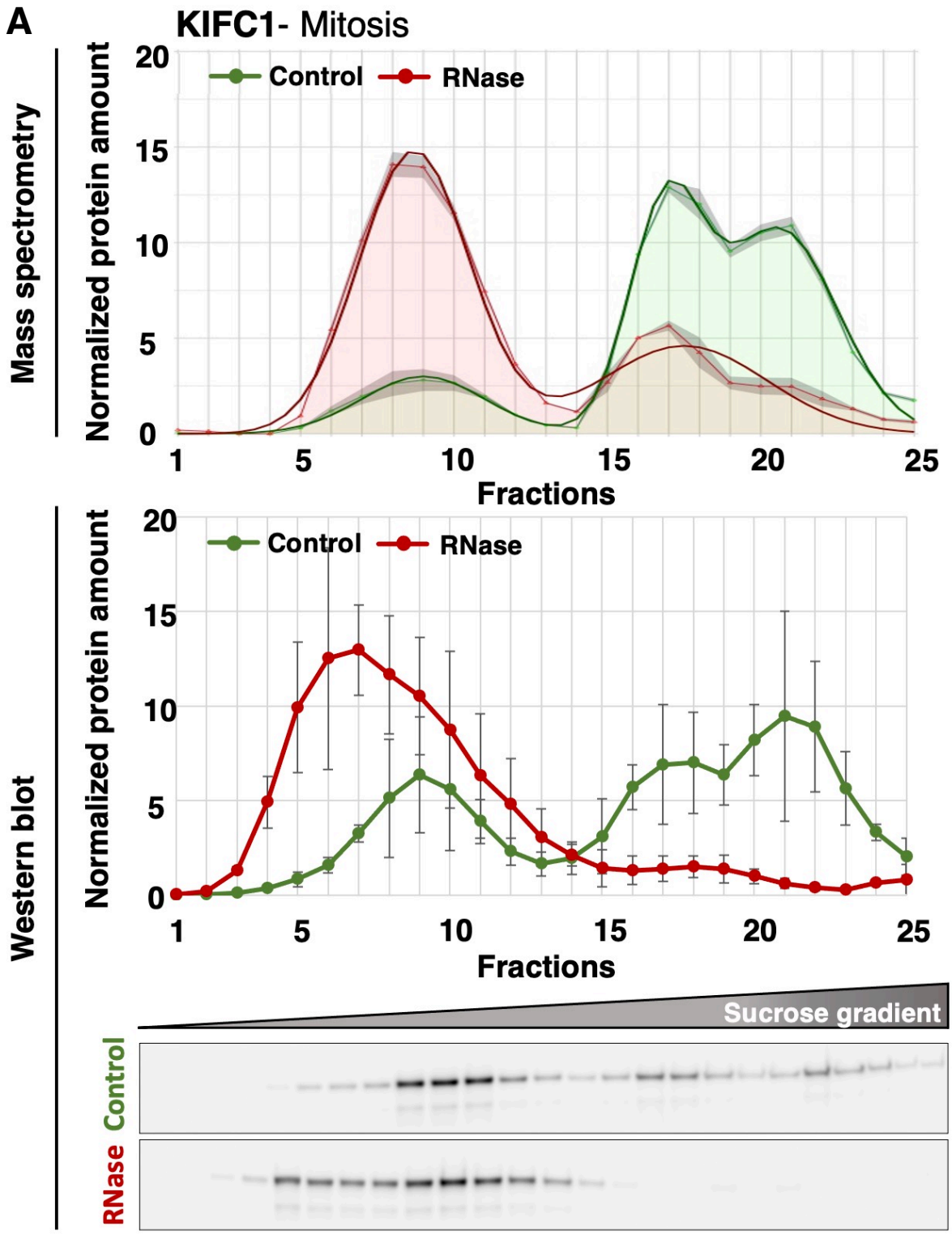


Figure 19: RNA-mediated proteins interactors of AURKA in mitosis

(A) Pie chart representing the percentage of RNase-sensitive (orange) and RNase-insensitive (blue) protein interactors of AURKA based on AURKA pulldown in prometaphase HeLa cells, and LC/MS-MS-based analysis (n=3).
(B) Pie chart representing the percentage of mitotic RNase-sensitive (orange) and RNase-insensitive (blue) protein interactors of AURKA, according to a GO analysis using the RBP2GO database (56). The proteins interactors were detected using LC-MS/MS-based analysis from AURKA pulldown in prometaphase HeLa cells (n=3).
(C) Schematic representation of a selected list of top RNase-sensitive and RNase-insensitive protein interactors of AURKA, as detected from the AURKA pulldown in presence or absence of RNase treatment.
The figure was adapted from (116).

4.2.3 RNA dependence of AURKA interactors KIFC1 and TPX2

Now that I selected two RNA-dependent interactors of AURKA of great interest based on the AURKA pulldown and mass spectrometry analysis performed above, I first verified whether KIFC1 and TPX2 were RNA dependent. To this aim, the control and RNase-treated gradient fractions from HeLa cells synchronised in prometaphase were loaded on SDS-PAGE, and western blot analysis was performed. Upon quantification of the western blot band intensities, good agreement between the mass spectrometry analysis and the western blot validation was observed. In the control gradients, the protein amount peak of KIFC1 was observed around fractions 6-11 and 16-23 and around fractions 2-7 and 17-24 for TPX2 (Figure 20A, B). Upon RNase treatment, the entire protein amount of KIFC1 and TPX2 shifted (left shift) to earlier fractions 6-11 (KIFC1) and 2-7 (TPX2) demonstrating a strong and significant partial RNA dependence (Figure 20A, B). This demonstrated the RNA dependence of KIFC1 and TPX2 in prometaphase. Upon closer observation, I recognised that AURKA, KIFC1 and TPX2 were detected in similar fractions in the control gradients with protein amount maxima at around fraction 21, indicating that these three proteins could be part of the same RNA-dependent protein complex in mitosis (Figure 17 and 20).


A
KIFC1- Mitosis
Mass spectrometry
Normalized protein amount
Control
RNase
Fractions
Western blot
Normalized protein amount
Control
RNase
Fractions
Sucrose gradient
Control
RNase

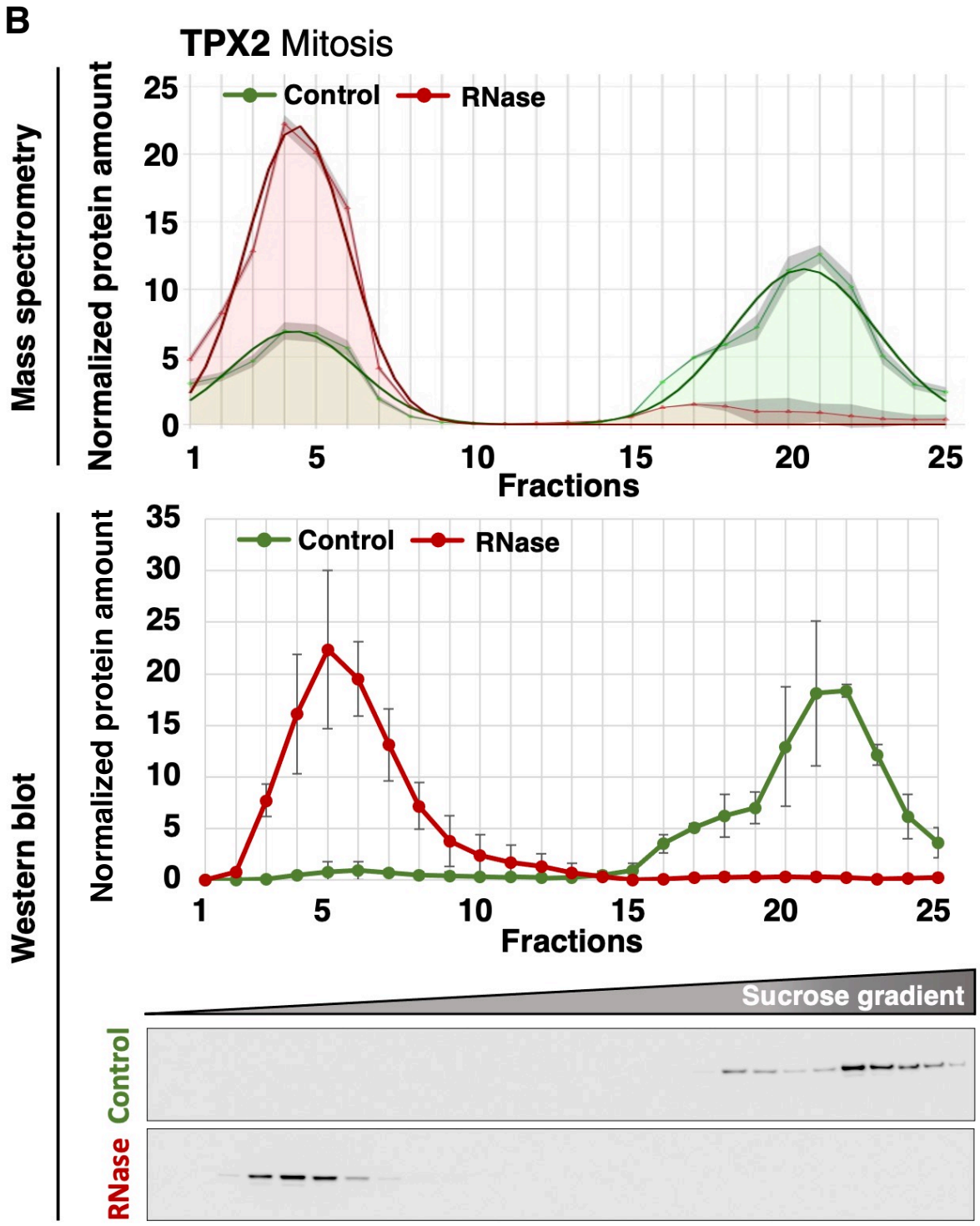
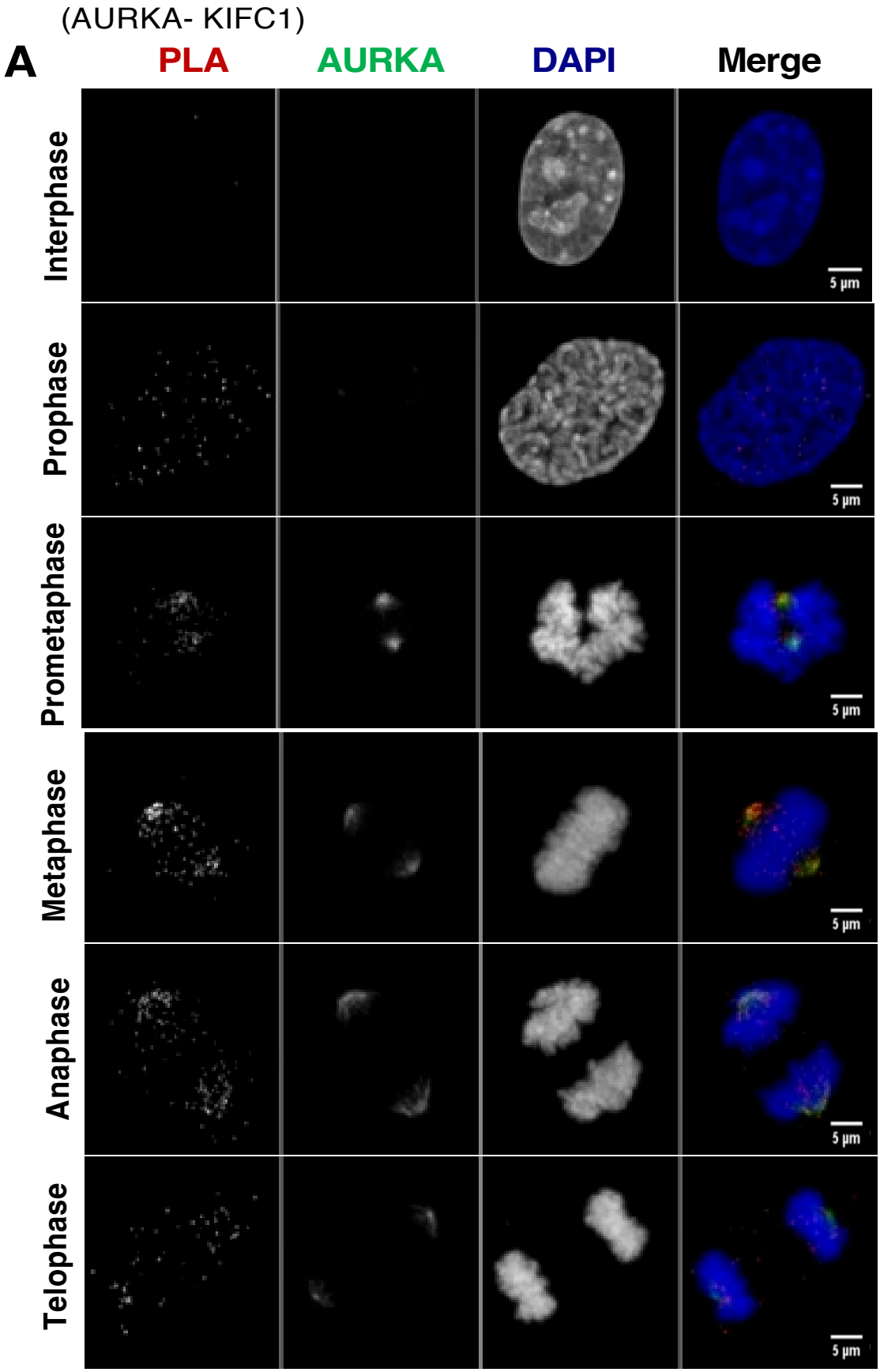


Figure 20: RNA dependence of KIFC1 and TPX2 in mitosis

(A) Top panel: Graphical representation of the mass spectrometry analysis showcasing KIFC1 amount in 25 different fractions of control (green) and RNase-treated (red) sucrose density gradients in HeLa cells synchronised in prometaphase. Line with markers is depicted by raw data (mean of three replicates). Smooth lines represent the respective Gaussian fit. The overall KIFC1 amount of the raw data was normalized to 100. Middle panel: graph showing the quantitative analysis of western blot replicates depicted by the mean of three replicates (SEM, n=3). Bottom panel: western blot representing the distribution of KIFC1 in 25 different fractions in control and RNase-treated mitotic gradients.
(B) R-DeeP profile and western blot validation of TPX2 (n=3). Same as (A). The figure was adapted from (116).

4.2.4 Interaction of AURKA-KIFC1 and TPX2-KIFC1 across cell cycle

In interphase, AURKA is a cytoplasmic protein, whereas TPX2 and KIFC1 are nuclear. Additionally, TPX2 and KIFC1 are bound to importin heterodimers until their release through the action of RanGTP. Hence, to understand the mechanism of their interaction and to identify the initial point of contact between AURKA-KIFC1 and TPX2-KIFC1, as a first step, their interactions were monitored from prophase till telophase. To visualize the proximity of two proteins in their native cellular environment, proximity ligation assay (PLA) was performed (172). This technique allows the visualization of *in situ* interactions between two proteins, at endogenous levels, represented by signal in the form of dots. First, the cells were seeded on a glass cover slip, followed by fixation, permeabilization, blocking and probing with specific primary antibody raised from different species against AURKA (mouse antibody) and KIFC1 (rabbit antibody) or against TPX2 (mouse antibody) and KIFC1 (rabbit antibody). Later, the plus and minus probes of the PLA kit were added that either recognised antibodies from rabbit or mouse. Once the probes were added, the ligase mediated ligation of the minus and the plus probes circularized the plasmid. This resulted in a signal, which indicated the proximity of two proteins. Further, upon amplification using a polymerase, the signal was amplified and the interaction between two proteins were visualised in the form of dots (see method section for more details). Here, I observed that AURKA interacted with KIFC1 from prophase until telophase and a peak of interaction was observed during prometaphase and metaphase (Figure 21A). A similar kind of interaction was observed between TPX2 and KIFC1 (Figure 21B), which further validated the potential interaction between these three proteins and also, strengthened the hypothesis that, AURKA, KIFC1 and TPX2 exited in the same complex.


(AURKA- KIFC1)
A
PLA
AURKA
DAPI
Merge
Interphase
Prophase
Prometaphase
Metaphase
Anaphase
Telophase
5 µm

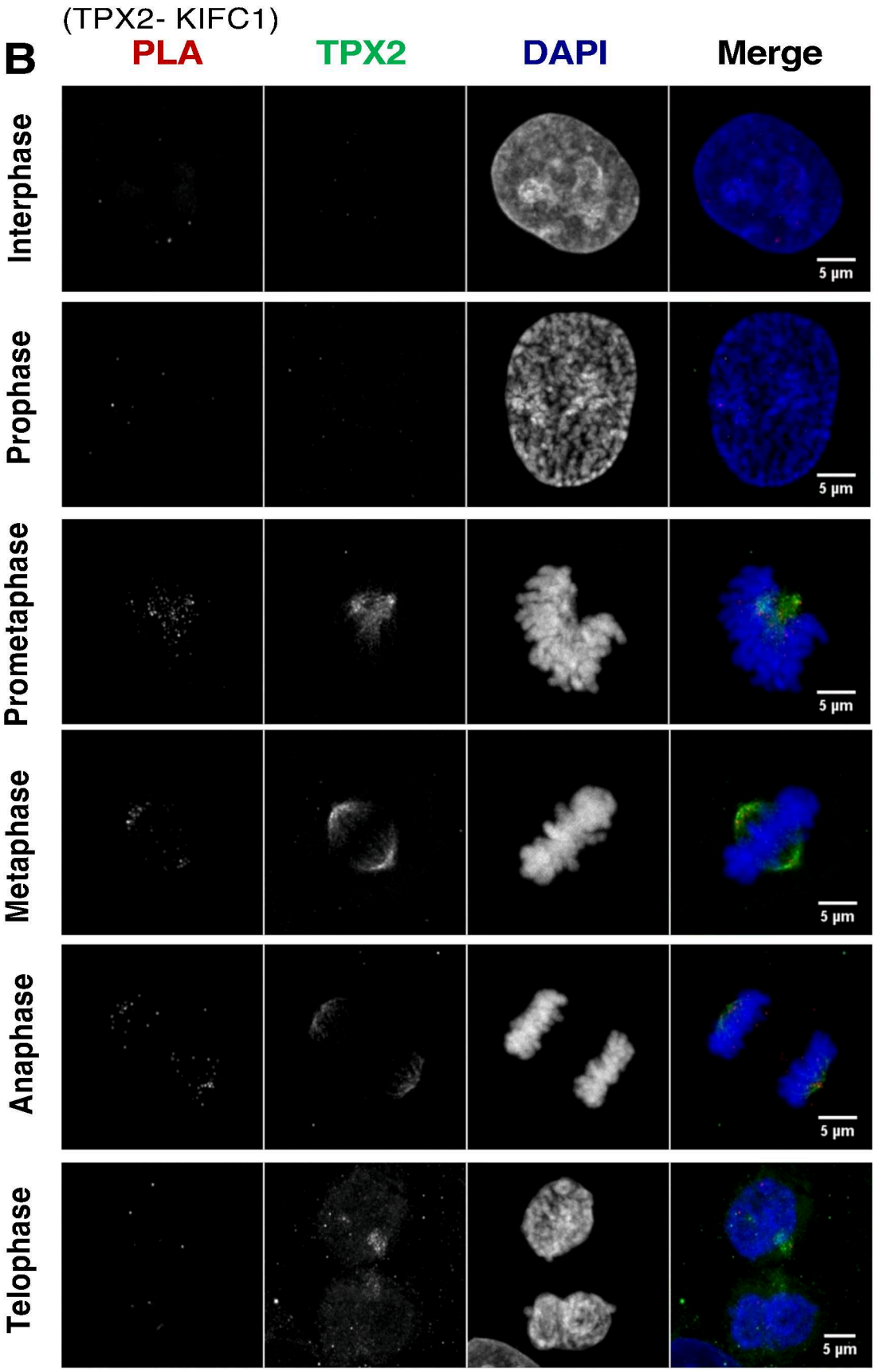

B
(TPX2- KIFC1)
PLA
TPX2
DAPI
Merge
Interphase
Prophase
Prometaphase
Metaphase
Anaphase
Telophase
5 µm

Figure 21: PLA assay representing the interaction of AURKA-KIFC1 and TPX2-KIFC1 across cell cycle

(A) Selected proximity ligation assay (PLA) images representing the in situ interaction of AURKA and KIFC1 in HeLa cells across the cell cycle (interphase to telophase as indicated). The interaction is represented by dots (PLA channel, red dots in the merge channel). In addition, AURKA was stained per immunofluorescence (green) and DNA was stained using DAPI (blue). Controls of the PLA assay and quantifications are seen in Figure 23M (see below). Scale bar, 5 µm, (n=3).
(B) Selected PLA images representing the in situ interaction of TPX2 and KIFC1. Same as (A).
The figure was adapted from (116).

4.2.5 Validating RNA-dependent interactions of KIFC1 and TPX2 with AURKA

Since we identified two new interactions of AURKA that was mediated by RNA and the RNA dependence of these interactors KIFC1 and TPX2 in mitosis, I focussed on validating their interaction using co-immunoprecipitation followed by western blot analysis. To this aim, the lysates prepared from HeLa and A549 prometaphase cells were used and treated with RNase or remained untreated (RNase-treated and control lysates). Both lysates were centrifuged, precleared and AURKA was further immunoprecipitated using a specific antibody. Next, the AURKA-bound complexes were pulled-down using magnetic beads. The complexes were eluted using 1X LDS containing DTT, loaded on a SDS-PAGE and western blot analysis was performed. Rabbit IgG was used as a negative control as the AURKA antibody was raised from rabbit. Here, I observed that, AURKA interacted with TPX2 and KIFC1 in prometaphase (Figure 22A-D). However, the interaction of AURKA with the two proteins was significantly lost upon RNase treatment on the lysate, validating the mass spectrometry results (Figure 22A-D). This proved that, AURKA interacted with KIFC1 in mitosis (Figure 22A, B). Importantly, AURKA interacted with KIFC1 and TPX2 in an RNA-dependent fashion which was not identified before (Figure 22A-D).

