



**HAL**  
open science

# Role of YTHDC1, nuclear reader of N6-methyladenosine (m6A) RNA modification, in the regulation of the cellular heat stress response.

Kalina Timcheva

## ► To cite this version:

Kalina Timcheva. Role of YTHDC1, nuclear reader of N6-methyladenosine (m6A) RNA modification, in the regulation of the cellular heat stress response.. Cellular Biology. Université Grenoble Alpes [2020-..], 2020. English. NNT: 2020GRALV011 . tel-03793808

**HAL Id: tel-03793808**

**<https://theses.hal.science/tel-03793808v1>**

Submitted on 2 Oct 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

## **THÈSE**

Pour obtenir le grade de

**DOCTEUR DE L'UNIVERSITE GRENOBLE ALPES**

Spécialité : **Biologie cellulaire**

Arrêté ministériel : 25 mai 2016

Présentée par

**Kalina TIMCHEVA**

Thèse dirigée par **Dr. André VERDEL** et  
codirigée par **Dr. Daphné SEIGNEURIN-BERNY**

préparée au sein de **Laboratoire CRI IAB - Centre de Recherche  
Epigenetics, Chronic Diseases, Cancer - Institute for Advanced  
Biosciences**  
dans l'**École Doctorale Chimie et Sciences du Vivant**

# **Role of YTHDC1, nuclear reader of N<sup>6</sup>-methyladenosine (m6A) RNA modification, in the regulation of the cellular heat stress response.**

Thèse soutenue publiquement le **30 Septembre 2020**  
devant le jury composé de :

**Pr. Pierre HAINAUT**

Professeur, IAB, Grenoble (Président)

**Dr. Valérie MEZGER**

Directrice de recherche, Paris Diderot (Rapporteur)

**Dr. Bertrand SERAPHIN**

Directeur de recherche, IGBMC, Strasbourg (Rapporteur)

**Dr. Ramesh PILLAI**

Directeur de recherche, UNIGE, Genève (Membre)

**Dr. André VERDEL**

Directeur de recherche, IAB, Grenoble (Membre)

**Dr. Daphné SEIGNEURIN-BERNY**

Chargée de recherche CNRS, IAB, Grenoble (Membre)



*To my family, my love and my friends !*



## REMERCIEMENTS

Tout d'abord, je tiens à remercier mes directeurs de thèse André Verdel et Daphné Berny. Merci pour tout ce que vous m'avez appris, merci pour votre soutien tout au long de ma thèse, merci de m'avoir donné une chance de poursuivre mes rêves. Vous êtes deux « très vieux » chercheurs qui ont beaucoup à apprendre les jeunes comme moi (au passage, merci d'avoir soutenu mes blagues !).

André, merci de m'avoir accueilli dans ton équipe. Je te remercie pour ton investissement dans mon projet de thèse, pour tes conseils et pour les nombreuses discussions qu'on a eues et qui ont été certainement au cœur du progrès de ce travail ! Merci de m'avoir appris à poser toujours les bonnes questions et à bien réfléchir à mes contrôles ;)

Daphné, merci pour tout ce que tu m'as appris au laboratoire, merci pour toutes les heures que tu as passées ensemble avec moi à discuter sur le projet. Je te remercie sincèrement de ton dévouement exceptionnel ! Merci d'avoir toujours été aussi gentille avec moi, de m'avoir soutenue dans les moments difficiles et d'avoir été à mes côtés à chaque étape. C'est sûr que notre travail ensemble (et te taquiner tout le temps) va beaucoup me manquer !

Je tiens à remercier toutes les personnes merveilleuses qui travaillent dans cette équipe - Claire V, Solenne, Daniel, Leila, Claire K, Auriane ! Merci pour tout ce que vous m'avez apporté, merci pour votre soutien, merci pour tous les rires que nous avons partagés, pour tous les croissants, pour les coups de champagne et tous les gâteaux que nous avons mangés ensemble !

Claire, je te remercie tout particulièrement de m'avoir montré pour la première fois ce qu'est l'épigénétique et merci de m'avoir transmis ta passion pour le domaine. C'est grâce à toi que je me suis retrouvée dans cette équipe et je te remercie. Merci d'avoir toujours été à mes côtés, de m'avoir beaucoup fait rire et d'avoir partagé avec moi des verres d'alcool de rose ;) Merci aussi pour toutes les heures que tu as passées avec moi au microscope.

Leila, merci beaucoup pour tes conseils qui m'ont certainement beaucoup apporté et pour ton soutien tout au long de ma thèse. Merci pour ton travail lié à mon projet de thèse et d'avoir toujours pris le temps pour répondre à mes questions et m'expliquer les choses.

Solenne, j'ai sûrement beaucoup rigolé avec toi ! Merci pour tout ce que tu m'as appris lorsque j'étais en M1 et notamment les techniques de microscopie et comment ne pas casser un microscope ;) Merci pour tes conseils et merci pour ton soutien.

Auriane, nous nous sommes bien amusés au laboratoire et en dehors des heures de travail ! Tu es le bébé de l'équipe, et pourtant j'ai hâte de venir à ta soutenance et de voir ton progrès.

Je tiens à remercier tous mes amis, ceux qui sont en France et ceux qui étaient loin mais toujours disponibles et à mes côtés ! Je remercie tous les étudiants de l'IAB- les « anciens » et les « nouveaux ». Merci à vous Mathieu, Mina, Keerthi, Tristan, Naghemh, Mathilde, Daphné, Amaris, Paul, Alex, Dorothy pour votre joie de vivre, pour tous les moments festifs qu'on a partagés ensemble et surtout pour avoir fait de ma thèse une période inoubliable.

Un merci très spécial de tout mon cœur à mes merveilleux parents et à toute ma famille ! Vous m'avez manqué chaque jour, mais vous avez toujours été dans mon cœur et vous m'avez soutenue (et supportée) tout au long de mon parcours malgré les distances. Je suis sûre que derrière mon entêtement et ma passion pour la science, il y a beaucoup de mécanismes (épi)génétiques et épitranscriptomiques. Merci pour votre amour et pour tout ce que vous m'avez donné !

Enfin, le dernier remerciement mais sûrement le plus touchant pour moi – merci à toi mon chéri ! Merci pour tout ce que tu es, pour ton soutien, pour tes encouragements, pour ton aide, pour ton sourire et pour tous les moments joyeux qu'on a passé ensemble. Merci pour ton investissement notamment pour la préparation de mes présentations et pour la mise en page de mon manuscrit ;) Tu as été ma véritable motivation chaque jour et sans toi je ne serais pas ici aujourd'hui.

Kalina

# Table of Contents

<b>Abbreviations.....</b>	<b>10</b>
<b>Abstract .....</b>	<b>14</b>
<b>Résumé.....</b>	<b>16</b>
<b>Introduction .....</b>	<b>18</b>
<b>Chapter I: Gene expression regulation .....</b>	<b>20</b>
I.I Genome organization .....	22
I.I.1 General introduction to chromatin.....	22
I.I.2 Euchromatin and heterochromatin .....	24
I.II Overview of gene expression regulation .....	26
I.II.1 Main mechanisms of gene expression regulation .....	26
I.II.2 Gene expression regulation through ncRNAs.....	40
I.III Gene expression regulation through RNA modifications .....	44
I.III.1 The epitranscriptome and the RNA modification m6A .....	44
I.III.2 Molecular players of m6A .....	46
I.III.3 Roles of m6A in gene expression regulation .....	51
I.III.4 Biological functions of m6A.....	59
<b>Chapter II: The Heat Shock Response (HSR) .....</b>	<b>63</b>
II.I General overview of the HSR .....	65
II.I.1 Cellular stress responses and the HSR .....	65
II.I.2 Discovery and evolutionary conservation of the HSR .....	68
II.I.3 HSR at the cellular level .....	69
II.II Gene expression regulation during HSR .....	71
II.II.1 General overview of gene expression regulation during HSR.....	71
II.II.2 Mechanisms of gene expression regulation during HSR .....	73
II.II.3 The Heat Shock Factor 1 (HSF1).....	84
II.III The transcriptional protein-coding response to HS.....	88
II.III.1 The Heat shock proteins (HSPs).....	88
II.III.2 HSPs in cellular homeostasis and disease .....	91
II.III.3 Other HS-induced protein-coding genes .....	93
II.III.4 Classification of the HS-induced and HS-repressed genes.....	95
II.IV The transcriptional non-coding response to HS .....	97



II.IV.1 Overview of the non-coding response to HS.....	97
II.IV.2 HSR, <i>SATIII</i> lncRNAs transcription and nuclear Stress Bodies formation.....	98
<b>Objectives of the PhD project .....</b>	<b>105</b>
<b>Materials and methods.....</b>	<b>109</b>
I. Cell culture and cell transfections.....	111
I.I Cell lines, culture conditions and heat stress treatment.....	111
I.II Transient transfections .....	111
I.II.1 siRNA transfection .....	111
I.II.2 DNA plasmid transfection .....	112
I.II.3 Rescue experiments .....	113
I.III Stable transfection.....	113
I.III.1 Stable cell lines establishment by CRISPR-Cas9.....	113
I.III.2 Rescue experiments in the stable cell lines .....	115
II. Subcellular fractionation.....	116
III. DNA constructs and mutagenesis .....	117
III.I DNA constructs for expression of YTH-domain GFP-fused proteins.....	117
III.II Mutagenesis of YTHDC1-coding siRNA non-degradable constructs.....	117
IV. Protein analyses.....	118
IV.I Western blot.....	118
IV.II Immunoprecipitation (IP).....	119
V. RNA analyses.....	121
V.I Total RNA extraction.....	121
V.II Reverse Transcription (RT) for cDNA synthesis .....	121
V.III Polymerase chain reaction (PCR) .....	122
V.IV Quantitative PCR (qPCR) .....	123
VI. Microscopy techniques.....	123
VI.I Immunofluorescence .....	123
VI.II Fluorescence in situ hybridization (FISH) .....	124
VI.II.1 RNA FISH .....	124
VI.II.2 DNA FISH .....	126
VI.III Images acquisition and processing.....	126
VII. Chromatin immunoprecipitation .....	127
VIII. Cell viability and cell cycle assays .....	128
VIII.I Cell growth curves .....	128
VIII.II Propidium iodide staining.....	129

VIII.III Assessment of apoptosis.....	129
VIII.IV Fluorescence-activated cell sorting (FACS) and data processing .....	130
IX. Antibodies and oligonucleotide sequences used in this study.....	131
Table 1: Antibodies used in this study .....	131
Table 2: Oligonucleotide sequences used in this study .....	133
<b>Results.....</b>	<b>137</b>
<b>Chapter I: YTHDC1 relocalization and role at nuclear Stress Bodies (nSBs).....</b>	<b>139</b>
I.I Characterization of YTHDC1 relocalization to nSBs .....	141
I.I.1 Localization of YTH-domain proteins in response to heat shock .....	141
I.I.2 YTHDC1 relocalizes to nSBs during the recovery period following heat shock .....	144
I.I.3 HSF1-dependent <i>SATIII</i> transcription is required for YTHDC1 relocalization .....	147
I.I.4 YTHDC1 relocalization to nSBs could be m6A-independent.....	149
I.II Role of YTHDC1 at nuclear Stress Bodies.....	153
I.II.1 Interactomes of YTHDC1 and <i>SATIII</i> lncRNAs.....	153
I.II.2 YTHDC1 and SATIII-dependent splicing events .....	158
<b>Chapter II: YTHDC1 and gene expression regulation in response to HS .....</b>	<b>162</b>
II.I HS induces YTHDC1 recruitment to heat responsive genes .....	164
II.II YTHDC1 regulates HSPs expression .....	170
II.III YTHDC1 regulation of HSPs could be m6A-dependent.....	175
II.IV Role of YTHDC1 in <i>HSP</i> mRNA processing and export .....	176
II.V YTHDC1 maintains Nuclear Speckles structure upon HS .....	186
<b>Chapter III: New cellular functions of YTHDC1 outside the stress context .....</b>	<b>192</b>
III.I YTHDC1 KD induces cell mortality through apoptosis .....	194
III.II YTHDC1 KD affects cell cycle progression .....	198
<b>Discussion and perspectives.....</b>	<b>202</b>
I. Regulation of cellular stress responses through the epitranscriptome ?.....	204
II. Heat shock recovery: YTHDC1 at nuclear Stress Bodies .....	205
II.I YTHDC1 relocalization to nSBs.....	205
II.II YTHDC1 and the stress-induced regulation of alternative splicing .....	207
III. YTHDC1 and the expression of stress-responsive genes .....	209
III.I YTHDC1 genomic targets following heat stress .....	209
III.II YTHDC1 is essential for the induction of HSPs upon HS .....	211
III.III YTHDC1 and the stability of Nuclear Speckles following stress .....	215
IV. YTHDC1 as regulator of fundamental cellular processes .....	217

V. m6A-dependent YTHDC1 functions? .....	219
<b>Conclusion</b> .....	<b>221</b>
<b>Bibliography</b> .....	<b>223</b>

# Abbreviations

8-oxo-G: 8-oxoguanosine

ac<sup>4</sup>C: N4-acetylcytidine

ALKBH5: Alk B homolog 5

APA: Alternative PolyAdenylation

aslncRNA: antisense long non-coding RNA

BET: Bromodomain Extra-Terminal domain

BRG1: Brahma-Related Gene 1

carRNAs: chromosome-associated regulatory RNAs

CBC: Cap-Binding Complex

CDK: Cyclin-Dependent Kinase

CFI/CFII: Cleavage Factors I and II

ChIP: Chromatin ImmunoPrecipitation

circRNAs: circular RNAs

CPA: Cleavage and PolyAdenylation complex

CPSF: Cleavage and Polyadenylation Specificity Factor

CREBP: cAMP Response Element-Binding Protein

CstF: Cleavage stimulation Factor

CTD: C-Terminal Domain

DBD: DNA-binding Domain

DNA: DeoxyriboNucleic Acid

DNMT: DNA MethylTransferase

DSB: Double Strand Break

DSBR: Double-Strand Break Repair

dsDNA: double-stranded DNA

eIF4: eukaryotic Initiation Factor 4

EJC: Exon Junction Complex

eRNAs: enhancer RNAs

ESE: Exon Splicing Enhancer

ESS: Exonic Splicing Silencer

FACS: Fluorescence-Activated Cell Sorting

FACT: FACilitate Chromatin Transcription

FISH: Fluorescence *In Situ* Hybridization

FMRP : Fragile X Mental Retardation Protein

FTO: FaT mass and Obesity associated factor

GB: Gene Body

GFP: Green Fluorescent Protein

GTF: General Transcription Factor

H3 or H4: Histone 3 or Histone 4

H3.3: Histone variant 3.3

H3K9Ac: Histone 3 Lysine 9 Acetylation

H3K9me3: Histone 3 Lysine 9 trimethylation

H4K20me3: Histone 4 Lysine 9 trimethylation

HAT: Histone AcetylTransferase

HDAC: Histone DeACetylase

HEK cells: Human Embryonic Kidney cells

hm<sup>5</sup>C: 2' -O-methyl-5-hydroxymethylcytidine

HMT: Histone MethylTransferase

hnRNPs: heterogeneous nuclear RiboNucleoProteins

HP1: Heterochromatin Protein-1

HR: Homologous Recombination

HS: Heat Shock

HSE: Heat Shock Element

HSF1: Heat Shock Factor 1

HSP: Heat Shock Proteins  
 HSR: Heat Shock Response  
 HSR-1: Heat Shock RNA 1  
 IF: ImmunoFluorescence  
 IGF2BP: Insulin-like Growth Factor 2 mRNA-Binding Proteins  
 IgG: Immunoglobulin G  
 IRES: Internal Ribosome Entry Site  
 IP: ImmunoPrecipitation  
 KD: Knock-Down  
 KO: Knock-Out  
 LC-MS: Liquid Chromatography-Mass Spectrometry  
 lincRNAs: long intergenic non-coding RNAs  
 lncRNAs: long non-coding RNAs  
 m<sup>1</sup>A: N<sup>1</sup>-methyladenosine  
 m<sup>5</sup>C: 5-methylcytosine  
 m<sup>6</sup>A: N<sup>6</sup>-methyladenosine  
 m<sup>6</sup>A<sub>m</sub>: N<sup>6</sup>, 2'-*O*-dimethyladenosine  
 m<sup>7</sup>G: 7-MethylGuanosine  
 m<sup>7</sup>G: N<sup>7</sup>-methylguanosine  
 MALAT1: Metastasis Associated Lung Adenocarcinoma Transcript 1  
 MEFs: Mouse Embryonic Fibroblasts  
 METTL: METHylTransferase Like  
 miRNAs: microRNAs  
 mRNA: messenger RNA  
 ncRNA: non-coding RNA  
 NEAT1: Nuclear Enriched Abundant Transcript 1  
 NEXT: Nuclear EXosome Targeting complex  
 NFE2L2: Nuclear Factor Erythroid 2-Like 2  
 NHS: Non Heat Shock  
 N<sub>m</sub>: 2'-*O*-methyl  
 NMD: Nonsense-Mediated mRNA Decay  
 NPC: Nuclear Pore Complex  
 nSBs: nuclear Stress Bodies  
 NSs: Nuclear Speckles  
 NTR: Nuclear Transport Receptor  
 P53: Tumor Protein 53  
 PABP: Poly-A-Binding Protein  
 PABPN1: Poly(A) Binding Protein Nuclear 1  
 PAP: Poly-A Polymerase  
 PARP: Poly (ADP-Ribose) Polymerase  
 PAS: PolyAdenylation Signal  
 PIC: Pre-Initiation Complex  
 piRNA: Piwi-interacting RNA  
 PRO-seq: Precision nuclear Run-On sequencing  
 pTEFb: positive Transcription Elongation Factor  
 PTM: PostTranslational Modification  
 R: Recovery period  
 RBP: RNA-Binding Proteins  
 RD: Regulatory Domain  
 RNA Pol II: RNA Polymerase II  
 RNA: RiboNucleic Acid  
 RNMT: RNA guanine-N7 MethylTransferase  
 rRNAs: ribosomal RNAs  
 SAFB: Scaffold Attachment Factor B  
 SAM: S-Adenosyl-Methionine  
 SGs: Stress Granules  
 SINE: Short Interspersed Nuclear Elements  
 siRNAs: small interference RNAs  
 snRNAs: small non-coding RNAs

snoRNAs: small nucleolar RNAs  
snRNAs: small nuclear RNAs  
snRNPs: small nuclear RiboNucleoProteins  
SR: Serine/Arginine-rich protein  
SRF: Serum Response Factor  
SRSF: Serine/Arginine-rich Splicing Factor  
TAD: TrAnsactivating Domain  
TD: Trimerization Domain  
TERRA: Telomeric Repeat containing RNA  
TES: Transcription End Site  
TF: Transcription Factor  
TREX: TRanscription/EXport complex  
TRM: Trimerization Repression Motif  
tRNAs: transfer RNAs  
TSS: Transcription Start Site  
UTR: UnTranslated Region  
WT: Wild Type  
WTAP: Wilms Tumor 1 Associated Protein  
XIST: X-Inactivation Specific Transcript  
YTHDC: YTH Domain-Containing  
YTHDF: YTH Domain Family



# Abstract

Over the years numerous studies unraveled the fascinating ability of the Heat Shock Response (HSR) to act on the genome and to transiently control the cell proteome in order to preserve cell viability in the context of hostile environments. Nevertheless, the implementation and orchestration of the HSR today still hide mysteries. The implication of N<sup>6</sup>-methyladenosine (m6A), the most abundant internal RNA modification, in the regulation of gene expression in response to stress has been recently uncovered. Yet, its importance, the associated functions and the involved mechanisms remain poorly characterized.

During my PhD thesis I have studied the role of the protein YTHDC1, the main known nuclear reader of m6A, in the conserved cellular stress response to heat shock (HS). Therefore, my thesis work is at the crossroad of two main research axes- the emerging field of the epitranscriptome and the broadly investigated domain of stress biology. In this work, we were able to identify the central role played by YTHDC1, and by extension of the m6A signaling, in the regulation of the HSR pathway in human cells. First, we have demonstrated that during the recovery period following HS YTHDC1 massively relocalizes to characteristic nuclear structures appearing upon stress, named nuclear Stress Bodies (nSBs). This striking relocalization relies on the transcription of repetitive pericentric heterochromatin regions that produces the lncRNAs *SATIII*, which are at the core of nSBs formation. Importantly, we found out that beyond localizing at nSBs, YTHDC1, together with *SATIII* lncRNAs, contributes to the stress-induced regulation of alternative splicing events of potentially several hundreds of mRNAs. Second, by conducting ChIP-seq experiments we discovered that, upon HS, YTHDC1 is recruited to new genomic sites. Noticeably, we identified many Heat Shock Protein (HSP)-coding genes to be targets of YTHDC1 upon stress. Follow-up analyses revealed that YTHDC1 is essential for the induction of HSPs expression following HS. The investigation of YTHDC1 molecular functions showed that, under heat stress, the protein may regulate proper transcription termination of *HSP* genes and the following nuclear export of their mRNAs. Moreover, we found that, in response to stress, YTHDC1 is required to maintain the structure of Nuclear Speckles, which are membrane-less bodies contributing to the control of mRNAs fate. Finally, m6A-dependency of YTHDC1 molecular and cellular functions was addressed in this work through the establishment of a genetically manipulated cellular model. This approach demonstrated that some of the roles of YTHDC1 that we identified are associated with its ability to recognize the m6A RNA mark.

In conclusion, this work has uncovered the epitranscriptomic reader YTHDC1 as a novel and central regulator of the HSR acting at various levels of the stress-induced genome reprogramming in human cells. Our findings expand the current knowledge on the implementation of stress responses with the addition of a broadly acting layer of gene expression control and open new area of research for future studies.





# Résumé

Au fil des années, de nombreuses études ont mis en évidence la capacité fascinante de la réponse au stress thermique à agir sur le génome et à contrôler de façon transitoire le protéome de la cellule afin de préserver la viabilité cellulaire dans un contexte environnemental défavorable. Néanmoins, la mise en œuvre et l'orchestration de cette réponse cellulaire cachent encore aujourd'hui des mystères. L'implication de N<sup>6</sup>-méthyladénosine (m6A), la modification interne de l'ARN la plus abondante, dans la régulation de l'expression des gènes en réponse au stress a été récemment découverte. Cependant, son importance, les fonctions associées et les mécanismes impliqués restent pas ou mal caractérisés.

Au cours de ma thèse, j'ai étudié le rôle de la protéine YTHDC1, le principal lecteur nucléaire connu de m6A, dans la voie conservée de réponse cellulaire au choc thermique. Par conséquent, mon travail est à la croisée de deux axes de recherche principaux : le domaine émergent de l'épitranscriptome et le domaine déjà largement étudié de la biologie du stress. Lors de ce travail, nous avons pu identifier le rôle central joué par YTHDC1, et par extension la signalisation m6A, dans la régulation de la réponse au stress thermique dans les cellules humaines. Tout d'abord, nous avons démontré que pendant la période de récupération, qui suit le choc thermique, YTHDC1 se relocalise dans des structures de stress caractéristiques et nucléaires, appelées « nuclear Stress Bodies » (nSB). Cette relocalisation repose sur la transcription de régions répétitives de l'hétérochromatine péricentrique qui produisent des ARN longs et non-codants, *SATIII*, qui sont au cœur de la formation des nSBs. De manière intéressante, nous avons découvert qu'en se localisant au niveau des nSBs, YTHDC1, tout comme les ARN *SATIII*, contribue à la régulation de l'épissage alternatif de potentiellement plusieurs centaines d'ARNm. Deuxièmement, en menant des expériences de CHIP-seq, nous avons découvert que suite au stress thermique, YTHDC1 est recrutée sur des nouveaux sites génomiques. Nous avons notamment identifié de nombreux gènes codant pour les protéines de choc thermique (Heat Shock Proteins, HSP) comme étant des cibles de YTHDC1 en réponse à un stress thermique. Les analyses effectuées ont révélé que YTHDC1 est essentielle à l'induction de l'expression des HSP après un choc thermique. L'étude des fonctions moléculaires de YTHDC1 a montré que, en cas de stress, la protéine régulerait la terminaison de la transcription des gènes *HSP* et l'export nucléaire de leurs ARNm. De plus, nous avons découvert, toujours en réponse au stress, que YTHDC1 est nécessaire pour maintenir la structure des Nuclear Speckles, qui sont des corps nucléaires contribuant au contrôle du métabolisme des ARNm. Enfin, l'importance de m6A pour les fonctions moléculaires et cellulaires de YTHDC1 a été abordée dans ce travail par l'établissement d'un modèle cellulaire modifié génétiquement. Cette approche a démontré que certains des rôles de YTHDC1 que nous avons identifiés sont liés à sa capacité à reconnaître la marque m6A.

En conclusion, ce travail a permis de découvrir qu'un acteur de l'épitranscriptome, la protéine YTHDC1, est un nouveau régulateur central de la réponse au choc thermique agissant à différents niveaux de la reprogrammation du génome induite par le stress. Nos résultats élargissent les connaissances actuelles sur la mise en œuvre de la voie de réponse au stress thermique par l'ajout d'un nouveau niveau de contrôle de l'expression génique et ouvrent de nombreuses opportunités pour des futures études.



# **Introduction**



# **Chapter I**

## **Gene expression regulation**



## **I.I Genome organization**

### **I.I.1 General introduction to chromatin**

The hereditary biological information of organisms is stored under the form of a macromolecule- DNA (DeoxyriboNucleic Acid). This molecule, keeper of the essential genetic material, is composed of nucleotide sequences, arranged in double helix (Watson and Crick, 1953).

In humans each cell contains 2 meters of DNA kept in the relatively small nucleus of approximately 10 $\mu$ m diameter. The space limitations are resolved through the electrostatic attraction and interaction of the negatively charged DNA sequence with basic, globular proteins called histones. The association of DNA and histones forms the chromatin (Kornberg, 1977). The basic unit, termed the nucleosome, is composed of a 147bp DNA sequence wrapped around an octamer of histones, which consists of two copies of the histone proteins H2A, H2B, H3, and H4. Nucleosomes are connected by linker DNA (20-80bp) to form repetitive motifs of ~200bp referred to as 10nm “beads on a string” fiber (Olins and Olins, 2003). The latter characterizes the first level of DNA compaction and can be further condensed to chromatin fiber and to metaphase chromosome, the highest level of compaction (Fig.1A).

The formation of chromatin allows not only the packing of the DNA material into the cell nucleus but the control of gene expression. The constitution of nucleosomes modulates the DNA accessibility and interaction with transcription, replication, recombination and repair machineries. Therefore the chromatin structure can either promote or hinder diverse DNA processes. The chromatin pattern is governed by dynamic covalent and reversible DNA and histone modifications.

