Optimizing the Conditions to Identify the DDX41 Interactome by BioID

A Thesis Submitted to the College of
Graduate and Postdoctoral Studies
In Fulfillment of the
Requirements for the Degree of
Master of Science
In the Department of Biochemistry, Microbiology, and Immunology
University of Saskatchewan
Saskatoon

By Ananaya Charaya

PERMISSION OF USE STATEMENT

I hereby present this thesis in partial fulfilment of the requirements for a postgraduate degree from the University of Saskatchewan and agree that the Libraries of this University may make it freely available for inspection. I further agree that permission for copying of this thesis in any manner, either in whole or in part, for scholarly purposes may be granted by the professor or professors who supervised this thesis or, in their absence, by the Head of the Department or the Dean of the College in which my thesis work was done. It is understood that any copying or publication or use of this thesis or parts of it for any financial gain will not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of Saskatchewan for any scholarly use which may be made of any material in my thesis.

Requests for permission to copy or to make other use of material in this thesis in whole or part should be addressed to:

Dr. Linda Chelico
Head of the Department of Biochemistry, Microbiology, and Immunology
University of Saskatchewan
Saskatoon, Saskatchewan, S7N 5E5

Dr. Debby Burshtyn
Dean
College of Graduate and Postdoctoral Studies
University of Saskatchewan
116 Thorvaldson Building, 110 Science Place
Saskatoon, Saskatchewan, S7N 5C9
Canada

ABSTRACT

DDX41 (DEAD-box 41) protein, an RNA helicase, is involved in several processes, including genome stability, mRNA splicing, ribosome biogenesis, translation, cell differentiation, and innate immunity. Mutations in DDX41 cause myeloid malignancies, especially myelodysplastic syndrome (MDS) and acute myeloid leukemia (AML). Our preliminary data showed that DDX41 experiences dynamic changes, including protein level and sub-cellular localization, when the cells are exposed to distinct stress conditions, including microbial infection (Herpes simplex virus) and radiation exposure (infrared radiations). However, the molecular mechanisms regulating DDX41 expression and sub-cellular localization are unknown. We hypothesize that DDX41-interacting proteins may regulate the dynamics of DDX41.

To identify the interacting proteins of DDX41 in living cells, we employed a proximity-dependent biotin identification (BioID) approach; hence, we used an improved biotin ligase, TurboID. We constructed a DDX41-TurboID fusion construct, where DDX41 is tagged with 3xFlag at its N-terminus and TurboID ligase at its C-terminus and a control construct that contains 3xFlag and TurboID. By transfection in HEK293T and HeLa DDX41 knockout cells, I observed the protein expression and nuclear localization for DDX41-TurboID fusion proteins and cytoplasmic localization of the 3xFlag-TurboID (control) proteins by Western blot and immunofluorescence microscopy respectively. I found that the labelling time of TurboID ligase of 30 min and a biotin concentration of 50 µM were optimal for the biotinylation activity of the TurboID ligase on proxisomes and hence the biotinylation of potential DDX41-interacting proteins. I optimized the immunoprecipitation (IP) of biotin-treated BioID fusion proteins and DDX41-interactome using Flag and streptavidin conjugated to protein A/G agarose beads as well as the streptavidin sepharose beads. Western blots validated the optimization of the IP, resulting in the presence of a potential biotinylated DDX41-interacting protein. In inference, we have constructed and validated the expression and localization of BioID fusion constructs and optimized the conditions to identify the interactome of DDX41.

In the future, mass spectrometric analysis will facilitate the identification of the interacting proteins of DDX41, and the validation of the DDX41-interacting proteins will lead us to a better understanding of the regulatory pathways and potential druggable targets for the treatment of DDX41-mutated disorders.

ACKNOWLEDGEMENTS

I bow down my head with utmost reverence before the Almighty whose eternal blessings enabled me to accomplish this endeavour. My vocabulary utterly fails to express my emotions towards my loving parents Mr. Ramesh Kumar Charaya and Mrs. Suman Charaya for their blessings and selfless sacrifices in making me capable of pursuing my dreams.

I veraciously realize the inadequacy of words at my command to extend my thanks to my supervisors, Dr. Yuliang Wu and Dr. Kiven Erique Lukong for their judicious guidance, keen supervision, and co-operation extended to me as and when required.

It is my privilege to express my sense of gratitude to my advisory committee members, Dr. Hong Wang and Dr. Harold Bull (Chair) for their encouragement, constructive criticism and valuable suggestions in the completion of my research work with satisfaction. It is my immense pleasure to be part of the Biochemistry Microbiology and Immunology Department, College of Medicine. I am grateful to Dr. Scot Stone (Graduate chair) and Dr. Bill Roesler (former Head of Department) for their leadership and immense support.

I would also like to extend my thanks to the funding sources, NSERC, CRS, LLSC and SHRF to support our research work.

I am thankful to my lab members, Shizhuo (Sarah) Yang, Lacey Winstone, Sohaumn Mondal, Yohan Jung and Nayoung Kim (from Dr. Lukong's lab) for their relentless help. Mark Boyd (Lab manager) and the members of the Cancer Cluster are entitled to my deepest sense of appreciation for their all-time guidance and comfort in all the phases of my journey during the program.

I owe my unbound affection towards my elder brother Vaibhav and his wife Dr. Sonam who supported and pushed me to achieve my goals.

All may not have been mentioned but none is forgotten.

DEDICATION

To My Heartbeats

My dearest Mumma and Papa

Mrs. Suman and Mr. Ramesh Kumar Charaya

And

To the strength of my soul

Shree RadhaKrishna

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AML	Acute myeloid leukemia
ATCC	American type culture collection
BCCP	Biotin carboxyl carrier protein
bioAMP	Biotinyl Adenosine mono phosphate
BioID	Biotin identification
cGAS	Cyclic GMP-AMP synthase
CML	Chronic myeloid leukemia
DBD	DNA binding domain
DDX41	DEAD-box helicase 41
DEAD	Asp-Glu-Ala-Asp
DMEM	Dulbecco's modified eagle's medium
DNA	Deoxyribose nucleic acid
EDTA	Ethylene diamine tetraacetic acid
FBS	Fetal bovine serum
GFP	Green fluorescent protein
HEK	Human embryonic kidney
HeLa	Henrietta Lacks
HRP	Horseradish peroxidase
HSV-1	Herpes simplex virus type 1
HU	Hydroxyurea
IP	Immunoprecipitation
IR	Ionizing radiation
IRF-3	Interferon regulatory factor 3
kb	Kilobases
kDa	Kilodalton
KO	Knock out
LB	Lysogeny broth
MDS	Myelodysplastic syndrome
min	Minutes
miRNA	Micro RNA
MNs	Myeloid neoplasms
MS	Mass spectrometry
NES	Nuclear export signal
NLS	Nuclear localization signal
NMD	Nonsense Mediated Decay
NTP	Nucleotide triphosphate
OD	Optical density
PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate-buffered saline
PCR	Polymerase chain reaction
Pi	Phosphate ion
PMSF	Phenylmethanesulfonylfluoride
PPI	Protein-protein interaction
PVDF	Polyvinylidene difluoride

RIPA	Radio immunoprecipitation assay
RNA	Ribonucleic acid
RNP	Ribonucleoproteins
RNAPII	RNA polymerase II
RT	Room temperature
SDS	Sodium dodecyl sulphate
SF	Superfamily
siRNA	Small interfering RNA
SSC	Saline sodium citrate
STING	Stimulator of interferon genes
TBK-1	TANK binding kinase-1
TEMED	N, N, N', N'- tetramethyl ethylenediamine
UV	Ultraviolet
V	Voltage
WCL	Whole cell lysate
WHO	World Health Organization
WT	Wild type
ZnF	Zinc finger

1. INTRODUCTION

1.1 Helicase

Helicases are defined as enzymes capable of unwinding nucleic acids (DNA and RNA) driven by the energy of ATP hydrolysis. They were first discovered in *E. coli* and reported in 1976 by two groups, Mackay and Linn at the University of California, Berkeley (Mackay & Linn, 1976) and Abdel-Monem and Hoffmann-Berling at the University of Heidelberg (Abdel-Monem et al., 1976; Abdel-Monem & Hoffmann-Berling, 1976). They were termed helicases by Hoffmann-Berling in 1978 (Geider, 1978). Subsequently, helicases were discovered across different species: eukaryotes (lily) (Hotta & Stern, 1978), bacteriophages (Venkatesan et al., 1982), viruses (Stahl et al., 1986), yeast (Sugino et al., 1986) and humans (Tuteja et al., 1990). Nearly 1% of the human genome encodes for the helicases, and about 200 have been identified. They are involved in different cellular functions like DNA replication (Pomerantz & O'Donnell, 2007), DNA repair (Delagoutte & Von Hippel, 2003; Truglio et al., 2006), recombination (Cobb & Bjergbaek, 2006; Hanada & Hickson, 2007; Ouyang et al., 2008), transcription (Clark et al., 2008), translation (Carrera et al., 2000), RNA splicing and its degradation (J. S. J. Anderson & R. Parker, 1998; Schwer, 2008) and ribosome biogenesis (Bernstein et al., 2006; Granneman et al., 2006).

The enzymatic activities of helicases are described in terms of directionality, processivity, and substrate. On the basis of directionality, helicases can unwind nucleic acids in 3′ to 5′ or 5′ to 3′ direction. The enzymatic rate of helicase is characterized by the number of nucleotide triphosphates (NTPs) hydrolyzed, per unit time and the number of base pairs unwound before the dissociation of helicase from the nucleic acid determines its processivity (Delagoutte & Von Hippel, 2002; von Hippel & Delagoutte, 2003). Based on the type of nucleic acids, helicases are categorized as DNA or RNA helicases. Nevertheless, some helicases can unwind both DNA and RNA (Pang et al., 2002; Zhang & Grosse, 1994), whereas some unwind RNA–DNA duplexes (Boule & Zakian, 2007).

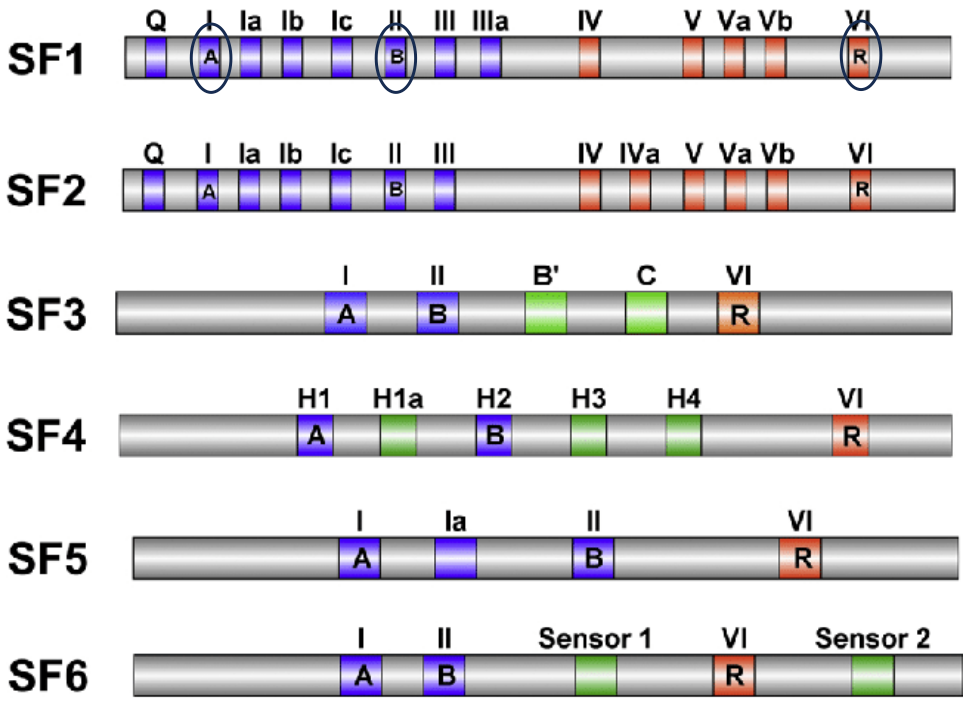


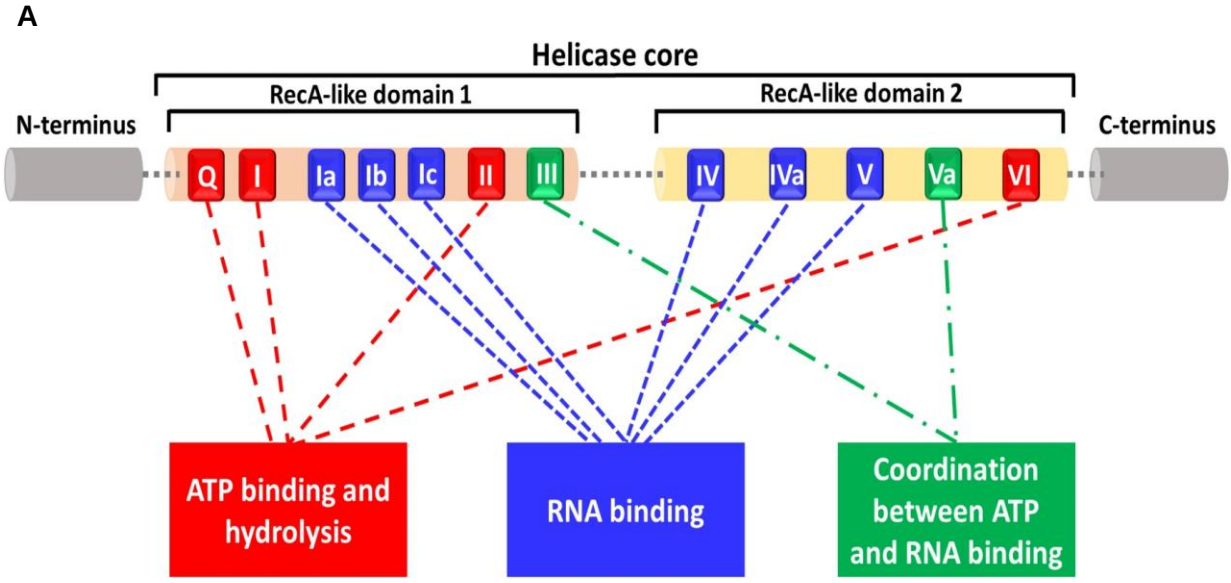
Figure 1. Classification of the helicases. Six superfamilies of helicases with their characteristic domains. SF1-2 share high similarity in their conserved motifs, while SF3-6 share more similarity in the presence of conserved motifs. All SFs have three motifs i.e. Walker A (designated as A), Walker B (B) and arginine finger (R in motif VI) in common. Referenced from (Sami et al., 2021).

Based on conserved motifs and their variations, helicases are divided into six superfamilies (SF1-6) (Gorbalenya & Koonin, 1993) (**Figure 1**). Further, based on their structural characteristics, the superfamilies can be classified into two major groups, one group (SF1-2) are monomeric (Dillingham, 2011; Lohman et al., 2008) and another group exist as hexamers and forms toroidal (ring-like) structure (Patel & Picha, 2000). Amongst all the SFs, SF1 and SF2 are recognized for possessing the highest number of DNA and RNA helicases (Jankowsky, 2011). The helicases of the two SFs are characterized by a conserved helicase core consisting of two tandemly placed RecA domains (Singleton et al., 2007). The domains are known as RecA-like because they are similar to the ATP-binding core of the RecA recombination protein, a DNA repair and maintenance protein in bacteria (Schmucker et al., 1997; Van Brabant et al., 2000). The RecA domains consist of nine short conserved amino acid sequences, named Q, I, Ia, Ib II, III, IV, V, and VI. These motifs are involved in crucial functions such as NTP binding, hydrolysis, and nucleic acid remodelling (Linder & Jankowsky, 2011; Putnam & Jankowsky, 2013). The helicase core domains exhibit two features, first, conserved amino acid residues involved in NTP binding and hydrolysis and second, an arginine residue that facilitates energy coupling resulting

in translocation and/or unwinding. In addition to the core domains, many helicases especially SF1 and SF2, possess C- and N-terminal accessory domains. These domains may determine the specific role of individual helicase, and their functions vary like the nuclease activity, RNA, DNA binding domains and promoting protein-protein interactions (Bernstein et al., 2003; Cui et al., 2008; He et al., 2010)

1.2 DEAD-box helicases

Of all the superfamilies, SF2 comprises the largest number of helicases including DEAD-box helicases (DDX) (Linder et al., 1989). DEAD-box family of RNA helicases derive their name from the presence of a conserved sequence of four amino acids in motif II (similar to the Walker B motif of ATPases). Different motifs of DEAD-box helicases perform distinct roles in nucleic acid metabolism (**Figure 2A)**. Motifs Q, I, II (DEAD motif) and VI are required for ATP binding and hydrolysis; motifs Ia, Ib, IV, and V participate in RNA interaction and intramolecular arrangement (Ali, 2021; Pause & Sonenberg, 1992). Motifs Q, I, Ia, Ib, Ic, II, and III form domain 1, and motifs IV, IVa, V, Va, VI form domain 2.



B

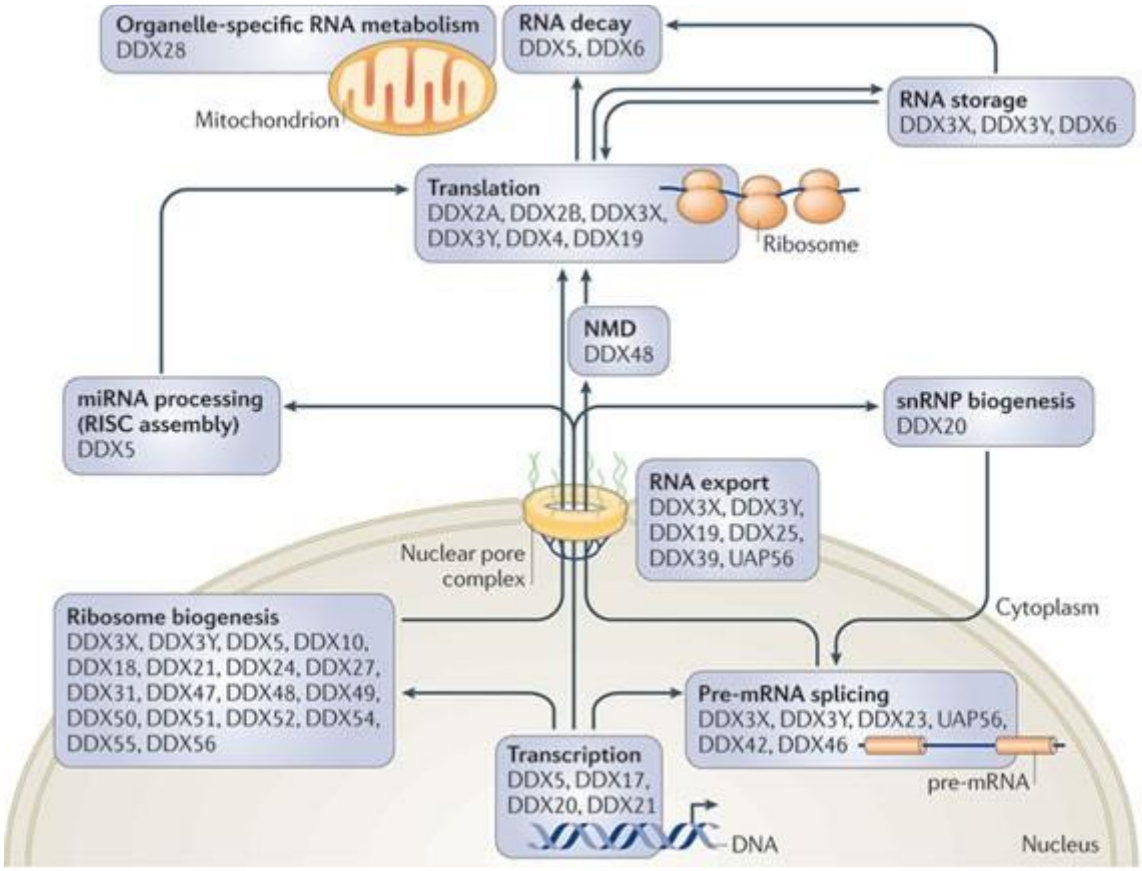


Figure 2. General structure and functions of DEAD-box helicases. (A) Structural features of DEAD-box (DDX) helicases. The helicase core of DEAD-box helicase comprises two domains (RecA-like domains 1 and 2). Conserved motifs Q, I, II and VI (red) function in ATP binding and hydrolysis; motifs Ia, Ib, Ic, IV, IVa and V (blue) bind to RNA substrate, and motifs III and Va communicate between ATP-binding and RNA-binding sites (green). Variable N- and the C-terminus are indicated in grey. Referenced from (Ali, 2021). **(B)** DEAD-box helicases in different processes of RNA metabolism. These include pre-mRNA splicing, transcription, ribosome synthesis in the nucleus; organelle-specific RNA metabolism, nonsense-mediated decay (NMD), protein translation, and microRNA (miRNA) processing in the cytoplasm. The DEAD-box proteins at the interface between the nucleus and the cytoplasm are required for the directional transport of mRNA molecules. Referenced from (Linder & Jankowsky, 2011)

Although containing highly conserved helicase core and conserved ATP-binding and RNA-binding sites, different DEAD-box helicases have diverse functions (**Figure 2B**). For instance, DDX21 is involved in transcription, DDX23 in pre-mRNA splicing (Auboeuf et al., 2002; Hönig et al., 2002; Liu, 2002), DDX10 and DDX18 in ribosome biogenesis (Venema &

Tollervey, 1999; Woolford & Baserga, 2013), DDX4 in translation initiation (Chuang et al., 1997), DDX6 in RNA degradation (J. S. Anderson & R. P. Parker, 1998), and DDX28 in organelle-specific gene expression (Missel et al., 1997). One of the DDX helicases, DDX3X is involved in nearly every aspect of RNA metabolism in the cell (Linder & Jankowsky, 2011). DEAD-box helicases have been associated with neurodegeneration, genetic diseases, and cancer (Fuller-Pace, 2013; Lasko, 2013). Deregulation or mutations in DDX genes manifest in the pre-disposition of human disorders like DDX1 in breast cancer (Germain et al., 2011), DDX3X in lung cancer (Wu et al., 2011), DDX5 in colon cancer (Shin et al., 2007) and DDX41 in leukemia (Ding et al., 2012).

1.3 DDX41 helicase

DDX41 (DEAD-box RNA helicase 41) is a member of the DEAD-box subfamily in SF2 (**Figure 3**). Abstrakt, an ortholog of DDX41, was first discovered in *Drosophila melanogaster* in 1999 (Irion & Leptin, 1999). It is a ubiquitous protein and has been found in *C. elegans* (Kim et al., 2012), shrimps and porcine (Zhu et al., 2014), chicken (Cheng et al., 2017), duck (Li et al., 2018), grass carp (Hu et al., 2020), mandarin fish (Qin et al., 2022), and mice and humans (Zhang et al., 2011). In humans, DDX41 is a 622 amino acid long protein that is characterized by the presence of a conserved helicase core, with a nuclear localization signal (NLS) and zinc finger (ZF) domain in the N- and C-terminal regions respectively **(Figure 3)**. It is interesting that DDX41, although belonging to RNA helicases, also acts as a DNA helicase (Shih & Lee, 2014; Singh et al., 2022).

Being a DEAD-box helicase, DDX41 is involved in various cellular functions like ATP-dependent unwinding DNA and ATP-independent rewinding of DNA (Singh et al., 2022), formation/deformation of R-loops (Mosler et al., 2021), DNA virus sensor (Zhang et al., 2011), and bacterial cyclic di-GMP/AMP sensor to activate type I interferons (IFNs) induced immune response (Parvatiyar et al., 2012). DDX41 has also been found to be involved in pre-mRNA splicing (Polprasert et al., 2015), transcriptional regulation (Qin et al., 2021), translation (Peters et al., 2017) and ribosome biogenesis (Tungalag et al., 2023). Thus, DDX41 is a multifunctional protein.

