UNIVERSITY OF CRETE

SCHOOL OF SCIENCES & ENGINEERING

DEPARTMENT OF BIOLOGY

Graduate Program in Molecular Biology and Biomedicine

“Epitranscriptomic control of DNA damage induced R-loops”

Nikoleta Pateraki A.M 448

Supervisor: Dr. Evgenia Ntini

Heraklion, 2023

Examination committee

Dr. Evgenia Ntini
Principal Investigator, Institute of Molecular biology and Biotechnology-Foundation of Research and Technology-Hellas.

Dr. Marina Vidaki
Assistant Professor of Cellular-Molecular Biology, Faculty of Medicine, Division of Basic Sciences, University of Crete

Dr. George Garinis
Professor of Genetics, School of Sciences and Engineering, Department of Biology, University of Crete

Acknowledgements

The present thesis was conducted in the RNA and Gene regulation Laboratory at the Institute of Molecular Biology and Biotechnology-FORTH. The completion of this thesis would not have been possible without the valuable help and support of some people.

This year I learned a lot; I adapted new skills and I'm grateful that I was part of that graduate program and that group. I owe a big thanks to my supervisor, Evgenia, who really trusted me with this project. I learned that starting a new project from scratch can be difficult, however, the knowledge I gained, and the independence overcame all the challenges. Thank you for including me in your team and thank you for exposing me to the "world of RNA modifications". I'm lucky to have found the field that I'm most passionate about.

In addition, I would like to express my gratitude towards Dr. Marina Vidaki and Dr. George Garinis. Thank you for overseeing and evaluating this thesis.

I would also like to thank all the colleagues and friends that I had on my side, along this journey. I owe my gratitude to all the members of the Ntini lab, especially Christos and Maria, but also the rest of the members, with whom I had a great collaboration but also fun. Thanks also to Christina and Evita, with whom we shared some nice moments and made the everyday lab routine more enjoyable. I would also like to thank Sofia and Nikoletta, who really supported me, both technically and emotionally, and shared their thoughts with me over constructive discussions.

Also, I want to thank my friends, thank you for your continuous support over the years, during both great and challenging times. Last but not least, with this MSc thesis, a cycle of 6 years of studying at the University of Crete comes to an end. I am grateful for the knowledge, experience, and the skills that I have gained all these years. The completion of these studies would not have been possible without the selfless support of my parents. Words can't describe how grateful I am for all you've done for me. Thank you.

Table of Contents

Abstract

The “epitranscriptome” refers to the total pool of biochemical modifications of RNA within a cell that do not alter the sequence itself. The most abundant RNA modification on RNA transcripts is N^6-methyladenosine (m6A), which plays a regulatory role in many biological processes, including transcription, splicing and RNA metabolism. Recent studies have focused on unraveling potential roles of m6A in response to various stress factors, including those that induce DNA damage. However, the mechanisms underlying the connection between m6A, and DNA damage response have not been fully elucidated. In this project, I aim to characterize the potential role of m6A in response to DNA damage and its involvement in regulating R-loop formation and/or resolution. To resolve this question, my primary focus is on the nuclear m6A reader YTHDC1, which has been shown to localize to double-strand DNA breaks. It is worth noting that this m6A reader also regulates responses to heat stress, suggesting a potential role in DNA damage response as well. In this Master’s thesis, bioinformatic, molecular and biochemical methods were used to approach this question. Firstly, reanalysis of published RNA-seq data upon YTHDC1 knockdown in HeLa cells was performed, in which several genes that take part in the DNA damage response were identified. Consequently, experiments using DNA damage induced methods (UV radiation) were performed to evaluate the expression levels as well as the localization of YTHDC1. Additionally, in order to better dissect the role of YTHDC1 in response to DNA damage, a specific protein degradation tag (dTAG) system against YTHDC1 was designed and incorporated in HCT116 colon cancer cells. This tool enables the targeted and controlled depletion of the protein of interest. Last but not least, the RChIP protocol was established in order to be able to map R-loops specifically by using dRNase H fused with V5 tag. The dRNase H has a mutation on its catalytic domain but not on the RNA:DNA binding domain, therefore, it can recognize the R-loops without cleaving the RNA moiety. These tools will be used in future experiments to unravel the role of YTHDC1 in stress responses.

Περίληψη

Το "επιμεταγράφωμα" (‘epitranscriptome’) αναφέρεται στο σύνολο των βιοχημικών τροποποιήσεων του RNA μέσα σε ένα κύτταρο, οι οποίες δεν μεταβάλλουν την ίδια την αλληλουχία του RNA. Η πιο συχνή τροποποίηση RNA στα μετάγραφα RNA είναι η N6-μεθυλαδενοσίνη (m6A), η οποία παίζει ρυθμιστικό ρόλο σε πολλές βιολογικές διεργασίες, συμπεριλαμβανομένης της μεταγραφής, του ματίσματος και του μεταβολισμού του RNA. Πρόσφατες έρευνες έχουν επικεντρωθεί στην μελέτη των πιθανών ρόλων της m6A στην απόκριση σε διάφορους παράγοντες στρες, συμπεριλαμβανομένων εκείνων που προκαλούν βλάβες στο DNA. Ωστόσο, οι μηχανισμοί που διέπουν τη σχέση μεταξύ του m6A και της απόκρισης σε βλάβες του DNA δεν έχουν διευκρινιστεί πλήρως. Σε αυτή τη μελέτη στοχεύω να χαρακτηρίσω τον πιθανό ρόλο της m6A στην απόκριση σε βλάβη του DNA και τη συμμετοχή της στη ρύθμιση του σχηματισμού ή/και της επίλυσης των R-loops. Για την επίλυση αυτού του ερωτήματος, εστιάζω κυρίως στον πυρηνικό m6A αναγνώστη (‘reader) YTHDC1, ο οποίος έχει αποδειχθεί ότι εντοπίζεται σε θραύσεις διπλής έλικας DNA. Αξίζει να σημειωθεί ότι αυτή η πρωτεΐνη ρυθμίζει επίσης τις αποκρίσεις στο θερμικό στρες, γεγονός που υποδηλώνει πιθανό ρόλο και στην απόκριση σε βλάβες του DNA. Στην παρούσα διατριβή χρησιμοποιήθηκαν βιοπληροφορικές, μοριακές και βιοχημικές μέθοδοι για την προσέγγιση αυτού του ερωτήματος. Αρχικά, πραγματοποιήθηκε εκ νέου ανάλυση των δημοσιευμένων δεδομένων RNA-seq από κύτταρα HeLa με σίγαση του YTHDC1, κατά την οποία εντοπίστηκαν αρκετά γονίδια που λαμβάνουν μέρος στην απόκριση σε βλάβες του DNA. Στη συνέχεια πραγματοποιήθηκαν πειράματα με μεθόδους επαγόμενης βλάβης του DNA (UV ακτινοβολία) και αξιολογήθηκαν τα επίπεδα καθώς και ο εντοπισμός του YTHDC1. Επιπλέον, προκειμένου να διερευνηθεί καλύτερα ο ρόλος του YTHDC1 στην απόκριση σε στρες, σχεδιάστηκε ένα ειδικό σύστημα αποικοδόμησης πρωτεϊνών (dTAG) έναντι του YTHDC1 που χρησιμοποιήθηκε σε κύτταρα HCT116. Αυτό το εργαλείο επιτρέπει τη στοχευμένη και ελεγχόμενη αποδόμηση της πρωτεΐνης ενδιαφέροντος. Τέλος, η μέθοδος RChIP εδραιώθηκε ως εναλλακτική μέθοδος χαρτογράφησης των δομών R-loop. Στην μέθοδο αυτή, χρησιμοποιείται το ένζυμο dRNase H, συνδεδεμένο με ετικέτα (tag) V5. Η dRNase H έχει μια μετάλλαξη στην καταλυτική της περιοχή, αλλά όχι στην περιοχή πρόσδεσης RNA:DNA, επομένως, μπορεί να αναγνωρίσει τα R-loops χωρίς να διασπάσει την αλυσίδα RNA. Αυτά τα εργαλεία θα χρησιμοποιηθούν σε μελλοντικά πειράματα για την αποκάλυψη του ρόλου της YTHDC1 στις αποκρίσεις στο στρες.

Introduction

1. Insights into the “epitranscriptome”

Like other major biomolecules, such as DNA, proteins and lipids, RNA can be chemically modified. And even though it was known for years that RNAs can undergo modifications, only in the last decade the advancement of technologies made their study feasible. The term “epitranscriptome” was given to describe the biochemical modifications of the RNA within a cell (Saletore et al, 2012). Up to date, over 170 different types of epitranscriptomic signatures have been described in RNA molecules, adding another regulatory layer in gene expression. The diversity of RNA modifications on the mRNAs is shown in the following schematic.

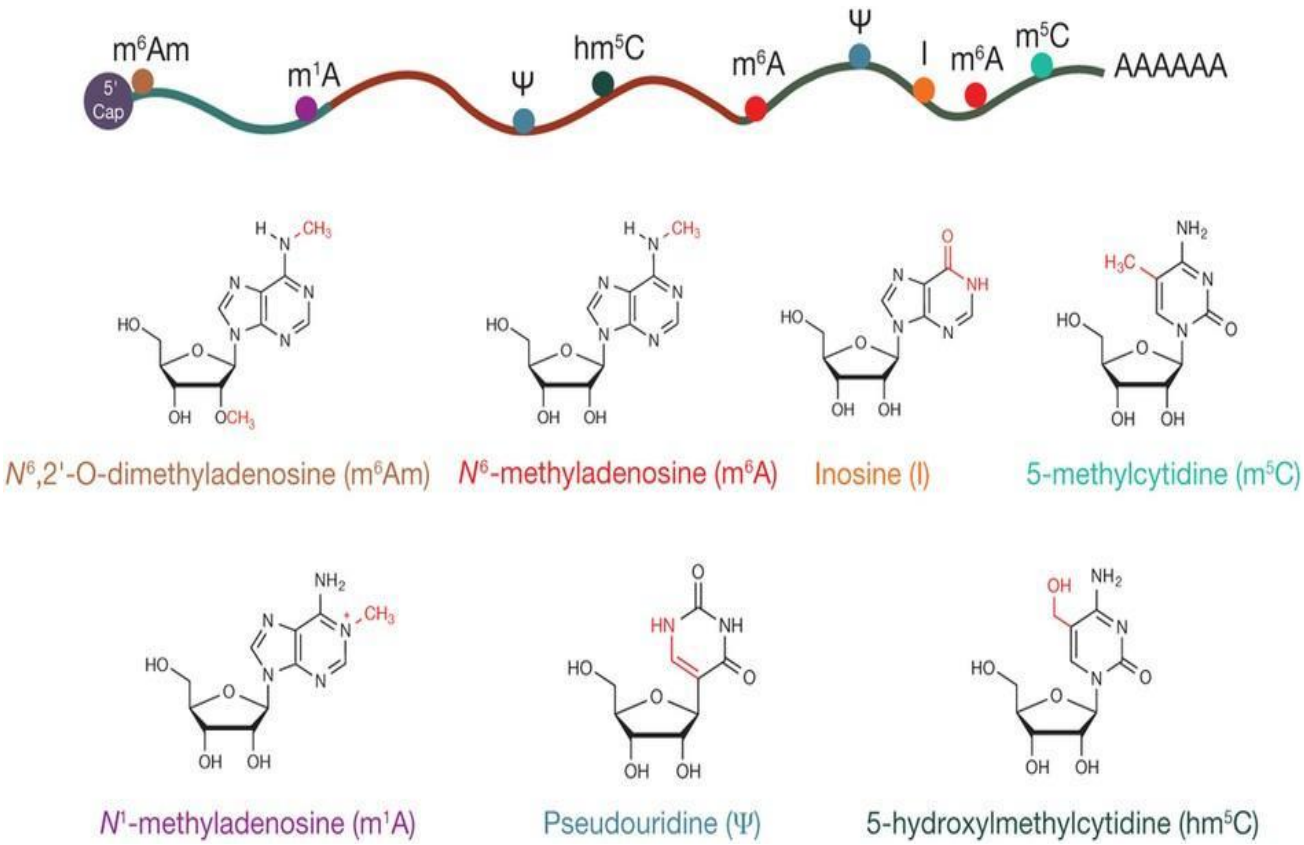


Figure 1. Chemical Modifications on mRNAs. Illustration of the chemical modifications that can occur on mRNA molecules. The upper part of the figure displays the distribution of these modifications along the mRNA molecule, based on current knowledge. Additionally, specific RNA modifications are highlighted using color coding. (Li et al, 2016)

Some of these modifications include the base isomerization to produce pseudouridine (Ψ), methylation of nucleotides to produce m6A, m1A, m5C, methylation of the ribose sugar (Nm, m6Am) and the oxidation of m5C to produce 5-hydroxymethylcytosine (hm5C) (Figure 1). Among these, the most abundant and well-studied epitranscriptomic mark is the methylation of the nitrogen on the 6th position of the adenosine or N6-Methyladenosine (m6A) (Arzumanian et al, 2022), which is estimated to decorate 20%-40% of the mammalian transcripts, while each transcript contains multiple methyl groups on its adenosines (Frye et al, 2018). In this study, the focus was on the m6A RNA modification whose actors and functional roles are presented below.

2. The RNA methylation is dynamic

Like DNA methylation and histone modification, RNA modifications are dynamic, and reversible and do not alter the RNA sequence itself. The dynamic nature of the RNA modifications as well as the output of their presence in the cellular context are ensured by specific proteins that can deposit, erase and recognize these modifications on the RNA molecules. Mainly three groups of proteins are required for epitranscriptomic control in order to maintain specific modification patterns in cell cycle and cell state dependent manner. Specific proteins, known as RNA-modifying proteins (RMPs), are responsible for depositing, reading, and removing the modification on the RNAs and can be divided in the following groups (Orsolic et al, 2022).

The m6A is deposited co-transcriptionally in the nucleus by specific proteins, which are termed as “writers”. The key writers of m6A are Methyltransferase-like protein 3 (METTL3) and methyltransferase-like protein 14 (METTL14). METTL3 is the first enzyme reported to catalyze the methylation of RNA molecules (Bokar et al, 1997). More precisely, METTL3 interacts with the METTL14 forming a heterodimer complex (Wang et al, 2016; Liu et al, 2014), the “m6a writer complex”, which is the main component that mediates co-transcriptional mRNA methylation (Figure 2) (Zaccara et al, 2019; Schwartz et al, 2012). The writers recognize a specific motif on the RNA sequence, the DRACH motif (D= A, G or U; R = A or G H = A, C or U) and deposit the mark there. Even though the DRACH motifs are widespread, only a portion of them is methylated *in vivo* (Meyer et al, 2012; Dominisini et al, 2012). Apart from METTL3 and METTL16, additional proteins have been characterized by their ability to methylate adenosines. For example, Methyltransferase-like protein 5 (METTL5) can deposit m6A on rRNAs (van Tran et al, 2019) and Methyltransferase-like protein 16 (METTL16) can methylate adenosines on the MAT2A mRNA (Doxtader et al, 2018)

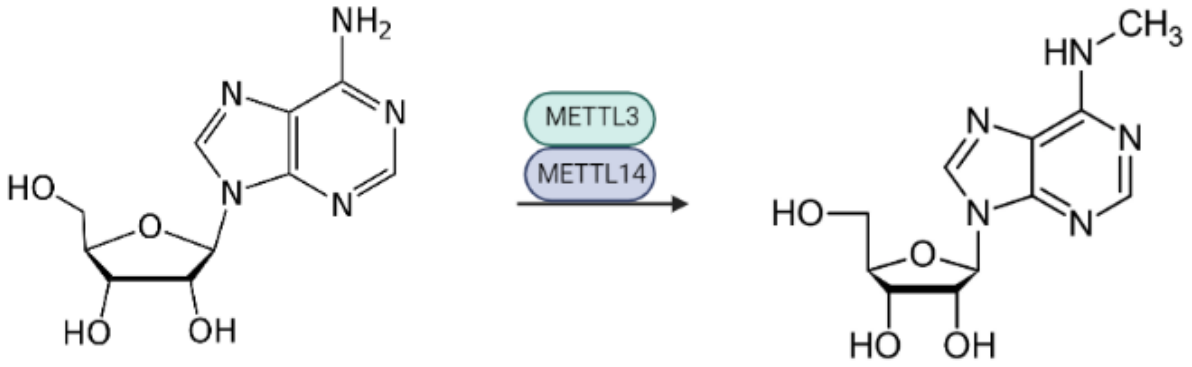


Figure 2. The m6A writer complex. The catalytic subunit METTL3 interacts with METTL14 to form a heterodimer writer complex. The METTL3-METTL14 heterodimer complex transfers the methyl group on the nitrogen on the 6th position of adenosine and therefore forms the m6A modification. (Designed at Biorender: https://app.biorender.com/)

The second important group consists of the proteins that recognize and bind the m6A and because they can “read” this modification across the transcripts, they are defined as “readers”. The “reading” can be executed either directly, by direct recognition from a binding pocket of YTH-domain proteins, or indirectly. The indirect recognition is facilitated through a

structural change that is caused by the modification. Briefly, the m6A readers can be divided into three categories based on the mechanism of m6A recognition. First, there is the direct m6A binding by the readers, YTH domain-containing proteins (YTHDC1, YTHDC2), YTH domain-containing family proteins (YTHDF1, YTHDF2, YTHDF13) and Eukaryotic initiation factor 3 (eIF3).

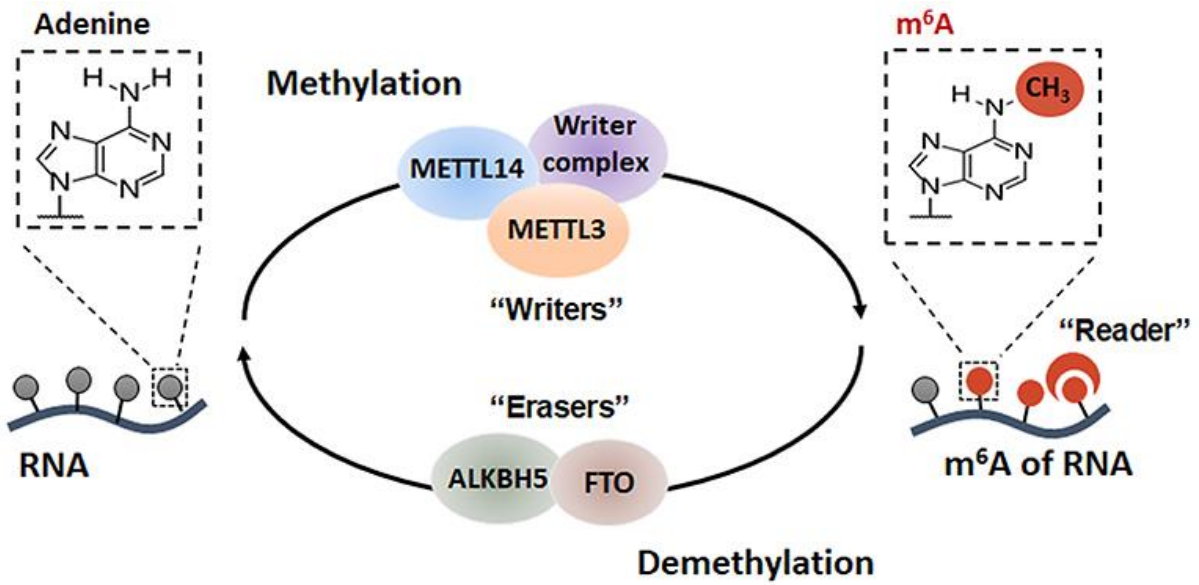


Figure 3. The dynamic methylation process of RNA m6A modification. The methylation of adenosine is installed by the "writers", demethylation is ensured by the "erasers". M6A modified transcripts are recognized and bound by the "readers" (Tang et al, 2015).

Also, there are the heterogeneous nuclear ribonucleoproteins (HNRNPC, HNRNPG, HNRNPA2B1) and insulin-like growth factor 2 mRNA binding proteins (IGF2BP1, IGF2BP2 and IGF2BP3), whose binding is regulated by m6A-induced structural changes. Last but not least, there is the category of proteins that attach to bona fide m6A binding proteins. The last category includes the fragile X retardation protein (FMRP) that directly binds YTHDF2 (Zaccara et al, 2019).

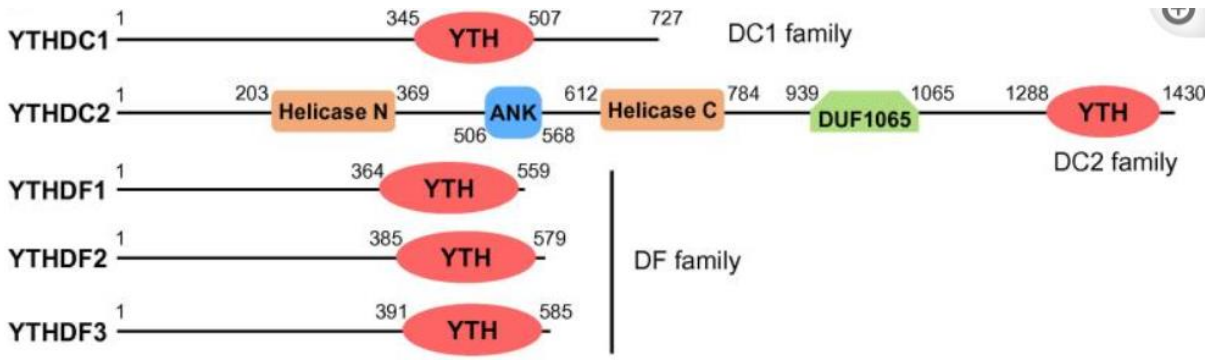


Figure 4. Domain organization of the YTH domain-containing proteins. The aromatic pocket by which the m6A is recognized is depicted in red. (adapted from Liao et al, 2018)

The third and last category of the m6A related factors includes the demethylating proteins, called "erasers" that secure the reversibility of the m6A modification (Liu et al, 2015). Two

m6A demethylases have been reported so far: Fat Mass and Obesity Associated protein (FTO) (Jia et al, 2011) and ALKBH5 (Zheng et al, 2013). FTO belongs to the Fe(II) and oxoglutarate-dependent AlkB oxygenase family, which catalyzes a wide range of biological oxidations (Fu et al, 2013). AlkB homolog 5 (ALKBH5) was also reported to oxidatively reverse m6A in mRNA in vitro and in vivo, with comparable enzymatic kinetics to FTO (Zheng et al, 2013).

2.1 YTH domain-containing protein 1 (YTHDC1)

As described above, the outcomes of RNA methylation and the downstream molecular and cellular effects are dictated by the “m6A readers”. Among the most well-studied m6A readers is the nuclear reader, YTH domain-containing protein 1 or YTHDC1. Early studies of this protein have unraveled the role of YTHDC1 in alternative splicing (Rafalska et al, 2004; Xiao et al, 2016). YTHDC1 facilitates exon inclusion in specific mRNAs through its interaction with pre-mRNA splicing factors. A mechanistic study showed that YTHDC1 can be competitively bound either by the Serine/arginine-rich splicing factor 3 (SRSF3) or by the Serine/arginine-rich splicing factor 10 (SRSF10) dictating different outcomes in each case. When SRSF3 binds to the YTHDC1 it enables the exon inclusion, however, when SRFS10 is bound that leads to exon skipping (Xiao et al, 2016). In the same study it was shown that YTHDC1 is needed for the nuclear speckle localization of splicing factors, including SRSF3 and SRSF10. Apart from its role in splicing regulation, YTHDC1 regulates mRNA export from the nucleus to the cytoplasm (Roundtree et al, 2017). An additional role of YTHDC1 in mRNA metabolism was described in oocytes, where YTHDC1 regulates alternative polyadenylation and therefore embryonic development (Kasowitz et al, 2018). Recent studies have shown that YTHDC1 cross talks with chromatin accessibility factors. In the study of Liu et al, the depletion of this protein was linked to chromatin accessibility and transcription activation in a m6A manner (Liu et al, 2020).