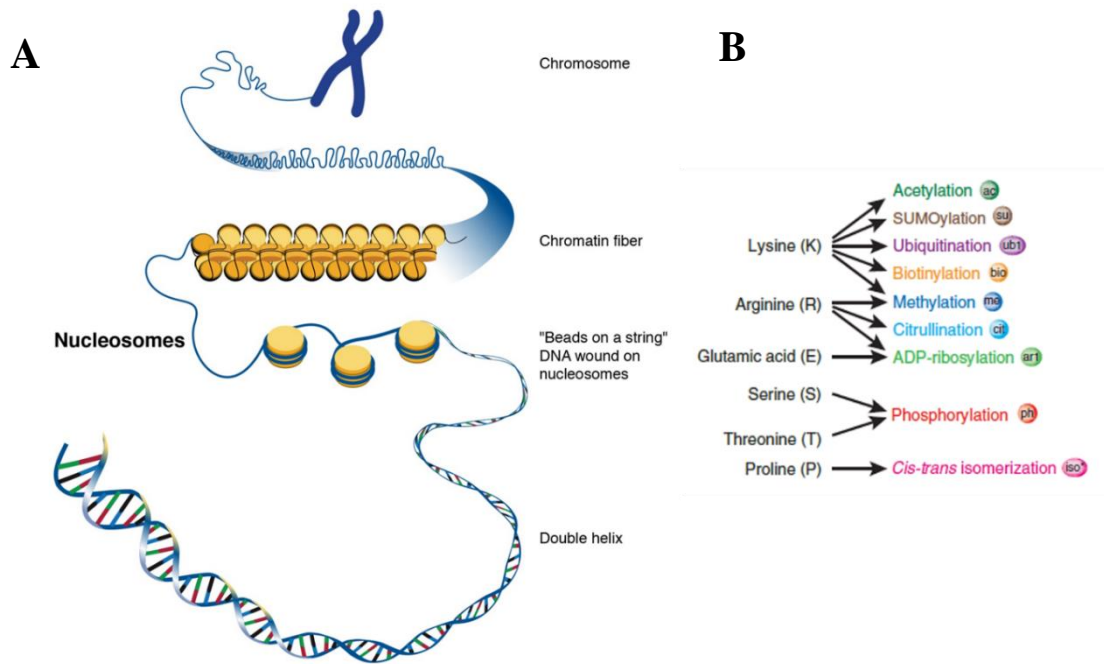
In vertebrates one of the most represented DNA modifications is methylation of cytosine residues at the level of short CpG-rich DNA stretches (or CpG islands) near promoter regions. DNA methylation plays an important role in gene expression regulation by recruiting proteins involved in gene repression or by inhibiting the binding of transcription factors to DNA (Moore et al., 2013).

The post-transcriptional modifications of histones reside predominantly in their *N*-terminal tail which is extended from the surface of nucleosomes and is accessible for interaction with various enzymes (Latham and Dent, 2007). The diversity of histone marks comprises dozens of different chemical modifications (Tan et al., 2011). The most representative include acetylation, methylation, ubiquitylation, sumoylation and biotinylation of lysines, methylation of arginines, phosphorylation



of serines and threonines as well as ADP ribosylation of glutamic acids, deamination of arginines and isomerization of prolines (Fig.1B). These various chemical modifications, which can moreover influence the deposition of one another, establish the “histone code” (Strahl and Allis, 2000).

The proteins responsible for the deposition of these histone marks are called “writers”, those who are able to remove them are the “erasers”, and the “readers” are the proteins which specifically recognize and interact with the modifications.



**Figure 1: Organization of chromatin and dynamic remodeling through diverse histone modifications.**

**A.** Structure and different states of compaction of chromatin- from the double helix DNA, through nucleosomes, “beads on a string” and chromatin fiber to the metaphase chromosome, the highest level of compaction. **B.** The main histone modifications found on different amino acid residues within the *N*-terminal region of histone proteins. *Adapted from Latham and Dent, 2007.*

The two best characterized modifications are acetylation and methylation. Specific enzymes, called Histone AcetylTransferases (HAT) acetylate lysine residues and are associated with chromatin decompaction while Histone DeAcetylases (HDAC) are the “erasers” of acetylation and promote tighter binding of DNA to histones. Histone MethylTransferases (HMT) carry out methylation which either support or inhibit transcription, depending on the targeted histone residue. Moreover a functional crosstalk between lysine methyltransferase families further mediating

chromatin compaction (Mozzetta et al., 2015). The histone demethylases can remove the methylation mark and counteract the HMTs (Kouzarides, 2007).

The *N*-terminal histone modifications are the sites of recruitment of different “readers”-proteins and complexes which can anchor to the histone and additionally induce the deposition or removal of chemical marks or serve as platforms for the binding of other regulatory proteins. The bromodomain-containing proteins recognize the acetylated residues and render DNA more accessible (Taniguchi, 2016). The chromodomain-containing proteins bind methylated lysine residues. The Heterochromatin Protein-1 (HP1), for example, interacts with the di- or trimethylated H3K9 (H3K9me<sub>2-3</sub>) and supports the compact and closed DNA packaging.

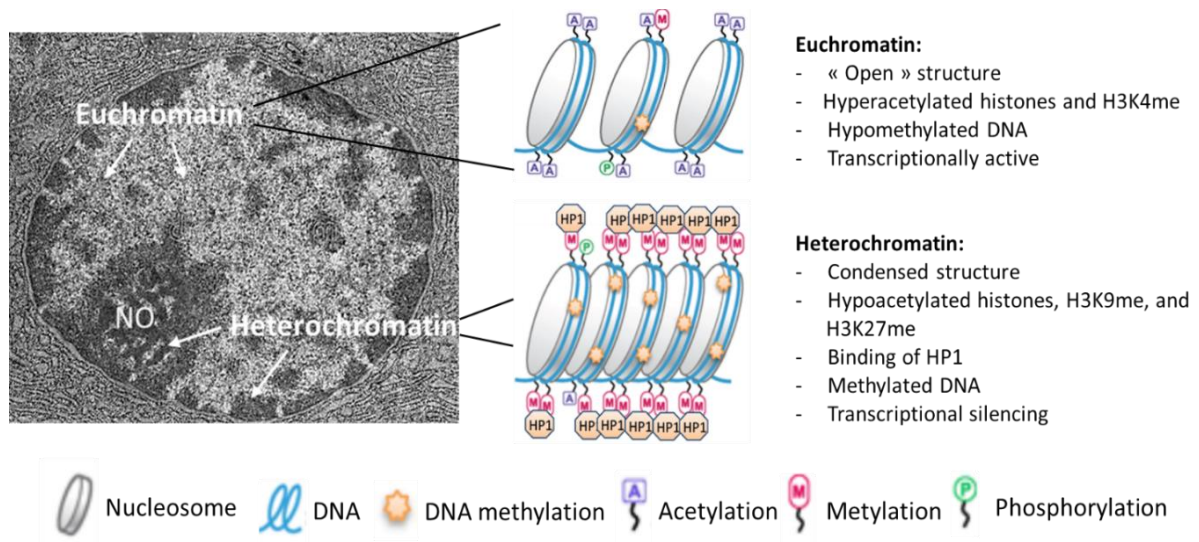
Moreover, the chromatin dynamics are modulated by histone chaperones and chromatin remodeling complexes which control the assembly, disassembly, displacement or removal of nucleosomes. For instance, the complex FACT (FACilitate Chromatin Transcription) is a histone chaperone which binds the H2A/H2B octamers and affects positively the transcription process (Gurova et al., 2018). The histone chaperones can also drive the incorporation of histone variants which can further influence the chromatin compaction at particular genomic regions (Li and Reinberg, 2011). The chromatin remodeling complexes can relocate nucleosomes through an ATP-dependent activity. One such complex is SWI/SNF who binds H3K14, induces the displacement of nucleosomes and promotes a more open chromatin conformation (Clapier and Cairns, 2009).

The tightly orchestrated combination between the DNA methylation landscape and the type of histone modifications at particular genomic regions describes two chromatin states: euchromatin and heterochromatin.

### **I.I.2 Euchromatin and heterochromatin**

The two major chromatin states, present in every eukaryote studied so far, are distinguished by their condensation degree, protein composition and also by their sub-nuclear localization. Heterochromatin is the highly packed, dense chromatin state and is found near the nuclear envelope and around the nucleolus while euchromatin is the “open” chromatin state and is dispersed in the nucleus (Fig.2). Another major difference between the two chromatin conformations is their DNA composition. The heterochromatin contains mainly repeated sequences and very low levels of protein-coding genes. Conversely, the euchromatin comprises relatively low degree of repeated sequences and is enriched for protein-coding genes.

The chromatin conformations differ moreover in the abundance and nature of their DNA and histone modifications which imply further distinct functional implications. The euchromatin is characterized by the presence of unmethylated DNA, the hyperacetylation of H3 and H4 and the trimethylation of H3K4 which govern the relaxed euchromatin conformation (Fig.2). On the other hand the inaccessible heterochromatin presents hypoacetylated H3 and H4 lysines, trimethylated H3 (K9 or K27 residues) and H4K20, and high DNA methylation levels (Fig.2). Therefore the euchromatin is considered to be the transcriptionally active form while heterochromatin is



**Figure 2: Chromatin organization in eukaryotes into two distinct forms: euchromatin and heterochromatin.** On the left is presented a transmission electron microscopy image of the nucleus of a plasma cell from hamster’s bone marrow (*from Fawcett, 1981*). On the right are schematically represented the characteristic DNA, histone marks and “readers” of euchromatin and heterochromatin.

associated with gene silencing and plays a general role in the genome stability (Strålfors and Ekwall, 2011).

Heterochromatin is furthermore subdivided into facultative and constitutive heterochromatin. The facultative form is found at genomic regions whose expression is limited to specific stages of development and cellular differentiation (Grewal and Jia, 2007). This type of heterochromatin allows the repression of genes in particular cellular contexts- environmental and developmental. For example in females facultative heterochromatin covers one of the two X chromosomes and promotes its inactivation as to assure the genetic equality with males (Barakat and Gribnau, 2012).

The constitutive heterochromatin is located at specific genomic loci- telomeres, centromeres and pericentromeres and is composed mainly of repeated sequences and for long time was taught to be constantly silenced. In addition, the constitutive heterochromatin is abundant in DNA mobile elements, named transposons which present a threat to the genome stability (Konkel and Batzer, 2010). Hence, one of the main roles of this form of heterochromatin is to maintain the transposons and repetitive sequences in a transcriptionally inactive state in order to preserve the genome stability. Importantly, recent studies demonstrated that constitutive heterochromatin harbors much more plasticity than previously assumed. There is an increasing evidence for active and regulated transcription of pericentromeres and telomeres in different physiopathological conditions such as proliferation, development, senescence, stress response and transformation during tumorigenesis. This particular activation is linked to specific and perhaps conserved functions of the transcribed ncRNAs (non-coding RNAs). The roles of some specific ncRNAs produced from constitutive heterochromatin regions in response to heat shock will be described in the next chapter.

Importantly, heterochromatin is maintained and persists after cellular divisions through epigenetic mechanisms. The broad field of epigenetics stand for the study of heritable gene expression variations which are not associated with changes of the DNA sequence (Bonasio et al., 2010). The epigenome comprises the DNA and histone modifications along with their specific “writers”, “erasers” and “readers”. Once established the epigenome can be modulated to control gene expression, the genome stability and all fundamental cellular processes according to the physiological, pathological or environmental context.

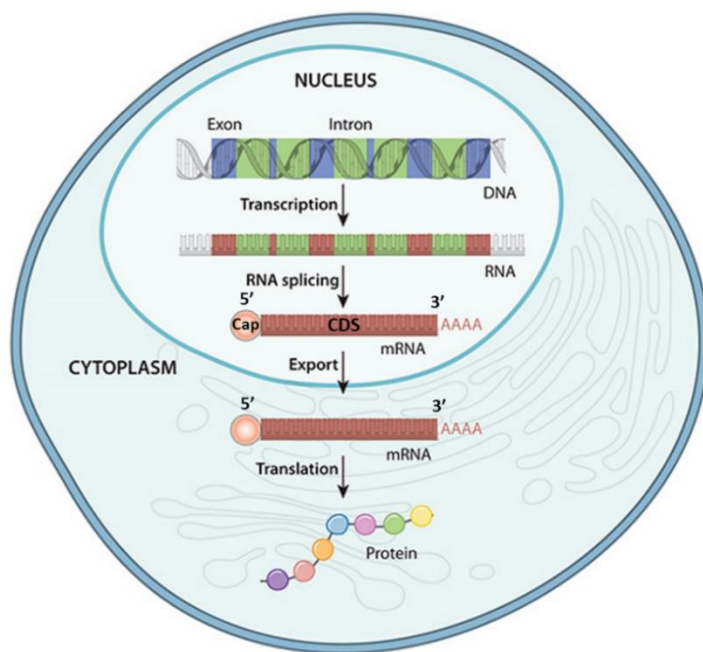
## **I.II Overview of gene expression regulation**

### **I.II.1 Main mechanisms of gene expression regulation**

The central dogma of molecular biology postulates that the genetic information flows from DNA to RNA and then to proteins. Described in this way the process appears to be straightforward. However, in the last several decades an extensive amount of research clearly depicted that the regulation of mammalian gene expression is more complicated than imagined initially.

Today we know that gene expression regulation occurs at various levels - transcriptional, co-transcriptional, post-transcriptional and translational levels, and involves complex

interconnected pathways. In order to be expressed a messenger RNAs (mRNA) needs to undergo several steps of maturation (capping, splicing and polyadenylation) prior to its nuclear export and its translation in the cytoplasm (Fig.3). The proper temporal and functional execution of all steps from transcription initiation to translation is critical to allow proper gene expression, and subsequently allow proper cell identity during development, to maintain it through lifetime, to coordinate cellular activity and to confer cell survival in variable and potentially stressful conditions. These major aspects of gene expression are not surprisingly often found deregulated in disease.



**Figure 3: Main steps of expression of a protein-coding gene.** Gene expression includes the transcription of coding (exon) and non-coding (intron) DNA sequences into RNA; the maturation of RNA through the addition of a cap at the 5'-end, splicing to remove the introns and express the coding DNA sequence (CDS) and polyadenylation at the 3'-end, the export from nucleus to cytoplasm and the final step- the translation into functional protein.

In the scope of my PhD work, we have investigated more in details some of the mechanisms of gene expression, notably transcription termination and polyadenylation, alternative splicing, transit through nuclear structures and nuclear export of particular mRNAs in normal and stress conditions. In this part of the introduction the main steps of expression of protein-coding genes, especially in mammals and in normal growing conditions, will be presented. Later on, in the second chapter, the modulation and execution of these steps specifically in the context of stress will be described.

### Transcription of protein-coding genes

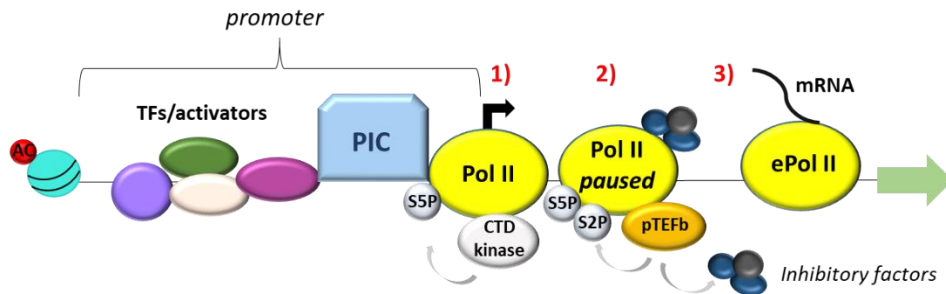
The process of transcription and its regulation have been the topic of important amount of functional and molecular studies. The transcription machinery in eukaryotes comprises three structurally very similar RNA Polymerases (RNA Pol) - I, II and III, which produce a large variety of RNAs (Cramer et al., 2008). Based on their nature RNAs can be divided into two main categories- coding and non-coding RNAs. The coding or messenger RNAs represent 1% to 2% of the total amount of RNAs expressed at any time in a cell and are the only one to be translated into proteins. The rest of the RNAs are non-coding RNAs (ncRNAs) which can have diverse regulatory roles. The RNA Pol I transcribes ribosomal RNAs (rRNAs) which are the essential components of ribosomes, while the RNA Pol III produces the transfer RNAs (tRNAs) needed for the translation of mRNAs. The RNA Pol II is the sole polymerase transcribing mRNAs but also a diversity of ncRNAs. Here, the mechanisms of action and the regulation of the activity of RNA Pol II will be presented.

The transcription process, initiated by any of the eukaryotic RNA Polymerases, comprises three main steps: initiation, elongation and termination which are tightly controlled by *cis*-acting and *trans*-acting elements. Regardless of the level of gene expression regulation the *cis*-acting elements are DNA or RNA sequences while the *trans*-acting elements are proteins which can influence the regulation of unlinked genes and RNAs via an interaction with the *cis*-elements (Louis and Sagarsky, 2018).

A transcriptionally active gene in mammalian cells is controlled by the interaction of transcription factors (TFs) with the promoter region, extending to approximately 1000bp upstream from the Transcription Start Site (TSS). The TFs can activate or block the recruitment of RNA Pol II to the TSS and are therefore key orchestrators of gene expression. The General TFs (GTFs) recognize conserved sequence motifs (e.g. TATA box) within the proximal TSS region and promote the formation of the Pre-Initiation Complex (PIC) (Sikorski and Buratowski, 2009). When formed, the PIC recruits the RNA Pol II in order to initiate transcription.

Gene expression can be constitutive or inducible, ubiquitous or cell-specific and is governed by a complex network of gene selective TFs. These TFs act as master regulators to switch ON or OFF gene expression accordingly to the perceived extra- and intracellular signals. Some of them operate on “open” promoters and regulate the constitutive transcription of housekeeping genes (Cairns, 2009). Others are present in cells in a latent form and specific internal stimuli

(activation of receptors, ROS, DNA damage, protein aggregation, etc.) are able to promote their activation. Examples of such primary TFs are HSF1, SRF, CREB, SMAD, p53 and many nuclear receptors. The classification of TFs can be further subdivided to transcriptional co-activators and co-repressors. The combination of DNA-binding TFs and co-activators/ co-repressors afford flexibility and precise control of both the magnitude and the specificity of the transcriptional response.



**Figure 4: Gene expression regulation at the level of transcription initiation and elongation.** When the chromatin structure is decondensed (displaced/acetylated nucleosomes), transcription can take place. General Transcription Factors (GTFs) associate to form the Pre-Initiation Complex (PIC) in the promoter region. The PIC stimulates the recruitment of RNA Pol II at the level of the transcription start site (black arrow) (1). Specific Transcription Factors (TFs) and activators participate to the transcription initiation. CTD kinases recruited by GTFs, specific TFs and activators phosphorylate serine 5 (S5P) of the C-Terminal Domain (CTD) of RNA Pol II and further promote transcription initiation (1). At the majority of genes the polymerase pauses shortly after initiation and is maintained in the paused state by inhibitory factors. Positive Transcription Elongation Factor b (pTEFb) is recruited by specific TFs, phosphorylates serine 2 of the CTD and induces the dissociation of the inhibitory factors (2). RNA Pol II continues into the elongation phase (ePol II) and transcribes the gene sequence into mRNA (3).

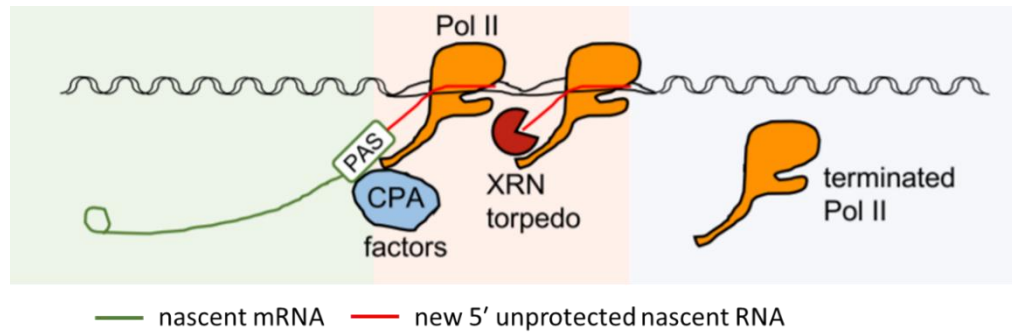
Another important player in the transcription control is the C-Terminal Domain (CTD) of RNA Pol II. During preinitiation Cyclin-Dependent Kinases (CDK) are recruited by the basal TFs, phosphorylate the CTD Ser5 (serine 5 residues) and allow the transcription initiation (Fig.4). However, 30% of the protein-coding genes and 70% of the developmental or inducible genes harbor a paused RNA Pol II at their promoter-proximal site (Adelman and Lis, 2012). It was demonstrated that after transcribing 10- 120 nucleotides RNA Pol II pauses and requires the activity of pTEFb or CDK9 (positive Transcription Elongation Factor b or Cyclin-Dependent Kinase 9) in order to continue into the elongation step (Chen et al., 2018). The paused polymerase state is

sustained by several inhibitory factors. Once recruited to the RNA Pol II, pTEFb phosphorylates Ser2 of the CTD, promotes the dissociation of the repressive factors and allows the release of the polymerase into the elongation phase (Fig.4). pTEFb is itself subjected to tight regulation and is maintained in an inactive complex in normal conditions and its activation is often dependent on the presence of specific TFs involved in the expression of inducible or developmental genes (Paparidis et al., 2017).

Furthermore, a broadly examined step of the process is the termination of transcription. It is well described in the literature that termination is closely coupled to the co-transcriptional 3'-end mRNA processing and consists in the recognition of the PolyAdenylation Signal (PAS), the recruitment of the 3'-end Cleavage and PolyAdenylation (CPA) complex, the cleavage and release of the nascent transcript at the PAS and the subsequent release of the RNA Pol II machinery. The CPA complex interacts initially with the RNA Pol II CTD and then assembles onto the PAS of the nascent transcript. Two models for transcription termination can be proposed- the allosteric model and the so called torpedo model (Proudfoot, 2016). The allosteric model refers to termination via conformational changes in the transcription complex (Epshtein et al., 2007). The process described by the torpedo model relates to the continuous synthesis of the RNA Pol II after the cleavage at the PAS and the consequent production of a new 5' unprotected nascent RNA (Fig.5). In this model the torpedo is XRN2 (5'-3' exoribonuclease 2), which is recruited to the PAS, progressively degrades the downstream transcript in a kinetic competition with the RNA Pol II and acts as a molecular trigger for the release of the polymerase from the gene body (Fig.5) (Crisp et al., 2018). Growing evidence suggests that these two models act in a cooperative manner and unify in one mechanism of transcription termination related to both conformational changes of the machinery and continuous transcription after the PAS (Eaton et al., 2020).

The termination of transcription is subjected to a quality control check. Its premature execution results in the production of truncated transcripts that are detected as such and rapidly degraded (Kamieniarz-Gdula and Proudfoot, 2019). On the other hand, improper cleavage and polyadenylation, and the subsequent abolished transcription termination lead to the production of readthrough transcripts resulting from the continuous transcription of downstream DNA sequence. Readthrough transcription can deregulate the expression of downstream genes and can endanger the genome stability (Oren et al., 2014).





**Figure 5: “Torpedo” model for transcription termination.** The Cleavage and PolyAdenylation (CPA) complex assembles on the PolyAdenylation Signal (PAS) of the nascent mRNA (green). The RNA Pol II continues to transcribe downstream from the PAS and produces a new 5' unprotected RNA. The 5'-3' exonuclease 2 is recruited to the PAS, degrades the new unprotected RNA and releases the RNA Pol II. *Adapted from Crisp et al., 2018.*

#### Co-transcriptional and Post-transcriptional mechanisms of gene expression

The posttranscriptional regulation of gene expression involves the closely linked mechanisms of mRNA processing and mRNA export. In order to be correctly expressed and later translated into a functional protein, a mRNA needs to undergo the steps of capping, splicing and polyadenylation which are regulated by various *cis*- and *trans*-acting elements. It is of note that the term posttranslational designates a functional rather than temporal or location distinction between transcription and mRNA maturation since many of the described processes take place at the site of transcription before the transcription termination and are broadly interconnected.

The first step, the **capping**, consists in the addition of a N<sup>7</sup>-methylguanosine cap at the 5' UnTranslated Region (5'UTR) of the nascent transcript. The process occurs early in transcription initiation and requires the activity of Capping Enzymes (CE) and an RNA guanine-N<sup>7</sup> methyltransferase (RNMT). The capping process presents a main checkpoint in the production of mature coding RNAs. The function of the cap is to stabilize and protect nascent transcripts from degradation by 5'-3' exonucleases. In the nucleus, the structure of the cap is recognized by Cap-Binding Complex (CBC). CBC as well as the capping enzymes interact with a spectrum of mRNA-processing factors and can affect gene expression at several levels- transcription, mRNA splicing, polyadenylation, stability and localization (Gonatopoulos-Pournatzis and Cowling, 2014). In the cytoplasm the cap is recognized by the translation factor eIF4 (eukaryotic Initiation Factor 4) and

the interaction is decisive for mRNA translation or decay. Many examples have emerged recently to indicate that capping is regulated by different signaling pathways and thus it can adapt mRNA fate to the precise context (Kachaev et al., 2020).

A second processing event is the **splicing** of mRNAs which is needed to remove the non-coding transcript sequences, the introns, from the pre-mRNA and to fuse contiguous coding regions, the exons. If splicing proceeds improperly the Nonsense-Mediated mRNA Decay (NMD) surveillance pathway is activated in order to degrade the aberrant transcripts and prevent the production of truncated proteins (Zheng, 2016).

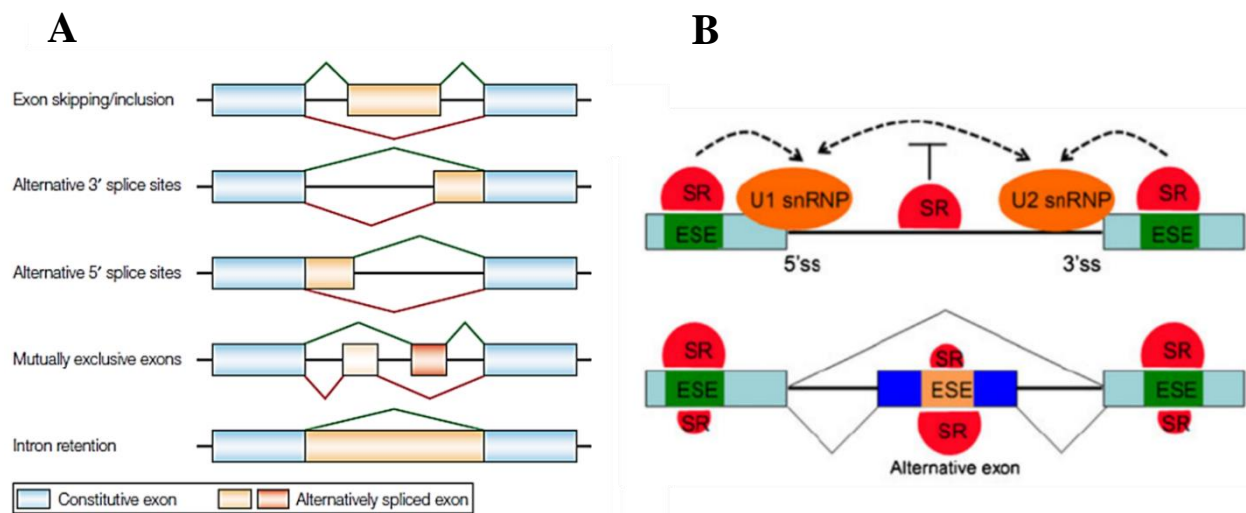
The splicing reaction requires the recognition of conserved signals at the ends of introns by the macromolecular complex named the spliceosome. The spliceosome is assembled in a sequential mode by more than 150 different proteins and comprises snRNPs (small nuclear RiboNucleoProteins), hnRNPs (heterogeneous nuclear RiboNucleoProteins), SR (Serine/Arginine-rich) proteins and many cofactors (Fica and Nagai, 2017). The snRNPs catalyze the transesterification reaction of splicing. The spliceosome is the central player in both constitutive and alternative splicing. While constitutive splicing describes the process of exon retention and intron elimination, alternative splicing refers to the potential differential splicing of exons and introns. Alternative splicing is the main mechanism able to give rise to a variety of proteins derived from a single gene and is thereby a crucial gene expression regulator. In humans, more than 90% of transcripts undergo alternative splicing (Wang et al., 2008).