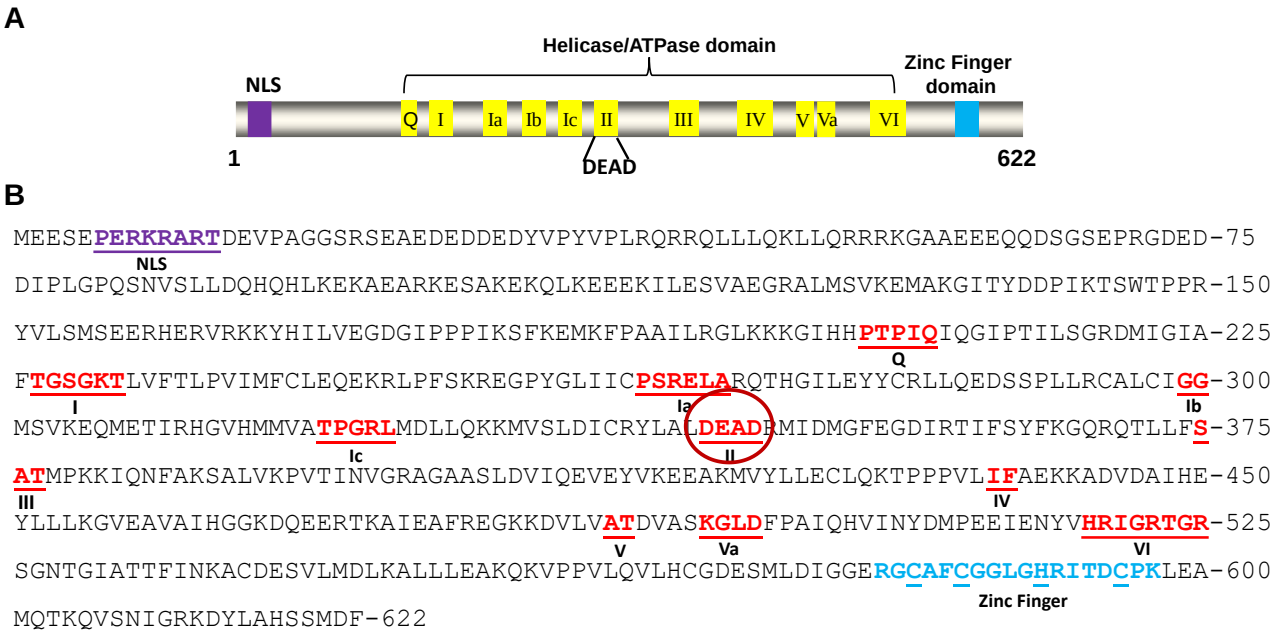


Figure 3. Structural and sequential representation of human DDX41 protein. (**A**) Cartoon depicting the nuclear localization sequence (NLS, purple), conserved helicase motifs (yellow), and zinc finger domain (ZF, blue) of DDX41 protein. (**B**) The amino acid sequence of the human DDX41 protein highlights the sequence region of its various conserved motifs

1.4 DDX41 mutations are associated with myelodysplastic syndrome and acute myeloid leukemia

Mutations in DDX41 manifest into myeloid neoplasms (MNs). The germline and acquired somatic mutations of DDX41 have been associated with myelodysplastic syndrome (MDS) and acute myeloid leukemia (AML) (Cardoso et al., 2016; Cheah et al., 2017; Churpek & Smith-Simmer, 1993; Ding et al., 2012; Sébert et al., 2019; Shin et al., 2022). MDS and AML are hematologic diseases, where MDS is characterized by unsuccessful hematopoiesis and peripheral blood cytopenia (a disorder in which the number of blood cells are lower than normal) whereas in AML, bone marrow is characterized by the spread of immature myeloid cells. Progression of MDS increases the risk of its evolution into AML. Amongst all adult cases of MDS and AML, the occurrence of myeloid neoplasms with germline predisposition accounts for nearly 5–15% (Kim et al., 2023). DDX41 germline mutations are one of the most frequently identified mutations and have been reported in nearly 1.5–3.8% of patients with myeloid neoplasms (Makishima et al., 2023; Quesada et al., 2019; Sébert et al., 2019)

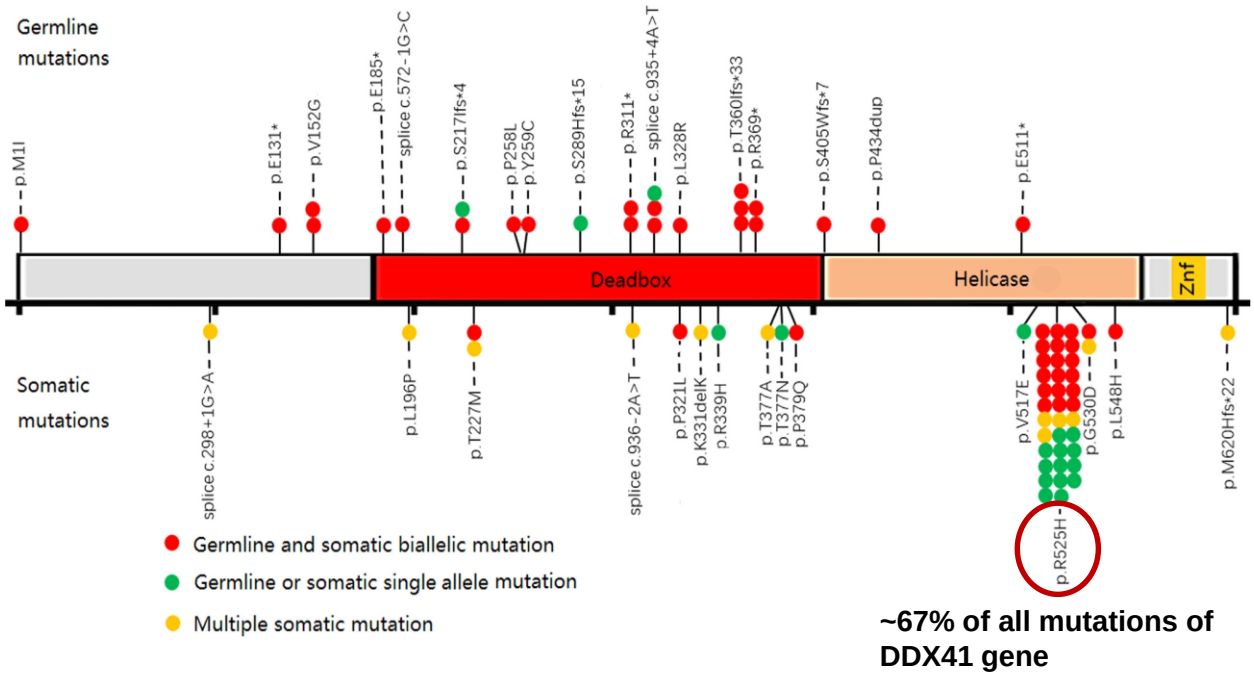


Figure 4. DDX41 germline and somatic mutations in MDS/AML patients. The R525H mutant is highlighted in the helicase domain. Referenced from (Qu et al., 2021).

In 2022, in the classification of haematolymphoid tumours, the World Health Organization (WHO) classified DDX41-mutated leukemias as genetic tumour syndromes in the subgroup of DNA repair and genomic stability defects (Khoury et al., 2022). The underlying cause for germline mutations might be truncation, frameshift mutations (Bannon et al., 2020; Kim et al., 2023; Lewinsohn et al., 2016) and most somatic mutations mostly account for missense mutation (Badar & Chlon, 2022; Chlon et al., 2021; Makishima et al., 2023) (**Figure 4**). The most frequently identified somatic mutation is c.1574G>A, p.R525H, which is mostly confined to highly conserved helicase motif VI (Arginine finger motif) of DDX41. This mutation is thought to contribute to nucleotide co-ordination, defects in pre-rRNA processing, and decreased protein synthesis due to ribosomal stress (Kadono et al., 2016; Li et al., 2016). Our group has reported that the p.R525H mutation is related to the reduced unwinding activity of 30-bp fork-duplex DNA substrate in comparison to the unwinding activity of WT DDX41 (Singh et al., 2022).

1.5 DDX41 and its interactome

In 2002, Jurica *et al* characterized the spliceosome using affinity purification under native conditions and delineated that it consists of a C complex containing splicing intermediates (Jurica et al., 2002). Upon purification of the C complex by size exclusion and affinity selection, they identified the presence of U2, U5, and U6 snRNAs, with minimal amounts of U1 or U4 snRNA. The presence of core snRNP proteins, U2 and U5 snRNP-specific proteins, as well as second step factors Prp16, Prp17, Slu7, and Prp22 were confirmed by tandem mass spectrometry analysis. They identified Abstrakt as one of the DEAD box proteins exclusively in the C complex, which was not previously associated with splicing. In 2008, Bessonov *et al.* performed affinity purification of the human spliceosomal C complexes and found that DDX41/Abstrakt was recruited to the catalytic C complex. In 2011, Agafonov *et al.* reported the presence of DDX41 (Abstrakt) in the spliceosomal C complex in *Drosophila* (Agafonov et al., 2011). The localization of DDX41 in the spliceosome, and its interaction with spliceosomal proteins, particularly proteins of the U2 and U5 snRNP, was demonstrated by Polprasert (Polprasert et al., 2015) (**Figure 5**). Furthermore, patient mutant R525H altered the native DDX41 interactome, especially major components in U2 (SF3B1, SF3B2, and SF3B3) and U5 (PRPF8 and SNRNP200) spliceosome. Studies of DDX41 orthologs in *C. elegans*, SACY-1, have shown that its depletion changes the transcriptome through both splicing-dependent and independent mechanisms (Osterhoudt et al., 2024; Tsukamoto et al., 2020). These results support the role of DDX41 in the spliceosome and its function in mRNA splicing.

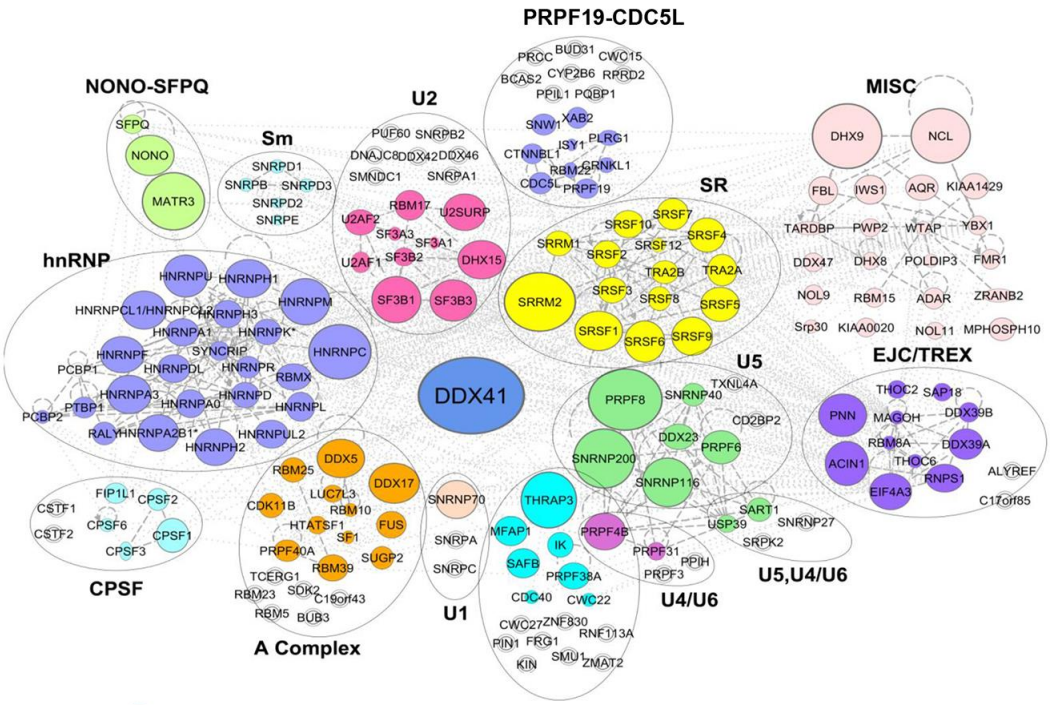


Figure 5. DDX41 in spliceosomal protein complexes. Spliceosomal proteins that coimmunoprecipitated with DDX41 were organized in coloured functional protein complexes based on Ingenuity pathway analysis. Individual protein enrichment was presented as total spectral counts and displayed by different circle sizes. Increased circle size indicates a higher number of total spectra counts for the protein. Unfilled double ring symbols indicate proteins that were not identified in DDX41 co-immunoprecipitation experiments, but which have been linked to the spliceosome. Referenced from (Polprasert et al., 2015).

1.6 Biotin identification (BioID) method

Immunoprecipitation or affinity purification followed by mass spectrometry (MS) has been one of the most used procedures to determine protein-protein interactions (PPI). The proximity labelling of proteins currently can be conducted using Biotin Identification (BioID). BioID is based on ligating an enzyme, biotin ligase, with the protein of interest (such as DDX41) and probing the interacting protein partners by MS (Roux et al., 2018) (**Figure 6**). The proteins present proximal to the protein of interest are called the "proxisome" (Roux et al., 2012). BioID overcomes the potential barriers of the conventional technique of studying PPI in that it allows the determination of the interactions between proteins in their natural environment as the protein biotinylation occurs preceding cell lysis and their solubilization. Also, this approach increases the possibility of locating the "proxisomes" in insoluble nuclear compartments like nucleoli or nuclear matrix. It facilitates the analysis of transient (short-time) interacting partners and the

limited abundant proteins (Roux et al., 2018; Ummethum & Hamperl, 2020). As biotinylation is a rare protein modification, it does not generate background signals during MS spectrometry (de Boer et al., 2003)

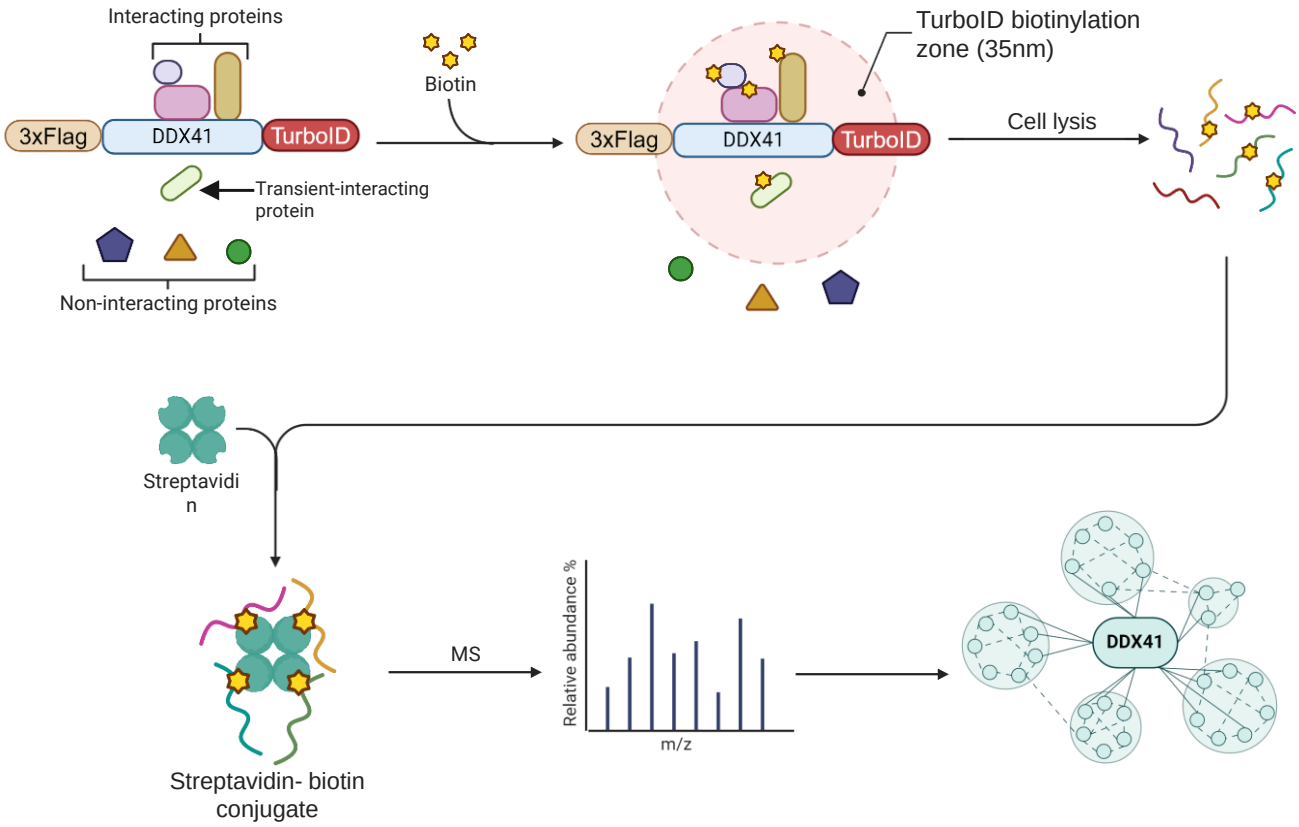


Figure 6. Basic mechanism and steps of BioID for DDX41. The protein of interest (human DDX41) is fused with a biotin ligase enzyme (TurboID) and expressed in human cells. TurboID converts exogenously added free biotin (yellow) to highly reactive and labile biotinyl-5′-AMP, which is released from the enzyme's active site to biotinylate on lysine residues on proximal proteins within the vicinity of labelling radius ~35 nm. Following biotin labelling, cells are lysed and proteins are extracted. Biotinylated proteins are then purified using streptavidin and identified by mass spectrometry**.** Created with BioRender.com

The advanced features like the increased labelling radius, short biotin labelling time or improved activity of the biotin ligase enzyme have manifested by the modifications in the BirA over the years. However, the basic mechanism of Bir A and its modified enzymes yet remains the same. The biotin ligase enzymes like Bir A catalyze the formation of biotinoyl adenylate (bioAMP) from biotin and ATP (Barker & Campbell, 1981a, 1981b; Eisenberg et al., 1982). The enzyme-bioAMP complex functions to biotinylate the lysine present at the biotinylation site in

biotin carboxyl carrier protein (BCCP). For the biotin identification of interacting proteins, the reaction was modified by fusing the biotin acceptor peptide (BAP) with a bait (protein of interest, such as DDX41), co-expressing with BirA, and the lysine residue on BAP is biotinylated by BirA (Smith et al., 1998); the biotinylated protein can then be purified using streptavidin cascade (de Boer et al., 2003). Another alteration to the methodology was done by mutated *E. coli* BirA* (R118G) by dissociation of BirA and bioAMP (Kwon & Beckett, 2000; Kwon et al., 2000), where biotinoyl adenylate reacts with a lysine residue based on the proximity of the substrates to BirA (Choi-Rhee et al., 2004; Cronan, 2005). A small-sized biotin ligase, BioID 2 of 28 kDa, was curated from the bacterium *Aquifex aeolicus* (Kim et al., 2016), and another mutated form of the enzyme known as TurboID was engineered which facilitates lessening the labelling time from 18 hours to 10 minutes (Branon et al., 2018).

1.7 Dynamic expression and subcellular localization of DDX41 under stress conditions

In Dr. Wu's lab, the former members examined DDX41 protein expression under stress conditions and found that DDX41 expression was increased after microbial, radiation and chemical stress treatments. For example, DDX41 started expressing at 1 hour after Herpes simplex virus type 1 (HSV-1) infection in HeLa cells, reached its peak at 4 hours, then decreased from 6 to 24 hours (**Figure 7A**). Similarly, DDX41's expression increased at 15 min after ionizing radiation (IR), UV, or hydroxyurea (HU) treatment, peaked at 1-2 hours, then declined to 8 hours, and even completely disappeared up to 72 hours. A similar pattern of endogenous and GFP-tagged DDX41 expression was noticed upon treatment with bleomycin, a type of antibiotic which induces breaks in DNA strands (**Figure 7B**). Both endogenous and GFP-tagged DDX41 formed a pattern of increasing foci until 4 h and decreased by 24 h. They also observed that transiently expressed GFP-tagged DDX41 in HeLa cells was localized predominately in the nucleus before poly (dA:dT) transfection, but translocated to the cytoplasm 4-6 h post-transfection (**Figure 7C**). Collectively, their results suggest that DDX41 undergoes dynamic expression and subcellular localization under stress conditions. However, the factors regulating the dynamic expression and subcellular localization of DDX41 under normal or stress conditions are unknown.

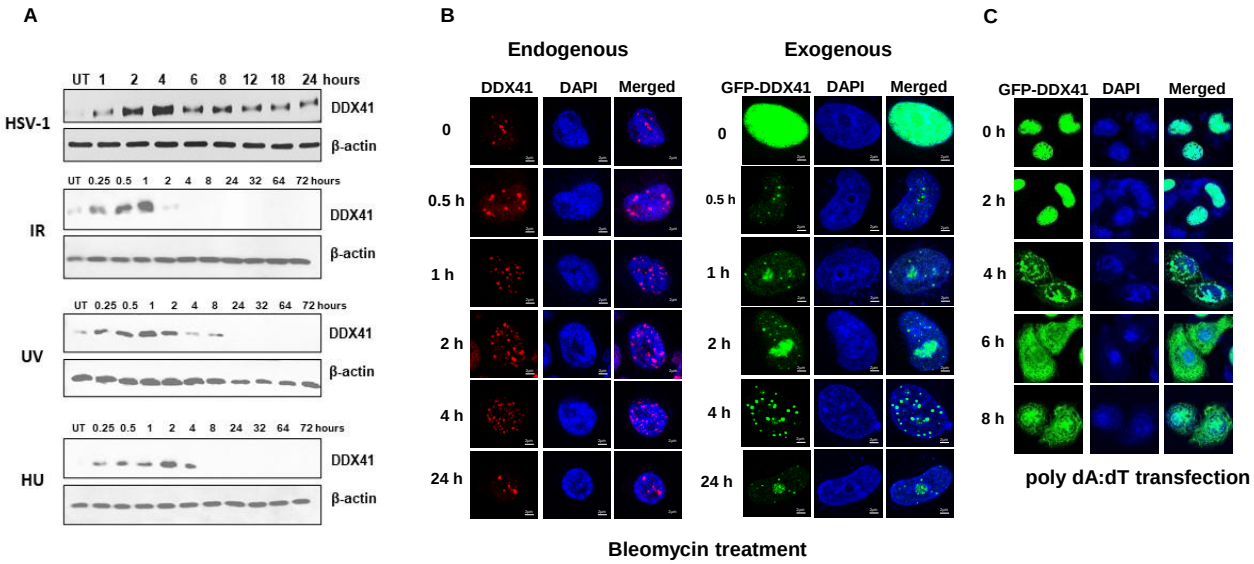


Figure 7. Dynamic expression and subcellular localization of DDX41 under stress. (**A**) Western blot analysis of DDX41 protein expression after stress treatment in HeLa cells. HSV-1, Herpes simplex virus type 1; IR, ionizing radiation, 5 Gy; UV, Ultraviolet, 50 J/m^2; HU, Hydroxyurea (10 mM). UT, untreated. β-actin serves as a loading control. (**B**) Immunofluorescence analysis of the expression of endogenous and exogenous DDX41 upon exposure to bleomycin induced stress. (**C**) Subcellular localization of GFP-tagged DDX41 after poly (dA:dT) transfection in HeLa cells (unpublished, experiments done by Dr. Ravi Shankar Singh, Aanchal Aggarwal and Ananna Arna).

2. HYPOTHESIS AND OBJECTIVES

2.1 Hypothesis: The functions and subcellular localization of DDX41 are regulated by its interacting proteins.

To test the hypothesis, our first and foremost priority was to establish a system that could facilitate the identification of DDX41 interactome *in vivo* under different tailored experimental conditions. Hence, we chose the BioID system and constructed a recombinant vector, consisting of target gene DDX41 that was tagged with Flag at one end and TurboID gene (BioID ligase protein) at another end. This study details the construction, optimization steps and purification of BioID fusion proteins by the following objectives.

2.2 Objectives:

I. Construct BioID fusion constructs [3xFlag-DDX41-TurboID and 3xFlag-TurboID (control)].

II. Validate the expression, subcellular localization of BioID fusion proteins and biotinylation activity of TurboID enzyme.

III. Isolate BioID fusion proteins by Flag immunoprecipitation (IP) and isolate the potential interacting proteins of DDX41 by streptavidin pulldown and identify the interacting proteins of DDX41 by mass spectrometry.

3. MATERIALS AND METHODS

3.1 Table 1: List of reagents used

Reagents, Cat No.	**Suppliers**
2-Mercaptoethanol, M3148	Sigma-Aldrich
Acrylamide, 0314	Amresco
Agarose, 1613101	Bio-Rad Laboratories
Ammonium persulfate (APS), A3678	Sigma-Aldrich
Ampicillin, A1593	Sigma-Aldrich
Biotin, BIO302	BioShop
Bovine serum albumin, A2058	Sigma-Aldrich
Bradford protein assay reagent, 500-0013	Bio-Rad Laboratories
Bromophenol blue, B0126	Sigma-Aldrich
Clarity Max Western ECL substrate, 1705062	Bio-Rad Laboratories
Dulbecco's Modified Eagle Medium (DMEM), 12800-082	Sigma-Aldrich
EDTA, 17892	Fisher Scientific
EcoRI HF, R3101	NEB
Fetal bovine serum (FBS), F1051	Sigma-Aldrich
Gibson assembly master mix kit, E2611L	NEB
Glycerol, 123170	Fisher Scientific
Glycine, BP381-5	Fisher Scientific
HindIII HF, R3104S	NEB
LB Broth with agar, L2897	Sigma-Aldrich
Magnesium chloride ($MgCl_2$), M8266	Sigma-Aldrich
Methanol, 154246	Sigma-Aldrich
Gibson Assembly Master Mix, E2611S	NEB
N,N,N',N'-Tetramethylethylenediamine (TEMED), 87689	Bio-Rad Laboratories
N,N'-Methylenebisacrylamide, AC164790250	Fisher Scientific
Opti-MEM I (1X) Reduced Serum Medium, 31985070	Gibco
Polyethylenimine (PEI 25K™), 23966	Polysciences
Penicillin-Streptomycin, SV30010	Cytiva
Phenylmethylsulfonyl fluoride (PMSF), P7626	Sigma-Aldrich
Pierce Protein A/G resin beads, PI20421	Thermo Fisher
Polyvinylidene difluoride membrane (PVDF), 10600023	Bio-Rad Laboratories
Precision Plus Protein Dual Color Standard, 1610374	BioRad Laboratories
ProLong Diamond Antifade Mountant with DAPI, P36962	Invitrogen
Pierce Protease Inhibitor Tablets, EDTA-free, A32965	Thermo Scientific
Sodium chloride (NaCl), S671	Fisher Scientific
Sodium dodecyl sulfate (SDS), L3771	Sigma-Aldrich
Streptavidin Sepharose high-performance beads, GE17-5113-01	Cytiva
Tris, 0826	AMRESCO
TritonTMX-100, X100	Sigma-Aldrich
Trypsin-EDTA, T4049	Sigma-Aldrich
Tryptone, TRP402	BioShop
Tween 20, BP337	Fisher Scientific

3.2 Table 2: List of primers used

Primer	Sequence (5´-3´)	Used in this study
AC-3FLAG-F	GCATAAGCTTGGCGGCACCATG GACTACAAAGACCATGAC	Forward primer to PCR amplify 3xFlag peptide DNA. Kozak consensus sequence is added before ATG. HindIII site (underlined) for pcDNA3.0 vector.
AC-3FLAG-R	GTTCGGGTTCCGACTCCTCCATC TTGTCATCGTCATCCTTGT	Reverse primer to PCR amplify 3xFlag peptide DNA. 20 nucleotides overlap with AC-DDX41-F.
AC-DDX41-F	ACAAGGATGACGATGACAAGAT GGAGGAGTCGGAACCCGAAC	Forward primer to PCR amplify DDX41 gene. 20 nucleotides overlap with AC-3FLAG-R
AC-DDX41-R	TTGGGGATGGGCTTGCCCATGA AGTCCATGGAGCTGTGGG	Reverse primer to PCR amplify DDX41 gene. 20 nucleotides overlap with AC-TurboID - F
AC-Bir A-F	CCCACAGCTCCATGGACTTCATG GGCAAGCCCATCCCCAA	Forward primer to PCR amplify TurboID gene. 20 nucleotides overlap with AC-DDX41-R.
AC-Bir A-R	GCATGAATTCTTAGTCCAGGGTC AGGCGCTC	Reverse primer to PCR amplify TurboID gene, EcoRI site (underlined) for pcDNA3.0 vector.
AC-BirA-F-fusion 3xFlag	ACAAGGATGACGATGACAAGAT GGGCAAGCCCATCCCCAA	Forward primer to PCR amplify BirA gene. 20 nucleotides overlap with AC-3FLAG-R.
AC-3FLAG-R-fusion BirA	TTGGGGATGGGCTTGCCCATCTT GTCATCGTCATCCTTGT	Reverse primer to PCR amplify3xFLAG peptide DNA. 20 nucleotides overlap with AC-BirA-F.
AC-NES-Del-F	TAAGAATTCTGCAGATATCCATC AC	Forward primer to remove NES at the C-terminal of the TurboID gene.
AC-NES-Del-R	CTTTTCGGCAGACCGCAG	Reverse primer to remove NES at the C-terminal of TurboID gene.