3. m6A is an additional regulatory layer of gene expression

The intense research on m6A, in combination with the advancement in m6A mapping methods, have uncovered regulatory roles of m6A in gene expression. The outcomes of RNA methylation are dictated by the m6A “readers”, the proteins that recognize and bind to the methylated adenosines. The “readers" can orchestrate the various functions of m6A within cells via their protein-protein interactions. One of the primary roles of m6A is in regulation of RNA metabolism, RNA processing, including splicing (Xiao et al, 2016; Louloupi et al, 2018), RNA decay (Wang et al, 2014), nuclear mRNA export (Roundtree et al, 2017) and translation (Lin et al, 2016). Via its role in RNA related functions, m6A regulates biological processes like cell differentiation (Batista et al, 2014), development (Frye et al, 2018), sex determination (Kan et al, 2017), chromosome X inactivation (Patil et al, 2016) and recently it was found to regulate gene dosage compensation in mammals (Rücklé et al, 2023). It is noteworthy that deregulation of m6A has also been implicated in disease development and progression. Abnormalities of m6A and its factors have been linked to cancer (Wang et al, 2020; Barbieri et al, 2017; Liu et al, 2018), neurodegenerative syndromes (Pupak et al, 2022; Han et al,

2020) and metabolic disorders (Zhang et al, 2021). Examples of the biological processes that are regulated by m6A are depicted in Figure 5.

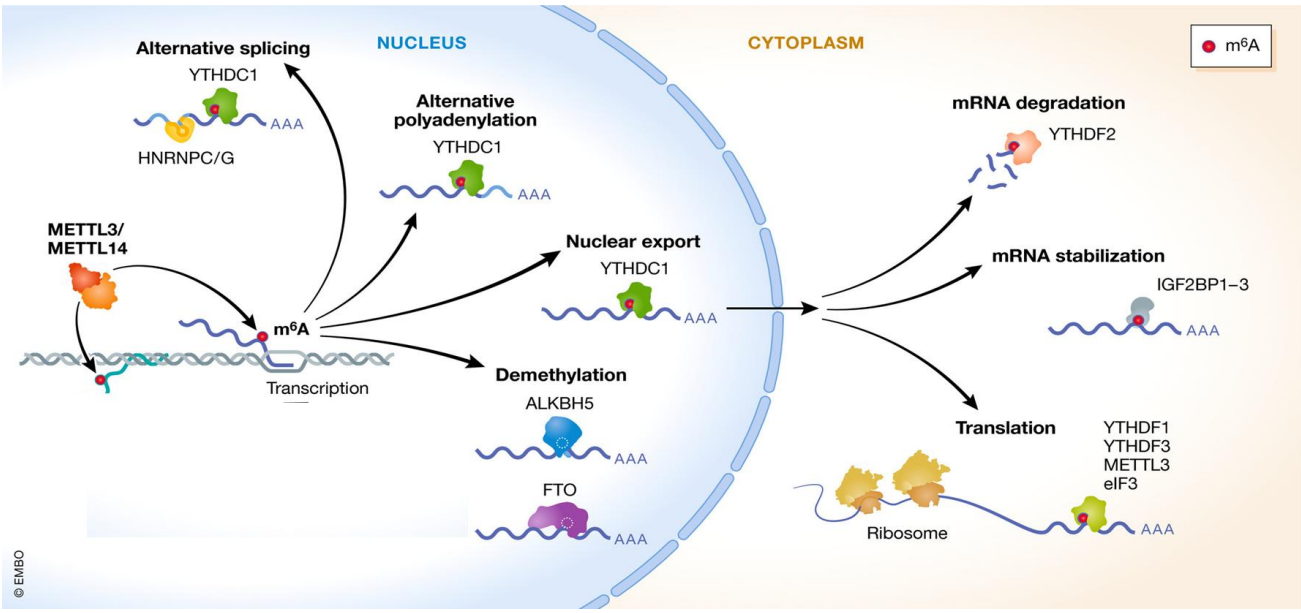


Figure 5. m6A is an additional regulatory element of gene expression (He and He, 2021). The outcomes of RNA methylation are dictated by the proteins that recognize and bind the m6A decorated transcripts. In this schematic representation the writer complex, the “erasers" and some “readers” are shown. Some of the processes that are regulated by m6A are alternative splicing, mRNA export, translation, and RNA stability.

4. m6A and DNA damage

m6A has attracted scientific interest during the past decade and researchers have started studying its role in many biological processes including DNA damage and repair. Recent studies have linked m6A with DNA damage response. Xiang et al, showed that in response to ultraviolet (UV) irradiation, m6A is accumulated rapidly at the DNA damage sites, while the knockdown of METTL3 led to delayed DNA repair unraveling a novel role of m6A in UV-induced DNA damage response (Xiang et al, 2017). A better insight of the mechanism came later in 2020, when another study showed that in response to double strand breaks, METTL3 is phosphorylated by ATM and localized to the damaged sites, where it deposits methyl groups on the DNA damage associated RNAs. In the same study it was depicted that YTHDC1 is recruited on the m6A-RNA methylated sites of the RNA strand, potentially engaged in RNA-DNA hybrid structures at the DNA damaged sites. Another indication of the significance of METTL3 in DNA damage response, is that METTL3 deficient cells showed impaired repair mechanisms and damage accumulation (Zhang et al, 2020). Last but not least, Svobodová Kovaríková et al, showed that after UVA microirradiation m6A is accumulated at the RNA in vicinity to damaged chromatin after two minutes post treatment and this indicates that (m6A) RNA methylation may be important for downstream events (Svobodová Kovaríková et al, 2020).

Even though there is emerging data suggesting that m6A could regulate DNA damage response through its recognizing RNA-binding protein readers and could act as an

intermediate regulatory player to DNA damage repair, the mechanisms and the factors that orchestrate these pathways, are not fully uncovered.

5. RNA:DNA hybrids (R-loops)

The formation of RNA:DNA hybrids occurs naturally either during replication or during transcription. However, not all RNA:DNA hybrids are categorized as R-loops: R-loops are trihybrid nucleic acid structures that are produced when the nascent RNA anneals with the template DNA, while dislocating the single stranded DNA. Notably R-loops are distinct structures from DNA-RNA triplex helices (Aguilera and Muse, 2012). One major characteristic of R-loops allowing their definition is that they are formed co-transcriptionally. It has been shown that R-loops are mainly formed while RNA polymerase II transcribes a C-rich DNA template so that a G-rich nascent RNA molecule is generated (Figure 6) (Daniels et al, 1995; Li and Manley 2005; Ginno et al, 2012). Those are found mostly at promoters and terminator regions in polyA encoding genes, which implies their contribution in regulating gene expression (Ginno et al, 2012). Over the last decade, the R-loop field has attracted scientific interest which has led to the discovery that RNA:DNA hybrids are like a “double-edged sword”, since they can act both as regulators of gene expression and as a menace to genomic stability (Skourti-Stathaki and Proudfoot, 2014). The factors that can contribute to R-loop formation and stabilization vary. Defects in the factors that can resolve R-loops (i.e. RNase H) (Wahba et al, 2011), DNA secondary structures (such as G-quadruplexes) (Sundquist et al, 1989) and nicks in the non-template strand downstream of a promoter (Roy et al, 2010) are some of the factors that can lead to the formation and stabilization of RNA:DNA hybrids (Hegazy et al, 2020).

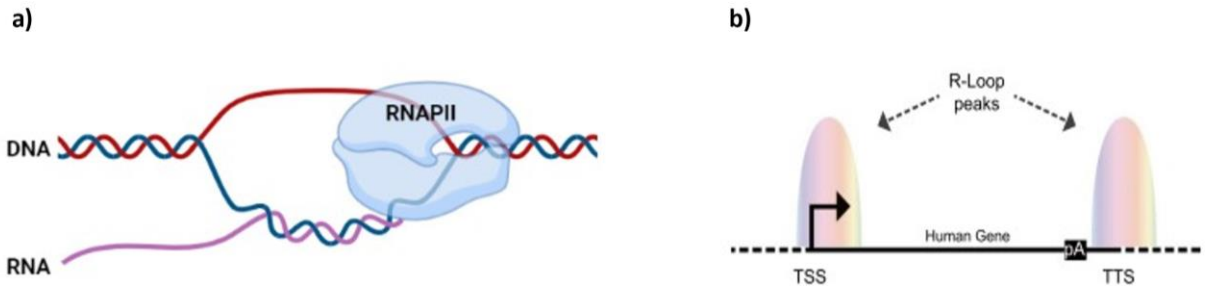


Figure 6. The formation of R-loops. a) During transcription R-loops can be formed when the nascent RNA anneals with the template DNA strand while dislocating the non-template ssDNA. (Designed at Biorender: https://app.biorender.com/), b) R-loops are enriched both at the gene start and the gene end (Skourti-Stathaki and Proudfoot, 2014).

6. Biological roles of R-loops

The formation of R-loops has regulatory effects on gene expression. As mentioned above, R-loops are enriched at the gene ends, both at a gene promoter region and at the 3’ end of a gene, implying their roles in gene expression regulation. At the 3’ end, RNA:DNA hybrids can contribute to transcription termination by modulating the pausing of RNA Polymerase II, which has been characterized as an essential step for transcription termination (Skourti-

Stathaki and Proudfoot, 2011). At the 5' end of genes, the RNA:DNA hybrid formation has been linked to gene activation by modulating chromatin accessibility (Sun et al, 2013; Powell et al, 2013; Xu et al, 2023).

One of the very first indications that R-loops regulate cellular functions came in 2003, when it was shown that R-loops are important for the class switch recombination mechanism during the activation of B cells (Yu et al, 2003).

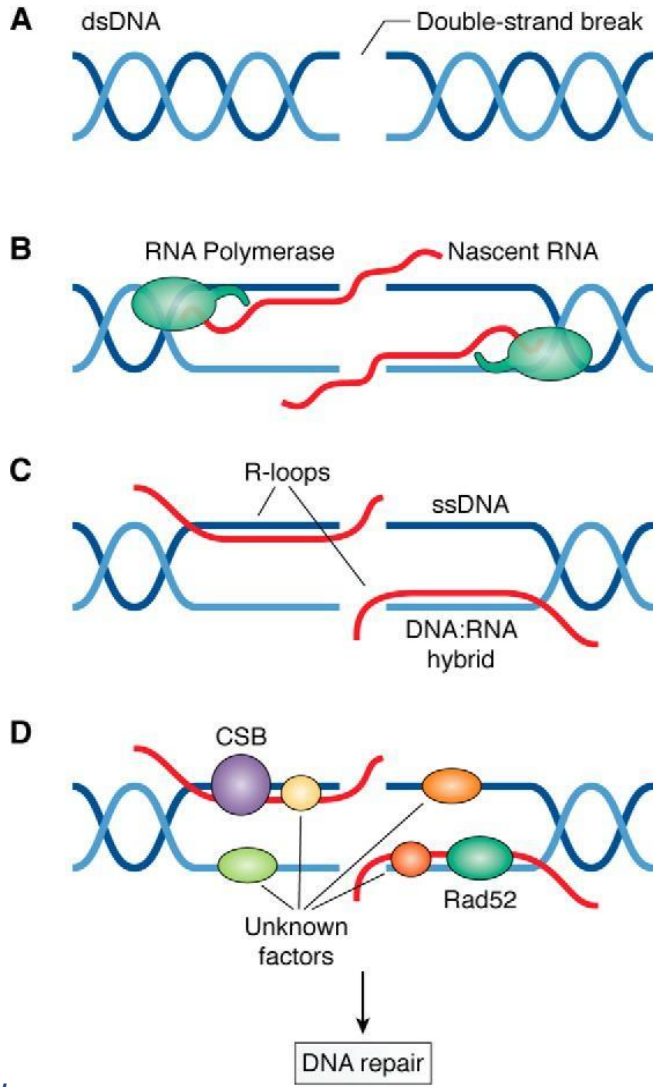


Figure 7. Proposed model for the role of R-loops in DSB repair. A and B) transcription is induced in double strand breaks, therefore, c) there is an accumulation of R-loops at DNA damaged sites. D) The presence of R-loops can recruit factors that take part in DSB repair, such as CSB and Rad52 (Hegazy et al, 2020).

Last but not least, R-loops are crucial intermediates in DNA damage repair mechanisms (Hegazy et al, 2020; Goulielmaki et al, 2021). It has been shown that R-loops can accumulate at damaged DNA sites (Ohle et al, 2016). It is noteworthy that two studies have shown that the single stranded (ssRNA) that is produced from transcriptional active double strand breaks, is involved in Double Strand Break repair (DSB repair). Both studies showed that the ssRNA can be recognized by Rad52, an important protein for homologous recombination and repair (Nogueira et al, 2019), which in turn activates DSB repair (Keskin et al, 2014; Mazina

et al, 2017). In line with these studies, it has been shown that the removal of the RNA from R-loops impairs Homologous Recombination (HR) and Non-Homologous End Joining (NHEJ) repair in human cells (Lu et al, 2018). There is also data supporting that RNA:DNA hybrids may recruit factors that take part in DSB repair, one proposed factor is *Cockayne syndrome protein B* (CSB) (Teng et al, 2018).

7. R-loops and genomic instability: What protects us from R-loops?

Even though R-loops have regulatory roles in gene expression regulation, their accumulation has been associated with genomic instability. By definition, R-loops are trihybrid RNA:DNA structures with a dislocated single stranded DNA. This displacement leaves the DNA strand exposed to nucleases and thus constitutes a threat to the genome, especially when these structures are accumulated. R-loops have also been associated with diseases like Prader-Willi Syndrome (Powell et al, 2013), Aicardi-Goutieres syndrome (AGS) (Lim et al, 2015) , chronic inflammation (Chatzidoukaki et al, 2021 ; Crossley et al, 2023) and different types of cancers (reviewed in Crossley et al, 2019).

It is therefore important to have a mechanism of R-loop homeostasis, i.e. a balance between the formation and the resolution of these structures. The cells have adapted mechanisms by which they can resolve R-loops when their presence inside the nucleus is not needed. Over the years different factors have been proposed to either resolve or prevent RNA:DNA hybrid accumulation, therefore, balancing R-loop homeostasis. Among these factors are nucleases, helicases, topoisomerases, RNA binding and processing proteins. Some of the protective mechanisms that have been described in the literature to regulate R-loop balance are briefly presented below (Figure 8).

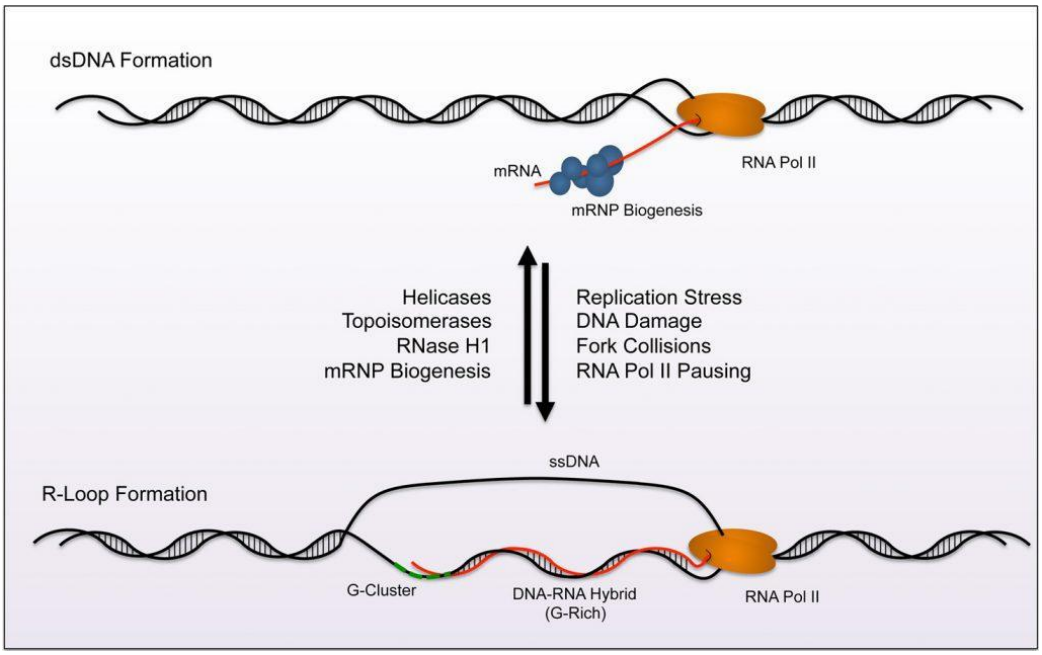


Figure 8. Model of R-loop homeostasis. R-loops can be formed as a result of biological stress such as replication stress, RNA polymerase II pausing, DNA damage stress. RNA:DNA hybrids can be resolved by helicases, topoisomerases and nucleases (Allison and Wang, 2019).

A significant group of ribonucleases that can resolve R-loops are RNase H1 and H2. RNase H1 which can recognize the RNA:DNA hybrid and cleave the RNA moiety, thus, removing the RNA from the loop (Parajuli et al, 2017). RNase H2 can also act as a ribonuclease excision repair and removes mis-incorporated ribonucleotides from DNA (Hyjek et al, 2019).

Complementary to these exonucleases, there are also helicases that can untangle the RNA:DNA hybrids by unwinding these structures, examples of these helicases are DExH-Box Helicase 9 (DHX9) (Chakraborty et al, 2011), aquarius (AQR) (Paulsen et al, 2009), senataxin (SETX) (Cohen et al, 2018; Hasanova et al, 2023) and an ATP-dependent DNA helicase PIF1 (Zhou et al, 2014).

8. Detection methods of R-loops

Over the years in the field of molecular biology, the RNA:DNA hybrids (or R-loops, both terms have been used interchangeably in the R-loop biology field) have attracted significant scientific interest as they can act both as regulators and as threats for genomic stability. Therefore, different techniques have been proposed that enable R-loop characterization, including techniques that allow to visualize and precipitate R-loops. The most widespread way to map R-loops is by using the monoclonal antibody S9.6 which recognizes the structure of RNA:DNA hybrids. The antibody can be coupled with next generation sequencing (DNA:RNA Immunoprecipitation sequencing-DRIP-Seq) in order to identify R-loops genome-wide, or used for immunofluorescence experiments to visualize R-loops in the nuclei of single cells (Skourti-Stathaki 2022). Even though the above methods had been used widely in the field, it was noted that the S9.6 antibody also has affinity to double stranded RNA, which can lead to false positive signals especially during immunofluorescence experiments (Hortono et al, 2018). Recently, new tools that leverage the ability of the RNase H to recognize and bind RNA:DNA hybrids have been proposed. As mentioned above, the RNase H recognizes R-loops and cleaves the RNA strand and, in recent works, researchers have taken advantage of this specific recognition to map R-loops (Chen et al, 2017; Crossley et al, 2021; Crossley et al, 2023). It is also important to note that RNase H1 is a monomeric nuclease making its manipulation easier. By mutating the RNaseH1 in its catalytic domain (D210N), the enzyme still retains the ability to bind R-loops without cleaving the RNA moiety, thus leaving these trihybrid nucleic acid structures intact. Some of the methods that have been described include the use of the catalytically inactive RNase H1 fused with GFP, as this tool enables the mapping of R-loops using immunofluorescence (Crossley et al, 2021). The GFP-dRNH1 was compared to the S9.6 antibody and was shown to bind specifically to R-loops and not dsRNA (Crossley et al, 2021). Additionally, MapR was described as an antibody-free method for R-loop identification, especially at the enhancer regions. MapR takes advantage of the ability of RNase H to bind R-loops and is used to guide micrococcal nuclease (MNase) on the site of the R-loops. The R-loops are then cleaved by MNase and sequenced. This tool enables us to study the abundance of R-loops especially at the promoter and enhancer regions (Yan et al, 2019). Another significant technique that enables the study of the R-loop proximity interactome was presented in the field in 2021 by Mosler and colleagues, in a study where they presented RNA–DNA Proximity Proteomics (RDProx), a tool that uses hybrid-binding

domain (HBD) of RNaseH1 and a variant of ascorbate peroxidase (APEX2) to identify R-loop interactors via Mass spectrometry.

One interesting tool was presented by Chen et al, 2017 and it is the RChIP (Chen et al, 2017). Again, this technique uses the catalytically inactive RNaseH1 (D210N) which is fused with a V5 tag to map R-loops genome wide and has been characterized as one robust R-loop detection method (Zhang et al, 2022). However, this tool has the limitation of the establishment of a stable cell line expressing the D210N-V5 RNase H.

9. m6A and R-loops

In 2020, Abakir et al, illustrated that m6A can be deposited on the RNA strand of R-loops in pluripotent stem cells and that this mark can be recognized by m6A reader, YTHDF2. Intriguingly, it was presented that the loss of this reader leads to accumulation of R-loops, hence, accumulation of DNA breaks, proposing that m6A, through YTHDF2, safeguards genomic stability by activating downstream events of R-loop resolution. On the contrary, another study in cancer cell lines showed that m6A promotes the formation and stability of R-loops through the binding of YTHDC1 (Zhang et al, 2020). The so far published data show a relationship between m6A and R-loops, however, the effects of m6A on DNA:RNA hybrids and the pathways that are activated by m6A deposition have not been elucidated yet. The regulatory effects of m6A on R-loops might depend on the cell type, cell cycle and/or the R-loop trigger event.

Aim of study

Aim of study

N6-Methyladenosine (m6A) has regulatory roles in many biological processes including splicing, mRNA export and RNA destabilization. Recent studies try to unravel possible roles of m6A upon stress factors, including factors that induce DNA damage. However, the functional link between m6A and DNA damage response and the mechanisms behind it have not been fully elucidated yet. In this project, I mainly focused on uncovering regulatory roles of m6A in RNA:DNA hybrids. Based on previous studies, which showed that the RNA moiety of R-loops can be m6A decorated, in this thesis I aim to address the potential role of N6-Methyladenosine (m6A) a) in response to DNA damage and b) in R-loop regulation. To induce DNA breaks and R-loop formation I used UVC radiation (20J/m^2).

To approach this question, I focused on a nuclear m6A reader, YTHDC1, which has been shown to on double strand breaks (Zhang et al, 2020). It is noteworthy that this reader regulates responses in heat stress (Timcheva et al, 2022), suggesting a potential role in the DNA damage response as well.

To address these questions, new tools were generated and established. Firstly, I established stable cell lines expressing inactive RNase H, a useful tool for specific R-loop precipitation and visualization (inspired by the protocol of Chen et al, 2017). Using a catalytically inactive RNase H that has a mutation in its catalytic center (-D210N-) and has the ability to recognize R-loops without cleaving the RNA moiety, the detection of RNA:DNA hybrids is feasible (Cerritelli et al, 2022). This tool will be used mostly to visualize R-loops using immunofluorescence. Also, to better dissect the role of YTHDC1, an acute and specific protein degradation tag (dTAG) system was designed and established in the colon cancer cell line HCT116. This system will be used to study the direct effects of the loss of YTHDC1 in human cells, both at steady state and in response to UVC-mediated DNA damage.