In addition to the consensus splice sites, exons and introns contain elements which can either enhance (exonic/intronic splicing enhancers) or repress (exonic/intronic splicing silencers) the use of the nearby consensus splice sites.

The differential selection of these additional sites can result in exon skipping/inclusion, alternative 5' or 3' splice site usage, mutually exclusive exons or intron retention (Fig.6A) (Cartegni et al., 2002). The most common mechanism is the exon skipping. The selection relies on the “strength” of the site defined by the affinity and competition of *trans*-elements. The SR and hnRNP proteins are the most prominent *trans*-acting elements involved in alternative splicing.

When SR proteins recognize a splicing site they induce the recruitment of snRNPs. The mechanisms of action of SR proteins are complex and their roles can be opposite depending on the recognized site. The binding of SR proteins to an Exon Splicing Enhancer (ESE) stimulates the

recognition of the nearby 5' and 3' splicing sites by the snRNPs U1 and U2, and results in exon inclusion (Fig.6B). However, the interaction of SR proteins with intronic regions inhibit splicing and leads to exon skipping (Zhou and Fu, 2013). If SR proteins bind to an alternative exon they promote its inclusion while recognition of the flanking competing exons causes exon skipping (Fig.6B) (Pandit et al., 2008). The roles of hnRNP proteins are relatively less understood. Some hnRNP can bind ESS (Exonic Splicing Silencer) sequences and inhibit exon splicing (Rothrock et al., 2005). HnRNP binding to splicing silencers occur frequently and have been found to influence constitutive and alternative splicing events throughout the human genome.



**Figure 6: Common modes of alternative splicing and regulation of alternative splicing by the SR proteins.** **A.** Presented are the five common modes of alternative splicing. Constitutive and alternative exons are indicated. Green lines show the retention of exon/intron in the mature mRNA, red lines indicate splicing and elimination of exon/intron. **B.** Upper panel: when SR proteins bind Exon Splicing Enhancer (ESE), they stimulate the recruitment of snRNPs U1/U2 to nearby splicing sites and promote exon inclusion. The interaction of SR proteins with intronic regions leads to exon skipping. Lower panel: SR proteins bound to alternative exon can promote either its inclusion or skipping depending on the competition between the sites. *Adapted from Cartegni et al., 2002; Zhou and Fu, 2013*

Notably, SR proteins and hnRNPs are also modular and their activity can be influenced by posttranslational modifications driven by various signaling pathways. In the case of SR proteins, the most influencing modification is their reversible phosphorylation at the C-terminal arginine/serine-rich (RS) domain. The phosphorylation of SR proteins regulates their localization and activity, and therefore governs their differential biological roles. Among the growing list of SR

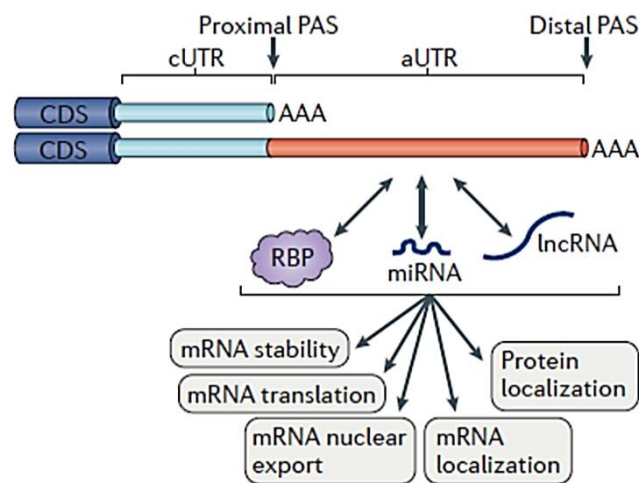
protein kinases, in eukaryotes the SRPK and CLK/STY families are best characterized (Zhong et al., 2009). During development and in response to stress, these specific kinases are (auto)-regulated and consequently tightly control and modulate the SR activity (Long et al., 2019). Such mechanism of regulation of splicing via an SR kinase and the phosphorylation state of SR proteins in the context of heat shock response will be thoroughly presented in the next chapter.

It is worth mentioning that splicing and transcription are largely interconnected and influence one another. Approximately 80% of pre-mRNAs are spliced before the transcription termination (Carrillo Oesterreich et al., 2010). The RNA Pol II elongation rate is critical for the kinetics of assembly of the spliceosome and influences the splice site selection. A fast movement of RNA Pol II stimulates strong splice site selection while slower transcription induces weak splicing site selection.

The **polyadenylation** of mRNAs is another process influencing gene expression. The poly-A tail is not encoded by the genome and requires the activity of specific factors. The core polyadenylation machinery, the Cleavage and PolyAdenylation (CPA) complex, comprises the Cleavage and Polyadenylation Specificity Factor (CPSF), the Cleavage stimulation Factor (CstF), the Cleavage Factors I and II (CFI and CFII) and the Poly-A Polymerase (PAP). CPSF functions as mRNA 3'-end-processing endonuclease and recognizes the PolyAdenylation Signal (PAS), a canonical sequence in the 3'UTR located 15-30nt upstream from the cleavage site. Once this sequence transcribed, the CPA is transferred from the CTD of the RNA Pol II to the nascent transcript. As mentioned previously this step of RNA maturation is functionally and physically linked to the transcription termination and the two mechanisms are largely interdependent. Following the cleavage, the PAP is recruited and drives the addition of approximately 200 adenines to the 3'terminus in an ATP-dependent manner. The polyA tail serves as a signal for the nuclear export, translation and stability of the mRNA. It is recognized by Poly-A-Binding Protein (PABP). On one hand PABP protects the mRNA from nuclease attack and extends its half-life and on the other hand it interacts with the 5'cap via eIF4 and enables ribosomal cycling and translation (Colgan and Manley, 1997).

Alternative PolyAdenylation (APA) is the mechanisms of production of diverse mRNA isoforms when multiple polyadenylation sites are present in a gene. Up to 70% of human genes have alternative polyadenylation sites which reside in the 3'UTR or in the coding region of

transcripts (Wang et al., 2008). When present in the coding region, upon selection, an alternative PAS generates protein isoforms with distinct C-terminal regions. The alternative sites differ in their relative distance from the start of the gene- a proximal PAS is closer to the TSS and the distal site is situated downstream from proximal APA sites (Fig.7). Lengthening or shortening of the 3'UTR modifies the *cis*-elements composition and further alters the interaction of the mRNA with RNA-binding proteins (RBPs), miRNAs and lncRNAs involved in the regulation of translation, export, stability and localization of mRNAs or the localization of the produced proteins (Fig.7) (Tian and Manley, 2017). Therefore, APA is decisive for mRNA fate and protein biogenesis.



**Figure 7: General mechanism of Alternative PolyAdenylation (APA).** A mRNA with two PolyAdenylation Sites (PAS)- one proximal, closer to the Transcription Start Site (TSS) and one distal-downstream from the proximal site. The selection of the distal PAS leads to lengthening of the 3'UTR region and increased abundance of *cis*-acting elements which can be recognized by RNA-Binding Proteins (RBP), miRNAs or lncRNAs. Consequently mRNA stability, translation, nuclear export, localization and the localization of the produced proteins can be affected. *Adapted from Tan and Manley, 2017.*

A primary determinant for PAS selection is its sequence identity to the canonical sequence and ability to recruit the polyadenylation machinery. Selection between alternative PAS is dependent on the level of core factors which can be either increased or reduced according to the activation of various signaling pathways. Furthermore, auxiliary factors can either inhibit the assembly of the polyadenylation machinery at canonical PAS or facilitate its recruitment to alternative sites. Among the factors influencing the polyadenylation pattern are some multifunctional RNA-binding proteins and mainly splicing factors (hnRNPs, snRNPs and SR

proteins). The interactions between elongation factors, splicing factors and CPA components are necessary for the recognition of the PAS, the definition of the last exon, the adapted cleavage, polyadenylation and transcription termination (Tellier et al., 2020).

### mRNA export

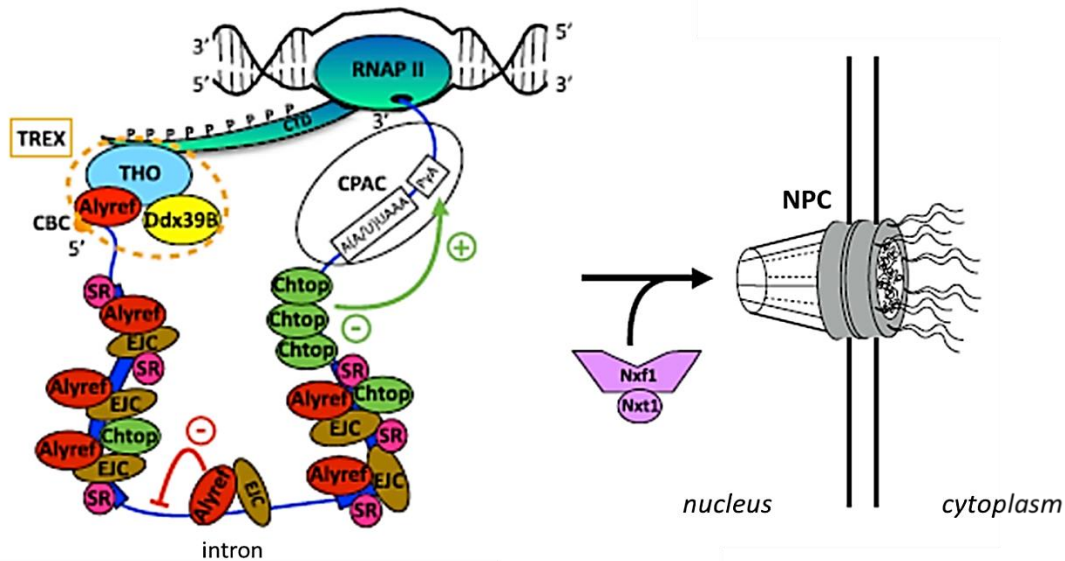
The export of mRNAs is a key stage of gene expression since it presents a major quality control checkpoint, crucial for the subsequent production of appropriately functioning proteins. The passage of mRNA through the Nuclear Pore Complex (NPC) is assisted by Nuclear Transport Receptors (NTRs) - NXF1/NXT1 or Tap/p15. The interaction requires the intermediate action of mRNA-Binding Proteins (RBPs) which associate on one end with the transcript and on the other end with the NTRs. The assembly of RBPs is essential for the selective export of properly matured mRNAs. The conserved multiprotein complex TREX (TRanscription/EXport complex) is the main orchestrator of mRNA export in mammals. It is composed of the sub-complex THO, the export adaptor Aly/Ref, the RNA helicase DDX39 and several co-adaptors (Katahira, 2012). The latest studies indicate that TREX subunits are recruited progressively and in a co-transcriptional manner in order to tightly couple transcription, mRNA processing and mRNA export (Fig.8), (Viphakone et al., 2019).

THO is recruited early via an interaction with the diphosphorylated S2-S5 CTD of RNA Pol II (Fig.8). Upon splicing, the Exon Junction Complex (EJC) is deposited closely to exon-exon junction sites and further supports the association of TREX components. Aly/Ref is recruited near the 5' end by the cap-binding complex and is rapidly transferred along the RNA to exon-exon junction sites in an EJC-dependent manner. EJC and Aly/Ref can be found also on introns and mediate intron retention (Fig.8).

At the last exon of the nascent transcript CHTOP (member of THO) is recruited also via splicing factors and participates in APA by the regulation of alternative PAS selection (Fig.8). Many of the TREX subunits cooperate with splicing factors, notably SR proteins which act as adaptors in order to assure the proper co-transcriptional splicing and the concomitant loading of the export machinery (Huang et al., 2003).

If TREX loading is achieved, the last step is the transfer of the mRNA from TREX to NXF1/NXT1 and their subsequent passage through the NPC (Fig.8). On the other hand, the

incorrectly processed mRNAs are unable to associate correctly with TREX. They are retained in the nucleus and degraded by NEXT (Nuclear EXosome Targeting complex) (Fan et al., 2017).



**Figure 8: Model for the co-transcriptional association of TREX subunits on the nascent mRNA.** THO associates with the diphosphorylated CTD of RNA Pol II. Exon Junction Complex (EJC) binds exon-exon junction upon splicing. Aly/Ref is recruited along the mRNA by EJC and can inhibit the splicing of introns and thus induce intron retention. CHTOP is recruited also via splicing and mediates APA. CHTOP stimulates proximal PAS usage and decreases the distal PAS selection. When the TREX components are correctly loaded, the complex interacts with NXF1/NXT1 and the mRNA is transferred to cytoplasm through the Nuclear Pore Complex (NPC). *Adapted from Viphakone et al., 2019.*

### mRNA turnover

The regulation of mRNA maturation and degradation is a crucial level of gene expression. The exosome complex is both a 3' to 5' exo- and endo-ribonuclease and is the main regulator of cytoplasmic and nuclear RNA turnover in all eukaryotes studied to date. The exosome is associated with co-factors that are essential for its RNA processing or degradation action. These co-factors promote the recruitment and recognition of the exosome to its targets, the unfolding of RNA, and the poly(A) polymerization of the targets, allowing the anchoring of the RNA to the exosome (Kim et al., 2016; LaCava et al., 2005; Zinder and Lima, 2017).

In mammals, various cofactors of the nuclear exosome have been identified. Among them are found the nucleolar-localized hTRAMP (human TRf4/Air2/Mtr4p Polyadenylation) complex and the nucleoplasm-specific Nuclear EXosome Targeting (NEXT). The NEXT complex is

composed of hMTR4, the RNA Recognition Motif (RRM)-containing RBM7 and the zinc-knuckle ZCCHC8 proteins (Fan et al., 2017). These two exosome cofactors act on different RNA substrates related to their distinct nuclear localization. The hTRAMP assists the exosome in rRNA maturation, while NEXT subunits assist the exosome in the degradation of the ncRNAs PROMPTs (PROMoter uPstream Transcripts), 3'- extended products from U1/U2 snRNA- and Replication-Dependent Histone (RDH) genes (Zinder and Lima, 2017). Thus, the exosome is responsible for the maturation or elimination of redundant transcripts and for the maintenance of the steady-state levels of diverse RNA species.

Cytoplasmic mRNA turnover is initiated by polyA tail removal and proceeds via the 3'–5' exoribonuclease activity of the exosome or decapping followed by 5'–3' degradation by the exoribonuclease Xrn1 (XRN1) (Tuck et al., 2020). Shortening of the polyA tail and deadenylation is mediated by the poly(A)-nuclease deadenylation complex (PAN2/3) and the carbon catabolite repression 4- negative on TATA-less (CCR4-NOT) complex (Ashworth et al., 2019). A broad variety of RNA factors recruit the deadenylase complexes to their targets. Furthermore, in the cytoplasm RNA decay is extensively coupled to translation through surveillance pathways such as NMD (Nonsense-Mediated Decay), translational roadblocks, or no stop codon decay (Roy and Jacobson, 2013).

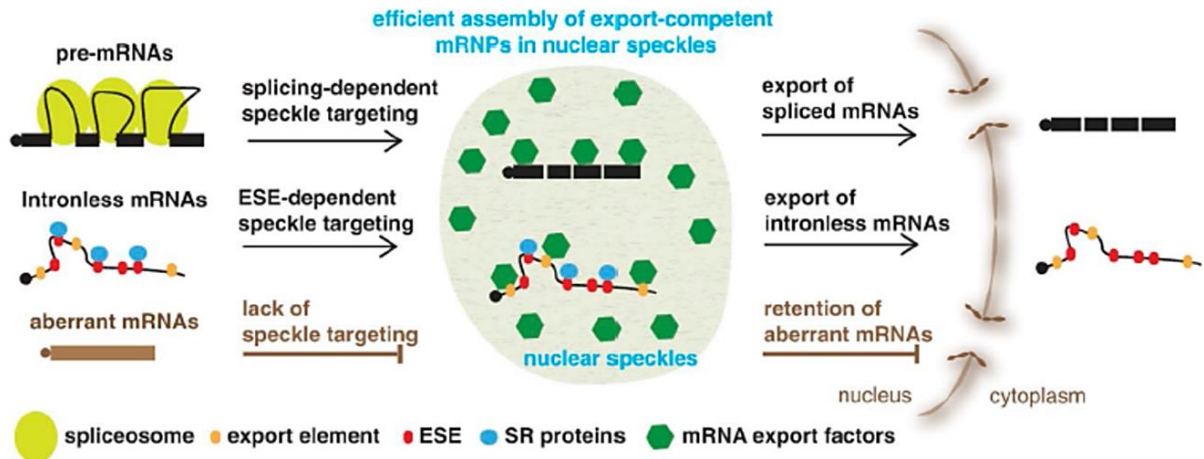
The equilibrium between transcription and RNA decay is constantly monitored through these various pathways and is vital for the function of the cell.

#### *Gene expression regulation at nuclear speckles*

Recent studies suggested that nuclear structures, termed Nuclear Speckles (NSs) facilitate the integrated regulation of gene expression in the nucleus. NSs are liquid droplet-like nuclear bodies, known also as splicing speckles or SC35 domains. They are enriched in various RNA processing and translation-associated factors (Kim et al., 2019) as well as proteins implicated in chromatin dynamics, splicing, 3'-end processing and export of mRNAs. These include notably the snRNPs, SR proteins and TREX subunits. NSs are presumably different from transcription factories since they are enriched in RNA Pol II elongation factors and relatively less initiation factors (Dow et al., 2010).



These nuclear compartments are considered to be central hubs for gene expression regulation where the coordination of the various broadly interconnected mRNA processing and export mechanisms can take place within spatially defined organization.



**Figure 9: Model for the role of nuclear speckles in gene expression regulation.** Nuclear Speckles (NSs) are enriched in various mRNA processing proteins, notably SR proteins and export factors. They present a quality control hub for splicing-dependent and intronless mRNA. The improperly matured, aberrant mRNAs are not targeted to NSs and are retained in the nucleus. *Adapted from Wang et al., 2018.*

Notably, recent studies demonstrated that in addition to splicing-dependent mRNAs, intronless mRNAs also associate with NSs and the interaction facilitates their nuclear export through enhanced loading of TREX (Wang et al., 2018). The intronless mRNAs are taught to be targeted to NSs through ESE sequences and the associated SR proteins (Fig.9). Hence, it was suggested that NSs are involved in the quality control of both splicing-dependent and intronless mRNAs. Conversely, aberrant mRNAs do not localize to nuclear speckles and are retained in the nucleus for degradation (Fig.9).

### Translational regulation

The translation of mRNAs is the final step of gene expression. The regulation of translation is a very flexible and quite selective process. The initiation step of translation involves many actors and therefore the molecular targets for modulating protein synthesis are numerous and can alter the expression from many to only few mRNAs. The activity of Initiation Factors (IFs) is often modified through phosphorylation/dephosphorylation reactions which adjust the translation efficiency. A

main player in many regulatory mechanisms is the cap binding protein eIF4 which can be bound by inhibitory proteins upon diverse internal and external stimuli and thus induce global translational repression (Lama and Verma, 2020). Ribosomes can be also subjected to posttranslational modifications which can influence the translation kinetics and/or productivity. In addition the translation relies on various *cis*-acting elements present on the mature transcript (Barreau et al. 2005). Different signals which alter the localization of mRNAs can further affect the translation process. Actively translating mRNAs are found at polysomes (ribosome factories). However, in certain conditions mRNAs can translocate to P-bodies or stress granules and either be degraded or stored for expression upon restoration of the normal growth conditions (Martin and Ephrussi, 2010)

In summary, all levels of gene expression regulation are broadly interconnected and implicate an important number of molecular actors with diverse functions that can often be regulated through various mechanisms. Mutations of the gene expression modulators and/or deregulations of the gene expression steps are associated with virtually all known human diseases. Unearthing the complex mechanisms of translation of the genotype to phenotype helps us understand the fundamental cellular machineries acting in gene expression, their impressive flexibility in response to distinct external and internal stimuli and their implication in pathology.

### **I.II.2 Gene expression regulation through ncRNAs**

In this work we have focused in the stress-induced activation of transcription of particular ncRNAs, termed *SATIII*. These non-coding RNAs (ncRNAs), which will be described in details in the second chapter are believed to be implicated in heterochromatin maintenance, the regulation of alternative splicing and have an important role in genome stability. Here, a more general description will be given on the main features of ncRNAs as well as some of their best characterized molecular functions, related to gene expression.

The realization that numerous RNAs are not translated into proteins was a major breakthrough of the last decades and defined a new category of gene expression regulators. In mammalian cells at least 75% of the genome produces RNAs and only 2% of them are coding – the mRNAs. The rest of the RNAs, transcribed by RNA Pol II, together with the rRNAs and tRNAs

transcribed respectively by RNA Pol I and RNA Pol III, constitute the population of ncRNAs (Andersson et al., 2014). The ncRNAs can be defined by their length, localization, and/or function. The lncRNAs are described as larger than 200 nucleotides while the small non-coding RNAs (sncRNAs) are less than 200 nucleotides.

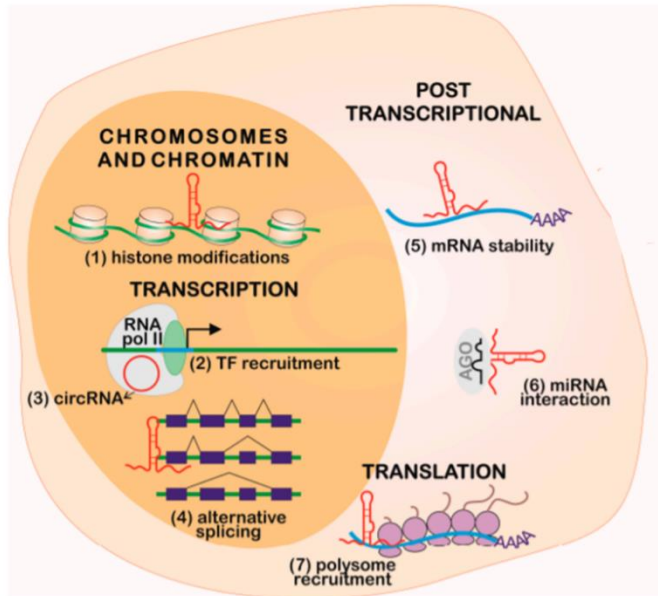
The lncRNAs can be subdivided also regarding their biogenesis into intergenic lncRNAs (lincRNAs), intronic lncRNAs, antisense lncRNAs (aslncRNAs), bidirectional lncRNAs and enhancer RNAs (eRNAs) (St. Laurent et al., 2015). The class of intronic lncRNAs includes the circular RNAs (circRNAs) which are produced via non-canonical splicing forming a closed circular loop of joined 3' and 5' RNA ends (Salzman et al., 2012). The eRNAs can be transcribed bidirectionally from enhancer regions. The lncRNAs can be also classified into functional and non-functional according to whether they present a specific role or not.

Most of the transcription steps leading to the production of lncRNAs are similar to the ones involved in the production of mRNAs. lncRNAs can be subjected to the different steps of mRNA processing. The mechanisms of action of regulatory lncRNAs rely on their cellular and temporal specificity and on the nature of their target. lncRNAs can associate with nucleic acids by base-pairing or can be recognized by RNA-binding proteins and thus can interact with various macromolecules. In general, lncRNAs are known to contribute to chromatin organization, transcription initiation and elongation, mRNA processing, export, and decay, translation and posttranslational regulation (Fig.10) (Fernandes et al., 2019).

The chromosome and chromatin structure can be modulated through interactions of lncRNAs with DNA by sequence complementarity or allocation in the helix. Representative examples of this function of lncRNAs include the X-inactivation specific transcript (*XIST*) (Rinn and Chang, 2012). The lncRNA *XIST* binds to one of the two X chromosomes, recruits silencing complexes and renders the chromosome transcriptionally inactive. Many other lncRNAs are found to regulate the general structure and the nuclear organization of the genome (Khalil et al., 2009).

At the transcriptional level the promoter regions of lncRNAs can act as enhancers and participate in the loading of the PIC. In addition, some lncRNAs can interact with the transcription machinery- circRNAs can bind the RNA Pol II and influence its processivity while eRNAs bind TFs and position them to particular promoters (Fig.10) (Li et al., 2015; Long et al., 2017).

Growing evidence unravels the functions of lncRNAs in mRNA processing. Recently the lncRNAs *MALAT1* and *NEAT1* have been observed respectively at the level of nuclear speckles and paraspeckles where they can have different roles in the coordination of mRNA maturation and export or the modulation of mRNA stability and degradation (Gonzalez et al., 2015).



**Figure 10: Roles of lncRNAs in gene expression regulation.** LncRNAs can regulate: chromatin structure via histone modifications (1); transcription initiation through TF recruitment (2) or binding of circRNAs to RNA Pol II (3); alternative splicing (4); mRNA stability (5); miRNA interaction with target sites (6); translation through polysomes recruitment (7). *Adapted from Fernandes et al., 2019*

Furthermore lncRNAs can either facilitate or repress the translation process (Fig.10). In some cases particular lncRNAs can promote the recruitment of polysomes to mRNAs and induce cap-independent translation while other participate in the recruitment of translational repressor and decrease the rate of newly synthesized proteins (Rashid et al., 2016).

Additionally, in a complex interplay lncRNAs can compete with miRNAs bound to argonaute proteins (AGO) for their target site and thereby influence RNA fate (Fig.10).