3.3 Table 3: List of antibodies used

Antibody, Cat No.	Suppliers
Flag, 20543-1-AP	Proteintech
DDX41 (D3F1Z) Rabbit mAb, 15076	Cell Signalling Technology (CST)
Bir A, 6C4c7	Novus Biologics
Streptavidin HRP, ab7403	Abcam
IRF-3 (D83B9) Rabbit mAb, 4302	Cell Signalling Technology (CST)
STING (D2P2F) Rabbit mAb, 13647	Cell Signalling Technology (CST)
cGAS (D1D3G) Rabbit mAb, 15102	Cell Signalling Technology (CST)

TBK-1, 3013	Cell Signalling Technology (CST)
DHX8, 50-155-8070	Fisher Scientific
β-actin-HRP, sc-47778	Santa Cruz
Goat anti-Rabbit IgG (H+L) Cross-Adsorbed Secondary Antibody, Alexa Fluor594, A-11012	Fisher Scientific
Goat anti-Rabbit IgG (H+L) Cross-Adsorbed Secondary Antibody, Alexa Fluor488, A-11008	Fisher Scientific
Goat anti-Mouse IgG (H+L) Highly Cross-Adsorbed Secondary Antibody, Alexa Fluor 594, A-11032.	Fisher Scientific
Goat anti-mouse IgG HRP-conjugated, 7076	NEB
Goat anti-rabbit IgG HRP-conjugated, 7074	NEB

3.4 Plasmid DNA

The individual DNA fragments, namely 3xFlag, DDX41 and TurboID, were amplified using PCR and were attempted to assemble by overlap PCR using their respective primers (Table 2) and cloned into pcDNA3.0 vector. The control BioID fusion construct, i.e. pcDNA 3.0-3xFlag-TurboID, was constructed using the assembled DNA fragments i.e 3xFlag-TurboID and cloned into pcDNA3.0 vector by traditional cloning using the restriction enzyme sites of HindIII at the N-terminus and EcoRI at the C-terminus. pcDNA3.0-3xFlag-DDX41-TurboID was constructed by using assembled DNA fragments of 3xFlag-DDX41 and TurboID, followed by cloning them using the Gibson Assembly (NEB). To remove the nuclear export sequence (NES) from the C-terminus region of TurboID, primers (Table 2) were designed according to the NEB website (https://nebasechanger.neb.com) and were synthesized by IDT. NES-deleted BioID fusion constructs were constructed using Q5 site-directed mutagenesis (NEB). All plasmid DNAs were verified by DNA sequencing.

3.5 Cell culture

Cells were cultured in 10-cm cell culture dishes at 5% CO_2 and 37°C incubator conditions. HEK293T (ATCC CRL-3216) and HeLa DDX41 knockout (KO) (Singh et al., 2022) were maintained in Dulbecco's Modified Eagle's Medium (DMEM) media. The DMEM media was supplemented with 10% fetal bovine serum (FBS) and 50 µg/mL penicillin-streptomycin (all from Sigma-Aldrich). HeLa DDX41 KO cell lines were generated by Dr. Ravi Shankar Singh in Dr. Wu's lab (Singh et al., 2022).

3.6 Cell transfection

Transient transfection was performed using a 1:3 ratio of the plasmid DNA to polyethyleneimine (PEI, Polysciences). For 10-cm cell culture dishes, 10 µg of plasmid DNA were mixed with 150 µl of Opti-MEM (Minimal Essential Medium, Gibco) media to form Solution A. Thirty microlitres of PEI were mixed with 150 µl of 0.15 M sterile NaCl to form Solution B. Then, the two solutions (A and B) were gently vortexed and incubated at room temperature (RT) for 10 min. After the incubation, solution A and solution B were combined by pipetting, and the DNA-PEI mixture was incubated at RT for 30 min. Meanwhile, cells were subjected to washing with 1XPBS (137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, and 1.8 mM KH_2PO_4) and replacing DMEM growth media with Opti-MEM (7.5 ml for 10 cm dishes and 1.5 ml for 6-well plates). After the incubation, the DNA-PEI mixture was slowly added to the cells from one side of the cell culture dishes (to prevent detachment of adherent cells). The transfected cells were incubated for 4-6 h in Opti-MEM growth media. After 4-6 h of incubation, the Opti-MEM media was changed back to DMEM growth media.

3.7 Biotinylation

A 200 mM stock solution was prepared by dissolving D-biotin (BioShop) in water and the pH was adjusted to pH 7.4 by 1 N NaOH. The biotin solution was filtered by a 0.45 µm syringe filter. For treatment, a working concentration of 50 µM biotin solution was added to fresh DMEM media, which was then gently dispensed onto the cells from the side of the dishes to prevent cell detachment. The treated cells were then incubated for 30 min before harvesting.

3.8 Western blotting

Cell lysates were collected at 4°C using cold RIPA (Radioimmunoprecipitation Assay) buffer (50 mM Tris-HCl, 150 mM NaCl, 1% Triton X-100, 0.5% sodium deoxycholate, 0.1% SDS, 1 mM EDTA, 10 mM NaF)], supplemented with 0.1 µM phenyl methyl sulfonyl fluoride (PMSF) and 1X protease inhibitor (from stock concentration of 25X, Thermo Scientific) at the time of cell lysis. The cells suspended in the lysis buffer were sonicated thrice using Branson Digital Sonifier SFX 250. The cells were sonicated using 30 A for 10 s per cycle, with 30 s of incubation on ice in between each cycle. After sonication, the cells were kept on ice for 30 min and subsequently centrifuged at 12,000 g for 30 min. The supernatant containing total proteins,

whole cell lysate (WCL), was transferred to fresh microcentrifuge tubes. The total protein was measured by Bradford assay, with RIPA buffer as a control. Optical density (OD) was measured at 595 nm for each protein sample. Bovine serum albumin (BSA) was used as a standard protein.

The protein samples were denatured at 95°C for 3-5 min. The denatured proteins were loaded on 8% polyacrylamide gel and electrophoresis was carried out at 90 V for 120 min. The proteins on the gel were then transferred onto the methanol-activated 0.45 µm polyvinylidene difluoride (PVDF) membrane (Bio-Rad), by a wet transfer method at 90 V for 150 min in the cold room. It was then gently rinsed in 1XPBST (1XPBS buffer with 0.1% Tween-20) solution followed by blocking the proteins in 3% BSA in 1XPBST on a shaker at RT for 45-60 min. After the blocking, the membrane was rinsed thoroughly with 1XPBST solution and incubated with the target primary antibodies (1:1000 dilution) overnight on a shaker in the cold room. The next day, the membranes were washed thrice with 1XPBST solution for 10 min each and then incubated with the respective species-specific secondary antibodies (1:6000 dilution) for 1 h at RT on the shaker. The membranes were washed thrice with 1XPBST for 10 min each time. The proteins were detected by Clarity Max Western ECL substrate and were visualized using the ChemiDoc Imaging System (both from Bio-Rad).

3.9 Immunoprecipitation to purify BioID fusion proteins

HEK293T cells were transfected with DDX41 and control BioID fusion constructs, and then treated with 50 µM biotin for 30 min prior to cell harvesting. WCLs were incubated with a Flag antibody or streptavidin conjugated to horseradish peroxidase (HRP) on a rotator at 4°C overnight. The protein-Flag antibody or biotin-streptavidin complexes were added to 10 µL of equilibrated protein A/G agarose beads, followed by incubation on a rotator at 4°C for 4 h. In the case of streptavidin-sepharose beads, the incubation process was performed overnight at 4°C. The bait protein-antibody-bead complexes were then washed thrice with 300 µL of washing buffer (20 mM Tris-Cl, 150 mM NaCl, 1 mM EDTA, 1% Triton X-100, 2.5 mM sodium pyrophosphate, 1 mM β-glycerophosphate, 1 mM sodium vanadate, pH 7.4, and 1X protease inhibitor cocktail) and finally eluted using 2X laemmlli buffer (0.125 M Tris-Cl, 4% SDS, 20% glycerol, 10% β-mercaptoethanol, 0.004% bromphenol blue, pH 6.8). The beads containing the purified biotinylated DDX41, and its interacting proteins were denatured and detected by Western

blotting.

3.10 Immunofluorescence microscopy

HeLa DDX41 KO cells were seeded in 6-well plates containing PBS-washed coverslips. After the cells adhered to the coverslips, 2 µg of BioID fusion constructs were transfected into the cells using PEI-based methods as described above in section **3.6**. After 24 h of transfection, the individual coverslips were collected and washed with cold 1XPBS. The cells were fixed with 100% ice-cold methanol (pre-chilled at -20°C) for 15 min, followed by rinsing with 1XPBS. The fixed cells were permeabilized by acetone (pre-chilled at -20°C) for a min followed by rinsing with 1XPBS. The proteins were blocked by incubating the cells with blocking buffer [3% BSA, 0.1% Tween 20 and 4X saline sodium citrate (SSC) (from 20X stock solution) in PBS] for 1 h at RT on a shaker, followed by three washings with 1XPBST for 5 min each time. The cells were then incubated with primary antibodies i.e. Flag, DDX41 and BirA antibodies respectively, prepared at a dilution ratio of 1:300 in the blocking buffer at 4°C overnight. The next day, the cells on the coverslips were washed with 1XPBST thrice for 5 min each time. Then, the cells were incubated with secondary antibodies conjugated to fluorescent dyes Alexa Fluor 488 and Alexa Fluor 594 at a dilution ratio of 1:900 and incubated at RT (in the dark) for an hour on the shaker. The cells were then washed with 1XPBST thrice for 5 min each time, followed by mounting with Prolong Diamond antifade reagent containing DAPI (Invitrogen) and rested at RT in the dark overnight. The cells were examined on a Zeiss LSM 700META inverted Axiovert 200 M laser scan microscope with a Plan-Apochromat 63x/1.4 oil DIC objective. Images were captured with a CCD camera and analyzed using LSM Browser software ZEN (Zeiss).

4. RESULTS

4.1 Construction of BioID fusion constructs

We chose the mammalian expression vector pcDNA3.0 as the backbone of the BioID fusion constructs. To prepare the BioID fusion constructs, pcDNA3.0-3xFlag-DDX41-TurboID and pcDNA3.0-3xFlag-TurboID (control), we amplified individual DNA fragments, namely 3xFlag (100 bp, lane 1), DDX41 (2 kb, lane 2) and TurboID (1.2 kb, lane 3, **Figure 9A**). Thereafter, the DNA fragments were assembled by overlap PCR (Hilgarth & Lanigan, 2020), we could assemble 3xFlag-TurboID (1.3 kb, lane 1& 2 **Figure 9B**) and 3xFlag-DDX41 (2.1 kb, lane 1 **Figure 9C**), but not 3xFlag- DDX41-TurboID . The 3xFlag-TurboID was cloned into pcDNA 3.0 using the traditional cloning method by using the restriction enzyme digestion sites of HindIII and *EcoR*I (**Figure 8A**). To clone the 3xFlag-DDX41 and TurboID into the pcDNA3.0 vector, we used the Gibson assembly method (**Figure 8B**). Finally, I digested the plasmid DNAs by the restriction enzymes HindIII and EcoRI and verified the inserts with their corresponding sizes, i.e. 1.3 kb for 3xFlag-Turbo ID (lane 1, **Figure 9D**) and 4 kb for 3xFlag-DDX41-TurboID (lane 2, **Figure 9D**). Furthermore, the sequences of the inserts in the constructs were verified by sequencing.

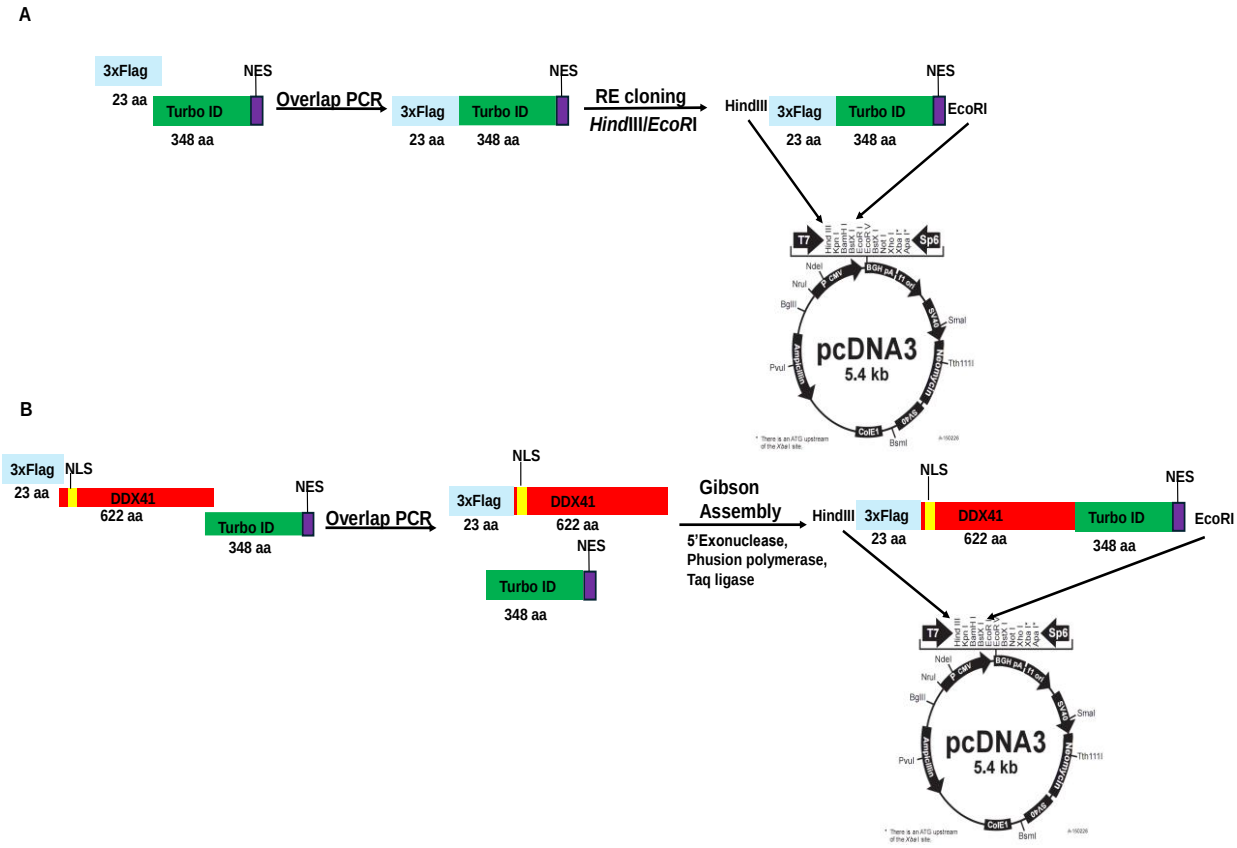


Figure 8. The strategy of assembly of BioID fusion constructs. The DNA fragments of 3xFlag, DDX41 and TurboID were amplified, assembled and cloned to form BioID fusion constructs using restriction cloning and Gibson assembly. **(A)** Cloning strategy of pcDNA3.0-3xFlag-TurboID: 3xFlag and TurboID were assembled to 3xFlag-Turbo ID by overlap PCR and cloned into the pcDNA 3.0 vector by using restriction enzymes HindIII and EcoRI at the N-terminus and C-terminus respectively. **(B)** Cloning strategy of pcDNA3.0-3xFlag-DDX41-TurboID: 3xFlag and DDX41 were assembled by overlap PCR. 3xFlag-DDX41 and TurboID were cloned into pcDNA3.0 by Gibson assembly.

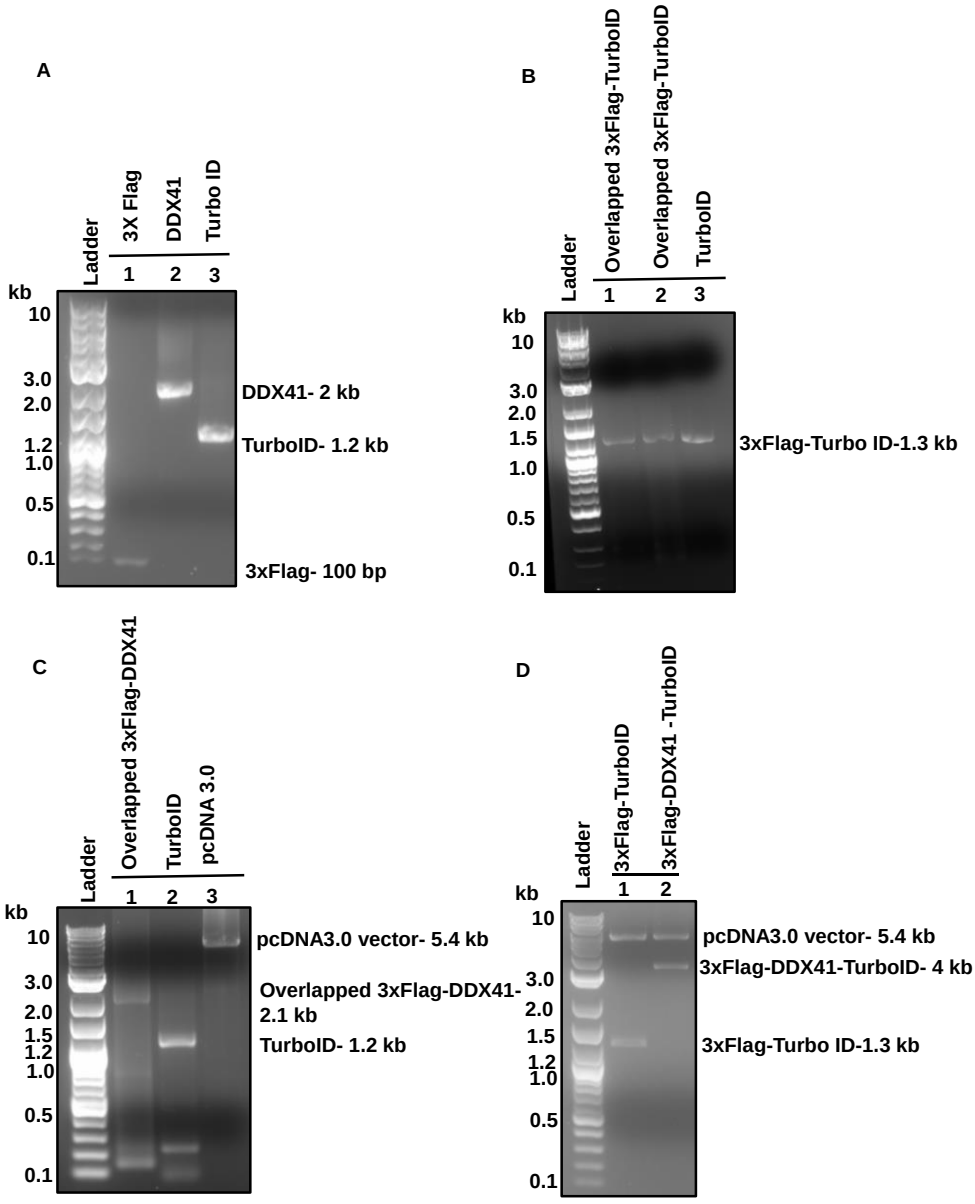


Figure 9. Amplified plasmids, overlap PCR plasmids and restriction enzymes' digestion products of BioID fusion constructs. (A) Amplified PCR products: amplified individual DNA fragments for 3xFlag aroundby 100 bp (lane 1), DDX41 of nearly 2 kb (lane 2) and TurboID at around 1.2 kb (lane 3) respectively. **(B)** Amplified overlap PCR products: overlapped 3xFlag-TurboID near 1.3 kb (lanes 1 & 2) and TurboID at nearly 1.2 kb (lane 3). **(C)** Amplified PCR products: overlapped 3xFlag-DDX41 at about 2.1 kb (lane 1), TurboID of nearly 1.2 kb (lane 2) and pcDNA3.0 of near to 5.4 kb (lane 3). **(D)** Restriction enzymes digestion products: Assembled BioID fusion constructs were digested by restriction enzymes HindIII and EcoRI. Lane 1 shows pcDNA3.0 vector of about 5.4 kb and 3xFlag-TurboID (control) of nearly 1.3 kb; lane 2 shows pcDNA3.0 vector of nearly 5.4 kb and 3xFlag-DDX41-TurboID nearly 4 kb respectively. DNA ladder in kb.

4.2 Expression of BioID fusion proteins

To verify the expression of the BioID fusion proteins, the fusion constructs were transfected into HEK293T cells and whole cell lysate (WCL) was collected after 24 h post-transfection. The expression of BioID fusion proteins i.e. 3xFlag-DDX41-TurboID and 3xFlag-TurboID (control) was validated using Flag, DDX41 and Bir A (against TurboID) antibodies respectively (**Figure 10 A-C**). I observed the expression of 3xFlag-DDX41-TurboID at 108 kDa with Flag, DDX41 and Bir A (TurboID) antibodies (lane 3 **Figure 10 A-C**), the expression of 3xFlag-Turbo ID (control) at 38 kDa with Flag and Bir A antibodies (lane 2 **Figure 10A and C),** and no expression for untransfected cells (lane 1 **Figure 10A-C)** by Flag (**Figure 10A)** and Bir A antibodies (**Figure 10C)**, respectively**,** except the presence of endogenous DDX41 at 70 kDa in response to DDX41 antibody (**Figure 10B)**. Overall, the observation of the DDX41 and control BioID fusion constructs suggests that they were effectively transfected and expressed inside the mammalian host cells.

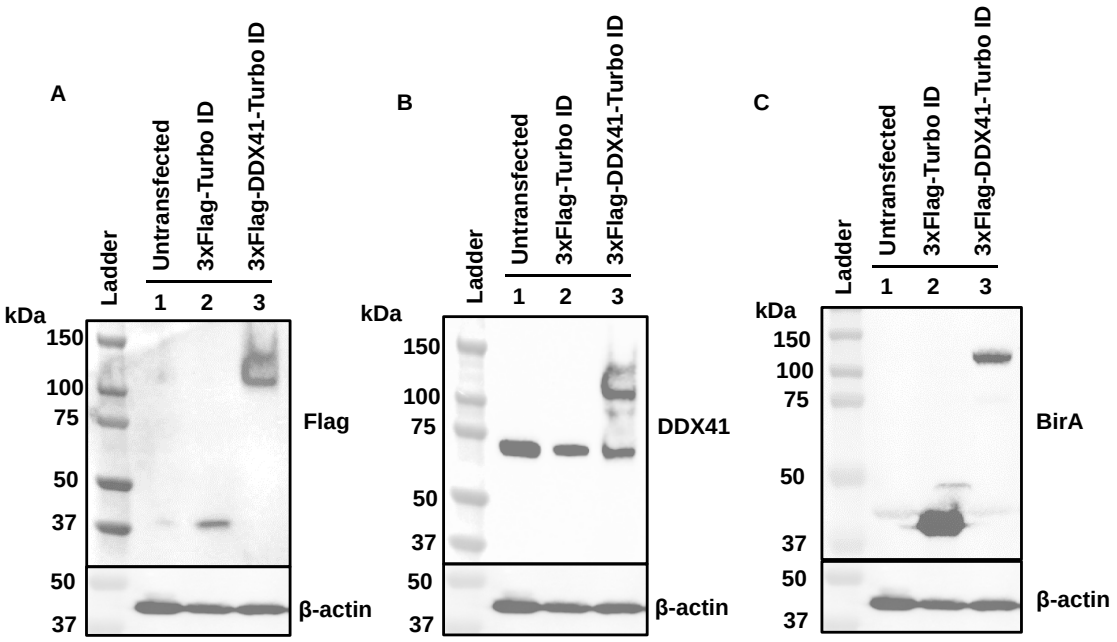


Figure 10. Expression of BioID fusion proteins. The fusion constructs of 3xFlag-TurboID and 3xFlag-DDX41-TurboID were transfected in HEK293T cells. The proteins from whole cell lysates (WCL) of untransfected HEK293T (lane 1), transfected 3xFlag-TurboID (lane 2) and transfected 3xFlag-DDX41-TurboID (lane 3) were identified by blotting with Flag (A), DDX41 (B), and Bir A (C) antibodies respectively. The WCL proteins of untransfected HEK293T proteins (lane 1) show expression of endogenous DDX41 around 70 kDa against anti-DDX41 antibody and no expression using Flag and Bir A antibodies, 3xFlag-TurboID (lane 2) expressed at 38 kDa against anti-Flag and anti-TurboID antibodies and 3xFlag-DDX41-TurboID (lane 3)

expressed at 108 kDa with anti-Flag, anti-DDX41 and anti-TurboID antibodies. β-actin serves as a loading control. Protein ladder in kDa.

Notably, while verifying the expression of BioID fusion proteins, I tried transient transfection in the HEK293T and HeLa DDX41 KO cells (previously established in our lab) (Singh et al., 2022). The HEK293T cells resulted in better transfection and expression of BioID fusion proteins relative to HeLa DDX41 KO cells. Therefore, we used HEK293T cells for most of our study experiments and HeLa DDX41 KO cells to determine the sub-cellular localization of BioID fusion proteins. Additionally, I tried to generate the BioID fusion proteins expressing HeLa DDX41 KO stable cell lines, but my attempts were unsuccessful.

4.3 Biotinylation activity of Turbo ID enzyme in BioID fusion proteins

To assess the effectiveness of the TurboID ligase enzyme in biotinylating the proximal proteins of DDX41 protein, I seeded HeLa DDX41 KO cells and transfected them with the BioID fusion constructs. Besides, I also used untransfected HeLa DDX41 KO cells as an additional control. After 24 hours of transfection, the cells were divided into two groups for the biotin and control treatments. Thereafter, the treatment group was treated with 50 µM biotin for 10 min before the cell harvesting and another group was left untreated, followed by the preparation of WCLs and Western blot. The biotinylated proteins were blotted using streptavidin-HRP. I observed strong expression of 3xFlag-TurboID at 38 kDa and the presence of biotinylated proteins (lane 4, **Figure 11**). For the WCL proteins from both untransfected cells and transfected cells with 3xFlag-DDX41-TurboID, I observed background biotinylation activity of TurboID irrespective of the biotin treatment (lanes 1-2 & 5-6, **Figure 11**)**.** Contrary to the expected results, I could not see a substantial expression of biotinylated 3xFlag-DDX41-TurboID by streptavidin-HRP or increased biotinylation activity of TurboID in biotin-treated proteins (lane 6, **Figure 11**).