Materials & Methods

Materials

1. Plasmids

Plasmids	
Name	**Addgene ID**
pAW62.YY1.FKBP.knock-in.mCherry	#104370
pAW63.YY1.FKBP.knock-in.BFP	#104371
pSpCas9(BB)-2A-Puro (PX459) V2.0	#62988
ppyCAG_RNaseH1_D210N	#111904
ppyCAG_RNaseH1_WKKD	#111905

2. Oligonucleotides used for the dTAG generation

Primers for isolation of vector fragments	
Name	**Sequence**
Backbone_For	GGATCCCCGGGTACCGAG
FKBP_Rev	TCAAGGAAAAACCAGACATCAACC
pUC19dTAG_Rev	GGATCCGGAGGAGTGCAG
mCherry_For	TTACTTGTACAGCTCGTCCATGC
mTagBFP2_For	TTAATTAAGCTTGTGCCCCAGTTTG

Oligonucleotides for isolation of YTHDC1 homology arms	
Name	**Sequence**
5' hom.arm Forward (Vector)	caagtgggttgatgtctggtttttccttgaATGATTTCCTTCGTCGCACACAAG
5' hom.arm Reverse (FKBP)	cactcctccggatccTCTTCTATATCGACCTCTCTCCC
3' hom.arm Forward (mCherry)	gagctgtacaagtaaTGGGCTTTTGGAAGCACTGATTG
3' hom.arm Forward (BFP)	cacaagcttaattaaTGGGCTTTTGGAAGCACTGATTG
3' hom.arm Reverse (Vector)	tgaattcgagctcggtacccggggatccTAAGAAAAAGATACAAAAGATAACCGTCAA

gRNAs for dTAG knock in for YTHDC1	
Name	**Sequence**
dc1 dTAG gRNA 1 oligo F	CACCAGTGTGATCGAGACAGAGACCGGTT
dc1 dTAG gRNA 1 oligo R	AAACAACCGGTCTCTGTCTCGATCACACT
dc1 dTAG gRNA 2 oligo F	CACCAGGAGGTCGATATAGAAGATAAGTT
dc1 dTAG gRNA 2 oligo R	AAACAACTTATCTTCTATATCGACCTCCT
dc1 dTAG gRNA 3 oligo F	CACCAGAGGTCGATATAGAAGATAATGTT
dc1 dTAG gRNA 3 oligo R	AAACAACATTATCTTCTATATCGACCTCT

Primers for sequencing verification	
Backbone_Forward	GAGGAAGGAGACACACTC
FKBP_Reverse	TGTCCCGGGAGGAATCAAC
FKBP_Forward	CAACAAACTTCTCTCTGCTGAAA
pUC19dTAG_Reverse	CTTCGCTATTACGCCAGCTG
pUC19dTAG_Reverse 2	AAGGCGATTAAGTTGGGTAAC
mCherry_Forward	GGCGCCTACAACGTCAAC
MTagBFP2_Forward	CTATGTGGACTACAGACTGG

3. Chemicals

Reagent	Company	Product Number
1Kb Plus DNA Ladder	NEB	N3200S
Acetic Acid	Honeywell	33209
Acrylamide/Bis Solution	Serva	10687.01
Agar-agar	Sigma	05040
Agarose	Biorad	161-3102
Bromophenol Blue	Sigma	B0126
Chloroform	Merck	1.02445.1000
Dimethyl sulfoxide	Sigma	D4540
Dithiothreitol	NEB	B1034A
dNTPs	Applied Biosystems	4367381
Dulbecco's Modified Eagle's Medium	Gibco	41966
Ethanol Abs	Merck	1.00983.2511
Ethylenediaminetetraacetic Acid	AppliChem	A5097
FBS	Gibco	10437-028
Glycerol	AppliChem	141339.1211
Glycine	AppliChem	A1067
Glycogen-carrier	Thermo Fisher Scientific	R0561
Hydrochloric Acid	Supelco	1.00317
Isopropanol	Merck	818766
Methanol	AppliChem	131091.1212
NP-40	Thermo Fisher Scientific	85125
PBS pellets	Gibco	18912-014
Phenol, acidic	Sigma	P4682
Potassium Chloride	Merck	1.04936
Prestain protein markers	ProteinTech	PL00001
RNasin Ribonuclease Inhibitor	Invitrogen	AM2696
SDS	Sigma	L5750
Sodium Chloride	Supelco	1.06404
TEMED	Sigma	1.10733.0100
Tris ultrapure	AppliChem	A1086
Triton-X-100	Sigma	9036-19-5
Trypsin EDTA 10X	Gibco	15400-054
Tryptone	Sigma	T.9410
Tween-20	Merck	8.22184
Xylene Cyanol	Sigma	X4126
Yeast Extract	Sigma	70161

4. Enzymes-Proteins

Reagent	Company	Product Number
Protein inhibitor cocktail	Sigma	58820-2 TAB
Phusion-HF DNA Polymerase	NEB	M0530S
T4 Polynucleotide Kinase	NEB	M0201S
Quick CIP	NEB	M0525S
T4 DNA Ligase	NEB	M0202S
Taq DNA Polymerase	NEB	M0273S
BbsI-HF	NEB	10119525
Q5-HF DNA Polymerase	NEB	M0491
HindIII-HF	NEB	R3104S
ApoI-HF	NEB	R3566S
AccI	NEB	R0161S
EarI	NEB	R0528S

5. Antibodies

Antibody	Company	Product Number
Anti-V5 Polyclonal antibody	ProteinTech	14440-1-AP
Anti-YTHDC1 Polyclonal antibody	ProteinTech	14392-1-AP
Anti- GAPDH Monoclonal antibody	ProteinTech	60004-1-Ig
Anti-γH2Ax Monoclonal antibody	Merck	05-636
Goat Anti-Mouse IgG	Jakson	115-035-146
Goat Anti-Rabbit IgG	Millipore	AP132
Alexa Fluor 488 anti-Rabbit IgG	Millipore	SAB4600234
Alexa Fluor 568 anti-Mouse IgG	Millipore	SAB4600312

6. Kits

Kit Name	Company	Product Number
Monarch® Plasmid Miniprep Kit	NEB	T1010S
Monarch® DNA Gel Extraction Kit	NEB	T1020S
Polyplus Transfection jetPEI	Polyplus	101000053/1ML

Methods

1. Cells and cell culture

In this study two different cell lines were used. HCT116 (human colon cancer cell line) and HEK293T (human embryonic kidney cells, which contain the SV40 large T antigen, that enables them to produce recombinant proteins within plasmid vectors containing the SV40 promoter).All cells were grown at 37 °C in a 5% CO2 incubator and were cultured with Dulbecco's Modified Eagle Medium (DMEM) supplemented with 10% Fetal Bovine Serum (FBS). When the cells reached a confluency of ~80% they were sub cultured

For long term storage, the cells were cryopreserved. When they reached a confluency of ~80%, they were trypsinized and transferred in a freezing medium (DMEM with FBS and 10% DMSO). The cells were aliquoted in cryovials and transferred in a freezing container with isopropanol at -80°C. The isopropanol is used to ensure that the cooling rate of cells is -1°C per minute. The vials were stored at -80°C and then transferred into liquid nitrogen.

For the DNA damage experiments, on the day of the experiments, the medium was removed from the plates and the cells were treated with UVC radiation ($^{20}J/ms^2$) and collected or fixed at specific time points after the treatment.

2. Cells' transfection

The plasmids used in this study were delivered to the cells through the process of transfection. For all the experiments that required the delivery of plasmid DNA, the transfection was performed using the jetPEI DNA transfection reagent, following the manufacturer's instructions. Prior to the day of the transfection, the cells were seeded in plates at the appropriate confluency (40-50%). On the day of the experiment, the jetPEI reagent was mixed with the DNA and incubated for 30 minutes at RT, in order to form the jetPEI/DNA complexes. Then the mix was transferred in the medium of cells. For the creation of the dTAG HCT116 cell line (described in paragraph 2.7), to increase the chances of successful knock in of the constructs, the transfection was repeated 48h post the first transfection hit.

3. Protein extraction

Proteins from the samples were extracted using RIPA buffer (50mM Tris-Cl pH 8, 150mM NaCl, 1% NP-40,0.1% SDS,0.5% sodium deoxycholate, PMSF, Protease inhibitors). Briefly, depending on the cell pellet 5-fold volumes of RIPA buffer were added in the cell pellet, after resuspending the pellet, 5 freezing - thawing cycles followed. In each cycle the lysate was first transferred to liquid nitrogen (-196°C) and then directly to 37°C. After the cycles were completed, the lysate was sonicated for 5 cycles x 5 seconds. The final step is centrifugation at 11,000rpm for 20 minutes to discard debris. The supernatant was kept for downstream analysis.

4. Western blotting

Western blotting was used in order to determine the levels of specific proteins. The ingredients used for the Western blotting are presented in the table below.

Gels for Western blot setup

Gel preparation		
Ingredients	**10 % Running gel**	**Stacking gel**
H_2O	4.1 ml	3.05ml
SDS	50ul	50ul
Acrylamide	3.3ml	0.65ml
TEMED	10ul	5ul
APS	100ul	50ul
1.5M Tris-HCL,pH 8.8/0.5M Tris-HCL pH 6.8 respectively	2.5ml	1.25ml

Buffers for SDS-PAGE

1x Running Buffer	
Ingredients	**Final concentration**
Tris-HCL	25mM
Glycine	200mM
SDS	0.1%(w/v)
Water	up to final volume

1x Transfer Buffer	
Ingredients	**Final concentration**
Tris	25mM
Glycine	192mM
Methanol pH 8.3	20% (v/v)
SDS	0.1% (v/v)
H_2O	up to final volume

Equal amounts of proteins were loaded on 10% (v/v) polyacrylamide gel, after the gel electrophoresis the proteins were transferred onto nitrocellulose membranes using the Mini Trans-Blot Cell (BioRad). The membranes were then blocked in a blocking buffer (1% v/v milk in tris-buffered saline with 0.01% Tween 20- TBST) for an hour at Room Temperature (RT). After the 60 minutes incubation, the primary antibodies were added to the membranes and were incubated overnight at 4℃. The next day, the membranes are washed with TBST and incubated with the secondary antibodies for 60 minutes at RT. Then the membranes are washed again with TBS-T, and the protein levels are detected after addition of the ECL (SuperSignal West Pico PLUS, Chemiluminescent Substrate, ThermoScientific), using the Chemidoc Imager (Bio-Rad).

5. Immunoprecipitation (IP)

Immunoprecipitation experiments were carried out in order to enrich the pull down of YTHDC1. In total 400 ug of total protein were used and precleared for 30 minutes at room temperature (RT) on rotating platform with 20 ul of IgA/IgG magnetic beads. During this incubation the appropriate amount of anti-IgG and anti-YTHDC1 antibodies were incubated with magnetic beads that would be used to isolate the protein of interest. The mix of antibodies and beads were incubated for 1h at RT on rotating platform. After the preclearing process was over, the protein lysate was kept. The precleared beads were washed 3 times with RIPA, loaded with 5x Laemmli and heated to 95℃. After the 1 h incubation of the antibodies with the beads, the beads were washed 3 times with RIPA and the lysate was transferred to the Antibodies-beads mixture and incubated for 2 h at RT on rotating platform. In parallel, 5% of the total protein was kept for input control, loaded with 5x Laemmli and heated to 95℃. After the 2 h incubation, 60 ul of the supernatant (flowthrough) was kept, loaded with 5x Laemmli and heated to 95℃. The beads were washed 3 times with RIPA and the antibodies and proteins were eluted from the beads with 2x Laemmli and 10-minute incubation at 80℃. The results of the procedure were analyzed using Western blotting.

6. Immunofluorescence (IF)

Before the day of the experiment the HCT116 cells were seeded in wells with slides. The cells were fixed using 4% (v/v) paraformaldehyde (PFA) for 15 min at 37 °C, washed once with PBS and then permeabilized using 0,3 % Triton- PBS for 3 min at RT. The slides were blocked for 1h at RT with 3% BSA-PBS. For the immunostaining procedure, primary antibodies (anti-YTHDC1 at a 1:500 dilution and gamma H2Ax at a 1:500 dilution) were applied and allowed to incubate overnight at 4°C. The following day, the slides were subjected to a series of three 5-minute washes with PBS. Secondary antibodies diluted in a 3% BSA-PBS solution were then added, and the slides were incubated at RT for 2 hours. The samples were washed 3 times for 5min with 1x PBS and then DAPI (1ul in 1ml) was added for 5 min incubation in the dark. The slides were washed 3 times with 1xPBS, and mounting took place using the Mowiol as a medium. After that step, the samples can be stored at 4°C or long term stored at -20°C. All samples were scanned using SP8 inverted confocal microscope (Leica). The figures were further processed using the platform Fiji (Schnindelin et al, 2012).

7. Acute protein depletion using the dTAG system

One of the most prevalent ways to study the function and biological role of a protein of interest is its depletion of the cells. One common way is to remove the gene via the CRISPR methodology, but this process is time consuming, often generates off-target effects and is not applicable to all genes. Thus, the degradation at the protein level is an effective tool to study the direct effects of a protein, in a time saving and reversible manner.

A widely recognized approach for achieving precise protein depletion is the degradation tag system (dTAG), which facilitates the targeted and specific removal of proteins through chemical means. The dTAG system pairs the degradation of $FKBP12^{F36V}$ with expression of $FKBP12^{F36V}$ in frame with the protein of interest. By CRISPR-mediated locus specific knock

in, the study of the direct effects of the protein depletion is feasible. The main principle of the system is as follows, the protein of interest is coupled with the FKBP12^{F36V}, which by adding the molecule degrader dTAG-47, is paired with the E3 ubiquitin ligase cereblon (CRBN). The FKBP12^{F36V} tagged protein is then degraded by triggering ubiquitin-mediated proteolysis (Figure 9).

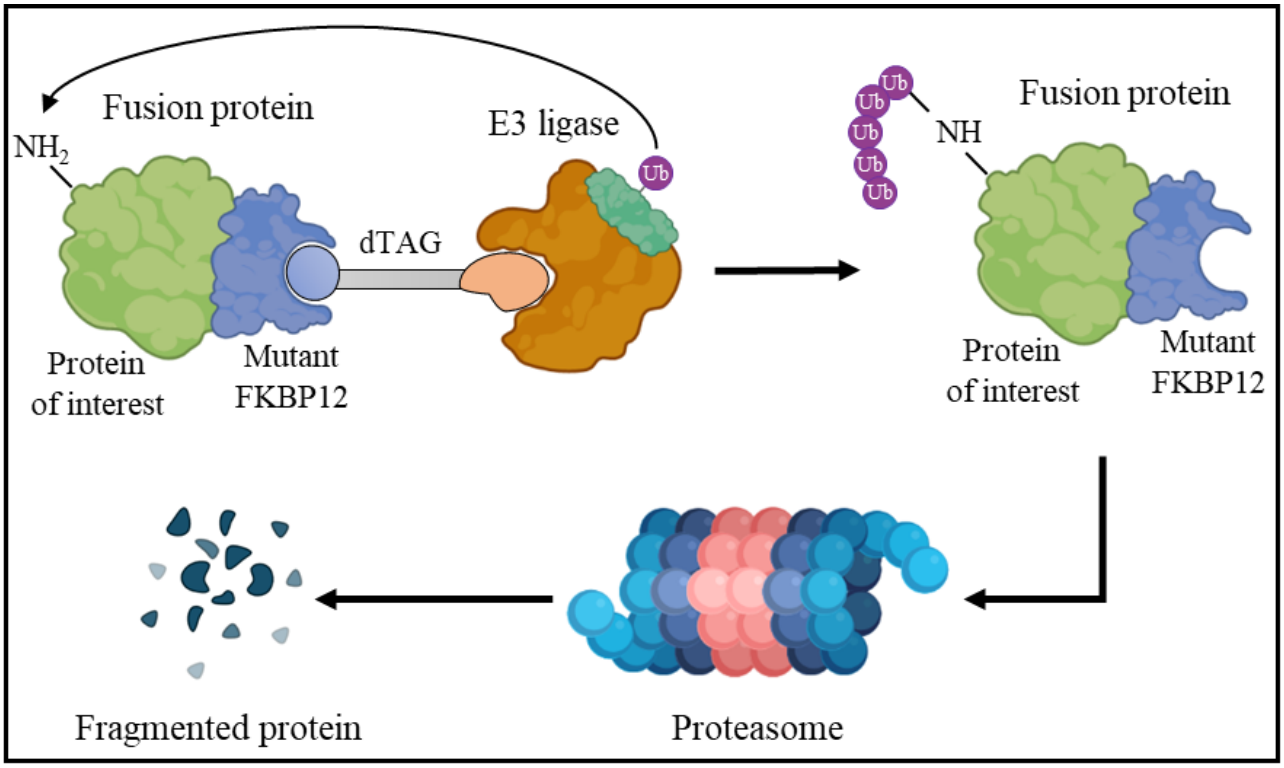


Figure 9. Schematic representation of the action mechanism of the dTAG depletion system. The protein of interest is fused with the mutant FKBP12 (FKBP12F36V), which in the presence of dTAG-47, is coupled to the cereblon (CRBN) ubiquitin protein ligase. After its ubiquitination the protein of interest is degraded by the proteasome.

7.1 Generation of constructs

For the generation of the construct's homology directed repair donor plasmids were used to ensure that the mutant FKBP12 will be inserted in frame and site specific to the locus of interest. For that reason, the design of specific oligos were used to clone target specific 5' and 3' homology arms to flank the mutant FKBP12. Fluorescent proteins were also fused downstream of the FKBP12^{F36V} to create the final construct which contained the FKBP12F36V-2xHA-P2A-fluorescent protein tag. Fluorescent proteins are important in order to check and select the cells that had successful insertion of the construct after the CRISPR mediated knock in. In this study, two fluorescent proteins were used, mCherry and Blue Fluorescent Protein (BFP), because the protein of interest is expressed by two alleles, therefore, two fluorescent proteins are needed to select cells that have undergone homozygous editing. For each gene, 2 constructs were generated, one containing the mCherry and one containing the BFP cassette (Figure 10).

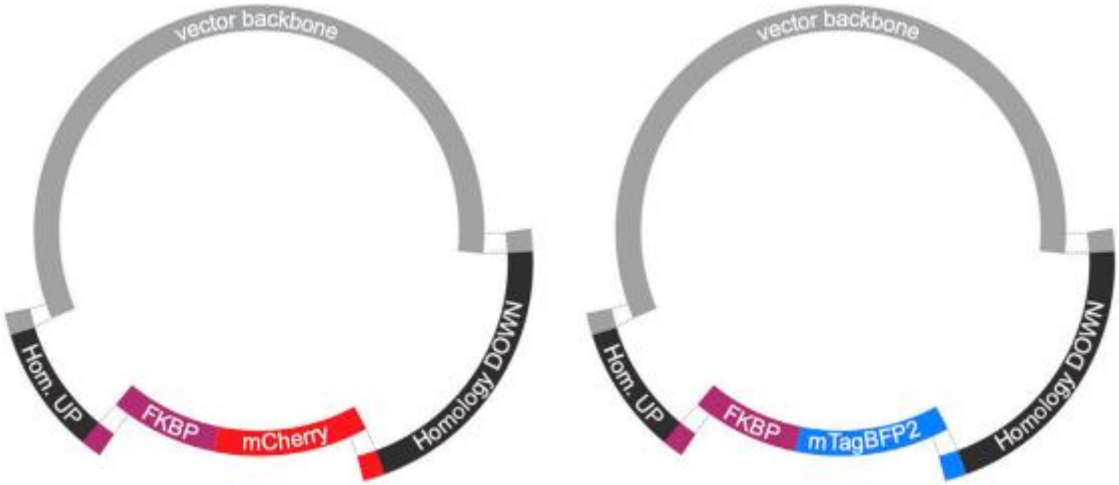


Figure 10. Cloning strategy for homology targeting vectors. In total two constructs were generated, each containing the 5' and 3 homology arms, the mutant FKBP12, the vector backbone and one fluorescent protein. In this study, two fluorescent proteins were used, mCherry and BFP. (Guarnaccia et al, 2021).

7.2 Genomic DNA isolation

The 5' and 3' homology arms were PCR amplified from genomic DNA of HCT116 cells. For the genomic DNA isolation, the protocol used was from Koh et al, 2017 and briefly the steps are the following: The cells were collected and pelleted through centrifugation, then an appropriate amount (depending on the cell pellet) of cell lysis buffer was added to the pellet. The lysis buffer ingredients and the final concentrations are: NaCl 100mM, Tris-HCl pH=8 10mM, EDTA pH=8 25mM, SDS 0.5%, Proteinase K 0.1mg/ml. The mix was incubated overnight at 50°C. The next day, the organic extraction was performed by adding an equal volume of phenol extraction buffer (Neutral phenol: SEVAG 1:1), to separate DNA from other components of the cells. The next step of the process is the precipitation of the DNA using 0.5 volume of 7.5M ammonium acetate and 2 volumes of 100 % ethanol. Finally, the DNA pellet is washed with 70% ethanol.

In total, 4 fragments had to be amplified via PCR, the vector backbone, the 5' homology arm, the 3' homology arm, the BFP or mCherry cassette. The PCR components and conditions are provided in the following tables.

PCR reaction setup	
Component	**Final concentration**
5x Q5 reaction buffer	1x
10mM dNTPs	200uM
Forward primer	0.5uM
Reverse primer	0.5uM
Template DNA	variable
Q5 polymerase	0.02U/ul
Nuclease free water	To final volume

PCR conditions for BFP and mCherry cassettes

PCR conditions					
	Cassettes (mCherry & BFP)		**Vector backbone**		
	Temperature	Time	Temperature	Time	Cycles
Initial denaturation	98ºC	30''	98ºC	30''	1
Denaturation	98ºC	10''	98ºC	15''	35
Annealing	**66ºC**	15''	**66ºC**	15''	
Extension	72ºC	30''	72ºC	2'	
Final extension	72ºC	2'	72ºC	2'	1
Hold	4ºC	∞	4ºC	∞	

PCR conditions for the isolation of the homology arms of YTHDC1

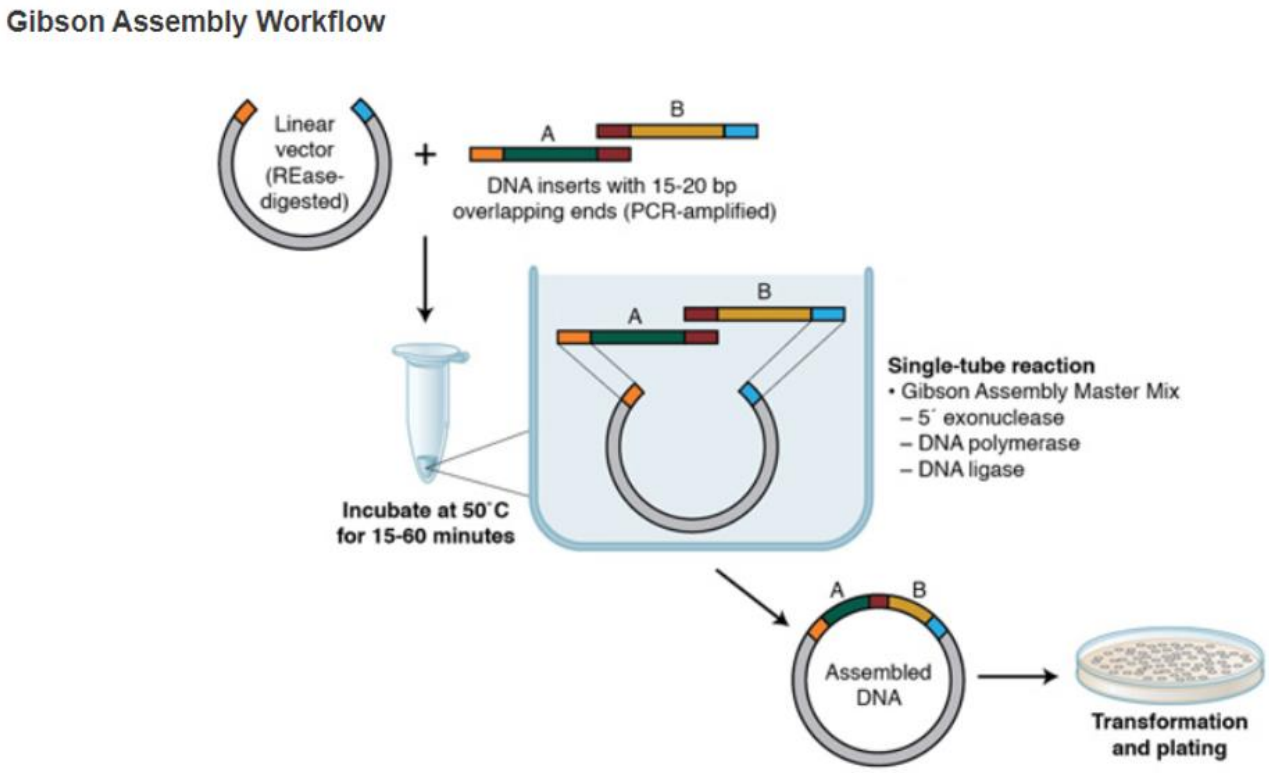
	5' homology arm YTHDC1		**3' homology arm YTHDC1**		
	Temperature	Time	Temperature	Time	Cycles
Initial denaturation	98ºC	30''	98ºC	30''	1
Denaturation	98ºC	15''	98ºC	15''	5
Annealing (without overhangs)	**65ºC**	15''	**63ºC**	15''	
Extension	72ºC	30''	72ºC	30''	
Denaturation	98ºC	15''	98ºC	15''	35
Annealing (with overhangs)	**72ºC**	15''	**72ºC**	15''	
Extension	72ºC	30''	72ºC	2'	
Final extension	72ºC	2'	72ºC	2'	1
Hold	4ºC	∞	4ºC	∞	

7.3 Gel extraction

After the PCR amplification the products were loaded on agarose gels and then extracted and purified using Monarch® DNA Gel Extraction Kit from New England Biolabs (NEB) following the manufacturer's instructions. Briefly, the steps are the following, the DNA fragment is excised and dissolved using the Dissolving Buffer. The sample was loaded on binding columns. The DNA is washed and eluted using the elution buffer.