The class of small non-coding RNAs (sncRNAs) is also significantly involved in diverse molecular processes from transcription to translation. In human cells miRNAs are the most extensively studied group of sncRNAs. They are found to be the central actors of the regulation of mRNA stability and translation through interaction with the 3'UTR of mRNAs. When a miRNA binds its target site in the 3'UTR of an mRNA it recruits the deadenylation complex CCR4-NOT which subsequently leads to the elimination of the polyA tail but also to the decapping of the mRNA by auxiliary enzymes and thereby makes the mRNA susceptible to rapid degradation by exoribonucleases (Gebert and MacRae, 2019). Other sncRNAs include the ribosomal RNAs

(rRNAs), transfer RNAs (tRNAs), small nuclear RNAs (snRNAs), small nucleolar RNAs (snoRNAs) and Piwi-interacting RNAs (piRNAs), with diverse regulatory roles, all greatly implicated in the precise control of gene expression (Srijyothi et al., 2018; van Wolfswinkel and Ketting, 2010).

In general, the expression, stability and localization of ncRNAs can be tightly regulated and adjusted to the physiopathological contexts in order to contribute to the genome flexibility and genome reprogramming during development and in response to various hostile conditions. As mentioned previously, this also relates to the *SATIII* lncRNAs, which are produced specifically in response to heat shock. Their role in the modulation of gene expression upon stress will be further described.

## **I.III Gene expression regulation through RNA modifications**

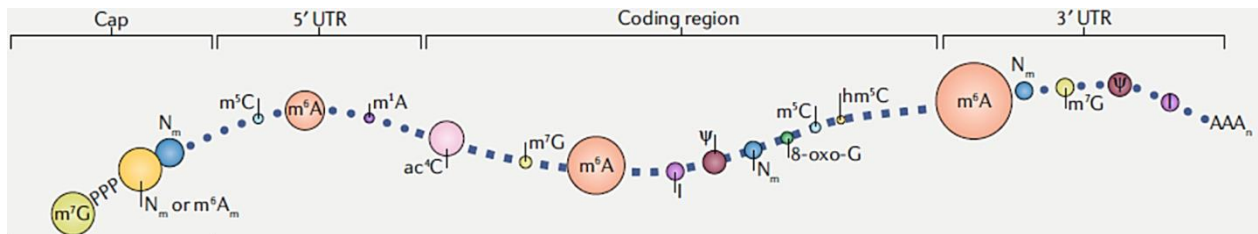
### **I.III.1 The epitranscriptome and the RNA modification m6A**

Lately the central dogma of molecular biology has been expanded and diversified by the realization that a broad range of RNA modifications are highly susceptible to control gene expression. The ensemble of RNA modifications describes the emerging field of the epitranscriptome. In the last decade the epitranscriptome has gained significant interest and a tremendous number of studies witnessed the key roles of RNA marks in virtually all steps of RNA metabolism.

Numerous RNA modifications have been identified during the last century. Nevertheless, the lack of analytical techniques limited the knowledge of their functions and RNA modifications were for long considered as constitutive, “basal” feature of RNA. The recent rapid innovations in epitranscriptomics allowed to link the presence of more than 150 different RNA modifications to various steps of gene expression (Boccaletto et al., 2018). The best characterized RNA mark to date is the methylation of adenosine residues on the sixth atom of nitrogen- N<sup>6</sup>-methyladenosine (m6A). This RNA modification has been detected for the first time in the 1970s (Desrosiers et al., 1974). However, four decades later, the discovery of the first specific m6A eraser reinforced the idea that RNA modifications can be reversible and dynamic with potential important implication in gene expression (Jia et al., 2011). The molecular players and the roles of m6A have been largely investigated ever since and will be thoroughly presented in the next parts.

Our current knowledge on m6A is due entirely to the development of novel biochemical approaches coupled with RNA sequencing which allowed the transcriptome-wide mapping of m6A on the level of single-nucleotide resolution (Linder et al., 2015). In the past few years through the application of innovative approaches the **main characteristics of m6A** have been successfully described: **1)** it presents more than 12, 000 sites in the transcripts of nearly 7, 000 coding and 300 non-coding human genes; **2)** its abundance is variable- from one to 20 m6A residues per mRNA; **3)** it resides within a consensus sequence- RRACH (R = purine, A = m6A, and H = A, C, or U); **4)** it is most strongly enriched in the 3’UTR region, around stop codons and at the level of long internal exons; **5)** it is selective and occurs in certain mRNAs (Dominianni et al., 2012; Linder et al., 2015; Meyer et al., 2012). Importantly, m6A is present in eukaryotes from yeast, plant, flies to mammals as well as in bacteria and virus strains (Deng et al., 2015; Fray and Simpson, 2015; Gokhale et al.,

2016; Schwartz et al., 2013). Even though m6A sites are broadly conserved between human and mouse, evolutionary differences in higher eukaryotes and in yeast have been discovered (Liu and Zhang, 2018; Ma et al., 2017). For instance, *Schizosaccharomyces pombe* and *Caenorhabditis elegans* are the sole yeast organisms to have evolved without m6A due to the lack of the core methylation components.



**Figure 11: Distribution and relative abundance of the currently mapped RNA modifications.** The different mRNA regions are represented (cap, 5'UTR, coding region and 3'UTR). The size of the circles indicates the relative abundance of each modification. The repertoire of mapped RNA modifications comprises N<sup>6</sup>-methyladenosine (m6A), N<sup>1</sup>-methyladenosine (m<sup>1</sup>A), 5-methylcytosine (m<sup>5</sup>C), 2'-O-methyl-5-hydroxymethylcytidine (hm<sup>5</sup>C), N4-acetylcytidine (ac<sup>4</sup>C), N<sup>7</sup>-methylguanosine (m<sup>7</sup>G), 2'-O-methyl (N<sub>m</sub>) or the doubly modified form – N<sup>6</sup>, 2'-O-dimethyladenosine (m6A<sub>m</sub>), 8-oxoguanosine (8-oxo-G), pseudouridine (Ψ) and Inosine (I). These modifications are dispersed along all regions of an mRNA with different abundance levels. *Adapted from Zaccara et al., 2019.*

Since the first studies on m6A, additional epitranscriptomic marks have been mapped mainly in mRNAs but also in ncRNAs. Their distribution along coding transcripts indicated that modified nucleotide bases can be dispersed in a differential manner along the cap, the 5'UTR, the coding region and the 3'UTR and can present distinct abundance levels (Fig.11) (Zaccara et al., 2019). In addition to the recent knowledge of the distribution of particular RNA modifications, the discovery of the enzymes responsible for the deposition and/or removal of some of these marks also allowed to unearth at least partially their implication in RNA metabolism. Today we know that pseudouridine (Ψ) regulates RNA structure and is involved in the translation of RNAs (Marianna Penzo et al., 2017). The 5-methylcytosine (m<sup>5</sup>C) mark is associated with RNA processing, mRNA stability, mRNA export and tRNA cleavage and translation (Trixl and Lusser, 2019). Another broadly studied RNA modification is N<sup>1</sup>-methyladenosine (m<sup>1</sup>A) was found to participate in translation and in the structural stability of tRNAs (Zhang and Jia, 2018). Increasing evidences show that the levels of many epitranscriptomic modifications vary during development,

stress responses and numerous pathological conditions. Thus, these novel regulators of gene expression are very likely essential for the integration and response to external and internal signals and could be driving factors in diverse human diseases (Kadumuri and Janga, 2018). Some of these nucleotide modifications are also present in mRNAs in prokaryotes, eukaryotes and viruses and can relate to largely conserved functional effects.

Moreover, a complex crosstalk between the RNA modifications was suggested. In analogy with the reversible epigenetic marks, RNA modifications establish an “epitranscriptomic code” and could influence the deposition, functions and stability of one another in order to provide fine-tuning of RNA fate in particular contexts.

The epitranscriptome arose very rapidly in the run of less than ten years and is now defined as an integral and central element in gene expression regulation. However, the prevalence, localization and roles of the majority of RNA modifications are still unknown and many intriguing questions need to be further addressed and deciphered in order to highlight the importance of this novel field.

### **I.III.2 Molecular players of m6A**

N<sup>6</sup>-methyladenosine is to date the only RNA modification with identified specific writers, erasers and readers regulating both its prevalence and location, and its implementation in different molecular and cellular functions. In the next section these specific m6A players, their mode of action and mechanisms of regulation will be presented with particular focus on the m6A reading process and the roles of the subject of my PhD thesis – the protein YTHDC1.

#### *Writers of m6A*

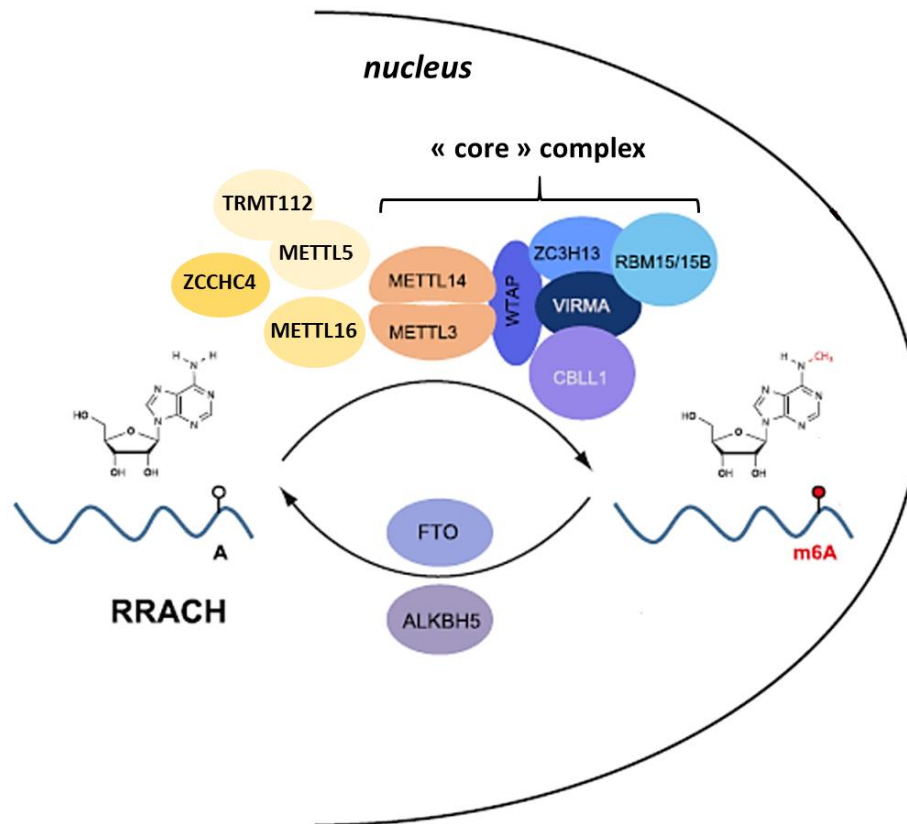
To date, four different nuclear methyltransferases are described to generate m6A RNA modifications. The majority of m6A residues in RNA Pol II-derived RNAs are primarily deposited by the heterodimer of METTL3 (METHylTransferase Like 3) and METTL14 (METHylTransferase Like 14), (Fig.12). METTL3 is the catalytically active component and METTL14 acts as an allosteric activator which also interacts with the RNA target (Wang et al., 2016). The enzymatic reaction consists in the transfer of a methyl group from the METTL3-bound S-Adenosyl-



Methionine (SAM) to the substrate. This methyltransferase “core complex” is assisted by other key subunits (Fig.12). The Wilms Tumor 1 Associated Protein (WTAP) is needed for the correct nuclear localization of the complex and maintains the interaction between METTL3 and METTL14 (Liu et al., 2014). The large protein VIRMA (originally KIAA1429) is considered also to have a regulatory role and promotes m6A installation particularly around stop codons and in the 3'UTR of mRNAs (Yue et al., 2018). The depletion of both WTAP and VIRMA leads to significant decrease of m6A total levels in human cells (Schwartz et al., 2014). The RNA binding proteins RBM15 and RBM15B interact with METTL3 and are suggested to assist its recruitment to consensus motifs (Patil et al., 2016). An important co-adaptor is the ZC3H13 which interacts with WTAP and RBM15/15B and coordinates the localization of the methyltransferases within the nucleus (Knuckles et al., 2018). Another potential constituent of the core methylation unit is the E3 ubiquitin ligase CBL1 (or HAKAI) whose precise role are yet unclear (Horiuchi et al., 2013).

The rest of the currently identified m6A writers are ZCCHC4, METTL5/ TRMT112 and METTL16 (Fig.12). They deposit the modification in a more restricted manner, to specific classes of RNAs. The ribosomal RNA (rRNA) methyltransferase ZCCHC4 forms the single m6A residue within the 28S ribosomal RNA (rRNA) (Ma et al., 2019). On the other hand the single m6A modification of the 18S rRNA is deposited by the most recently discovered complex of the catalytically active METTL5 (METHylTransferase Like 5) and the adaptor TRMT112 (van Tran et al., 2019). The finding of METTL16 (METHylTransferase Like 16) as specific m6A enzyme initially led to the hypothesis that it is one of the main m6A writers. Nevertheless, METTL16 turned out to drive the formation of m6A mostly on the U6 snRNA and the MAT2A mRNA, which encodes the enzyme for SAM biosynthesis, and on only a small number of other mRNAs and ncRNAs (Pendleton et al., 2017; Shima et al., 2017; Warda et al., 2017).

Importantly, the genomic deletion of METTL3 was found to result in the loss of approximately 70% of all m6A sites in mouse germ cells while the concomitant deletion of METTL3 and METTL14 leads to the loss of nearly 99% of m6A residues in polyadenylated RNAs in human cells (Geula et al., 2015; Lin et al., 2017). Taken together these results confirm the status of the “core complex” as central m6A writer.



**Figure 12: Writers and erasers of m6A.** m6A is deposited in the nucleus on a consensus sequence RRACH (R = purine, A = m6A, and H = A, C, or U). The writers “core complex” comprises the methyltransferases METTL3 and METTL14 (METHyltAnserase Like 3 and 14) together with their co-adaptors WTAP, VIRMA, RBM15/RBM15B and ZC3H13. Additional m6A methylases, acting on specific subsets of RNAs are ZCCHC4, METTL5 and its adaptor TRMT112 and METTL16. The mark can be erased by the nuclear demethylases FTO (FaT-mass and Obesity associated protein) or ALKBH5 (Alk B homolog 5). *Adapted from Coker et al., 2019.*

It has been demonstrated that virtually all METTL3-dependent m6A residues are deposited at the pre-mRNA level during transcription and mRNA processing (Ke et al., 2017). In the same line, METTL3 was detected bound to RNA Pol II suggesting that the writing complex could be recruited to nascent transcripts by the transcription machinery (Slobodin et al., 2017). In addition, the terminal exon-exon junction in mRNA is proposed to be a key structural feature for the enrichment of m6A particularly around stop codons (Ke et al., 2015).

One of the main open questions regarding m6A “writing” is the selectivity of the process. Only certain mRNAs contain the mark, not all potential m6A consensus sites present within a transcript are modified and m6A is restricted to certain regions of mRNAs. Since m6A occurs co-

transcriptionally, one possibility is that the transcription and mRNA-processing factors are needed for the targeted deposition of m6A early in the mRNA life cycle. However, the enigma of m6A selectivity needs further investigation to be resolved.

### Erasers of m6A

As mentioned previously, the reversible pattern of m6A has been revealed through the discovery of its erasers- the Fat mass and Obesity associated protein (FTO) and ALKB Homolog 5 (ALKBH5), (Fig.12), (Jia et al., 2011; Zheng et al., 2013).

The nuclear enzyme FTO, from the ALKB family, erases the m6A mark and generates *in vitro* and *in vivo* the intermediate N<sup>6</sup>-hydroxymethyladenosine (hm<sup>6</sup>A) whose functions have not been yet clarified (Fu et al., 2013). The knockdown of FTO leads to increased m6A peaks while its overexpression reduces them (Jia et al., 2011). However, the biological relevance of FTO *in vivo* is quite unclear yet since it can demethylate both DNA and RNA and preferentially erases the N<sup>7</sup>-methylguanosine cap (Mauer et al., 2017).

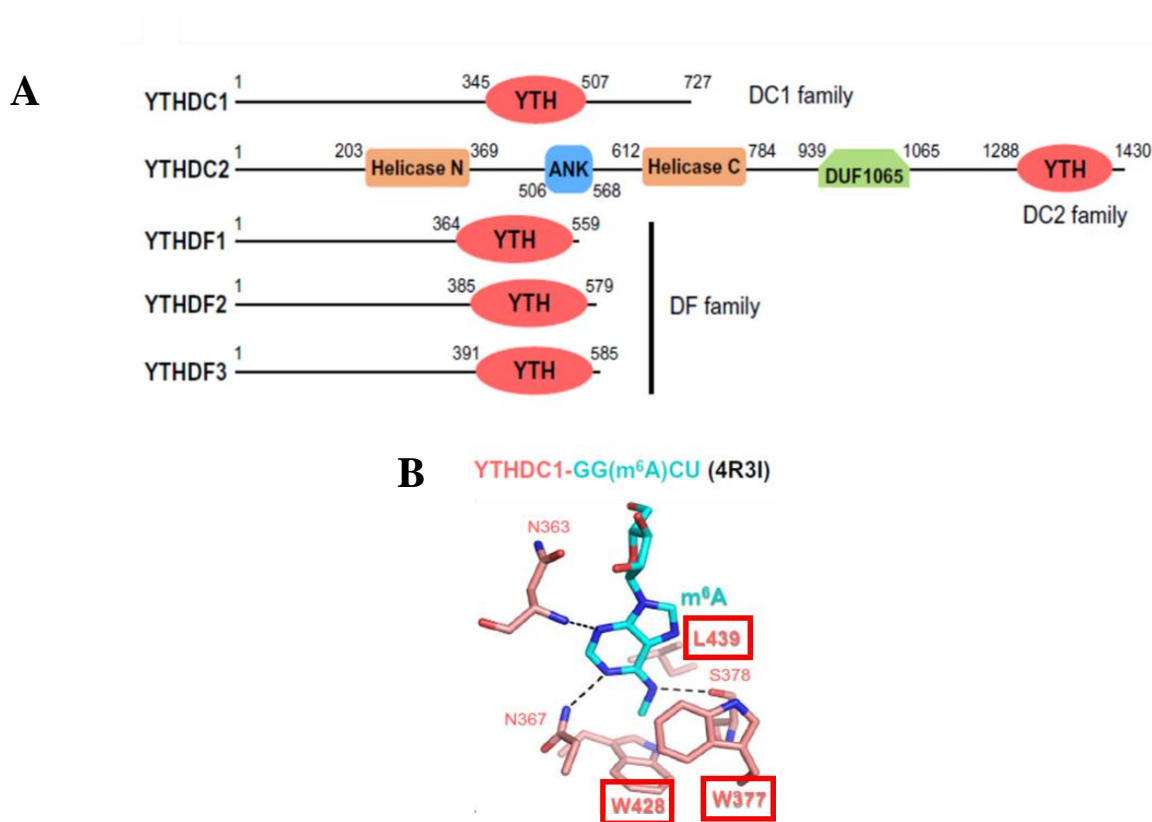
The nuclear demethylase ALKBH5 is believed to be a better candidate for m6A removal *in vivo*. Recently, loss of ALKBH5 was associated with reduced m6A abundance and deregulated molecular m6A functions (Tang et al., 2018). Nevertheless, the data on the role of ALKBH5 demethylation *in vivo* is limited to certain transcripts examined in specific cellular contexts.

The demethylation of m6A is less understood compared to its formation and has been often the subject of scientific debates. It is suggested that demethylation is functionally irrelevant for RNAs with relatively short half-life. In the same setting, m6A dynamics might be regulated via the control of its writers and their activity, rather than through the recognition by specific erasers. Future studies are needed to better understand the general mechanisms of regulation of m6A levels.

### Readers of m6A

The downstream molecular and cellular effects of m6A are conferred by its interaction with reader proteins that specifically recognize the modification and bind to it. The most well-known readers of m6A correspond to the members of the family of YTH domain-containing proteins. In humans there are five YTH domain-containing proteins which can be classified into three categories on the basis of their domain organization: **YTHDC1** (DC1 family), YTHDC2 (C2 family) and YTHDF1, YTHDF2 and YTHDF3 (DF family), (Fig.13A) (Liao et al., 2018). The

characteristic YTH domain is the module for recognition of m<sup>6</sup>A. The five YTH-domain proteins are conserved in mammalian genomes and the YTH domain is evolutionary conserved across eukaryotes. YTHDC1 is the only exclusively nuclear m<sup>6</sup>A reader from this family (Fig.13A). YTHDC2 is cytoplasmic and YTHDF proteins are cytoplasmic in particular in normal conditions. YTHDC2 has a different domain organization with an helicase domain, ankryn repeats and a DUF1065 domain (Fig.13A). The YTH domain-containing proteins are ubiquitous except for YTHDC2 which is found enriched in testes (Wojtas et al., 2017).



**Figure 13: The family of YTH domain-containing proteins and the structure of YTHDC1 with a bound m<sup>6</sup>A containing probe. A.** The YTH domain-containing family of proteins comprises five proteins in humans which are classified in three categories: YTHDC1 (DC1 family), YTHDC2 (C2 family) and YTHDF1, YTHDF2 and YTHDF3 (DF family). YTHDC1 is exclusively nuclear, YTHDC2 is cytoplasmic and YTHDF1-3 are cytoplasmic in normal conditions. YTHDC2 has an helicase domain, ankryn repeats and a DUF1065 domain. **B.** Interaction of YTHDC1 with an m<sup>6</sup>A-containing RNA probe (GG(m<sup>6</sup>A)CU). The essential residues for m<sup>6</sup>A recognition are W377, W428 and L439 situated in a hydrophobic cage in the YTH domain. *Adapted from Liao et al., 2018.*

The specific mode of recognition of m6A remained unknown until the studies of the structure of YTHDC1 YTH domain alone and together with an m6A containing RNA probe (GG(m6A)CU) (Xu et al., 2014; Xu et al., 2015). These works demonstrated that the bound m6A is buried within a hydrophobic cage of the YTH domain formed by three essential residues- W377, W428 and L439 in the case of human YTHDC1 (Fig.13B). Mutations of either W377 or W428 to alanine completely abolish the recognition of m6A emphasizing the importance of the hydrophobic cage for m6A binding. Notably, the YTH domain of YTHDC1 exhibits binding affinity preferentially toward m6A-modified RNAs. Additional electrostatic interactions between YTHDC1 and the RNA promote the formation of the complex. The structure of the YTH domain of YTHDF1 and YTHDF2 was also resolved and indicated a similar manner of recognition of the m6A modified residues (Xu et al., 2015). Taken together these studies identified the family of YTH-domain proteins as specific and direct m6A readers.

In the past years several other putative m6A-binding proteins have been also discovered. The presence of m6A reduces the ability of RNA to form structures and favors the linear, unfolded form of RNA. This “m6A structural switch” promotes the interaction with RNA-binding proteins such as the nuclear hnRNPs (HNRNPG, HNRNPA2B1 and HNRNPC) (Liu et al., 2015). The Fragile X Mental Retardation Protein (FMRP) is considered to be also an m6A reader in the nucleus (Zhang et al., 2018). In the cytoplasm m6A can be recognized also by the Insulin-like Growth Factor 2 mRNA-Binding Proteins (IGF2BP1, IGF2BP2, and IGF2BP3) and the translation factor eIF3 (Huang et al., 2018). It remains unclear today whether these proteins bind m6A directly and if they recognize the consensus sequence or a different RNA sequence motif. The IGF2BP proteins and FMRP might also interact with m6A through the association with YTH domain-containing proteins (Youn et al., 2018). The available analyses of these questions are often controversial and how the specificity of the m6A binding is achieved needs to be further examined.

### **I.III.3 Roles of m6A in gene expression regulation**

The interest toward the biological significance of m6A revealed rapidly its numerous and diverse roles in gene expression regulation. The molecular and cellular functions of this RNA mark rely on the functioning and flexibility of its readers. Here, the main functions of the nuclear reader YTHDC1, the subject of this work, will be presented first followed by the description of the roles of the other YTH domain-containing proteins and the putative m6A readers.

## YTHDC1

Along with the rapid development of the epitranscriptome, in the scope of only few years, various important roles of YTHDC1 in gene expression regulation have been unraveled. These recently described functions of the protein will be presented following the steps of gene expression regulation- from chromatin organization to mRNA export, rather than in chronological order.

Two studies indicated that YTHDC1 associates with chromatin structure and the transcription control of gene expression. Firstly, YTHDC1 was linked to **chromatin silencing** in human cells through an interaction with the m6A-modified lncRNA *XIST* (Patil et al., 2016). It has been shown that the m6A residues on *XIST* are preferentially recognized by YTHDC1 and that the protein mediates the *XIST*-dependent X chromosome inactivation presumably by recruiting chromatin silencing factors. Upon loss of m6A, the artificial tethering of YTHDC1 to *XIST* rescued the *XIST*-mediated silencing. These results highlighted the essential role of YTHDC1 in the fundamental developmental process of sex-dosage compensation in females.

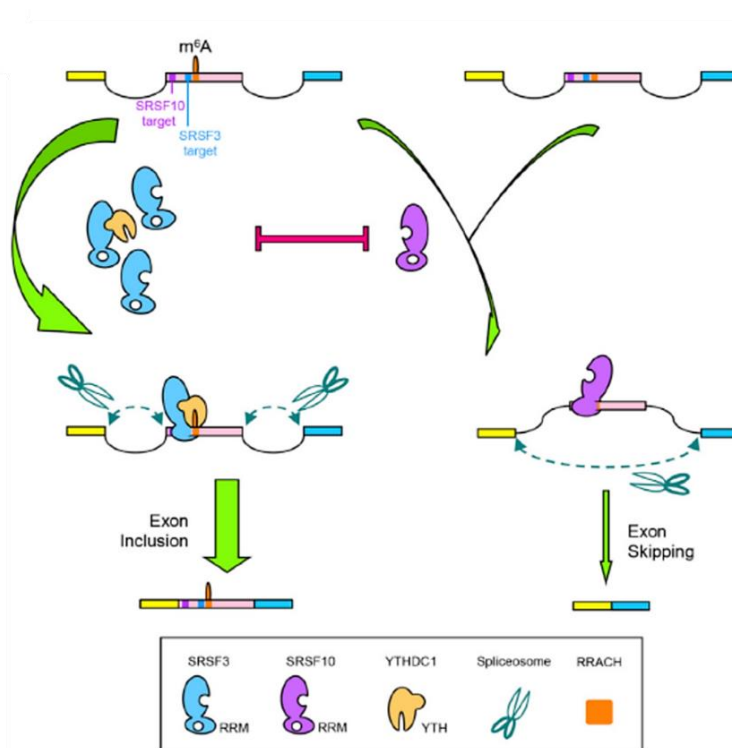
Secondly, a recent work demonstrated that the conditional knockout of METTL3 or YTHDC1 in mouse embryonic stem cells increases **chromatin accessibility** and activates **transcription** in an m6A-dependent manner (Liu et al., 2020). The depletion of YTHDC1 or METTL3 was associated to chromatin openness through the elevated levels of H3K4me3 and H3K27ac. Furthermore, the examination of chromosome-associated regulatory RNAs (carRNAs) indicated that m6A destabilizes carRNAs. This effect was linked to the **degradation of the m6A-modified carRNAs** through an interaction of YTHDC1 with NEXT (Nuclear EXosome Targeting complex) components. The depletion of YTHDC1 affected most strongly the long interspersed element 1 (LINE1) family, a class of mouse retrotransposons, expressed in embryos and critical for early development. Collectively, the results from this work identified the role of YTHDC1 in the regulation of chromatin structure and subsequent transcription activation as well as in the m6A-driven degradation of carRNAs.