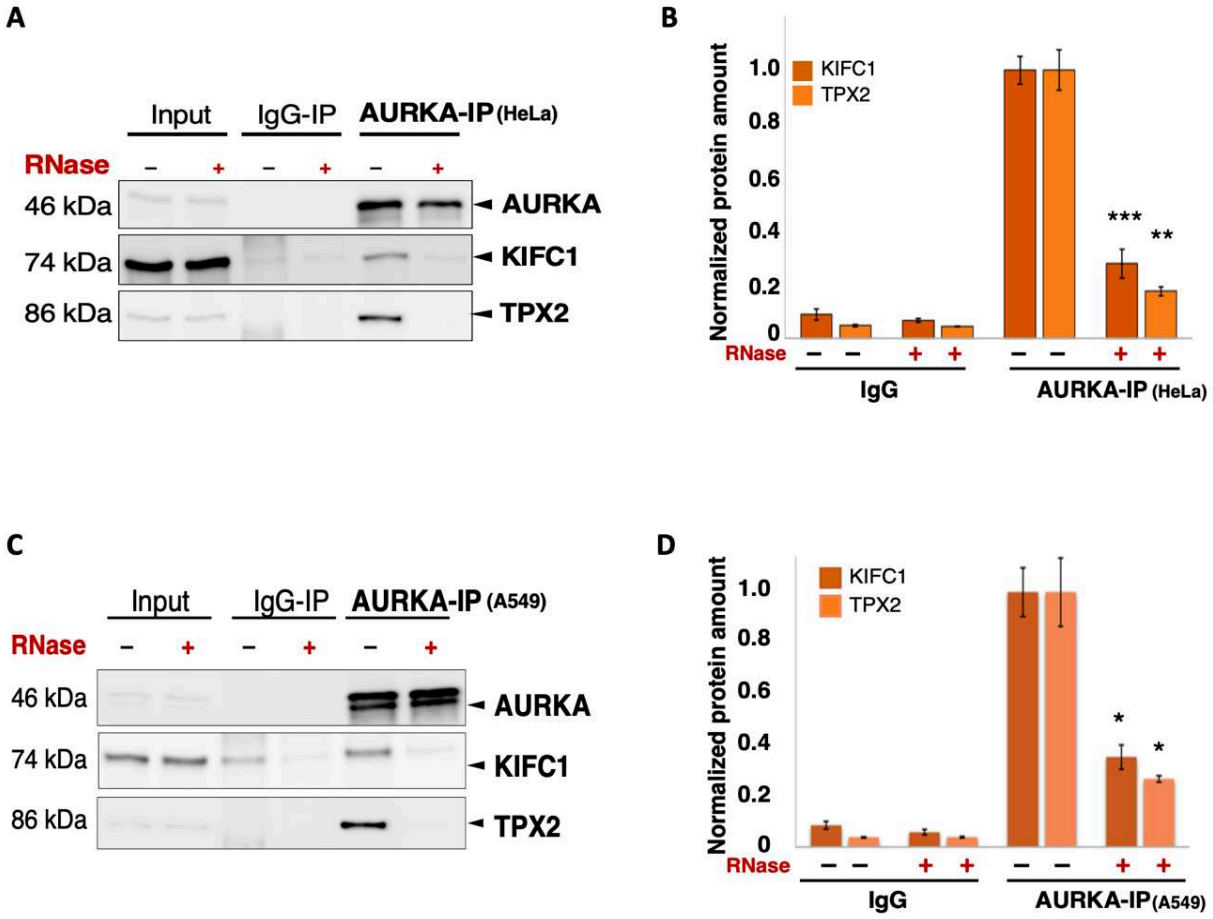
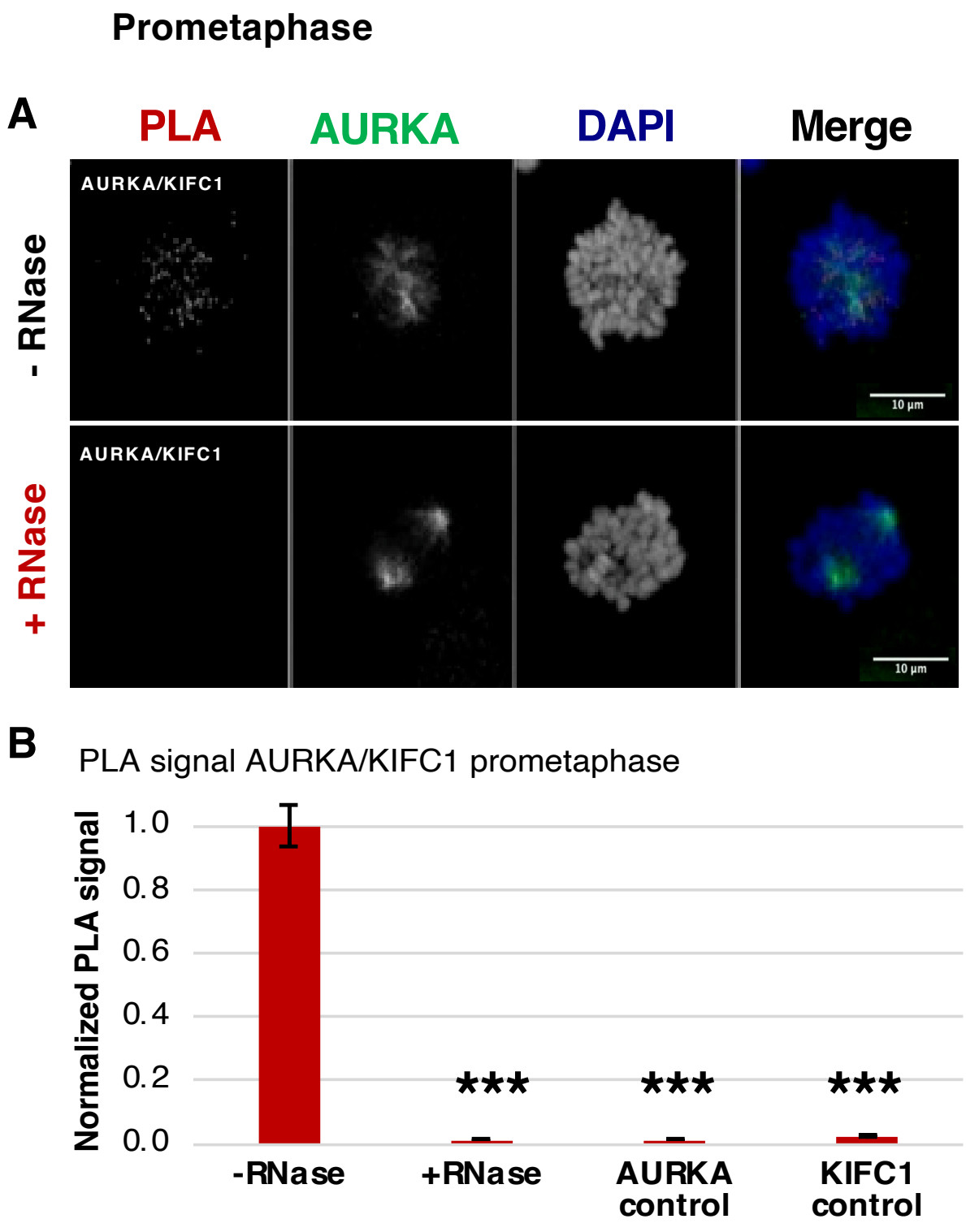


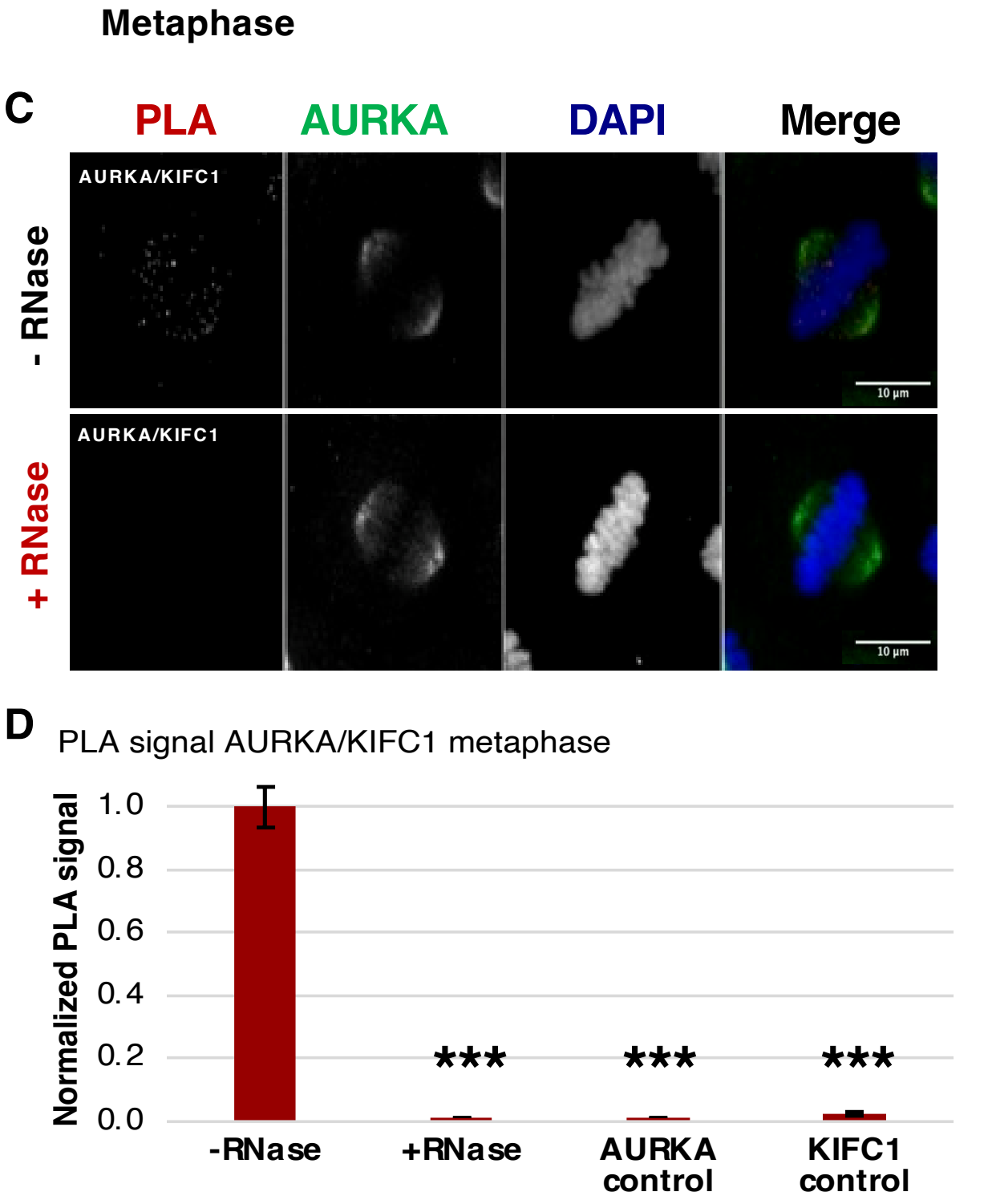
Figure 22: AURKA immunoprecipitation validating the RNA-dependent interactions of KIFC1 and TPX2 with AURKA in mitosis

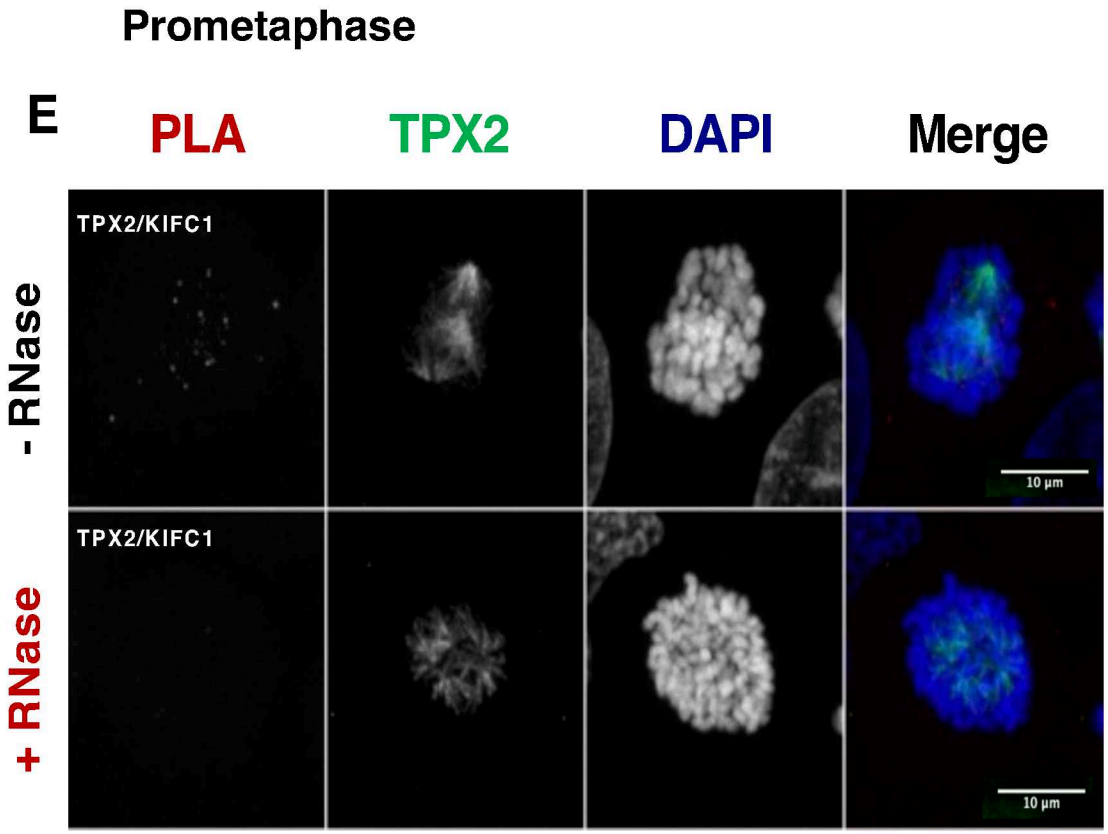
(A) Representative western blot demonstrating the interaction of AURKA (46 kDa) with KIFC1 (74 kDa) and TPX2 (86 kDa) in the presence of RNA, and the loss of interaction between these proteins upon RNase treatment in HeLa cells synchronised in prometaphase. IgG was used as a negative control (n=3).
(B) Quantitative graph representing the normalized protein amount of KIFC1 and TPX2 present in IgG and AURKA pulldown samples treated with or without RNase I in HeLa cells synchronised in prometaphase. The WB band intensities were quantified using ImageJ and the signal intensities are represented as a bar graph with SEM (n=3). P-values were evaluated using a two-tailed, paired t-test (P-value <0.05).*
(C) Same as (A) in A549 cells synchronised in prometaphase (n=3).
(D) Same as (B) in A549 cells synchronised in prometaphase (n=3).
The figure was adapted from (116).

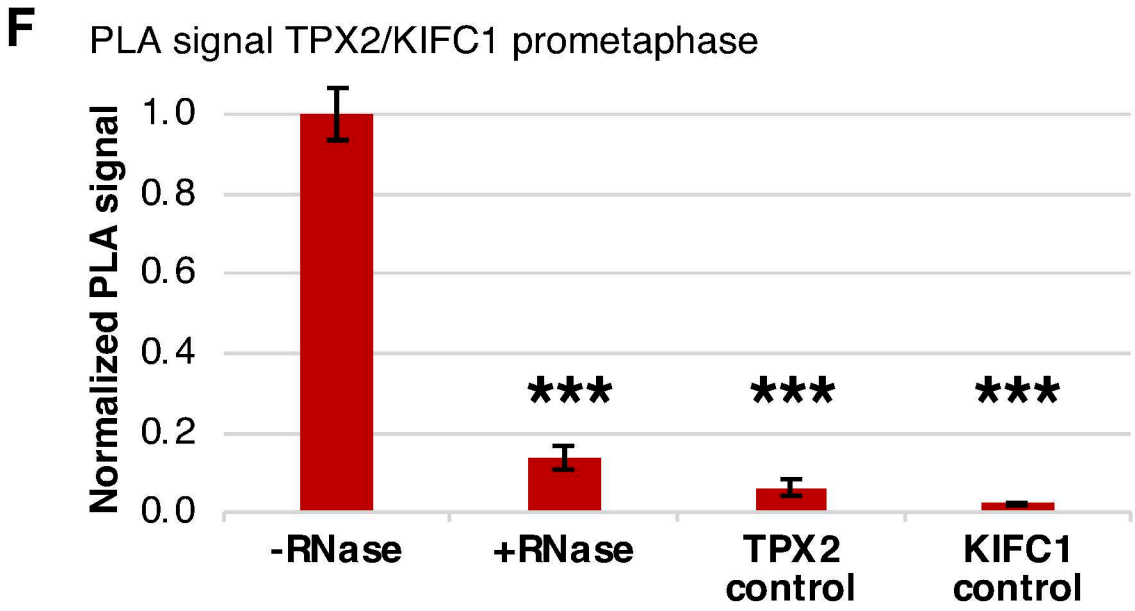
Similarly, the RNA-dependent interaction between AURKA-KIFC1-TPX2 was observed pair-wise using complementary PLA assay (Figure 23A-L). Here, the PLA assay was repeated in the presence or absence of RNase treatment before fixing the cells. The resulting signal was quantified using an ImageJ-based analysis, and a clear PLA signal in the control cells in prometaphase and metaphase was observed around the centrosomal asters and the spindles, indicating the proximity of the proteins as depicted in the Figures (Figure 23A-L). Upon RNase treatment, the PLA signal strongly decreased, reaching a level comparable to the background level produced by the antibodies alone (PLA control experiments). Altogether, these

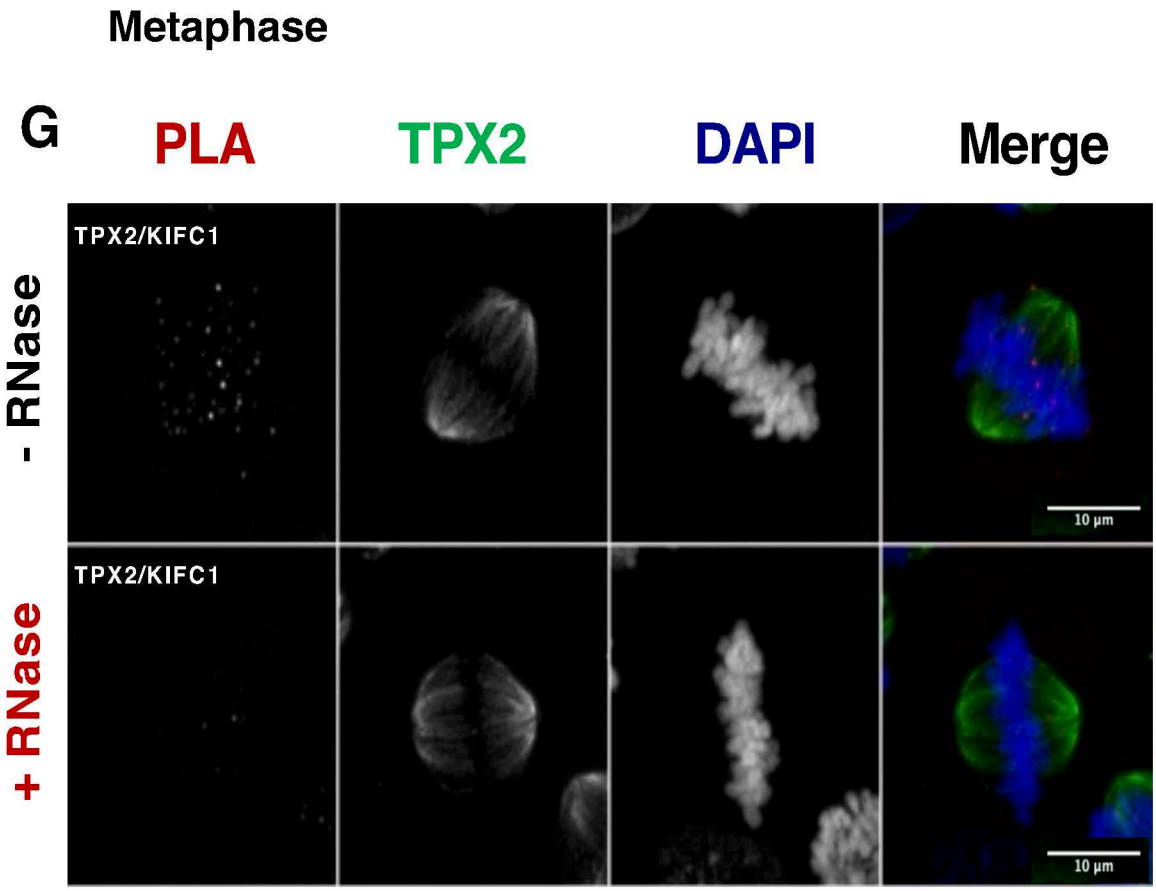
observations further confirmed the mass spectrometry results demonstrating that AURKA, TPX2 and KIFC1 interactions were mediated by RNA (Figure 23A-L).