7.4 Gibson assembly

Gibson assembly is an exonuclease-dependent method to assemble DNA fragments in correct order. The reaction is carried out under isothermal conditions using three enzymes which were all included in the Gibson assembly reaction mix. The 5' exonuclease generates long overhangs, the polymerase which fills in the gaps and a ligase which seals the nicks that have been produced during the previous enzymatic reactions. The following figure depicts the principle of Gibson assembly.



Source: https://www.neb.com/en/applications/cloning-and-synthetic-biology/dna-assembly-and-cloning/gibson-assembly

Figure 11. Gibson assembly workflow. The Gibson assembly employs three enzymatic reactions in a single tube. 5' exonuclease generates overhangs, DNA polymerase fills the gaps, and a ligase seals all nicks that have occurred. The figure was retrieved from NEB's website.

7.5 Bacterial transformation and plasmid isolation

For all the transformations and the plasmids amplification competent DHα *E.coli* cells were used. For the bacterial transformation, competent cells were transferred from -80°C to 4℃ for 5 minutes. After the cells are thawed, the appropriate amount of plasmid DNA (10pg-100 ng) was added to the cells and followed by a 20-minute incubation on ice. In order to achieve efficient transformation, the cells were transferred to 42℃ for 45 seconds for brief heat shock and then on ice for 2 minutes. 500 ul of pre-warmed LB was added to the cell-plasmid mixture and incubated at 37℃ for 1 hour while shaking. The cells were then plated on LB-plates with appropriate antibiotics for selection and were grown overnight at 37°C. The plasmids were isolated using the Monarch® Plasmid Miniprep Kit from NEB, following the manufacturer's instructions. In short, the cells were pelleted and lysed using the Plasmid Lysis Buffer. Then the Plasmid Neutralization Buffer was added, and the samples were centrifuged. The supernatant was transferred to a DNA binding column to bind the plasmid DNA and washed using the washing buffers. The plasmid DNA was eluted using the DNA elution buffer.

7.6 Validation of constructs

The validation of the constructs was performed using three different methods: restriction digestion analysis, PCR and Sanger sequencing. The first step was to linearize the plasmids and keep only the ones of the expected size. Then the plasmids were digested using restriction enzymes that gave bands of specific size depending on whether all fragments are inserted in the construct. An additional verification was performed with PCR and agarose gel analysis of the products. The best candidates for the correctly assembled fragments were further validated using Sanger sequencing.

7.7 Generation of gRNA constructs

To introduce the constructs into the cells, CRISPR-Cas9 technology was employed. To achieve this, custom guide RNAs were created. The steps for generating these gRNA constructs are as follows:

a) *Preparation of vector*

In order to achieve the knock in of the construct inside the cells, the CRISPR-Cas9 technology was used. For that reason, specific guide RNAs were inserted to the pSpCas9(BB)-2A-Puro (PX459) V2.0 plasmid. To ensure that the guide RNAs will be inserted into the vector, the plasmid was first linearized by restriction digestion with BsgI for an hour at 37℃. The linear plasmids were loaded on agarose gel and gel extracted as described above. To ensure that the vector won't re-circularize, a dephosphorylation process took place using the Alkaline Phosphatase Calf Intestinal (CIP) for 10 minutes at 37℃ with the following setup. The enzyme was inactivated for 2 minutes at 80°C. The dephosphorylated plasmids were then loaded on agarose gel, extracted, and purified.

Phosphorylation reaction setup

Dephosphorylation process	
Component	**Volume**
DNA	1 pmol of DNA ends
rCutSmart buffer (NEB)	2ul
Quick CIP	1ul
Nuclease free water	up to final volume (in this case 20ul)

b) *Selection of design tool and gRNA sequence*

The guide RNAs for the knock in were selected to fulfill the following criteria: The targeting region must be immediately 5' of a NGG PAM sequence and the cut sites should be within 30bp of the stop codon of the genetic locus of interest. A range of 100 bp flanking the stop codon was selected and imported on the online guide tool

CRISPOR (Concordet and Haeussler,2018). Three pairs of guide RNAs were selected having high score of efficiency and minimum chances of off target effects.

c) *Preparation of gRNAs*

The single-stranded guide RNAs were converted into duplexes by combining equal quantities of the two complementary guide RNAs. This mixture was then heated to 95℃ for 2 minutes, followed by a gradual cooling to room temperature (RT). This process guarantees the annealing of the complementary RNA strands, as heat disrupts existing hydrogen bonds, and cooling enables the formation of new bonds between the sequences. Phosphoryl groups were added to the ends of the guide RNA duplexes to ensure their insertion to the dephosphorylated vector. The phosphorylation occurred using the T4 PNK with the following set up. The reaction mix was incubated at 37°C for 30 minutes and the enzyme was inactivated after 20-minute incubation at 60°C.

Phosphorylation reaction setup

Phosphorylation process	
Component	**Volume**
DNA	up to 300 pmol of 5'termini
T4 PNK Reaction Buffer (10x)	5ul (1x)
ATP (10mM)	5ul(1mM)
T4 PNK	1ul (10 units)
Nuclease free water	up to final volume (in this case 50ul)

d) *Ligation of the gRNAs inside the vector*

The guide RNAs (insert DNA) were integrated into the vector backbone (vector DNA) via the ligation process. The details of this procedure are outlined in the provided table. The ligation was performed using T4 Ligase and left to incubate at 16℃ overnight. Subsequently, the ligated plasmids were introduced into DHα *E.coli* bacteria through transformation.

Ligation reaction setup

Ligation process	
Component	**Volume**
T4 DNA Ligase Buffer (10X)	2ul
Vector DNA	0.020 pmol
Insert DNA	0.060 pmol
T4 DNA Ligase	1ul
Nuclease free water	up to final volume (in this case 20ul)

e) *Transformation*

The transformation protocol can be found in paragraph 2.5.4.

f) Colony PCR and miniprep

To ensure that the bacterial colonies contain the full desired insert, colony PCR was performed. The set up is presented in the table below.

Colony PCR	
Component	**Final concentration**
10x Standard *Taq* Reaction Buffer	1x
10mM dNTPs	200uM
10uM Forward primer	0.2uM
10uM Reverse primer	0.2uM
Template DNA	<1000 ng
Taq DNA polymerase	1.25U/50 ul PCR
Nuclease free water	up to final reaction volume

g) Sanger sequencing

To verify the sequence of the cloned plasmids, Sanger sequencing was performed. For the sequencing the appropriate concentration of sequencing primers and plasmid DNA were mixed, according to the company's instructions.

7.8 Fluorescent activated cell sorting (FACS)

Fluorescence-activated cell sorting (FACS) is a method for sorting cells based on their fluorescence and physical traits. To isolate the cell population with dual integration of the constructs, FACS was used. To ensure effective degradation of the target proteins, the constructs needed to be inserted into both genomic loci. As a result, two constructs were created: one carrying the mCherry cassette and another with the BFP cassette. Cells exhibiting both fluorescent proteins were then chosen.

8. Bioinformatic analysis (RNA-seq analysis)

The YTHDC1 KD raw data were retrieved from GSE74397 from the study of Roundtree et al,2017. The fastq files were aligned to the genome of reference (GRCh38), human assembly hg38 with annotation from Gencode version 41 using STAR (Dobin et al, 2013). The .sam files were converted to .bam files using samtools (Li et al, 2009) and then to .bed files using bedtools (Quinlan et al, 2012). DESeq2 (Love et al, 2014) was used to normalize reads and analyze the gene expression between three biological replicates and two conditions (control and YTHDC1 KD). The volcano plot was performed using the "ggplot2" package in R (version 4.3.1) and using only the results of DESeq2 with p-adjusted value <0.001.

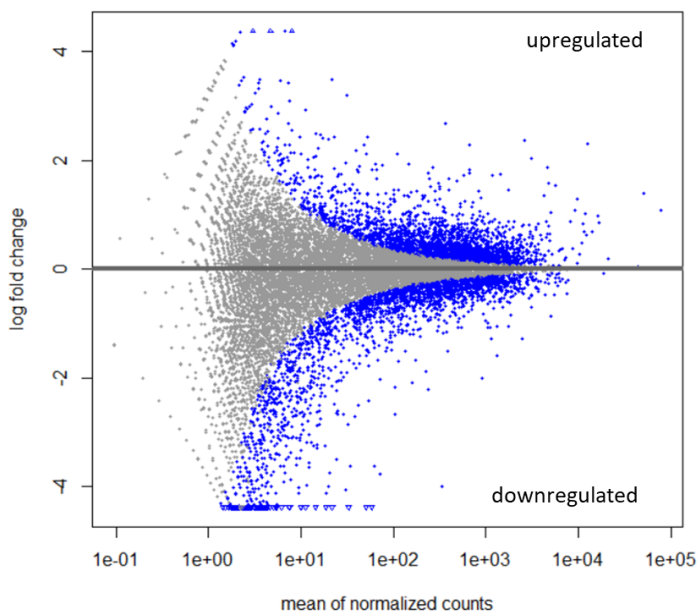
Results

1. Differentially expression analysis of HeLa YTHDC1 knockdown cells

Reanalysis of published dataset upon knockdown of the m6A reader, YTHDC1, was performed in order to get an insight into the genes whose expression is affected. The raw data come from HeLa cells and the differential expression analysis was performed using DESeq2. GEO accession number: GSE74397.

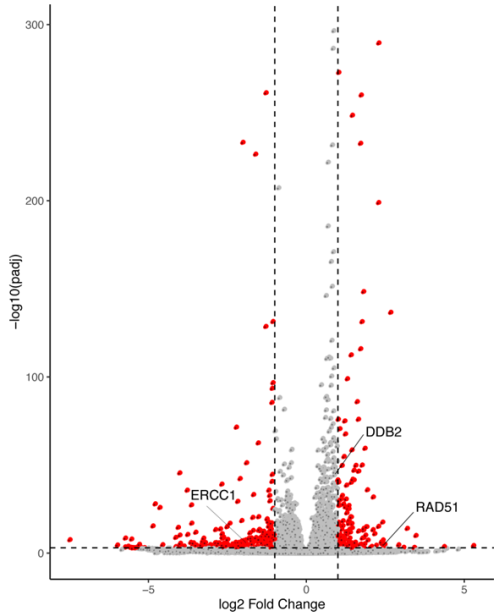
a)

MA plot after DESeq2, padj<0.05



b)

Volcano plot showing DEGs with padj<0.001



c)

gene_id	ensembl_id	log2FoldChange	padj	DNA repair pathway
RAD51	**ENSG00000051180**	**2.438887809**	**0.000908**	**HR**
ERCC2	ENSG00000104884	-0.449795271	0.000226	NER
ERCC1	**ENSG00000012061**	**-2.033632443**	**7.77E-07**	**NER**
DDB2	ENSG00000134574	0.89704676	3.93E-45	NER
RAD54L2	ENSG00000164080	0.288773277	1.38E-05	HR
ATR	ENSG00000175054	-0.407637283	8.56E-05	HR
DHX9	ENSG00000135829	0.346445854	2.28E-07	R-loop processing

Figure 12. Differential expression analysis upon YTHDC1 KD. a) MAplot after DESeq2 illustrates the genes that are differentially expressed (DEGs) in the knockdown compared to the untreated HeLa cells. The plot shows the upregulated (log 2-fold change > 0) and the downregulated genes (log 2-fold change < 0) in the knockdown condition. b) Volcano plot of the DEGs upon treatment, the significant threshold is padj < 0.001. The red dots show the genes whose expression is significantly changed whereas the gray dots show non-significant changes. Some of the genes whose expression is altered upon YTHDC1 KD and take part in DNA damage response, are labeled in the graph.c) Table with representative DNA damage responsive genes whose expression is significantly (padj < 0.001) altered in YTHDC1 KD cells.

2. The protein levels YTHDC1 are increased upon UVC radiation

To dissect the role of YTHDC1 in response to stress, UV mediated DNA damage was employed, as described below. Before the day of the experiment HCT116 cells were seeded in plates with coverslips and when they reached a confluency of ~70%, they were used for the UVC experiments. On the day of the experiment, the medium was removed and kept. The cells were then hit with 20J/m^2 of UVC, after the hit the medium was returned in the plates and returned to the incubator.

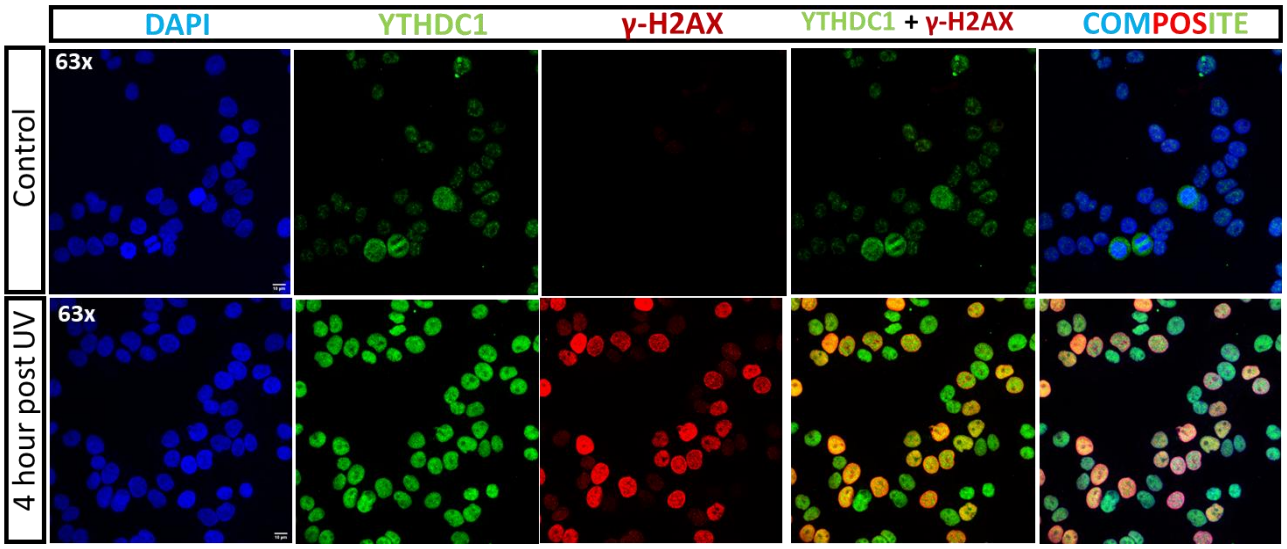


Figure 13. Immunofluorescence experiments showing the YTHDC1 and γH2Ax DNA damage marker. In the figure the DAPI staining for nuclei is shown in blue, the m6A reader is shown in green whereas with red the marker γH2Ax is shown. The upper panel shows the untreated with UVC radiation cells (control) and the lower shows the cells treated with UVC and fixed 4h post treatment. Scalebar is 10 um.

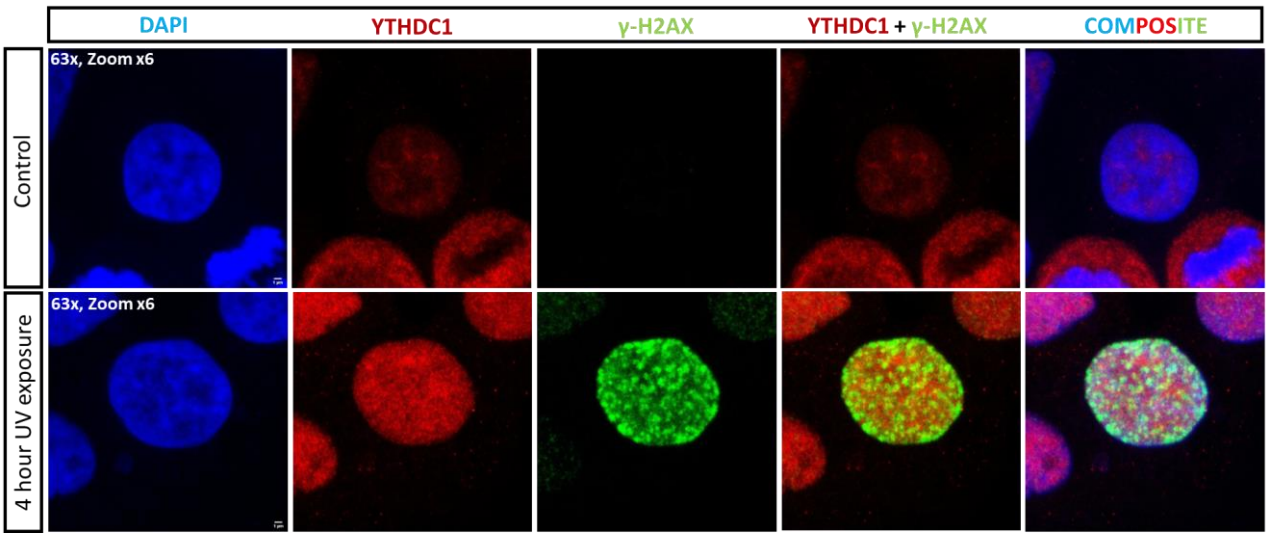


Figure 14. Immunofluorescence experiments showing the YTHDC1 and γH2Ax DNA damage marker (Zoom). In the figure the DAPI staining for nuclei is shown in blue, the m6A reader is shown in green whereas with red the marker γH2Ax is shown. The upper panel shows the untreated with UVC radiation cells (control) and the lower shows the cells treated with UVC and fixed 4h post treatment. Scalebar is 1 um.

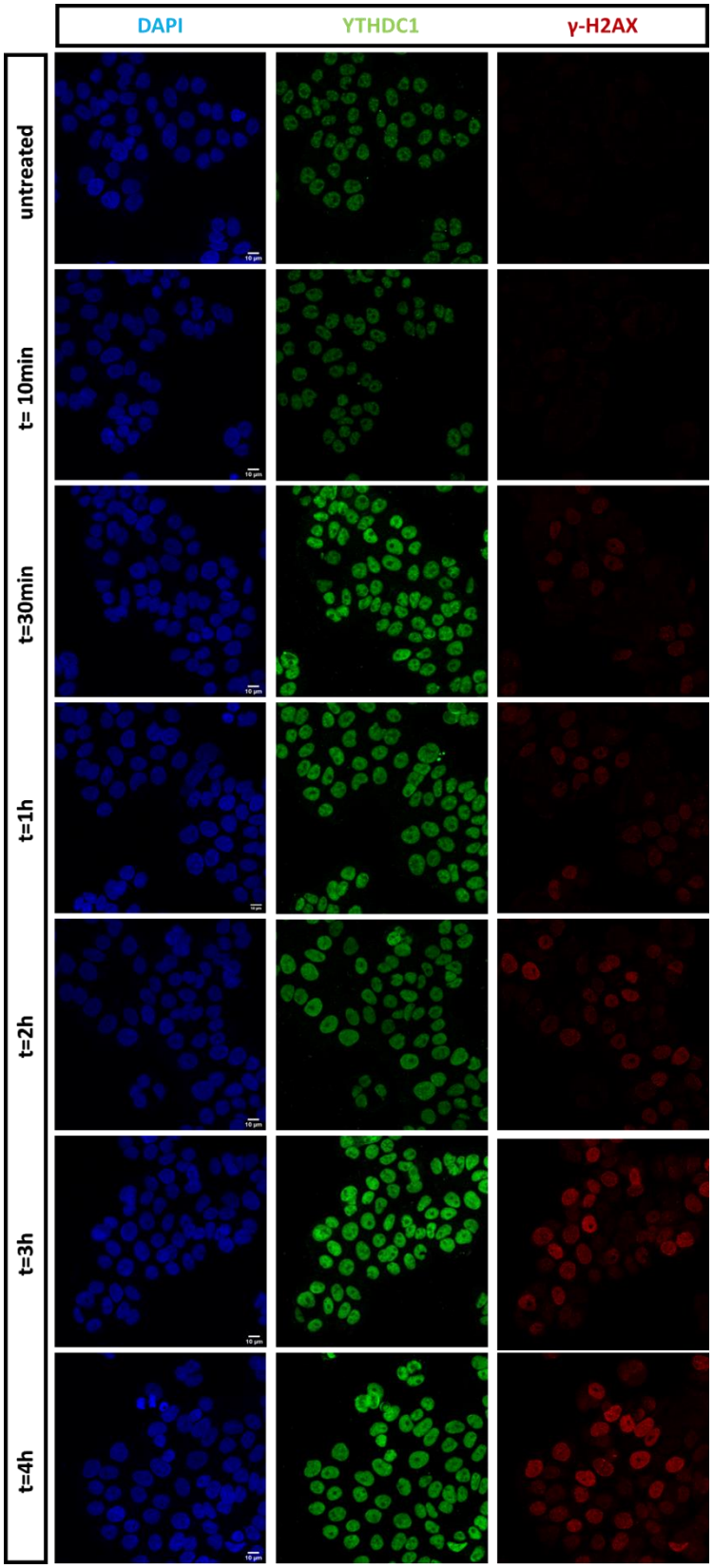


Figure 15. Immunofluorescence experiments showing the YTHDC1 and γH2Ax DNA damage marker, In the figure the DAPI staining for nuclei is shown in blue, the m6A reader is shown in green whereas with red the marker γH2Ax is shown. Untreated, and specific timepoints post treatment are shown in the figure. The cells were fixed 10', 30', 1h, 2h, 3h and 4h post UV treatment. Scalebar is 10 um.