The latest work on m6A identified METTL3 and YTHDC1 as important actors in the regulation of **Double-Strand Break Repair (DSBR)** pathways (Zhang et al., 2020). It was shown that both the methyltransferase and YTHDC1 are recruited to Double-Strand Break (DSB) sites following X-ray irradiation or exposure to chemical agents. METTL3 was found to methylate nascent RNAs at DSBs, which are subsequently recognized and protected by YTHDC1. The axis METTL3-m6A-YTHDC1 was associated with an increased accumulation of DNA-RNA hybrids at DSBs and with the implementation of homologous recombination-mediated DSB repair. Hence, this study demonstrated a novel role of YTHDC1 in the repair of DNA lesions and thereby in the preservation of genome integrity.

Long before its description as an m6A reader, the first discovered function of YTHDC1 was its regulation of **alternative splicing**. Early studies indicated that YTHDC1 is able to change alternative splice site selection and is essential for the splicing of several mRNAs (Rafalska et al., 2004).

In addition to this initial finding, more recently the mechanistic view of the process and the genome-wide significance of YTHDC1-mediated alternative splicing events were provided (Xiao et al., 2016). It was shown that YTHDC1 interacts with pre-mRNA splicing factors among which many SR proteins. In HeLa cells the knockdown (KD) of YTHDC1, SRSF3 (SRp20) or SRSF10 (SRp38) led to more than 2,000 cassette exons (exon skipping) events. YTHDC1 bound to exonic m6A residues was found to promote the recruitment of SRSF3 and to block the binding of SRSF10 to mRNAs (Fig.14). The two SR proteins presented a competitive binding to YTHDC1 resulting in either exon inclusion (YTHDC1-SRSF3 binding) or exon skipping (SRSF10 binding) (Fig.14). This role of YTHDC1 in the regulation of alternative splicing was furthermore defined as m6A-dependent since in YTHDC1 KD cells the regular splicing pattern was rescued using wild type but not m6A-recognition defective YTHDC1.

Moreover, this study revealed that YTHDC1 is needed for the nuclear speckle localization of SRSF3 and SRSF10 which was disrupted upon KD of YTHDC1 and restored by the introduction of the WT protein but not the m6A-binding mutant.



**Figure 14. YTHDC1 promotes exon inclusion through the recruitment of SRSF3 and the blockage of SRSF10 binding.** YTHDC1 interacts with SRSF3 and promotes its recruitment to exons while inhibiting the binding of SRSF10. YTHDC1 recognizes m<sup>6</sup>A residues within exons and together with SRSF3 leads to exon inclusion and abolishes the SRSF10-induced exon skipping. *Adapted from Xiao et al., 2016.*

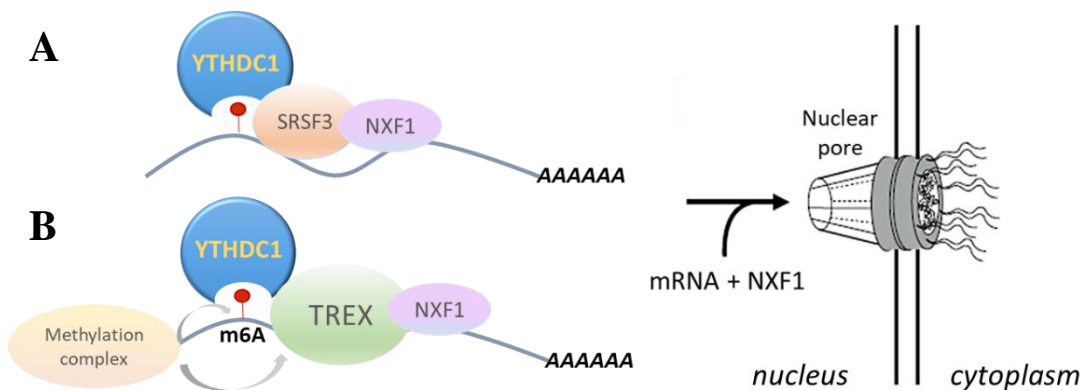
Recently, the role of YTHDC1 in splicing was extended to the level of circular RNAs (circRNAs). The production of circRNAs requires the so called back-splicing which consists in the fusion of two splice sites. The reaction is driven by the canonical splicing machinery but in an alternative way to the regular linear splicing of transcripts. It was shown that METTL3 and YTHDC1 are needed for the biogenesis of a subset of circRNAs (Di Timoteo et al., 2020). YTHDC1 was found to promote the back-splicing pathway and to have a positive effect on the production of a class of circRNAs in an m<sup>6</sup>A-dependent manner.

Another intriguing molecular role of YTHDC1 is its function in **mRNA export** from nucleus to cytoplasm. In HeLa cells the KD of the protein leads to nuclear retention of m<sup>6</sup>A-containing mRNAs observed through increased nuclear m<sup>6</sup>A levels and decreased cytoplasmic m<sup>6</sup>A content, accompanied by a global increase of the levels of polyadenylated RNAs in nucleus (Roundtree et al., 2017). These effects were not attributed to nuclear decay of mRNAs. In line with



the effect of the KD of the protein, the overexpression of YTHDC1 resulted in a drop of m6A nuclear levels and suggested the stimulated export of m6A-modified mRNAs. This work revealed that YTHDC1 regulates the subcellular localization of its mRNA targets in an m6A-dependent manner through an interaction with SRSF3. The proposed model suggested that SRSF3 facilitates the nuclear export of m6A-modified mRNAs via an interaction with YTHDC1 on one hand and the canonical export receptor NXF1 on the other hand (Fig.15A).

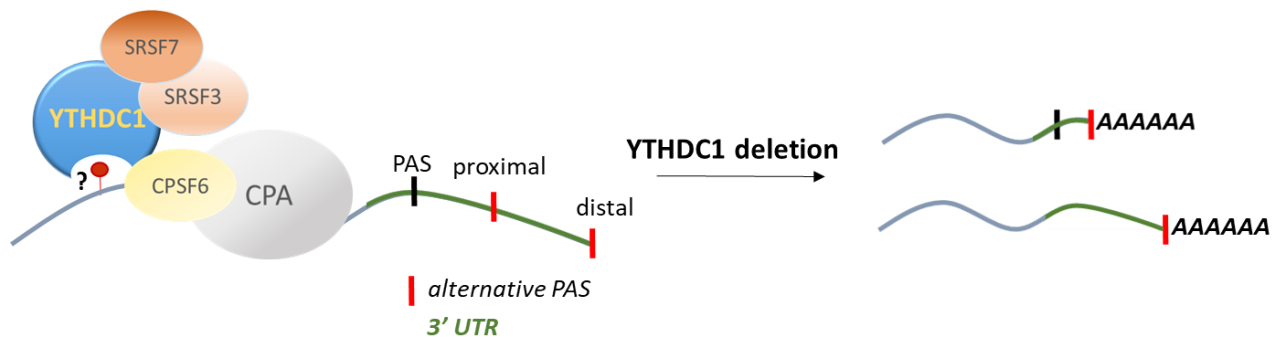
Further, recently it was demonstrated that the m6A methylation core complex is able to recruit TREX to specific mRNAs and in this way control their nuclear export (Lesbirel et al., 2018). The methyltransferases co-adaptors WTAP and VIRMA were found to be essential for the interaction of nascent m6A-modified mRNAs with TREX. Importantly, this work showed that YTHDC1 interacts with the main TREX components (Aly/Ref, DDX39B and CHTOP). Based on additional analysis a functional link between the methylation complex, TREX and YTHDC1 was proposed. The methylation complex recruits TREX which subsequently stimulates the recruitment of YTHDC1 and NXF1 to mediate the efficient mRNA export (Fig.15B). It is of note that this study suggested a more global role of YTHDC1 in mRNA export, beyond m6A signaling, since the majority of identified YTHDC1 targets were not m6A-modified.



**Figure 15: YTHDC1 is essential for the nuclear export of m6A-modified mRNAs.** **A.** YTHDC1 promotes the nuclear export of m6A-modified mRNAs via an interaction with SRSF3 which recruits NXF1 to the targeted transcripts and leads to their subsequent nuclear export through the nuclear pore. **B.** The core m6A methylation complex methylates nascent mRNAs and recruits TREX. TREX promotes the recruitment of YTHDC1 and NXF1 to targeted mRNAs and leads to their subsequent nuclear export through the nuclear pore.

Taken together these two studies indicate that the association of YTHDC1 with both SR proteins and TREX is essential for the interaction of newly synthesized m6A-containing mRNAs with NXF1 and is responsible for their precise nuclear to cytoplasmic export (Fig.15). Future studies are needed to decipher the association of these two YTHDC1-dependent export pathways and their significance outside the range of m6A-modified mRNAs.

In addition, YTHDC1 plays an important role in the **alternative polyadenylation of mRNAs**. It was demonstrated that YTHDC1-deficient mouse oocytes experience massive alternative splicing defects and an extensive polyadenylation linked to altered 3'UTR length (Kasowitz et al., 2018). The analysis of WT and YTHDC1-depleted oocytes revealed more than thousand APA events among nearly 900 different mRNAs. The APA patterns associated relatively equally with longer 3'UTR and shorter 3'UTR. YTHDC1 was shown to interact with CPSF6, one of the main components of the cleavage and polyadenylation complex. (Fig.16) Moreover, two of YTHDC1 partners- SRSF3 and SRSF7 are known to couple APA to mRNA export and might promote YTHDC1's role in alternative PAS selection (Fig.16). Whether YTHDC1-driven APA events are mediated by CPSF6, supplementary CPA subunits or other mRNA-processing factors remains to be resolved. Another important open question is if this molecular function of YTHDC1 is m6A-dependent or not.



**Figure 16: YTHDC1 is essential for the alternative polyadenylation of mRNAs.** The deletion of YTHDC1 leads to alternative PolyAdenylation Site (PAS) selection and subsequently to shortened or lengthened 3'UTRs. YTHDC1 interacts with CPSF6, SRSF3 and SRSF7 which probably mediate the role of YTHDC1 in APA. Whether this molecular function of the protein relies on the recognition of m6A residues is not understood.

Prior to the lately identified functions of YTHDC1 in mRNA processing and export, it was proposed that the protein resides within characteristic nuclear compartment termed “YT bodies” which were found to be different from all other known nuclear structures (Nayler et al., 2000). However, with the recent description of YTHDC1 as central regulator of mRNA metabolism, the so termed YT bodies are now believed to be **nuclear speckles** (NSs) (Galganski et al., 2017). This suggestion is based on the interaction of YTHDC1 with SRSF proteins and its importance for their NSs localization. Moreover, YTHDC1 contains a large low-complexity domain which can be involved in phase separation properties of the protein. Therefore, NS could be the site of **orchestration and interconnection of YTHDC1-dependent mRNA-processing events** such as alternative splicing, alternative polyadenylation and export. In addition, other m6A players-METTL3, METTL14, WTAP, ALKBH5 and FTO were recently found to localize to NSs and this structures could present a global regulatory site for m6A signaling.

#### *YTHDC2 and YTHDF proteins*

The m6A reader YTHDC2 was shown to mediate mRNA degradation through the recruitment of the 5' – 3' exoribonuclease XRN1 which was also confirmed by the increased expression of mRNAs upon YTHDC2 depletion in mouse testes (Wojtas et al., 2017). Further, it was proposed that YTHDC2 is able to promote mRNA translation by the activity of its RNA helicase domain and through the recognition of m6A sites located in the coding region of mRNAs (Mao et al., 2019). Overall, YTHDC2 might have a central role in the balance between mRNA translation and degradation, however its functions have not been fully elucidated yet.

The three members of the YTHDF family have interconnected roles in mRNA metabolism (Fig.17) (Zhao et al., 2019). YTHDF1 was found to bind to the eukaryotic translation initiation factor eIF3 and to interact with the translation machinery in order to enhance the protein production from dynamic m6A-marked transcripts (Wang et al., 2015). In line with this, it was demonstrated that YTHDF1 knock-down in human cells leads to decreased translation of its mRNA targets. Nevertheless, the precise molecular mechanisms of the process are still not known.

It was revealed that YTHDF2 recognizes m6A preferentially in the 3'UTR and promotes the localization of mRNA to decay sites such as P-bodies. Further investigations found that

YTHDF2 recruits the mRNA deadenylation complex CCR4-NOT in order to degrade its targets (Du et al., 2016). In the same context, it was indicated that the m6A-modified transcripts present a shorter half-life compared to the unmodified mRNAs (Zhao et al., 2017). Additional roles have been assigned to YTHDF2 specifically under stress conditions and will be later on presented in the next chapter.

YTHDF3 seems to have overlapping functions with YTHDF1 and YTHDF2 and was proposed to work with them in a cooperative fashion on one hand to stimulate mRNA translation and on the other one to induce the degradation of targeted transcripts (Shi et al., 2017).

The latest study on YTHDF proteins indicates that in addition to their individual functions, these specific m6A readers act in a cooperative fashion in order to induce mRNA degradation (Zaccara and Jaffrey, 2020). It is of note that this work revealed for the first time that YTHDF proteins act on the same subset of mRNAs and can compensate the functions of one another. In addition to mRNA degradation YTHDF1, 2 and 3 were found to mediate cell differentiation, highlighted through the depletion of the three paralogs.

As YTHDC1, all members of the YTHDF family have also a low-complexity domain and are able to phase separate into liquid droplets with mRNAs containing multiple m6A modifications. These RNA-protein droplets then partition into endogenous membraneless compartments such as stress granules, P-bodies or neuronal RNA granules and participate respectively in the storage, degradation or transport of m6A-containing mRNAs (Ries et al., 2019).

In conclusion, the YTH domain-containing proteins, have diverse molecular roles and are implicated at multiple steps of the gene expression process- chromatin condensation/silencing, transcription of circRNAs, DSB repair, mRNA processing and export, translation, decay and phase separation of mRNAs (Fig.17). However, many intriguing questions are unanswered today including the selective recognition of particular mRNAs and of precise m6A sites. It is highly possible that the direct m6A readers present many other functions in addition to the once which have been already revealed. Thus, future analyses are needed to decipher the complex network of YTH domain-containing proteins and their broad implication in gene expression regulation.

nucleus		m6A reader
Chromatin silencing		YTHDC1
Chromatin condensation Degradation of carRNAs		YTHDC1
HR-mediated DSB repair		YTHDC1
mRNA alternative splicing mRNA APA mRNA export	NS (?)	YTHDC1
Back-splicing (biogenesis circRNAs)		YTHDC1
cytoplasm		
mRNA translation		YTHDC2, YTHDF1, YTHDF3
mRNA degradation		YTHDC2, YTHDF2, YTHDF3
mRNA phase separation		YTHDF1, YTHDF2, YTHDF3

**Figure 17: Roles of m6A direct readers in gene expression regulation.** The YTH domain-containing proteins are involved in chromatin organization, transcription control, homologous recombination-mediated double-strand break repair, mRNA processing and export, mRNA translation, degradation and phase separation.

#### Roles of the other putative m6A readers

The other putative m6A readers have also several important molecular functions. HNRNPA2B1 was shown to influence the m6A-dependent alternative splicing and the biogenesis of miRNAs (Alarcón et al., 2015). The IGF2BP proteins and FMRP were shown to enhance m6A-containing mRNA's stability (Huang et al., 2018; Zhang et al., 2018). The direct binding of eIF3 to m6A residues in the 5'UTR of mRNAs promotes translation in a cap-independent manner without the implication of eIF4E (Meyer et al., 2015a). Thus, this mechanism was proposed as essential for the expression of mRNAs when the eIF4E functions are impaired.

It was revealed that all of these proteins (HNRNPA2B1, IGF2BP, FMRP and eIF3) interact with the YTHDF family members (Youn et al., 2018). Therefore, since the binding of these putative readers is unclear it is possible that they act indirectly as YTHDF partners rather than present specific functions on their own.

#### **I.III.4 Biological functions of m6A**

The various molecular functions of m6A molecular players constitute a dynamic regulatory network, essential for the cellular homeostasis and for the proper implementation of fundamental processes. The biological significance of m6A has been revealed through the examination of

phenotypes derived from perturbations of its writers, erasers and readers. These studies indicated that m6A is essential for development, fertility, sex determination, stress responses, circadian clock and moreover its deregulations were found to be lethal.

The experimental elimination of METTL3 and METTL4 in human and mouse embryonic stem cells was associated with the inability of cells to terminate their naïve state and to differentiate, consequently leading to restricted lineage priming (Batista et al., 2014; Geula et al., 2015). Further, depletion of METTL3 elicited circadian period elongation (Fustin et al., 2013). Importantly, the deletion of METTL3 or WTAP were also found to result in early embryonic lethality (Geula et al., 2015).

The genetic deletion of m6A erasers induces growth retardations and impaired pre-adipocytes differentiation (FTO) as well as defects in spermatogenesis and impaired fertility (ALKBH5) in mice (Zhao et al., 2014; Zheng et al., 2013).

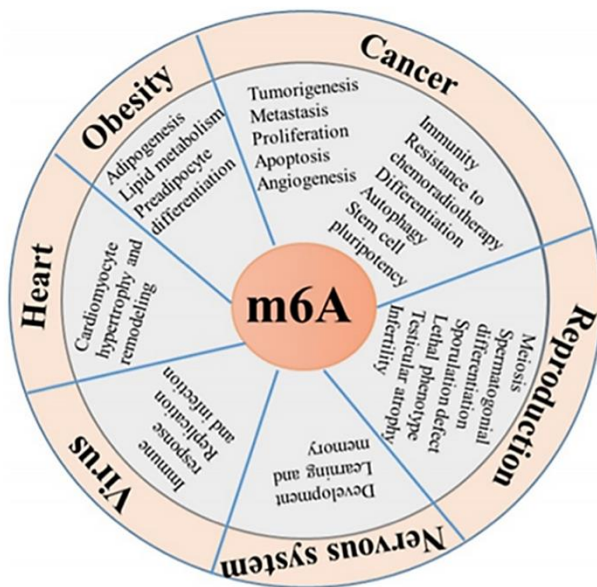
YTHDC2 knockout mice were shown to exhibit defects in spermatogenesis due to the entry of germ cells in a premature and aberrant metaphase promoting apoptosis. In general, it was demonstrated that YTHDC2 is needed for male and female mice fertility (Wojtas et al., 2017).

Studies revealed that **YTHDC1** is essential for sex determination in *Drosophila* through the modulation of the alternative splicing of particular regulatory transcripts (Kan et al., 2017). Furthermore, the deletion of YTHDC1 was found to induce early embryonic lethality in mice (Kasowitz et al., 2018). The APA and splicing defects following conditional knockout of YTHDC1 in mice were associated with broad germline developmental defects and resulted in male and female infertility.

In humans m6A signaling has been extensively studied in the context of chronic diseases. The deregulation of m6A levels and/or the deviant expression levels of the m6A actors are linked to the inactivation or overexpression of downstream carcinogenic or tumor suppressor genes and participate to the oncogenic transformation and tumor progression in many types of cancers (Liu et al., 2018). Analysis of tumorigenesis in mice models and cancer cell lines depicted the central

role of METTL3 in the highly proliferative state of cancer cells, in the formation of metastasis and in the progression of breast cancer, prostate cancer, glioblastoma and lung cancer (Cai et al., 2019; Cui et al., 2017; Lin et al., 2016; Wang et al., 2020). Therefore, METTL3 is currently investigated as a potential therapeutic target in oncology. Overall, deregulated levels of m6A and FTO, ALKBH5, METTL3 and METTL14 are observed in cancer, type 2 diabetes, infertility, obesity, neurological disorders and inflammation (Fig.18) (Zhu et al., 2020; Kadumuri and Janga, 2018).

Importantly, m6A was described as dynamic RNA modification since its levels can vary not only in the context of diseases but also among human tissues, during development and in response to stress (Knuckles et al., 2017; Sánchez-Vásquez et al., 2018; Zhang et al., 2020b; Zhou et al., 2015). The m6A flexibility is considered to relate greatly to the plasticity of the genome reprogramming observed among cell types, during development and under hostile environmental conditions, nevertheless many aspects of these dynamics are unclear today. The role of m6A, particularly in the implementation of the adaptive heat stress response, will be presented in the next chapter.



**Figure 18: Examples of some of m6A-dependent disease mechanisms in humans.** m6A is associated with cancer, obesity, reproduction, virus infections, heart and nervous system development and morphology. *From Zhu et al., 2020*

An exciting and relatively less studied enigma is the role of m6A signaling at the level of ncRNAs which can present a model for modulation of gene expression also in particular conditions, at certain stages of development or along with the progression of diseases (Coker et al., 2019).

All of these crucial biological functions demonstrate the central role of m6A in gene expression regulation, cellular and organismal homeostasis. Yet, many mysteries in the field of the epitranscriptome need to be further investigated in order to understand the molecular implication of this and potentially other RNA modifications in the modulation of cell fate in normal and stress conditions and in the development and progression of human pathologies.

In summary, this part of the introduction provided a general overview of the molecular mechanisms implicated in gene expression. The regulation of the process allowing the flow of information from DNA to proteins is governed at multiple levels, which have been the subject of tremendous number of studies over the past decades. Interestingly, recent technological advances allowed to unravel the importance of a new exciting layer of regulation of gene expression- the epitranscriptome. Altogether, the studies on RNA modifications, and their molecular actors have diversified the central dogma of molecular biology and continuously unravel novel pathways for regulation of RNA fate, cellular functions and cellular homeostasis. One of the main actors of the epitranscriptome is the m6A nuclear reader, the protein YTHDC1, whose functions and role in gene expression regulation have been examined in this work in both normal and stress conditions.



## **Chapter II**

### **The Heat Shock Response (HSR)**



## **II.I General overview of the HSR**

Cellular stress responses are of great importance and are often applied in order to properly modulate gene expression and subsequent cellular processes in the context of variable external conditions. In my PhD work we have used the cellular stress response to heat shock as a model to unravel new ways of gene expression regulation and adaptation involving both well-studied molecular mechanisms as well as novel and relatively less understood epitranscriptomic pathways. In this second chapter of the introduction, I will present the implementation and the main characteristics of the Heat Shock Response with particular focus on the processes investigated in this study.

### **II.I.1 Cellular stress responses and the HSR**

All living organisms are constantly exposed to various environmental changes which could be harmful and potentially deleterious. Over the years the broadly used term “stress” evolved to describe situations and particular contexts which present a threat to organisms.

To a great extent, the mammal’s evolutionary success is due to their extraordinary capacities to cope and adapt with exposure to a variety of stressors. Behind these great capacities stand numerous molecular mechanisms which support cellular functions and maintain the microenvironment (Kültz, 2005). Such mechanisms of cellular adaptation to different forms of stresses describe the cellular stress response. From an evolutionary point of view, the stress response system derived from natural selection to adjust physiology and behavior to diverse circumstances. Remarkably, the defending pathways have been broadly conserved among species over hundreds of millions of years (Nesse et al., 2016).

Adaptive responses depend on the natural environment, the organism and its physiological state. In response to stress almost every aspect of cell physiology is modulated- gene expression regulation is modified together with the cellular metabolism, cell cycle progression, protein homeostasis and enzymatic activities. In eukaryotes, sophisticated sensing mechanisms allow to transduce external stimuli into internal signals in order to activate stress response pathways with the sole goal of preserving the cellular homeostasis.

Examples of some of the most common external stresses leading to multiple cellular alterations in mammals include hypoxia, nutrition deprivation, heat shock, exposure to chemical toxins, DNA damaging agents or infectious agents.

In each scenario the stimulated stress response pathway leads to damage repair and cell survival or, if the cell cannot cope with the consequences of the stress, results in cellular senescence and programmed cell death (Samali et al., 2010). When the stress stimulus is severe and/or when the stress exposure is prolonged, the activated mechanisms may be insufficient to tackle the perturbations. In that case, the seriously damaged cells are functionally neutralized or eliminated in order to preserve the systemic homeostasis. Importantly, the inability to cope with stress can have severe organismal impact and trigger the initiation and progression of numerous pathological states (inflammation, cancer, ischemia, necrosis, etc.) (Zhang, 2018).

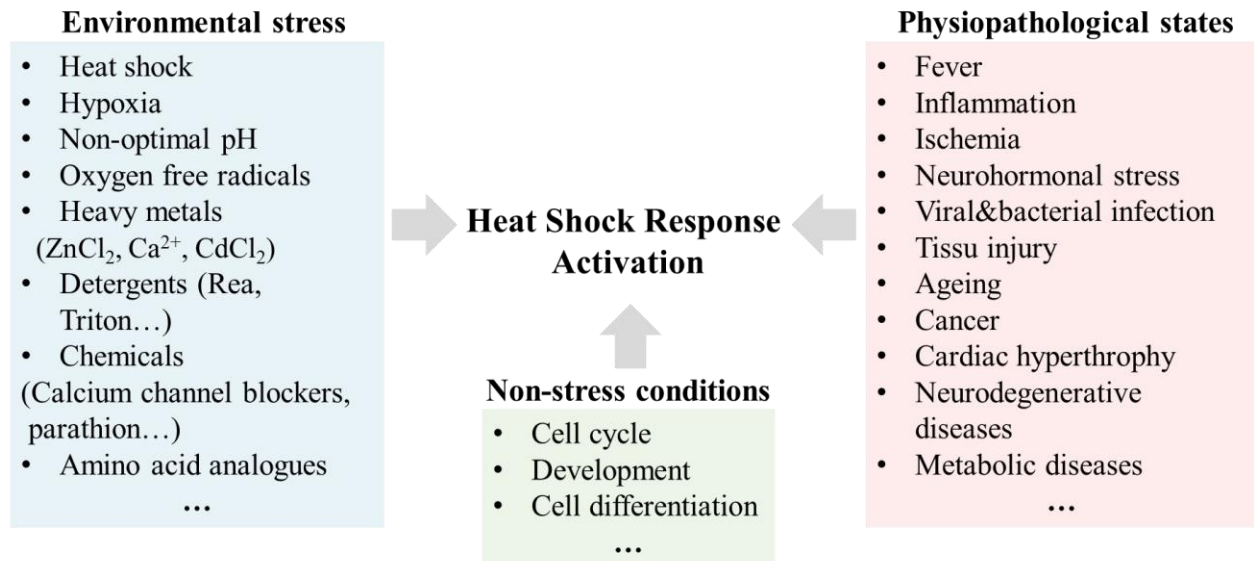
It is of note that cellular stress responses present a complex interplay and redundancy that should be taken into account. Diverse external stimuli can modify in the same way the cell behavior. However, the outcomes at the cellular level are often multiple with more than one stress response pathway being activated simultaneously in order to protect vital functions. Therefore, many aspects of cellular stress responses are not specific for a given stressor because cells surveil stress based on damage at the macromolecular scale. Subsequently the activated mechanisms are rather interconnected and share common elements.