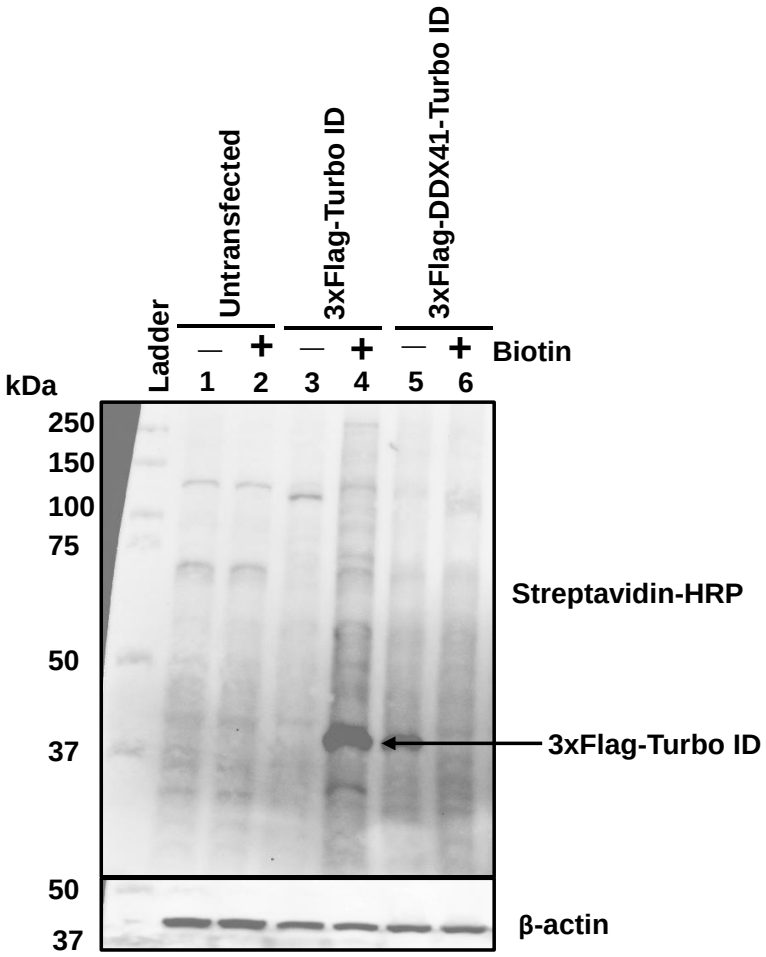


Figure 11. Biotinylation activity of TurboID after 10 min of biotin treatment. The fusion constructs of 3xFlag-TurboID and 3xFlag-DDX41-TurboID were transfected in HeLa DDX41 KO cells. The cells were treated with 50 μM biotin for 10 min. The biotinylation activity of TurboID and expression of biotinylated BioID fusion proteins were detected using streptavidin-HRP. The proteins from untreated and biotin-treated whole cell lysates of untransfected HeLa DDX41 KO (lanes 1 and 2) show the background activity of the TurboID enzyme, biotinylated expression of 3xFlag-TurboID at 38 kDa (lane 4) and no expression of 3xFlag-DDX41-TurboID expected at 108kDa (lane 6). β-actin serves as a loading control. "+" and "-" denote treatments with and without biotin, respectively. Protein ladder in kDa.

To optimize the biotinylation activity, I increased the biotin labelling time from 10 min to 30 min. In addition, I used HEK293T cells and transfected them with BioID fusion constructs. As mentioned above, I used the same treatment and experiment conditions. The biotinylated WCL proteins were detected by using streptavidin-HRP and the expression of BioID fusion proteins was examined using anti-Flag, anti-DDX41 and anti-TurboID antibodies.

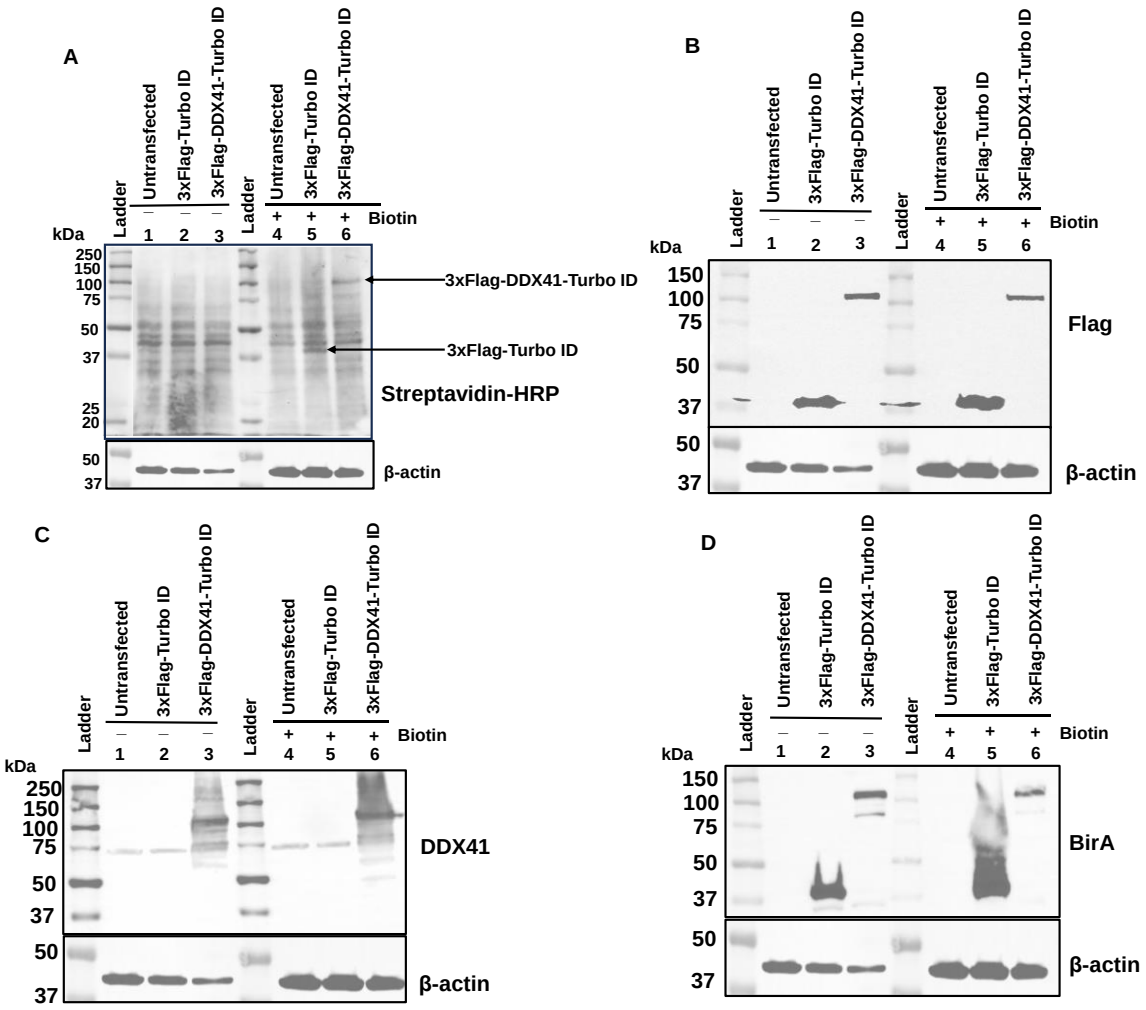


Figure 12. Biotinylation activity of TurboID after 30 min of biotin treatment and expression of BioID fusion proteins. The HEK293T cells were transient transfected with 3xFlag-TurboID and 3xFlag-DDX41-TurboID fusion constructs; and treated with 50 µM biotin for 30 min. The presence of putative biotinylated proteins was detected using Streptavidin-HRP and the expression of untreated and biotin-treated BioID fusion proteins was verified using antibodies against Flag, DDX41 and TurboID respectively. (**A**) Validation of the biotinylation activity of TurboID and expression of biotinylated proteins with streptavidin-HRP: Biotinylated 3xFlag-TurboID proteins expressing at 38 kDa (lane 5) and biotinylated 3xFlag-DDX41-TurboID proteins nearly at 108 kDa (lane 6). (**B-D**) Validation of expression of BioID fusion proteins using Flag (B), DDX41 (C), and TurboID (D) antibodies: The 3xFlag-TurboID proteins express at 38 kDa (lanes 2 and 5) and 3xFlag-DDX41-TurboID at 108 kDa (lanes 3 and 6). β-actin serves as a loading control. "+" and "-" denote treatments with and without biotin, respectively. Protein ladder in kDa.

With an increase in the biotin-treatment time, I observed a rise in the biotinylated expression of 3xFlag-DDX41-TurboID at 108 kDa (lane 6, **Figure 12A**) and 3xFlag-TurboID at

38 kDa (lane 5, **Figure 12A)**. The TurboID showed background biotinylation activity in the untreated proteins (lanes 1-3 **Figure 12A**) and an increased biotinylation activity in the biotin-treated proteins (lanes 4-6 **Figure 12A**). The background biotinylation activity in the untreated proteins represents endogenous biotinylated proteins of cells which utilize the biotin in the growth medium to maintain the growth of cells. Additionally, I checked the expression of BioID fusion proteins with Flag, DDX41, and BirA antibodies and the expression of 3xFlag-DDX41-TurboID and 3xFlag-TurboID was confirmed by the presence of expected protein bands at 108 kDa and 38 kDa (lanes 2 & 5 and 3 & 6), respectively (**Figure 12B-D)**. The increase in biotin treatment time resulted in the detection of biotinylated expression of DDX41 BioID fusion protein bands in contrast to negligible expression after 10 min of biotin incubation (**Figure 11**). Thus, I used 30 min of biotin incubation for the rest of the study.

4.4 Cytoplasmic localization of BioID fusion proteins

To determine the sub-cellular localization of BioID fusion proteins, I performed an immunofluorescence (IF) experiment for BioID fusion proteins. The BioID fusion constructs were transfected into HeLa DDX41 KO cells (Singh et al., 2022) and the IF protocol was performed as detailed in **section 3.10**.

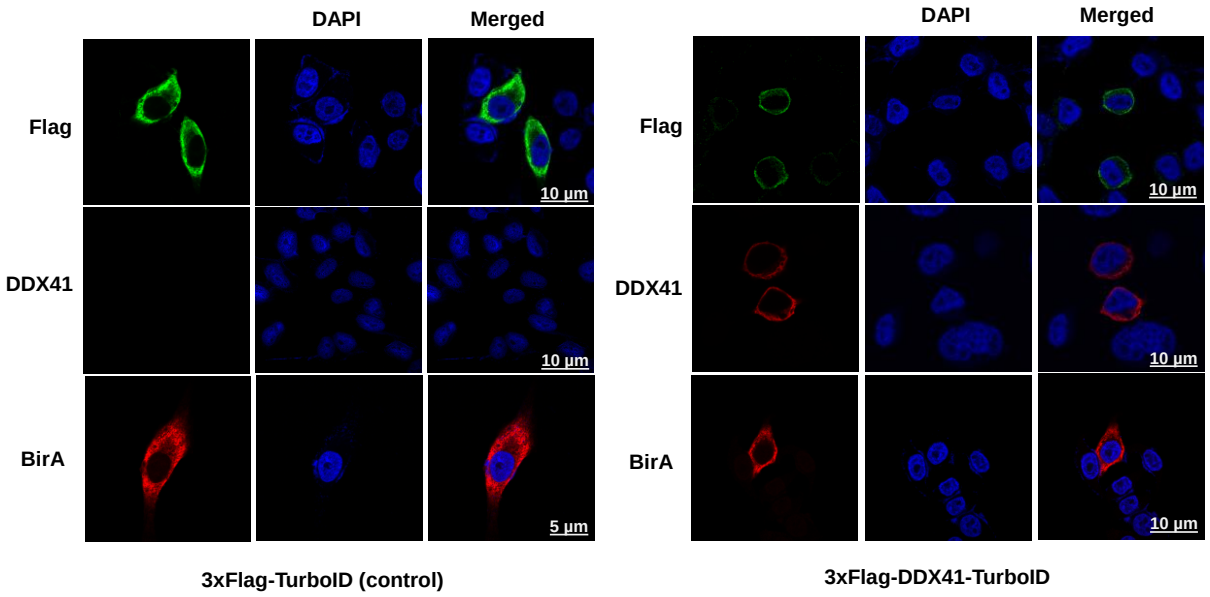


Figure 13. Sub-cellular localization of BioID fusion proteins. The sub-cellular localization of BioID fusion proteins was determined by transfecting 3xFlag-TurboID and 3xFlag-DDX41-TurboID constructs in HeLa DDX41 KO cells. The cells were collected at 24 h post-transfection

and the localization of BioID fusion proteins was identified using antibodies against Flag-tag, DDX41, and TurboID respectively. DNA was stained with DAPI (blue).

I found that both 3xFlag-DDX41-TurboID and 3xFlag-TurboID (control) were localized in the cytoplasm of the cells (**Figure 13**). Kadono *et al*. reported the nuclear localization of DDX41 (Kadono et al., 2016), and the observed subcellular localization of our DDX41 fusion protein was different from its endogenous localization. The observed cytoplasmic localization of DDX41 suggests that its potential interactome might differ from its nuclear interactome. Thus, we looked for the probable reasons for the unexpected cytoplasmic localization of the 3xFlag-DDX41-TurboID fusion protein.

4.5 Generation of modified BioID fusion constructs with a deleted nuclear export signal (NES)

One of the possible reasons for the cytoplasmic localization of DDX41 BioID fusion protein could be the dominant effect of nuclear export signal (NES, LPPLERLTL) at the C-terminus end of the TurboID, over the function of nuclear localization sequence (NLS, PERKRART) at the N-terminus of DDX41 (**Figure 8B and 14**). The NES functions to export proteins from the nucleus to the cytoplasm and the opposite function is performed by the NLS. As we aimed to observe the native interactome of DDX41 i.e. nuclear interactome, the DDX41 BioID fusion protein must be expressed in the nucleus as the native DDX41. Therefore, we removed the NES from the two TurboID constructs by site-directed mutagenesis as a trouble-shooting step (**Figure 14),** and the modified plasmids (NES-deleted BioID fusion plasmids) were verified by sequencing.

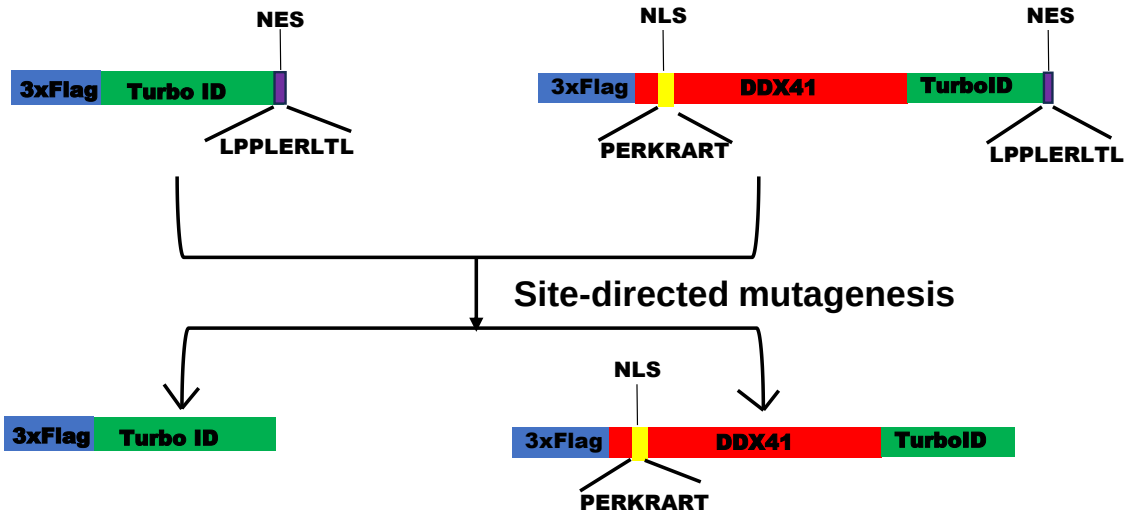


Figure 14. Construction of truncated BioID fusion constructs by deletion of NES. The nuclear localization sequence (NLS in yellow) in DDX41 and nuclear export sequence (NES in violet) in TurboID are indicated. The NES was removed by site-directed mutagenesis to obtain NES-deleted BioID fusion constructs (work done by Dr. Yuliang Wu and Shizhuo Yang).

4.6 Nuclear and cytoplasmic localization of BioID fusion proteins

To check the cytoplasmic/nuclear localization of the NES-deleted DDX41 BioID fusion constructs, they were transfected in HeLa DDX41 KO cells and their expression was checked using antibodies against Flag, DDX41 and TurboID. I found that the NES-deleted 3xFlag-DDX41-TurboID proteins were localized within the nucleus and NES-deleted 3xFlag-TurboID proteins were distributed in the cytosol (**Figure 15**). Thus, we decided to use these NES-deleted constructs for further work.

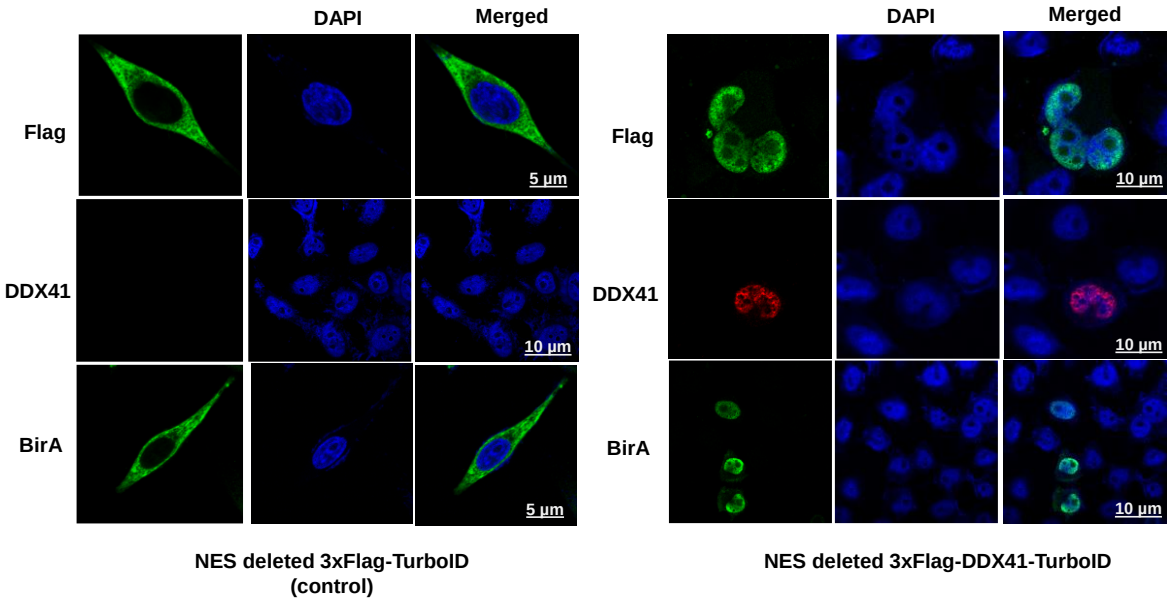


Figure 15. Sub-cellular localization of NES-deleted BioID fusion proteins. NES-deleted 3xFlag-TurboID and NES-deleted 3xFlag-DDX41-TurboID constructs were transfected in HeLa DDX41 KO cells. The cells were collected at 24 h post-transfection and the localization of NES-deleted BioID fusion proteins was detected using antibodies against Flag-tag, DDX41, and TurboID respectively. DNA was stained with DAPI (blue).

4.7 Expression of modified (with a deleted NES) BioID fusion proteins

To determine the expression of modified NES-deleted BioID fusion proteins, I transfected the modified fusion constructs into HEK293T cells and WCL was collected after 24 h post-transfection. I observed the expression of 3xFlag-DDX41-TurboID at 108 kDa using Flag, DDX41 and Bir A antibodies (lane 3, Figure 10A-C), the expression of 3xFlag-Turbo ID (control) at 38 kDa using Flag and Bir A antibodies **(**lane 2, **Figure 10A and C)**, and no expression for untransfected HEK293T proteins (lane 1, **Figure 10A-C)** by Flag **(Figure 10A)** and Bir A antibodies **(Figure 10C)**, except for the presence of endogenous DDX41 at 70 kDa in response to DDX41 antibody **(Figure 10B)**. Overall, the observation of the DDX41 and control BioID fusion constructs indicates that they were successfully transfected and expressed within the mammalian host cells.

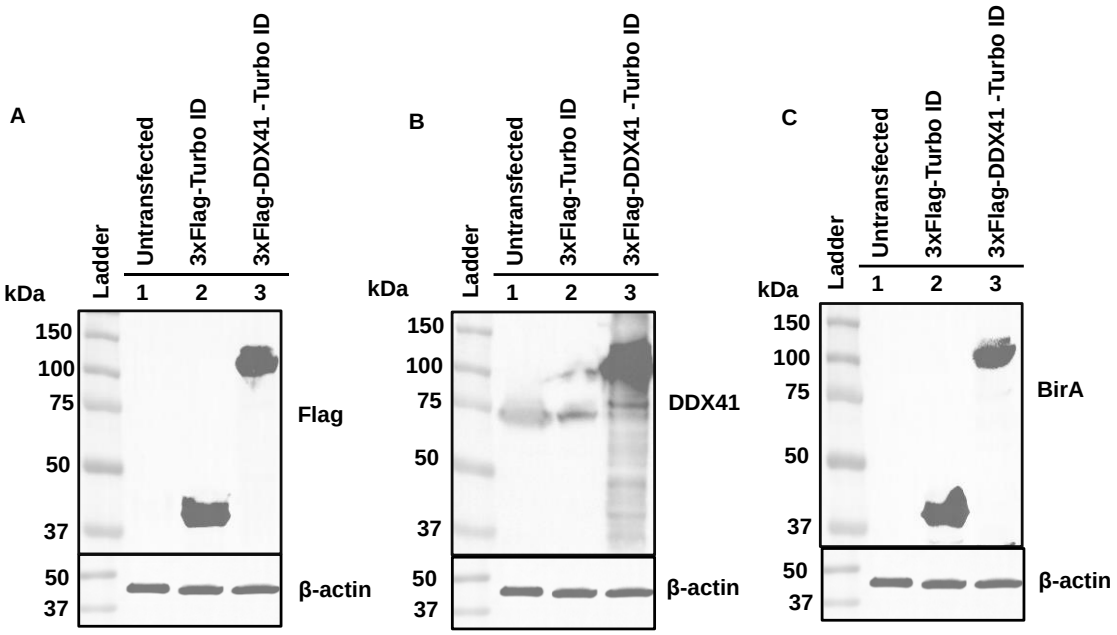


Figure 16. Expression of NES-deleted BioID fusion proteins. The HEK293T cells were transfected with 3xFlag-TurboID and 3xFlag-DDX41-TurboID constructs. The expression of proteins of untransfected HEK293T proteins, 3xFlag-TurboID and transfected 3xFlag-DDX41-TurboID from their respective whole cell lysate proteins were identified with Flag (A), DDX41 (B), and Bir A (C) antibodies respectively. The immunoblots showed no expression of untransfected HEK293T proteins (lane 1) against anti-Flag and anti-TurboID and endogenous DDX41 around 70 kDa against anti-DDX41 antibodies, 3xFlag-TurboID (lane 2) expressed at 38 kDa on blotting with anti-Flag and anti-TurboID antibodies and 3xFlag-DDX41-TurboID (lane 3) expressed at 108 kDa on blotting with anti-Flag, anti-DDX41 and anti-TurboID antibodies. β-actin serves as a loading control. Protein ladder in kDa.

4.8 Biotinylation activity of Turbo ID enzyme in the modified (with a deleted NES) BioID fusion proteins

To validate the biotinylation activity of the TurboID enzyme in the NES-deleted BioID fusion proteins, I used the same experiment and treatment conditions as described in section 4.3 i.e. 50 µM of biotin concentration and 30 minutes of biotin labelling time. The presence of biotinylated proteins was detected by using streptavidin-HRP. The Western blots resulted in the observation of biotinylated expression of 3xFlag-DDX41-TurboID at 108 kDa (lane 6, **Figure 17A**) and 3xFlag-TurboID at 38 kDa (lane 5, **Figure 17A**). Besides the biotinylated BioID fusion proteins, some additional biotinylated protein bands on the blot indicate the presence of potential interacting proteins of DDX41 (lane 6, **Figure 17A**). The additional biotinylated protein bands were not visible in the previous Western blot (lane 6, **Figure 12A**). Another observation was the background biotinylation activity of the TurboID in the untreated cells (lanes 1-3, **Figure 17A**)

and an increased biotinylation activity in the biotin-treated samples (lanes 4-6, **Figure 17A**). The expression of untreated and biotin-treated 3xFlag-DDX41-TurboID and 3xFlag-TurboID at 108 kDa and 38 kDa respectively (lanes 2 & 5 and 3 & 6, **Figure 17 B-D**) was confirmed by using antibodies against Flag-tag, DDX41, and TurboID.

Hence, the data suggested the optimum biotinylation activity of the TurboID enzyme at optimized biotin labelling time and concentration. The results also indicated the presence of potential interacting proteins of DDX41.