The cells were fixed or selected at specific time points after the treatment. In the IF experiment below, the cells were fixed and prepared for immunofluorescence following the protocol described in section 2.5. The cells were fixed 4h post treatment. In Figure 13, untreated with UVC (control) and UVC treated cells 4h post treatment is shown. The gamma-H2Ax is used as a marker for DNA double strand breaks (Kuo and Yang 2008). With green is depicted the m6A reader whereas the marker is shown with red.

In order to follow the process and the implication of YTHDC1 in response to UVC radiation, the experiment was repeated using HCT116 and treated with UVC (20 J/m^2). The cells were fixed 10', 30', 1h, 2h, 3h and 4h. The analysis was performed using immunofluorescence (Figure 15).

The levels of YTHDC1, the DNA damage marker γH2AX and the histone H2AX were evaluated using Western blot, following the same experimental set up, as described above. HCT116 cells were treated with UV radiation and the cells were collected at specific time points post treatment as shown in the following Western blot (Figure 16).

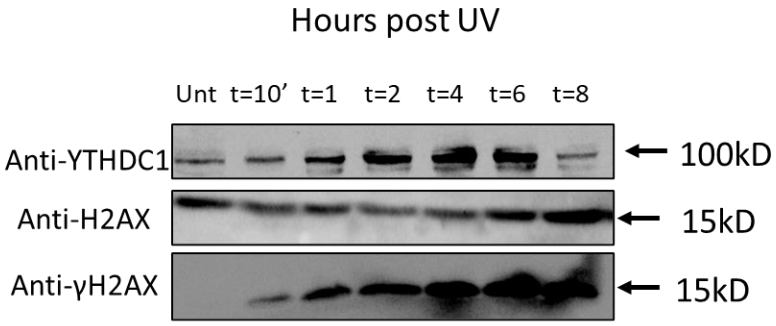


Figure 16. Western blot evaluating the levels of YTHDC1 and γH2Ax DNA damage marker. In the figure the protein levels of YTHDC1, γH2AX and the histone H2AX were evaluated at specific time points post UV radiation. The cells were collected at 10 minutes, 1h, 2h, 4h, 6h and 8h post treatment. Untreated cells were used as well.

3. Design and establishment of the dTAG depletion system for YTHDC1

To better dissect the role of YTHDC1 in UVC response, I employed the dTAG system for the acute depletion of YTHDC1. This tool will be used to dissect and characterize the primary effects of YTHDC1 depletion in response to DNA damage.

The first step for the generation of the dTAG-YTHDC1 constructs was the *in silico* design of primers that would ensure the insertion of the mutant FKPB in frame to the gene of interest, the YTHDC1. In total, 2 constructs were designed and generated, one containing the mCherry fluorescent protein and the second containing the BFP. The fluorescent proteins were used to validate the integration of the constructs inside the cells after Cas9 mediated knock in, using Flow cytometry and cell sorting. Two plasmids were used as templates in order to isolate the mutant FKBP followed by different fluorophore coding sequences (see table 2.1.1). The constructs were inserted in the genomic loci of interest through the homology directed repair (HDR), in which the double strand breaks can be repaired using a template strand. In order for the HDR to take place and ensure the insertion of the constructs after the Cas9 cut, there must have been a homology between the arms of the constructs and the gene of interest in this case, YTHDC1.

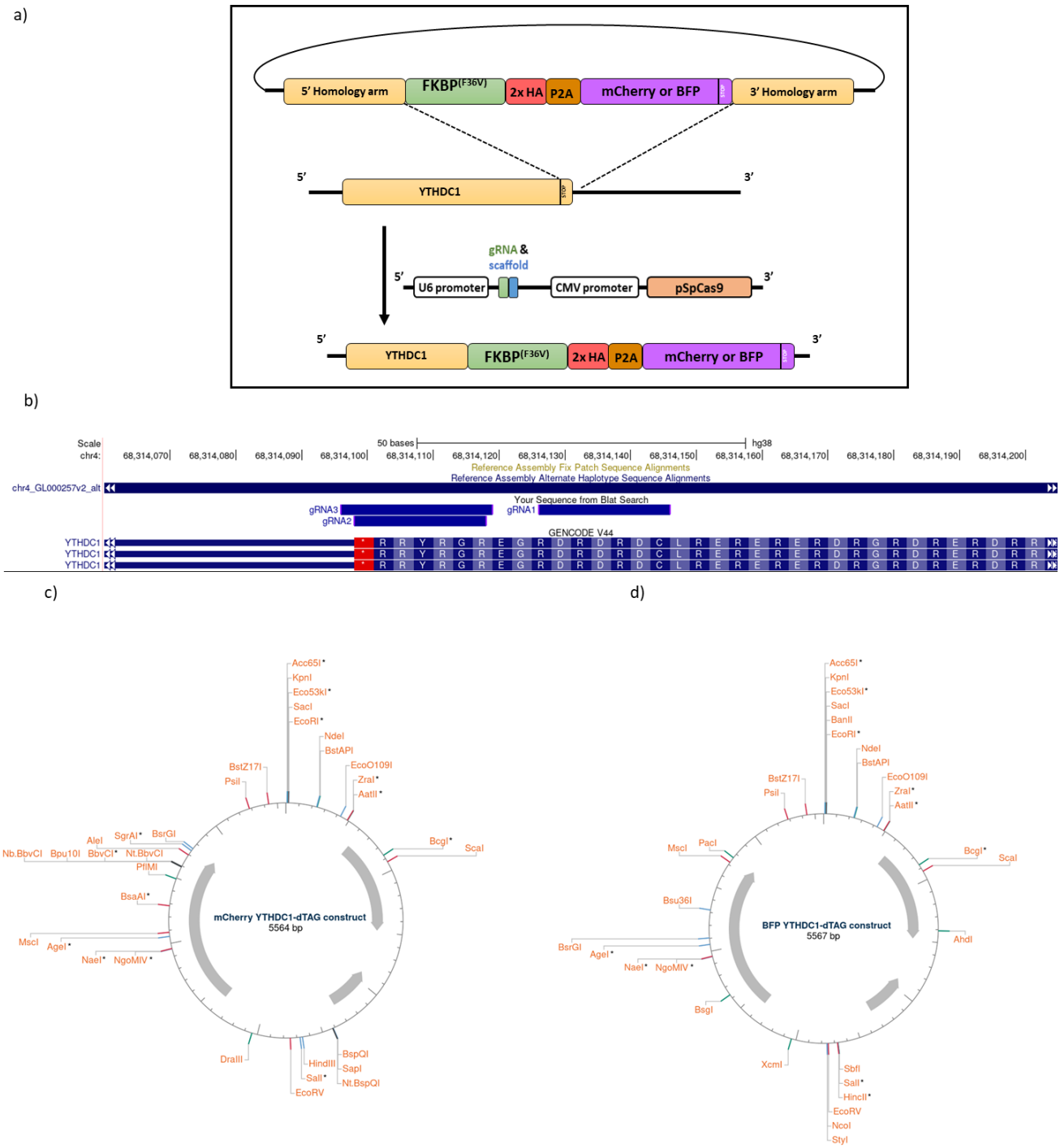


Figure 17. In silico design of the dTAG-YTHDC1 constructs. a) Design homology directed repair donor plasmid, schematic illustration of the insertion of FKBP12F36V-2xHA-P2A-fluorescent protein tag in the genetic locus of interest, b) screenshot from the UCSC genome browser of YTHDC1 stop codon, the target regions of the designed gRNAs are also shown. The stop codon is shown in red. I designed asymmetrical homology arms as described in Guarnaccia et al, 2021. For the 5' homology arm selected 200 bp upstream of the stop codon and for the 3' homology arm 800 bp downstream of the stop codon. These arms flank but exclude the stop codon itself. c, d) The expected plasmids as generated from the NEBcutter online tool. I designed 2 plasmids containing the 5' and 3' homology arms for YTHDC1, the vector fragment and a fluorescent protein. One plasmid contains the mCherry protein (c) and the other contains the blue fluorescent protein (BFP) (d).

For that aim, a 5' and a 3' homology arm was designed as depicted in Figure 17a. The sequence was retrieved from NCBI (https://www.ncbi.nlm.nih.gov/geo/) using the genome reference (GRCh38). The 5' homology arm stretched 200 bp upstream of the stop codon and the 3' homology arm stretched 800 bp downstream of the stop codon (Figure 17b). The stop codon is shown in red. Taking all these into consideration, primers were designed via the NEBuilder tool.

Regarding the selection of the appropriate guide RNAs that would ensure a successful CRISPR-Cas9 knock in of the constructs in the genomic loci of YTHDC1, three guide RNAs were designed. In order to maximize the efficiency of the insertion, the guide RNAs were designed to target a region of ~20 bp of the stop codon. The sequence of ~100 bp centered on the stop codon was retrieved from NCBI. The most promising guide RNAs with higher specificity and less off-target effects were selected using the online tool CRISPOR (http://crispor.tefor.net/).

The steps that were followed to design the 5' and 3' homology arms for the gene of interest (YTHDC1) are briefly described in Figure 17. The schematic representation of the steps and homology directed repair donor plasmid is shown, as well as a screenshot of the UCSC genome browser for the genetic locus of YTHDC1 and the selected gRNA. The oligos for the amplification of the homology arms as well as the guide RNA sequences were designed by taking into consideration the sequences upstream and downstream of the stop codon. The expected constructs are shown in Figure 17c-d.

4. Generation of the dTAG - YTHDC1 constructs

For the construction of the dTAG plasmids 5 different fragments were isolated through PCR. The FKBP-mCherry cassette and the FKBP-BFP cassette were isolated from the plasmids pAW62.YY1.FKBP.knock-in.mCherry and pAW63.YY1.FKBP.knock-in.BFP, respectively. The vector backbone was isolated from the same plasmids. The PCR products were loaded on 1% agarose gels (Figure 18) and the size of the bands were compared to the expected ones. Since the bands had the correct size, they were gel extracted and the DNA was purified. The homology arms were isolated using specific oligos as described in the Materials and Method section. The homology arms were isolated from HCT116 cells, in order to include cell line specific alterations in their sequence. The agarose gels of the 5' and 3' homology arms are depicted in Figure 18.

After having extracted and purified each product, the different fragments were Gibson assembled to create the final constructs. The assembled constructs were transferred to chemically competent DH5α *E.coli*. Plasmid DNA was extracted from the colonies and was further tested to check whether they contain all fragments and if the fragments were inserted in the expected way. The validation of the constructs is presented in the next paragraph.

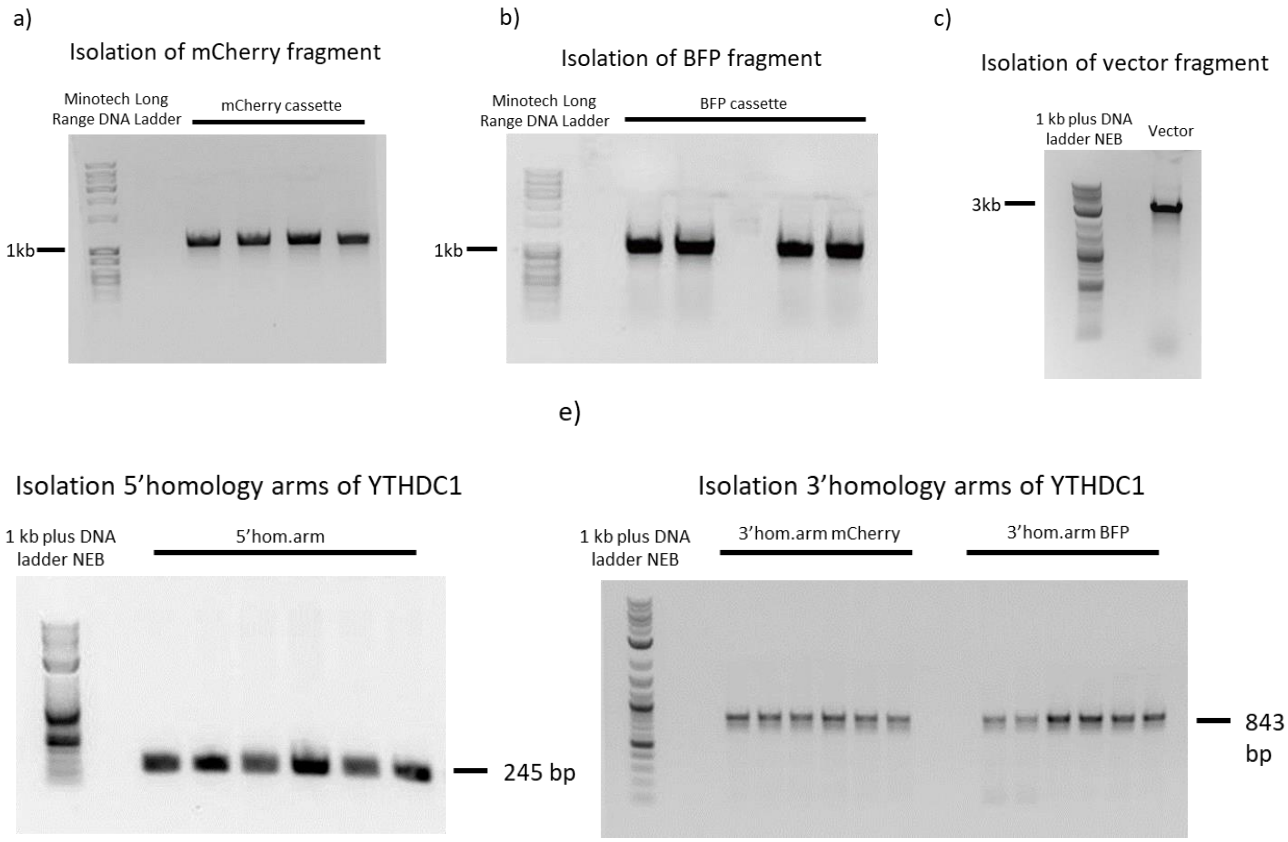


Figure 18. Isolation and extraction of the fragments required for the dTAG system. The different fragments were amplified via PCR reactions using specific primers. The PCR products were loaded in 1% agarose gels and the products are shown. a) isolation of the mCherry cassette, expected size 1kb, b) isolation of the BPF cassette, the expected size is 1kb, c) isolation of the vector fragment, the expected size is ~3kb., d) the 5' homology arm was extracted from HCT116 cells, the expected PCR product is 245 bp size long., e) the 3' homology arm was isolated from HCT116 cells, the expected size is approximately 840 bp.

5. Validation of the dTAG - YTHDC1 constructs

To verify the results of the Gibson assembly three different methods were employed: restriction digestion analysis, PCR and Sanger sequencing. Representative results are depicted in the following figures. First, the two plasmids were digested with restriction enzymes giving specific band sizes according to their cutting site. For the digestion of the BFP-YTHDC1 plasmid the enzymes: 1. AccI (with expected band sizes 768 bp, 796 bp, 1196 bp, 2807 bp, 2. EarI (with expected band sizes 30 bp, 446 bp, 488 bp, 1173 bp, 1626 bp, 1804 bp), 3. BbsI (40 bp, 141 bp, 470 bp, 942 bp, 3971 bp). (Figure 19). For the digestion of the mCherry-YTHDC1 plasmid the enzymes: HindIII (expected profile linearization of plasmids), ApoI (expected fragments 446 bp, 488 bp, 1173 bp, 1653 bp, 1804 bp) and EarI (expected fragments 211 bp, 228 bp, 534 bp, 1685 bp, 2906 bp) (Figure 20).

This analysis via the restriction enzymes showed which of the plasmids tested could serve as the best candidates for further analysis. The plasmids that gave the expected band sizes after the different digestions, were further validated via PCR using the sequencing primers to

test whether they contain the 5’ and 3’ homology arm and whether their total length is correct (Figure 21).

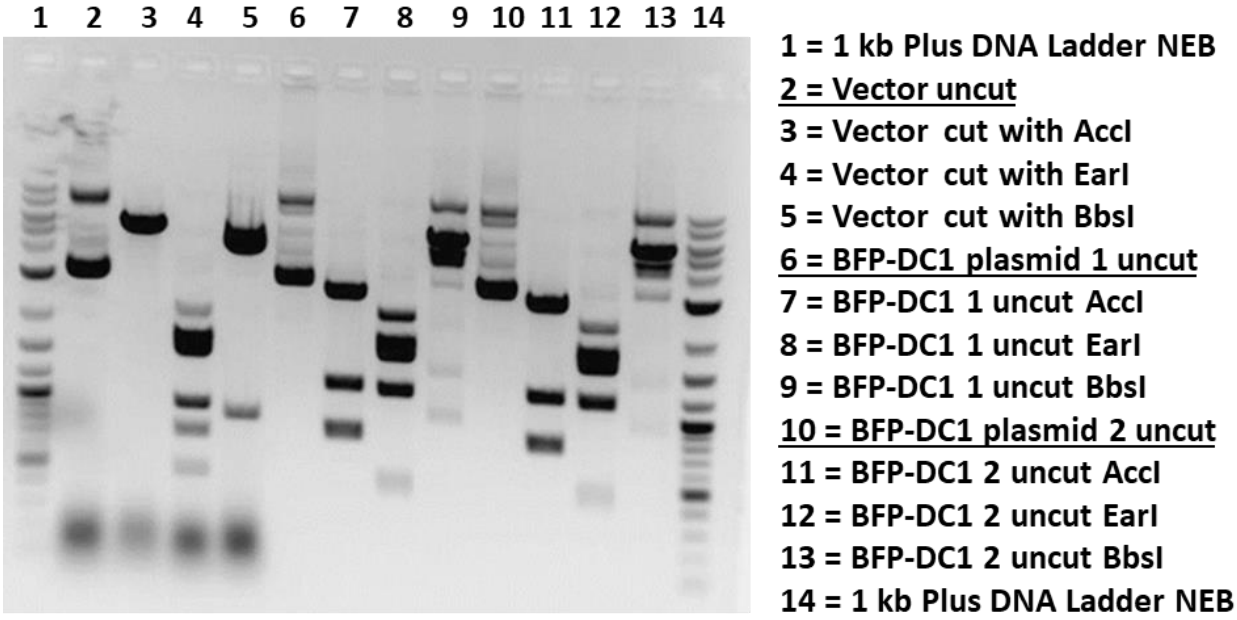


Figure 19. Validation of the BFP-YTHDC1 constructs using restriction enzymes. For testing the assembled plasmids, the enzymes AccI, EarI and BbsI were used. For AccI the expected band sizes were 768 bp, 796 bp, 1196 bp, 2807 bp, for EarI the expected band sizes: 30 bp, 446 bp, 488 bp, 1173 bp, 1626 bp, 1804 bp and for BbsI the expected band sizes: 40 bp, 141 bp, 470 bp, 942 bp, 3971 bp.

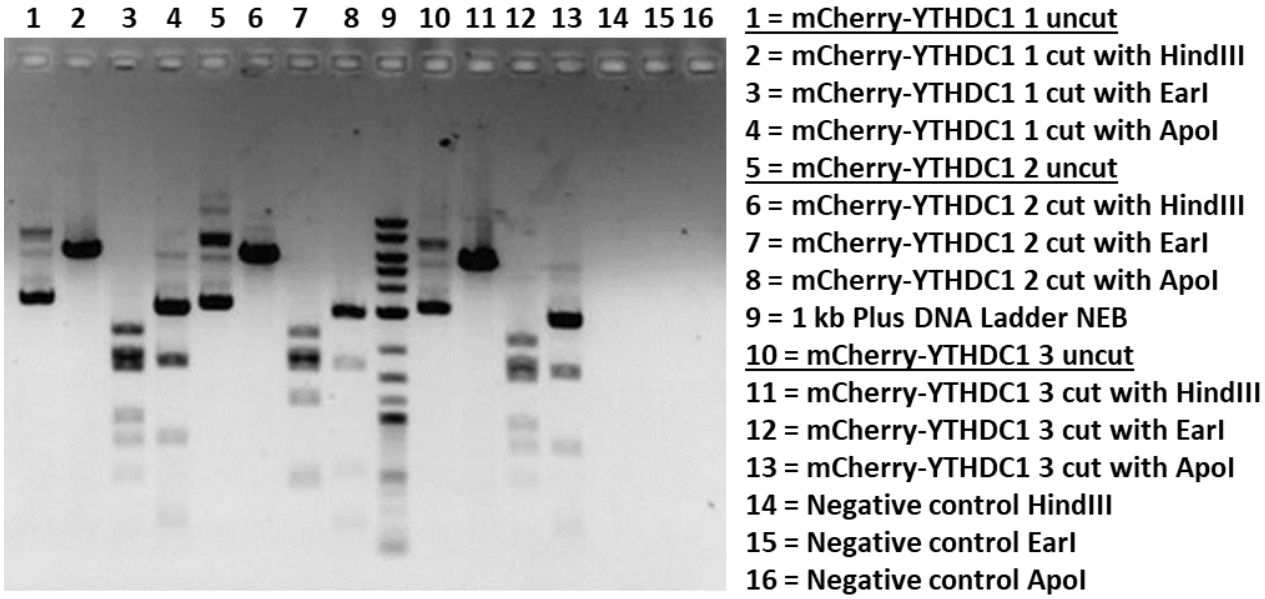


Figure 20. Validation of the mCherry-YTHDC1 constructs using restriction enzymes. For testing the assembled plasmids the enzymes HindIII,ApoI and EarI were used. For HindIII the expected profile is the linearization of the plasmid, for ApoI the expected band sizes are 446 bp, 488 bp, 1173 bp, 1653 bp, 1804 bp, for EarI the expected band sizes: 211 bp, 228 bp, 534 bp, 1685 bp, 2906 bp.

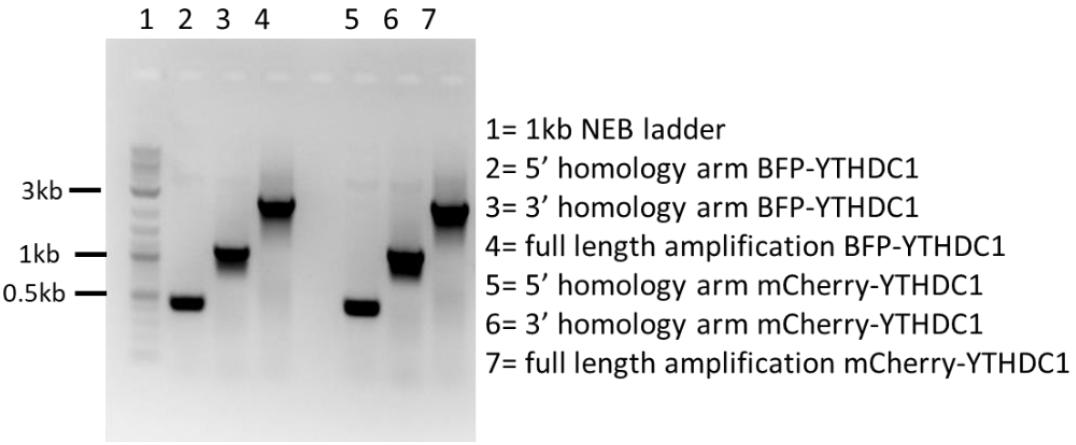


Figure 21. PCR amplification to validate the insertion of the 5', 3' homology arms and the full length of the candidate plasmids. The expected sizes of the PCR products are for the 5’ homology arm ~430 kb, for the 3’hom arm ~1kb and for the full length ~2,3kb.