Apart from the environmental stress-inducing fluctuations, diverse physiopathological states can set off cellular stress responses (Poljšak and Milisav, 2012). While in physiological conditions the protective mechanisms insure survival, in the context of disease, stress response pathways can be deregulated and often support pathological development.

A great example of a cellular stress response triggered by multiple external stimuli, in different physiopathological states or even in non-stress conditions is the Heat Shock Response (HSR), which will be thoroughly presented in the next parts of this chapter. The HSR (as defined by the activation of specific pathways) is activated when protein denaturation, unfolding and aggregation are perceived by the cells. Indeed, proteins present a certain sensitivity to external perturbations and the accumulation of damaged/unfolded proteins is considered as a signal to the cell to start a protective program.

An increase of the temperature of only few degrees triggers very rapidly the HSR survival pathway (Fig.19). Besides temperature increase, another classical HSR trigger is proteotoxic stress (Morimoto, 1998). Proteotoxic stressors include metal-ions, toxic chemicals, heavy metals, solvents, amino acid analogs and some pharmacologically active molecules. More recent studies identified multiple pathological contexts to correlate with the expression of specific heat shock-

induced proteins. Ischemia, fever, inflammation, cancer, aging, cardiac hypertrophy, neurodegenerative diseases, metabolic diseases and tissue damage disrupt the protein homeostasis and result in the activation of the HSR independently from external environmental changes (Fig.19), (Jolly & Morimoto, 2000). Interestingly, the HSR is also stimulated in non-stress conditions– during cell cycle, cell differentiation and development (Fig.19) and in this case is tightly linked to the accurate regulation of physiological processes.



**Figure 19: Triggers of the cellular Heat Shock Response.** Different environmental stresses, physiopathological states or non-stress conditions lead to the activation of the Heat Shock Response in mammals.

The list of triggers responsible for the deployment of the HSR pathway is constantly increasing. Yet, apart from heat shock and proteotoxic stress, the activation of the HSR in other contexts is relatively less understood and might be less specific and rather reflect the execution of other general stress pathways.

All living organism are challenged by the ever-changing variables of nature. Studies on the mechanisms used to cope with external perturbations are of great importance today more than ever. They serve as crucial indicators of global environmental changes and could be further applied to monitor such events. The research on cellular stress pathways allows us to gain great insight into evolution and understand complex genetic and molecular mechanisms conserved among species. More importantly, the field of stress biology is needed as well to address diverse health problems,

decipher disease mechanisms and clinically utilize the properties of the cellular stress responses for the development of novel predictive biomarkers and targeted therapies.

In eukaryotes the activation of the HSR is characterized by a massive genome remodeling. Ever since its discovery, the HSR has been a model of choice for the studies of gene expression regulation due to its easy implementation and its reversible pattern. In the next parts of this section of the introduction, the discovery and conservation of the HSR as well as the deleterious effects particularly of heat stress will be presented. The following chapters will focus on the key mechanisms of gene expression regulation accompanying the HSR and the roles of the best characterized coding and non-coding targeted genes.

### **II.1.2 Discovery and evolutionary conservation of the HSR**

The pioneering discovery of the HSR by the Italian researcher Ferruccio Ritossa dates from nearly sixty years ago. At the early 1960's Ritossa was studying the synthesis of nucleic acids in "puffs" (chromosomal enlargements visible through optical microscope) in *Drosophila*. Through an extraordinary combination of knowledge and as a result of a fortunate coincidence Ritossa and colleagues described for the first time environmental-driven changes in gene expression. An accidental shifting of the incubator's temperature led to the crucial observation of novel puffing pattern in *Drosophila*'s salivary glands (Ritossa, 1962). The newly synthesized RNAs were observed very rapidly, only in a few minutes, while previously detected puffs were shut off.

In the scope of two decades the HSR was found to be the driver of major genetic alterations and presented an ideal situation to study gene regulation in eukaryotes. Tissières' lab described originally the so called "puffs" as active loci, corresponding to heat shock genes (Tissières et al., 1972). These genes were among the first to be cloned and to have their organization established. Their vigorous activation in response to stress was found to result in the synthesis of Heat Shock Proteins (HSPs) while the expression of many active genes was found to be drastically inhibited (Pauli et al., 1992). The cloning of heat shock genes led to the discovery of a specific transcription factor, named HSF1 for Heat Shock Factor 1, recognizing particular DNA motifs and driving the transcription of *HSP* genes (Pelham, 1982; Wu et al., 1984).

Following the early work on *Drosophila*, it has been described that virtually all organisms from archaeobacteria to human show a universal response to thermal stress (Kelley and Schlesinger,

1978; Lemaux et al., 1978). The important evolutionary conservation of the HSR is thought to reflect an adaptive mechanism needed for the survival of eukaryotes and prokaryotes under hostile environmental conditions. Nevertheless, some of the heat stress-related proteins have evolved and can have species-specific characteristics, also the kinetics of expression of heat-inducible genes can be quite diverse from one organism to another. These differences represent an evolutionary optimization preserving the homeostasis at the specific growth temperatures of each organism (Richter et al., 2010). However, the main actors of the HSR, the transcription factor HSF1 and the HSPs as well as the underlying protective mechanisms are highly conserved in eukaryotes and the transient gene expression program driven by heat stress is believed to be as ancient as life.

The great scientific discoveries of the last century are the milestones of our current knowledge on the molecular complexity and physiological importance of this conserved pathway.

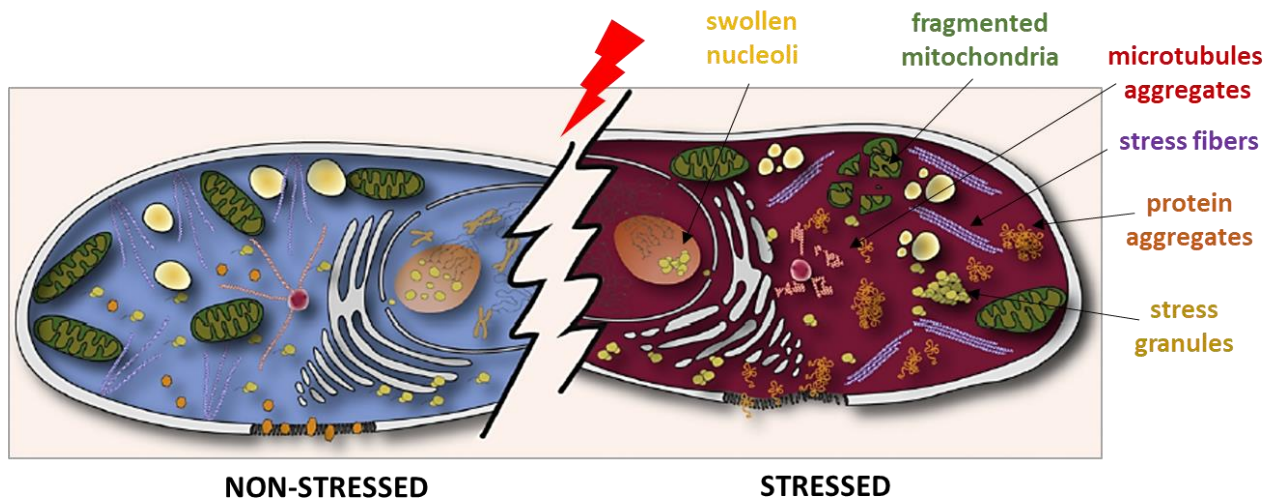
### **II.1.3 HSR at the cellular level**

An interesting aspect of the HSR is its rapid activation triggered not by temperatures modifications per se but by the general imbalance of protein homeostasis. Since proteins are the main building blocks of cells, their dysregulation is associated with multiple morphological and phenotypic alterations. Here, the deleterious effects of the classical trigger of the HSR- heat shock (HS), will be presented.

In eukaryotic cells, beyond the unfolding of individual proteins, HS alters significantly the internal cellular organization. One of the main perturbations occurs at the level of the cytoskeleton. Even mild heat stress can induce the reorganization of actin filaments, into stress fibers (Toivola et al., 2010). Moreover, severe HS leads to the aggregation of vimentin and other filament-maintaining proteins, thus resulting in the collapse of actin and tubulin networks (Welch and Suhan, 1985). The vicious disruption of the cytoskeleton associates with loss of the correct organelles localization, failure of intracellular transport, reorganization of the nucleolar fibrillar reticulum and general deterioration of cellular compartments (Fig.20) (Welch and Suhan, 1986). Under stress conditions the endoplasmic reticulum (ER) and the Golgi system become fragmented while the number of mitochondria and lysosomes decrease. Heat stress causes as well a drastic drop in the cellular ATP levels induced by the uncoupling of oxidative phosphorylation and the loss of mitochondria (Fig.20) (Patriarca & Maresca, 1990).

In these conditions the morphology of cellular membranes is also severely altered. The ratio of proteins to lipids is modified resulting in higher membrane fluidity (Fig.20) (Vigh et al., 2007).

Diverse nuclear processes are affected by HS. The changes in transcription and the translation processes can be visualized even at the cellular level (Fig.20). In the cytosol, stress granules, containing non-translating mRNAs, RNA-protein structures, and other RNA-processing factors, are forming (Buchan and Parker, 2009). Simultaneously, in the nucleus, at the sites of ribosome assembly, the nucleoli, large granules containing incorrectly processed ribosomal RNAs and aggregated ribosomal proteins, can be observed (Welch and Suhan, 1985).



**Figure 20: Deleterious effects of heat shock on eukaryotic cell and restoration of the normal state through the activation of the HSR.** On the left is represented an unstressed eukaryotic cells, on the right a cell under heat stress. HS leads to reorganization of actin filaments (purple) into stress fibers and the aggregation of microtubule (red). Some organelles disassemble and become fragmented (mitochondria, in green; Endoplasmic Reticulum, in white). The nucleoli (yellow) swell and large depositions of ribosomal proteins are visible. Stress granules (darker yellow), formed by assemblies of protein and RNA, are observed in the cytoplasm as well as numerous protein aggregates (dark orange). The membrane morphology is strongly altered. The multitude of dysregulations leads to cell cycle arrest (non-condensed chromosomes in the nucleus). *Adapted from Richter et al., Molecular Cell, 2010.*

Altogether, these harmful effects lead to an arrest of the cell cycle and decreased cellular growth and proliferation (Lindquist, 1986). If the HS duration is prolonged or the increase of the temperature is drastic, the accumulation of cellular damages can result in cell death and can further alter the organismal microenvironment.



Hence, the correct implementation of the HSR is essential for all aspect of a cell- from adapted gene expression regulation and conservation of cellular functions to cell survival and preservation of viable physiological processes. When non-stressful growing conditions are restored, the cells reacquire their original morphology and internal organization and the conventional cell phenotype and functioning are reestablished due to the “healing” capacities of the HSR.

## **II.II Gene expression regulation during HSR**

### **II.II.1 General overview of gene expression regulation during HSR**

The most important and extensively studied characteristic of the HSR is the particular gene expression program which is activated to combat the cellular perturbations. For a long time, the stress-induced genomic landscape has been a key interest in the domain of molecular biology and the control of gene expression.

Today it is known that the HSR associates with the upregulation of specific stress-responsive coding and non-coding genes and the downregulation of the majority of active genes. Many studies have addressed the transcriptomic pattern of the HSR, activated specifically upon heat exposure, on the genome-wide level. It is of note that following the stress conditions (time of heat shock and temperature increase), the model organism and the applied techniques, the obtained results might vary significantly.

Regardless of the experimental differences, the genome-wide studies performed in human cells collectively indicate that more than half of the genome is shutdown while a much smaller fraction (5 to 10%) is vigorously activated in response to HS (Trinklein et al., 2004; Tabuchi et al., 2008; Vihervaara et al., 2017). Earlier, the examination of the global HS-transcription patterns was often carried out using microarray and RNA-seq experiments and did not reveal the involved regulatory mechanism (Trinklein et al., 2004). However, a recent study carried out precision nuclear run-on sequencing (PRO-seq), an assay that maps the transcriptionally engaged polymerase at nucleotide resolution, in mouse embryonic fibroblasts subjected to HS (Mahat et al., 2016). The application of this specific and sensitive technique confirmed the significant upregulation of ~ 1500 (10%) and downregulation of ~ 8000 (55%) of all active genes during the first hour of HS in mammalian cells. More importantly, this work unraveled at a precedence time scale and at a

genome-wide scale the extensive mechanisms underlying the HSR-associated program of gene expression regulation and highlighted the importance of new central actors, besides the well-known transcription factor HSF1. In fact, HSF1 turned out to have a rather limited role in the activation of stress genes and was not associated with the stress-related downregulation of active genes.

The HSR has the powerful ability to overturn the cellular genome and to transiently control the cell's proteome. Among the stress-responsive genes most significantly induced are the Heat Shock Protein-coding genes. Heat Shock Proteins (HSPs), or chaperones, are the central players of the HSR since they are needed to restore the proteome homeostasis and insure cell survival. In addition various other proteins and specific ncRNAs are upregulated in order to assist the HSR implementation and to sustain cellular processes.

On the other hand, some of the significantly downregulated genes include metabolic enzymes, cell cycle regulators, transcription and mRNA processing factors and transport proteins. The very rapid shutdown of canonical genes and the resulting global arrest of cellular functions are required to have enough free resources for the immediate production of stress proteins. Moreover, the downregulation of the transcriptome is essential to prevent the production of new proteins which can be damaged by the stress and will further need the assistance of HSPs to recover.

Notably, the HSR-related genome remodeling is a reversible process which is precisely coordinated in time. For instance, most of the chaperones are rapidly required and their expression is activated immediately after heat exposure while other proteins and ncRNAs with various functions are induced later, during the recovery period following the HS. When the proteome homeostasis is restored, the HSR gets progressively inactivated and the regular transcriptomic characteristics are reinstated.

The rapid and robust genome remodeling relies on the remarkable capacity of the HSR to modulate virtually all molecular mechanisms implicated in gene expression. The stress exposure is accompanied by changes of chromatin organization, transcription initiation and termination, mRNA-processing, mRNA export and translation. These events will be presented next in parallel for the downregulated and the upregulated genes, in order to allow comparison of the different regulatory processes underlying the drastic genome reprogramming.

## **II.II.2 Mechanisms of gene expression regulation during HSR**

### *Chromatin dynamics during HSR*

The HSR involves fast and global modifications of histone epigenetic marks and thus controls the compaction and accessibility of chromatin. At the level of euchromatin heat stress induces wide-spread deacetylation which is followed by reacetylation during the recovery period (Fritah et al., 2009). For instance, in HeLa cells H4K16 and H3K9 are deacetylated rapidly and present reacetylation after a few hours of recovery at 37°C. The global deacetylation of euchromatic regions was suggested to correlate with the downregulation of the transcribed genome and can be explained by the enhanced activity of the Histone DeACetylases 1 and 2 (HDAC1 and 2) upon stress induction (Fig.21). Importantly, the transcription factor HSF1 was found to interact with HDAC1 and 2 and participate to this stress-induced deacetylation. In addition, during the HSR, H3 and H4 phosphorylation and methylation marks are also altered (increased or decreased levels depending on the mark) which is thought to promote transcriptional silencing at specific genomic loci (Fritah et al., 2009; Li et al., 2019).

For a longtime HSF1 has been considered to be the central regulator of the HSR and therefore its mode of action and the induction of its targets are the best studied to date. The particular opening of the chromatin structure at the promoter elements of HSF1-dependent stress genes relies on the collaborative action of the transcription factor with various chromatin remodeling complexes. HSF1 was found to interact with Brahma-Related Gene 1 (BRG1) transcription factor (part of the chromatin remodeling complex SWI/SNF) and the Tip60 histone acetylating complex (Fig.21) (Sullivan et al., 2001). The recruitment of Tip60 to stress-responsive genes results in histone acetylation and spreading along the gene to maintain the transcriptional active compartment (Miozzo et al., 2015). Further, HSF1 cooperates with PARP (poly(ADP)-Ribose Polymerase) which can contribute to histones displacement and nucleosomes destabilization (Fig.21). The histone chaperone Facilitates Chromatin Transcription (FACT), is also recruited to heat-shock genes and displaces H2A-H2B dimer to open the nucleosomal DNA (Fig.21) (Fujimoto et al., 2012).

In addition to the modulation of euchromatin, HS and proteotoxic stressors modify the epigenetic pattern of constitutive heterochromatin. When the HSR is induced, the transcription of pericentric DNA regions is activated and results in the production of lncRNAs, termed *SATIII*. At

this inaccessible locus the opening of chromatin is accompanied by the loss of the H3K9me3 mark and the dissociation of its reader- the Heterochromatin Protein 1 (HP1). The transcription at these regions is additionally supported by the recruitment of histone acetyltransferase and the presence of phosphorylated RNA Pol II at Ser 2 (Jolly et al., 2004, Col et al. 2017). As mentioned previously, *SATIII* lncRNAs have several intriguing molecular functions which will be presented later on in the chapter on the transcriptional non-coding response to HS.

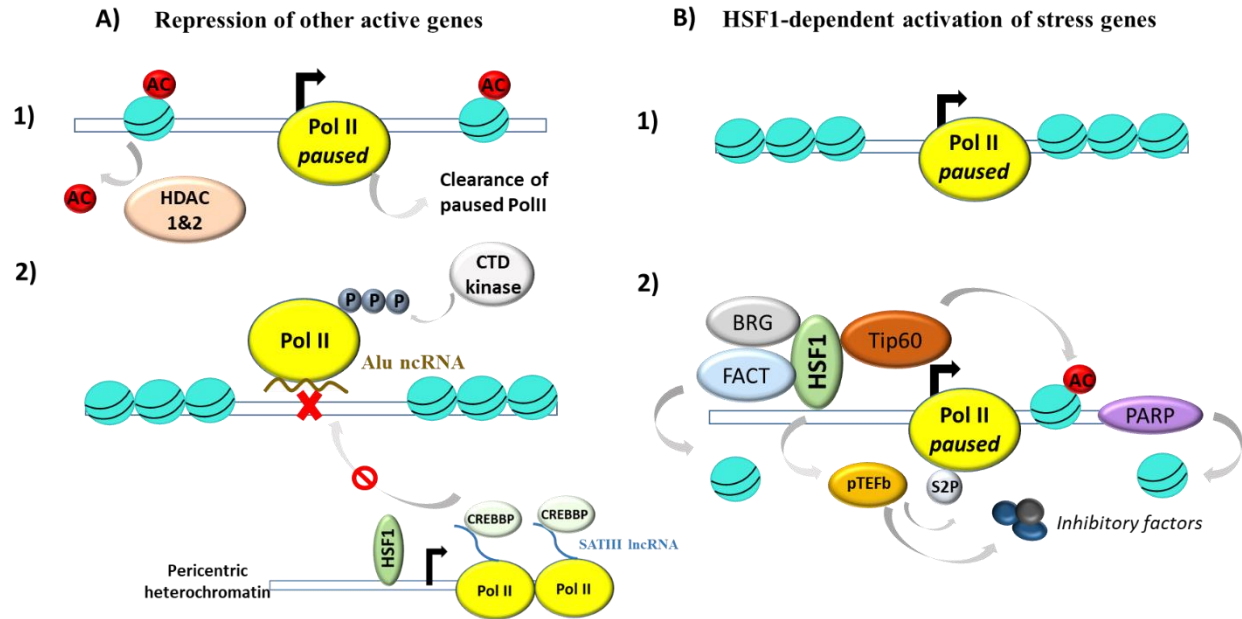
### Transcription during HSR

The transcription steps- initiation, elongation and termination are tightly controlled during the HSR and responsible to a great extent for the gene-specific reprogramming of gene expression. The genome-wide transcriptional repression occurs via the inhibition of the release of promoter-proximal paused RNA Pol II into elongation, subsequently leading to the clearance of the transcriptional machinery from the core of thousands of genes (Mahat et al., 2016).

In addition, the recruitment of RNA Pol II to genes which are not stress-induced is impeded through several mechanisms. Recent analyses indicated the important role of ncRNAs in this process. It has been demonstrated that in human cells the *Alu* ncRNAs, transcribed from Short Interspersed Elements (SINE), regulate different aspects of the heat-induced transcriptional repression (Allen et al., 2004). The *Alu* RNAs interact directly with the RNA Pol II at the promoter of downregulated genes and prevent the appropriate contacts between the polymerase and the promoter upstream and downstream from the TATA box (Fig.21) (Mariner et al., 2008). The transcription of these ncRNAs is regulated by HSF1 and is induced in response to HS.

Furthermore, the lncRNAs *SATIII* were found to sequester the transcription co-activator CREB Binding Protein (CREBBP) and render it unavailable for the transcription of its targets (Fig.21) (Goenka et al., 2016). The massive stress-driven transcription of *SATIII* is HSF1- and RNA Pol II-dependent. In this situation it has been proposed that the repression mediated by *SATIII* is supported through the titration of RNA Pol II at pericentric regions (Metz, 2004).

Experiments performed in HeLa cells showed that the RNA Pol II C-terminal domain (CTD) is hyperphosphorylated upon HS by stress-induced CTD kinases while the CTD-specific phosphatase gets inactivated and denatured (Fig.21) (Dubois, 1999). It seems that the RNA Pol II with hyperphosphorylated CTD has reduced affinity for the “normal” genes and increased efficiency for the transcription from the promoter of heat-responsive genes.



**Figure 21: Chromatin dynamics and transcription initiation during HSR.** On the right is represented a gene which is repressed upon HS (A). The activity of HDAC1&2 increases upon HS, the acetylated active regions get progressively deacetylated and the chromatin is compacted (A,1). The release of the paused RNA Pol II is inhibited and shortly after the transcriptional machinery is cleared off repressed gene's promoters (A,1). *Alu* ncRNAs abolish the contact of newly arriving RNA Pol II with the promoter of non-stress responsive genes (A,2). The activity of CTD kinases is stress-induced and the RNA Pol II CTD is hyperphosphorylated (A,2). *SATIII* lncRNAs, transcribed by HSF1 at pericentric heterochromatin sequester the transcriptional co-activator CREBBP and RNA Pol II and further promote the downregulation of active genes (A,2). On the left is represented a stress-responsive gene with compacted chromatin and a transcriptionally engaged and paused RNA Pol II in normal conditions (B,1). Upon HS, HSF1 binds the promoter of stress genes, interacts with Tip60, BRG, FACT, recruits PARP and promotes the decompaction of chromatin (B,2). HSF1 recruits pTEFb which phosphorylates the RNA Pol II and dissociates the inhibitory factors (B,2).

The study of the activation of HSP's transcription led for the first time to the discovery of promoter proximal pausing of RNA Pol II (Rougvie and Lis, 1991). Today, the involved mechanisms are believed to be a regulatory step at many mammalian genes and particularly at the level of inducible and developmental genes (Adelman and Lis, 2012). This permanent "starting position" of the RNA Pol II provides a highly reactive transcriptional response leading to the rapid production of stress proteins upon HS. At the level of HSF1-dependent stress genes, the binding of HSF1 to the promoter region induces the recruitment of p-TEFb. Further p-TEFb drives the

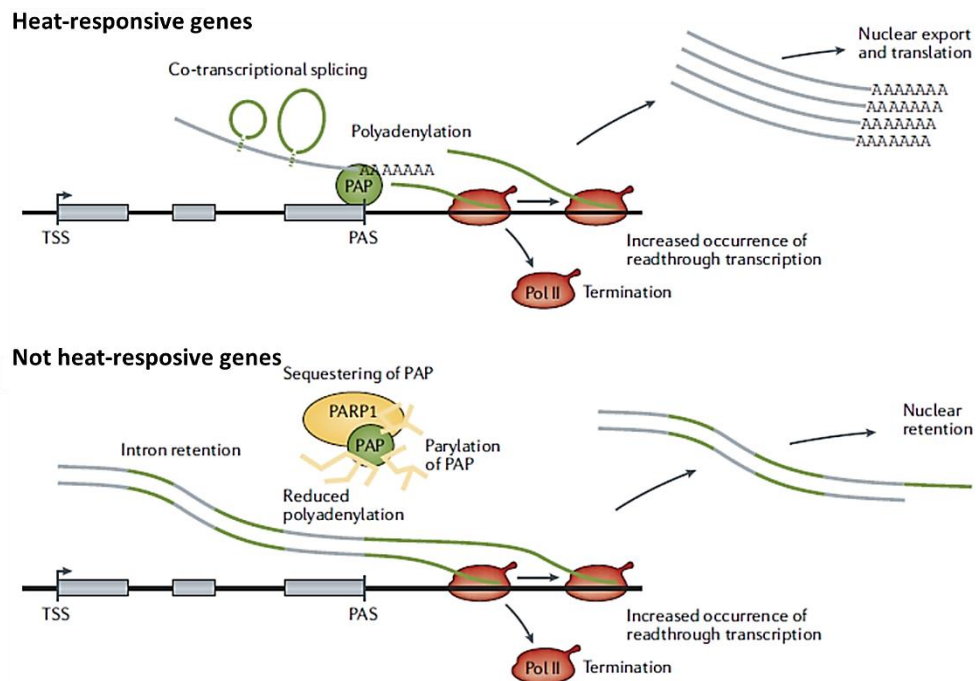
phosphorylation of the RNA Pol II CTD repeated motif at serine 5, the dissociation of the inhibitory factors associated with RNA Pol II and its subsequent release into the elongation phase (Fig.21).

In response to HS the occupancy of RNA Pol II decreases across thousands of genes and increases only at a few hundreds of others. However, most striking is the loss of RNA Pol II peaks in the 3' end of genes, coupled with its extended occupancy tens of kilobases downstream of 3' ends (Cardiello et al., 2018). It has been shown that this pattern results in a global and reversible loss of normal transcription termination during the HSR. Upon stress the occupancy of some termination factor was strongly reduced at the 3' end of representative genes, suggesting impairment of the classical molecular mechanisms of transcription termination. Furthermore, reduced polyadenylation following HS involves PARP1-mediated poly(ADP-ribose)ylation and subsequent removal of Poly-A Polymerase (PAP) from the 3' ends of the genes which are not stress-responsive (Fig.22) (Di Giammartino et al., 2013). The reduced recruitment of the termination machinery and the sequestration of PAP contribute to the widespread formation of shock-induced readthrough transcripts (Fig.22) (Cardiello et al., 2018). The termination defects were described as reversible since the usual RNA Pol II occupancy was largely restored after one hour of recovery.

How the stress-induced genes escaped the altered transcription termination remained an enigma until recently. One study demonstrated the importance of hTHO (subcomplex of the TREX) for the selective transcription termination and release of heat-responsive transcripts. Several factors from the hTHO were found to interact with the *HSPA1A* gene, a member of the HSP70 chaperone family (Katahira et al., 2019). The study of Katahira et al., revealed that the interaction of hTHO and *HSPA1A* gene, restrains the formation of R-loops and thus facilitates the transcriptional termination and release of the *HSPA1A* mRNA from its locus. On the other hand, the aberrant formation of R-loops prevents the processing and elimination of readthrough transcripts at *HSPA1A* and abolished the transcription termination. It is yet unknown whether this is a common mechanism for all stress-responsive genes.