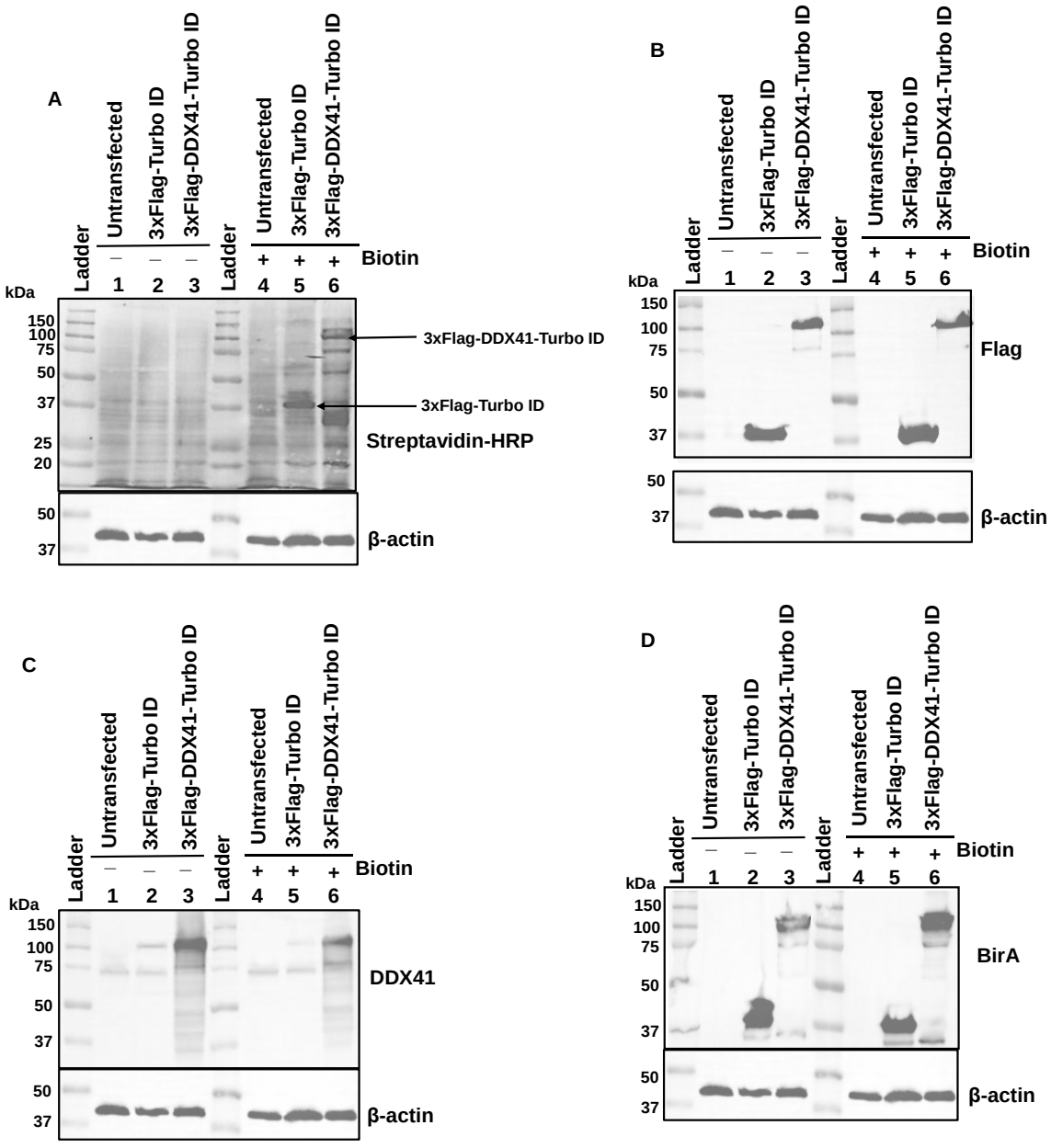


Figure 17. Biotinylation activity of TurboID and expression of NES-deleted BioID fusion proteins. The biotinylation activity of TurboID, the presence of biotinylated potential interacting proteins and the expression of BioID fusion proteins were investigated in HEK293T cells. The cells were transient transfected with BioID fusion constructs and treated with 50 µM biotin for 30 min before cell harvesting. (**A**) Validation of the biotinylation activity of TurboID and expression of biotinylated proteins was done with streptavidin-HRP: Biotinylated 3xFlag-TurboID express at 38 kDa (lane 5) and biotinylated 3xFlag-DDX41-TurboID express at 108 kDa (lane 6). (**B-D**) Detection of expression of the untreated and biotin-treated BioID fusion proteins by Flag, DDX41, and Bir A antibodies: 3xFlag-TurboID proteins express at 38 kDa (lanes 2 & 5) and 3xFlag-DDX41-TurboID at 108 kDa (lanes 3 & 6) against anti-Flag (B), anti-DDX41 (C), and anti-TurboID (D) antibodies. β-actin serves as a loading control. "+" and "-" denote treatments with and without biotin, respectively. Protein ladder in kDa.

4.9 Immunoprecipitation (IP) of BioID fusion proteins

4.9.1 Optimization of IP using a Flag antibody conjugated to protein A/G agarose beads

To purify the biotinylated DDX41 BioID fusion proteins, it was essential to optimize the IP protocol. Hence, I transfected HEK293T cells with NES-deleted BioID fusion constructs and treated them with 50 µM biotin for 30 min. To optimize the purity of elute proteins, I tried the pre-clearing of the WCLs because pre-clearing can help in the removal of lipids, carbohydrates, or nucleic acids from WCL (https://www.ptglab.com/news/blog/8-top-tips-for-immunoprecipitation). Therefore, to immunoprecipitate the biotinylated BioID fusion proteins, I divided WCLs into two groups, a pre-clearing WCL and a control WCL group.

In the pre-clearing group, The representative Western blots of pre-cleared 3xFlag-TurboID and 3xFlag-DDX41-TurboID **(Figure 18A and B)** showed the expression of several protein bands in WCL, pre-cleared lysate and flowthrough lanes, at 108 kDa and 38 kDa respectively and a single band at 108 kDa and 38 kDa in the elute lanes Similar results were corroborated by the anti-Flag antibody in the no pre-clearing group **(Figure 18C)**. The presence of biotinylated BioID fusion proteins was detected by using Streptavidin-HRP **(Figure 18D)**, by through the expression of biotinylated 3xFlag-DDX41-TurboID at 108 kDa and 3xFlag-TurboID at 38 kDa in WCL, flowthrough and elute lanes. Besides, a faint protein band (marked in circle) at approximately 80 kDa was observed, indicating the potential interacting protein of DDX41.

Overall, the data suggested that IP by the Flag conjugated to agarose beads could purify the Flag-tagged BioID fusion proteins but not the biotinylated potential interacting proteins of DDX41. The unyielding of biotinylated potential interacting proteins of DDX41 may be either because of their loss during the washing process or the sparseness of the proteins for elution.

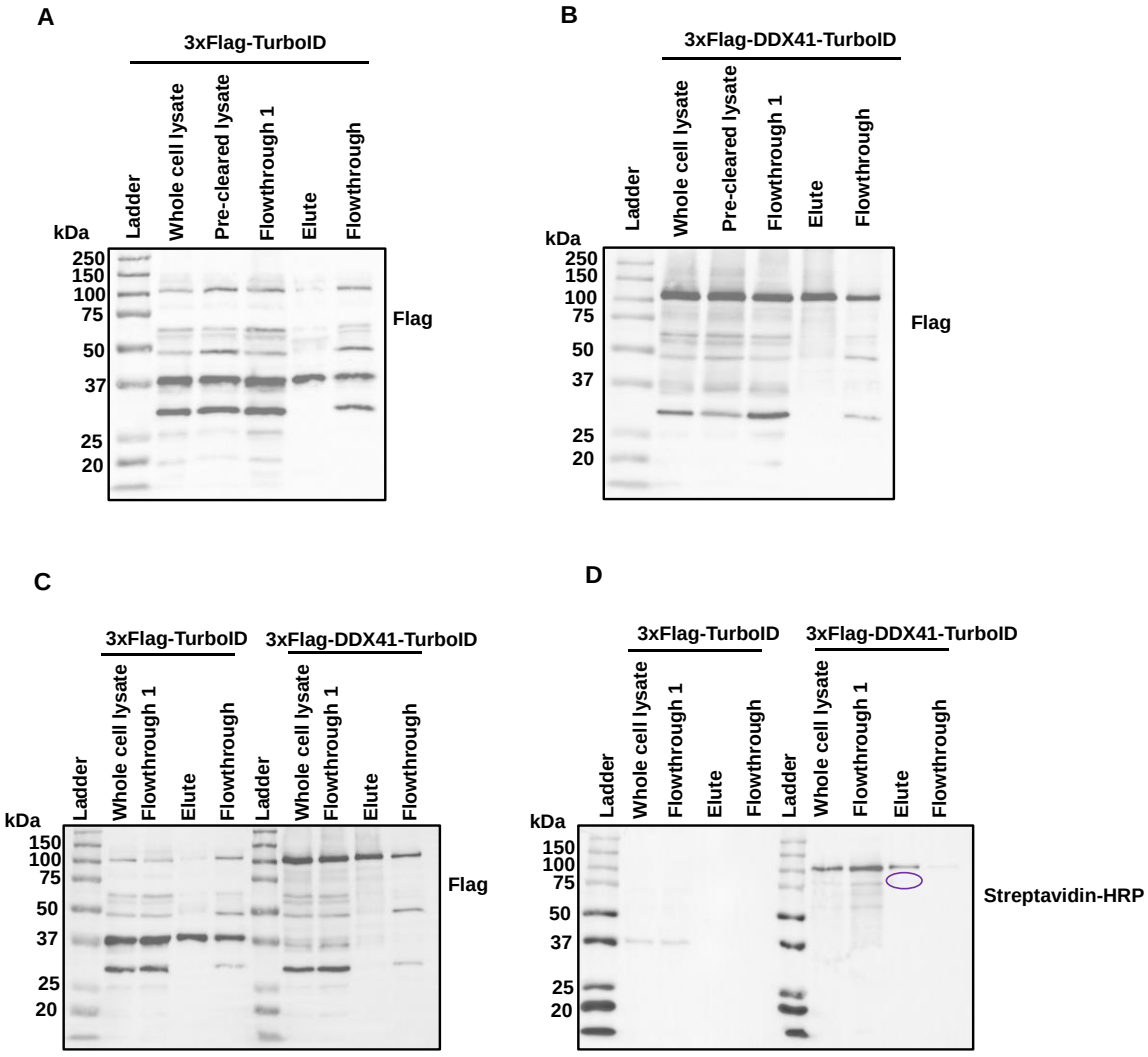


Figure 18. Optimization of immunoprecipitation of BioID fusion proteins. The immunoprecipitation protocol was optimized to isolate biotin-treated BioID fusion proteins and potential biotinylated interacting proteins of DDX41. The BioID fusion proteins were isolated by tagging to Flag conjugated with agarose beads. Immunoblots A and B represent pre-clearing groups of immunoprecipitated lysates of 3xFlag-TurboID and 3xFlag-DDX41-TurboID respectively. Immunoblot C and D represent the control group (no pre-clearing) of 3xFlag-TurboID and 3xFlag-DDX41-TurboID together on each immunoblot. The WCL, flowthrough and elutes of BioID fusion proteins were detected using Flag antibody (**A-C**). Potential biotinylated proteins were detected using streptavidin-HRP. **(A)** Pre-cleared lysate proteins of 3xFlag-TurboID express at 38 kDa. **(B)** Pre-cleared lysate proteins of 3xFlag-DDX41-TurboID express at 108 kDa. **(C)** Control lysates (no pre-clearing) of 3xFlag-TurboID and 3xFlag-DDX41-TurboID show expression at 38 kDa and 108 kDa respectively. **(D)** Biotinylated 3xFlag-TurboID express at 38 kDa in WCL and flowthrough 1 and biotinylated 3xFlag-DDX41-TurboID proteins express at 108 kDa in WCL, flowthrough 1 and elute. A faint band of around 80 kDa, indicated by circle depicts the presence of a potential DDX41-interacting protein. Protein ladder in kDa.

4.9.2 Repetition of IP by Flag antibody conjugated to protein A/G agarose beads

To follow up the immunoprecipitation of biotinylated BioID fusion proteins and to seek additional bands of interacting proteins of DDX41, I repeated the IP using the same experimental procedure as in **4.9.1** except for the pre-clearing step. I observed the expression of 3xFlag-DDX41-TurboID at 108 kDa by antibodies against Flag, DDX41, and TurboID **(Figure 19A-C)** and biotinylated expression by streptavidin-HRP (**Figure 19D);** expression of 3xFlag-TurboID at 38 kDa by antibodies against Flag and TurboID (**Figure 19A and C**) and biotinylated expression by streptavidin-HRP (**Figure 19D).** However, I could not see the previously visible protein bands of potential DDX41-interacting proteins (**Figure 19D)**, which might be due to loss in stringent washings.

Hence, the repetition of IP and Western blot analysis indicated that this experimental procedure could purify the biotinylated Flag-tagged BioID fusion proteins but was not suitable for isolating the biotinylated potential interactome of DDX41.

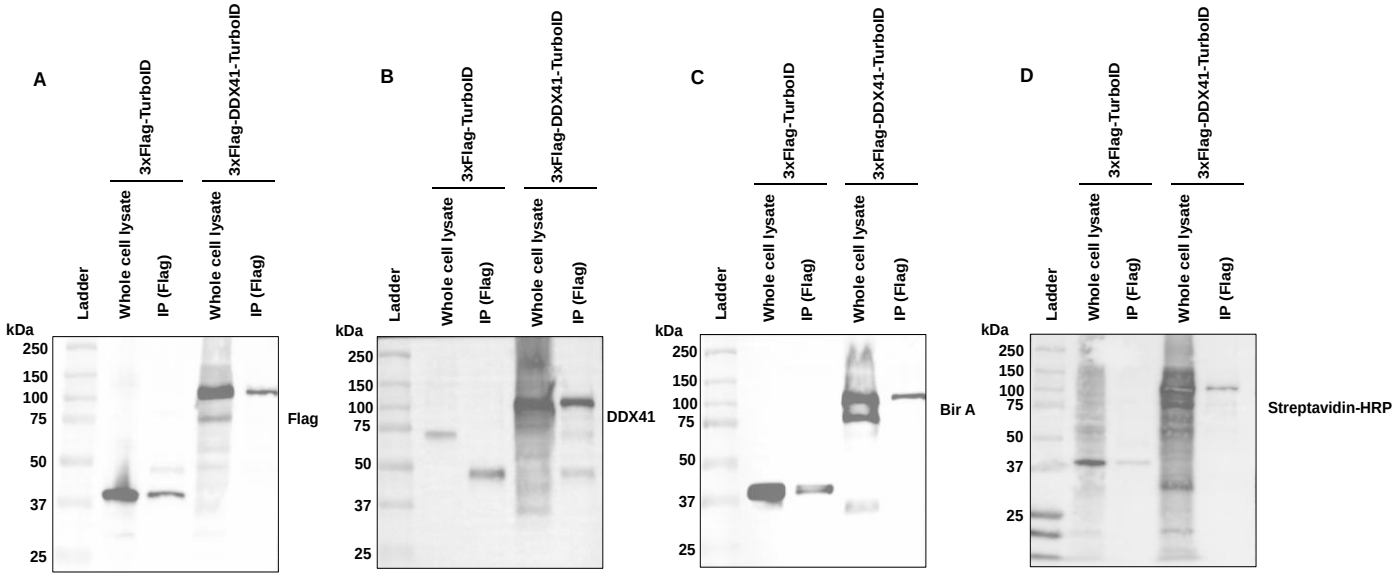


Figure 19. Immunoprecipitation of BioID fusion proteins. HEK293T cells were transfected with 3xFlag-TurboID and 3xFlag-DDX41-TurboID constructs followed by 50µM biotin treatment for 30 min. The whole cell lysate proteins were immunoprecipitated with the anti-Flag antibody in conjugation with the protein A/G agarose beads. The expression of WCL and purified proteins of biotin-treated 3xFlag-TurboID and 3xFlag-DDX41-TurboID was verified using Flag, DDX41 and Bir A antibodies and biotinylated proteins were detected by streptavidin-HRP. The WCL and elute (IP) proteins of 3xFlag-TurboID and 3xFlag-DDX41-TurboID express at 38 kDa and 108 kDa respectively by Flag (A), DDX41 (B), Bir A (C) antibodies, and biotinylated 3xFlag-

TurboID and 3xFlag-DDX41-TurboID proteins show expression at 38 kDa and 108 kDa by streptavidin-HRP (D). Protein ladder in kDa.

4.10 Precipitation of biotinylated BioID fusion proteins and potential interactome of DDX41

4.10.1 Optimization of pull-down of biotinylated BioID fusion proteins by using streptavidin sepharose beads

To purify the biotinylated BioID fusion proteins and potential interacting proteins, I tried to optimize the pull-down of biotinylated proteins using streptavidin-sepharose beads. I employed biotin-treated WCL proteins of 3xFlag-TurboID and 3xFlag-DDX41-TurboID and used the same protocol described in section **4.9** to purify the biotinylated proteins. The expression of purified biotinylated proteins from WCL, streptavidin-purified proteins (IP) and post-incubation supernatant (supernatant collected after the incubation of WCL with the streptavidin-sepharose beads overnight at 4^{o}C) was verified using streptavidin-HRP. To examine the abundance of biotinylated proteins in the elute, I tried two concentrations (30 µg and 60 µg) of elute proteins (lanes 2 & 3 and 6 & 7, **Figure 20**).

The Western blot showed strong expression of 3xFlag-TurboID at 38 kDa and 3xFlag-DDX41-TurboID at 108 kDa in WCL (lanes 1 & 5, **Figure 20**). In the elute/ IP lanes of 3xFlag-TurboID, another biotinylated protein at approximately 54 kDa (lanes 2 & 3) can be discerned. Additionally, weak biotinylation of purified proteins of 3xFlag-DDX41-TurboID (lanes 6 & 7) was seen with no prominent expression of the biotinylated protein of 3xFlag-DDX41-TurboID. In post-incubation supernatants, a single band at 38 kDa (lane 4) and 108 kDa (lane 8) corresponding to 3xFlag-TurboID and 3xFlag-DDX41-TurboID were observed, indicating the loss of some of the 3xFlag-DDX41-TurboID proteins with the supernatant before the washing steps.

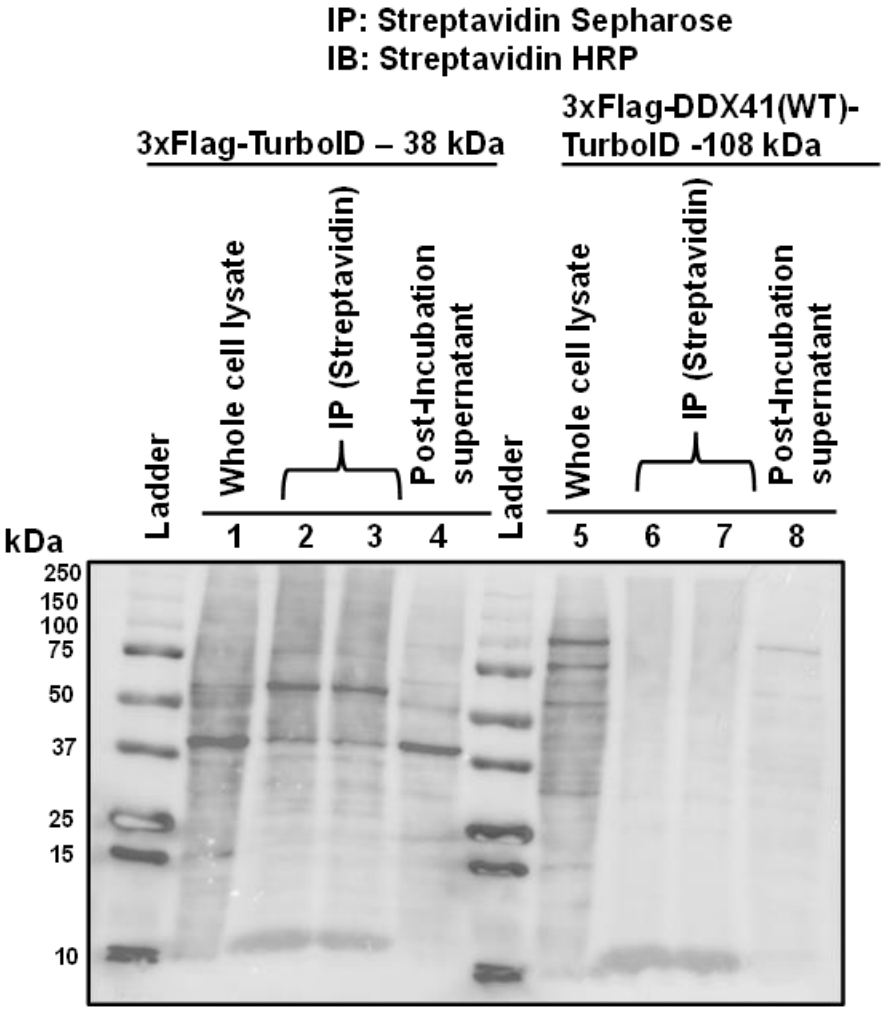


Figure 20. Optimization of pull-down of the BioID fusion proteins and biotinylated potential interactome of DDX41 by streptavidin sepharose beads. The putative biotinylated proteins from the biotin-treated whole cell lysates of 3xFlag-TurboID and 3xFlag-DDX41-TurboID (transfected in HEK293T cells) were pulled down by the bait "streptavidin" immobilized on sepharose. The biotinylated BioID fusion proteins and the potential interactome of DDX41 in the whole cell lysate and purified elutes (IP) were detected using streptavidin-HRP. The 3xFlag-TurboID protein band is expected to be about 38 kDa (lanes 1-4) and 3xFlag-DDX41-TurboID proteins to be about 108 kDa (lanes 5 and 8). A potential DDX41-interacting protein is visible at nearly 80 kDa (lane 5). Lanes 2 & 3 and 6 & 7 indicate different loading volumes of elute proteins. "+" and "-" denote treatments with and without biotin, respectively. Protein ladder in kDa.

Altogether, the Western blot analysis suggested that the volume of the streptavidin-sepharose beads was enough to purify the biotinylated 3xFlag-DDX41-TurboID proteins but not the 3xFlag-DDX41-TurboID and potential interactome. Therefore, I tested the volume of streptavidin-sepharose beads for purifying the biotinylated BioID fusion proteins and interacting proteins at the next step.

4.10.2 Immunoprecipitation of the biotinylated DDX41 interactome and validation of overexpression of biotinylated NES-deleted BioID fusion proteins

After the first trial, the pull-down procedure of biotin-treated and untreated BioID fusion proteins was repeated using streptavidin-sepharose beads followed by immunoblotting of the WCL and elute proteins. To assess the biotinylated interacting proteins of DDX41, Western blot analysis was done by using streptavidin-HRP (**Figure 21A)** and to verify the expression of BioID fusion proteins, the immunoblots were examined with the anti-Flag antibody (**Figure 21B)**.

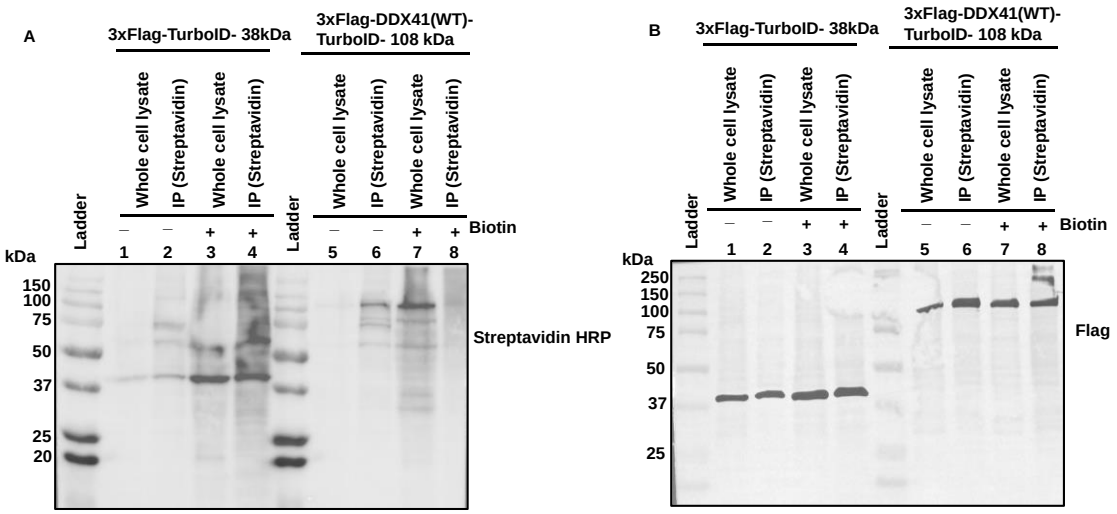


Figure 21. Pull-down of the BioID fusion proteins and biotinylated potential interactome of DDX41 by streptavidin sepharose beads. The putative biotinylated 3xFlag-TurboID and 3xFlag-DDX41-TurboID were pulled down from the untreated and biotin-treated whole cell lysates by using "streptavidin"- immobilized on sepharose. The BioID fusion proteins and the potential DDX41-interacting proteins in the whole cell lysate and purified elutes (IP) were detected using streptavidin-HRP. **(A)** Proteins detected using streptavidin-HRP. Expression of 3xFlag-TurboID WCL proteins (lanes 1 and 3) and IP proteins of about 38 kDa (lanes 2 and 4) and expression of 3xFlag-DDX41-TurboID WCL proteins (lanes 5 and 7) and IP proteins (lanes 6 and 8) at 108 kDa respectively. **(B)** Proteins detected using anti-Flag antibody. Expression of 3xFlag-TurboID WCL proteins (lanes 1 and 3) and IP proteins of about 38 kDa (lanes 2 and 4) and expression of 3xFlag-DDX41-TurboID WCL proteins (lanes 5 and 7) and IP proteins (lanes 6 and 8) at 108 kDa respectively. "+" and "-" denote treatments with and without biotin, respectively. Protein ladder in kDa.

From **Figure 21A**, the immunoblot displayed strong expression of biotinylated 3xFlag-TurboID in the biotin-treated WCL (lane 3) and purified IP (lane 4). A repeated appearance of a biotinylated protein band was visible at a size little above 50 kDa in lanes 2-4, which may be a biotinylated non-interacting protein.

In biotin-treated 3xFlag-DDX41-TurboID proteins, a faint expression at 108 kDa protein band in the elute/IP of biotin-treated (lane 8) relative to stronger expression in the WCL (lane 7) was still problematic. The weaker expression was yet suggestive of either the inefficient binding of biotinylated proteins with streptavidin or loss during the washings. Weak expression of 3xFlag-TurboID and 3xFlag-DDX41-TurboID at 38 kDa and 108 kDa in the untreated group of proteins suggests background biotinylation activity of TurboID. In **Figure 21B**, the expression of both biotin-treated and control group of 3xFlag-TurboID proteins and 3xFlag-DDX41-TurboID proteins was confirmed at their corresponding size of 38 kDa and 108 kDa respectively.

Collectively, we can infer two conclusions. First, the detection of the faint band of biotinylated 3xFlag-DDX41-TurboID in the elute/IP (lane 8, **Figure 21A**), indicates an improvement in the volume of streptavidin-sepharose beads used. Second, the expression of the BioID fusion proteins (untreated and biotin-treated) was validated with the anti-Flag antibody.

4.11 An alternative approach to purify the interacting proteins of DDX41 by streptavidin HRP conjugated to protein A/G agarose resin beads

To yield the biotinylated interacting proteins of the DDX41 as well as to cross-check the failed efficacy of the streptavidin-sepharose beads to pull-down biotinylated 3xFlag-DDX41-TurboID proteins, I performed a trial of adding streptavidin-HRP to protein A/G agarose beads. Hence, I performed the pull-down protocol by using the same experimental procedure by replacing Flag with streptavidin-HRP to purify the biotinylated BioID fusion proteins and interacting proteins of DDX41. Although, by the scientific rationale, this approach had more chances of failure and no expectation of obtaining results because, first, the agarose beads are coated with protein A and protein G that bind to the antibodies raised in animal species like rabbit and mouse. Secondly, streptavidin was not linked to the agarose by any spacer arm, like it is covalently linked in commercially available streptavidin-agarose beads, but interestingly the method worked.