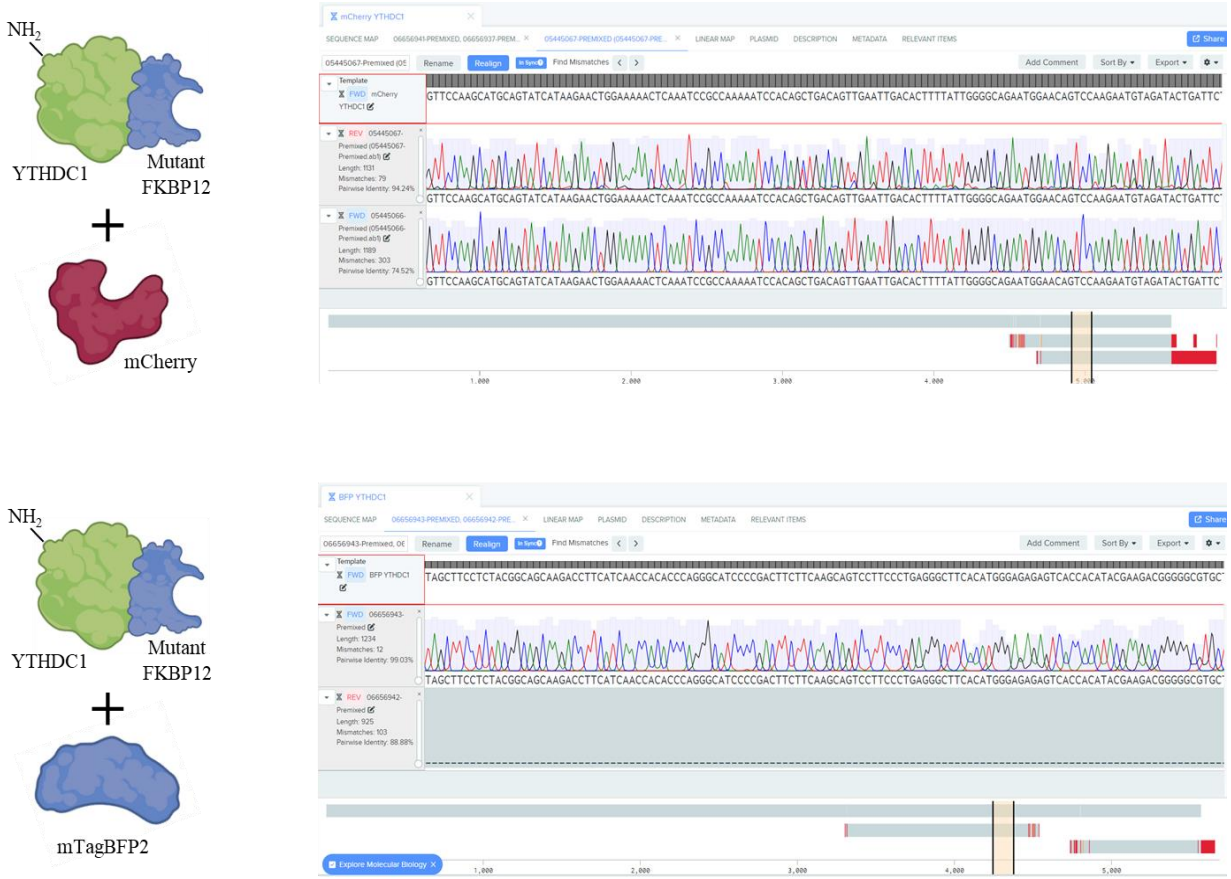


Figure 22. Validation of YTHDC1 plasmids via Sanger sequencing. The sequences of the plasmids were further validated via Sanger sequencing, the results were aligned to the expected sequence of the designed constructs. In the figure snapshots of the alignment using the Benchling online tool, is depicted, along with the electropherograms of the Sanger sequencing. The upper alignment corresponds to the mCherry-YTHDC1, the lower corresponds to the BFP-YTHDC1.

The best candidates were further validated using Sanger sequencing and aligned the products of the sequencing with that of the expected constructs. For the alignment the online tool Benchling was used as well as the Basic Local Alignment Search Tool (BLAST) from

National Center for Biotechnology Information (NCBI). Representative screenshots of the alignments using Benchling are depicted above (Figure 22).

After validating the sequence of constructs with Sanger sequencing, the preparation of the guide RNAs started. As described in the Methods section, the guide RNAs were cloned into the Cas9 plasmid after the latter had been linearized with BbsI digestion (Figure 23). After the ligation process and the transformation of the bacteria, the colonies were tested to see if they contained the expected constructs with colony PCR (Figure 23). The best candidates for containing the right constructs were validated with Sanger sequencing.

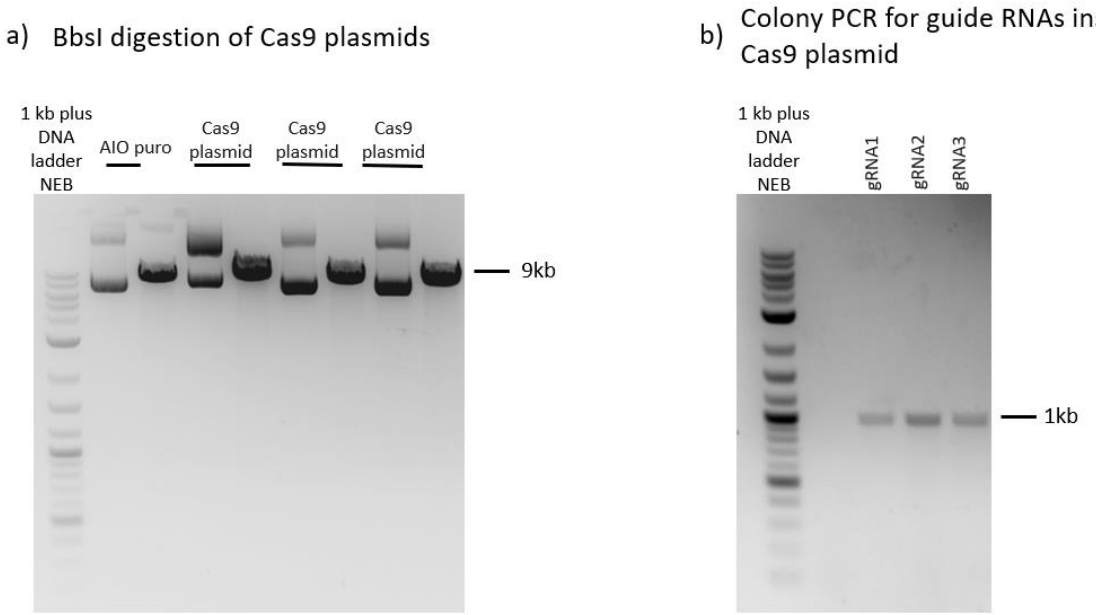


Figure 23. Preparation and validation of the Cas9 plasmids containing the appropriate guide RNAs for the knock in of the dTAG plasmid in HCT116 cells. The Cas9 plasmid was linearized using BbsI to generate sticky ends. The guide RNAs were designed to have complement ends to the sticky ends of the plasmid after BbsI digestion, to ensure the successful insertion of the guide RNAs to the appropriate position inside the plasmid. The linear plasmids were then loaded on agarose gels and extracted. The single stranded guide RNAs were annealed and formed duplexes. The duplexes were then phosphorylated and ligated to the vector. The ligated product was transformed into bacteria. The colonies were tested via colony PCR to check if the colonies contained the right constructs. a) 1% agarose gel of the Cas9 linear plasmids after BbsI digestion, AIO puro plasmid was used as a positive control, b) products of the colony PCR of the Cas9-guide RNA construct. The constructs were also validated via Sanger sequencing.

After the validation of both the constructs and the Cas9 plasmids containing the gRNAs with the correct sequence, HCT116 cells were transfected. The transfection included the delivery of in total 5 plasmids inside the cells, the 3 different Cas9 plasmids with the inserted gRNAs that were designed (as described in the paragraph 2.7.6 in the methods section, the sequences of the gRNAs are depicted in the 2.1.2 paragraph of the materials section) and the two constructs that were generated, one with the mCherry and the other with the BFP fluorophore. The double transfected cells were sorted and selected using Fluorescence-activated Cell Sorting (FACS). The results of the sorting are depicted in Figure 24. After the sorting analysis, approximately 1000 double positive cells were collected and further grown to establish homogenous cell lines containing dTAG insertion on both alleles of YTHDC1.

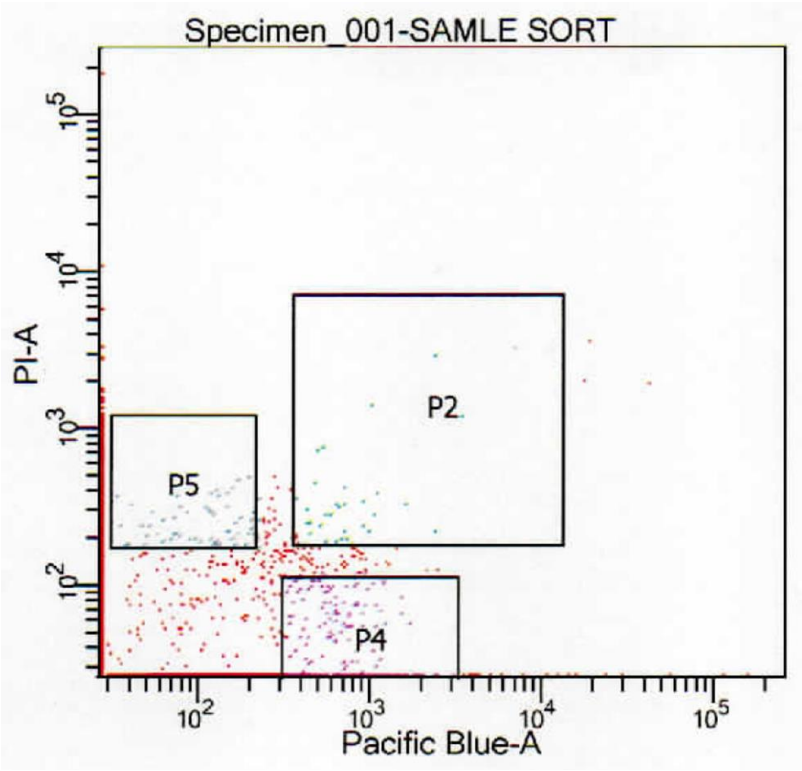


Figure 24. Flow cytometry analysis of genomic integration of fluorescent markers. Distinct populations of BFP-positive, mCherry-positive, and double-positive cells are observed. The BFP-positive are indicated in the P4 box, the mCherry-positive in the P5 box and lastly, the double positive cells are gated in green, P2 box.

6. Using catalytically inactive Rnase H to map R-loops

The most widespread tool in order to detect and map R-loops is by the use of the S9.6 monoclonal antibody (Boguslawski et al, 1986). However, its use especially for imaging purposes has been found to be problematic as S9.6 can also bind to double-stranded RNA (dsRNA) in *vitro* and in *vivo*. Therefore, researchers have focused on the development of alternative tools to image and map R-loops. A new technique is based on the ability of RNase H to recognize R-loops. In *vivo*, RNase H binds to R-loops and cleaves the RNA moiety, leading to the release of these trihybrid structures. Chen et al, in 2017 presented R-ChIP and proposed that the catalytically inactive RNase H can be used to recognize RNA:DNA hybrids without cleaving the RNA moiety, therefore, this tool can be used to detect R-loops by detecting the mutant RNase H. In this case, the dead RNase H is fused with a V5 tag, which is targeted by an anti-V5 antibody. A schematic representation of this tool is depicted in the figure below (Figure 25).

This tool was established in HEK 293T cells which were transfected with the plasmids and the cells successfully transfected were selected via Hygromycin B (Addgene id#111904 corresponds to the catalytically inactive RNase H, #111905 corresponds to double mutated RNase H, which serves as a negative control).

The tool was also validated via immunofluorescence experiments. By using HEK293T cells untransfected (control) and transfected with the dRNase H1 (dRNaseH) were able to observe the localization of the dRNaseH (D210N) and the double mutated dRNase H (WKKD).

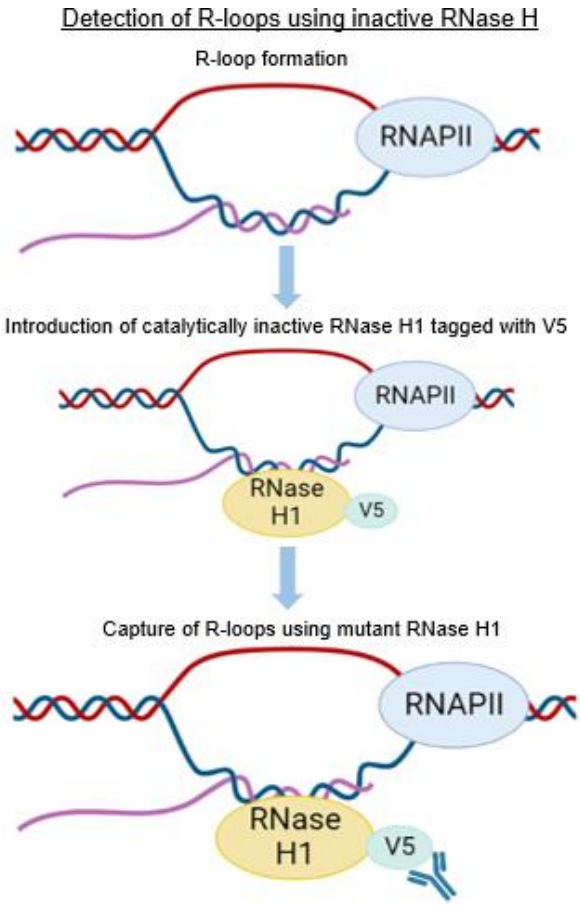


Figure 13. Schematic presentation of R-ChIP. In this strategy, a catalytically inactive RNase H is overexpressed and fused with a V5 tag. The mutant RNase H can still recognize RNA:DNA hybrids without cleaving the RNA moiety. The R-loops can then be detected using an antibody against the V5 tag. The Figure was designed at Biorender: https://www.biorender.com/

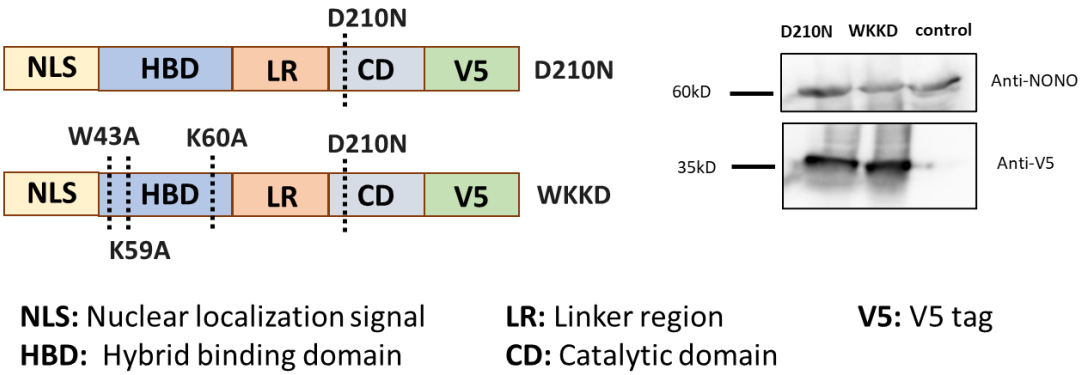


Figure 14. Validation of stable HEK293T expressing the inactive RNase H (dRNaseH). The cells containing the plasmids with the inactive RNaseH were selected via Hygromycin B. The design of the vectors containing the inactive forms of RNase H. The D210N contains a mutation in the catalytic domain of the nuclease, therefore, the RNAse can recognize R-loops but cannot cleave the RNA moiety. The WKKD plasmid contains four mutations, one is the D210N on the catalytic activity and three in the RNA-DNA hybrid binding domain (W43A, K59A, K60A), therefore, this strain cannot recognize nor bind R-loops, serving as negative control. On the right, the western blot targeting anti-V5, showing the presence of the plasmids in the cells is depicted.

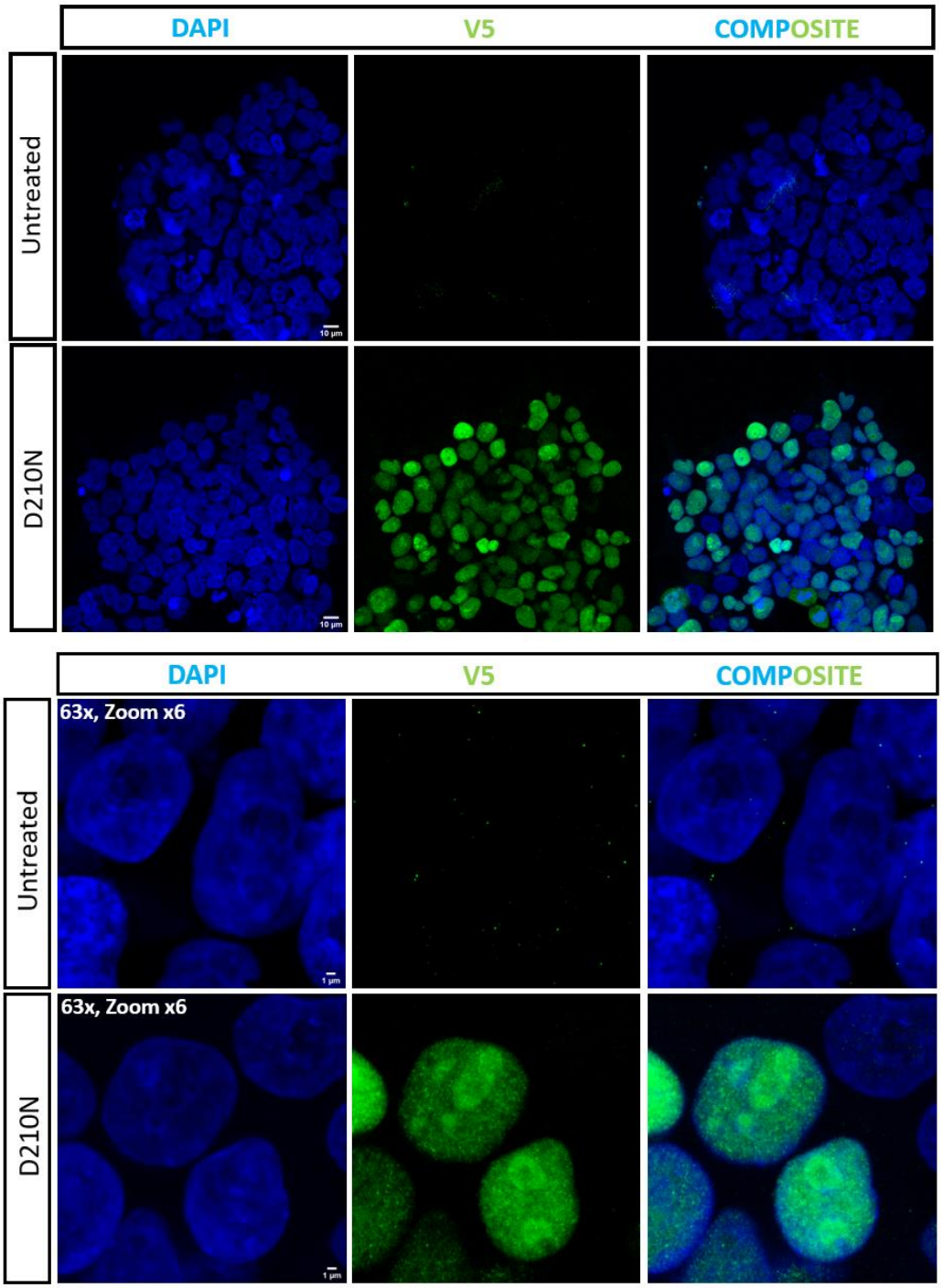


Figure 15. Immunofluorescent experiments studying the localization of the inactive RNase H (dRNaseH). Upper panel: For this experiment untreated (control, not transfected with plasmids) and treated (transfected and selected cells containing the D210N dRNase H) were used. DAPI is used to stain nuclei and is shown in blue, dRNase H is targeted via the V5 tag, shown in green. Scale bar 10 um. Below panel: zoom figures of control and D210N cells. Scale bar is 1um.

Discussion

Conclusion

The aim of this project was to unravel possible roles of m6A deposition in response to UV stress, focusing both on the relationship between (m6A) RNA methylation and DNA damage repair as well as the possible connection between m6A and UV induced R-loop regulation. Most specifically, I focused on the nuclear m6A reader YTHDC1, which regulates alternative splicing, mRNA export, and recently was also implicated in stress response (Timcheva et al, 2022). The contribution of YTHDC1 to heat stress response along with recent publications which show the recruitment of YTHDC1 on DNA damaged sites (Zhang et al, 2020), led me to examine further its role in response to DNA damage, and its possible regulatory role in DNA damage-induced R-loops.

First, I re-analyzed available datasets of raw RNA-seq data from YTHDC1 KD HeLa cells (Roundtree et al, 2017, GEO: GSE74397). In this analysis I focused on the genes whose expression is affected upon siRNA-mediated depletion of the nuclear m6A reader, YTHDC1. I found that among the genes that are significantly affected (at p-adjusted < 0.001), are genes that take part in DNA damage response, such as Rad51 that takes part in homologous recombination, (Bhattacharya et al, 2017) and ERCC1 that is a key component of the nucleotide excision repair pathway (McNeil and Melton, 2012). Even though this reanalysis itself does not prove the implication of YTHDC1 in DNA damage response, it served as a first indication that prompted me to study the role of this protein in response to DNA damage.

To better understand the possible implication of YTHDC1 in DNA damage response, UVC ($20J/m^2$) induced damage was utilized. Our preliminary data so far, both from the Western blots and the IF experiments suggested a possible role of YTHDC1 in response to DNA damage stress. Specifically, both techniques show the same trend, an increase of the levels of the m6A reader at specific time points post treatment suggesting a possible role of YTHDC1 in response to UVC radiation. However, confirming the localization of YTHDC1 on double strand breaks, which has been previously reported by Zhang et al, 2020, thus the co-localization of the m6A reader with the γH2AX, was not feasible using this technique. The signal from the m6A reader was diffused in the nuclei of the cells, therefore, co-localization could not be measured. A more optimal direction to check for localization of YTHDC1 on damaged sites is to use microirradiation, inducing localized DNA damage on chromatin with a beam of light (Svobodová Kovaříková et al, 2020). Microirradiation has been proposed as a technique that can cause DNA damage in defined regions of chromosomes (Berns et al, 1969; Berns et al, 1981). This enables the study of the spatiotemporal dynamics of damaged chromatin, including the recruitment of factors on the DNA lesions (Kong et al, 2009; Kruhlak et al, 2009; Ferrando-May et al, 2013; Kong et al, 2018). By focusing on the damaged chromatin, thus a specific region inside the nucleus, rather than the diffused signal that I observed in fixed nuclei (Figures 13, 14), the evaluation of the potential localization of YTHDC1 on damaged sites is possible in high resolution.