The different studies on the HSR transcriptional modulation reveal the tight coordination of the mechanisms involved in the initiation and the termination of transcription (Vihervaara et al., 2018). The correct knowledge of the regulation of these events and the selective expression of

stress-responsive genes are needed to better understand the precise execution of the protective response.



**Figure 22: Transcription termination and splicing at heat-responsive and not heat-responsive genes following heat shock.** Upper panel: at heat-responsive genes co-transcriptional splicing, polyadenylation and transcriptional termination are taking place correctly and the mature transcripts are successfully exported for translation. Lower panel: at not heat-responsive genes PARP1-mediated parylation of PAP reduces polyadenylation, transcription termination is abolished and readthrough transcripts are forming and are retained in the nucleus for degradation. In addition, at these genes splicing is improper and leads to intron retention. *Adapted from Vihervaara et al., 2018.*

### mRNA processing during HSR

One of the key consequences of the HSR are altered splicing events or global inhibition of splicing which have been reported in numerous studies. HS leads particularly to intron retention which promotes the introduction of premature termination codon and the degradation of the mis-spliced mRNA through the Non-sense Mediated Degradation (NMD) pathway (Fig.21) (Hussong et al., 2017). Early studies have indicated that the intron retention at many mRNAs results from the modulated phosphorylation of SR proteins which is disrupted upon HS (Shalgi et al., 2014; Hussong et al., 2017). A novel mechanism, involving *SATIII* lncRNAs as indirect modulators of the phosphorylation of SR proteins was very recently discovered and will be presented in the

chapter on stress-induced non-coding RNAs (Ninomiya et al., 2020). The *SATIII* lncRNAs may contribute to the splicing alterations through the common titration of important splicing factors which has been reported earlier (Biamonti and Vourc'h, 2010; Denegri et al., 2001; Metz, 2004).

An intriguing and relatively uninvestigated question is how the stress-induced mRNAs get correctly spliced in order to be properly expressed? It is believed that the majority of *HSPs* escape the stress-induced splicing inhibition since they are intronless. However, some chaperones and other stress-responsive transcripts are subjected to splicing and are successfully expressed following HS (Fig.22). Indeed, studies demonstrated that in response to stress the co-transcriptional splicing of hundreds of genes involved in protein folding occurs correctly (Shalgi et al., 2014).

Another mechanism of mRNA processing involved in gene expression regulation is the Alternative PolyAdenylation (APA) of transcripts. Even though perturbations in APA have been already described in various stress conditions, genome-wide studies on APA upon HS specifically in human cells are still lacking (Sadek et al., 2019). Historically, broad APA events on different mRNAs, including some *HSP* mRNAs, have been revealed in *Drosophila* after thermal shock (Moran, 1986; Vazquez, 1991).

One study demonstrated that stress-induced APA of the *HSPA1A* gene results in the accumulation of transcripts with shorter 3'UTR in mice cells (Tranter et al., 2011). Furthermore, this shortening was associated with the elimination of suppressive miRNA binding site and the consequent robust increase of the chaperone expression. Afterwards, it was observed that the APA-mediated truncation of *HSPA1A* mRNA promotes the protein induction through enhanced polyribosome loading (Kraynik et al., 2015).

In this context, in human cells APA could be a mechanism of a tailored expression of certain genes, altering the function and regulatory traits of mRNAs during the HSR. Future investigations are needed to reveal this aspect of mRNA processing in stress conditions at the genome-wide level.

#### *Export of mRNAs during HSR*

In response to stress, besides the perturbations of transcription and the altered mRNA processing, the nuclear to cytoplasmic export of mRNAs is also regulated towards the blockage of “housekeeping” mRNAs and the export of stress mRNAs. The underlying diverse mechanisms



have been extensively studied in yeast, however studies on this subject in human cells are still rare. In *Saccharomyces cerevisiae* it was shown that the incompletely or incorrectly processed mRNAs are retained at or near their site of transcription by the nuclear exosome complex (Hilleren et al., 2001). In addition, more recent studies indicated that the stress-induced dissociation of the key export regulator Mex67 (human TAP-p15) with its adaptor proteins prevents the general mRNA export (Zander et al., 2016). In the same way this association is sustained at the level of *HSP* mRNAs in an HSF1-dependent manner and allows the export specifically of the stress-induced transcripts.

Regarding the stress-induced mRNAs, the export process following HS was studied for the chaperone *HSP70* mRNA. In human HeLa cells it has been observed that the interaction of an essential nuclear pore-associating protein (named TRP) with HSF1, brings it into proximity of the *HSP70* promoter and thus assists the export of the chaperone mRNAs in a stress-responsive mode (Skaggs et al., 2007). Later on, it was demonstrated that THOC5 (subunit of the TREX) is needed for the export of *HSP70* mRNA in HeLa cells (Katahira et al., 2009). The authors showed that THOC5 together with adaptor proteins, Aly/REF and TAP/NXF1, support the transcription-dependent loading of the mRNA export machinery onto *HSP70* nascent transcripts and drive their nuclear export. Indeed, this work showed that the knock-down of THOC5, Aly/REF or NXF1 leads to the accumulation of *HSP70* mRNAs near their transcription site indicating a nuclear retention of the transcripts. Along with these observations the latest study of Katahira et al. supports the idea of the involvement of THO/TREX in the transcription termination and the tightly linked mRNA export of stress-responsive transcripts (Katahira et al., 2019).

#### *mRNA processing and export through Nuclear Speckles*

A novel potential mechanism for the precise orchestration of the expression of stress-induced genes involves the previously presented membraneless mRNA processing and export factories, named Nuclear Speckles (NSs). Early studies have indicated the colocalization of *HSP70* mRNAs with a well-known marker of NSs- the splicing factor SC35 (or SRSF2) following HS (Metz, 2004). It was shown that the association of *HSP70* mRNA with NSs is critical for the HS-induced transcriptional activation of *HSP70* and leads to a net increase of gene expression correlating with an increase of the mRNA levels (Khanna et al., 2014; Kim et al., 2019). Moreover,

others demonstrated that the transit through NSs of many intronless mRNAs, including some isoforms of *HSP70*, is a critical step for their proper nuclear export (Wang et al., 2018). This work indicated that the passage through NSs promotes the recruitment of the TREX complex and supports the TREX-dependent export of *HSP70* and various intronless mRNAs. Importantly, the experiments in this study were performed through DNA microinjections and not in HS conditions, therefore the relevance of these results should be confirmed also in cells subjected to stress.

Overall, the association with NSs could facilitate the processing, the export and the expression of heat-induced mRNAs. Hence, the transit through these confined structures could present a general mechanism of regulation of stress-responsive transcripts during the HSR.

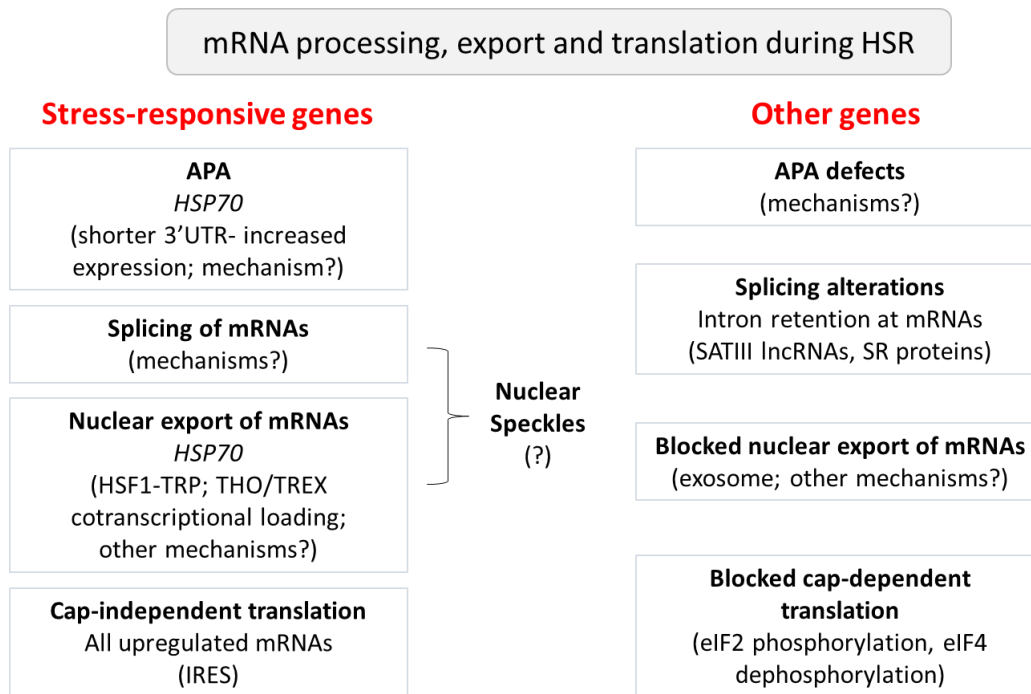
### *Translation during HSR*

The translation process, the final step of gene expression, is not an exception and is also modified and adapted to the cellular needs in response to HS. One of the most common mechanisms of inhibition of translation during HSR in mammalian cells is through the phosphorylation and consequent inactivation of the eukaryotic Initiation Factor 2 (eIF2). Various stress-induced kinases promote the phosphorylation of eIF2 while HSPs act as chaperones and prevent their misfolding (Holcik and Sonenberg, 2005). Another mechanism of translational suppression involves the dephosphorylation of eIF4E-binding proteins which promotes a block of the interaction of eIF4 with the 5' cap of mRNAs (Rhoads and Lamphear, 1995).

Importantly, during HSR the global translational inhibition is coupled with the formation of cytoplasmic Stress Granules (SGs). These stress-induced non-membraneous cytosolic foci are composed of non-translating mRNAs, translation initiation factors and other RNA-binding proteins implicated in mRNA function (Buchan and Parker, 2009; Douglas and Martin, 2009). When formed, SGs act as sites of triage where mRNAs can be stored upon the restoration of normal conditions or conversely primed for degradation (Anderson and Kedersha, 2008). In parallel, the preexisting polysomes, the hubs of proteins production, are either decomposed or the translation on them is stalled (Panniers, 1994). New polysomes form exclusively on stress-induced mRNAs arriving from the nucleus.

To escape the translational repression, the stress-induced transcripts possess a characteristic sequence, called Internal Ribosome Entry Site (IRES) which is often found in the 5'UTR but can also occur elsewhere in mRNAs (Leppek et al., 2018). The IRES binds with the preinitiation

complex in the 40S ribosomal subunit and positions the mRNA start codon in a cap-independent manner without prior 5'UTR scanning (Kim and Jang, 2002). Therefore, translation of the HS-upregulated mRNAs consists in the so called “jumping” of the 40S subunit rather than the regular complete scanning of the 5'UTR during initiation of translation.



**Figure 23: Gene expression regulation via modulation of mRNA processing, export and translation in response to heat shock.** The HSR-induced genome reprogramming is associated with the modulation of mRNA processing (APA, splicing), nuclear to cytoplasmic export and translation of both stress-responsive and other genes. The expression of stress-responsive genes is achieved through particular mechanisms, some of which are still poorly understood. The rest of the genome is shutdown via diverse rapid and reversible changes of the regulatory processes.

Overall, a general deregulation of all fundamental steps of mRNA processing, export and translation can be observed upon stress induction and activation of the conserved HSR pathway (Fig.23). The underlying mechanisms demonstrate the remarkable orchestration of these steps upon HSR, set up in order to pause all the non-essential processes ongoing in the cell, to prevent future damages of newly synthesized proteins and to selectively upregulate the production only of the essential ones, needed to preserve the cellular homeostasis.

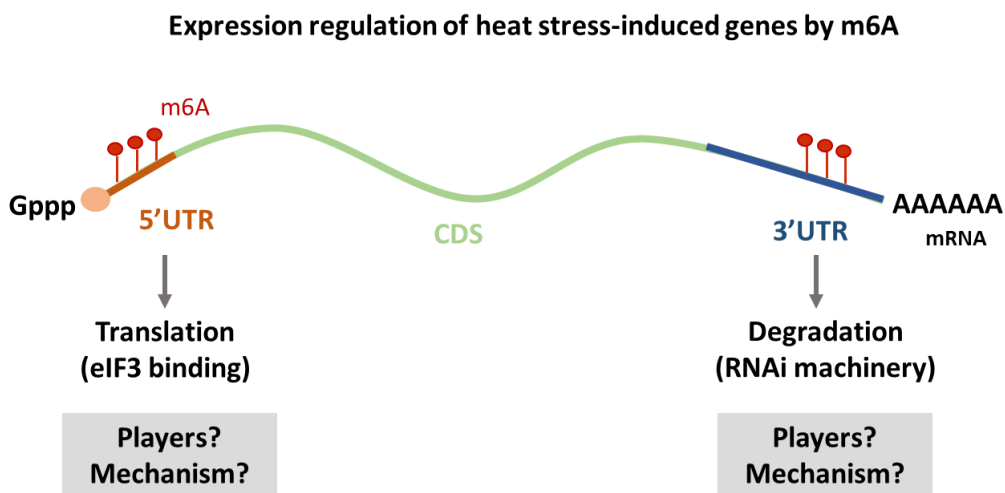
### Role of the epitranscriptome in gene expression regulation during HSR

Recently, the emerging field of the epitranscriptome provided additional information regarding the stress-induced gene expression modulation. Several studies addressed the implication of RNA modifications, and more precisely of m6A, in the context of heat stress, oxidative stress, UV exposure, X-ray irradiation or chemical exposure and neuronal stress (Anders et al., 2018; Engel et al., 2018; Knuckles et al., 2017; Xiang et al., 2017; Zhou et al., 2015; Zhang et al., 2020). In the majority of these studies the performed analyses revealed the dynamic pattern of m6A. Following UV exposure m6A levels were found to rapidly and transiently increase at the DNA damage sites (Xiang et al., 2017). Moreover, m6A turn out to be essential for the execution of the UV-induced DNA damage response. In addition, as described previously, the more recent work on METTL3 and YTHDC1 unraveled their role in homologous recombination-mediated repair of double-strand breaks following X-ray irradiation or chemical exposure (Zhang et al., 2020). The neuronal stress exposure and subsequent analysis of the epitranscriptome unraveled the importance of m6A in synaptic plasticity in adult brain (Engel et al., 2018). In response to oxidative stress it was demonstrated that some mRNAs are additionally modified in their 5'UTR which drives their relocation from the translatable pool to Stress Granules (SGs) (Anders et al., 2018). Recently, the members of the YTHDF family of m6A readers were found to be essential for the triaging of m6A-modified mRNAs to the SGs formed upon oxidative stress (Fu and Zhuang, 2020).

In response to HS it has been observed that residues within the 5'UTR of newly transcribed stress-responsive genes can be preferentially methylated. However, conversely to the effects detected following oxidative stress, the increased m6A levels upon HS are believed to promote the expression of stress-responsive transcripts. The upregulated m6A levels in the 5'UTR of *HSP70* mRNA were found to support the cap-independent translation mechanism of the chaperone (Zhou et al., 2015). The increased m6A abundance correlated with the elevated polysome occupancy and translation efficiency at *HSP70* mRNA. It was suggested that this process relies on the cytoplasmic m6A reader, the protein YTHDF2. The proposed mechanism indicated that YTHDF2 translocates to the nucleus in response to HS, binds the methylated residues of HSPs and prevents them from demethylation by FTO, thus further allowing the m6A-dependent and cap-independent translation. However, this model was contradicted by more recent work, which showed that YTHDF2 persists in the cytoplasm and does not relocalize to the nucleus under stress conditions (Ries et al., 2019). Although, the precise mechanism is still poorly understood, the role of 5'UTR m6A residues has

been confirmed as an essential component of the cap-independent translation of stress-induced mRNAs. Particularly, the increased m6A levels in the 5'UTR of *HSP70* mRNA were shown to serve as a signal for the selective binding of the eukaryotic Initiation Factor 3 (eIF3) followed by the recruitment of the ribosomal 40S subunit and the concomitant translation activation (Meyer et al., 2015). Along with this, decreased m6A levels were found to be associated with reduced cap-independent translation following HS. Therefore, these results validated the role of m6A in the particular stress-related mechanism of translation and highlighted the importance of the RNA modification for the expression of heat-induced transcripts.

Conversely, others assigned a role of m6A in the degradation of *HSP* transcripts. The levels of m6A were shown to increase in the 3' UTR of mRNAs upon HS (Knuckles et al., 2017). This increase was found to promote the co-transcriptional interaction of METTL3 and components of the RNA interference (RNAi) machinery (Knuckles et al., 2017). Since the timing and the magnitude of the HSR need to be precisely regulated, it was proposed that acute heat stress cotranscriptionally marks *HSP70* mRNAs for subsequent RNA degradation by RNAi components and prevents the potentially harmful accumulation of stress-related transcripts.



**Figure 24: Model for the implication of m6A in the expression regulation of heat stress-induced genes.** The levels of the RNA modification m6A are found to increase upon HS at the 5'UTR and 3'UTR regions of stress-responsive mRNAs. It is proposed that m6A at the 5'UTR allows the cap-independent translation of the stress-induced genes. The m6A increase at the 3'UTR could mark the stress-responsive transcripts for degradation by the RNAi machinery. The exact mechanisms and molecular players involved in both processes are still poorly understood.

Based on these observations, the levels of m6A seem to play a central role in the production of HS-responsive genes (Fig.24). However, the exact mechanisms need to be further resolved. The controversial results on the implication of YTHDF2 in the cap-independent translation of m6A-moidied transcripts suggest that other m6A readers could be involved in the process. It is of note that since RNA modifications are novel central regulators of gene expression, it is not yet known whether their abundance and distribution might correlate with various stress-related processes and provide fine tuning of the global genome reprogramming. The implication of m6A in the massive shutdown of the genome upon HS, cannot be excluded, however this question has not been addressed yet. Therefore, additional studies are needed to decipher the role of RNA modifications in the context of cellular stress and could bring new significant insight for the establishment and the proper implementation of the HSR.

### **II.II.3 The Heat Shock Factor 1 (HSF1)**

As mentioned previously for longtime the main orchestrator of the stress-induced expression of particular coding and non-coding genes was considered to be HSF1. This ubiquitous transcription factor is capable of precisely integrating the stress signals and translating them into a protective program. It is of note that although for several decades HSF1 has been described as the key actor of the HSR, recent transcriptomic studies suggest novel regulators of the stress response program (Mahat et al., 2016). Mahat et al. confirmed that HSF1 is critical for the expression of HSPs, other chaperones and additional 200 genes during HS. However, the modulation of the majority of stress-responsive genes turned out to be HSF1-independent. This study identified a novel regulator of particular stress-responsive genes- the Serum Response Factor (SRF), completely uncharacterized previously. Therefore, this recent discovery highlights the complexity of the HSR, which involves multiple layers of regulation mechanisms, potentially driven by diverse molecular actors, the majority of which remain to be uncovered. In this work we have been interested in the expression of HSF1-dependent stress-induced coding and non-coding genes and here the structure, the activation and the main functions of HSF1 will be described. In the next parts, the most characteristic HSF1 coding targets- the chaperones, will be introduced followed later on by the presentation of the non-coding HSF1-dependent genes such as the lncRNAs *SATIII*.

In mammals, HSF1 belongs to the HSF family composed of four different members, HSF1 to HSF4. All of these factors possess unique and overlapping functions, tissue specific expression and can be subject of diverse posttranslational modifications. The transcription factor from this family which responds to the classical heat shock stimuli is HSF1, (Morimoto, 1998).

### Structure of HSF1

HSF1 is composed of four different functional domains- DNA-binding domain (DBD), Trimerization Domain (TD), Regulatory Domain (RD), and a TrAnsactivation Domain (TAD).

The DBD is the best preserved domain in evolution and was identified through studies of the activation of the transcription of HSPs. It was shown that DBD binds HSEs (Heat Shock Elements) - specific motifs found at the proximal promoter regions of stress-induced genes (Flick et al., 1994). The HSE is composed of at least three contiguous inverted repeats: nTTCnnGAAnnTTCn (Xiao and Lis, 1988). Each DNA-Binding Domain (DBD) of HSF1 recognizes the HSE motif and associates with it (Wu, 1995).

The trimerization domain of the protein is responsible for the trimerization-dependent activation of HSF1. In normal growth conditions HSF1 is maintained in a monomeric state while in stress conditions the monomers undergo conformational change to form trimers. This multimerization is essential for the high affinity DNA binding in eukaryotes (Sarge et al., 1993).

The RD prevents HSF1 activation in the absence of protein damage. The RD domain is the subject of numerous PostTranslational Modifications (PTMs)- hyperphosphorylation, sumoylation and acetylation which were shown to support its repressive ability in normal conditions (Anckar and Sistonen, 2011).

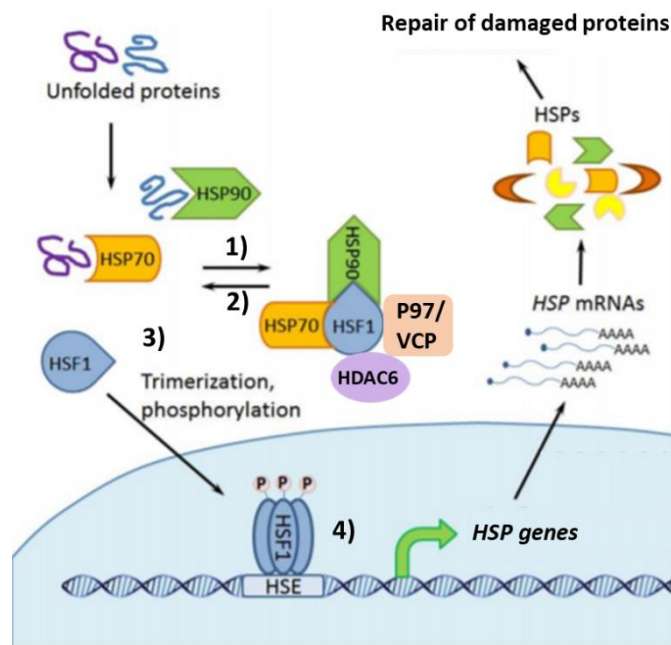
The TrAnsactivation Domain (TAD) regulates the magnitude of activation of HSF1 and facilitates the transcriptional activation of its targets. The transactivation of HSF1 is controlled mainly by the RD which is able to repress the TAD (Newton et al., 1996).

In addition to the four conserved domains, HSF1 presents a motif able to suppress the spontaneous trimerization, named Trimerization Repression Motif (TRM). The repressive mechanism is essential for keeping HSF1 in an inactive state in normal conditions (Wu, 1995).

### Activation of HSF1

The important variety of stimuli able to activate the HSR support the hypothesis that several regulation pathways are involved in the activation of HSF1. To be transcriptionally competent HSF1 needs to be released from its repressor complex, to be hyperphosphorylated, to trimerize and to accumulate in the nucleus (Fig.25).

HSF1 is mainly cytoplasmic in its monomeric form and is incapable of binding DNA. In this state the protein is inactivated and sequestered in the cytoplasm via the interaction with a repressive complex formed by two chaperones- HSP70 and HSP90, the Histone DeAcetylase 6 (HDAC6) and the AAA ATPase factor P97/VCP (Fig.25) (Baler et al., 1992; Pernet et al., 2014). Following HS, HDAC6 dissociates (together with P97/VCP) from HSF1 and binds to polyubiquitinated proteins. Following this, HSP70 and HSP90 bind to the unfolded proteins, liberating HSF1. HSF1 trimerizes, is hyperphosphorylated and translocates to the nucleus to activate the transcription of *HSP* genes. The newly synthesized HSPs continue to repair the damaged proteins until protein homeostasis is reinstated.



**Figure 25: Activation of HSF1 upon heat stress and accumulation of unfolded proteins.** In normal conditions HSF1 is maintained inactive in a repressive complex with HDAC6, P97/VCP, HSP90 and HSP70. Upon heat stress, HDAC6 and subsequently P97/VCP dissociate from HSF1. The unfolded proteins accumulate and HSP70 and HSP90 bind them, liberating HSF1. HSF1 trimerizes, is hyperphosphorylated and translocates to the nucleus to activate the transcription of *HSP* genes. The newly synthesized HSPs continue to repair the damaged proteins until protein homeostasis is reinstated.



Meanwhile HSP70 and HSP90 bind the accumulating unfolded proteins and liberate HSF1. Importantly, the HS-dependent upregulation of HSP70 and HSP90 acts as negative feedback loop on HSF1 transactivation capacity (Shi et al., 1998). This model explains the inactivation of HSF1 in the presence of “unemployed” chaperones and its rapid activation when chaperones bind unfolded proteins and the repressive complex dissociates (Fig.25).

The exposure of cells to heat stress, makes HSF1 homotrimerize, translocate to the nucleus and acquire DNA-binding and transcription-stimulating capacities. A layer of complexity regarding HSF1 regulation is introduced by the multitude of PTMs of the protein. Notably, hyperphosphorylation of HSF1, occurring mainly at the level of the regulatory domain, renders it transcriptionally active (Fig.25) (Budzyński et al., 2015).

HSF1 activation is not strictly dictated by the increased temperature but moreover dependent on the cellular context. Several studies demonstrated that the HSF1 activation threshold could be adapted to the cell type. For instance, in mammalian cells HSF1 can be triggered after exposure for one hour at 39°C in T-lymphocytes or at 42°C in fibroblasts (Gothard et al., 2003). Therefore, the conditions of HS should be optimized according to the used cell line in order to assure proper recovery and prevent lethality.

The reversible character of the HSR relies on the ability of HSF1 to become rapidly inactivated once the molecular actors who tackle the heat-driven perturbations are no longer needed. Sumoylation of particular residues in the regulatory domain promotes DNA dissociation and consequently the nuclear export of HSF1 which becomes again sequestered in the cytoplasm through the interaction with HSP70 and HSP90 (Liebelt et al., 2019).

### Functions of HSF1

In addition to protecting cells against proteotoxic stress, HSF1 has an important role in many physiological processes and pathological states. In development HSF1 and HSF2 are needed for oogenesis, spermatogenesis and brain development (Åkerfelt et al., 2010). In mouse, *HSF1* depletion leads to growth retardation, female infertility or prenatal lethality and inability to trigger the HSR (Xiao, 1999).

HSF1, as well as HSPs have a general positive impact on lifespan by maintaining the proteome homeostasis and assuring health of the organism (Westerheide et al., 2009).

Importantly, elevated levels of HSF1 are detected in various human cancers (Mendillo et al., 2012). In line with this, HSF1 deficient mice present lower incidence of tumors and increased survival in carcinogenesis models (Dai et al., 2007). The enhanced activity of the main regulator of the HSR is proposed to be essential promoter of tumorigenesis and is therefore a potential therapeutic target for cancer treatment (Whitesell and Lindquist, 2009).

The enigma of the controversial functions of HSF1 lies in the importance of its appropriate and tightly balanced activation. The high levels of active HSF1 are protective in stress conditions but can also promote the cancerous cell phenotype. On the other hand, its incorrect or insufficient activation leads to the development of pathologies associated with the accumulation of protein aggregates.