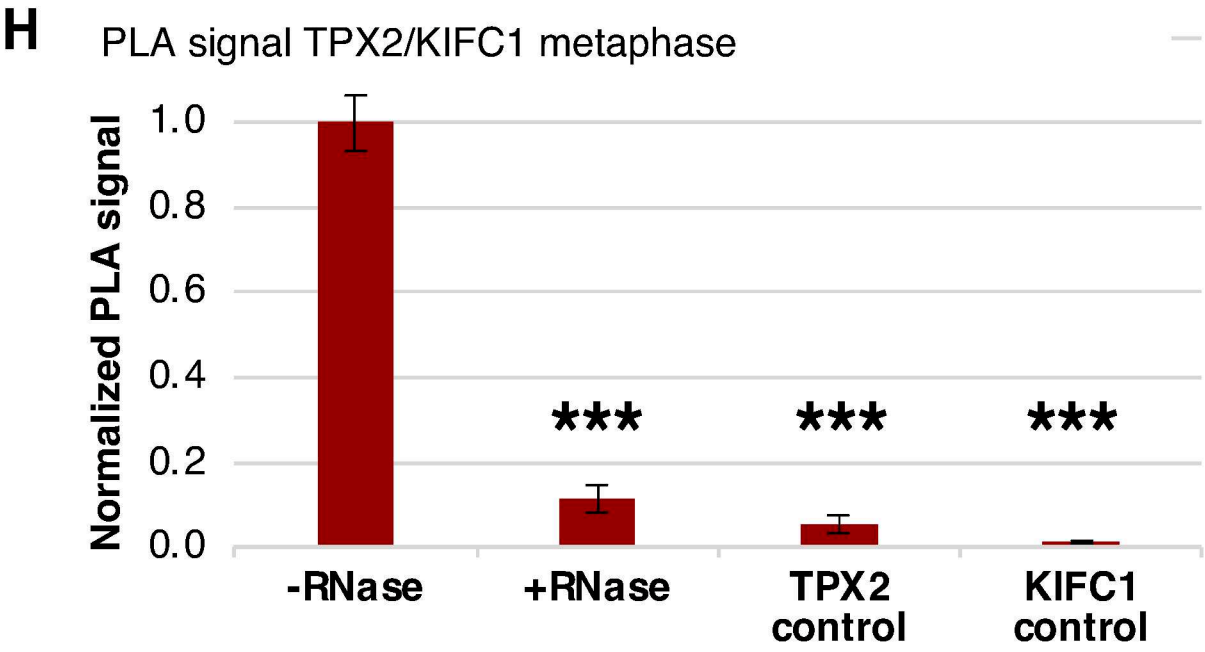


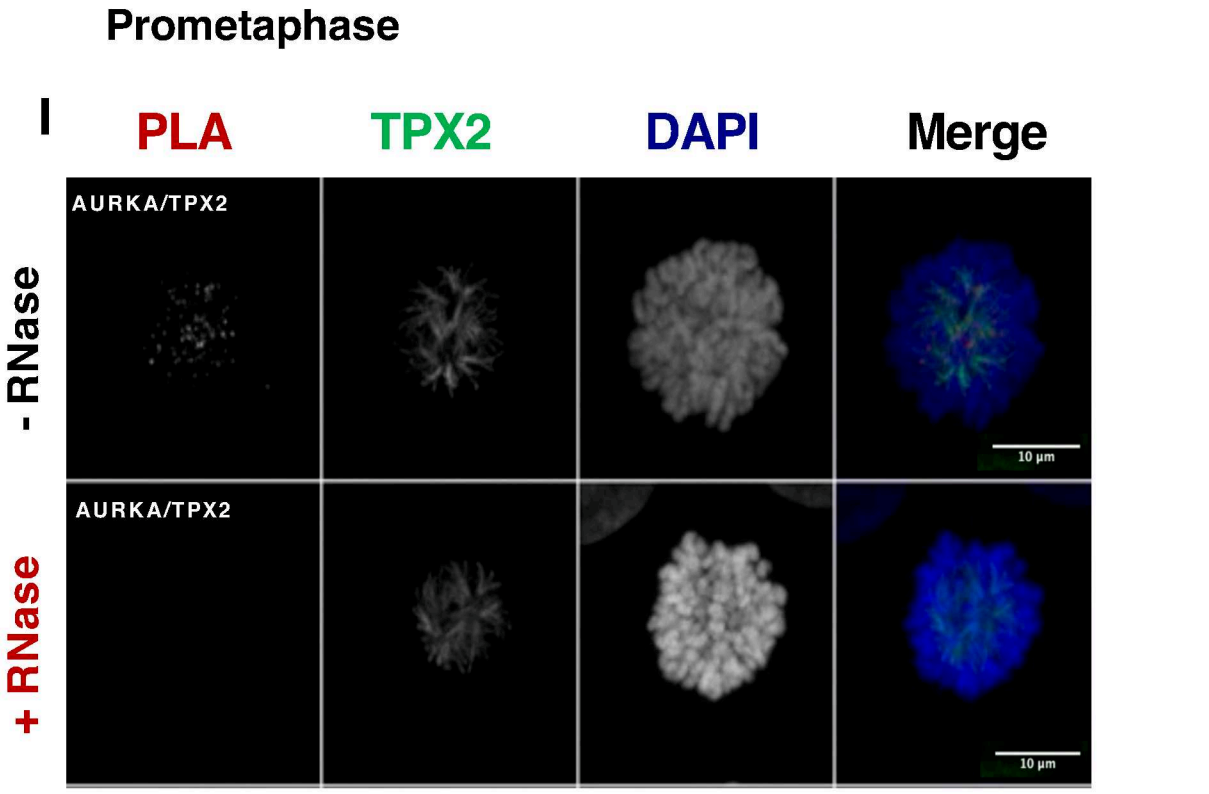

Metaphase
C
PLA
AURKA
DAPI
Merge
AURKA/KIFC1
- RNase
AURKA/KIFC1
+ RNase
10 µm
10 µm
D
PLA signal AURKA/KIFC1 metaphase
Normalized PLA signal
1.0
0.8
0.6
0.4
0.2
0.0
***
***
***
-RNase
+RNase
AURKA control
KIFC1 control

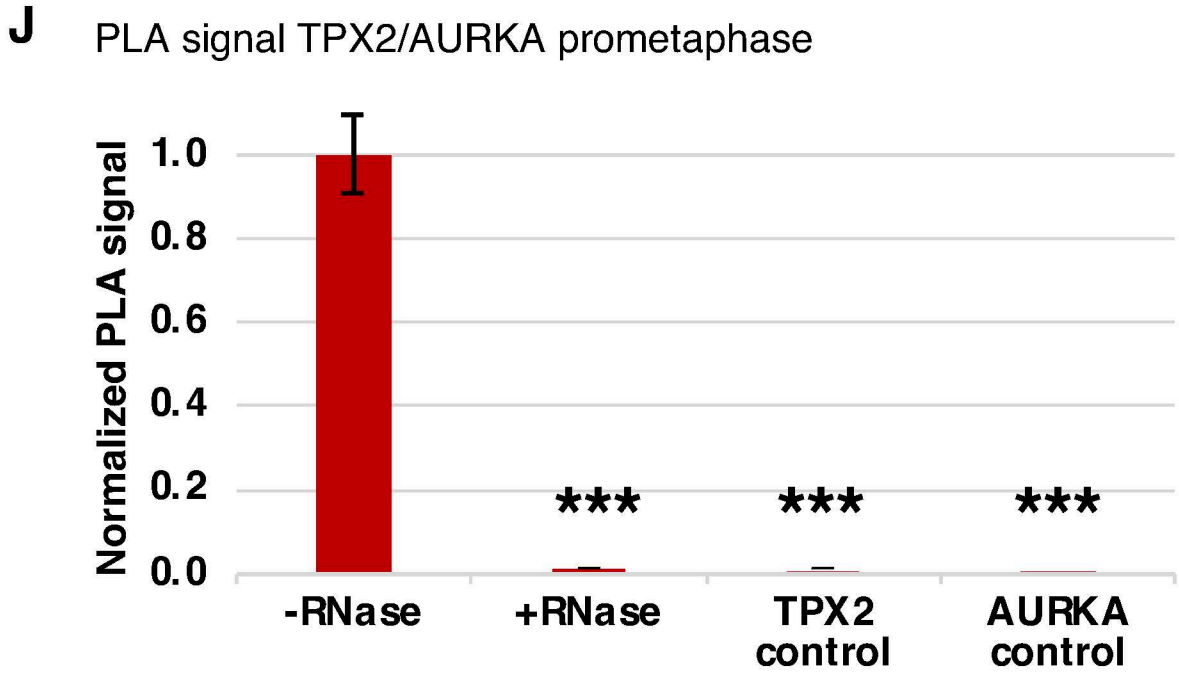

Prometaphase
E
PLA
TPX2
DAPI
Merge
- RNase
+ RNase
TPX2/KIFC1
TPX2/KIFC1
10 µm
10 µm

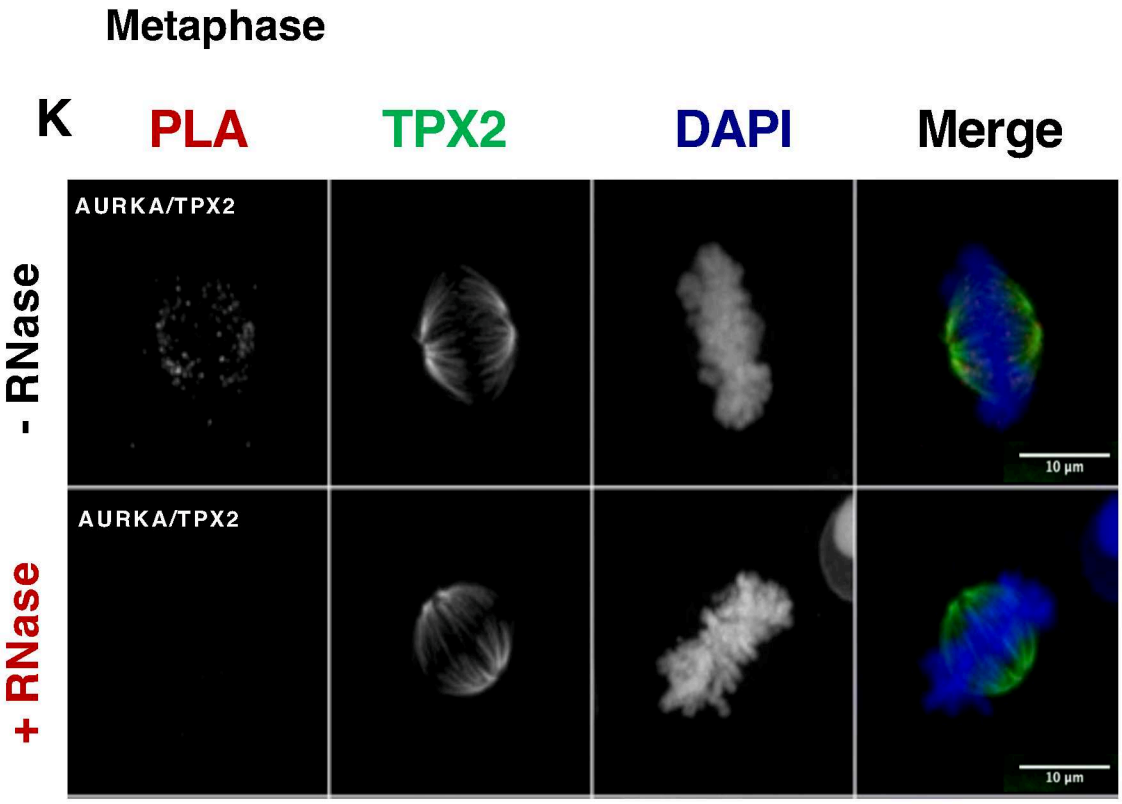

F
PLA signal TPX2/KIFC1 prometaphase
Normalized PLA signal
1.0
0.8
0.6
0.4
0.2
0.0
***
***
***
-RNase
+RNase
TPX2 control
KIFC1 control

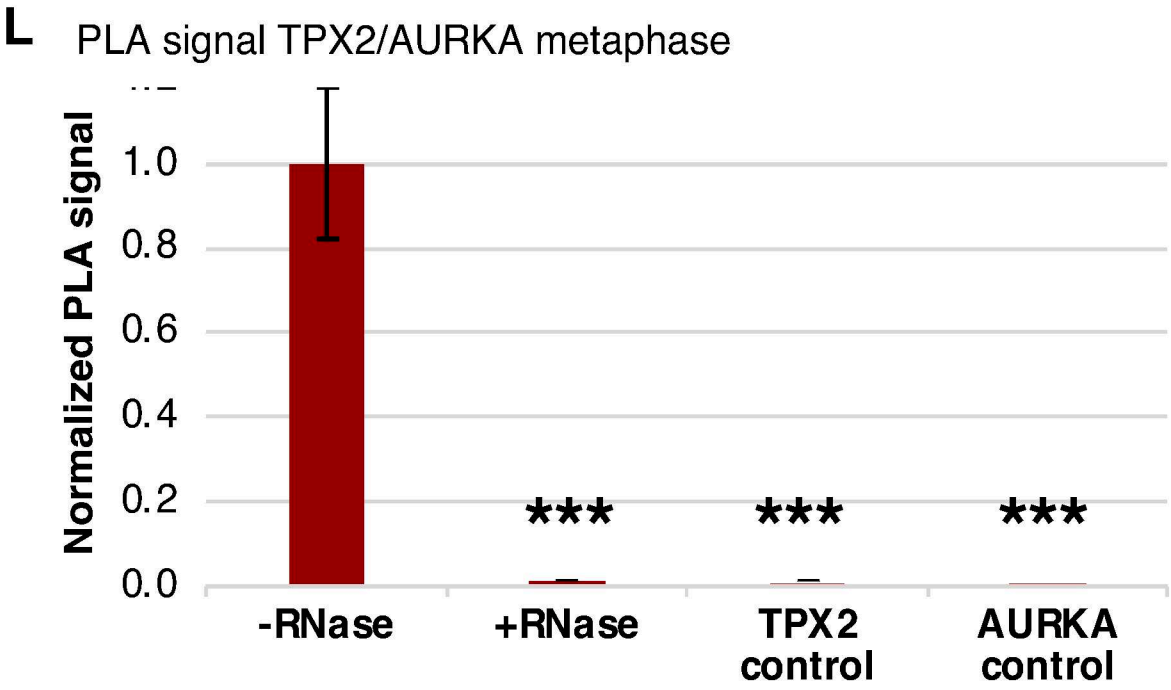

Metaphase
G
PLA
TPX2
DAPI
Merge
- RNase
+ RNase
TPX2/KIFC1
TPX2/KIFC1
10 µm
10 µm


H
PLA signal TPX2/KIFC1 metaphase
Normalized PLA signal
1.0
0.8
0.6
0.4
0.2
0.0
***
***
***
-RNase
+RNase
TPX2 control
KIFC1 control


Prometaphase
I
PLA
TPX2
DAPI
Merge
- RNase
+ RNase
AURKA/TPX2
AURKA/TPX2
10 µm
10 µm


J
PLA signal TPX2/AURKA prometaphase
Normalized PLA signal
1.0
0.8
0.6
0.4
0.2
0.0
***
***
***
-RNase
+RNase
TPX2 control
AURKA control


Metaphase
K
PLA
TPX2
DAPI
Merge
- RNase
+ RNase
AURKA/TPX2
AURKA/TPX2
10 µm
10 µm


L
PLA signal TPX2/AURKA metaphase
Normalized PLA signal
1.0
0.8
0.6
0.4
0.2
0.0
***
***
***
-RNase
+RNase
TPX2 control
AURKA control

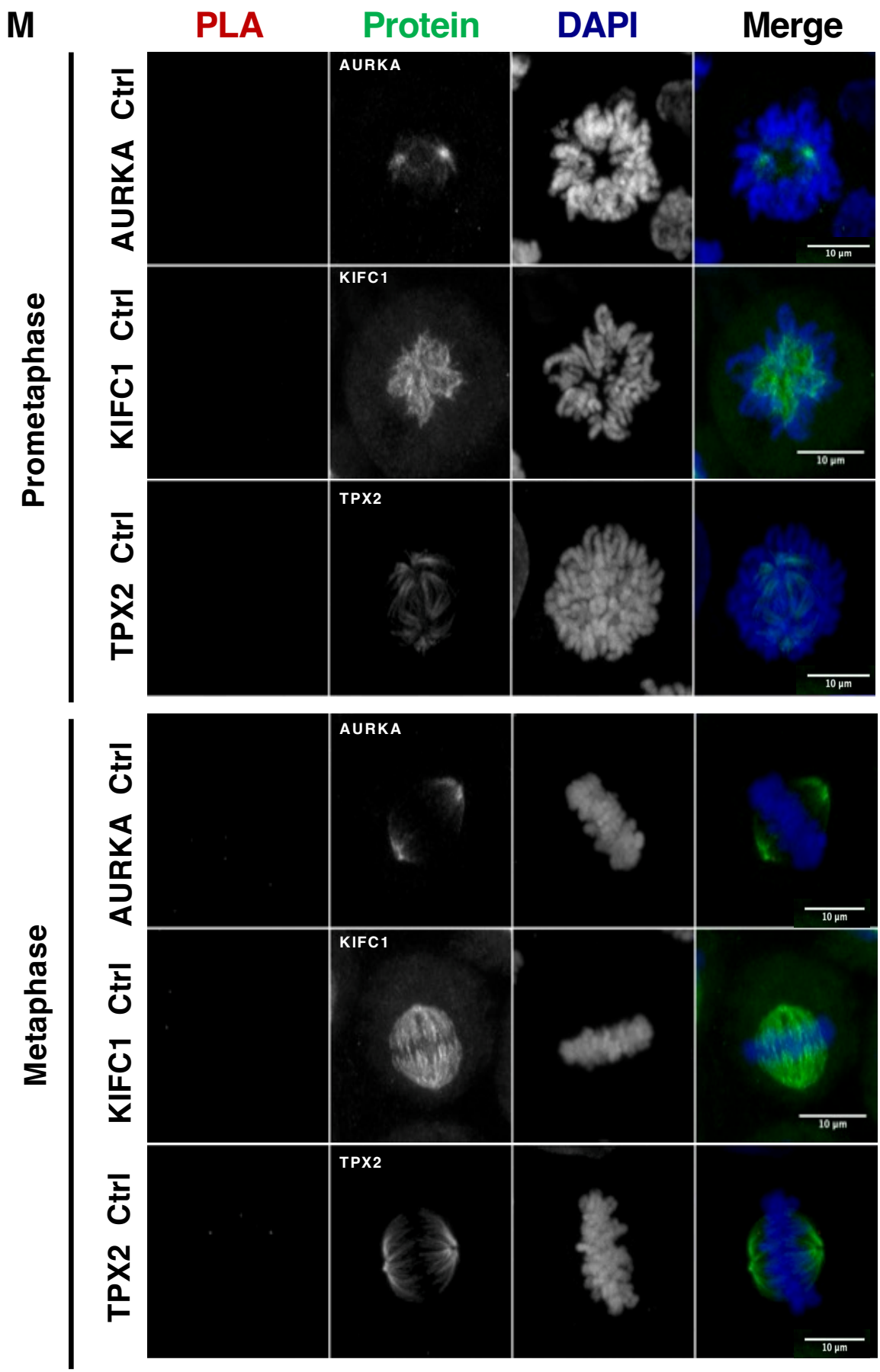

M
PLA
Protein
DAPI
Merge
Prometaphase
AURKA Ctrl
KIFC1 Ctrl
TPX2 Ctrl
Metaphase
AURKA Ctrl
KIFC1 Ctrl
TPX2 Ctrl
AURKA
KIFC1
TPX2
10 µm

Figure 23: PLA assay validating the RNA-dependent interactions of KIFC1 and TPX2 with AURKA in mitosis

(A, C) Representative PLA images showing the proximity of AURKA and KIFC1 at prometaphase and metaphase HeLa cells, (PLA channel, red dots in the merge channel). AURKA is shown in green (immunostaining) and DNA was stained using DAPI (blue). The top panel depicts representative images in control cells. Bottom panel depicts representative images in RNase-treated cells. Scale bar, 10 µm.
*(B, D) Quantification bar graph of the PLA signal for the proximity of AURKA and KIFC1 in HeLa prometaphase and metaphase cells in absence (-RNase) or presence (+RNase) of RNase treatment, as well as in the individual antibody control PLA assays (Figure 23M). The signal intensities in each sample were normalized to the signal intensity of the -RNase sample. The error bars indicate the SEM (n=3). The P-value were calculated using a two-tailed, unpaired t-test (*** P-value <0.001).*
(E-H) PLA demonstrating the proximity of TPX2 and KIFC1. Same as (A-D) (n=3).
(I-L) PLA demonstrating the proximity of TPX2 and AURKA. Same as (A-H) (n=3)
(M) Representative PLA images of the PLA assay control samples in HeLa cells synchronized in prometaphase and metaphase. These images estimate the amount of background PLA signal generated by the antibody alone and not because of the proximity of the two proteins. Here, the cells were probed with only one antibody against AURKA or KIFC1 or TPX2 (in green) and further processed for the PLA assay. It shows that each antibody alone did not produce any strong PLA signal. Scale bar, 10 µm.
The figure was adapted from (116).

To confirm that the loss of interaction between AURKA and its interaction partners KIFC1 and TPX2 was due to the lack of RNA, and not due to the adverse effects of RNase itself, the interaction between α-and β-tubulin was used as a negative control. α-tubulin and β-tubulin are the building blocks of the MTs that are at the heart of mitotic progression. They are dynamic in nature and play essential roles in chromosome capture, congression, and segregation. Thus, I hypothesized that the interaction between α/β-tubulin should persist, remaining unperturbed by the presence or absence of RNA, as long as MTs are visible in the cell. The overall PLA signal between the control and the RNase-treated samples remained equal, confirming the unperturbed interaction between α/β-tubulin upon RNase treatment. However, I observed a modification in the α/β-tubulin PLA signal distribution (Figure 24A, B). Upon RNase treatment, the PLA signal which represented the interaction between α/β-tubulin appeared less homogeneous (Figure 24A, B). To quantify the distribution of the α/β-tubulin PLA signal, the collected images were rotated and rescaled and finally superposed. This allowed to calculate the variance of the signal and to evaluate this variance pixel per pixel. This analysis exposed an increased variance of the α/β-tubulin PLA signal throughout the metaphasic spindle structure after RNase treatment, in particular at the poles of the structure (dark orange) (Figure 24C), potentially indicating a perturbation in the organization of the MTs.

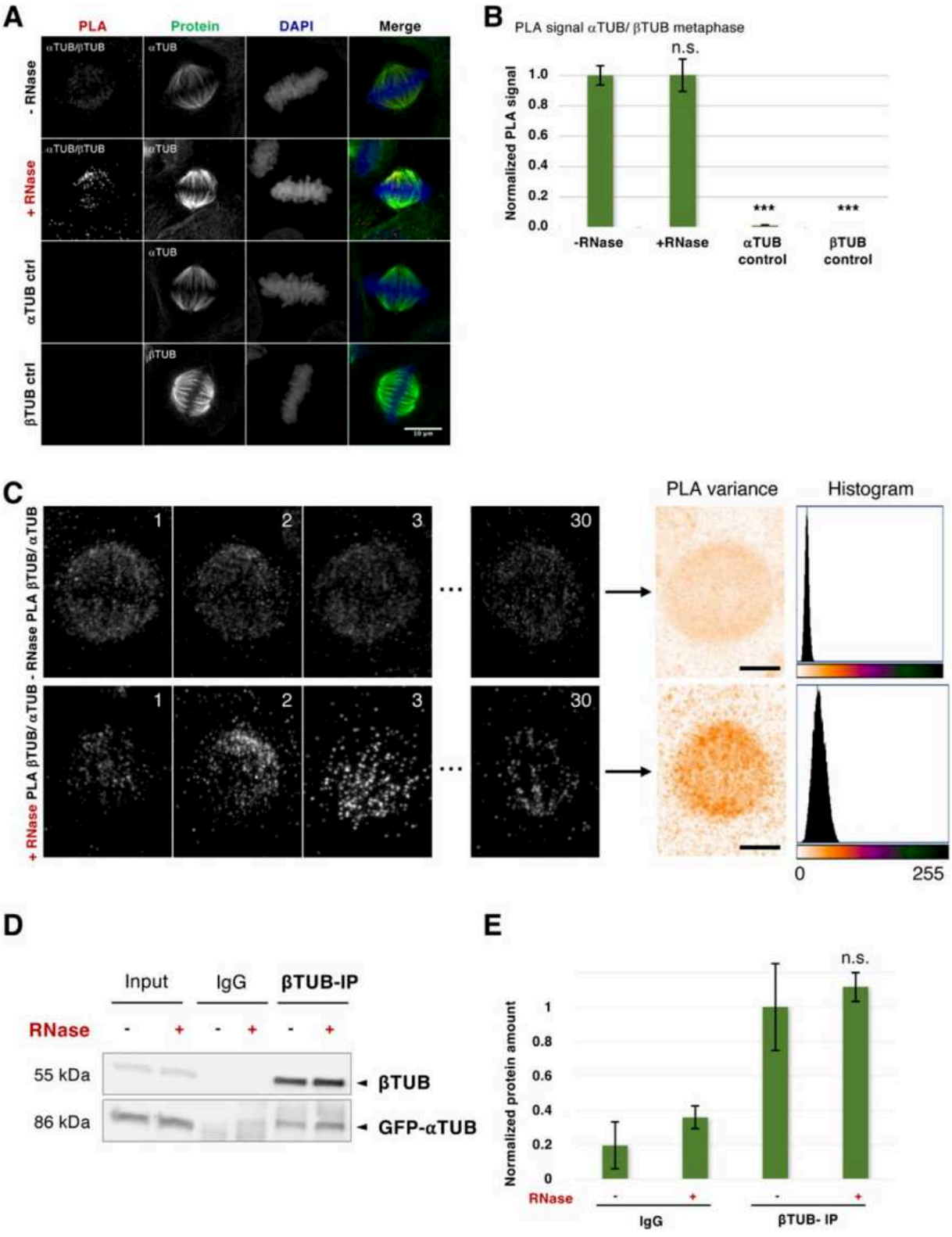


Figure 24: Interaction between α-tubulin and β-tubulin upon RNase treatment

(A) Representative PLA images indicating the proximity of α– and β– tubulin in HeLa cells synchronized in metaphase (PLA channel, red dots in the merge channel). The protein (tubulin) is stained in green and DNA was stained using DAPI (blue). The top image panels depict representative images in control cells and RNase-treated cells. The bottom two image panels depict representative images of the individual antibodies PLA controls. Scale bars, 10 μm.