One of the most common ways to study the role and function of a protein in the cells is by removing this protein and studying the functional consequences of this depletion. The complete depletion of YTHDC1 is lethal, therefore, the CRISPR-Cas9 mediated knockout is not feasible (Kasowitz et al, 2018). It is worth noting that the depletion of m6A related factors, like METTL3 and METTL14, via siRNAs and/or CRISPR-Cas9 methods, had conflicting

results regarding the primary and secondary effects of that protein loss. Since m6A regulates RNA stability and degradation (Wang et al, 2014), the incomplete loss of function makes it particularly challenging to disentangle the potential functions of these proteins. Therefore, acute depletion systems targeting m6A related factors, could help studying the primary effects of the depletion of the proteins in the cellular and molecular context. In line with other studies, where acute protein depletion systems for m6A related factors, specifically for the m6A writer METTL3, were employed (Wei et al, 2021), a technical aim of this thesis was to utilize the dTAG system, for the acute degradation of YTHDC1 in HCT116 cell line. YTHDC1 is located on chromosome 4 and HCT116 cell line is diploid for this. Therefore, two plasmids were constructed, each one containing a different fluorescent protein, either mCherry or BFP. The use of both constructs and their successful transfection inside the cells is important for the validation and selection of cell populations that contain both constructs. Using FACS, cell populations with double knock-in were selected and further cultivated. Thus, I can report the successful design and construction of the necessary plasmids, as well as the efforts for the generation of a stable cell line in HCT116 colon cancer cells.

To examine whether the m6A reader YTHDC1 has a regulatory role in UV-induced R-loops, I established a new tool in the lab to detect R-loops. The most widespread way to map and detect RNA:DNA hybrids is by the use of the S9.6 monoclonal and structure specific antibody. However, recent studies have shown that S9.6 can produce artifacts, especially during IF experiments, as the antibody has significant affinity for dsRNA (Hortono et al, 2018), thus new detection methods depending on the ability of RNase H to bind R-loops have to be utilized. During this thesis, I was able to establish the RChIP method (Chen et al, 2017), by generating stable cell lines expressing the D210N- V5 and WKKD-V5 mutants of RNase H1, which was validated both via Western blot and immunofluorescence experiments. In the Western blot, I used an anti V5 antibody to target the V5 tag of the mutant RNases and I detected the V5 band for both HEK293T cell lines, expressing D210N and WKKD RNase H respectively. To examine the localization of the D210N mutant RNaseH, immunofluorescence experiments were performed. As expected, the D210N and WKKD RNase are localized exclusively in the nuclei of the cells, as shown by its colocalization with DAPI, which is a fluorescent stain for nuclei. It is worth noting that most of the signal coming from D210N-V5 was located in the nucleoli, where ribosomal DNA is stored. Ribosomal DNA has been reported to contain high levels of R-loop formation (El Hage et al, 2010, Wahba et al, 2016). These tools will be further used to study the R-loop formation especially after UV radiation and at specific genomic loci, especially when coupled with deep sequencing. The R-loops can then be identified by using ChIP peak calling algorithms, as described in Chen et al, 2017.

Future directions - Outlook

It has been suggested that, upon DNA damage, i.e in the case of UV irradiation, the RNA polymerase II (RNAPII) remains stalled at DNA lesions and therefore, transcription is perturbed (Donahue et al, 1994; Lavigne et al, 2017; Williamson et al, 2017). Ongoing elongation of transcription is important for the survival of cells upon stress (Lavigne et al, 2017). Additionally, m6A levels are affected by the elongation rate of RNAPII (Slobodin et al, 2017). It was shown that higher m6A deposition is observed in slowly transcribed RNAs compared to the rapidly synthesized transcripts (Slobodin et al, 2017). This could imply that the RNA from the damaged chromatin can be m6A decorated, while the RNAPII is stalled at DNA lesions. This hypothesis is supported by Svobodová Kovaříková and colleagues, who showed that m6A is preferentially bound on RNA at damaged chromatin regions (Svobodová Kovaříková et al, 2020).

One possible role of m6A in DNA damage response is that newly transcribed RNA, whose transcription is stalled at damaged sites, undergoes m6A modification, attracting the nuclear reader YTHDC1. This interaction might lead to a crosstalk or interplay with factors involved in DNA damage response, potentially activating the repair process through a non-canonical pathway. This hypothesis comes in line with Zhang et al, 2020 showing that m6A is found on the nascent RNA involved in DNA-RNA hybrids (R-loops) at damaged sites, the YTHDC1 is also recruited on the sites and the recruitment of BRCA1 and Rad51 modulate the homologous recombination repair at the damaged regions (Zhang et al, 2020; Gomez-Gonzalez and Aguilera 2020; Marnef and Legube, 2020). However, the mechanisms and the implicated downstream factors that contribute to R-loop mediated DNA damage response, have not been identified.

To better assess the role of RNA methylation, in the context of triggered R-loops, in DNA damage response, it is important to evaluate the molecular mechanism behind it. It is important to note that the outcome of m6A deposition is dictated by the “reader” proteins. Therefore, in order to study the molecular mechanism, it is crucial to identify the RNA-dependent protein interactome of YTHDC1 upon stress, possibly through Mass Spectrometry. This will unravel the proteins with which YTHDC1 interacts after UVC radiation and evaluate its role in DNA damage response.

To effectively isolate the protein of interest, immunoprecipitation experiments were carried out. Supplementary Figure 1 illustrates the efforts made to optimize the immunoprecipitation protocol for enriching YTHDC1. The next step, after the enrichment of the protein, is to carry out experiments in control and after UVC radiation conditions, co-immunoprecipitation of YTHDC1 and Mass Spectrometry analysis. The differential RNA-dependent protein interactome of the m6A reader YTHDC1 upon UV stress will unravel the possible roles of YTHDC1 and RNA methylation (m6A) in response to DNA damage. The identification of YTHDC1 RNA-dependent protein interactors is crucial to approach the working hypothesis, and mechanistically characterize its role in m6A-mediated DNA-damage response, which at least partially involves the regulation, formation, and eventual resolution of triggered R-loops at sites of DNA damage.

Another suggestion for future experiments is to resolve whether the finding supporting involvement of YTHDC1 in the DNA damage response, that is observed in IF confocal

microscopy and increased protein levels (Figure 13, 16) is m6A dependent. A more straightforward approach is to check how the YTHDC1 protein levels are altered upon METTL3 depletion, which is expected to affect (reduce) the levels of the m6A RNA modification. To assess that, dTAG plasmids have been designed and created for the conditional acute protein degradation of METTL3. Since METTL3 is the major writer of m6A, an acute depletion of the protein is expected to decrease the RNA methylation levels. The first step is to check if and how the expression levels of YTHDC1 are affected upon depletion of METTL3. Next, the dTAG-METTL3 stable cells could be treated with UVC radiation to gain an insight into the role of METTL3 and m6A in UVC response. In this condition, the YTHDC1 levels can be assessed, and therefore, unravel whether the participation of the m6A reader is m6A dependent. The acute protein depletion of METTL3 can be accompanied by its catalytic inhibition. The drug STM2457 (Yankova et al, 2021) can inhibit the catalytic activity of METTL3, thus its ability to transfer methyl groups to adenosines. The combination of these experiments will determine whether it is the protein and/or its catalytic activity that is important for DNA damage response to UVC radiation.

The dTAG system for acute depletion of the m6A reader YTHDC1 that has been established during this thesis, will enable better understanding of the role of YTHDC1 in DNA damage response and could be used to evaluate the levels of gamma H2AX, the DNA damage marker, upon acute depletion of YTHDC1. If the levels of this marker are upregulated when the reader is depleted compared to the wild type cell line, this will be a strong indication that this reader is important for the genome integrity of the cells, in the absence of any other exogenous treatment or stress. This experiment can also be coupled with UVC radiation in dTAG-YTHDC1 and WT cells: there the levels of the γH2AX marker will assess whether the dTAG-YTHDC1 treated cells (i.e., in the absence of YTHDC1) are more susceptible to DNA damage compared to WT. In addition, the role of YTHDC1 in DNA damage response can be assessed by restoring the protein levels of the reader in dTAG-YTHDC1 cells (by removal of the chemical dTAG-47); thereby, if the γH2AX levels are restored (reduced) upon restoring the protein levels of YTHDC1, this could further support that YTHDC1 and m6A regulate response to DNA damage, and help towards the maintenance of genomic stability.

It is worth noting, that, as presented in paragraph 1.2.1, one major role of YTHDC1 is in regulation of alternative splicing, which is essentially mediated by reading the m6A mark deposited co-transcriptionally around or near transcript splice sites (Xiao et al, 2016; Roundtree et al, 2016, Louloupi et al, 2018). Alternative splicing can regulate responses upon different stress stimuli, by affecting the alternative isoform production of RNA transcripts. Alternative RNA splicing and different isoform formation was observed upon UV damage response as well (Williamson et al, 2017). Thus, questions arise about how YTHDC1 affects alternative splicing genome-wide upon UV radiation. It could be that the splicing of DNA damage responsive genes is altered upon UVC stress. This could potentially assist or enhance the repair of DNA breaks. Additional studies have supported the link between splicing and DNA damage, while recent studies showed that the splicing factor XAB2 interacts with repair factors and processes R-loops under specific conditions that induce the formation of R-loops (Goulielmaki et al, 2021). UV can interfere with alternative splicing (Muñoz et al, 2009). However, the interplay between alternative splicing and DNA damage response, has not been fully uncovered. To study the potential role of splicing, further experiments using UV

and long-read sequencing such as Oxford Nanopore (direct RNA or cDNA) sequencing, could reveal the changes of splicing efficiency at individual splice sites and at transcript level upon stress, as well as the different isoforms of each gene that are being expressed and the specific genomic loci that could be affected. This could show whether the splicing rate and efficiency is affected upon UV treatment; furthermore, implementing the newly established dTAG-YTHDC1 cell lines shall unravel if any observed changes in splicing are YTHDC1 dependent.

In summary, this thesis has achieved the generation of dTAG-YTHDC1 HCT116 cells, the establishment of RChIP (dRNase H-V5) and the evaluation of YTHDC1 levels upon UVC radiation. Defects in RNA processing, including splicing and modifications have been linked to neurodegeneration (Hutton et al, 1998; Han et al, 2020; Pupak et al, 2022), metabolic disorders (Sen et al, 2013; Zhang et al, 2021) and cancer (Barbieri et al, 2017; Liu et al, 2018; Wang et al, 2020). Abnormalities in DNA damage repair pathways have also been implicated in the development of cancer and neurodegenerative disorders (reviewed in Kamileri et al, 2012). Therefore, unraveling the functional link between RNA processing and DNA damage repair is important for understanding their role in disease development and progression.

References

Abakir, A., Giles, T. C., Cristini, A., Foster, J. M., Dai, N., Starczak, M., Rubio-Roldan, A., Li, M., Eleftheriou, M., Crutchley, J., Flatt, L., Young, L., Gaffney, D. J., Denning, C., Dalhus, B., Emes, R. D., Gackowski, D., Corrêa, I. R., Garcia-Perez, J. L., … Ruzov, A. (2020). N 6-methyladenosine regulates the stability of RNA:DNA hybrids in human cells. *Nature Genetics*, *52*(1), 48–55. https://doi.org/10.1038/s41588-019-0549-x

Aguilera, A., & García-Muse, T. (2012). R Loops: From Transcription Byproducts to Threats to Genome Stability. *Molecular Cell*, *46*(2), 115–124. https://doi.org/10.1016/j.molcel.2012.04.009

Allison, D. F., & Wang, G. G. (2019). R-loops: Formation, function, and relevance to cell stress. In *Cell Stress* (Vol. 3, Issue 2, pp. 38–46). Shared Science Publishers OG. https://doi.org/10.15698/cst2019.02.175

Arzumanian, V. A., Dolgalev, G. V., Kurbatov, I. Y., Kiseleva, O. I., & Poverennaya, E. V. (2022). Epitranscriptome: Review of Top 25 Most-Studied RNA Modifications. *International Journal of Molecular Sciences*, *23*(22). https://doi.org/10.3390/ijms232213851

Barbieri, I., Tzelepis, K., Pandolfini, L., Shi, J., Millán-, G., Robson, S. C., Aspris, D., Migliori, V., Andrew, J., Han, N., Braekeleer, E. De, Ponstingl, H., Hendrick, A., Vakoc, C. R., Vassiliou, G. S., & Kouzarides, T. (2018). *Europe PMC Funders Group Promoter-bound METTL3 maintains myeloid leukaemia via m6A- dependent translation control*. *552*(7683), 126–131. https://doi.org/10.1038/nature24678.Promoter-bound

Batista, P. J., Molinie, B., Giallourakis, C. C., Howard, Y., Batista, P. J., Molinie, B., Wang, J., Qu, K., Zhang, J., Li, L., Bouley, D. M., Dedon, P., Wernig, M., Mullen, A. C., Xing, Y., Giallourakis, C. C., & Chang, H. Y. (2014). m 6 A RNA Modification Controls Cell Fate Transition in Mammalian Embryonic Stem Cells Article m 6 A RNA Modification Controls Cell Fate Transition in Mammalian Embryonic Stem Cells. *Stem Cell*, *15*(6), 707–719. https://doi.org/10.1016/j.stem.2014.09.019

Berns, M. W., Aist, J., Edwards, J., Strahs, K., Girton, J., McNeill, P., Rattner, J. B., Kitzes, M., Hammer-Wilson, M., Liaw, L. H., Siemens, A., Koonce, M., Peterson, S., Brenner, S., Burt, J., Walter, R., Bryant, P. J., Van Dyk, D., Coulombe, J., … Berns, G. S. (1981). Laser microsurgery in cell and developmental biology. *Science*, *213*(4507), 505–513. https://doi.org/10.1126/science.7017933

Bokar, J. A., Shambaugh, M. E., Polayes, D., Matera, A. G., & Rottman, F. M. (1997). Purification and cDNA cloning of the AdoMet-binding subunit of the human mRNA (N6-adenosine)-methyltransferase. *Rna*, *3*(11), 1233–1247.

Chakraborty, P., & Grosse, F. (2011). Human DHX9 helicase preferentially unwinds RNA-containing displacement loops (R-loops) and G-quadruplexes. *DNA Repair*, *10*(6), 654–665. https://doi.org/10.1016/j.dnarep.2011.04.013

Chatzidoukaki, O., Stratigi, K., Goulielmaki, E., Niotis, G., Akalestou-Clocher, A., Gkirtzimanaki, K., Zafeiropoulos, A., Altmüller, J., Topalis, P., & Garinis, G. A. (2021).

R-loops trigger the release of cytoplasmic ssDNAs leading to chronic inflammation upon DNA damage. *Science Advances*, *7*(47). https://doi.org/10.1126/sciadv.abj5769

Chen, L., Chen, J. Y., Zhang, X., Gu, Y., Xiao, R., Shao, C., Tang, P., Qian, H., Luo, D., Li, H., Zhou, Y., Zhang, D. E., & Fu, X. D. (2017). R-ChIP Using Inactive RNase H Reveals Dynamic Coupling of R-loops with Transcriptional Pausing at Gene Promoters. *Molecular Cell*, *68*(4), 745-757.e5. https://doi.org/10.1016/j.molcel.2017.10.008

Cohen, S., Puget, N., Lin, Y. L., Clouaire, T., Aguirrebengoa, M., Rocher, V., Pasero, P., Canitrot, Y., & Legube, G. (2018). Senataxin resolves RNA:DNA hybrids forming at DNA double-strand breaks to prevent translocations. *Nature Communications*, *9*(1). https://doi.org/10.1038/s41467-018-02894-w

Crossley, M. P., Bocek, M., & Cimprich, K. A. (2019). R-Loops as Cellular Regulators and Genomic Threats. *Molecular Cell*, *73*(3), 398–411. https://doi.org/10.1016/j.molcel.2019.01.024

Crossley, M. P., Brickner, J. R., Song, C., Zar, S. M. T., Maw, S. S., Chédin, F., Tsai, M. S., & Cimprich, K. A. (2021). Catalytically inactive, purified rnase h1: A specific and sensitive probe for rna–dna hybrid imaging. *Journal of Cell Biology*, *220*(9). https://doi.org/10.1083/jcb.202101092

Crossley, M. P., Song, C., Bocek, M. J., Choi, J., Sathirachinda, A., Lin, C., Brickner, J. R., Bai, G., Lans, H., Vermeulen, W., Abu-remaileh, M., & Cimprich, K. A. (2023). *immune response*. *613*(7942), 187–194. https://doi.org/10.1038/s41586-022-05545-9.R-loop-derived

Dominissini, D., Moshitch-Moshkovitz, S., Schwartz, S., Salmon-Divon, M., Ungar, L., Osenberg, S., Cesarkas, K., Jacob-Hirsch, J., Amariglio, N., Kupiec, M., Sorek, R., & Rechavi, G. (2012). Topology of the human and mouse m6A RNA methylomes revealed by m6A-seq. *Nature*, *485*(7397), 201–206. https://doi.org/10.1038/nature11112

Doxtader, K. A., Wang, P., Scarborough, A. M., Seo, D., Conrad, N. K., & Nam, Y. (2018). Structural Basis for Regulation of METTL16, an S-Adenosylmethionine Homeostasis Factor. *Molecular Cell*, *71*(6), 1001-1011.e4. https://doi.org/10.1016/j.molcel.2018.07.025

El Hage, A., French, S. L., Beyer, A. L., & Tollervey, D. (2010). Loss of Topoisomerase I leads to R-loop-mediated transcriptional blocks during ribosomal RNA synthesis. *Genes and Development*, *24*(14), 1546–1558. https://doi.org/10.1101/gad.573310

Ferrando-May, E., Tomas, M., Blumhardt, P., Stöckl, M., Fuchs, M., & Leitenstorfer, A. (2013). Highlighting the DNA damage response with ultrashort laser pulses in the near infrared and kinetic modeling. *Frontiers in Genetics*, *4*(JUL), 1–8. https://doi.org/10.3389/fgene.2013.00135

Frye, M., T. Haranda, B., Behm, M., & He, C. (2018). Expression During Development. *Science*, *361*(September), 1346–1349.

Ginno, P. A., Lott, P. L., Christensen, H. C., Korf, I., & Chédin, F. (2012). R-Loop Formation Is a Distinctive Characteristic of Unmethylated Human CpG Island Promoters. *Molecular Cell*, *45*(6), 814–825. https://doi.org/10.1016/j.molcel.2012.01.017

Gómez-González, B., & Aguilera, A. (2020). Looping the (R) Loop in DSB Repair via RNA Methylation. *Molecular Cell*, *79*(3), 361–362. https://doi.org/10.1016/j.molcel.2020.07.015

Goulielmaki, E., Tsekrekou, M., Batsiotos, N., Ascensão-Ferreira, M., Ledaki, E., Stratigi, K., Chatzinikolaou, G., Topalis, P., Kosteas, T., Altmüller, J., Demmers, J. A., Barbosa-Morais, N. L., & Garinis, G. A. (2021). The splicing factor XAB2 interacts with ERCC1-XPF and XPG for R-loop processing. *Nature Communications*, *12*(1). https://doi.org/10.1038/s41467-021-23505-1

Guarnaccia, A. D., Weissmiller, A. M., & Tansey, W. P. (2021). Gene-specific quantification of nascent transcription following targeted degradation of endogenous proteins in cultured cells. *STAR Protocols*, *2*(4), 101000. https://doi.org/10.1016/j.xpro.2021.101000

Han, M., Liu, Z., Xu, Y., Liu, X., Wang, D., Li, F., Wang, Y., & Bi, J. (2020). Abnormality of m6A mRNA Methylation Is Involved in Alzheimer's Disease. *Frontiers in Neuroscience*, *14*(February), 1–9. https://doi.org/10.3389/fnins.2020.00098

Hasanova, Z., Klapstova, V., Porrua, O., Stefl, R., & Sebesta, M. (2023). Human senataxin is a bona fide R-loop resolving enzyme and transcription termination factor. *Nucleic Acids Research*, *51*(6), 2818–2837. https://doi.org/10.1093/nar/gkad092

He, P. C., & He, C. (2021). m 6 A RNA methylation: from mechanisms to therapeutic potential . *The EMBO Journal*, *40*(3), 1–15. https://doi.org/10.15252/embj.2020105977

Hegazy, Y. A., Fernando, C. M., & Tran, E. J. (2020). The balancing act of R-loop biology: The good, the bad, and the ugly. *Journal of Biological Chemistry*, *295*(4), 905–913. https://doi.org/10.1074/jbc.REV119.011353

Hyjek, M., Figiel, M., & Nowotny, M. (2019). RNases H: Structure and mechanism. *DNA Repair*, *84*(April), 102672. https://doi.org/10.1016/j.dnarep.2019.102672

Jia, G., Fu, Y., Zhao, X., Dai, Q., Zheng, G., Yang, Y., Yi, C., Lindahl, T., Pan, T., Yang, Y., He, C., Group, S., Genomics, D., District, C., Laboratories, C. H., & Mimms, S. (2012). *HHS Public Access*. *7*(12), 885–887. https://doi.org/10.1038/nchembio.687.N

Kamileri, I., Karakasilioti, I., & Garinis, G. A. (2012). Nucleotide excision repair: New tricks with old bricks. *Trends in Genetics*, *28*(11), 566–573. https://doi.org/10.1016/j.tig.2012.06.004

Kasowitz, S. D., Ma, J., Anderson, S. J., Leu, N. A., Xu, Y., Gregory, D., Schultz, R. M., & Wang, P. J. (2018). *Nuclear m 6 A reader YTHDC1 regulates alternative polyadenylation and splicing during mouse oocyte development*. 1–28.