The experimental setup was like the experimental conditions under section **4.10**. In the biotin-treated WCL and elutes/IP proteins of 3xFlag-TurboID and 3xFlag-DDX41-TurboID (lanes 3 & 4 and 7 & 8, **Figure 22**), I observed strong expression of 3xFlag-TurboID at 38 kDa and 3xFlag-DDX41-TurboID at 108 kDa, respectively relative to their untreated counterparts (lanes 1&2 and 5 & 6 **Figure 22).** The second crucial observation is the presence of strong

expression of a biotinylated protein band at nearly 80 kDa in the elute proteins of 3xFlag-DDX41-TurboID (lane 8, **Figure 22**).

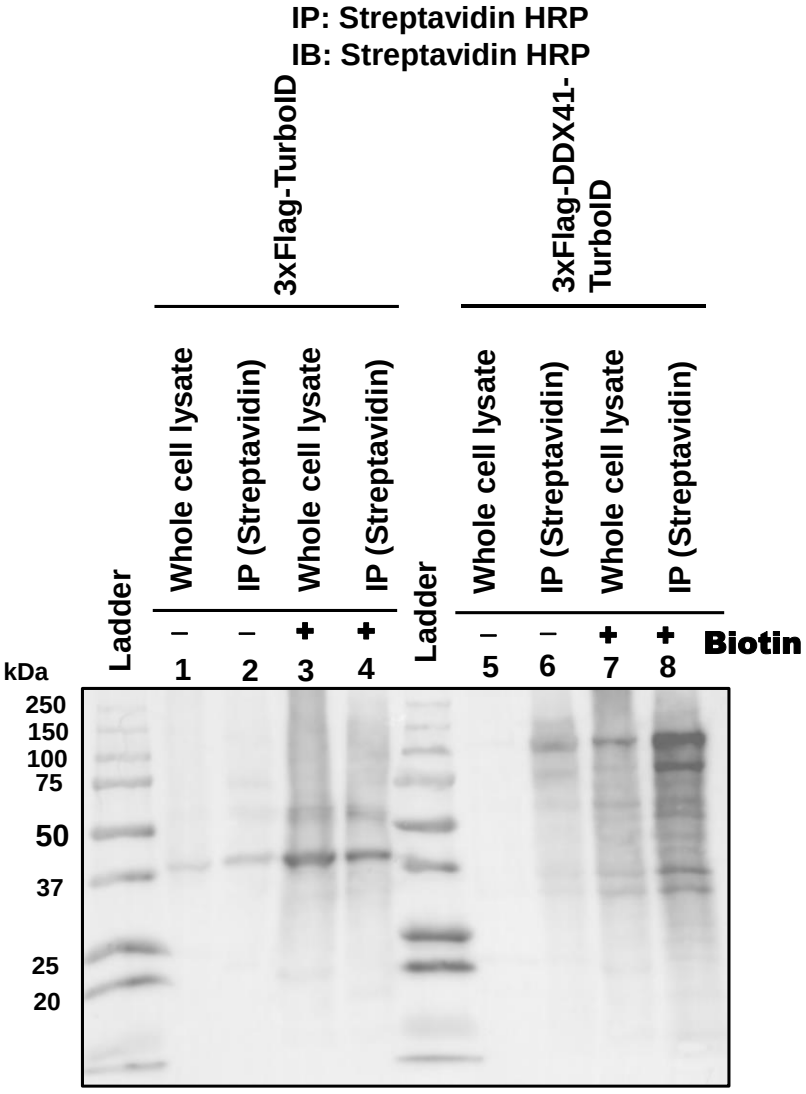


Figure 22. Proteins pulled down using streptavidin-HRP immobilized on protein A/G agarose beads. Proteins were pulled down from the whole cell lysates of untreated and biotin-treated 3xFlag-TurboID and 3xFlag-DDX41-TurboID (transfected in HEK293T cells) using streptavidin-HRP immobilized on protein A/G agarose beads. The proteins in the whole cell lysate and purified elutes (IP) were detected with streptavidin HRP. The 3xFlag-TurboID protein is about 38 kDa (lanes 1-4) and the 3xFlag-DDX41-TurboID protein is about 108 kDa (lanes 5-8). A potential DDX41-interacting protein is visible at about 80 kDa (lane 8). "+" and "-" denote treatments with and without biotin, respectively. Protein ladder in kDa.

Therefore, the key points from the experimental observations are the elution of the

biotinylated BioID fusion proteins purified by streptavidin-HRP and a strong presence of a potential interacting protein of nearly 80 kDa in the purified elute of NES-deleted 3xFlag-DDX41-TurboID proteins. The strong expression of DDX41 BioID fusion proteins and potential interactome protein bands suggest the enrichment of biotinylated proteins in elutes by this method.

4.12 Determination of the presence of potential interacting proteins of DDX41

To determine the presence of potential biotinylated interacting proteins of DDX41, I chose a few candidates (indicated below) that have been reported to interact with DDX41, in the context of two of DDX41 well-known functions, i.e. innate immunity and pre-mRNA splicing, which occurs in the cytoplasm and nucleus respectively (Andreou, 2021; Ma & Ross, 2024; Winstone et al., 2024).I used the WCL and IP proteins (obtained by streptavidin-sepharose pull-down of WCLs) of untreated and biotin-treated BioID fusion proteins. I immunoblotted them with antibodies against the stimulator of interferon genes (STING) (Zhang et al., 2011), cyclic GMP-AMP synthase (cGAS), interferon regulatory factor 3 (IRF-3), TANK-binding kinase 1 (TBK-1) (Jiang et al., 2017) and DEAH-box helicase 8 (DHX8)/PRP22) (Dybkov et al., 2023).

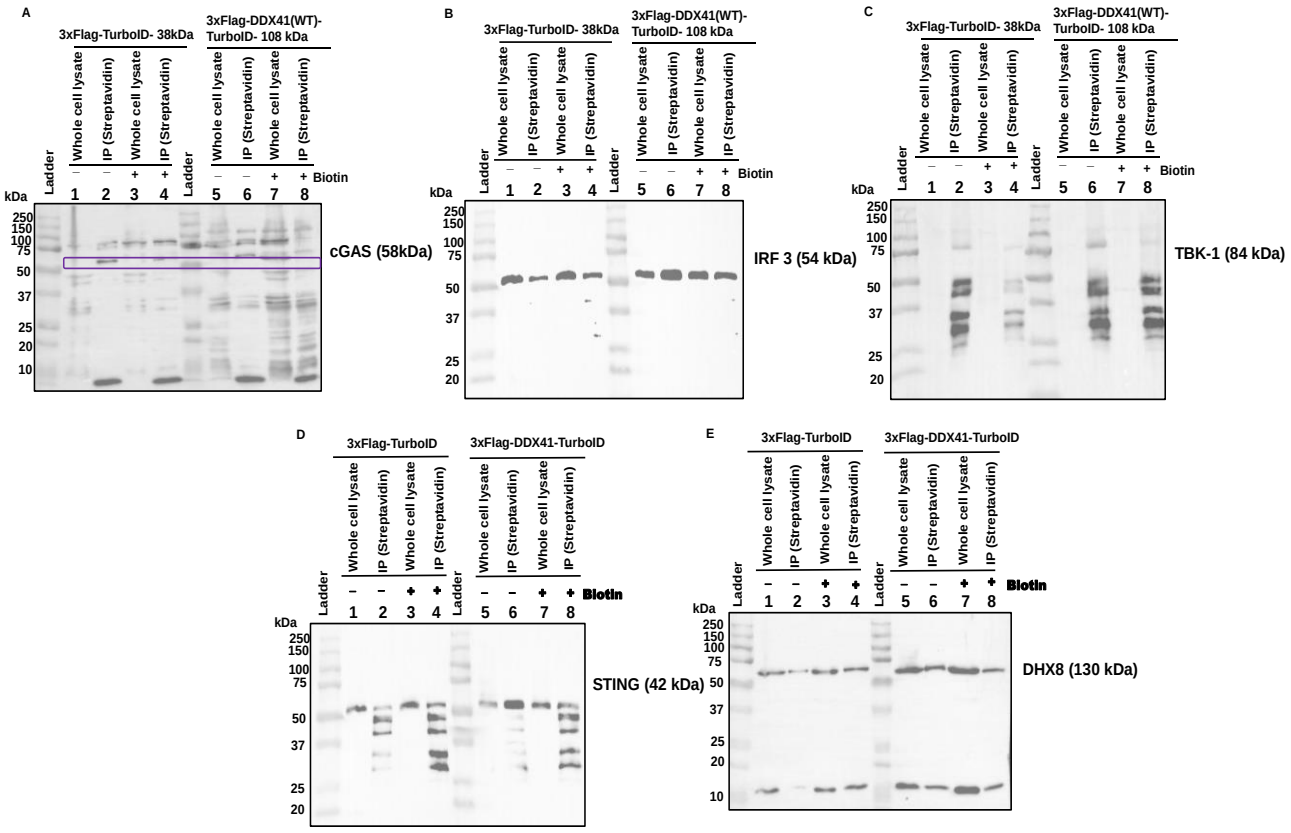


Figure 23. Probing the presence of potential interacting proteins of DDX41. The proteins from the whole cell lysate and "streptavidin-sepharose" pulled-down proteins (IP) of 3xFlag-TurboID and 3xFlag-DDX41 -TurboID were immunoblotted to identify the presence of potential biotinylated interactome of DDX41 by antibodies against few known DDX41's interacting proteins. Whole cell lysate and pulled down proteins (IP) of 3xFlag-TurboID (lanes 1-4) and 3xFlag-DDX41-TurboID (lanes 5-8) blotted against cGAS (A), IRF-3 (B), TBK-1 (C), STING (D), and DHX8 (E), respectively. The molecular weight of antibodies against different DDX41 interacting proteins is indicated with their name. "+" and "-" denote treatments with and without biotin, respectively. Protein ladder in kDa.

The presence of multiple bands in the immunoblots from using anti-cGAS, anti-STING, and anti-TBK-1 antibodies indicated no clear results (**Figure 23A-C**). When I probed against IRF-3, the blots showed strong expression of protein bands at 54 kDa corresponding to the size of IRF-3 in the WCLs and elutes of BioID fusion proteins, because IRF-3 is present in both cytoplasm and nucleus of the cell as given in the literature (**Figure 23B**) (Saikruang et al., 2022). The DHX8 antibody did not work as it yielded two non-specific bands (**Figure 23E)**. Consequently, this experiment yielded no clear interpretations of the immunoblot analysis.

5. DISCUSSION

5.1 The potential effects of fusion of DDX41 and TurboID on the helicase/ligase activity

DDX41, a member of DEAD-box helicases, has RNA as well as DNA helicase activity. It is a multi-functional protein, including pre-mRNA splicing (Polprasert et al., 2015), transcriptional regulation (Qin et al., 2021), translation (Peters et al., 2017) and ribosome biogenesis (Tungalag et al., 2023). While performing different functions inside the living cells, the enzymatic activity of DDX41 (as a helicase or splicing enzyme) might play a crucial role; meantime, DDX41 might be involved with distinct kinds of interacting proteins. The somatic mutation DDX41 (p.R525H), which is confined to highly conserved helicase motif VI (Arginine finger motif) of DDX41, is thought to contribute to nucleotide co-ordination (Qu et al., 2021). This mutant manifests as defects in pre-rRNA processing and decreased protein synthesis due to ribosomal stress (Kadono et al., 2016; Li et al., 2016). In addition, the mutation results in reduced unwinding activity of duplex DNA substrate relative to the unwinding activity of WT DDX41 (Singh et al., 2022). Considering that the helicase activity of DDX41 is critical for its functions, including recruitment of binding proteins, it will be ideal to examine the helicase activity of 3xFlag-DDX41 (WT/R525H)-TurboID, 3xFlag-DDX41 (WT) and 3xFlag-DDX41 (R525H) fusion proteins. The comparative results will provide important insights into the effect of TurboID on the helicase activity of DDX41. Importantly, it would confirm whether the R525H mutant protein has reduced unwinding activity than WT (Singh et al., 2022; Yoneyama-Hirozane et al., 2017). In fact, it has been reported that the R525H has altered interactome (Polprasert et al., 2015). The short 3xFlag tag (23 amino acids) at their N-terminus usually would not affect DDX41 and TurboID's enzymatic activities. In case, DDX41-BioID fusion protein has lower helicase activity than DDX41, which could be due to the steric hinderance by the TurboID in the folding conformation of the fusion protein, this will boil down to probability of change in the interactions of DDX41 with its interactome. Therefore, introducing a linker sequence between DDX41 and TurboID can be considered to reduce steric hindrance, giving DDX41 enough space to perform its biological function while still allowing TurboID to biotinylate interacting partners.

Protein linkers are sequences of amino acids which connect different protein domains. The linkers often adopt random coil configurations, allowing flexibility, essential for proteins to function correctly in the dynamic cellular environment (Patel et al., 2022). There are diverse types of linkers like rigid protein linkers, flexible protein linkers, short protein linkers, long protein

linkers, cleavable protein linkers, non-cleavable protein linkers. some linkers like cell membrane linkers, transmembrane linkers, membrane-binding linkers and lipid-anchored linkers facilitate communication between the extracellular environment and the cell's interior. The desired spatial arrangement between functional units can be optimized by length of the linkers used. Short linkers (5-10 amino acids) are ideal for proximity applications like BioID, minimizing spacing for efficient interactions. Medium linkers (15-20 amino acids) are suitable for balancing flexibility and separation, suitable for moderate spatial distance between protein domains. Long linkers (25-30 amino acids or more) which are optimal for large spatial separation, providing significant flexibility for dynamic conformational changes or large constructs. The most used protein-protein linkers are flexible and often consist of small, polar or non-polar amino acids like glycine (Gly) or serine (Ser) (Reddy Chichili et al., 2013). Glycine-rich linkers are indeed quite versatile in stabilizing protein-protein interactions, especially when those interactions are weak or transient. By forming covalent bonds, they help create stable complexes. These linkers are also useful for forming covalently linked dimers and connecting independent domains to create functional sites, such as ligand-binding or recognition sequences. The amino acid composition such as glycine, serine, and proline influences properties like flexibility and rigidity, contributing to conformational freedom in the domain-domain interaction and protein folding into the peculiar conformations.

For our study, flexible linkers (such as those composed of glycine-serine repeats) could provide the necessary flexibility for the fusion protein. The introduction of flexible linkers could enhance the functional integrity of DDX41 by providing the necessary space for the helicase domain to perform its enzymatic activity while still allowing for the efficient biotinylation of interacting proteins. A linker would ensure that the TurboID does not interfere with DDX41 enzymatic activity, allowing better interaction of the enzyme with its binding partners. Assays of the helicase activity followed by the BioID activity of linker-added BioID fusion proteins under both stressed and non-stressed conditions will provide a comprehensive view of how linkers facilitate the helicase activity and recruitment of specific binding partners. Different stress conditions (for instance DNA damage, heat shock, and oxidative stress) may result in capturing of physiological and stress-induced interactomes of DDX41.

In summary, examining the helicase activity of both the WT and R525H mutant of DDX41 BioID proteins and assessing the impact of TurboID fusion on protein recruitment,

optimizing fusion constructs with flexible linkers to minimize steric hindrance without compromising helicase function, will provide valuable insight into how DDX41 helicase activity contributes to its interactome. Assessing the biotinylation activity will provide thorough understanding of the role of helicase activity in DDX41 interaction network and its potential to recruit DDX41 interacting proteins under different conditions.

5.2 Role of NLS and NES in the subcellular localization of Flag-DDX41-TurboID fusion protein

The nuclear localization signal (NLS) is an amino acid sequence that acts as a "tag" for a protein, facilitating its import into the cell nucleus through nuclear transport. Typically, this signal comprises one or more short sequences of positively charged lysine or arginine exposed on the protein's surface. An NLS can be categorized as monopartite or bipartite. The key distinction between the two lies in the arrangement of basic amino acid clusters. In bipartite NLSs, these clusters are separated by a short spacer sequence, whereas monopartite NLSs lack this feature. The nuclear export signal (NES) is a concise peptide sequence within a protein, typically comprising four hydrophobic residues. This signal is responsible for directing the export of the protein from the cell nucleus to the cytoplasm, a process facilitated by the nuclear pore complex using nuclear transport. Thus, an NES functions opposite to that of an NLS.

DDX41 is ubiquitously found across a diverse range of species and functions in a diversity of cellular processes. It is predicted to have two NLSs i.e. NLS1 (bipartite), NLS2 (monopartite) and an NES motif in the DEAD domain (Ma et al., 2018); however, no experimental evidence has yet confirmed the presence of the NES motif. Human DDX41 (*Hs*DDX41) share 79–92% of an overall high amino acid identity with zebrafish DDX41 (*Dr*DDX41), suggesting the significance in the functional activities as well. Ma *et al*., reported that the N-terminal portion (1–190 amino acids) contributes to the nuclear localization of the DrDDX41 protein and exhibits translocation under poly (dA:dT) stimulation (Ma et al., 2018) and also reported another experimental observation about cytoplasmic localization of a modified construct comprised of putative NLS2 sequence from 191-411 amino acid sequence. They further provided the experimental evidence of localization of a potential NES mutant construct (316-MDLLNKKMVSLDI-329) and observed the mutant protein localizes in the nucleus and does not localize into the cytosol upon poly (dA:dT) stimulation. Another study (Kadono et al., 2016)

confirmed the nuclear localization of human full-length 70-kDa DDX41, while they observed both cytoplasmic and nuclear localization of a shorter isoform of 52-kDa DDX41 (127-622 amino acid sequence of DDX41), lacking the putative NLS, suggesting that DDX41 might function as cytosolic DNA sensor, a function of DDX41 previously delineated (Zhang et al., 2011). Recently, our group reported the translocation of full-length human DDX41 from the nucleus to cytoplasm upon exposure to poly (dA:dT), and the cytoplasmic localization of mutant DDX41 (lacking two putative NLS 6-PERKRART-13, 97-KAEARK-102) irrespective of the external DNA stimuli (Singh et al., 2022). In the sum of the experimental observations by different studies, DDX41 is a trafficking protein primarily localized in the nucleus and might translocate into the cytoplasm in response to foreign DNA. Another factor influencing the translocation activity of DDX41 is the sequence of predictive NLS2 and the NES. Therefore, more experimental evidence of the exact molecular mechanism will help uncover the mystery of the cytoplasmic localization and translocation activity of DDX41.

The early forms of the biotin ligase enzymes, such as Bir A and BioID had long periods of biotin labelling time i.e. 18 h, therefore, Branon *et al* engineered the TurboID (Branon et al., 2018) from BioID to overcome the slow labelling kinetics. Resultantly, TurboID emerged as a biotin ligase with the highest activity and shorter labelling time that enabled the probing of dynamic biological processes with a higher temporal resolution. This improvement made TurboID an ideal biotin ligase for *in vivo* application. As TurboID can easily be directed into the different cellular organelles, different researchers have tagged it with NLS or NES sequence or simply with their protein of interest and generated target constructs of their choice according to their experimental goals. For instance, Mair *et al* fused TurboID with their bait nuclear protein and profiled its nuclear proteome (Mair et al., 2019). Another group used TurboID-NES to identify only the cytoplasmic proteome in brain cells (Sunna et al., 2023). Hence, TurboID can become an effective tool for probing the PPI and investigating various downstream applications of PPI like exploring the metabolic pathways of novel proteins, uncovering the physiological functions associated with the proteins, unfolding the diseases and their cure and so on.

We found the unexpected localization of the 3xFlag-DDX41-TurboID fusion proteins in the cytoplasm (**Figure 12**) as we accidentally used the TurboID plasmid with an NES sequence tagged at its C-terminus. The mislocalization might be due to the dominant effect of the cytoplasmic directing function of an NES at the C-terminus of TurboID. After the removal of the

NES sequence, the DDX41 BioID fusion proteins were localized in the nucleus, suggesting that the NES sequence had affected the localization of the DDX41 BioID fusion protein. However, it remains unknown whether such a dominant effect or modification can be applied to other proteins of interest that are fused with TurboID.

5.3 Transient and inducible stable expression of BioID-DDX41 in cells

To identify the protein interactome of a bait protein, transient or stable cell lines can be generated depending on the protein of interest or the choice of biotin ligase to be fused. While transient transfection yields asymmetrical/dissimilar protein expression in the cells, stable transfection might be advantageous regarding similar levels of protein expression.

I have tried to generate stable cell lines of our 3xFlag-DDX41-TurboID fusion protein using the HeLa DDX41 KO cell line. However, I could not succeed in generating stable cells because of low cell viability, which might to due to the toxicity of our target BioID fusion protein. It has been observed that mammalian cells constitutively expressing TurboID result in cell toxicity and interfere with the protein function (Branon et al., 2018). Another group, May *et al* also confirmed the cellular toxicity and instability of proteins involving constitutive expression of TurboID in live cells of *C. elegans* and *D. melanogaster* (May et al., 2020). In contrast, Sunna *et al* were successful in profiling the cytosolic proteome using murine neuroblastoma and microglial cells constitutively expressing TurboID-NES, without observing any cellular toxicity (Sunna et al., 2023). Therefore, with contrasting results about the cellular toxicity caused by TurboID stable cell lines, it will be judicious to carefully consider testing TurboID transient expression or stable expression with the protein of interest, cell lines and experimental model organism.

I have used HeLa DDX41 KO and HEK293T cell lines with transient transfection for the experiments. However, due to the lower transfection rate and lower protein expression of HeLa DDX41 KO relative to HEK293T, I could use this cell line only for the immunofluorescence experiment to determine the localization of BioID fusion proteins. Therefore, in our work, I used the HEK293T cell line which has a higher rate of transient transfection and a higher level of protein expression to verify the expression, biotinylation activity of TurboID as well as to isolate biotinylated proteins.

In fact, transient transfection has several advantages over stable expression in terms of BioID. Since TurboID has a higher labelling potential of biotinylation and a shorter labelling time

relative to other biotin ligase enzymes. Therefore, adopting transient transfection over stable expression can likely mitigate the scope of continuous biotinylation of proximate proteins by TurboID under conventional culture conditions. Also, transient transfection might be beneficial as it decreases the probability of capturing false positives due to overexpression in stable cells.

On the other hand, stable cell lines can be favourable as they can capture comparatively more direct, indirect and other transient interactions, i.e. proteins binding directly or transiently to a bait protein along with proteins present in the labelling radius with no direct or transient interaction to the protein of interest (Roux et al., 2012). This is due to the covalent binding of biotin to the lysine residues of proximal proteins. Moreover, stable cells become the first line of choice to identify the *in vivo* proteome (Sunna et al., 2023), where the cells are not triggered by any external or internal stimuli. They are advantageous over transient transfection as they can depict more detailed proteomic profile of various dynamic physiological states both *in vitro* and *in vivo.*

In summary, for optimum BioID experiment, either the transient cell lines or the inducible stable cell lines can be used based on the consideration of issues such as cellular toxicity, overexpression, biotinylation activity and other factors.

5.4 Potential reasons why DDX41-interacting proteins could not be detected

I have tried to validate the presence of some previously known interacting proteins of DDX41 such as STING, TBK-1, IRF-3 and cGAS **(Figure 22)**. Nevertheless, the interpretation of the results is indeterminate because of the technical limitations like immunoblotting was only performed once, thus, the reliability of the results is questionable because of the lack of experimental replicates. The effectiveness of antibodies remains questionable, particularly due to the detection of multiple bands in the immunoblots with cGAS, TBK-1, and STING antibodies.

The identity of potential interacting proteins of DDX41 is expected to vary because of distinct variable factors like sub-cellular localization of DDX41 (cytoplasm or nucleus) or metabolic functions of DDX41 (pre-mRNA splicing or innate immunity). DDX41 has been found in the nucleus as a splicing protein (Cheah et al., 2017) and also functions as a cytosolic DNA sensor (Zhang et al., 2011). The nuclear localization of DDX41 has been observed in HEK293, murine lung fibroblast, and THP1 cells (Kadono et al., 2016; Lewinsohn et al., 2016). On

stimulation of extracellular DNA, the sub-cellular localization of human DDX41 (**Figure 7C**) (Singh et al., 2022) as well as Zebrafish DDX41 (Ma et al., 2018) have been reported to shift from the nucleus to the cytoplasm The potential interactome of DDX41 is likely to vary depending on its prominent localization and cellular function. Therefore, to identify true DDX41 interactors, it is crucial to analyze nuclear and cytosolic interactors separately.

It is noteworthy that we may risk losing the identification of some potential interacting proteins because of several reasons such as the highly dynamic structure of the spliceosome, dynamic shift in the sub-cellular localization of DDX41, limited labelling radius of TurboID, physiological and functional status of DDX41 BioID fusion proteins at time of biotin labelling or loss of interacting proteins during the immunoprecipitation. Therefore, BioID might provide information about unknown protein partners, but it can also lead to potential loss in identifying interactors. Another significant drawback of proximity labelling methods is that the biotin ligase enzymes fail to discriminate between two interactome proteins bound directly or merely present adjacent to each other (Chen & Perrimon, 2017).

5.5 Optimization for a successful BioID approach

BioID has proved to be a revolutionary approach to study protein-protein interactions (PPI) (Roux et al., 2018), RNA-protein interactions (Ramanathan et al., 2018), and cell-cell interactions (Ge et al., 2019). The first inception of the idea of BioID happened in Brian Burke's lab in 2012 (Roux et al., 2012) to detect the weak and transient PPI to overcome the limitations of conventional PPI methods. The success of a BioID experiment relies on thorough optimization and critical considerations at every stage, from the initial setup to the final step of identifying the interacting proteins. In our selection of a biotinylation agent, we opted for TurboID because of its non-toxic nature towards living cells, wide 35 nm labelling radius, short 10-minute labelling time and fast biotin labelling activity. To ensure the efficient stability and activity of the fusion protein, we linked TurboID, which contains a nuclear export sequence at its C-terminus, to the C-terminus of DDX41. This was done to prevent any unnecessary structural, steric, or functional conflicts between the NES in the C-terminal of TurboID and the NLS in the N-terminal of DDX41. In our BioID experiment involving DDX41, we optimized various factors including cell lines, transfection reagents, biotin treatment (30 minutes and 50 µM), antibody dilutions, cell quantities, and immunoprecipitation. For instance, we preferred the HEK293T cell line over the

HeLa DDX41 KO cell line because of their higher transfection efficiency and a higher level of BioID fusion protein expression; we optimized the biotin labelling time as 30 min in contrast to the standard 10 min of biotin labelling (Branon et al., 2018). Interestingly, a recent study reported their optimized biotin incubation time for TurboID of 18 h in the zebrafish model (Xiong et al., 2021). Thus, even though TurboID, originally reported as one of highly robust biotin ligase, may vary in its proximity labelling in different model systems or for different proteins. The optimization of immunoprecipitation of BioID fusion proteins with the anti-Flag antibody conjugated to protein A/G agarose beads. The optimization of the pull-down of putative biotinylated proteins was carried out using two different methods, first by using streptavidin-sepharose beads and second by using streptavidin-HRP with protein A/G agarose beads. The first method could not succeed in pulling down the biotinylated 3xFlag-DDX41-TurboID proteins and other potential interacting proteins of DDX41 from the WCL. This may be attributed to the reason that the streptavidin-sepharose beads were saturated with either non-specific proteins or free biotin. On the other hand, the second approach was successful in the enrichment of biotinylated 3xFlag-DDX41-TurboID proteins and a few potential interacting proteins of DDX41 **(Figure 22)**. The enrichment was indicative of a high affinity of streptavidin-HRP towards capturing the biotinylated proteins, and an unexpected display of the linking of the added streptavidin to the agarose beads.