*(B) Quantification bar graph of the PLA signal for the proximity of α– and β– tubulin in HeLa metaphase cells in absence (-RNase) or presence (+RNase) of RNase treatment, and the individual antibody controls. The signal intensities in each sample were normalized to the signal intensity of the -RNase sample. The P-value were calculated using a t-test (*** P-value <0.001, ns not significant), SEM, (n=3). There is no significant decrease in the PLA signal upon RNase treatment, as compared to the untreated (-RNase) sample.*
(C) Differential analysis of the PLA signal for the proximity of α– and β– tubulin. Left panel: images of three replicates (3 x 10 images) were rotated, scaled to the average pole-to-pole distance, cropped and superposed to compute an image (orange) representing the variance of the PLA signal in the two series of images (right panel). As seen from the PLA images, the PLA variances and the histograms, the PLA signal for the proximity of α- and β-tubulin after RNase treatment is remarkably heterogeneously distributed, especially throughout the spindle area (greater variance values), compared to the cells not treated with RNase (-RNase). Scale bar, 5 µm.
(D) WB analysis of the pulldown of β-tubulin (βTUB) and co-immunoprecipitation of α-tubulin tagged with GFP (GFP-αTUB) in HeLa prometaphase cell lysates treated with or without RNase. IgG was used as a negative control. αTUB was detected using an anti-GFP antibody.
(E) Bar plot representing the quantification of the pulldown of β-tubulin (βTUB) and co-pulldown of α-tubulin tagged with GFP (GFP-αTUB) in HeLa prometaphase cell lysates treated with or without RNase, as in D. The P-value was calculated using a two-tailed, unpaired t-test (n.s., not significant), SEM, (n=3).
The figure was adapted from (116).

In addition to the PLA analysis, immunoprecipitation of β-tubulin was performed in mitotic lysates from cells that stably expressed GFP-α-tubulin, followed by the western blot analysis. The immunoprecipitation also confirmed the persistence of the interaction between β- and α-tubulin even after RNase treatment (Figure 24D, E).

Taken together, I demonstrated that AURKA interacted with KIFC1 from the onset of prophase until telophase. Importantly, we observed an RNA-dependent interaction between AURKA, TPX2 and KIFC1 throughout mitosis. Though the interaction between α/β-tubulin persisted upon RNase treatment, the distribution of the PLA signal pattern became heterogenous, indicating perturbations in the MT-based structure. Thus, these results indicated that RNA differentially regulates protein-protein interactions during mitosis.

4.2.6 KIFC1 is an RNA-binding protein

Since KIFC1 was RNA dependent and interacted with AURKA and TPX2 essentially in the presence of RNA, I questioned if KIFC1 was an RBP. Can KIFC1 directly bind to RNA? To answer these questions, the iCLIP2 assay was repeated with KIFC1 in HeLa and A549 prometaphase cells synchronised in prometaphase. The cells were lysed, treated with varying concentrations of RNase I, KIFC1 was pulled down, and RNA labelling was performed using [γ-32P] ATP. Finally, autoradiography revealed that KIFC1 can directly bind to RNA. Here, IgG and non-crosslinked (NCL) samples were used as negative controls. At highest RNase concentration, the signal was observed at the height of the protein (74 kDa). With decrease in

the RNase I concentration, an increasingly dispersed smear was observed towards higher molecular weight further substantiating that KIFC1 is a true RBP (Figure 25A, C). The immunoprecipitation of KIFC1 was verified by western blotting as a control to demonstrate equal protein amounts in all the samples (Figure 25B, D). Similarly, iCLIP2 technique was performed to identify the direct RNA-binding capacity of TPX2 by my colleague and they observed similar shift in the RNA signal upon RNase dilution, proving that TPX2 is an RBP (data not shown).

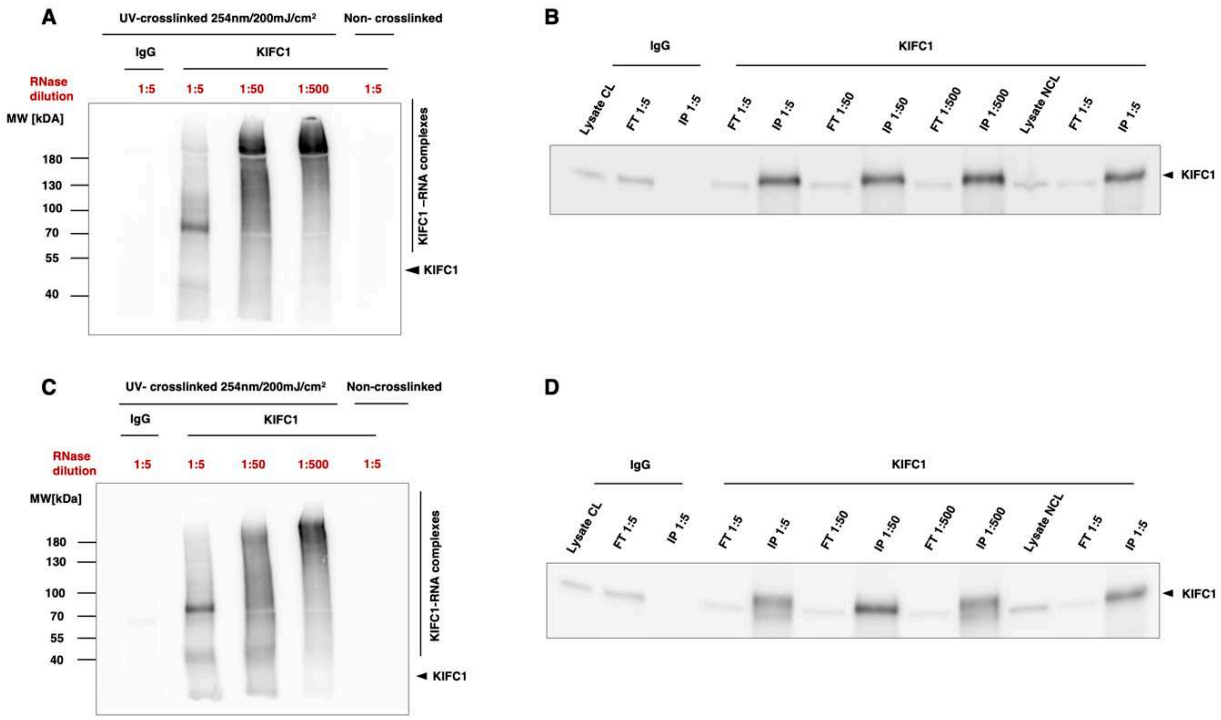


Figure 25: KIFC1 is an RNA-binding protein

(A) Autoradiography presentation the direct binding of KIFC1 to RNA (74 kDa) by iCLIP2 assay in HeLa prometaphase cells, as indicated by the shift in radioactive signal toward higher molecular weights with decreasing RNase I concentrations (n = 3).
(B) Western blot validating the equal amount of KIFC1 in all the samples. 2.5% of FT and total lysate from the CL and NCL samples were loaded on the gel for validating the immunoprecipitation efficiency and size confirmation (n = 3).
(C, D) Autoradiography and western blot images of KIFC1 iCLIP2 assay (n=3) in A549 prometaphase cells proving the direct binding of KIFC1 to RNA. Same as (A, B).
The figure was partially adapted from (116).

Since KIFC1 was able to directly bind RNA, I focussed on identifying the interacting RNAs as the next step. The iCLIP2 for KIFC1 presented a high radioactive signal of interacting RNAs (Figure 25A, C). Upon isolating and sequencing the KIFC1-bound RNAs, the sequences of the KIFC1-associated RNAs were analysed based on the published pipeline and protocol

(124,128). About 80 million reads were obtained from each of the four replicates, which correlated well with each other (Figure 26A). Following quality control and alignment, a proportion of up to 45% of rRNA reads was observed (Figure 26B). Furthermore, the analysis of the uniquely mapped reads showed that the vast majority of the binding sites (~80%, 29465 out of 36934) were associated to protein-coding mRNA transcripts (Figure 26C). In total, KIFC1-associated binding sites were localized in 5687 genes, in particular in the coding, intronic and 3'UTR sequences (Figure 26D). The analysis of the pentamer frequency at the binding sites did not disclose a particular sequence specificity for the strong binding sites, which were defined according to the PureCLIP scoring system (126). The most enriched pentamers contained U-stretches, most likely reflecting the uridine bias of UV crosslinking inherent to the method, and were rather localized in the lower 20% binding sites (Figure 26E) (173). Since, TPX2 was RNA binding, iCLIP2-Seq was performed by my colleague to identify TPX2 binding RNAs and similar results were observed (data not included). TPX2 predominantly bound rRNAs and protein-coding RNAs with no sequence specificity or motif enrichment. Also, we observed a huge overlap between the RNA targets of KIFC1 and TPX2 (data not included). These parallel and comparable findings from my colleague strengthened the KIFC1 iCLIP2-seq data and proved that RNA binding of KIFC1 is not by chance. This results also suggested that these unconventional RBPs, mainly KIFC1 and TPX2 behaved in a similar manner binding to same RNA targets without a clear mechanism.

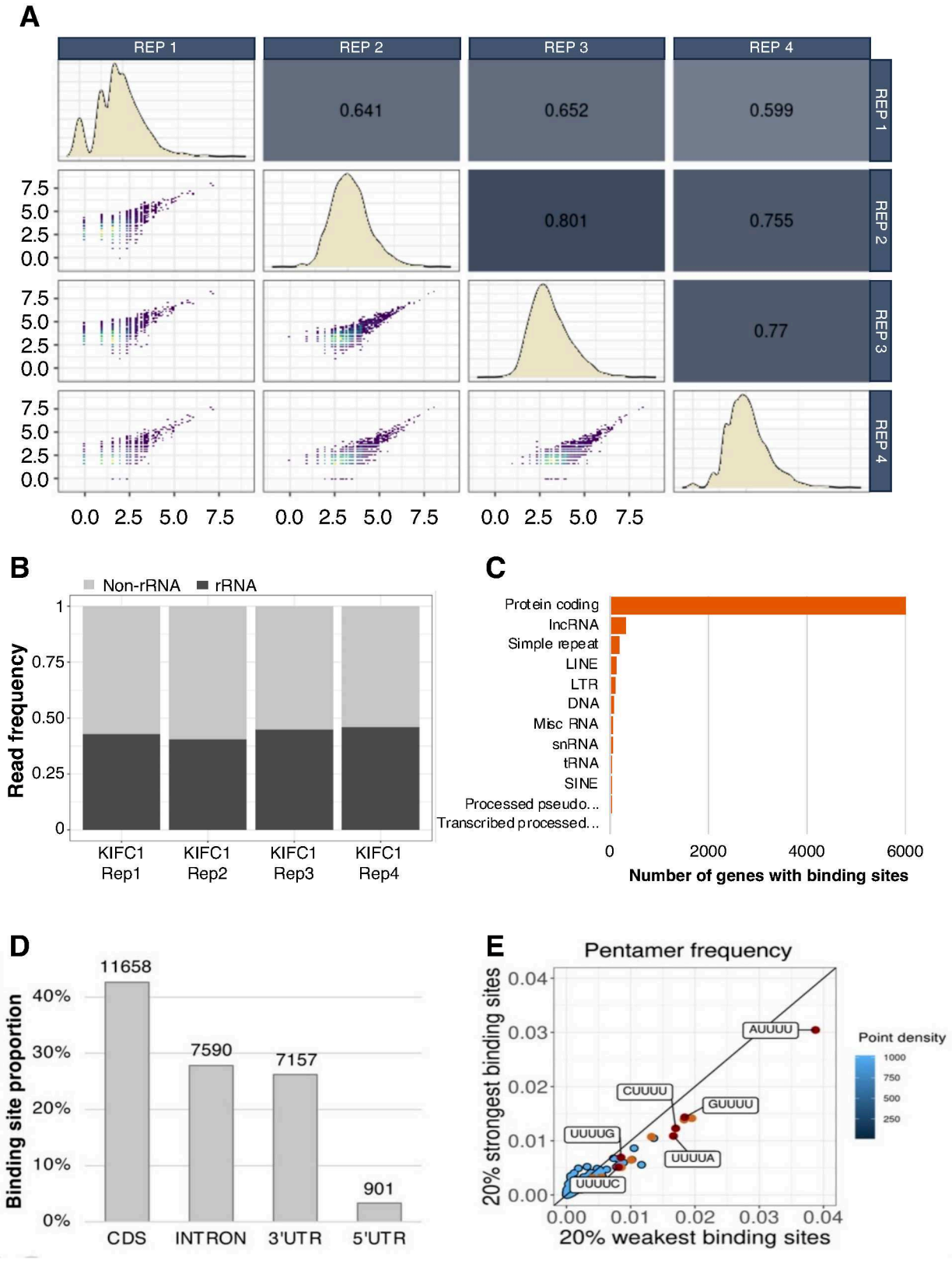


Figure 26: KIFC1 binds to ribosomal RNA and protein-coding RNA transcripts in mitosis

(A) Pairwise correlations between the KIFC1 iCLIP2-Seq replicates (n=4) as scatter plot (bottom left), the Pearson correlation coefficient (upper right) and the coverage distribution as density (diagonal), represented by composite graph.

(B) Bar plot representing the ribosomal RNA (rRNA) and non-ribosomal RNA (non-rRNA) read frequencies in KIFC1 iCLIP2-Seq replicates in HeLa prometaphase cells.
(C) Horizontal bar plot showing the KIFC1 target RNA gene spectrum with number of genes identified in decreasing order in iCLIP-Seq data. The genes are classified as protein coding, lncRNA, simple repeat, LINE, LTR, DNA, misc RNA, snRNA, tRNA, SINE, processed pseudogene and transcribed processed pseudogene. The KIFC1 iCLIP2-Seq was performed using HeLa prometaphase cells.
(D) Bar plot representing the binding site proportion in the respective transcript regions of protein-coding genes.
(E) Scatter plot comparing the pentamer frequency within the 7-nt binding sites within the 20% strong binding sites vs. 20% weakest binding sites as defined by the PureCLIP score. The pentamers with the most extreme frequencies are coloured in orange and red. These extreme pentamers contain U-stretches.
The figure was adapted from (116).

Further, to validate the iCLIP2-Seq results, an RNA affinity purification was performed. The top 20 KIFC1 target RNAs based on PureCLIP score is presented in Table 6. The top four KIFC1 interacting protein-coding RNAs were selected based on (i) high PureCLIP score (≥ 80) with 91.1 being the highest score, (ii) relevance of each candidate to mitosis, (iii) binding site outside of the intronic regions with the aim to study the functional role of the specific RNA in mitosis. Further, gene blocks (gblocks) were designed around the 7-nt binding site on the target RNAs, with the T7 polymerase promoter site for *in vitro* transcription. Additionally, the transcripts were or not labelled with biotin 16-UTP every 6 nt for RNA pulldown in the following steps. Lysates prepared from HeLa cells synchronised in prometaphase were used. The lysates were incubated with biotin-labelled or unlabelled RNA (used as control RNAs) for complex formation with the interacting proteins. Next, the RNA was pulled down with streptavidin coated beads and the beads washed with wash buffer to get rid of unspecific binding. The purified complexes were then eluted using 1X LDS containing DTT. Finally, the samples were loaded on SDS-PAGE and western blot was performed. Further, the blots were stained for proteins of interest: KIFC1, AURKA and TPX2. Here, I observed that KIFC1-target RNAs interacted not only with KIFC1, but also with its interacting partners, AURKA and TPX2 (Figure 27).

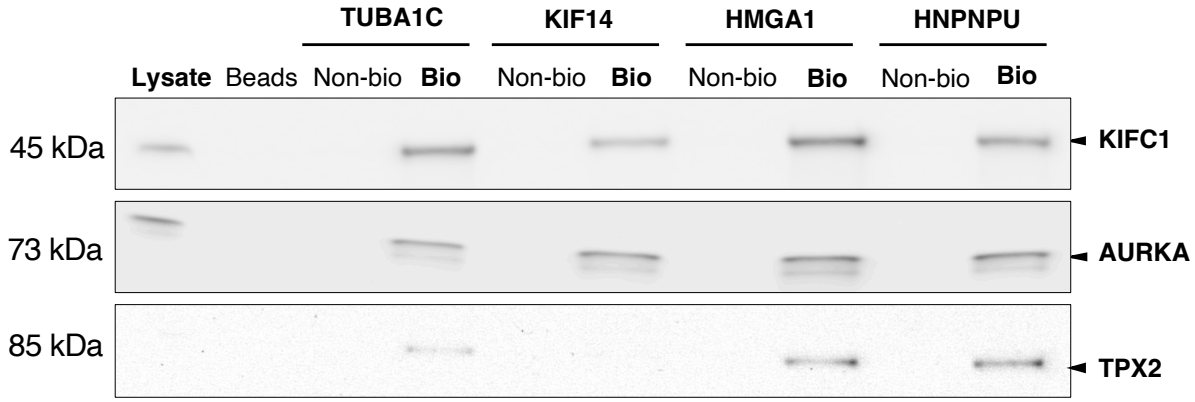


Figure 27: RNA affinity purification of KIFC1-bound transcripts in HeLa prometaphase cells

The western blot depicts the interaction of KIFC1 (74 kDa), AURKA (46 kDa) and TPX2 (85 kDa) with biotin 16-UTP labelled KIFC1-target RNAs (Tubulin mRNA, KIF14 mRNA, HMGA1 mRNA and hnRNPu mRNA) identified through KIFC1-iCLIP2-Seq analysis. Here, the unlabelled RNAs were used as a negative control. As expected, the corresponding lanes do not depict any protein signal. The experiment was performed in triplicates.

Table 6: Top 150 KIFC1-bound RNAs in prometaphase

Gene Name	Gene Type	Transcript Region	PureCLIP score	Binding Sequence
PUM1	protein_coding	INTRON	91.1	CTATAGT
PABPC4	protein_coding	INTRON	91.1	TTGTTTT
FSCN1	protein_coding	UTR3	90.8	CCTCTTT
ODC1	protein_coding	INTRON	90.5	AACCATT
MOSMO	protein_coding	INTRON	90.5	CCCAGGC
PRRC2A	protein_coding	CDS	90.5	TTCTATG
C1QBP	protein_coding	CDS	90.5	GCCTCTG
TRIM9	protein_coding	INTRON	90.4	CGGCTCG
KCNB1	protein_coding	INTRON	90.3	GATAAAA
NDUFS7	protein_coding	REPEATS	90.3	GCGCGTT
ZNF131	protein_coding	UTR3	89.9	TAAAAGT
NOS1	protein_coding	INTRON	89.7	AGACTCC
RPL23	protein_coding	CDS	89.6	TCTTTAT
NCL	protein_coding	CDS	89.6	ATTTTGA
CLPTM1L	protein_coding	UTR3	89.5	AATTTTT
RPL23A	protein_coding	INTRON	89.2	TATGTTA
SLCO5A1	protein_coding	INTRON	89.2	GGCATGG
P4HB	protein_coding	CDS	89.1	CAGATGC
KANSL2	protein_coding	INTRON	89.1	GAATCAA
CD46	protein_coding	CDS	88.9	GTAATTT
PABIR2	protein_coding	UTR3	88.7	TAAAATT
HSP90B1	protein_coding	CDS	88.6	AACTGTT
SNRPD3	protein_coding	UTR3	88.5	TCCTTTT
SAMD1	protein_coding	INTRON	88.5	GCATATT
TM4SF1	protein_coding	UTR5	88.4	CTAATTT
EIF4B	protein_coding	CDS	88.4	CATTTGG
ATG16L1	protein_coding	INTRON	88.3	TTTTTCC
NALF1	protein_coding	INTRON	88.3	AGAGTGA
APP	protein_coding	UTR3	88.2	GACTTTT
SHROOM1	protein_coding	CDS	88.1	ACCTAGA
ENSG00000266086	protein_coding	INTRON	88.1	ACATTTA
TUBA1C	protein_coding	CDS	88.0	CAATACC
SMPD3	protein_coding	INTRON	87.9	CCCCTTC
HSP90B1	protein_coding	UTR5	87.7	GGTGTAG
ATP5F1B	protein_coding	INTRON	87.7	TTCTTCT

LAMA2	protein_coding	INTRON	87.6	ATTTGTC
SLC16A3	protein_coding	CDS	87.5	TCTTCGT
CAPZA1	protein_coding	UTR3	87.5	TTTATTT
YBX1	protein_coding	CDS	87.4	TATTTGT
MYDGF	protein_coding	INTRON	87.4	CAACAGC
EEF1G	protein_coding	CDS	87.1	CTGGATC
ENO1	protein_coding	CDS	87.1	ATCCATG
ACTB	protein_coding	CDS	87.0	GATCATT
ATN1	protein_coding	INTRON	87.0	GGGCATT
MCF2L	protein_coding	INTRON	87.0	ACACTCC
NOL11	protein_coding	INTRON	86.9	ATTTACC
MFSD12	protein_coding	INTRON	86.9	TTGGTAG
SLC16A3	protein_coding	CDS	86.9	CACAAGC
ANKRD30BL	protein_coding	INTRON	86.8	AATAACA
TAF1D	protein_coding	INTRON	86.6	TAGAGAA
EEF1E1-BLOC1S5	protein_coding	INTRON	86.6	AATAATT
PHB2	protein_coding	INTRON	86.6	CATGGGC
IGF2BP1	protein_coding	INTRON	86.6	GGTGTTC
NDUFB2	protein_coding	CDS	86.5	AGTATGT
LAMP1	protein_coding	UTR3	86.5	GGGGTGC
HECTD4	protein_coding	UTR5	86.5	CTCCTGA
HSPA9	protein_coding	INTRON	86.3	TATTTGT
TM4SF1	protein_coding	CDS	86.3	CCTTTGC
FSTL3	protein_coding	UTR3	86.3	AACCACC
GAPDH	protein_coding	CDS	86.1	CGGGAAA
PTMA	protein_coding	UTR3	86.1	AATCTAA
RACK1	protein_coding	CDS	85.9	CAGTTTG
EGLN1	protein_coding	UTR3	85.6	GTGATTG
HNRNPU	protein_coding	CDS	85.5	GAATTTT
UBAP2	protein_coding	INTRON	85.5	CCAGAAA
POLD1	protein_coding	INTRON	85.3	CGGCTCC
ENSG00000262526	protein_coding	UTR3	85.2	CAGCTTT
PHB2	protein_coding	INTRON	85.2	ACTACAG
CTSZ	protein_coding	UTR3	85.1	GGGATCC
HNRNPA2B1	protein_coding	CDS	85.1	AACTTTG
RBM26	protein_coding	UTR3	85.0	AGCATTT
GAPDH	protein_coding	UTR5	84.9	CTTCTTT
B2M	protein_coding	UTR5	84.9	CAGCATT
PRDX1	protein_coding	CDS	84.9	AAATATG
ATF1	protein_coding	CDS	84.7	CATATTG
PFN1	protein_coding	UTR3	84.7	CCATTTT
FHL1	protein_coding	INTRON	84.7	GAATCTG