Kiehn, O., & Car. (2017). 乳鼠心肌提取 HHS Public Access. *Physiology & Behavior*, *176*(3), 139–148. https://doi.org/10.1007/978-1-59745-190-1

Kong, X., Cruz, G. M. S., Silva, B. A., Wakida, N. M., Khatibzadeh, N., Berns, M. W., & Yokomori, K. (2018). Laser microirradiation to study in vivo cellular responses to simple and complex DNA damage. *Journal of Visualized Experiments*, *2018*(131), 1–10. https://doi.org/10.3791/56213

Kong, X., Mohanty, S. K., Stephens, J., Heale, J. T., Gomez-Godinez, V., Shi, L. Z., Kim, J. S., Yokomori, K., & Berns, M. W. (2009). Comparative analysis of different laser systems to study cellular responses to DNA damage in mammalian cells. *Nucleic Acids Research*, *37*(9). https://doi.org/10.1093/nar/gkp221

Lavigne, M. D., Konstantopoulos, D., Ntakou-Zamplara, K. Z., Liakos, A., & Fousteri, M. (2017). Global unleashing of transcription elongation waves in response to genotoxic stress restricts somatic mutation rate. *Nature Communications*, *8*(1). https://doi.org/10.1038/s41467-017-02145-4

Layden, H. M., Eleuteri, N. A., Hiebert, S. W., & Stengel, K. R. (2021). A protocol for rapid degradation of endogenous transcription factors in mammalian cells and identification of direct regulatory targets. *STAR Protocols*, *2*(2), 100530. https://doi.org/10.1016/j.xpro.2021.100530

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., & Durbin, R. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics*, *25*(16), 2078–2079. https://doi.org/10.1093/bioinformatics/btp352

Li, X., & Manley, J. L. (2005). Inactivation of the SR protein splicing factor ASF/SF2 results in genomic instability. *Cell*, *122*(3), 365–378. https://doi.org/10.1016/j.cell.2005.06.008

Li, X., Xiong, X., & Yi, C. (2016). Epitranscriptome sequencing technologies: Decoding RNA modifications. *Nature Methods*, *14*(1), 23–31. https://doi.org/10.1038/nmeth.4110

Liao, S., Sun, H., & Xu, C. (2018). YTH Domain: A Family of N6-methyladenosine (m6A) Readers. *Genomics, Proteomics and Bioinformatics*, *16*(2), 99–107. https://doi.org/10.1016/j.gpb.2018.04.002

Lim, Y. W., Sanz, L. A., Xu, X., Hartono, S. R., & Chédin, F. (2015). Genome-wide DNA hypomethylation and RNA:DNA hybrid accumulation in Aicardi–Goutières syndrome. *ELife*, *4*(JULY2015), 1–21. https://doi.org/10.7554/eLife.08007

Lin, S., Choe, J., Du, P., Triboulet, R., & Gregory, R. I. (2016). The m6A Methyltransferase METTL3 Promotes Translation in Human Cancer Cells. *Molecular Cell*, *62*(3), 335–345. https://doi.org/10.1016/j.molcel.2016.03.021

Liu, J., Yue, Y., Han, D., Wang, X., Fu, Y., Zhang, L., Jia, G., Yu, M., Lu, Z., Deng, X., Dai, Q., Chen, W., & He, C. (2014). *HHS Public Access*. *10*(2), 93–95. https://doi.org/10.1038/nchembio.1432.A

Liu, J., Dou, X., Chen, C., Chen, C., Liu, C., Michelle Xu, M., Zhao, S., Shen, B., Gao, Y., Han, D., & He, C. (2020). N6-methyladenosine of chromosome-associated regulatory RNA regulates chromatin state and transcription. *Science*, *367*(6477), 580–586. https://doi.org/10.1126/science.aay6018

Liu, Z. X., Li, L. M., Sun, H. L., & Liu, S. M. (2018). Link Between m6A Modification and Cancers. *Frontiers in Bioengineering and Biotechnology*, *6*(July), 1–11. https://doi.org/10.3389/fbioe.2018.00089

Louloupi, A., Ntini, E., Conrad, T., & Ørom, U. A. V. (2018). Transient N-6-Methyladenosine Transcriptome Sequencing Reveals a Regulatory Role of m6A in Splicing Efficiency. *Cell Reports*, *23*(12), 3429–3437. https://doi.org/10.1016/j.celrep.2018.05.077

Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*, *15*(12), 1–21. https://doi.org/10.1186/s13059-014-0550-8

Lu, W. T., Hawley, B. R., Skalka, G. L., Baldock, R. A., Smith, E. M., Bader, A. S., Malewicz, M., Watts, F. Z., Wilczynska, A., & Bushell, M. (2018). Drosha drives the formation of DNA:RNA hybrids around DNA break sites to facilitate DNA repair. *Nature Communications*, *9*(1). https://doi.org/10.1038/s41467-018-02893-x

Marnef, A., & Legube, G. (2020). m6A RNA modification as a new player in R-loop regulation. *Nature Genetics*, *52*(1), 27–28. https://doi.org/10.1038/s41588-019-0563-z

Mazina, O. M., Keskin, H., Hanamshet, K., Storici, F., & Mazin, A. V. (2017). Rad52 Inverse Strand Exchange Drives RNA-Templated DNA Double-Strand Break Repair. *Molecular Cell*, *67*(1), 19-29.e3. https://doi.org/10.1016/j.molcel.2017.05.019

Meyer, K. D., Saletore, Y., Zumbo, P., Elemento, O., Mason, C. E., & Jaffrey, S. R. (2012). Comprehensive analysis of mRNA methylation reveals enrichment in 3′ UTRs and near stop codons. *Cell*, *149*(7), 1635–1646. https://doi.org/10.1016/j.cell.2012.05.003

Mosler, T., Conte, F., Longo, G. M. C., Mikicic, I., Kreim, N., Möckel, M. M., Petrosino, G., Flach, J., Barau, J., Luke, B., Roukos, V., & Beli, P. (2021). R-loop proximity proteomics identifies a role of DDX41 in transcription-associated genomic instability. *Nature Communications*, *12*(1), 1–17. https://doi.org/10.1038/s41467-021-27530-y

Muñoz, M. J., Santangelo, M. S. P., Paronetto, M. P., de la Mata, M., Pelisch, F., Boireau, S., Glover-Cutter, K., Ben-Dov, C., Blaustein, M., Lozano, J. J., Bird, G., Bentley, D., Bertrand, E., & Kornblihtt, A. R. (2009). DNA Damage Regulates Alternative Splicing through Inhibition of RNA Polymerase II Elongation. *Cell*, *137*(4), 708–720. https://doi.org/10.1016/j.cell.2009.03.010

Nabet, B., Roberts, J. M., Buckley, D. L., Paulk, J., Yang, A., Leggett, A. L., Erb, M. A., Lawlor, M. A., Scott, T. G., Vittori, S., Perry, J. A., Qi, J., Winter, G. E., Wong, K., Gray, N. S., & Bradner, J. E. (2018). *degradation*. *14*(5), 431–441. https://doi.org/10.1038/s41589-018-0021-8.The

Nogueira, A., Fernandes, M., Catarino, R., & Medeiros, R. (2019). RAD52 functions in homologous recombination and its importance on genomic integrity maintenance and cancer therapy. *Cancers*, *11*(11). https://doi.org/10.3390/cancers11111622

Orsolic, I., Carrier, A., & Esteller, M. (2023). Genetic and epigenetic defects of the RNA modification machinery in cancer. *Trends in Genetics*, *39*(1), 74–88. https://doi.org/10.1016/j.tig.2022.10.004

Parajuli, S., Teasley, D. C., Murali, B., Jackson, J., Vindigni, A., & Stewart, S. A. (2017). Human ribonuclease H1 resolves R-loops and thereby enables progression of the DNA replication fork. *Journal of Biological Chemistry*, *292*(37), 15216–15224. https://doi.org/10.1074/jbc.M117.787473

Paulsen, R. D., Soni, D. V., Wollman, R., Hahn, A. T., Yee, M. C., Guan, A., Hesley, J. A., Miller, S. C., Cromwell, E. F., Solow-Cordero, D. E., Meyer, T., & Cimprich, K. A. (2009). A Genome-wide siRNA Screen Reveals Diverse Cellular Processes and Pathways that Mediate Genome Stability. *Molecular Cell*, *35*(2), 228–239. https://doi.org/10.1016/j.molcel.2009.06.021

Powell, W. T., Coulson, R. L., Gonzales, M. L., Crary, F. K., Wong, S. S., Adams, S., Ach, R. A., Tsang, P., Yamada, N. A., Yasui, D. H., Chédin, F., & LaSalle, J. M. (2013). R-loop formation at Snord116 mediates topotecan inhibition of Ube3a-antisense and allele-specific chromatin decondensation. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(34), 13938–13943. https://doi.org/10.1073/pnas.1305426110

Pupak, A., Singh, A., Sancho-Balsells, A., Alcalá-Vida, R., Espina, M., Giralt, A., Martí, E., Ørom, U. A. V., Ginés, S., & Brito, V. (2022). Altered m6A RNA methylation contributes to hippocampal memory deficits in Huntington's disease mice. *Cellular and Molecular Life Sciences*, *79*(8), 1–21. https://doi.org/10.1007/s00018-022-04444-6

Quinlan, A. R., & Hall, I. M. (2010). BEDTools: A flexible suite of utilities for comparing genomic features. *Bioinformatics*, *26*(6), 841–842. https://doi.org/10.1093/bioinformatics/btq033

Rafalska, I., Zhang, Z., Benderska, N., Wolff, H., Hartmann, A. M., Brack-Werner, R., & Stamm, S. (2004). The intranuclear localization and function of YT521-B is regulated by tyrosine phosphorylation. *Human Molecular Genetics*, *13*(15), 1535–1549. https://doi.org/10.1093/hmg/ddh167

Roundtree, I. A., Luo, G. Z., Zhang, Z., Wang, X., Zhou, T., Cui, Y., Sha, J., Huang, X., Guerrero, L., Xie, P., He, E., Shen, B., & He, C. (2017). YTHDC1 mediates nuclear export of N6-methyladenosine methylated mRNAs. *ELife*, *6*, 1–28. https://doi.org/10.7554/eLife.31311

Rücklé, C., Körtel, N., Basilicata, M. F., Busch, A., Zhou, Y., Hoch-Kraft, P., Tretow, K., Kielisch, F., Bertin, M., Pradhan, M., Musheev, M., Schweiger, S., Niehrs, C., Rausch, O., Zarnack, K., Keller Valsecchi, C. I., & König, J. (2023). RNA stability controlled by m6A methylation contributes to X-to-autosome dosage compensation in mammals.

Nature Structural and Molecular Biology, *30*(8), 1207–1215. https://doi.org/10.1038/s41594-023-00997-7

Saletore, Y., Meyer, K., Korlach, J., Vilfan, I. D., Jaffrey, S., & Mason, C. E. (2012). The birth of the Epitranscriptome: deciphering the function of RNA modifications. *Genome Biology*, *13*(10), 175. https://doi.org/10.1186/gb-2012-13-10-175

Schindelin, J., Arganda-Carrera, I., Frise, E., Verena, K., Mark, L., Tobias, P., Stephan, P., Curtis, R., Stephan, S., Benjamin, S., Jean-Yves, T., Daniel, J. W., Volker, H., Kevin, E., Pavel, T., & Albert, C. (2009). Fiji - an Open platform for biological image analysis. *Nature Methods*, *9*(7). https://doi.org/10.1038/nmeth.2019.Fiji

Schwartz, S., Mumbach, M. R., Jovanovic, M., Wang, T., Maciag, K., Bushkin, G. G., Mertins, P., Ter-Ovanesyan, D., Habib, N., Cacchiarelli, D., Sanjana, N. E., Freinkman, E., Pacold, M. E., Satija, R., Mikkelsen, T. S., Hacohen, N., Zhang, F., Carr, S. A., Lander, E. S., & Regev, A. (2014). Perturbation of m6A writers reveals two distinct classes of mRNA methylation at internal and 5' sites. *Cell Reports*, *8*(1), 284–296. https://doi.org/10.1016/j.celrep.2014.05.048

Sen, S., Jumaa, H., & Webster, N. J. G. (2013). Splicing factor SRSF3 is crucial for hepatocyte differentiation and metabolic function. *Nature Communications*, *4*(May 2012). https://doi.org/10.1038/ncomms2342

Skourti-Stathaki, K., & Proudfoot, N. J. (1984). Lee and Clayton 1996; Carles-Kinch and Kreuzer 1997), the mitochondrial origin of replication (Xu and Clayton 1996), and the mammalian Ig class Ó 2014 Skourti-Stathaki and Proudfoot This article. *Baker and Kornberg*, 1384–1396. https://doi.org/10.1101/gad.242990.114.Freely

Skourti-Stathaki, K., Proudfoot, N. J., & Gromak, N. (2011). Human Senataxin Resolves RNA/DNA Hybrids Formed at Transcriptional Pause Sites to Promote Xrn2-Dependent Termination. *Molecular Cell*, *42*(6), 794–805. https://doi.org/10.1016/j.molcel.2011.04.026

Slobodin, B., Han, R., Calderone, V., Vrielink, J. A. F. O., Loayza-Puch, F., Elkon, R., & Agami, R. (2017). Transcription Impacts the Efficiency of mRNA Translation via Co-transcriptional N6-adenosine Methylation. *Cell*, *169*(2), 326-337.e12. https://doi.org/10.1016/j.cell.2017.03.031

Sollier, J., Stork, C. T., García-Rubio, M. L., Paulsen, R. D., Aguilera, A., & Cimprich, K. A. (2014). Transcription-Coupled Nucleotide Excision Repair Factors Promote R-Loop-Induced Genome Instability. *Molecular Cell*, *56*(6), 777–785. https://doi.org/10.1016/j.molcel.2014.10.020

Sun, Q., Csorba, T., Skourti-Stathaki, K., Proudfoot, N. J., & Dean, C. (2016). Europe PMC Funders Group R-Loop Stabilization Represses Antisense Transcription at the Arabidopsis FLC Locus. *Science*, *340*(6132), 619–621. https://doi.org/10.1126/science.1234848.R-Loop

Svobodová Kovaříková, A., Stixová, L., Kovařík, A., Komůrková, D., Legartová, S., Fagherazzi, P., & Bártová, E. (2020). N6-Adenosine Methylation in RNA and a

Reduced m3G/TMG Level in Non-Coding RNAs Appear at Microirradiation-Induced DNA Lesions. *Cells*, *9*(2), 12–14. https://doi.org/10.3390/cells9020360

Sweatt,S.K, Gower, B.A, Chieh, A.Y, Liu, Y, Li, L. (2016). 乳鼠心肌提取 HHS Public Access. *Physiology & Behavior*, *176*(1), 139–148. https://doi.org/10.1038/nature13682.Transcript

Teng, Y., Yadav, T., Duan, M., Tan, J., Xiang, Y., Gao, B., Xu, J., Liang, Z., Liu, Y., Nakajima, S., Shi, Y., Levine, A. S., Zou, L., & Lan, L. (2018). ROS-induced R loops trigger a transcription-coupled but BRCA1/2-independent homologous recombination pathway through CSB. *Nature Communications*, *9*(1). https://doi.org/10.1038/s41467-018-06586-3

Timcheva, K., Dufour, S., Touat-Todeschini, L., Burnard, C., Carpentier, M. C., Chuffart, F., Merret, R., Helsmoortel, M., Ferré, S., Grézy, A., Couté, Y., Rousseaux, S., Khochbin, S., Vourc'h, C., Bousquet-Antonelli, C., Kiernan, R., Seigneurin-Berny, D., & Verdel, A. (2022). Chromatin-associated YTHDC1 coordinates heat-induced reprogramming of gene expression. *Cell Reports*, *41*(11). https://doi.org/10.1016/j.celrep.2022.111784

Tornaletti, S., Reines, D., & Hanawalt, P. C. (1999). Structural characterization of RNA polymerase II complexes arrested by a cyclobutane pyrimidine dimer in the transcribed strand of template DNA. *Journal of Biological Chemistry*, *274*(34), 24124–24130. https://doi.org/10.1074/jbc.274.34.24124

Traub, W., & Yonath, A. (1969). © 1969 Nature Publishing Group. *Group*, *224*, 177–178. http://adsabs.harvard.edu/abs/1969Natur.224..177K

Van Tran, N., Ernst, F. G. M., Hawley, B. R., Zorbas, C., Ulryck, N., Hackert, P., Bohnsack, K. E., Bohnsack, M. T., Jaffrey, S. R., Graille, M., & Lafontaine, D. L. J. (2019). The human 18S rRNA m6A methyltransferase METTL5 is stabilized by TRMT112. *Nucleic Acids Research*, *47*(15), 7719–7733. https://doi.org/10.1093/nar/gkz619

Wahba, L., Costantino, L., Tan, F. J., Zimmer, A., & Koshland, D. (2016). S1-DRIP-seq identifies high expression and polyA tracts as major contributors to R-loop formation. *Genes and Development*, *30*(11), 1327–1338. https://doi.org/10.1101/gad.280834.116

Wang, P., Doxtader, K. A., & Nam, Y. (2016). Structural Basis for Cooperative Function of Mettl3 and Mettl14 Methyltransferases. *Molecular Cell*, *63*(2), 306–317. https://doi.org/10.1016/j.molcel.2016.05.041

Wang, T., Kong, S., Tao, M., & Ju, S. (2020). The potential role of RNA N6-methyladenosine in Cancer progression. *Molecular Cancer*, *19*(1), 1–18. https://doi.org/10.1186/s12943-020-01204-7

Wei, G., Almeida, M., Pintacuda, G., Coker, H., Bowness, J. S., Ule, J., & Brockdorff, N. (2021). Acute depletion of METTL3 implicates N6-methyladenosine in alternative intron/exon inclusion in the nascent transcriptome. *Genome Research*, *31*(8), 1395–1408. https://doi.org/10.1101/gr.271635.120

Williamson, L., Saponaro, M., Boeing, S., East, P., Mitter, R., Kantidakis, T., Kelly, G. P., Lobley, A., Walker, J., Spencer-Dene, B., Howell, M., Stewart, A., & Svejstrup, J. Q. (2017). UV Irradiation Induces a Non-coding RNA that Functionally Opposes the Protein Encoded by the Same Gene. *Cell*, *168*(5), 843-855.e13. https://doi.org/10.1016/j.cell.2017.01.019

Xiang, Y., Laurent, B., Hsu, C., Nachtergaele, S., Lu, Z., Sheng, W., Xu, C., Chen, H., Ouyang, J., Wang, S., Ling, D., Hsu, P., Zou, L., Jambhekar, A., & He, C. (2017). m6A RNA methylation regulates the UV-induced DNA damage response. *Nature*, *543*(7646), 573–576. https://doi.org/10.1038/nature21671.m

Xiao Wang1, Zhike Lu1, Adrian Gomez1, Gary C. Hon2, Yanan Yue1, Dali Han1, Ye Fu1, Marc Parisien3, Qing Dai1, Guifang Jia1, 4, Bing Ren2, Tao Pan3, and C. H. (2014). m6A-dependent regulation of messenger RNA stability. *Nature*, *505*(7481), 1–20. https://doi.org/10.1038/nature12730.m

Xiao, W., Adhikari, S., Xiao, W., Adhikari, S., Dahal, U., Chen, Y., Hao, Y., Sun, B., & Sun, H. (2016). Article Nuclear m 6 A Reader YTHDC1 Regulates mRNA Splicing. *Molecular Cell*, *61*(4), 507–519. https://doi.org/10.1016/j.molcel.2016.01.012

Yan, Q., Shields, E. J., Bonasio, R., & Sarma, K. (2019). Mapping Native R-Loops Genome-wide Using a Targeted Nuclease Approach. *Cell Reports*, *29*(5), 1369-1380.e5. https://doi.org/10.1016/j.celrep.2019.09.052

Yangguang Ou, Rachael E Wilson, and S. G. W. (2018). 乳鼠心肌提取 HHS Public Access. *Annu Rev Anal Chem (Palo Alto Calif)*, *11*(1), 509–533. https://doi.org/10.1038/nature14234.N

Yankova, E., Blackaby, W., Albertella, M., Rak, J., De Braekeleer, E., Tsagkogeorga, G., Pilka, E. S., Aspris, D., Leggate, D., Hendrick, A. G., Webster, N. A., Andrews, B., Fosbeary, R., Guest, P., Irigoyen, N., Eleftheriou, M., Gozdecka, M., Dias, J. M. L., Bannister, A. J., … Kouzarides, T. (2021). Small-molecule inhibition of METTL3 as a strategy against myeloid leukaemia. *Nature*, *593*(7860), 597–601. https://doi.org/10.1038/s41586-021-03536-w

Yu, K., Chedin, F., Hsieh, C. L., Wilson, T. E., & Lieber, M. R. (2003). R-loops at immunoglobulin class switch regions in the chromosomes of stimulated B cells. *Nature Immunology*, *4*(5), 442–451. https://doi.org/10.1038/ni919

Zaccara, S., Ries, R. J., & Jaffrey, S. R. (2019). Reading, writing and erasing mRNA methylation. *Nature Reviews Molecular Cell Biology*, *20*(10), 608–624. https://doi.org/10.1038/s41580-019-0168-5

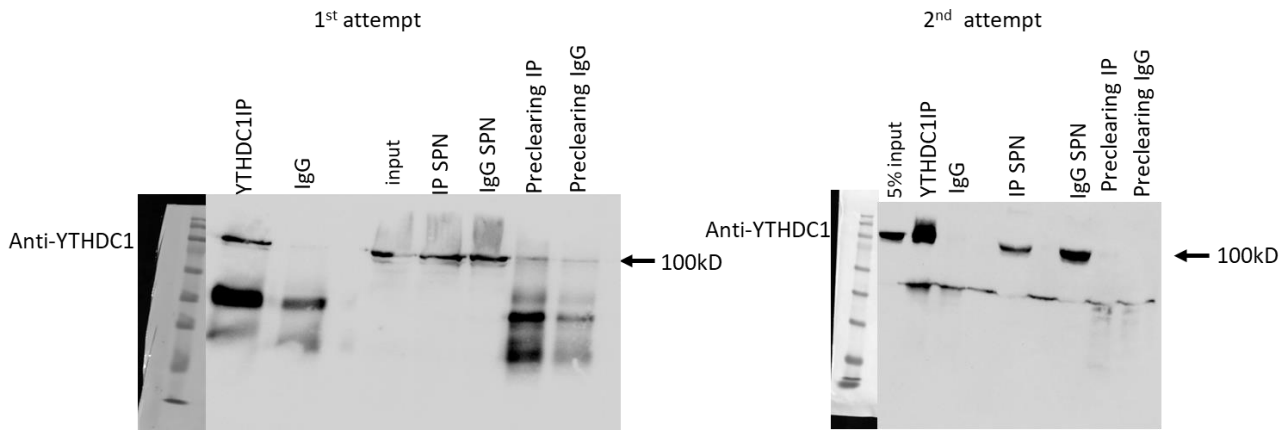
Zhang, C., Chen, L., Peng, D., Jiang, A., He, Y., Zeng, Y., Xie, C., Zhou, H., Luo, X., Liu, H., Chen, L., Ren, J., Wang, W., & Zhao, Y. (2020). METTL3 and N6-Methyladenosine Promote Homologous Recombination-Mediated Repair of DSBs by Modulating DNA-RNA Hybrid Accumulation. *Molecular Cell*, *79*(3), 425-442.e7. https://doi.org/10.1016/j.molcel.2020.06.017

Zhang, Y., Chen, W., Zheng, X., Guo, Y., Cao, J., Zhang, Y., Wen, S., Gao, W., & Wu, Y. (2021). Regulatory role and mechanism of m6A RNA modification in human metabolic diseases. *Molecular Therapy - Oncolytics*, *22*(85), 52–63. https://doi.org/10.1016/j.omto.2021.05.003

Zheng, G., Dahl, J. A., Niu, Y., Fedorcsak, P., Huang, C. M., Li, C. J., Vågbø, C. B., Shi, Y., Wang, W. L., Song, S. H., Lu, Z., Bosmans, R. P. G., Dai, Q., Hao, Y. J., Yang, X., Zhao, W. M., Tong, W. M., Wang, X. J., Bogdan, F., … He, C. (2013). ALKBH5 Is a Mammalian RNA Demethylase that Impacts RNA Metabolism and Mouse Fertility. *Molecular Cell*, *49*(1), 18–29. https://doi.org/10.1016/j.molcel.2012.10.015

Zhou, R., Zhang, J., Bochman, M. L., Zakian, V. A., & Ha, T. (2014). Periodic DNA patrolling underlies diverse functions of Pif1 on R-loops and G-rich DNA. *ELife*, *3*, 1–16. https://doi.org/10.7554/elife.02190

Appendix



Supplementary Figure 1. Optimization of the immunoprecipitation protocol against YTHDC1. Two biological and technical attempts were carried out in order to pull down sufficient amounts of the protein of interest. In the first attempt (Left), the YTHDC1 is pulled down, however, a significant amount was found in the flowthrough, therefore, the experiment was repeated. In the second attempt (right), the same amount of the initial protein was used (400ug), however, the amount of magnetic beads as well as the amount of the antibody used to precipitate the protein, were increased (see text above). The YTHDC1 was successfully pulled down and the flowthrough contained less amount of protein compared to the first attempt. IgG was used as negative control.

For the IP experiments, in the first two technical replicates 400ug of total protein were used. In the first attempt, we could observe that the protein is immunoprecipitated, however, some portion of the protein was also found in the flowthrough. To further optimize it, I used the same amount of protein, but I increased the volume of magnetic beads used as well as the antibody to bind YTHDC1. It is noteworthy that the signal I got after these changes, I saw an enriched signal of my pulled down protein compared to the input control (5% of the IP). An additional step to optimize the pull-down procedure could be to reduce the amount of the total protein (the suggested amount could be 300ug), since the signal of the immunoprecipitated protein is very strong. This will enable the sufficient enrichment of the protein with less starting material.