Although I optimized the basic parameters for DDX41-BioID, it will be prudent to consider some other vital factors to obtain a representative set of the true interactome of DDX41 under relevant physiological conditions. During the process of biotin labelling, the possibility of biotinylation of non-interacting proteins that are in close proximity to the interacting proteins of POI, therefore the use of supplementary controls plays a vital role in ruling out the non-interacting proteins (Mair et al., 2019). For our BioID fusion proteins, the use of controls such as a nuclear-localized DDX41 control as opposed to the DDX41, an empty TurboID enzyme for biotin ligase enzyme, untransfected endogenous proteins for transfected DNA, and non-biotinylated proteins for the biotin. It is noteworthy that the use of TurboID with NLS can be challenging as it can pose cellular toxicity because of the potential of TurboID to biotinylate gene regulatory proteins, thus, altering the structure and function of these proteins. Therefore, it would be crucial to consider the possibility of toxicity with a nuclear direction of TurboID. Hence, comparing proteomes using TurboID-NES to TurboID without NES to observe the cellular toxicity might be considered

(Sunna et al., 2023). Besides, exposing the cells to stress conditions like external DNA stimuli [poly (dA:dT)] might lead to the unravelling of different protein interactors of DDX41, given the changes in sub-cellular localization and biological function of DDX41 under stress. Some other technical optimizations may include the desalting of biotinylated pull-down proteins (Lin et al., 2024) to remove free biotin from the purified proteins; and optimization of the abundance and purity of the biotinylated proteins.

Implementing these additional approaches and conducting biological replicates exclusively for nucleus and cytoplasmic BioID in the DDX41 MS study might lead to the identification of true interacting proteins of DDX41.

6. CONCLUSIONS AND FUTURE WORK

6.1 Conclusions

In conclusion, we have generated DDX41 BioID fusion construct, i.e. 3xFlag-DDX41-TurboID into the pcDNA 3.0 vector and a control BioID fusion construct, i.e. 3xFlag-TurboID. Immunoblots against Flag, DDX41 and TurboID validated the translational expression of BioID fusion proteins. I also found the optimum biotinylation time of 30 minutes and 50 µM of biotin concentration for biotinylation activity by TurboID. In contrary to the expected nuclear localization of native DDX41, our data from the immunofluorescence experiments revealed the cytoplasmic localization of DDX41 BioID fusion proteins. Therefore, we removed the NES from TurboID by site-directed mutagenesis and constructed NES-deleted BioID fusion constructs. The immunoblotting and immunofluorescence experiments confirmed the translational expression and nuclear localization of the NES-deleted DDX41 BioID fusion proteins.

The IP experiments using Flag antibody resulted in the detection of BioID fusion proteins and the faint presence of potential DDX41 interacting protein at nearly 80 kDa. The same observation of potential DDX41 interacting protein was also made earlier in the WB analysis of biotin-treated WCL of NES-deleted DDX41 BioID fusion protein. When NES-deleted DDX41 BioID fusion proteins were pulled-down by streptavidin-HRP, and the immunoblot showed not only the biotinylated BioID fusion proteins but also the repeated presence of the same DDX41 potential interacting protein in the elute. Therefore, we have successfully observed and confirmed the strong presence of at least one potential DDX41 interacting protein in three different experiments, the identity of which will be verified upon mass spectrometric analysis.

The Western blot experiments to seek the validation of a few known interacting proteins of DDX41 yielded no clear results, as these WBs were performed once. Another factor is the antibodies' non-specific binding with the immunoprecipitated biotinylated proteins. Hence, these results will be interpreted more clearly as validation of the MS results of DDX41 interactome.

6.2 Future work

Although our current system needs further fine-tuning, I believe that I have laid the foundation of the BioID system to identify DDX41-interacting proteins. In the future, we will:

1. Identify potential interacting proteins (both cytoplasmic and nuclear) of DDX41 through mass spectrometric analysis with and without stress conditions.

2. Validate the top five interacting proteins by Co-IP.
3. Define the interacting domain, motif and amino acid by truncation and site-directed mutagenesis.
4. Uncover the biological functions of DDX41-interacting proteins and their associated pathways with and without stress conditions.
5. Discover druggable target proteins/pathways to ameliorate DDX41-related leukemia, haematological and other disorders.

7. REFERENCES

Abdel-Monem, M., Dürwald, H., & Hoffmann-Berling, H. (1976). Enzymic unwinding of DNA. 2. Chain separation by an ATP-dependent DNA unwinding enzyme. *Eur J Biochem*, *65*(2), 441-449. https://doi.org/10.1111/j.1432-1033.1976.tb10359.x

Abdel-Monem, M., & Hoffmann-Berling, H. (1976). Enzymic unwinding of DNA. 1. Purification and characterization of a DNA-dependent ATPase from Escherichia coli. *Eur J Biochem*, *65*(2), 431-440. https://doi.org/10.1111/j.1432-1033.1976.tb10358.x

Agafonov, D. E., Deckert, J., Wolf, E., Odenwälder, P., Bessonov, S., Will, C. L., Urlaub, H., & Lührmann, R. (2011). Semiquantitative proteomic analysis of the human spliceosome via a novel two-dimensional gel electrophoresis method. *Mol Cell Biol*, *31*(13), 2667-2682. https://doi.org/10.1128/mcb.05266-11

Ali, M. A. M. (2021). The DEAD-box protein family of RNA helicases: sentinels for a myriad of cellular functions with emerging roles in tumorigenesis. *Int J Clin Oncol*, *26*(5), 795-825. https://doi.org/10.1007/s10147-021-01892-1

Anderson, J. S., & Parker, R. P. (1998). The 3' to 5' degradation of yeast mRNAs is a general mechanism for mRNA turnover that requires the SKI2 DEVH box protein and 3' to 5' exonucleases of the exosome complex. *Embo j*, *17*(5), 1497-1506. https://doi.org/10.1093/emboj/17.5.1497

Anderson, J. S. J., & Parker, R. (1998). The 3′ to 5′ degradation of yeast mRNAs is a general mechanism for mRNA turnover that requires the SKI2 DEVH box protein and 3′ to 5′ exonucleases of the exosome complex. *The EMBO journal*.

Andreou, A. Z. (2021). DDX41: a multifunctional DEAD-box protein involved in pre-mRNA splicing and innate immunity. *Biological Chemistry*, *402*(5), 645-651. https://doi.org/doi:10.1515/hsz-2020-0367

Auboeuf, D., Hönig, A., Berget, S. M., & O'Malley, B. W. (2002). Coordinate regulation of transcription and splicing by steroid receptor coregulators. *Science*, *298*(5592), 416-419. https://doi.org/10.1126/science.1073734

Badar, T., & Chlon, T. (2022). Germline and somatic defects in DDX41 and its impact on myeloid neoplasms. *Current hematologic malignancy reports*, *17*(5), 113-120.

Bannon, S. A., Routbort, M. J., Montalban-Bravo, G., Mehta, R. S., Jelloul, F. Z., Takahashi, K., Daver, N., Oran, B., Pemmaraju, N., Borthakur, G., Naqvi, K., Issa, G., Sasaki, K., Alvarado, Y., Kadia, T. M., Konopleva, M., Shamanna, R. K., Khoury, J. D., Ravandi, F., . . . DiNardo, C. D. (2020). Next-Generation Sequencing of DDX41 in Myeloid Neoplasms Leads to Increased Detection of Germline Alterations. *Front Oncol*, *10*, 582213. https://doi.org/10.3389/fonc.2020.582213

Barker, D. F., & Campbell, A. M. (1981a). The birA gene of Escherichia coli encodes a biotin holoenzyme synthetase. *J Mol Biol*, *146*(4), 451-467. https://doi.org/10.1016/0022-2836(81)90042-5

Barker, D. F., & Campbell, A. M. (1981b). Genetic and biochemical characterization of the birA gene and its product: evidence for a direct role of biotin holoenzyme synthetase in repression of the biotin operon in Escherichia coli. *J Mol Biol*, *146*(4), 469-492. https://doi.org/10.1016/0022-2836(81)90043-7

Bernstein, D. A., Zittel, M. C., & Keck, J. L. (2003). High-resolution structure of the E.coli RecQ helicase catalytic core. *Embo j*, *22*(19), 4910-4921. https://doi.org/10.1093/emboj/cdg500

Bernstein, K. A., Granneman, S., Lee, A. V., Manickam, S., & Baserga, S. J. (2006). Comprehensive mutational analysis of yeast DEXD/H box RNA helicases involved in large ribosomal subunit biogenesis. *Molecular and cellular biology*, *26*(4), 1195-1208.

Boule, J.-B., & Zakian, V. A. (2007). The yeast Pif1p DNA helicase preferentially unwinds RNA–DNA substrates. *Nucleic acids research*, *35*(17), 5809-5818.

Branon, T. C., Bosch, J. A., Sanchez, A. D., Udeshi, N. D., Svinkina, T., Carr, S. A., Feldman, J. L.,

Perrimon, N., & Ting, A. Y. (2018). Efficient proximity labeling in living cells and organisms with TurboID. *Nat Biotechnol*, *36*(9), 880-887. https://doi.org/10.1038/nbt.4201

Cardoso, F., van't Veer, L. J., Bogaerts, J., Slaets, L., Viale, G., Delaloge, S., Pierga, J. Y., Brain, E., Causeret, S., DeLorenzi, M., Glas, A. M., Golfinopoulos, V., Goulioti, T., Knox, S., Matos, E., Meulemans, B., Neijenhuis, P. A., Nitz, U., Passalacqua, R., . . . Piccart, M. (2016). 70-Gene Signature as an Aid to Treatment Decisions in Early-Stage Breast Cancer. *N Engl J Med*, *375*(8), 717-729. https://doi.org/10.1056/NEJMoa1602253

Carrera, P., Johnstone, O., Nakamura, A., Casanova, J., Jäckle, H., & Lasko, P. (2000). VASA mediates translation through interaction with a Drosophila yIF2 homolog. *Molecular cell*, *5*(1), 181-187.

Cheah, J. J. C., Hahn, C. N., Hiwase, D. K., Scott, H. S., & Brown, A. L. (2017). Myeloid neoplasms with germline DDX41 mutation. *Int J Hematol*, *106*(2), 163-174. https://doi.org/10.1007/s12185-017-2260-y

Chen, C. L., & Perrimon, N. (2017). Proximity-dependent labeling methods for proteomic profiling in living cells. *Wiley Interdiscip Rev Dev Biol*, *6*(4). https://doi.org/10.1002/wdev.272

Cheng, Y., Liu, Y., Wang, Y., Niu, Q., Gao, Q., Fu, Q., Ma, J., Wang, H., Yan, Y., & Ding, C. (2017). Chicken DNA virus sensor DDX41 activates IFN-β signaling pathway dependent on STING. *Developmental & Comparative Immunology*, *76*, 334-342.

Chlon, T. M., Stepanchick, E., Hershberger, C. E., Daniels, N. J., Hueneman, K. M., Davis, A. K., Choi, K., Zheng, Y., Gurnari, C., & Haferlach, T. (2021). Germline DDX41 mutations cause ineffective hematopoiesis and myelodysplasia. *Cell Stem Cell*, *28*(11), 1966-1981. e1966.

Choi-Rhee, E., Schulman, H., & Cronan, J. E. (2004). Promiscuous protein biotinylation by Escherichia coli biotin protein ligase. *Protein Sci*, *13*(11), 3043-3050. https://doi.org/10.1110/ps.04911804

Chuang, R. Y., Weaver, P. L., Liu, Z., & Chang, T. H. (1997). Requirement of the DEAD-Box protein ded1p for messenger RNA translation. *Science*, *275*(5305), 1468-1471. https://doi.org/10.1126/science.275.5305.1468

Churpek, J. E., & Smith-Simmer, K. (1993). DDX41-Associated Familial Myelodysplastic Syndrome and Acute Myeloid Leukemia. In M. P. Adam, H. H. Ardinger, R. A. Pagon, S. E. Wallace, L. J. H. Bean, K. W. Gripp, G. M. Mirzaa, & A. Amemiya (Eds.), *GeneReviews(®)*. University of Washington, Seattle

Copyright © 1993-2022, University of Washington, Seattle. GeneReviews is a registered trademark of the University of Washington, Seattle. All rights reserved.

Clark, E. L., Fuller-Pace, F. V., Elliott, D. J., & Robson, C. N. (2008). Coupling transcription to RNA processing via the p68 DEAD box RNA helicase androgen receptor co-activator in prostate cancer. *Biochemical Society Transactions*, *36*(3), 546-547.

Cobb, J. A., & Bjergbaek, L. (2006). RecQ helicases: lessons from model organisms. *Nucleic acids research*, *34*(15), 4106-4114.

Cronan, J. E. (2005). Targeted and proximity-dependent promiscuous protein biotinylation by a mutant Escherichia coli biotin protein ligase. *J Nutr Biochem*, *16*(7), 416-418. https://doi.org/10.1016/j.jnutbio.2005.03.017

Cui, S., Eisenächer, K., Kirchhofer, A., Brzózka, K., Lammens, A., Lammens, K., Fujita, T., Conzelmann, K. K., Krug, A., & Hopfner, K. P. (2008). The C-terminal regulatory domain is the RNA 5'-triphosphate sensor of RIG-I. *Mol Cell*, *29*(2), 169-179. https://doi.org/10.1016/j.molcel.2007.10.032

de Boer, E., Rodriguez, P., Bonte, E., Krijgsveld, J., Katsantoni, E., Heck, A., Grosveld, F., & Strouboulis, J. (2003). Efficient biotinylation and single-step purification of tagged transcription factors in mammalian cells and transgenic mice. *Proc Natl Acad Sci U S A*, *100*(13), 7480-7485. https://doi.org/10.1073/pnas.1332608100

Delagoutte, E., & Von Hippel, P. H. (2002). Helicase mechanisms and the coupling of helicases within macromolecular machines part I: structures and properties of isolated helicases. *Quarterly reviews of biophysics*, *35*(4), 431-478.

Delagoutte, E., & Von Hippel, P. H. (2003). Helicase mechanisms and the coupling of helicases within

macromolecular machines Part ii: Integration of helicases into cellular processes. *Quarterly reviews of biophysics*, *36*(1), 1-69.
Dillingham, M. S. (2011). Superfamily I helicases as modular components of DNA-processing machines. *Biochem Soc Trans*, *39*(2), 413-423. https://doi.org/10.1042/bst0390413
Ding, L., Ley, T. J., Larson, D. E., Miller, C. A., Koboldt, D. C., Welch, J. S., Ritchey, J. K., Young, M. A., Lamprecht, T., McLellan, M. D., McMichael, J. F., Wallis, J. W., Lu, C., Shen, D., Harris, C. C., Dooling, D. J., Fulton, R. S., Fulton, L. L., Chen, K., . . . DiPersio, J. F. (2012). Clonal evolution in relapsed acute myeloid leukaemia revealed by whole-genome sequencing. *Nature*, *481*(7382), 506-510. https://doi.org/10.1038/nature10738
Dybkov, O., Preußner, M., El Ayoubi, L., Feng, V. Y., Harnisch, C., Merz, K., Leupold, P., Yudichev, P., Agafonov, D. E., Will, C. L., Girard, C., Dienemann, C., Urlaub, H., Kastner, B., Heyd, F., & Lührmann, R. (2023). Regulation of 3' splice site selection after step 1 of splicing by spliceosomal C* proteins. *Sci Adv*, *9*(9), eadf1785. https://doi.org/10.1126/sciadv.adf1785
Eisenberg, M. A., Prakash, O., & Hsiung, S. C. (1982). Purification and properties of the biotin repressor. A bifunctional protein. *J Biol Chem*, *257*(24), 15167-15173.
Fuller-Pace, F. V. (2013). DEAD box RNA helicase functions in cancer. *RNA biology*, *10*(1), 121-132.
Ge, Y., Chen, L., Liu, S., Zhao, J., Zhang, H., & Chen, P. R. (2019). Enzyme-Mediated Intercellular Proximity Labeling for Detecting Cell-Cell Interactions. *J Am Chem Soc*, *141*(5), 1833-1837. https://doi.org/10.1021/jacs.8b10286
Geider, K. (1978). The single-stranded DNA phages.(eds. Denhardt, Dressler & Ray) 378–387. *Cold Spring Habor Laboratory Press, New York*.
Germain, D. R., Graham, K., Glubrecht, D. D., Hugh, J. C., Mackey, J. R., & Godbout, R. (2011). DEAD box 1: a novel and independent prognostic marker for early recurrence in breast cancer. *Breast cancer research and treatment*, *127*, 53-63.
Gorbalenya, A. E., & Koonin, E. V. (1993). Helicases: amino acid sequence comparisons and structure-function relationships. *Current opinion in structural biology*, *3*(3), 419-429.
Granneman, S., Bernstein, K. A., Bleichert, F., & Baserga, S. J. (2006). Comprehensive mutational analysis of yeast DEXD/H box RNA helicases required for small ribosomal subunit synthesis. *Molecular and cellular biology*.
Hanada, K., & Hickson, I. (2007). Molecular genetics of RecQ helicase disorders. *Cellular and molecular life sciences*, *64*, 2306-2322.
He, Y., Andersen, G. R., & Nielsen, K. H. (2010). Structural basis for the function of DEAH helicases. *EMBO Rep*, *11*(3), 180-186. https://doi.org/10.1038/embor.2010.11
Hilgarth, R. S., & Lanigan, T. M. (2020). Optimization of overlap extension PCR for efficient transgene construction. *MethodsX*, *7*, 100759. https://doi.org/10.1016/j.mex.2019.12.001
Hönig, A., Auboeuf, D., Parker, M. M., O'Malley, B. W., & Berget, S. M. (2002). Regulation of alternative splicing by the ATP-dependent DEAD-box RNA helicase p72. *Mol Cell Biol*, *22*(16), 5698-5707. https://doi.org/10.1128/mcb.22.16.5698-5707.2002
Hotta, Y., & Stern, H. (1978). DNA unwinding protein from meiotic cells of Lilium. *Biochemistry*, *17*(10), 1872-1880. https://doi.org/10.1021/bi00603a011
Hu, J., Xu, X., Wang, S., & Ge, G. (2020). Ctenopharyngodon idellus DDX41 initiates IFN I and ISG15 expression in response to GCRV infection. *Fish & Shellfish Immunology*, *106*, 149-160.
Irion, U., & Leptin, M. (1999). Developmental and cell biological functions of the Drosophila DEAD-box protein abstrakt. *Current biology*, *9*(23), 1373-1381.
Jankowsky, E. (2011). RNA helicases at work: binding and rearranging. *Trends Biochem Sci*, *36*(1), 19-29. https://doi.org/10.1016/j.tibs.2010.07.008
Jiang, Y., Zhu, Y., Liu, Z. J., & Ouyang, S. (2017). The emerging roles of the DDX41 protein in immunity and diseases. *Protein Cell*, *8*(2), 83-89. https://doi.org/10.1007/s13238-016-0303-4
Jurica, M. S., Licklider, L. J., Gygi, S. R., Grigorieff, N., & Moore, M. J. (2002). Purification and characterization of native spliceosomes suitable for three-dimensional structural analysis. *Rna*, *8*(4), 426-439. https://doi.org/10.1017/s1355838202021088

Kadono, M., Kanai, A., Nagamachi, A., Shinriki, S., Kawata, J., Iwato, K., Kyo, T., Oshima, K., Yokoyama, A., Kawamura, T., Nagase, R., Inoue, D., Kitamura, T., Inaba, T., Ichinohe, T., & Matsui, H. (2016). Biological implications of somatic DDX41 p.R525H mutation in acute myeloid leukemia. *Exp Hematol*, *44*(8), 745-754.e744. https://doi.org/10.1016/j.exphem.2016.04.017

Khoury, J. D., Solary, E., Abla, O., Akkari, Y., Alaggio, R., Apperley, J. F., Bejar, R., Berti, E., Busque, L., Chan, J. K. C., Chen, W., Chen, X., Chng, W. J., Choi, J. K., Colmenero, I., Coupland, S. E., Cross, N. C. P., De Jong, D., Elghetany, M. T., . . . Hochhaus, A. (2022). The 5th edition of the World Health Organization Classification of Haematolymphoid Tumours: Myeloid and Histiocytic/Dendritic Neoplasms. *Leukemia*, *36*(7), 1703-1719. https://doi.org/10.1038/s41375-022-01613-1

Kim, D. I., Jensen, S. C., Noble, K. A., Kc, B., Roux, K. H., Motamedchaboki, K., & Roux, K. J. (2016). An improved smaller biotin ligase for BioID proximity labeling. *Mol Biol Cell*, *27*(8), 1188-1196. https://doi.org/10.1091/mbc.E15-12-0844

Kim, K., Ong, F., & Sasaki, K. (2023). Current Understanding of DDX41 Mutations in Myeloid Neoplasms. *Cancers (Basel)*, *15*(2). https://doi.org/10.3390/cancers15020344

Kim, S., Govindan, J. A., Tu, Z. J., & Greenstein, D. (2012). SACY-1 DEAD-Box helicase links the somatic control of oocyte meiotic maturation to the sperm-to-oocyte switch and gamete maintenance in Caenorhabditis elegans. *Genetics*, *192*(3), 905-928.

Kwon, K., & Beckett, D. (2000). Function of a conserved sequence motif in biotin holoenzyme synthetases. *Protein Sci*, *9*(8), 1530-1539. https://doi.org/10.1110/ps.9.8.1530

Kwon, K., Streaker, E. D., Ruparelia, S., & Beckett, D. (2000). Multiple disordered loops function in corepressor-induced dimerization of the biotin repressor. *J Mol Biol*, *304*(5), 821-833. https://doi.org/10.1006/jmbi.2000.4249

Lasko, P. (2013). The DEAD-box helicase Vasa: evidence for a multiplicity of functions in RNA processes and developmental biology. *Biochimica et Biophysica Acta (BBA)-Gene Regulatory Mechanisms*, *1829*(8), 810-816.

Lewinsohn, M., Brown, A. L., Weinel, L. M., Phung, C., Rafidi, G., Lee, M. K., Schreiber, A. W., Feng, J., Babic, M., Chong, C. E., Lee, Y., Yong, A., Suthers, G. K., Poplawski, N., Altree, M., Phillips, K., Jaensch, L., Fine, M., D'Andrea, R. J., . . . Scott, H. S. (2016). Novel germ line DDX41 mutations define families with a lower age of MDS/AML onset and lymphoid malignancies. *Blood*, *127*(8), 1017-1023. https://doi.org/10.1182/blood-2015-10-676098

Li, R., Sobreira, N., Witmer, P. D., Pratz, K. W., & Braunstein, E. M. (2016). Two novel germline DDX41 mutations in a family with inherited myelodysplasia/acute myeloid leukemia. *Haematologica*, *101*(6), e228-231. https://doi.org/10.3324/haematol.2015.139790

Li, Y., Li, H., Su, N., Liu, D., Luo, R., & Jin, H. (2018). Molecular cloning and functional characterization of duck DDX41. *Developmental & Comparative Immunology*, *88*, 183-189.

Lin, D., Kuang, J., & Gao, C. (2024). Identification of Neighboring Proteins of Endosomal Regulators by Using TurboID-Based Proximity Labeling. *Methods Mol Biol*, *2841*, 121-130. https://doi.org/10.1007/978-1-0716-4059-3_11

Linder, P., & Jankowsky, E. (2011). From unwinding to clamping - the DEAD box RNA helicase family. *Nat Rev Mol Cell Biol*, *12*(8), 505-516. https://doi.org/10.1038/nrm3154

Linder, P., Lasko, P. F., Ashburner, M., Leroy, P., Nielsen, P. J., Nishi, K., Schnier, J., & Slonimski, P. P. (1989). Birth of the D-E-A-D box. *Nature*, *337*(6203), 121-122. https://doi.org/10.1038/337121a0

Liu, Z. R. (2002). p68 RNA helicase is an essential human splicing factor that acts at the U1 snRNA-5' splice site duplex. *Mol Cell Biol*, *22*(15), 5443-5450. https://doi.org/10.1128/mcb.22.15.5443-5450.2002

Lohman, T. M., Tomko, E. J., & Wu, C. G. (2008). Non-hexameric DNA helicases and translocases: mechanisms and regulation. *Nature Reviews Molecular Cell Biology*, *9*(5), 391-401. https://doi.org/10.1038/nrm2394

Ma, J., & Ross, S. R. (2024). Multifunctional role of DEAD-box helicase 41 in innate immunity, hematopoiesis and disease. *Front Immunol*, *15*, 1451705. https://doi.org/10.3389/fimmu.2024.1451705

Ma, J. X., Li, J. Y., Fan, D. D., Feng, W., Lin, A. F., Xiang, L. X., & Shao, J. Z. (2018). Identification of DEAD-Box RNA Helicase DDX41 as a Trafficking Protein That Involves in Multiple Innate Immune Signaling Pathways in a Zebrafish Model. *Front Immunol*, *9*, 1327. https://doi.org/10.3389/fimmu.2018.01327

Mackay, V., & Linn, S. (1976). Selective inhibition of the dnase activity of the recBC enzyme by the DNA binding protein from Escherichia coli. *Journal of Biological Chemistry*, *251*(12), 3716-3719.

Mair, A., Xu, S. L., Branon, T. C., Ting, A. Y., & Bergmann, D. C. (2019). Proximity labeling of protein complexes and cell-type-specific organellar proteomes in Arabidopsis enabled by TurboID. *Elife*, *8*. https://doi.org/10.7554/eLife.47864

Makishima, H., Saiki, R., Nannya, Y., Korotev, S., Gurnari, C., Takeda, J., Momozawa, Y., Best, S., Krishnamurthy, P., & Yoshizato, T. (2023). Germ line DDX41 mutations define a unique subtype of myeloid neoplasms. *Blood*, *141*(5), 534-549.