NAP1L1	protein_coding	CDS	84.6	TATTTTA
GAPDH	protein_coding	CDS	84.6	TCGTATT
C6orf62	protein_coding	UTR3	84.6	CATTTTT
CCT7	protein_coding	CDS	84.4	CCAAACT
RPS16	protein_coding	CDS	84.4	CAGTATG
KIF14	protein_coding	UTR3	84.4	TTAAAGA
CORO1C	protein_coding	UTR3	84.0	CGTATTT
SLC39A7	protein_coding	UTR3	83.9	GTGGGAG
ALG9	protein_coding	INTRON	83.8	ACTCTTA
SEC14L1	protein_coding	INTRON	83.8	GGGTTTT
KTN1	protein_coding	CDS	83.6	AATTTTA
MFSD12	protein_coding	INTRON	83.5	AGCGAGC
PABPC1	protein_coding	CDS	83.5	GGTTATG
PTP4A1	protein_coding	UTR3	83.4	ATATTTT
EEF1D	protein_coding	INTRON	83.4	TTGGGTA
CALR	protein_coding	UTR3	83.3	TTCATTT
BSG	protein_coding	UTR3	83.3	AGCTCTG
TM4SF1	protein_coding	CDS	83.3	CATTCTC
ACTG1	protein_coding	UTR5	83.2	TTCTCTG
CHD8	protein_coding	INTRON	83.0	ATTTTGC
RTN4RL2	protein_coding	UTR3	83.0	CTCCTTG
NOP58	protein_coding	INTRON	82.9	TTCTTGG
KNL1	protein_coding	UTR3	82.8	TTTTTTA
P4HB	protein_coding	UTR3	82.6	ATTTTTG
DAB1	protein_coding	INTRON	82.6	ACTCCTC
HOMER1	protein_coding	INTRON	82.5	TGGATTT
TAF12	protein_coding	INTRON	82.4	CACCGCT
NACA	protein_coding	CDS	82.4	GAAAAGA
CBX4	protein_coding	UTR3	82.3	TTTCATT
BSG	protein_coding	CDS	82.3	GAGTACT
EIF3A	protein_coding	INTRON	82.3	ATTTCTT
B4GALT1	protein_coding	INTRON	82.2	CGAAGTT
RPS29	protein_coding	CDS	82.2	GACTAAA
LARP1	protein_coding	UTR3	82.2	GTGGGTA
HSPD1	protein_coding	CDS	82.1	GATTATT
HMGA1	protein_coding	UTR3	82.1	CACTTAG
GRM5	protein_coding	INTRON	82.1	AAATTGG
DPY19L1	protein_coding	UTR3	82.1	TTATCGT
B4GALT1	protein_coding	INTRON	82.1	CATTTTT
ILF3	protein_coding	INTRON	82.1	CCCTAGG
DDX17	protein_coding	UTR3	81.9	CATCATC
FUBP1	protein_coding	CDS	81.8	AACAGTG

RPL10	protein_coding	CDS	81.8	AAAAAGG
HNRNPU	protein_coding	CDS	81.7	GTGTTTT
ZC3H7B	protein_coding	UTR3	81.7	GTGGGGA
CDYL	protein_coding	UTR3	81.5	AACGTCA
EIF4G2	protein_coding	CDS	81.5	CAGAAGG
RPL4	protein_coding	CDS	81.5	CATCAGA
ZNF326	protein_coding	CDS	81.3	GTGGGAG
APLP2	protein_coding	CDS	81.3	TTATTGG
STOM	protein_coding	INTRON	81.3	TTTTTTA
ATP1B1	protein_coding	UTR3	81.3	TCTTTCC
FLNA	protein_coding	CDS	81.2	ACACATT
CCT7	protein_coding	CDS	81.1	ACCCAGT
DHX30	protein_coding	CDS	81.0	CCAGCCT
NAP1L1	protein_coding	CDS	80.9	GCAATTT
ACTG1	protein_coding	UTR3	80.8	CCTTTAG
TRAPPC3	protein_coding	UTR3	80.7	TGGGGAG
TPT1	protein_coding	UTR5	80.7	TGCACCG
CPNE1	protein_coding	INTRON	80.6	CCCAATG
GAPDH	protein_coding	UTR5	80.5	CGACAGT
LAMA1	protein_coding	CDS	80.4	GTGCCTG
PRDX1	protein_coding	UTR3	80.4	TGTATTT
ENSG00000260342	protein_coding	INTRON	80.4	CAGTTTT
ILRUN	protein_coding	UTR5	80.4	GGGGGGA
SLC37A3	protein_coding	INTRON	80.1	GGAATAC
CTNNB1	protein_coding	UTR3	80.0	GTAATCT
NSUN2	protein_coding	INTRON	79.9	GAGGGGC
HSPH1	protein_coding	CDS	79.9	GTATATT
ARIH1	protein_coding	INTRON	79.9	CATATTT
POR	protein_coding	CDS	79.8	GGCAAGG
HNRNPC	protein_coding	INTRON	79.8	ACCATTG
PTPRF	protein_coding	CDS	79.8	ATCCATG

In summary, these results confirm that KIFC1 is an RBP, that interacts predominantly with rRNA and protein-coding RNAs, and lacks sequence specificity when binding to RNA. The RNA affinity purification results further validated the iCLIP2-Seq results and importantly proved that AURKA, KIFC1 and TPX2 interacted with the same RNAs, reenforcing the assumption that they could be part of the same complex in mitosis, highlighting their cooperative function in mitosis.

4.2.7 KIFC1 - Yet another substrate of AURKA in mitosis

Since the interaction between AURKA and KIFC1 in mitosis was uncovered, I hypothesized that KIFC1 could be a substrate of AURKA, as suggested by earlier large-scale quantitative phosphoproteome analyses (174). AURKA is one of the major mitotic kinases, whose main function is to phosphorylate and regulate several mitotic factors for their effective activation and promotion of cell division (144). Nevertheless, the interaction between KIFC1 and AURKA remained so far uncharacterized. Hence, I tested this hypothesis with the help of published consensus sequence for AURKA (174,175).

I selected eight serine and threonine residues that were distributed across the KIFC1 sequence and localized in various domains such as disordered regions and the kinesin motor domain (Figure 28A, B). Later, KIFC1 WT protein was mutated at its shortlisted serine and threonine residues to alanine – a non-phosphorylatable residue – using site directed mutagenesis. Next, KIFC1 WT protein and the KIFC1 mutants were overexpressed with a flag-HA tag at their N-terminal. Later, the proteins were pulled down from HeLa cell lysates and an *in vitro* kinase assay was performed in the presence or absence of purified AURKA and [γ-32P] ATP. The assay disclosed two interesting residues in KIFC1: S349 and T359, whose mutation to alanine induced a strong reduction in the phosphorylation signal by AURKA (Figure 28C). The immunoprecipitation of flag-HA-KIFC1 WT and mutants was verified by western blotting as a control to demonstrate equal protein amounts in all the samples. Here, empty vector expressing the flag-HA tag was used as negative control (Figure 28D).

Further, a mutant version of KIFC1, which contains both mutations (S349A and T359A) was generated and the kinase assay was repeated using HeLa cells overexpressing the KIFC1 WT or the KIFC1 double mutant. The kinase assay results demonstrated a >90% loss of phosphorylation level for the KIFC1 double mutant (S349A/T359A) as compared to the KIFC1 WT protein (Figure 29A, E). Here, a purified kinase dead mutant of AURKA (D274A) was used to determine the specificity of KIFC1 phosphorylation by AURKA. The immunoprecipitation of flag-HA-KIFC1 and the presence of kinase dead and WT AURKA protein was verified by Coomassie staining to demonstrate equal protein amounts in all the samples (Figure 29B).

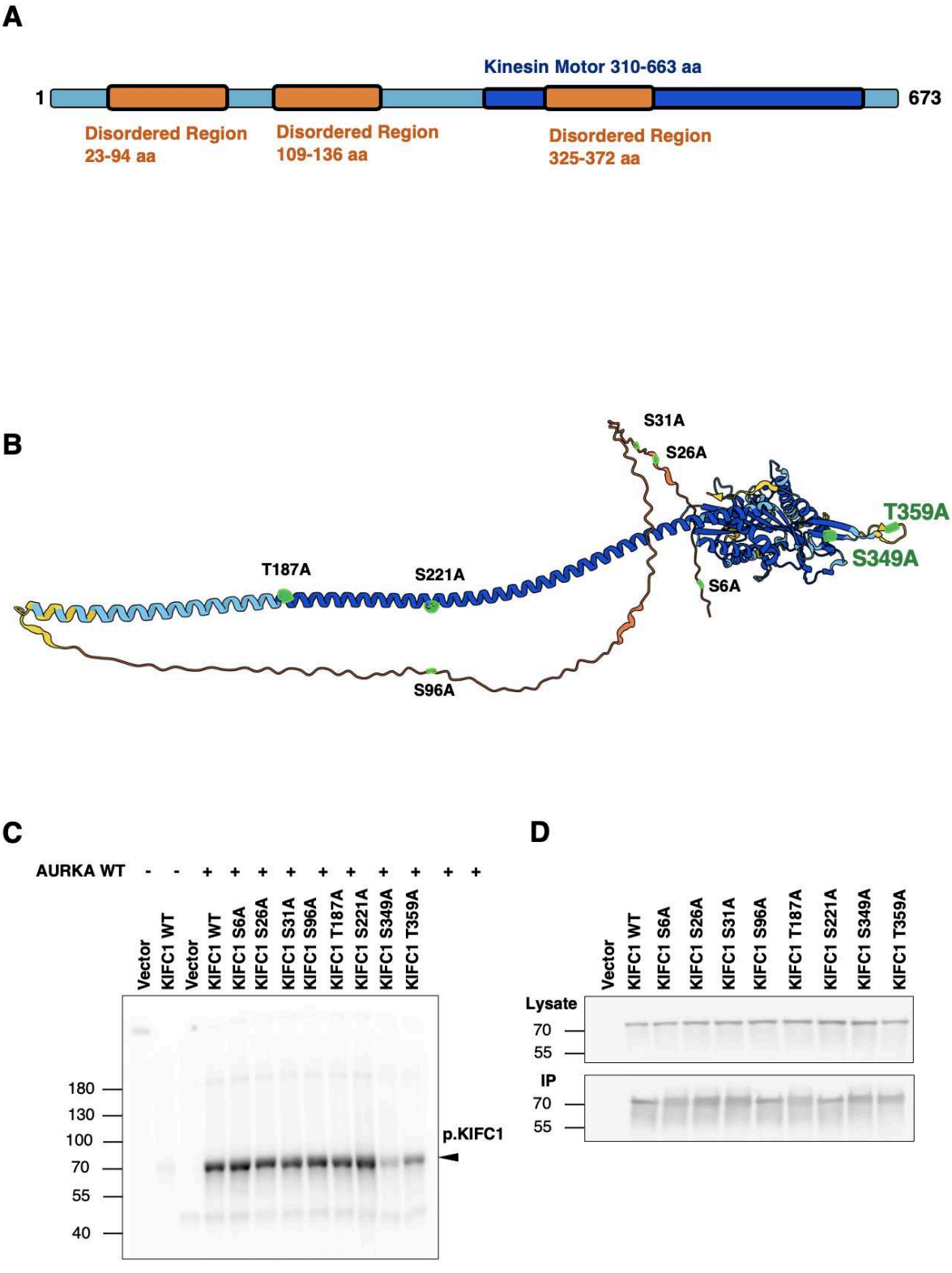


Figure 28: Kinase assay showing the phosphorylation of KIFC1 by AURKA

(A) Graphic depicting the distribution of the kinase domain and the disordered regions across the KIFC1 sequence.

(B) Schematic representation of KIFC1 structure highlighting the position of eight potential phosphorylation sites (green) that were identified based on the consensus sequence of AURKA. The schematic was generated using AlphaFold (alphafold.ebi.ac.uk).
(C) Autoradiography blot showing the phosphorylation level of KIFC1 wild-type (WT) and non-phosphorylatable KIFC1 mutants in the presence of purified AURKA WT and [γ-32P] ATP. The in vitro kinase assay was performed using KIFC1 pulled down with an N-terminal flag-HA tag. An empty vector was used as a negative control (n=3).
(D) The western blot depicting the immunoprecipitation efficiency of KIFC1 wild-type (WT) and non-phosphorylatable KIFC1 mutants (74 kDa) with N-terminal flag-HA tag, over expressed in HeLa cells. Empty vector plasmid containing flag-HA tag without protein was used as a negative control (n=3).
The figure was adapted from (116).

Since, the interaction of KIFC1 with AURKA is mediated by RNA, I next questioned if the phosphorylation could also be RNA dependent. Hence, I repeated the kinase assay with the cells overexpressing KIFC1 WT protein with N-terminal flag-HA tag. The cells were lysed, the lysates were treated with or without RNase and the KIFC1 WT protein was immunoprecipitated using anti-flag-HA beads. Later, the beads containing the pulled down proteins were washed to get rid of unspecific binding and incubated with purified AURKA WT or kinase dead AURKA D274A proteins and [γ-32P] ATP. The kinase assay demonstrated a 30% decrease in the phosphorylation signal of KIFC1 beads treated with RNase compared to the untreated KIFC1 beads (Figure 29C, F). The immunoprecipitation of flag-HA-KIFC1 and the presence of kinase dead and WT purified AURKA was verified by Coomassie staining as a control to demonstrate equal protein amounts in all the samples. Here, empty vector expressing the flag-HA tag was used as negative control (Figure 29 D). This proved that phosphorylation of KIFC1 by AURKA is RNA dependent.

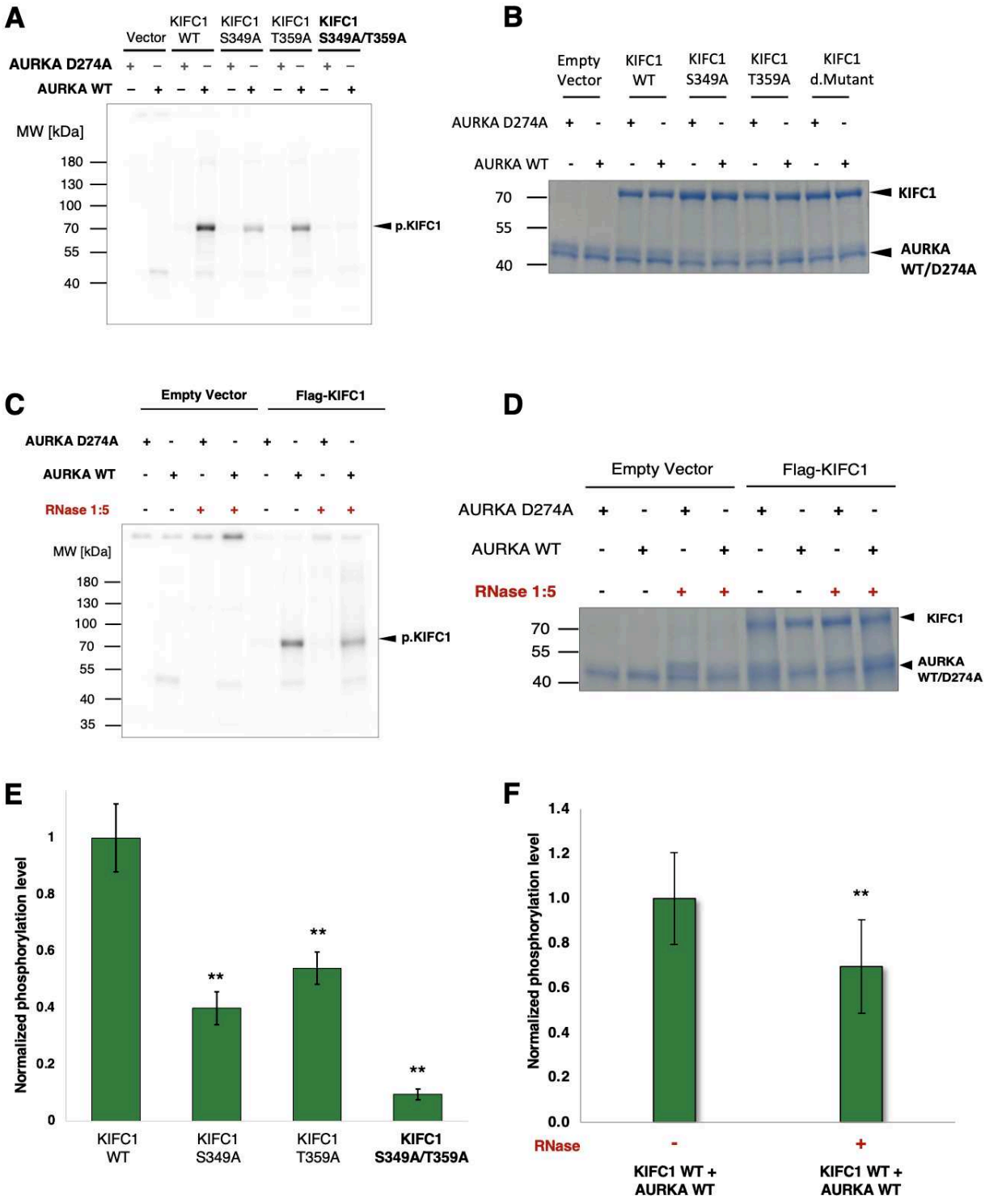


Figure 29: AURKA phosphorylates KIFC1 at S349 and T359 in an RNA-dependent manner

(A) Autoradiography blot representing the phosphorylation level of KIFC1 wild-type (WT) and non-phosphorylatable KIFC1 mutants (S349A, T359A and S349A/T359A) in the presence of purified AURKA WT or AURKA kinase-dead mutant (D274A). The in vitro kinase assay was performed using KIFC1 pulled down from HeLa prometaphase lysates overexpressing the WT or mutant KIFC1 proteins with an N-terminal flag-HA tag. An empty vector was used as a negative control.

(B) The Coomassie image depicting the protein amount of KIFC1 wild-type (WT) and non-phosphorylatable KIFC1 mutants (74 kDa) with N-terminal flag tag, over expressed in HeLa cells Also the image indicates the protein amount of AURKA WT and kinase dead AURKA D274A. HeLa cells transfected with empty vector plasmid containing flag-HA tag without expressing the protein was used as a negative control (n=3).
(C) Autoradiography blot representing the phosphorylation level of KIFC1 wild-type (WT) treated with or without RNase in the presence of purified AURKA WT or AURKA kinase-dead mutant (D274A). The in vitro kinase assay was performed using KIFC1 pulled down from HeLa prometaphase lysates overexpressing the WT or mutant KIFC1 proteins with an N-terminal flag-HA tag. An empty vector was used as a negative control (n=3).
(D) Same as (C) n=3.
*(E) Quantification of the autoradiography image (as in A), representing the KIFC1 phosphorylation level in the form of a bar graph with SEM (n=3). P-values were calculated using two-tailed, paired t-test (** P-value < 0.01).*
*(F) Quantification of the autoradiography image (as in C), representing the KIFC1 phosphorylation level in the presence or absence of RNA, in the form of a bar graph with SEM (n=3). P-values were calculated using two-tailed, paired t-test (** P-value < 0.01).*
The figure was adapted from (116).

Altogether, these results show that AURKA phosphorylates KIFC1 at amino acid residues S349 and T359. Importantly, the phosphorylation of KIFC1 by AURKA is RNA dependent which is a novel finding, that was never reported in previous studies.

In summary, in the second part of my PhD project, I discovered that three major mitotic factors AURKA, KIFC1 and TPX2 were RNA dependent, using the R-DeeP gradients from HeLa cells synchronized in prometaphase. Upon careful observation, I noticed that all the three proteins were detected in similar fractions in the control gradients, suggesting that, AURKA, KIFC1 and TPX2 are a part of the same complex. Later, I uncovered that, AURKA and KIFC1 were not just RNA dependent, also an RBP and the three mitotic proteins interacted with each other in an RNA-dependent manner. Though the interaction between AURKA and TPX2 was well characterized for decades by numerous researchers, the RNA-dependent interaction between these well-known mitotic partners was unknown. On the other hand, the interaction between AURKA and KIFC1 was uncharacterized and I identified a novel interaction that was RNA dependent in mitosis. Further, to identify the RNA candidates mediating the interaction between these mitotic factors, I performed iCLIP2-Seq of KIFC1 binding RNAs in HeLa cells synchronized in prometaphase. Upon, analysis it was clear that KIFC1 was predominantly bound to rRNAs and protein-coding RNAs with no specific motif enrichment. Though KIFC1 lacked sequence specificity, the top RNA targets bound to KIFC1 also interacted with AURKA and TPX2 in prometaphasic HeLa cell lysates. This further strengthened the hypothesis that AURKA, KIFC1 and TPX2 were a part of the same complex. The functional significance of interaction between AURKA and TPX2 is well studied, whereas the implications of AURKA-KIFC1 interaction is unknown. Hence, to understand the importance of their interaction, I first

performed an *in vitro* kinase assay and identified that AURKA phosphorylated KIFC1 at S349 and T359 aa residues. Since the interaction between AURKA and KIFC1 was RNA dependent, I repeated the *in vitro* kinase assay in the untreated cell lysates or cell lysates treated with RNase and I observed that the phosphorylation of KIFC1 by AURKA was RNA dependent. All these findings have generated a manuscript, which is already available for public knowledge in Biorxiv (116). Also, the manuscript is currently under revision in a well renowned peer reviewed journal Nature Communications.

This study identified that major mitotic regulators AURKA, KIFC1 and TPX2 are RBPs. This highlights that RNA is essential for the interaction of such vital proteins which is indispensable for spindle formation and faithful cell division. This highlights that RNAs and RBPs are crucial for mitosis. In future, this data could be utilised to identify the RNA-binding sites on these proteins and characterize the impact on mitosis when these proteins can no longer bind RNA. This could provide deeper knowledge on the molecular details of the role of RNA and RBPs in regulating mitosis. This study would set stage for further investigation into several well known and unknown mitotic factors and their role as an RBP in mitosis, adding another layer to an already complex, yet crucial biological pathway.

5 Discussion

The field of RNA biology faces a number of exciting new challenges related to the discovery and understanding of an increasing number of RBPs (25,56). Several new RBPs were uncovered using the established techniques that were mostly based on affinity purification, UV-crosslinking, RNA pulldown or organic phase separation that all bear their own advantages and limitations (30,41,46,49,50,53,176-180). Hence, there is still a strong requirement for new orthogonal and alternative strategies to identify RBPs or RNA-dependent proteins without any bias linked to enrichment procedures such as R-DeeP (59,60,116). The R-DeeP screen is based on the concept of RNA dependence, where a protein is classified as "RNA dependent" if its interactome is dependent on the presence of RNA (59,60,63). R-DeeP offers an independent approach devoid of potential biases from pulldown or separation based on physicochemical properties or the presence or absence of conventional RBDs (59,60,63). At the same time, R-DeeP also offers quantitative information on the RNA-dependent fraction of the protein. Using this screen, it is possible to identify proteins that are part of the same complexes based on the proteins shifting out of the similar control fractions (59,60,63). Although R-DeeP offers various advantages over other techniques, it does not provide any information about the RNA candidates or the RNA-binding sites on the protein, as the detected protein might bind directly or indirectly to RNA (59,60,63).

In the first part of my project, I took advantage of the R-DeeP strategy to investigate RNA-dependent proteins in lung cancer. Using this screening method, I identified 1189 RNA-dependent proteins including 170 novel RNA-dependent proteins that were not identified in any of the previous 43 human proteome-wide studies (63). However, 1894 proteins that were previously linked to RNA in at least one out of 43 human proteome-wide studies did not show significant RNA dependence. This could be due to the loss of weak interactions during cell lysis using detergents or the long centrifugation steps (60,63). There is also a possibility that some fraction of these RBPs could be false positives from the previous screens or due to the difference in the cell line models used for the experiment (60,63). This could be the case especially if they had been identified only in one study and if they were associated to a low RBP2GO score, as indicated by the analysis on the properties of RBPs in the shifting vs non-shifting protein groups (56,63). Additionally, the criteria for a shift required a minimum shifting distance of strictly more than one fraction. Hence, proteins with very short differences in their protein peaks between the control and RNase-treated fractions could have been missed (59,60,63).

The R-DeeP screen in A549 cells detected 3743 proteins in total, out of which I quantified 1525 left shifts and 241 right shifts. Left shifts toward lower molecular weights indicate the loss of interaction partners in the absence of RNA. One of the established functions of RNA is that, it acts as a scaffolding or docking platforms for proteins (22,26). In this screen, the cell lysates were treated with RNase leading to a loss of the RNA from the complex, which would result in decrease in the apparent molecular weight of the complex. Hence, I expected to see a majority of left shifts. In contrast, I also detected significant right shifts that indicated the gain of new protein interactors upon RNase treatment. In this scenario, this could be due to an increased accessibility of regions on the protein that are otherwise occupied by RNA. Additionally, the use of cellular extracts could also support artificial interactions between proteins that do not necessarily exist in the cell (63). However, the analysis of the RBP2GO score, enrichment of RBDs and IDRs between shifting vs. non-shifting proteins confirms a strong and specific enrichment of RNA-binding proteins in the shifting proteins (56,63). Here, the presence of RBDs and IDRs in the shifting proteins are considered as an important parameter for RNA binding as RBPs are majorly known to bind RNA through one or both the regions.

From the group of newly identified RNA-dependent proteins, I validated the RNA dependence of three proteins DOCK5, ELMO2 and ABRAXAS1, that are involved in cancer progression (132-139). Additionally, I demonstrated the RNA-binding capacity of DOCK5 and ABRAXAS1 in A549 cells using the iCLIP2 technique, further strengthening the results of the R-DeeP screen (63). Interestingly, I observed that DOCK5 shifted to its monomeric size after RNase treatment, whereas the two other proteins that were validated ELMO2 and ABRAXAS1 remained in a complex that was larger than their respective monomeric sizes (63). This suggests that after RNase treatment proteins could be completely discharged from an RNA-dependent complex like DOCK5 or they could still persist in a complex mediated through protein–protein interactions that are RNA independent or through RNA-independent oligomerization of the protein (63).

Furthermore, the complete datasets of RNA-dependent proteins from both HeLa S3 and A549 cells are now available in the R-DeeP 2.0 database (63). With both the datasets available to the community, it is possible to compare the RBPs between the cancer types. Also, with the lists of co-segregating proteins in control and RNase-treated fractions, it is possible to predict potential interaction between the proteins and reconstruct complexes. The database would serve as an effective tool for the community to gain information about the protein of interest, especially cancer targets that are overexpressed to understand the disease mechanism better.

In brief, In the first part of my PhD thesis, I identified 170 novel RNA-dependent proteins in lung cancer cells and validated the RNA dependence of three proteins involved in cancer progression. Additionally, I demonstrated the direct RNA binding of DOCK5 and ABRAXAS1 (63). With the remaining new shifting candidates in A549 cells, it opens up opportunities for future research to understand and discover new RNP complexes and their implications in lung cancer (63). Also, this data could be utilised to investigate the difference between the RBPome of healthy cells and lung cancer cells, which would shed light on the change in interactome of proteins leading to cancer. Along with the possibility to reconstruct the complexes based on co-segregation, it is possible to device a therapeutic strategy to target or interrupt signalling pathway that drives cancer progression.

The RBPs are classified based on the presence or absence of RBDs (25). The unconventional RBPs, which interact with RNA in absence of canonical RBDs, represent a large proportion (> 75%) of the RBPs in many species (20,21,181). Interestingly, the cellular functions of these unconventional RBPs are usually well-characterized but their interaction with RNA is often overlooked and thus remained unexplored. In a traditional view, RNA-protein interactions and dynamic RNP assemblies are mainly seen as a means for proteins to control the fate of cellular RNAs, from transcription to degradation. With the expanding concept of riboregulation, according to which RNA transcripts bind to RBPs and regulate their localization, conformation, interactions and function, this traditional view is being challenged (21,31,182,183), particularly in unexpected pathways such as cell division. Here, I consider cell division as a particular unexpected pathway as one might not anticipate major activity of RNA and RBPs due to the global transcriptional and translational repression during cell division (184,185).

In the second part of my project, I focused on characterization of such unconventional RBPs in mitosis. Due to the enrichment of mitosis-related terms in the GO analysis of the shifting proteins from the original R-DeeP screen in unsynchronized HeLa cells, I decided to focus on the mitotic factors that are RNA dependent (59). AURKA emerged as a strong RNA-dependent candidate with a clear left shift and similar shifting profile in the R-DeeP screen in unsynchronised HeLa cells as well as in the HeLa cells synchronised in mitosis. Due to the vital role of AURKA in mitosis as a key regulatory kinase that activates and drives various interacting proteins for spindle assembly and faithful cell division as well as its imperative role as an anti-cancer candidate, I decided to further investigate the RNA-dependent role of AURKA in mitosis (140,142-144,151,152,155). As indicated by the R-DeeP screens, AURKA is a novel RNA-dependent protein which can as well directly interact with RNA as indicated by the iCLIP2 results in mitotic HeLa and A549 cells. Interestingly, AURKA is also an

unconventional RBP that lacks specific RBDs. The functions and interactors of AURKA in mitosis has been well characterized for decades (140,142-145,151,152,154,155). Though the protein-protein interactions of AURKA in mitosis are investigated in detail, the perspective of RNA-dependent interactions are totally new to the mitotic community. Using the mass spectrometry analysis of AURKA immunoprecipitation samples in the presence and absence of RNase I found that 87% of AURKA interactors are RNA dependent. To be more specific, the interaction between AURKA and TPX2 have been thoroughly investigated. Certain studies have reported the interaction between AURKA and TPX2 at amino acid level along with its functional implications, binding residues and crystal structures (144,145,170,186,187). Yet, none of these studies have reported any involvement of RNA. Surprisingly, in this project, I identified TPX2 as an RNA-dependent interactor of AURKA which was not known before. Additionally, I also identified new uncharacterized interactions of AURKA such as its interaction with KIFC1, the motor protein involved in centrosome clustering (164,165). Since TPX2 and KIFC1 interacted with AURKA in an RNA-dependent fashion, I speculated whether these two interactors could also be RNA dependent. As expected, TPX2 and KIFC1 demonstrated an RNA-dependent shift similar to AURKA. Noteworthily, all the three proteins shifted towards earlier fractions upon RNase treatment from the same control fractions implying that AURKA, KIFC1 and TPX2 could be a part of same complex. It is well known that AURKA, TPX2 and their interactions are essential for MT nucleation and spindle assembly (144,145,186). Similar to TPX2, KIFC1 is a nuclear protein bound by importins. Upon nuclear envelop breakdown and RanGTP efflux from the chromosomes, KIFC1 is released from the importins (146,156,161,163,166). Interestingly, a recent study has demonstrated that excess soluble (non-MT) tubulin stimulates aster formation in HeLa cells overexpressing KIFC1 during mitosis (188). Additionally, KIFC1 is known to be involved in MT bundling, crosslinking and centrosome clustering to form bipolar spindle structure (156,163,166). Hence, it is plausible that AURKA, KIFC1 and TPX2 are part of the same complex that functions in mitosis for MT nucleation and bipolar spindle assembly. In line with this speculation, I observed that, KIFC1 not only interacted with AURKA, but also interacted with TPX2 throughout mitosis and more specifically in an RNA-dependent fashion, proving that AURKA, KIFC1 and TPX2 were RNA-dependent protein interactors, possibly, functioning together as a complex in mitosis. Here, RNA plays a crucial role in assembling these proteins together which is in turn, important for their respective phosphorylation and activation, including the phosphorylation of KIFC1 at S349 and T359 by AURKA. Though I identified phosphorylation sites of KIFC1 by AURKA, I did not observe so far any adverse mitotic phenotype upon

overexpression of this double phosphor-mutant KIFC1 in HeLa cells. This could be due to the fact that KIFC1 is essential for cells that go through acentrosomal cell division and cells with centrosome amplifications, to cluster and focus the spindles for forming two poles that is essential for bipolar cell division (156,163-165). Whereas, in the cells that contain centrosomes without centrosomal amplification, KIFC1 appears to be redundant (156,189). Since, HeLa cell do not contain centrosome amplification, it is not an optimal cell line model to study the mitotic defects caused by non-phosphorylatable KIFC1 mutant (156,189). In future, it is possible to characterize the role of KIFC1 non-phosphorylatable mutant in different cancer cell model that contain high centrosome amplifications such as MDA-MD-231 (167). On the other hand, I uncovered that AURKA phosphorylates KIFC1 in an RNA-dependent manner. The difference in the phosphorylation signal between the RNase-treated and RNase untreated samples are not as striking as the loss of the interactions between the proteins upon RNase treatment. This could be due to the reason that; the kinase reaction is performed using isolated proteins using pulldown and purified AURKA incubated in very small volumes. This creates an artificial close proximity between the proteins leading to relatively less difference in the phosphorylation signal in the presence or absence of RNA, which might not necessarily be the case in an intact cell. Nevertheless, I observed significant reduction (30%) in the phosphorylation signal in the samples treated with RNase compared to the untreated samples, showing that AURKA interacts and phosphorylates KIFC1 in an RNA-dependent fashion.

Given the central role of RNA in mediating complex formation, that contains major mitotic factors involved in spindle assembly, and knowing that AURKA and KIFC1 directly bound to RNA, I identified the KIFC1-interacting RNAs using iCLIP2-Seq (124,128). Following the analysis of the sequencing results, I found that, KIFC1 interacted predominantly with ribosomal RNAs and protein-coding transcripts, with a lack of specific binding site sequences. The interaction with such a huge amount of ribosomal RNAs was first misunderstood as contamination. However, recent publications have clearly highlighted the enrichment of rRNA around the mitotic chromosomes and their role in chromosome clustering (190-192). Moreover, rRNAs are also known to be associated with the mitotic spindles (193,194). Collectively, these data emphasize an essential role of rRNA in spindle assembly and mitotic progression, although the molecular mechanisms are yet to be uncovered. Similar to the rRNAs, mRNAs are also localized at the spindles, habitually the mRNAs of mitotic factors, coherent with the concept of spindle-localized protein synthesis (195-198). Likewise, a recent study also highlighted the role of RNAs in regulating protein localization to the mitotic spindle (199). Altogether, these previous findings confirm the localization of rRNA and mRNA species at the mitotic spindles

suggesting the interaction of these RNAs with mitotic factors, as indicated from our KIFC1 iCLIP2 data. Further, the RNA affinity purification showed that all three proteins, KIFC1, AURKA and TPX2 interacted with similar RNAs, further validating that these three proteins were part of same complex, whose formation is mediated by RNA.

Though KIFC1 lacks sequence specificity for binding to RNA, all the three proteins AURKA, KIFC1 and TPX2 contain high isoelectric points. This could be due to their positively charged amino acids, which might promote their interaction with RNA based on mere charge-based attraction. Additionally, these proteins contain IDRs, which might support RNA binding by possibly stabilizing their structure via a disorder-to-order transition. This in turn could be crucial for their interaction with other proteins or RNAs (200,201). While the RNA-binding characteristics of such proteins are not well understood, a recent study reported the lack of sequence specificity for 492 investigated unconventional RBPs and suggested that their identification in RNA interactome studies could occur via weak non-sequence-specific interactions with RNA (202). This correlates well with the KIFC1 iCLIP2-sequencing results, demonstrating the lack of sequence specificity for RNA binding. However, further investigations on the RNA-protein interactions for such unconventional RBPs might help understanding the driving mechanisms that might not be simply limited to sequence characteristics. With time, the functions of RNA have been increasingly clear specially in the context of protein regulation (183). Numerous studies have investigated the role of single regulatory RNA: protein interaction followed by mechanism by which phenotypic effects are perpetrated (183). This fuels an expectation on how regulatory mechanisms should function. Though it is accurate in many instances, it is also important to consider that certain regulatory functions are implemented by collective group of RNAs rather than one single transcript (183). One such example would be the regulation of enolase-1 enzyme. Huppertz *et al.,* reported that enolase-1 bound about 2000 different RNA transcripts, that collectively inhibited the enzyme activity, thereby altering glycolytic metabolite concentrations during stem cell differentiation (31,183). This study shed light on the concept and coined the term ”crowd-control” which basically reflects the regulation of protein functions via the collective action of several RNAs (“crowd”) and is characterized by extreme redundancy of the individual regulators (31,183). This concept can hold true for several RBPs at several instances (31,183). Hence, it is conceivable that multiple RNA transcripts collectively “crowd-control” the localization and interactions of AURKA, KIFC1 and TPX2. This view may be also expanded to posttranslational modifications such as phosphorylation, mediated by RNA, that is required for the activation of proteins, as observed in the case of KIFC1 phosphorylation by AURKA.

In summary, in the second part of the project, I demonstrated the RNA dependence of major mitotic factors AURKA, KIFC1 and TPX2 (116). The detailed analysis on the RNA dependence of AURKA, TPX2 and KIFC1 discloses the importance of RNA in riboregulating mitotic protein-protein interactions and its role in regulating posttranslational modifications such as phosphorylation (116). This study provides enormous scope to understand such RNA mediated interactions and its functional implications in various subcellular structures and vital pathways.

Altogether, during my PhD, I focussed on identifying and characterizing the role of RNA-dependent proteins in cancer using R-DeeP screen (63). Importantly, I highlighted the role of RNA in mediating protein-protein interactions that are crucial for the spindle assembly and bipolar cell division (116). This, study serves as pioneer in deciphering the role of RNA-protein interactions and their implications in cellular pathways. Thus, opening up new research avenues that integrates RNA-protein interactions as a new layer of regulation in our understanding of cell division and disease mechanisms. Given that AURKA, KIFC1 and TPX2 are overexpressed in cancer, in future with additional experiments, their expression could be compared to their RNA-binding activity. By mutating the RNA-binding residues in these proteins, it is possible to study the implications of such mutants in the context of mitosis and cancer. Since the interaction between AURKA, KIFC1 and TPX2 are RNA dependent, it possible to investigate further on mitotic phenotypes or therapy response to cancer drugs when the interactions between these proteins and RNAs are interrupted. On the whole, these findings bridge the field of RNA biology to the field of cell division and cancer and provides a new perspective on understanding complex mechanisms of cell proliferation. Importantly, it promises the development of new therapeutic strategies in future, that focuses not only on targeting a single protein, but also protein-RNA interactions which could be the trigger point for RBP dysregulation in cancer.

Both the parts of my PhD project have resulted in a manuscript. My first manuscript “Proteome-Wide Identification of RNA-Dependent Proteins in Lung Cancer Cells” was published in the journal *Cancers* on 12 December 2022 (63). My second manuscript “An atlas of RNA-dependent proteins in cell division reveals the riboregulation of mitotic protein-protein interactions” is available for the community on bioRxiv since 26 September 2024 and is currently under revision in the journal *Nature Communications (116)*.

6 Acknowledgements

I take this opportunity to thank all the people who helped me through my PhD journey. It wouldn't be possible without the help of such great people. First, I would like to thank PD. Dr. Maiwen Caudron-Herger and Prof. Dr. Sven Diederichs for giving me the opportunity to work in their labs. I am grateful for their support throughout this project, specially through the difficult phases and for the freedom to explore different avenues of the project as a researcher. Next, I would like to thank Prof. Dr. Gislene Pereira and Prof. Dr. Oliver Gruß for their valuable feedback and support as members of my thesis committee.

Importantly, I would like to thank all present and former B150 members with whom I had the chance to work with over the years, for their immense support. A special gratitude to Jeanette Seiler, who not only was my project partner, but was also a great mentor and greatest support through my journey as a PhD student. I also would like to thank Jana Theiß and Dr. Simona Cantarella, who helped me with a lot of experiments and analysis which were crucial for the project. I also would like to thank all my collaborators, especially within the DKFZ the Light microscopy core facility, Manuela Brom and Dr. Felix Bestvater. Collaborators from Proteomic core facility Dr. Dominic Helm and Martin Schneider. And our collaborator from the Cellular tools core facility Dr. Rainer Will.

I am especially grateful to my parents, grandparents and my husband Ganesh for the motivation, support and for trusting me in career choices, without whom nothing would have been possible. I take this opportunity to extend my gratitude to my parents in-law and our entire family for their patience and belief in me. I also thank my House owner Frau. Elisabeth Grün and my friends at home and abroad for their greatest help and understanding throughout my journey.

Most importantly, I submit all of my accomplishments in the feet of my Guru: Sri Satchidananda Sadguru Sainath Maharaj.

"All I know is that I know nothing"- Socrates

7 References

1. Morris, K.V. and Mattick, J.S. (2014) The rise of regulatory RNA. *Nat Rev Genet*, **15**, 423-437.
2. Spitale, R.C. and Incarnato, D. (2023) Probing the dynamic RNA structurome and its functions. *Nat Rev Genet*, **24**, 178-196.
3. Nam, J.W., Choi, S.W. and You, B.H. (2016) Incredible RNA: Dual Functions of Coding and Noncoding. *Mol Cells*, **39**, 367-374.
4. Candeias, M.M., Malbert-Colas, L., Powell, D.J., Daskalogianni, C., Maslon, M.M., Naski, N., Bourougaa, K., Calvo, F. and Fahraeus, R. (2008) P53 mRNA controls p53 activity by managing Mdm2 functions. *Nat Cell Biol*, **10**, 1098-1105.
5. Hombach, S. and Kretz, M. (2016) Non-coding RNAs: Classification, Biology and Functioning. *Adv Exp Med Biol*, **937**, 3-17.
6. Li, J. and Liu, C. (2019) Coding or Noncoding, the Converging Concepts of RNAs. *Front Genet*, **10**, 496.
7. Bartel, D.P. (2009) MicroRNAs: target recognition and regulatory functions. *Cell*, **136**, 215-233.
8. Iwasaki, Y.W., Siomi, M.C. and Siomi, H. (2015) PIWI-Interacting RNA: Its Biogenesis and Functions. *Annu Rev Biochem*, **84**, 405-433.
9. Gutschner, T. and Diederichs, S. (2012) The hallmarks of cancer: a long non-coding RNA point of view. *RNA Biol*, **9**, 703-719.
10. Mattick, J.S., Amaral, P.P., Carninci, P., Carpenter, S., Chang, H.Y., Chen, L.L., Chen, R., Dean, C., Dinger, M.E., Fitzgerald, K.A. *et al.* (2023) Long non-coding RNAs: definitions, functions, challenges and recommendations. *Nat Rev Mol Cell Biol*, **24**, 430-447.
11. Gutschner, T., Hämmerle, M. and Diederichs, S. (2013) MALAT1 — a paradigm for long noncoding RNA function in cancer. *Journal of Molecular Medicine*, **91**, 791-801.
12. Gutschner, T., Hammerle, M., Eissmann, M., Hsu, J., Kim, Y., Hung, G., Revenko, A., Arun, G., Stentrup, M., Gross, M. *et al.* (2013) The noncoding RNA MALAT1 is a critical regulator of the metastasis phenotype of lung cancer cells. *Cancer Res*, **73**, 1180-1189.
13. Rupaimoole, R. and Slack, F.J. (2017) MicroRNA therapeutics: towards a new era for the management of cancer and other diseases. *Nat Rev Drug Discov*, **16**, 203-222.