May, D. G., Scott, K. L., Campos, A. R., & Roux, K. J. (2020). Comparative Application of BioID and TurboID for Protein-Proximity Biotinylation. *Cells*, *9*(5). https://doi.org/10.3390/cells9051070

Missel, A., Souza, A. E., Nörskau, G., & Göringer, H. U. (1997). Disruption of a gene encoding a novel mitochondrial DEAD-box protein in Trypanosoma brucei affects edited mRNAs. *Mol Cell Biol*, *17*(9), 4895-4903. https://doi.org/10.1128/mcb.17.9.4895

Mosler, T., Conte, F., Longo, G. M., Mikicic, I., Kreim, N., Möckel, M. M., Petrosino, G., Flach, J., Barau, J., & Luke, B. (2021). R-loop proximity proteomics identifies a role of DDX41 in transcription-associated genomic instability. *Nature communications*, *12*(1), 7314.

Osterhoudt, K., Bagno, O., Katzman, S., & Zahler, A. M. (2024). Spliceosomal helicases DDX41/SACY-1 and PRP22/MOG-5 both contribute to proofreading against proximal 3' splice site usage. *Rna*, *30*(4), 404-417. https://doi.org/10.1261/rna.079888.123

Ouyang, K. J., Woo, L. L., & Ellis, N. A. (2008). Homologous recombination and maintenance of genome integrity: cancer and aging through the prism of human RecQ helicases. *Mechanisms of Ageing and Development*, *129*(7-8), 425-440.

Pang, P. S., Jankowsky, E., Planet, P. J., & Pyle, A. M. (2002). The hepatitis C viral NS3 protein is a processive DNA helicase with cofactor enhanced RNA unwinding. *The EMBO journal*.

Parvatiyar, K., Zhang, Z., Teles, R. M., Ouyang, S., Jiang, Y., Iyer, S. S., Zaver, S. A., Schenk, M., Zeng, S., Zhong, W., Liu, Z. J., Modlin, R. L., Liu, Y. J., & Cheng, G. (2012). The helicase DDX41 recognizes the bacterial secondary messengers cyclic di-GMP and cyclic di-AMP to activate a type I interferon immune response. *Nat Immunol*, *13*(12), 1155-1161. https://doi.org/10.1038/ni.2460

Patel, D. K., Menon, D. V., Patel, D. H., & Dave, G. (2022). Linkers: A synergistic way for the synthesis of chimeric proteins. *Protein Expr Purif*, *191*, 106012. https://doi.org/10.1016/j.pep.2021.106012

Patel, S. S., & Picha, K. M. (2000). Structure and function of hexameric helicases. *Annu Rev Biochem*, *69*, 651-697. https://doi.org/10.1146/annurev.biochem.69.1.651

Pause, A., & Sonenberg, N. (1992). Mutational analysis of a DEAD box RNA helicase: the mammalian translation initiation factor eIF-4A. *Embo j*, *11*(7), 2643-2654.

Peters, D., Radine, C., Reese, A., Budach, W., Sohn, D., & Jänicke, R. U. (2017). The DEAD-box RNA helicase DDX41 is a novel repressor of p21WAF1/CIP1 mRNA translation. *Journal of Biological Chemistry*, *292*(20), 8331-8341.

Polprasert, C., Schulze, I., Sekeres, M. A., Makishima, H., Przychodzen, B., Hosono, N., Singh, J., Padgett, R. A., Gu, X., Phillips, J. G., Clemente, M., Parker, Y., Lindner, D., Dienes, B., Jankowsky, E., Sauntharajah, Y., Du, Y., Oakley, K., Nguyen, N., . . . Maciejewski, J. P. (2015). Inherited and Somatic Defects in DDX41 in Myeloid Neoplasms. *Cancer Cell*, *27*(5),

658-670. https://doi.org/10.1016/j.ccell.2015.03.017
Pomerantz, R. T., & O'Donnell, M. (2007). Replisome mechanics: insights into a twin DNA polymerase machine. *Trends in microbiology*, *15*(4), 156-164.
Putnam, A. A., & Jankowsky, E. (2013). DEAD-box helicases as integrators of RNA, nucleotide and protein binding. *Biochim Biophys Acta*, *1829*(8), 884-893. https://doi.org/10.1016/j.bbagrm.2013.02.002
Qin, K., Jian, D., Xue, Y., Cheng, Y., Zhang, P., Wei, Y., Zhang, J., Xiong, H., Zhang, Y., & Yuan, X. (2021). DDX41 regulates the expression and alternative splicing of genes involved in tumorigenesis and immune response. *Oncology reports*, *45*(3), 1213-1225.
Qin, X.-W., Luo, Z.-Y., Pan, W.-Q., He, J., Li, Z.-M., Yu, Y., Liu, C., Weng, S.-P., He, J.-G., & Guo, C.-J. (2022). The interaction of mandarin fish DDX41 with STING evokes type I interferon responses inhibiting ranavirus replication. *Viruses*, *15*(1), 58.
Qu, S., Li, B., Qin, T., Xu, Z., Pan, L., Hu, N., Huang, G., Peter Gale, R., & Xiao, Z. (2021). Molecular and clinical features of myeloid neoplasms with somatic DDX41 mutations. *Br J Haematol*, *192*(6), 1006-1010. https://doi.org/10.1111/bjh.16668
Quesada, A. E., Routbort, M. J., DiNardo, C. D., Bueso-Ramos, C. E., Kanagal-Shamanna, R., Khoury, J. D., Thakral, B., Zuo, Z., Yin, C. C., Loghavi, S., Ok, C. Y., Wang, S. A., Tang, Z., Bannon, S. A., Benton, C. B., Garcia-Manero, G., Kantarjian, H., Luthra, R., Medeiros, L. J., & Patel, K. P. (2019). DDX41 mutations in myeloid neoplasms are associated with male gender, TP53 mutations and high-risk disease. *Am J Hematol*, *94*(7), 757-766. https://doi.org/10.1002/ajh.25486
Ramanathan, M., Majzoub, K., Rao, D. S., Neela, P. H., Zarnegar, B. J., Mondal, S., Roth, J. G., Gai, H., Kovalski, J. R., Siprashvili, Z., Palmer, T. D., Carette, J. E., & Khavari, P. A. (2018). RNA-protein interaction detection in living cells. *Nat Methods*, *15*(3), 207-212. https://doi.org/10.1038/nmeth.4601
Reddy Chichili, V. P., Kumar, V., & Sivaraman, J. (2013). Linkers in the structural biology of protein-protein interactions. *Protein Sci*, *22*(2), 153-167. https://doi.org/10.1002/pro.2206
Roux, K. J., Kim, D. I., Burke, B., & May, D. G. (2018). BioID: A Screen for Protein-Protein Interactions. *Curr Protoc Protein Sci*, *91*, 19.23.11-19.23.15. https://doi.org/10.1002/cpps.51
Roux, K. J., Kim, D. I., Raida, M., & Burke, B. (2012). A promiscuous biotin ligase fusion protein identifies proximal and interacting proteins in mammalian cells. *J Cell Biol*, *196*(6), 801-810. https://doi.org/10.1083/jcb.201112098
Saikruang, W., Ang Yan Ping, L., Abe, H., Kasumba, D. M., Kato, H., & Fujita, T. (2022). The RNA helicase DDX3 promotes IFNB transcription via enhancing IRF-3/p300 holocomplex binding to the IFNB promoter. *Sci Rep*, *12*(1), 3967. https://doi.org/10.1038/s41598-022-07876-z
Sami, A. A., Arabia, S., Sarker, R. H., & Islam, T. (2021). Deciphering the role of helicases and translocases: A multifunctional gene family safeguarding plants from diverse environmental adversities. *Current Plant Biology*, *26*, 100204.
Schmucker, D., Jäckle, H., & Gaul, U. (1997). Genetic analysis of the larval optic nerve projection in Drosophila. *Development*, *124*(5), 937-948.
Schwer, B. (2008). A conformational rearrangement in the spliceosome sets the stage for Prp22-dependent mRNA release. *Molecular cell*, *30*(6), 743-754.
Sébert, M., Passet, M., Raimbault, A., Rahmé, R., Raffoux, E., Sicre de Fontbrune, F., Cerrano, M., Quentin, S., Vasquez, N., Da Costa, M., Boissel, N., Dombret, H., Peffault de Latour, R., Socié, G., Itzykson, R., Fenaux, P., Soulier, J., Adès, L., & Clappier, E. (2019). Germline DDX41 mutations define a significant entity within adult MDS/AML patients. *Blood*, *134*(17), 1441-1444. https://doi.org/10.1182/blood.2019000909
Shih, J. W., & Lee, Y. H. (2014). Human DExD/H RNA helicases: emerging roles in stress survival regulation. *Clin Chim Acta*, *436*, 45-58. https://doi.org/10.1016/j.cca.2014.05.003
Shin, S., Rossow, K. L., Grande, J. P., & Janknecht, R. (2007). Involvement of RNA helicases p68 and p72 in colon cancer. *Cancer research*, *67*(16), 7572-7578.

Shin, W. Y., Yoon, S. Y., Park, R., Kim, J. A., Song, H. H., Bang, H. I., Won, J. H., & Kim, J. (2022). A novel bi-alleleic DDX41 mutations in B-cell lymphoblastic leukemia: case report. *BMC Med Genomics*, *15*(1), 46. https://doi.org/10.1186/s12920-022-01191-2

Singh, R. S., Vidhyasagar, V., Yang, S., Arna, A. B., Yadav, M., Aggarwal, A., Aguilera, A. N., Shinriki, S., Bhanumathy, K. K., Pandey, K., Xu, A., Rapin, N., Bosch, M., DeCoteau, J., Xiang, J., Vizeacoumar, F. J., Zhou, Y., Misra, V., Matsui, H., . . . Wu, Y. (2022). DDX41 is required for cGAS-STING activation against DNA virus infection. *Cell Rep*, *39*(8), 110856. https://doi.org/10.1016/j.celrep.2022.110856

Singleton, M. R., Dillingham, M. S., & Wigley, D. B. (2007). Structure and mechanism of helicases and nucleic acid translocases. *Annu Rev Biochem*, *76*, 23-50. https://doi.org/10.1146/annurev.biochem.76.052305.115300

Smith, P. A., Tripp, B. C., DiBlasio-Smith, E. A., Lu, Z., LaVallie, E. R., & McCoy, J. M. (1998). A plasmid expression system for quantitative in vivo biotinylation of thioredoxin fusion proteins in Escherichia coli. *Nucleic Acids Res*, *26*(6), 1414-1420. https://doi.org/10.1093/nar/26.6.1414

Stahl, H., Dröge, P., & Knippers, R. (1986). DNA helicase activity of SV40 large tumor antigen. *Embo j*, *5*(8), 1939-1944. https://doi.org/10.1002/j.1460-2075.1986.tb04447.x

Sugino, A., Ryu, B. H., Sugino, T., Naumovski, L., & Friedberg, E. C. (1986). A new DNA-dependent ATPase which stimulates yeast DNA polymerase I and has DNA-unwinding activity. *J Biol Chem*, *261*(25), 11744-11750.

Sunna, S., Bowen, C., Zeng, H., Rayaprolu, S., Kumar, P., Bagchi, P., Dammer, E. B., Guo, Q., Duong, D. M., Bitarafan, S., Natu, A., Wood, L., Seyfried, N. T., & Rangaraju, S. (2023). Cellular Proteomic Profiling Using Proximity Labeling by TurboID-NES in Microglial and Neuronal Cell Lines. *Mol Cell Proteomics*, *22*(6), 100546. https://doi.org/10.1016/j.mcpro.2023.100546

Truglio, J. J., Croteau, D. L., Van Houten, B., & Kisker, C. (2006). Prokaryotic nucleotide excision repair: the UvrABC system. *Chemical reviews*, *106*(2), 233-252.

Tsukamoto, T., Gearhart, M. D., Kim, S., Mekonnen, G., Spike, C. A., & Greenstein, D. (2020). Insights into the Involvement of Spliceosomal Mutations in Myelodysplastic Disorders from Analysis of SACY-1/DDX41 in Caenorhabditis elegans. *Genetics*, *214*(4), 869-893. https://doi.org/10.1534/genetics.119.302973

Tungalag, S., Shinriki, S., Hirayama, M., Nagamachi, A., Kanai, A., Inaba, T., & Matsui, H. (2023). Ribosome profiling analysis reveals the roles of DDX41 in translational regulation. *International Journal of Hematology*, *117*(6), 876-888.

Tuteja, N., Tuteja, R., Rahman, K., Kang, L. Y., & Falaschi, A. (1990). A DNA helicase from human cells. *Nucleic Acids Res*, *18*(23), 6785-6792. https://doi.org/10.1093/nar/18.23.6785

Ummethum, H., & Hamperl, S. (2020). Proximity Labeling Techniques to Study Chromatin [Review]. *Frontiers in Genetics*, *11*. https://doi.org/10.3389/fgene.2020.00450

Van Brabant, A. J., Stan, R., & Ellis, N. A. (2000). DNA helicases, genomic instability, and human genetic disease. *Annual review of genomics and human genetics*, *1*(1), 409-459.

Venema, J., & Tollervey, D. (1999). Ribosome synthesis in Saccharomyces cerevisiae. *Annu Rev Genet*, *33*, 261-311. https://doi.org/10.1146/annurev.genet.33.1.261

Venkatesan, M., Silver, L. L., & Nossal, N. G. (1982). Bacteriophage T4 gene 41 protein, required for the synthesis of RNA primers, is also a DNA helicase. *J Biol Chem*, *257*(20), 12426-12434.

von Hippel, P. H., & Delagoutte, E. (2003). Macromolecular complexes that unwind nucleic acids. *Bioessays*, *25*(12), 1168-1177.

Winstone, L., Jung, Y., & Wu, Y. (2024). DDX41: exploring the roles of a versatile helicase. *Biochem Soc Trans*, *52*(1), 395-405. https://doi.org/10.1042/bst20230725

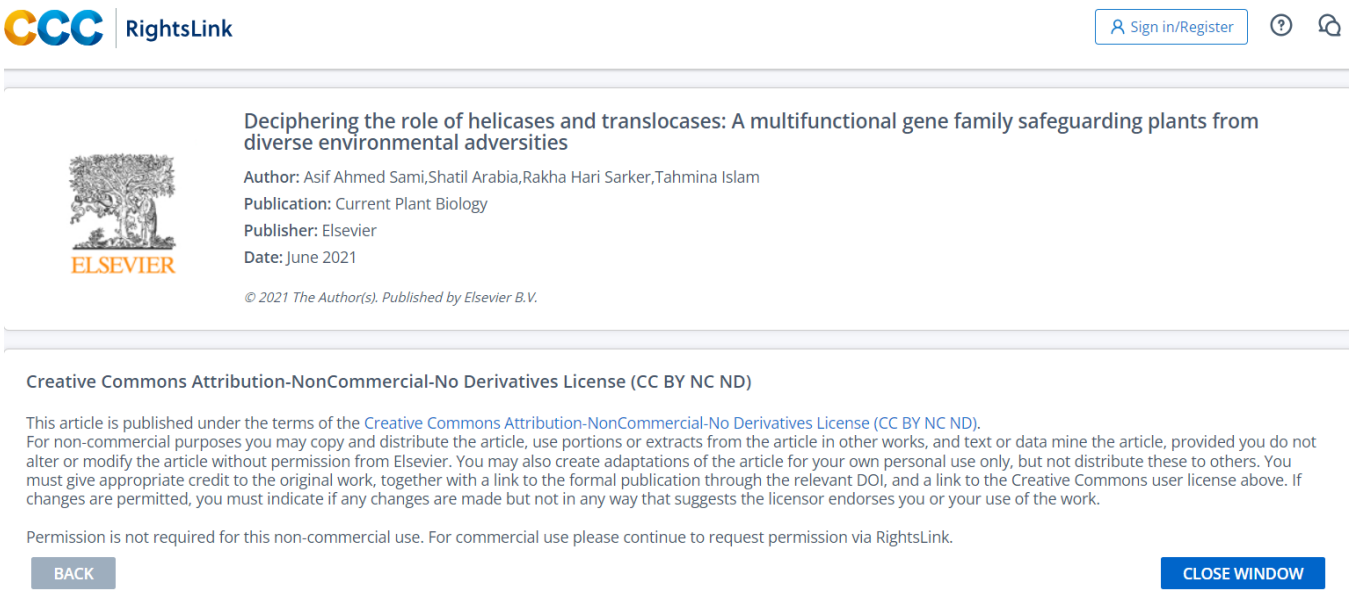
Woolford, J. L., Jr., & Baserga, S. J. (2013). Ribosome biogenesis in the yeast Saccharomyces cerevisiae. *Genetics*, *195*(3), 643-681. https://doi.org/10.1534/genetics.113.153197

Wu, D.-W., Liu, W.-S., Wang, J., Chen, C.-Y., Cheng, Y.-W., & Lee, H. (2011). Reduced p21WAF1/CIP1 via Alteration of p53-DDX3 pathway is associated with poor relapse-free survival in early-stage human papillomavirus–associated lung cancer. *Clinical Cancer*

Research, *17*(7), 1895-1905.
Xiong, Z., Lo, H. P., McMahon, K. A., Martel, N., Jones, A., Hill, M. M., Parton, R. G., & Hall, T. E. (2021). In vivo proteomic mapping through GFP-directed proximity-dependent biotin labelling in zebrafish. *Elife*, *10*. https://doi.org/10.7554/eLife.64631
Yoneyama-Hirozane, M., Kondo, M., Matsumoto, S. I., Morikawa-Oki, A., Morishita, D., Nakanishi, A., Kawamoto, T., & Nakayama, M. (2017). High-Throughput Screening to Identify Inhibitors of DEAD Box Helicase DDX41. *SLAS Discov*, *22*(9), 1084-1092. https://doi.org/10.1177/2472555217705952
Zhang, S., & Grosse, F. (1994). Nuclear DNA helicase II unwinds both DNA and RNA. *Biochemistry*, *33*(13), 3906-3912.
Zhang, Z., Yuan, B., Bao, M., Lu, N., Kim, T., & Liu, Y. J. (2011). The helicase DDX41 senses intracellular DNA mediated by the adaptor STING in dendritic cells. *Nat Immunol*, *12*(10), 959-965. https://doi.org/10.1038/ni.2091
Zhu, X., Wang, D., Zhang, H., Zhou, Y., Luo, R., Chen, H., Xiao, S., & Fang, L. (2014). Molecular cloning and functional characterization of porcine DEAD (Asp–Glu–Ala–Asp) box polypeptide 41 (DDX41). *Developmental & Comparative Immunology*, *47*(2), 191-196.

PERMISSION TO RE-USE FIGURES

Figure 1: Classification of the helicases



CCC RightsLink

Sign in/Register

Deciphering the role of helicases and translocases: A multifunctional gene family safeguarding plants from diverse environmental adversities

Author: Asif Ahmed Sami,Shatil Arabia,Rakha Hari Sarker,Tahmina Islam

Publication: Current Plant Biology

Publisher: Elsevier

Date: June 2021

© 2021 The Author(s). Published by Elsevier B.V.

Creative Commons Attribution-NonCommercial-No Derivatives License (CC BY NC ND)

This article is published under the terms of the Creative Commons Attribution-NonCommercial-No Derivatives License (CC BY NC ND).
For non-commercial purposes you may copy and distribute the article, use portions or extracts from the article in other works, and text or data mine the article, provided you do not alter or modify the article without permission from Elsevier. You may also create adaptations of the article for your own personal use only, but not distribute these to others. You must give appropriate credit to the original work, together with a link to the formal publication through the relevant DOI, and a link to the Creative Commons user license above. If changes are permitted, you must indicate if any changes are made but not in any way that suggests the licensor endorses you or your use of the work.

Permission is not required for this non-commercial use. For commercial use please continue to request permission via RightsLink.

BACK

CLOSE WINDOW

Figure 2A: Structural features of DEAD-box helicases

This is a License Agreement between University of Saskatchewan ("User") and Copyright Clearance Center, Inc. ("CCC") on behalf of the Rightsholder identified in the order details below. The license consists of the order details, the Marketplace Permissions General Terms and Conditions below, and any Rightsholder Terms and Conditions which are included below.

All payments must be made in full to CCC in accordance with the Marketplace Permissions General Terms and Conditions below.

Order Date	23-Aug-2024	Type of Use	Republish in a thesis/dissertation
Order License ID	1518864-2		
ISSN	1437-7772	Publisher	Springer Japan KK
		Portion	Image/photo/illustration

LICENSED CONTENT

Publication Title	International journal of clinical oncology	Publication Type	e-Journal
		Start Page	795
Article Title	The DEAD-box protein family of RNA helicases: sentinels for a myriad of cellular functions with emerging roles in tumorigenesis	End Page	825
		Issue	5
		Volume	26
		URL	http://www.springerlink.com/content/103370/
Author / Editor	Japan Society of Clinical Oncology.		
Date	01/01/1996		
Language	English		
Country	Japan		
Rightsholder	Springer Nature BV		

REQUEST DETAILS

Portion Type	Image/photo/illustration	Distribution	Worldwide
Number of Images / Photos / Illustrations	1	Translation	Original language of publication
Format (select all that apply)	Electronic	Copies for the Disabled?	No
		Minor Editing Privileges?	No
Who Will Republish the Content?	Academic institution	Incidental Promotional Use?	No
Duration of Use	Life of current edition	Currency	CAD
Lifetime Unit Quantity	Up to 499		
Rights Requested	Main product		

NEW WORK DETAILS

Title	Professor	Institution Name	University of Saskatchewan
Instructor Name	Dr. Yuliang Wu		

Expected Presentation Date	2024-10-31

ADDITIONAL DETAILS

The Requesting Person / Organization to Appear on the License	University of Saskatchewan

REQUESTED CONTENT DETAILS

Title, Description or Numeric Reference of the Portion(s)	Figure 1
Editor of Portion(s)	Ali, Mohamed A. M.
Volume / Edition	26
Page or Page Range of Portion	795-825
Title of the Article / Chapter the Portion Is From	The DEAD-box protein family of RNA helicases: sentinels for a myriad of cellular functions with emerging roles in tumorigenesis
Author of Portion(s)	Ali, Mohamed A. M.
Publication Date of Portion	2021-04-30

Figure 2B: DEAD-box helicases in different processes of RNA metabolism

This Agreement between University of Saskatchewan ("You") and Springer Nature ("Springer Nature") consists of your license details and the terms and conditions provided by Springer Nature and Copyright Clearance Center.

License Number	5863291082992
License date	Sep 06, 2024
Licensed Content Publisher	Springer Nature
Licensed Content Publication	Nature Reviews Molecular Cell Biology
Licensed Content Title	From unwinding to clamping — the DEAD box RNA helicase family
Licensed Content Author	Patrick Linder et al
Licensed Content Date	Jul 22, 2011
Type of Use	Thesis/Dissertation
Requestor type	academic/university or research institute
Format	electronic
Portion	figures/tables/illustrations
Number of figures/tables/illustrations	1

Would you like a high resolution image with your order?	no
Will you be translating?	no
Circulation/distribution	50000 or greater
Author of this Springer Nature content	no
Title of new work	Optimising the Conditions to Identify the DDX41 Interactome by BioID
Institution name	University of Saskatchewan
Expected presentation date	Oct 2024
Portions	Figure 1
The Requesting Person / Organization to Appear on the License	University of Saskatchewan
Requestor Location	University of Saskatchwan 107 Wiggins road Health Sciences building University of Saskatchewan Saskatoon, SK S7N 5E5 Canada Attn: University of Saskatchwan
Billing Type	Invoice
Billing Address	University of Saskatchwan 107 Wiggins road Health Sciences building University of Saskatchewan Saskatoon, SK S7N 5E5 Canada Attn: Ananaya Charaya
Total	0.00 CAD

Figure 4: DDX41 germline and somatic mutations in MDS/AML patients

This Agreement between University of Saskatchewan ("You") and John Wiley and Sons ("John Wiley and Sons") consists of your license details and the terms and conditions provided by John Wiley and Sons and Copyright Clearance Center.

License Number	5855001435558
License date	Aug 23, 2024
Licensed Content Publisher	John Wiley and Sons
Licensed Content Publication	British Journal of Haematology
Licensed Content Title	Molecular and clinical features of myeloid neoplasms with somatic DDX41 mutations
Licensed Content Author	Shiqiang Qu, Bing Li, Tiejun Qin, et al
Licensed Content Date	Apr 19, 2020
Licensed Content Volume	192
Licensed Content Issue	6
Licensed Content Pages	5
Type of use	Dissertation/Thesis
Requestor type	University/Academic

End Page	670
Type of Use	reuse in a thesis/dissertation
Portion	figures/tables/illustrations
Number of figures/tables/illustrations	1
Format	electronic
Are you the author of this Elsevier article?	No
Will you be translating?	No
Title of new work	Optimising the Conditions to Identify the DDX41 Interactome by BioID
Institution name	University of Saskatchewan
Expected presentation date	Oct 2024
Portions	Figure 5A
The Requesting Person / Organization to Appear on the License	University of Saskatchewan
Requestor Location	University of Saskatchwan 107 Wiggins road Health Sciences building University of Saskatchewan Saskatoon, SK S7N 5E5 Canada Attn: University of Saskatchwan
Publisher Tax ID	GB 494 6272 12

Figure 5: DDX41 interactions with spliceosomal protein complexes

This Agreement between University of Saskatchewan ("You") and Elsevier ("Elsevier") consists of your license details and the terms and conditions provided by Elsevier and Copyright Clearance Center.

License Number	5855010684022
License date	Aug 23, 2024
Licensed Content Publisher	Elsevier
Licensed Content Publication	Cancer Cell
Licensed Content Title	Inherited and Somatic Defects in DDX41 in Myeloid Neoplasms
Licensed Content Author	Chantana Polprasert,Isabell Schulze,Mikkael A. Sekeres,Hideki Makishima,Bartlomiej Przychodzen,Naoko Hosono,Jarnail Singh,Richard A. Padgett,Xiaorong Gu,James G. Phillips,Michael Clemente,Yvonne Parker,Daniel Lindner,Brittney Dienes et al.
Licensed Content Date	May 11, 2015
Licensed Content Volume	27
Licensed Content Issue	5
Licensed Content Pages	13
Start Page	658

End Page	670
Type of Use	reuse in a thesis/dissertation
Portion	figures/tables/illustrations
Number of figures/tables/illustrations	1
Format	electronic
Are you the author of this Elsevier article?	No
Will you be translating?	No
Title of new work	Optimising the Conditions to Identify the DDX41 Interactome by BioID
Institution name	University of Saskatchewan
Expected presentation date	Oct 2024
Portions	Figure 5A
The Requesting Person / Organization to Appear on the License	University of Saskatchewan
Requestor Location	University of Saskatchwan 107 Wiggins road Health Sciences building University of Saskatchewan Saskatoon, SK S7N 5E5 Canada Attn: University of Saskatchwan
Publisher Tax ID	GB 494 6272 12