14. Chen, Y., Li, Z., Chen, X. and Zhang, S. (2021) Long non-coding RNAs: From disease code to drug role. *Acta Pharm Sin B*, **11**, 340-354.
15. Esquela-Kerscher, A. and Slack, F.J. (2006) Oncomirs - microRNAs with a role in cancer. *Nat Rev Cancer*, **6**, 259-269.
16. Zhu, X., Wu, Y.B., Zhou, J. and Kang, D.M. (2016) Upregulation of lncRNA MEG3 promotes hepatic insulin resistance via increasing FoxO1 expression. *Biochem Biophys Res Commun*, **469**, 319-325.
17. Cheng, Y. and Zhang, C. (2010) MicroRNA-21 in cardiovascular disease. *J Cardiovasc Transl Res*, **3**, 251-255.
18. Tokumaru, S., Suzuki, M., Yamada, H., Nagino, M. and Takahashi, T. (2008) let-7 regulates Dicer expression and constitutes a negative feedback loop. *Carcinogenesis*, **29**, 2073-2077.
19. Cech, T.R. and Steitz, J.A. (2014) The noncoding RNA revolution-trashing old rules to forge new ones. *Cell*, **157**, 77-94.
20. Gebauer, F., Schwarzl, T., Valcarcel, J. and Hentze, M.W. (2021) RNA-binding proteins in human genetic disease. *Nat Rev Genet*, **22**, 185-198.
21. Hentze, M.W., Castello, A., Schwarzl, T. and Preiss, T. (2018) A brave new world of RNA-binding proteins. *Nat Rev Mol Cell Biol*, **19**, 327-341.
22. Corley, M. (2020) How RNA-Binding Proteins Interact with RNA: Molecules and Mechanisms. *Molecular Cell*, **78**, 9-29.
23. Dreyfuss, G., Kim, V.N. and Kataoka, N. (2002) Messenger-RNA-binding proteins and the messages they carry. *Nat Rev Mol Cell Biol*, **3**, 195-205.
24. Moore, M.J. (2005) From birth to death: the complex lives of eukaryotic mRNAs. *Science*, **309**, 1514-1518.
25. Wassmer, E., Koppany, G., Hermes, M., Diederichs, S. and Caudron-Herger, M. (2024) Refining the pool of RNA-binding domains advances the classification and prediction of RNA-binding proteins. *Nucleic Acids Res*, **52**, 7504–7522.
26. Lunde, B.M., Moore, C. and Varani, G. (2007) RNA-binding proteins: modular design for efficient function. *Nat Rev Mol Cell Biol*, **8**, 479-490.
27. Saldana-Meyer, R., Gonzalez-Buendia, E., Guerrero, G., Narendra, V., Bonasio, R., Recillas-Targa, F. and Reinberg, D. (2014) CTCF regulates the human p53 gene through direct interaction with its natural antisense transcript, Wrap53. *Genes Dev*, **28**, 723-734.
28. Saldana-Meyer, R., Rodriguez-Hernaez, J., Escobar, T., Nishana, M., Jacome-Lopez, K., Nora, E.P., Bruneau, B.G., Tsirigos, A., Furlan-Magaril, M., Skok, J. *et al.* (2019)

RNA Interactions Are Essential for CTCF-Mediated Genome Organization. *Mol Cell*, **76**, 412-422 e415.

29. Hansen, A.S., Hsieh, T.S., Cattoglio, C., Pustova, I., Saldana-Meyer, R., Reinberg, D., Darzacq, X. and Tjian, R. (2019) Distinct Classes of Chromatin Loops Revealed by Deletion of an RNA-Binding Region in CTCF. *Mol Cell*, **76**, 395-411 e313.
30. Castello, A., Fischer, B., Frese, C.K., Horos, R., Alleaume, A.M., Foehr, S., Curk, T., Krijgsveld, J. and Hentze, M.W. (2016) Comprehensive Identification of RNA-Binding Domains in Human Cells. *Mol Cell*, **63**, 696-710.
31. Huppertz, I., Perez-Perri, J.I., Mantas, P., Sekaran, T., Schwarzl, T., Russo, F., Ferring-Appel, D., Koskova, Z., Dimitrova-Paternoga, L., Kafkia, E. *et al.* (2022) Riboregulation of Enolase 1 activity controls glycolysis and embryonic stem cell differentiation. *Molecular Cell*, **82**, 2666-2680.e2611.
32. Lukong, K.E., Chang, K.W., Khandjian, E.W. and Richard, S. (2008) RNA-binding proteins in human genetic disease. *Trends Genet*, **24**, 416-425.
33. Castello, A., Fischer, B., Hentze, M.W. and Preiss, T. (2013) RNA-binding proteins in Mendelian disease. *Trends Genet*, **29**, 318-327.
34. Wang, X., Liu, R., Zhu, W., Chu, H., Yu, H., Wei, P., Wu, X., Zhu, H., Gao, H., Liang, J. *et al.* (2019) UDP-glucose accelerates SNAI1 mRNA decay and impairs lung cancer metastasis. *Nature*, **571**, 127-131.
35. Greenberg, J.R. (1979) Ultraviolet light-induced crosslinking of mRNA to proteins. *Nucleic Acids Res*, **6**, 715-732.
36. Wagenmakers, A.J., Reinders, R.J. and van Venrooij, W.J. (1980) Cross-linking of mRNA to proteins by irradiation of intact cells with ultraviolet light. *Eur J Biochem*, **112**, 323-330.
37. Lindberg, U. and Sundquist, B. (1974) Isolation of messenger ribonucleoproteins from mammalian cells. *J Mol Biol*, **86**, 451-468.
38. Adam, S.A., Nakagawa, T., Swanson, M.S., Woodruff, T.K. and Dreyfuss, G. (1986) mRNA polyadenylate-binding protein: gene isolation and sequencing and identification of a ribonucleoprotein consensus sequence. *Mol Cell Biol*, **6**, 2932-2943.
39. Choi, Y.D. and Dreyfuss, G. (1984) Isolation of the heterogeneous nuclear RNA-ribonucleoprotein complex (hnRNP): a unique supramolecular assembly. *Proc Natl Acad Sci U S A*, **81**, 7471-7475.
40. Hafner, M., Landthaler, M., Burger, L., Khorshid, M., Hausser, J., Berninger, P., Rothballer, A., Ascano, M., Jr., Jungkamp, A.C., Munschauer, M. *et al.* (2010)

Transcriptome-wide identification of RNA-binding protein and microRNA target sites by PAR-CLIP. *Cell*, **141**, 129-141.

41. Baltz, A.G., Munschauer, M., Schwanhausser, B., Vasile, A., Murakawa, Y., Schueler, M., Youngs, N., Penfold-Brown, D., Drew, K., Milek, M. *et al.* (2012) The mRNA-bound proteome and its global occupancy profile on protein-coding transcripts. *Mol Cell*, **46**, 674-690.
42. Conrad, T., Albrecht, A.-S., Costa, V.R.d.M., Sauer, S., Meierhofer, D. and Ørom, U.A. (2016) Serial interactome capture of the human cell nucleus. *Nature Communications*, **7**, 11212.
43. Kwon, S.C., Yi, H., Eichelbaum, K., Fohr, S., Fischer, B., You, K.T., Castello, A., Krijgsveld, J., Hentze, M.W. and Kim, V.N. (2013) The RNA-binding protein repertoire of embryonic stem cells. *Nat Struct Mol Biol*, **20**, 1122-1130.
44. Perez-Perri, J.I., Rogell, B., Schwarzl, T., Stein, F., Zhou, Y., Rettel, M., Brosig, A. and Hentze, M.W. (2018) Discovery of RNA-binding proteins and characterization of their dynamic responses by enhanced RNA interactome capture. *Nat Commun*, **9**, 4408.
45. Beckmann, B.M., Horos, R., Fischer, B., Castello, A., Eichelbaum, K., Alleaume, A.-M., Schwarzl, T., Curk, T., Foehr, S., Huber, W. *et al.* (2015) The RNA-binding proteomes from yeast to man harbour conserved enigmRBPs. *Nature Communications*, **6**, 10127.
46. Castello, A., Fischer, B., Eichelbaum, K., Horos, R., Beckmann, B.M., Strein, C., Davey, N.E., Humphreys, D.T., Preiss, T., Steinmetz, L.M. *et al.* (2012) Insights into RNA biology from an atlas of mammalian mRNA-binding proteins. *Cell*, **149**, 1393-1406.
47. Liao, Y., Castello, A., Fischer, B., Leicht, S., Foehr, S., Frese, C.K., Ragan, C., Kurscheid, S., Pagler, E., Yang, H. *et al.* (2016) The Cardiomyocyte RNA-Binding Proteome: Links to Intermediary Metabolism and Heart Disease. *Cell Rep*, **16**, 1456-1469.
48. Liepelt, A., Vries, I.S.N.-d., Simons, N., Eichelbaum, K., Föhr, S., Archer, S.K., Castello, A., Usadel, B., Krijgsveld, J., Preiss, T. *et al.* (2016) Identification of RNA-binding Proteins in Macrophages by Interactome Capture. *Molecular & Cellular Proteomics*, **15**, 2699-2714.
49. Urdaneta, E.C., Vieira-Vieira, C.H., Hick, T., Wessels, H.-H., Figini, D., Moschall, R., Medenbach, J., Ohler, U., Granneman, S., Selbach, M. *et al.* (2019) Purification of

cross-linked RNA-protein complexes by phenol-toluol extraction. *Nature Communications*, **10**, 990.

50. Trendel, J., Schwarzl, T., Horos, R., Prakash, A., Bateman, A., Hentze, M.W. and Krijgsveld, J. (2019) The Human RNA-Binding Proteome and Its Dynamics during Translational Arrest. *Cell*, **176**, 391-403 e319.
51. Queiroz, R.M.L., Smith, T., Villanueva, E., Marti-Solano, M., Monti, M., Pizzinga, M., Mirea, D.-M., Ramakrishna, M., Harvey, R.F., Dezi, V. *et al.* (2019) Comprehensive identification of RNA–protein interactions in any organism using orthogonal organic phase separation (OOPS). *Nature Biotechnology*, **37**, 169-178.
52. Urdaneta, E.C., Vieira-Vieira, C.H., Hick, T., Wessels, H.H., Figini, D., Moschall, R., Medenbach, J., Ohler, U., Granneman, S., Selbach, M. *et al.* (2019) Purification of cross-linked RNA-protein complexes by phenol-toluol extraction. *Nat Commun*, **10**, 990.
53. Bao, X., Guo, X., Yin, M., Tariq, M., Lai, Y., Kanwal, S., Zhou, J., Li, N., Lv, Y., Pulido-Quetglas, C. *et al.* (2018) Capturing the interactome of newly transcribed RNA. *Nat Methods*, **15**, 213-220.
54. Huang, R., Han, M., Meng, L. and Chen, X. (2018) Capture and Identification of RNA-binding Proteins by Using Click Chemistry-assisted RNA-interactome Capture (CARIC) Strategy. *J Vis Exp*, **115**, E3879–E3887.
55. Brannan, K.W., Jin, W., Huelga, S.C., Banks, C.A., Gilmore, J.M., Florens, L., Washburn, M.P., Van Nostrand, E.L., Pratt, G.A., Schwinn, M.K. *et al.* (2016) SONAR Discovers RNA-Binding Proteins from Analysis of Large-Scale Protein-Protein Interactomes. *Mol Cell*, **64**, 282-293.
56. Caudron-Herger, M., Jansen, R.E., Wassmer, E. and Diederichs, S. (2021) RBP2GO: a comprehensive pan-species database on RNA-binding proteins, their interactions and functions. *Nucleic Acids Res*, **49**, D425-D436.
57. Ray, D., Kazan, H., Cook, K.B., Weirauch, M.T., Najafabadi, H.S., Li, X., Gueroussov, S., Albu, M., Zheng, H., Yang, A. *et al.* (2013) A compendium of RNA-binding motifs for decoding gene regulation. *Nature*, **499**, 172-177.
58. Cook, K.B., Kazan, H., Zuberi, K., Morris, Q. and Hughes, T.R. (2011) RBPDB: a database of RNA-binding specificities. *Nucleic Acids Res*, **39**, D301-308.
59. Caudron-Herger, M., Rusin, S.F., Adamo, M.E., Seiler, J., Schmid, V.K., Barreau, E., Kettenbach, A.N. and Diederichs, S. (2019) R-DeeP: Proteome-wide and Quantitative

Identification of RNA-Dependent Proteins by Density Gradient Ultracentrifugation. *Mol Cell*, **75**, 184-199 e110.

60. Caudron-Herger, M., Wassmer, E., Nasa, I., Schultz, A.S., Seiler, J., Kettenbach, A.N. and Diederichs, S. (2020) Identification, quantification and bioinformatic analysis of RNA-dependent proteins by RNase treatment and density gradient ultracentrifugation using R-DeeP. *Nat Protoc*, **15**, 1338-1370.
61. Mikulits, W., Pradet-Balade, B., Habermann, B., Beug, H., Garcia-Sanz, J.A. and Mullner, E.W. (2000) Isolation of translationally controlled mRNAs by differential screening. *FASEB J*, **14**, 1641-1652.
62. Hock, J., Weinmann, L., Ender, C., Rudel, S., Kremmer, E., Raabe, M., Urlaub, H. and Meister, G. (2007) Proteomic and functional analysis of Argonaute-containing mRNA-protein complexes in human cells. *EMBO Rep*, **8**, 1052-1060.
63. Rajagopal, V., Loubal, A.S., Engel, N., Wassmer, E., Seiler, J., Schilling, O., Caudron-Herger, M. and Diederichs, S. (2022) Proteome-Wide Identification of RNA-Dependent Proteins in Lung Cancer Cells. *Cancers (Basel)*, **14**, 6109.
64. Hollin, T., Abel, S., Banks, C., Hristov, B., Prudhomme, J., Hales, K., Florens, L., Stafford Noble, W. and Le Roch, K.G. (2024) Proteome-Wide Identification of RNA-dependent proteins and an emerging role for RNAs in Plasmodium falciparum protein complexes. *Nat Commun*, **15**, 1365.
65. Schafer, K.A. (1998) The cell cycle: a review. *Vet Pathol*, **35**, 461-478.
66. Vermeulen, K., Van Bockstaele, D.R. and Berneman, Z.N. (2003) The cell cycle: a review of regulation, deregulation and therapeutic targets in cancer. *Cell Prolif*, **36**, 131-149.
67. Salazar-Roa, M. and Malumbres, M. (2017) Fueling the Cell Division Cycle. *Trends Cell Biol*, **27**, 69-81.
68. Norbury, C. and Nurse, P. (1992) Animal cell cycles and their control. *Annu Rev Biochem*, **61**, 441-470.
69. Matthews, H.K., Bertoli, C. and de Bruin, R.A.M. (2022) Cell cycle control in cancer. *Nat Rev Mol Cell Biol*, **23**, 74-88.
70. Limas, J.C. and Cook, J.G. (2019) Preparation for DNA replication: the key to a successful S phase. *FEBS Lett*, **593**, 2853-2867.
71. Lockhead, S., Moskaleva, A., Kamenz, J., Chen, Y., Kang, M., Reddy, A.R., Santos, S.D.M. and Ferrell, J.E., Jr. (2020) The Apparent Requirement for Protein Synthesis during G2 Phase Is due to Checkpoint Activation. *Cell Rep*, **32**, 107901.

72. Malumbres, M. and Barbacid, M. (2009) Cell cycle, CDKs and cancer: a changing paradigm. *Nat Rev Cancer*, **9**, 153-166.
73. Ong, J.Y. and Torres, J.Z. (2019) Dissecting the mechanisms of cell division. *J Biol Chem*, **294**, 11382-11390.
74. McIntosh, J.R. (2016) Mitosis. *Cold Spring Harb Perspect Biol*, **8**, a023218.
75. Batty, P. and Gerlich, D.W. (2019) Mitotic Chromosome Mechanics: How Cells Segregate Their Genome. *Trends Cell Biol*, **29**, 717-726.
76. Ferreira, L.T. and Maiato, H. (2021) Prometaphase. *Semin Cell Dev Biol*, **117**, 52-61.
77. Carlton, J.G., Jones, H. and Eggert, U.S. (2020) Membrane and organelle dynamics during cell division. *Nat Rev Mol Cell Biol*, **21**, 151-166.
78. Scholey, J.M., Rogers, G.C. and Sharp, D.J. (2001) Mitosis, microtubules, and the matrix. *J Cell Biol*, **154**, 261-266.
79. Mortuza, G.B., Cavazza, T., Garcia-Mayoral, M.F., Hermida, D., Peset, I., Pedrero, J.G., Merino, N., Blanco, F.J., Lyngso, J., Bruix, M. *et al.* (2014) XTACC3-XMAP215 association reveals an asymmetric interaction promoting microtubule elongation. *Nat Commun*, **5**, 5072.
80. Telzer, B.R., Moses, M.J. and Rosenbaum, J.L. (1975) Assembly of microtubules onto kinetochores of isolated mitotic chromosomes of HeLa cells. *Proc Natl Acad Sci U S A*, **72**, 4023-4027.
81. Petry, S. and Vale, R.D. (2015) Microtubule nucleation at the centrosome and beyond. *Nat Cell Biol*, **17**, 1089-1093.
82. Petry, S., Groen, A.C., Ishihara, K., Mitchison, T.J. and Vale, R.D. (2013) Branching microtubule nucleation in Xenopus egg extracts mediated by augmin and TPX2. *Cell*, **152**, 768-777.
83. Heald, R., Tournebize, R., Blank, T., Sandaltzopoulos, R., Becker, P., Hyman, A. and Karsenti, E. (1996) Self-organization of microtubules into bipolar spindles around artificial chromosomes in Xenopus egg extracts. *Nature*, **382**, 420-425.
84. Gadde, S. and Heald, R. (2004) Mechanisms and molecules of the mitotic spindle. *Curr Biol*, **14**, R797-805.
85. Vukusic, K. and Tolic, I.M. (2021) Anaphase B: Long-standing models meet new concepts. *Semin Cell Dev Biol*, **117**, 127-139.
86. Asbury, C.L. (2017) Anaphase A: Disassembling Microtubules Move Chromosomes toward Spindle Poles. *Biology (Basel)*, **6**, 15.

87. Vagnarelli, P. (2021) Back to the new beginning: Mitotic exit in space and time. *Semin Cell Dev Biol*, **117**, 140-148.
88. Skibbens, R.V., Skeen, V.P. and Salmon, E.D. (1993) Directional instability of kinetochore motility during chromosome congression and segregation in mitotic newt lung cells: a push-pull mechanism. *J Cell Biol*, **122**, 859-875.
89. Vukusic, K., Buda, R. and Tolic, I.M. (2019) Force-generating mechanisms of anaphase in human cells. *J Cell Sci*, **132**, jcs231985.
90. Afonso, O., Matos, I., Pereira, A.J., Aguiar, P., Lampson, M.A. and Maiato, H. (2014) Feedback control of chromosome separation by a midzone Aurora B gradient. *Science*, **345**, 332-336.
91. Guttinger, S., Laurell, E. and Kutay, U. (2009) Orchestrating nuclear envelope disassembly and reassembly during mitosis. *Nat Rev Mol Cell Biol*, **10**, 178-191.
92. Pollard, T.D. and O'Shaughnessy, B. (2019) Molecular Mechanism of Cytokinesis. *Annu Rev Biochem*, **88**, 661-689.
93. Carmena, M., Wheelock, M., Funabiki, H. and Earnshaw, W.C. (2012) The chromosomal passenger complex (CPC): from easy rider to the godfather of mitosis. *Nat Rev Mol Cell Biol*, **13**, 789-803.
94. Liu, X., Chen, Y., Li, Y., Petersen, R.B. and Huang, K. (2019) Targeting mitosis exit: A brake for cancer cell proliferation. *Biochim Biophys Acta Rev Cancer*, **1871**, 179-191.
95. Matellan, L. and Monje-Casas, F. (2020) Regulation of Mitotic Exit by Cell Cycle Checkpoints: Lessons From Saccharomyces cerevisiae. *Genes (Basel)*, **11**, 195.
96. Fouad, Y.A. and Aanei, C. (2017) Revisiting the hallmarks of cancer. *Am J Cancer Res*, **7**, 1016-1036.
97. Hanahan, D. and Weinberg, R.A. (2000) The hallmarks of cancer. *Cell*, **100**, 57-70.
98. Hanahan, D. and Weinberg, R.A. (2011) Hallmarks of cancer: the next generation. *Cell*, **144**, 646-674.
99. Hanahan, D. (2022) Hallmarks of Cancer: New Dimensions. *Cancer Discov*, **12**, 31-46.
100. J, F. (2020) Global Cancer Observatory: Cancer Today. Lyon: International Agency for Research on Cancer.
101. Liu, B., Zhou, H., Tan, L., Siu, K.T.H. and Guan, X.Y. (2024) Exploring treatment options in cancer: Tumor treatment strategies. *Signal Transduct Target Ther*, **9**, 175.
102. Debela, D.T., Muzazu, S.G., Heraro, K.D., Ndalama, M.T., Mesele, B.W., Haile, D.C., Kitui, S.K. and Manyazewal, T. (2021) New approaches and procedures for cancer treatment: Current perspectives. *SAGE Open Med*, **9**, 20503121211034366.

103. Frankel, R.I. (1996) Centennial of Rontgen's discovery of x-rays. *West J Med*, **164**, 497-501.

104. Grillo-Lopez, A.J., White, C.A., Dallaire, B.K., Varns, C.L., Shen, C.D., Wei, A., Leonard, J.E., McClure, A., Weaver, R., Cairelli, S. *et al.* (2000) Rituximab: the first monoclonal antibody approved for the treatment of lymphoma. *Curr Pharm Biotechnol*, **1**, 1-9.

105. Dagher, R., Cohen, M., Williams, G., Rothmann, M., Gobburu, J., Robbie, G., Rahman, A., Chen, G., Staten, A., Griebel, D. *et al.* (2002) Approval summary: imatinib mesylate in the treatment of metastatic and/or unresectable malignant gastrointestinal stromal tumors. *Clin Cancer Res*, **8**, 3034-3038.

106. Porter, D.L., Levine, B.L., Kalos, M., Bagg, A. and June, C.H. (2011) Chimeric antigen receptor-modified T cells in chronic lymphoid leukemia. *N Engl J Med*, **365**, 725-733.

107. Cohen, M.H., Williams, G.A., Sridhara, R., Chen, G. and Pazdur, R. (2003) FDA drug approval summary: gefitinib (ZD1839) (Iressa) tablets. *Oncologist*, **8**, 303-306.

108. Cohen, M.H., Gootenberg, J., Keegan, P. and Pazdur, R. (2007) FDA drug approval summary: bevacizumab (Avastin) plus Carboplatin and Paclitaxel as first-line treatment of advanced/metastatic recurrent nonsquamous non-small cell lung cancer. *Oncologist*, **12**, 713-718.

109. Institute, N.C. (2024) SEER*Explorer: An interactive website for SEER cancer statistics.

110. The, L. (2024) Lung cancer treatment: 20 years of progress. *The Lancet*, **403**, 2663.

111. Wurth, L. and Gebauer, F. (2015) RNA-binding proteins, multifaceted translational regulators in cancer. *Biochim Biophys Acta*, **1849**, 881-886.

112. Qin, H., Ni, H., Liu, Y., Yuan, Y., Xi, T., Li, X. and Zheng, L. (2020) RNA-binding proteins in tumor progression. *J Hematol Oncol*, **13**, 90.

113. Schmitt, A.M. and Chang, H.Y. (2016) Long Noncoding RNAs in Cancer Pathways. *Cancer Cell*, **29**, 452-463.

114. Xu, Y., Huangyang, P., Wang, Y., Xue, L., Devericks, E., Nguyen, H.G., Yu, X., Oses-Prieto, J.A., Burlingame, A.L., Miglani, S. *et al.* (2021) ERalpha is an RNA-binding protein sustaining tumor cell survival and drug resistance. *Cell*, **184**, 5215-5229 e5217.

115. Dominguez-Brauer, C., Thu, K.L., Mason, J.M., Blaser, H., Bray, M.R. and Mak, T.W. (2015) Targeting Mitosis in Cancer: Emerging Strategies. *Mol Cell*, **60**, 524-536.

116. Rajagopal, V., Seiler, J., Nasa, I., Cantarella, S., Theiss, J., Herget, F., Kaifer, B., Schneider, M., Helm, D., Konig, J. *et al.* (2024) An atlas of RNA-dependent proteins

in cell division reveals the riboregulation of mitotic protein-protein interactions. *bioRxiv*, https://doi.org/10.1101/2024.1109.1125.614981.

117. Blower, M.D., Nachury, M., Heald, R. and Weis, K. (2005) A Rae1-containing ribonucleoprotein complex is required for mitotic spindle assembly. *Cell*, **121**, 223-234.
118. Jambhekar, A., Emerman, A.B., Schweidenback, C.T. and Blower, M.D. (2014) RNA stimulates Aurora B kinase activity during mitosis. *PLoS One*, **9**, e100748.
119. Gene Ontology, C. (2021) The Gene Ontology resource: enriching a GOld mine. *Nucleic Acids Res*, **49**, D325-D334.
120. Shevchenko, A., Tomas, H., Havlis, J., Olsen, J.V. and Mann, M. (2006) In-gel digestion for mass spectrometric characterization of proteins and proteomes. *Nat Protoc*, **1**, 2856-2860.
121. Tyanova, S., Temu, T. and Cox, J. (2016) The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. *Nat Protoc*, **11**, 2301-2319.
122. Schwanhausser, B., Busse, D., Li, N., Dittmar, G., Schuchhardt, J., Wolf, J., Chen, W. and Selbach, M. (2011) Global quantification of mammalian gene expression control. *Nature*, **473**, 337-342.
123. Schroeder, A.B., Dobson, E.T.A., Rueden, C.T., Tomancak, P., Jug, F. and Eliceiri, K.W. (2021) The ImageJ ecosystem: Open-source software for image visualization, processing, and analysis. *Protein Sci*, **30**, 234-249.
124. Buchbender, A., Mutter, H., Sutandy, F.X.R., Kortel, N., Hanel, H., Busch, A., Ebersberger, S. and Konig, J. (2020) Improved library preparation with the new iCLIP2 protocol. *Methods*, **178**, 33-48.
125. Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M. and Gingeras, T.R. (2013) STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*, **29**, 15-21.
126. Krakau, S., Richard, H. and Marsico, A. (2017) PureCLIP: capturing target-specific protein-RNA interaction footprints from single-nucleotide CLIP-seq data. *Genome Biol*, **18**, 240.
127. de Oliveira Freitas Machado, C., Schafranek, M., Bruggemann, M., Hernandez Canas, M.C., Keller, M., Di Liddo, A., Brezski, A., Blumel, N., Arnold, B., Bremm, A. *et al.* (2023) Poison cassette exon splicing of SRSF6 regulates nuclear speckle dispersal and the response to hypoxia. *Nucleic Acids Res*, **51**, 870-890.

128. Busch, A., Bruggemann, M., Ebersberger, S. and Zarnack, K. (2020) iCLIP data analysis: A complete pipeline from sequencing reads to RBP binding sites. *Methods*, **178**, 49-62.

129. Townsend, M.H., Anderson, M.D., Weagel, E.G., Velazquez, E.J., Weber, K.S., Robison, R.A. and O'Neill, K.L. (2017) Non-small-cell lung cancer cell lines A549 and NCI-H460 express hypoxanthine guanine phosphoribosyltransferase on the plasma membrane. *Onco Targets Ther*, **10**, 1921-1932.

130. Liu, X., Zou, H., Zhao, Y., Chen, H., Liu, T., Wu, Z., Yang, C., Li, Q. and Li, Y. (2020) Frontiers | Tanshinone Inhibits NSCLC by Downregulating AURKA Through Let-7a-5p. *Frontiers in Genetics*, **11**, a838.

131. Garcia-de-Alba, C. (2021) Repurposing A549 Adenocarcinoma Cells: New Options for Drug Discovery. *Am J Respir Cell Mol Biol*, **64**, 405-406.

132. Gadea, G. and Blangy, A. (2014) Dock-family exchange factors in cell migration and disease. *Eur J Cell Biol*, **93**, 466-477.

133. Frank, S.R., Kollmann, C.P., van Lidth de Jeude, J.F., Thiagarajah, J.R., Engelholm, L.H., Frodin, M. and Hansen, S.H. (2017) The focal adhesion-associated proteins DOCK5 and GIT2 comprise a rheostat in control of epithelial invasion. *Oncogene*, **36**, 1816-1828.

134. Liu, C., Li, G., Zheng, S., She, L., Lu, S., Wang, Y., Huang, D., Zhang, X., Sun, L., Liu, Y. *et al.* (2023) PHF5A regulates the expression of the DOCK5 variant to promote HNSCC progression through p38 MAPK activation. *Biol Direct*, **18**, 39.

135. Weng, Z., Situ, C., Lin, L., Wu, Z., Zhu, J. and Zhang, R. (2019) Structure of BAI1/ELMO2 complex reveals an action mechanism of adhesion GPCRs via ELMO family scaffolds. *Nat Commun*, **10**, 51.

136. Xu, X. and Jin, T. (2019) ELMO proteins transduce G protein-coupled receptor signal to control reorganization of actin cytoskeleton in chemotaxis of eukaryotic cells. *Small GTPases*, **10**, 271-279.

137. Wang, Y., Li, H. and Li, F. (2020) ELMO2 association with Galphai2 regulates pancreatic cancer cell chemotaxis and metastasis. *PeerJ*, **8**, e8910.

138. Castillo, A., Paul, A., Sun, B., Huang, T.H., Wang, Y., Yazinski, S.A., Tyler, J., Li, L., You, M.J., Zou, L. *et al.* (2014) The BRCA1-interacting protein Abraxas is required for genomic stability and tumor suppression. *Cell Rep*, **8**, 807-817.

139. Mok, M.T. and Henderson, B.R. (2012) The in vivo dynamic organization of BRCA1-A complex proteins at DNA damage-induced nuclear foci. *Traffic*, **13**, 800-814.

140. Barr, A.R. and Gergely, F. (2007) Aurora-A: the maker and breaker of spindle poles. *J Cell Sci*, **120**, 2987-2996.

141. Fu, J., Bian, M., Jiang, Q. and Zhang, C. (2007) Roles of Aurora kinases in mitosis and tumorigenesis. *Mol Cancer Res*, **5**, 1-10.

142. Vader, G. and Lens, S.M. (2008) The Aurora kinase family in cell division and cancer. *Biochim Biophys Acta*, **1786**, 60-72.

143. Cowley, D.O., Rivera-Perez, J.A., Schliekelman, M., He, Y.J., Oliver, T.G., Lu, L., O'Quinn, R., Salmon, E.D., Magnuson, T. and Van Dyke, T. (2009) Aurora-A kinase is essential for bipolar spindle formation and early development. *Mol Cell Biol*, **29**, 1059-1071.

144. Joukov, V. and De Nicolo, A. (2018) Aurora-PLK1 cascades as key signaling modules in the regulation of mitosis. *Sci Signal*, **11**, eaar4195.

145. Kufer, T.A., Sillje, H.H., Korner, R., Gruss, O.J., Meraldi, P. and Nigg, E.A. (2002) Human TPX2 is required for targeting Aurora-A kinase to the spindle. *J Cell Biol*, **158**, 617-623.

146. Schatz, C.A., Santarella, R., Hoenger, A., Karsenti, E., Mattaj, I.W., Gruss, O.J. and Carazo-Salas, R.E. (2003) Importin alpha-regulated nucleation of microtubules by TPX2. *EMBO J*, **22**, 2060-2070.

147. Gruss, O.J., Carazo-Salas, R.E., Schatz, C.A., Guarguaglini, G., Kast, J., Wilm, M., Le Bot, N., Vernos, I., Karsenti, E. and Mattaj, I.W. (2001) Ran induces spindle assembly by reversing the inhibitory effect of importin alpha on TPX2 activity. *Cell*, **104**, 83-93.

148. Joukov, V., De Nicolo, A., Rodriguez, A., Walter, J.C. and Livingston, D.M. (2010) Centrosomal protein of 192 kDa (Cep192) promotes centrosome-driven spindle assembly by engaging in organelle-specific Aurora A activation. *Proc Natl Acad Sci U S A*, **107**, 21022-21027.

149. Meng, L., Park, J.E., Kim, T.S., Lee, E.H., Park, S.Y., Zhou, M., Bang, J.K. and Lee, K.S. (2015) Bimodal Interaction of Mammalian Polo-Like Kinase 1 and a Centrosomal Scaffold, Cep192, in the Regulation of Bipolar Spindle Formation. *Mol Cell Biol*, **35**, 2626-2640.

150. Joukov, V., Walter, J.C. and De Nicolo, A. (2014) The Cep192-organized aurora A-Plk1 cascade is essential for centrosome cycle and bipolar spindle assembly. *Mol Cell*, **55**, 578-591.

151. Mou, P.K., Yang, E.J., Shi, C., Ren, G., Tao, S. and Shim, J.S. (2021) Aurora kinase A, a synthetic lethal target for precision cancer medicine. *Exp Mol Med*, **53**, 835-847.

152. Du, R., Huang, C., Liu, K., Li, X. and Dong, Z. (2021) Targeting AURKA in Cancer: molecular mechanisms and opportunities for Cancer therapy. *Mol Cancer*, **20**, 15.

153. Wang, J., Nikhil, K., Viccaro, K., Chang, L., Jacobsen, M., Sandusky, G. and Shah, K. (2017) The Aurora-A-Twist1 axis promotes highly aggressive phenotypes in pancreatic carcinoma. *J Cell Sci*, **130**, 1078-1093.

154. Chang, S.S., Yamaguchi, H., Xia, W., Lim, S.O., Khotskaya, Y., Wu, Y., Chang, W.C., Liu, Q. and Hung, M.C. (2017) Aurora A kinase activates YAP signaling in triple-negative breast cancer. *Oncogene*, **36**, 1265-1275.

155. Shah, K.N., Bhatt, R., Rotow, J., Rohrberg, J., Olivas, V., Wang, V.E., Hemmati, G., Martins, M.M., Maynard, A., Kuhn, J. *et al.* (2019) Aurora kinase A drives the evolution of resistance to third-generation EGFR inhibitors in lung cancer. *Nat Med*, **25**, 111-118.

156. Mountain, V., Simerly, C., Howard, L., Ando, A., Schatten, G. and Compton, D.A. (1999) The kinesin-related protein, HSET, opposes the activity of Eg5 and cross-links microtubules in the mammalian mitotic spindle. *J Cell Biol*, **147**, 351-366.

157. Walczak, C.E., Vernos, I., Mitchison, T.J., Karsenti, E. and Heald, R. (1998) A model for the proposed roles of different microtubule-based motor proteins in establishing spindle bipolarity. *Curr Biol*, **8**, 903-913.

158. Ems-McClung, S.C., Zheng, Y. and Walczak, C.E. (2004) Importin alpha/beta and Ran-GTP regulate XCTK2 microtubule binding through a bipartite nuclear localization signal. *Mol Biol Cell*, **15**, 46-57.

159. Wittmann, T., Wilm, M., Karsenti, E. and Vernos, I. (2000) TPX2, A novel xenopus MAP involved in spindle pole organization. *J Cell Biol*, **149**, 1405-1418.

160. Gruss, O.J., Wittmann, M., Yokoyama, H., Pepperkok, R., Kufer, T., Sillje, H., Karsenti, E., Mattaj, I.W. and Vernos, I. (2002) Chromosome-induced microtubule assembly mediated by TPX2 is required for spindle formation in HeLa cells. *Nat Cell Biol*, **4**, 871-879.

161. Weaver, L.N., Ems-McClung, S.C., Chen, S.H., Yang, G., Shaw, S.L. and Walczak, C.E. (2015) The Ran-GTP gradient spatially regulates XCTK2 in the spindle. *Curr Biol*, **25**, 1509-1514.

162. Ems-McClung, S.C., Emch, M., Zhang, S., Mahnoor, S., Weaver, L.N. and Walczak, C.E. (2020) RanGTP induces an effector gradient of XCTK2 and importin alpha/beta for spindle microtubule cross-linking. *J Cell Biol*, **219**, e201906045.

163. Kleylein-Sohn, J., Pollinger, B., Ohmer, M., Hofmann, F., Nigg, E.A., Hemmings, B.A. and Wartmann, M. (2012) Acentrosomal spindle organization renders cancer cells dependent on the kinesin HSET. *J Cell Sci*, **125**, 5391-5402.

164. Zhang, C., Wu, B.Z., Di Ciano-Oliveira, C., Wu, Y.F., Khavkine Binstock, S.S., Soria-Bretones, I., Pham, N.A., Elia, A.J., Chari, R., Lam, W.L. *et al.* (2024) Identification of KIFC1 as a putative vulnerability in lung cancers with centrosome amplification. *Cancer Gene Ther*, **31**, 1559-1570.

165. Fan, G., Sun, L., Meng, L., Hu, C., Wang, X., Shi, Z., Hu, C., Han, Y., Yang, Q., Cao, L. *et al.* (2021) The ATM and ATR kinases regulate centrosome clustering and tumor recurrence by targeting KIFC1 phosphorylation. *Nat Commun*, **12**, 20.

166. Chavali, P.L., Chandrasekaran, G., Barr, A.R., Tatrai, P., Taylor, C., Papachristou, E.K., Woods, C.G., Chavali, S. and Gergely, F. (2016) A CEP215-HSET complex links centrosomes with spindle poles and drives centrosome clustering in cancer. *Nat Commun*, **7**, 11005.

167. Li, Y., Lu, W., Chen, D., Boohaker, R.J., Zhai, L., Padmalayam, I., Wennerberg, K., Xu, B. and Zhang, W. (2015) KIFC1 is a novel potential therapeutic target for breast cancer. *Cancer Biol Ther*, **16**, 1316-1322.

168. Sun, M., Jia, M., Ren, H., Yang, B., Chi, W., Xin, G., Jiang, Q. and Zhang, C. (2021) NuMA regulates mitotic spindle assembly, structural dynamics and function via phase separation. *Nat Commun*, **12**, 7157.

169. Naso, F.D., Polverino, F., Cilluffo, D., Latini, L., Stagni, V., Asteriti, I.A., Rosa, A., Soddu, S. and Guarguaglini, G. (2024) AurkA/TPX2 co-overexpression in nontransformed cells promotes genome instability through induction of chromosome mis-segregation and attenuation of the p53 signalling pathway. *Biochim Biophys Acta Mol Basis Dis*, **1870**, 167116.

170. van Gijn, S.E., Wierenga, E., van den Tempel, N., Kok, Y.P., Heijink, A.M., Spierings, D.C.J., Foijer, F., van Vugt, M. and Fehrmann, R.S.N. (2019) TPX2/Aurora kinase A signaling as a potential therapeutic target in genomically unstable cancer cells. *Oncogene*, **38**, 852-867.

171. Jung, J., Jeong, H., Choi, J.W., Kim, H.S., Oh, H.E., Lee, E.S., Kim, Y.S. and Lee, J.H. (2021) Increased expression levels of AURKA and KIFC1 are promising predictors of progression and poor survival associated with gastric cancer. *Pathol Res Pract*, **224**, 153524.

172. Fredriksson, S., Gullberg, M., Jarvius, J., Olsson, C., Pietras, K., Gustafsdottir, S.M., Ostman, A. and Landegren, U. (2002) Protein detection using proximity-dependent DNA ligation assays. *Nat Biotechnol*, **20**, 473-477.

173. Sugimoto, Y., Konig, J., Hussain, S., Zupan, B., Curk, T., Frye, M. and Ule, J. (2012) Analysis of CLIP and iCLIP methods for nucleotide-resolution studies of protein-RNA interactions. *Genome Biol*, **13**, R67.

174. Kettenbach, A.N., Schweppe, D.K., Faherty, B.K., Pechenick, D., Pletnev, A.A. and Gerber, S.A. (2011) Quantitative phosphoproteomics identifies substrates and functional modules of Aurora and Polo-like kinase activities in mitotic cells. *Sci Signal*, **4**, rs5.

175. Ferrari, S., Marin, O., Pagano, M.A., Meggio, F., Hess, D., El-Shemerly, M., Krystyniak, A. and Pinna, L.A. (2005) Aurora-A site specificity: a study with synthetic peptide substrates. *Biochem J*, **390**, 293-302.

176. Beckmann, B.M., Horos, R., Fischer, B., Castello, A., Eichelbaum, K., Alleaume, A.M., Schwarzl, T., Curk, T., Foehr, S., Huber, W. *et al.* (2015) The RNA-binding proteomes from yeast to man harbour conserved enigmRBPs. *Nat Commun*, **6**, 10127.

177. Conrad, T., Albrecht, A.S., de Melo Costa, V.R., Sauer, S., Meierhofer, D. and Orom, U.A. (2016) Serial interactome capture of the human cell nucleus. *Nat Commun*, **7**, 11212.

178. Mullari, M., Lyon, D., Jensen, L.J. and Nielsen, M.L. (2017) Specifying RNA-Binding Regions in Proteins by Peptide Cross-Linking and Affinity Purification. *J Proteome Res*, **16**, 2762-2772.

179. He, C., Sidoli, S., Warneford-Thomson, R., Tatomer, D.C., Wilusz, J.E., Garcia, B.A. and Bonasio, R. (2016) High-Resolution Mapping of RNA-Binding Regions in the Nuclear Proteome of Embryonic Stem Cells. *Mol Cell*, **64**, 416-430.

180. Queiroz, R.M.L., Smith, T., Villanueva, E., Marti-Solano, M., Monti, M., Pizzinga, M., Mirea, D.M., Ramakrishna, M., Harvey, R.F., Dezi, V. *et al.* (2019) Comprehensive identification of RNA-protein interactions in any organism using orthogonal organic phase separation (OOPS). *Nat Biotechnol*, **37**, 169-178.

181. Panhale, A., Richter, F.M., Ramirez, F., Shvedunova, M., Manke, T., Mittler, G. and Akhtar, A. (2019) CAPRI enables comparison of evolutionarily conserved RNA interacting regions. *Nat Commun*, **10**, 2682.

182. Spizzichino, S., Di Fonzo, F., Marabelli, C., Tramonti, A., Chaves-Sanjuan, A., Parroni, A., Boumis, G., Liberati, F.R., Paone, A., Montemiglio, L.C. *et al.* (2024) Structure-

based mechanism of riboregulation of the metabolic enzyme SHMT1. *Mol Cell*, **84**, 2682-2697.e2686.

183. Bracken, C.P. (2023) "Crowd-control" by RNA: a pervasive theme in biology. *RNA*, **29**, 885-888.
184. Fan, H. and Penman, S. (1970) Regulation of protein synthesis in mammalian cells. II. Inhibition of protein synthesis at the level of initiation during mitosis. *J Mol Biol*, **50**, 655-670.
185. Contreras, A. and Perea-Resa, C. (2024) Transcriptional repression across mitosis: mechanisms and functions. *Biochem Soc Trans*, **52**, 455-464.
186. Eyers, P.A. and Maller, J.L. (2004) Regulation of Xenopus Aurora A activation by TPX2. *J Biol Chem*, **279**, 9008-9015.
187. Bayliss, R., Sardon, T., Vernos, I. and Conti, E. (2003) Structural basis of Aurora-A activation by TPX2 at the mitotic spindle. *Mol Cell*, **12**, 851-862.
188. Norris, S.R., Jung, S., Singh, P., Strothman, C.E., Erwin, A.L., Ohi, M.D., Zanic, M. and Ohi, R. (2018) Microtubule minus-end aster organization is driven by processive HSET-tubulin clusters. *Nat Commun*, **9**, 2659.
189. Cai, S., Weaver, L.N., Ems-McClung, S.C. and Walczak, C.E. (2009) Kinesin-14 family proteins HSET/XCTK2 control spindle length by cross-linking and sliding microtubules. *Mol Biol Cell*, **20**, 1348-1359.
190. Ma, K., Luo, M., Xie, G., Wang, X., Li, Q., Gao, L., Yu, H. and Yu, X. (2022) Ribosomal RNA regulates chromosome clustering during mitosis. *Cell Discov*, **8**, 51.
191. Hayashi, Y., Kato, K. and Kimura, K. (2017) The hierarchical structure of the perichromosomal layer comprises Ki67, ribosomal RNAs, and nucleolar proteins. *Biochem Biophys Res Commun*, **493**, 1043-1049.
192. Hernandez-Armendariz, A., Sorichetti, V., Hayashi, Y., Koskova, Z., Brunner, A., Ellenberg, J., Saric, A. and Cuylen-Haering, S. (2024) A liquid-like coat mediates chromosome clustering during mitotic exit. *Mol Cell*, **84**, 3254-3270 e3259.
193. Sharp, J.A., Plant, J.J., Ohsumi, T.K., Borowsky, M. and Blower, M.D. (2011) Functional analysis of the microtubule-interacting transcriptome. *Mol Biol Cell*, **22**, 4312-4323.
194. Hussain, S., Benavente, S.B., Nascimento, E., Dragoni, I., Kurowski, A., Gillich, A., Humphreys, P. and Frye, M. (2009) The nucleolar RNA methyltransferase Misu (NSun2) is required for mitotic spindle stability. *J Cell Biol*, **186**, 27-40.

195. Groisman, I., Huang, Y.S., Mendez, R., Cao, Q., Theurkauf, W. and Richter, J.D. (2000) CPEB, maskin, and cyclin B1 mRNA at the mitotic apparatus: implications for local translational control of cell division. *Cell*, **103**, 435-447.

196. Eliscovich, C., Peset, I., Vernos, I. and Mendez, R. (2008) Spindle-localized CPE-mediated translation controls meiotic chromosome segregation. *Nat Cell Biol*, **10**, 858-865.

197. Waldron, A. and Yajima, M. (2020) Localized translation on the mitotic apparatus: A history and perspective. *Dev Biol*, **468**, 55-58.

198. Blower, M.D., Feric, E., Weis, K. and Heald, R. (2007) Genome-wide analysis demonstrates conserved localization of messenger RNAs to mitotic microtubules. *J Cell Biol*, **179**, 1365-1373.

199. Sharp, J.A., Perea-Resa, C., Wang, W. and Blower, M.D. (2020) Cell division requires RNA eviction from condensing chromosomes. *J Cell Biol*, **219**, e201910148.

200. Calabretta, S. and Richard, S. (2015) Emerging Roles of Disordered Sequences in RNA-Binding Proteins. *Trends Biochem Sci*, **40**, 662-672.

201. Zeke, A., Schád, É., Horváth, T., Abukhairan, R., Szabó, B. and Tantos, A. (2022) Deep structural insights into RNA-binding disordered protein regions. *WIREs RNA*, **5**, e1714.

202. Ray, D., Laverty, K.U., Jolma, A., Nie, K., Samson, R., Pour, S.E., Tam, C.L., von Krosigk, N., Nabeel-Shah, S., Albu, M. *et al.* (2023) RNA-binding proteins that lack canonical RNA-binding domains are rarely sequence-specific. *Sci Rep*, **13**, 5238.