## **II.III The transcriptional protein-coding response to HS**

### **II.III.1 The Heat shock proteins (HSPs)**

In response to HS the most specific genomic targets of HSF1 are the Heat Shock Protein-coding genes. Even though HSPs account for less than 5% of the HS-upregulated genes, their induction is the strongest (2 to 40-fold increase depending on the chaperone) and they are considered to be the hallmark of the HSR in mammals. Recently, it has been demonstrated that all *HSP* genes presenting an HSF1 enrichment at their promoter region are robustly induced upon HS (Mahat et al., 2016). Strikingly, many genes in the HSP family which are not bound by HSF1 were found to be transcriptionally repressed following thermal shock suggesting another level of regulation of HSPs or a potential implication of some chaperones in other stress responses.

The HSPs are the cellular “warriors” against protein misfolding and aggregation in stress conditions and the “guardians” of the precise protein folding in normal conditions. The threedimensional fold of proteins governs their biological functions. The exact folding is directed by the primary amino acid sequence and both weak and covalent interactions support the formation of the native conformation. Interestingly, chaperones are expressed even at physiological temperatures since there is a constant need for assistance during *de novo* protein folding and refolding as the stability of cellular proteins is relatively low and competes with folding (Mayer, 2010).

The family of HSPs comprises five broadly conserved classes, established based on the molecular weight of the proteins- small heat shock proteins (sHSPs, HSPB, HSPE), HSP60s (HSPE), HSP70s (HSPA) (and HSP40s), HSP90s (HSPC) and HSP110s. Together the HSPs account for approximately 100 different proteins with central implication in the proteostasis network.

All HSPs interact promiscuously with a broad variety of unfolded proteins. When proteins start to unfold they get recognized by chaperones through the increased exposure of hydrophobic amino acids. The binding occurs to hydrophobic segments, specific peptide sequences or structural elements of the altered proteins (Wu et al., 2019). The molecular chaperones prevent unwanted intermolecular interactions rather than conveying structural information. The mechanism is based on the constant binding and release of nonnative proteins, which is modulated by the affinity of HSPs for their substrate (Weibezahn et al., 2005). The switch of affinity states is controlled through the binding and hydrolysis of ATP, except for the sHSPs. ATP-independent chaperones may function as additional “holdases”, which hold the substrate proteins and bring them to the folding competent state by preventing their aggregation. The efficient folding is achieved when the rate of folding is faster than the rates of aggregation and chaperones rebinding. If the folding is slower, the chaperones “client” can be transferred to a different chaperone-based mechanistic system or alternatively to the degradation machinery. If the concentration of folding intermediates exceeds the available chaperone capacity, aggregation occurs (Kim et al., 2013). In such situation, the cellular stress response is triggered and the chaperones expression is induced in order to increase their abundance and restore the balance.

HSPs are ubiquitous and their localization and function vary within the cell. Moreover within each subfamily exist many isoforms which can be either inducible or constitutive. The inducible chaperones are present at very low levels in normal conditions and are strongly induced following stress. The constitutive forms are present usually at relatively high levels and their expression is additionally increased upon HS. The cellular localization (in normal and stress conditions), the particular roles and the mode of expression of the five HSPs classes are summarized in Fig. 26.

Class of HSPs	Cellular localization	Function	Expression
sHSPs	Mitochondria, Plasma membrane	Protein folding (« holdases »)	Inducible
HSP60s	Mitochondria	Protein folding; Protein assembly by forming hetero-oligomeric protein complex	Constitutive and Inducible
HSP70s (& HSP40s)	Cytosolasm, Mitochondria Endoplasmic reticulum, Nucleus (stress)	Protein <i>de novo</i> folding and refolding of denatured proteins; protein transport; protein degradation in	Constitutive and Inducible
HSP90s	Cytoplasm	Folding of native proteins	Constitutive
HSP110s	Cytoplasm, Nucleus (stress)	Solubilization of aggregates	Constitutive and Inducible

**Figure 26: Localization, functions and expression pattern of the different Heat Shock Proteins.** The HSPs are presented in five classes according to their molecular weight (small heat shock proteins (sHSPs), HSP60s, HSP70s, HSP90s and HSP110s). The cellular localization (in normal or exclusively in stress conditions) and the corresponding functions of the members of each class are indicated.

The HSP70s class is the most extensively studied group of stress proteins, involved in wide range of quality control functions. They assist *de novo* folding, refolding of denatured proteins, protein transport and degradation. Most HSP70s are strictly stress-induced and present at very low levels in normal conditions. HSP70 present many isoforms which share a remarkable homology of 95% identity of the DNA sequence. The functioning of HSP70s is assisted by the HSP40s and nucleotide exchange factors (NEF). In the group of 40 kDa chaperones more than 40 different actors have been described up to date (Kampinga and Craig, 2010). HSP40s (also named DnaJ proteins) comprise a characteristic J domain which can bind the ATPase domain of HSP70s. The HSP40-HSP70 system protects nascent chains against irregular interactions and moreover assist the co- and posttranslational protein folding. A small proportion of the HSP40s can function as chaperones independently and recruit HSP70s to nonnative proteins, therefore even if HSP70s are located in the cytoplasm, some HSP40s can be found in mitochondria or at the endoplasmic reticulum (Kampinga and Craig, 2010). In response to stress HSP70s can be found also in the nucleus.

The other classes of chaperones work also mainly in a cooperative manner with each other in order to assure the proper folding and to prevent aggregation (Hattendorf, 2002).

Importantly, chaperones assist also protein degradation when the repair mechanisms are insufficient and the accumulation of damaged proteins becomes harmful for the cells. Various HSPs are known to deliver target proteins to the proteasomal subunits. For instance, HSP70 and HSP90 cooperate with several E3 ubiquitin ligases in order to mark unfolded and/or aggregated substrates for degradation (Young et al., 2003).

The cellular proteome is governed constantly by the complex network of chaperones. It is of note that the kinetics of chaperones expression appear to be subjected to strict controls during the HSR. Although some chaperone's transcription is activated in only few minutes, their protein levels can increase later during the recovery period following HS suggesting a transcription-translation uncoupling and possibly specific processing steps prior to their expression (Abravaya et al., 1991). The differential expression patterns can reflect the various roles of HSPs- some are needed very rapidly to assist the chaperones already present in cells, while others, induced later reinforce the combat against protein perturbations (Tutar et al., 2006).

While major research advances unraveled the functions of individual chaperones, we are far from understanding their vast common mechanisms. The functioning of all HSPs is interconnected and their action is cooperative, therefore dysregulation of one sole chaperone might lead to the collapse of the protective HSR system and endanger the cellular homeostasis and survival.

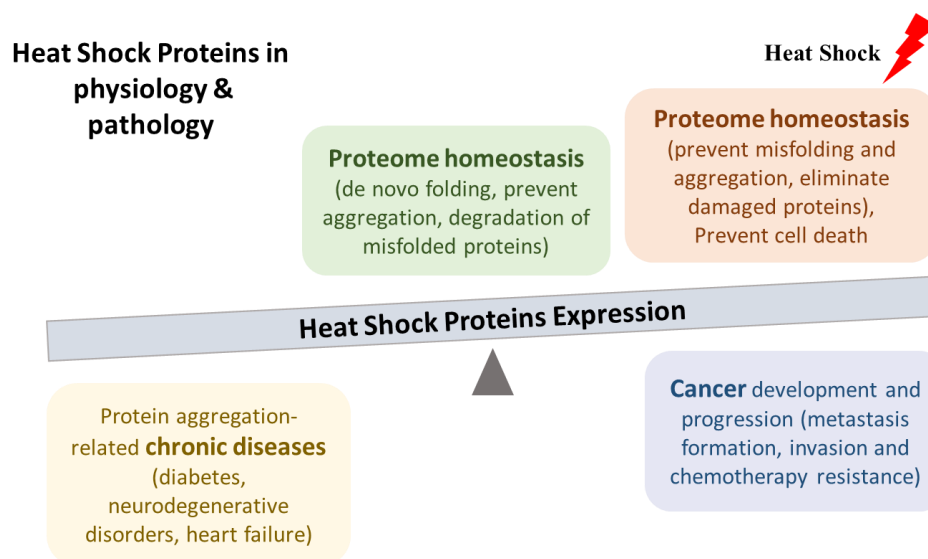
### **II.III.2 HSPs in cellular homeostasis and disease**

Beyond their crucial role in protein homeostasis HSPs are involved in numerous stress-related, physiological and pathological processes. Indeed, accumulating evidence during the past two decades demonstrates that in order to cope with stress perturbations chaperones are implicated as well in the control of cell cycle and apoptosis. For instance, HSP70 and HSP90 are needed for the proper cell cycle progression and prevent blockage in G1 or G2/M phases (Antonova et al., 2019; Truman et al., 2012).

In the regulation of cell death HSP27, HSP70 and HSP90 have been identified as anti-apoptotic proteins acting at multiples levels of the intrinsic and extrinsic pathways (Wang et al., 2014). These properties are needed to preserve cell survival in stress conditions since the cellular damages often result in rapid activation of apoptosis.

In the physiological context HSPs are extensively studied for their implication in development. It is known that various chaperones assist gametogenesis, the pre-implantation development, the extra-embryonic development and organogenesis, however many of the relying mechanisms are still poorly understood (Christians et al., 2003).

On the other hand, HSPs are found deregulated in diverse pathologies. Many chronic diseases associate with protein aggregation- type 2 diabetes, neurodegenerative diseases (Parkinson, Alzheimer’s and Huntington’s), autoimmune diseases, obesity and chronic heart failure. Numerous studies are revealing the beneficial role of HSPs in the combat against protein aggregation and the prevention of the onset of protein conformational disorders. Significant evidence indicates that chaperones are suppressors of neurodegeneration, and therefore promising therapeutic targets (Edkins et al., 2018; Muchowski and Wacker, 2005).



**Figure 27: Heat Shock Protein expression balance in physiology and pathology.** In physiology HSPs assure the protein homeostasis in normal growth conditions and prevent protein misfolding and aggregation and cell death after heat stress. In pathology the decreased HSPs levels associate with chronic diseases, while their increased expression promotes cancer development and progression.

The role of HSPs in chronic diseases is dual since in the case of cancer increased expression of HSPs promotes the development and progression of the disease. Cancer cells are rapidly dividing and need to survive a strongly hypoxic microenvironment. Therefore the expression of multiple HSPs is found increased in nearly every type of solid tumor and associates with poor prognosis (Calderwood and Gong, 2016). The abnormally high levels of chaperones support cancer cell's survival, metastasis formation and confer resistance to chemotherapy by interfering with apoptosis signaling.

The diverse key functions of HSPs in physiology and pathology reveal the capital importance of their correct expression regulation (Fig.27). Hence, understanding the implication of various molecular actors in the processes of chaperone expression is crucial not only for the study of fundamental cellular processes in normal and stress conditions but also in the context of pathology initiation and progression, targeted therapies and biomarkers development.

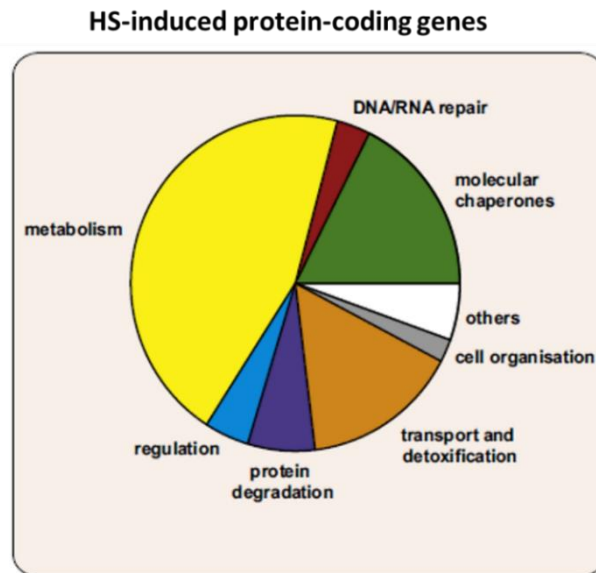
### **II.III.3 Other HS-induced protein-coding genes**

The dynamic research on the HSR and its implication in physiopathology led to the identification of numerous HS-upregulated coding genes in addition to the classical stress-induced chaperones. The stress-modulated proteins, targeted by HSF1 or other still unknown modulators, present a large functional diversity and their increased expression is intuitive in the scenario of cells combating the numerous molecular perturbations arising in stress conditions (Fig.28).

Many actors of the proteolytic system, needed for the elimination of aggregated and potentially harmful proteins from the cell, are rapidly upregulated during HS (proteasomal subunit PSMD10, Siah-interacting protein and others). A second group of stress-responsive genes includes RNA and DNA-modifying enzymes which repair the non-physiological covalent modifications of DNA and maintain proper DNA processing (Jantschitsch and Trautinger, 2003). The transcription of metabolic enzymes, essential for the restoration of the cellular energy supply and for the stabilization of key catabolic reactions, is activated during HSR (for example arylamine N-methyltransferase and a cytochrome B5 reductase) (Malmendal et al., 2006).

The HS-upregulated proteins comprise different regulatory proteins- transcription factors and kinases involved in the HSR pathway or in the inhibition of cellular processes arrested upon stress (Murray et al., 2004). Importantly, in this group are found also various regulators of cell

cycle and apoptosis such as MEN, p21, c-jun, btg-1, ERBB3, and Caspase 9. Their expression increases during the recovery period following HS in order to reestablish the regular cell proliferation. One of the most extensively studied heat-induced genes code for cytoskeleton proteins and extracellular matrix components, involved in sustaining the cellular morphology, the general cell organization and the microenvironment, which are found severely deregulated after stress. This class of stress-induced proteins includes vinculin, keratin, N-cadherin, integrin, collagens, elastin, fibrillin and others. In addition, particular proteins implicated in intracellular transport and the elimination of toxic compounds are also strongly expressed in response upon HS (G proteins, monoamine transporters and leukocyte receptor cluster 4) (Murray et al., 2004).



**Figure 28: Functional categories of Heat Shock-induced protein-coding genes.** The stress-induced proteins include chaperones but also proteins involved in: metabolism, DNA/RNA repair, regulation of diverse processes (cell cycle and cell death, transcription, etc.), protein degradation, transport and detoxification and cell organization (cytoskeleton genes). Others implicated in different cell functions can also be upregulated following heat stress. *Adapted from Richter et al., Molecular Cell, 2010*

The variety of stress-induced proteins highlights once again the complexity of the HSR and the importance of the tight gene expression regulation accompanying this protective response (Fig.28). Nevertheless, for longtime the stress-related gene expression time-scale and the mechanisms of transcription regulation of HS-modified genes, have not been addressed at the genome-wide level.



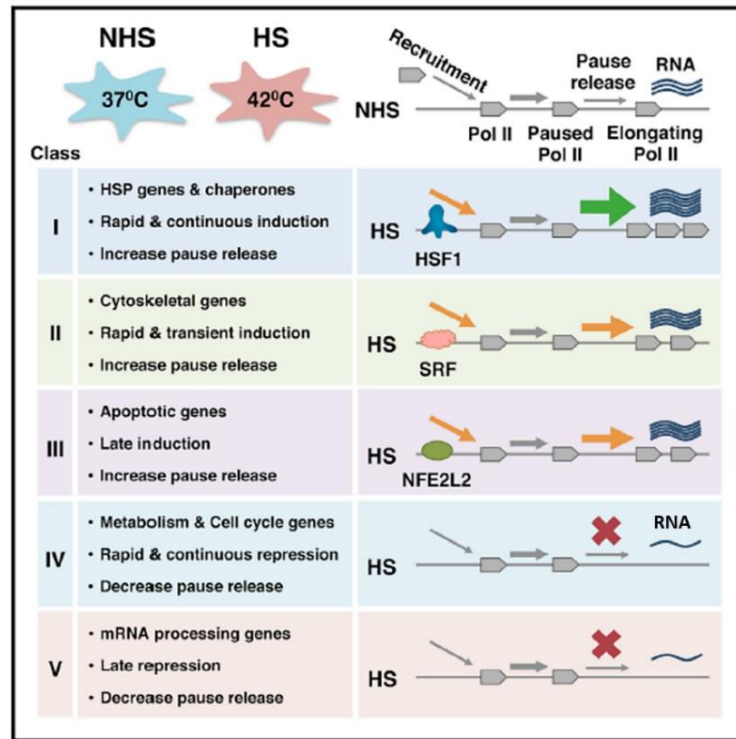
#### **II.III.4 Classification of the HS-induced and HS-repressed genes**

A general picture of the HS-induced transcriptomic reprogramming of both upregulated and downregulated genes in MEFs (Mouse Embryonic Fibroblasts) was recently unraveled by the study of Mahat and colleagues and will be discussed here (Mahat et al., 2016).

Using Precision nuclear Run-On sequencing (PRO-seq), this work presented a new mechanistic understanding for the coordinated regulation of the genome upon HS. The modulated genes have been grouped into five functional classes based on their nature, expression kinetics and their transcriptional regulation (Fig.29). In the proposed model the first three classes concern the stress-induced genes which are needed to counteract the consequences of heat stress. Class I includes HSPs and chaperones, mainly targeted by HSF1 and expressed rapidly and continuously during the thermal shock (Fig.29). As already described, the transcription of these genes is dependent on the HSF1-stimulated RNA Pol II release from paused to elongating state. The second class comprises the early and transiently induced genes (Fig.29). In this category are found many cytoskeletal genes (such as stress fibers and actin filaments) which remarkably turn out to be activated and upregulated only after 2.5min of HS. The SRF (Serum Response Factor) transcription factor turned out to be strongly enriched at the promoter region of cytoskeletal genes and was therefore described for the first time as novel central actor of the HSR. The third proposed class involves mainly apoptotic genes whose expression increases late during the HSR and mostly in an HSF1-independent manner (Fig.29). Some of these genes presented enrichment of the Nuclear Factor Erythroid 2-Like 2 (NFE2L2) at their promoter region and thus were suggested to be regulated via the binding of this transcription factor. In general this study confirmed that all HS-induced genes are characterized by an increased pause release of the RNA Pol II, allowing robust activation of transcription upon stress perturbations.

On the other hand the downregulated genes can also be grouped in two functional classes according to their nature, repression kinetics and polymerase occupancy. In this study, the class IV refers to the genes which are immediately downregulated upon HS and present an inhibition of the RNA Pol II release into elongation (Fig29). These genes include various regulators of metabolism, of the cell cycle and of protein synthesis. The last class, described in this work, relates to the genes which are repressed during the late phase of the HSR also via decreased pause release of RNA Pol

II (Fig.28). This V<sup>th</sup> category comprises factors implicated in splicing, mRNA processing and nuclear transport.



**Figure 29: Classes of Heat Shock-induced and Heat Shock-repressed protein-coding genes in MEF cells.** Five classes of stress-modulated genes, grouped according to their nature, their stress- inducible or repressed pattern, their expression kinetics and the RNA Pol II occupancy at their promoter region. The main (known) regulators of the inducible genes are indicated (HSF1, SRF, and NFE2L2). *From Mahat et al., 2016*

This broad classification of the HS-regulated genes demonstrates once again the complexity of the HSR and the importance of its precise orchestration. Even though the use of novel techniques shed light onto the stress-related transcriptional program, the mechanisms of gene expression regulation, independent from the activation of HSF1, are still poorly understood and largely unknown. Additional studies will be needed to identify novel HSR actors and to further comprehend the network supporting cell survival in the context of stress.

## II.IV The transcriptional non-coding response to HS

### II.IV.1 Overview of the non-coding response to HS

A growing amount of evidence now indicates that effectors of the HSR include not only protein coding-genes but also specific ncRNAs, part of which are already known to have regulatory functions. Upon HS, HSF1 was found to activate the transcription of *SATIII* lncRNAs, enhancer RNAs (eRNAs), Short Interspersed Nuclear Elements (*SINEs*), Telomeric Repeat containing RNA (*TERRA*) and the Nuclear Enriched Abundant Transcript 1 (*NEATI*). The transcriptional control and molecular roles of these ncRNAs have been thoroughly reviewed by Vourc'h et al., 2020. In addition to the HSF1-dependent transcription of ncRNAs, induced expression of other non-coding genes (Heat Shock RNA-1 (*HSRI*), *7SK* lncRNA and some miRNAs) has been observed in response to HS (Peterlin et al., 2012; Shamovsky and Nudler, 2009). The diversity and abundance of ncRNAs, upregulated or expressed particularly in the context of the HSR must be taken into account in the studies on the implication of this conserved pathway in physiology and pathology. The transcription of non-coding genes is involved in the correct execution of the HSR, the cell recovery process and the global protection of cell survival (Goenka et al., 2016; Traner et al., 2011; Vourc'h et al., 2020). More importantly, the HS-modulated ncRNAs can participate to the long-term effects of stress through the generation of genomic instability and could unravel still poorly known disease mechanisms linked to the deregulated expression of ncRNAs. Still many open questions need to be further examined to resolve the gene expression regulation and the precise roles of the non-coding HSR which at this stage remains mostly a black box.

One of the most extensively-studied lncRNAs expressed following HS are the *SATIII* RNAs. These lncRNAs are at the center of several questions addressed during my PhD work. Therefore, in this last part of the introduction, I will present the known characteristics and molecular functions of *SATIII*.

#### **II.IV.2 HSR, SATIII lncRNAs transcription and nuclear Stress Bodies formation**

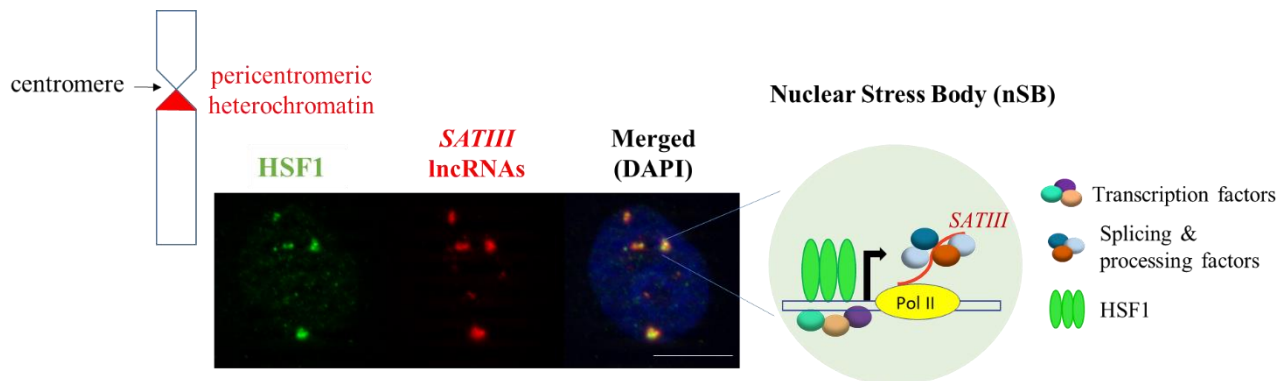
Historically, it was observed that in response to heat shock HSF1 presents a characteristic localization into bright nuclear foci. Surprisingly HSF1 foci did not correspond as expected to HSPs genomic loci (Jolly et al., 1997). This significant discovery indicated for the first time the existence of HSF1 targets, other than chaperones.

Consistent with its activation, HSF1 translocates to the nucleus following HS and transiently forms these large irregularly shaped granules which turned out to be distinct from other known nuclear structures (Cotto et al., 1997; Jolly et al., 1997). The number of HSF1 foci correlated with the ploidy of the studied cells confirming the chromosomal nature of the target (Denegri et al., 2001; Jolly et al., 1997). Interestingly HSF1 granules were observed in a large variety of human and primate cells only, suggesting an evolutionary distinction between humans and mammals such as rodents (Biamonti, 2004; Jolly et al., 1997). These curious novel nuclear structures were termed nuclear Stress Bodies (nSBs) (Biamonti, 2004).

Fluorescence *in situ* hybridization assays showed that nSBs form at the level of pericentric repetitive regions of chromosomes (Denegri et al., 2001; Jolly et al., 1997). Further analyses indicated a direct interaction between HSF1 and very large pericentric heterochromatic regions (up to 5Mb) enriched in epigenetic repressive marks (Amor et al., 2004). In humans the pericentric heterochromatin is composed of highly repeated DNA satellite sequences (*satI*, *satII* and *satIII*). Among them the *satIII* sequences present a consensus motif- GGAAT which resembles the canonical HSE present on HSPs gene promoters. The *satIII* sequences are enriched mostly at chromosomes 1, 9, Y and acrocentric chromosomes (Eymery et al., 2010). Following HS the pericentric region of chromosome 9 (locus 9q12) was characterized as primary HSF1 target and formed large nSBs, while secondary, smaller nSBs are formed at chromosomes 16 and Y, and acrocentric chromosomes (Fig.30) (Jolly et al., 2002).

Strikingly, the interaction of HSF1 with *satIII* sequences resulted in the rapid activation of the transcription of lncRNAs, named *SATIII* (Jolly et al., 2004). It has been rapidly revealed that nSBs correspond in fact to transcription factories where *SATIII* lncRNAs are transcribed in an HSF1 and RNA Pol II-dependent manner. The produced lncRNAs are very heterogeneous in size and cover lengths from 2 to over 5kb. nSBs are very dynamic structures indicated by the fast and reversible relocalization of HSF1. When activated, HSF1 accumulates at the level of *satIII* repeats during the HS induction, persist in the first three hours of recovery and slowly dissociates later while

the *SATIII* lncRNAs remain close to their transcription site for several hours during the recovery period (Biamonti, 2004). Furthermore, it was demonstrated that transcription and splicing factors are massively recruited to nSBs. Ever since nSBs have been described not as HSF1 granules, but rather as structures formed by HSF1-transcribed *SATIII* lncRNAs, various transcription, splicing and processing factors (Fig.30) (Biamonti and Vourc'h, 2010).



**Figure 30: Formation of nuclear Stress Bodies in human cells.** On a chromosome scheme is indicated the location of pericentromeric heterochromatin (red). RNA Fluorescent *in situ* hybridization (FISH) using a specific probe against *SATIII* RNA (red), coupled with immunodetection of HSF1 (green) in HeLa cells subjected to HS (1h at 43°C) and 3h of recovery period (at 37°). The composition of the nSBs, including HSF1 and RNA Pol II-dependent transcription of *SATIII* and the recruitment of transcription and processing factors, is indicated schematically on the right. *Unpublished microscopy image from our laboratory.*

Importantly in heat-shocked cells, nSBs present a pool of factors involved in *SATIII* transcription- including RNA Pol II, several histone acetyl transferases (Gcn5, CBP/p300, Tip60) and chromatin readers such as Bromodomain Extra-Terminal domain (BET)-containing proteins (Col et al., 2017; Jolly et al., 2004). Moreover numerous RNA processing factors were described to be recruited to nSBs. Among them are found particularly splicing factors- components of ribonucleoprotein complexes (hnRNP M, hnRNP HAP), small nuclear RiboNucleoProteins (snRNP), SAM68 (regulator of alternative splicing) and members of the SR family (SRp30c/ SRSF9 and SF2/ASF/SRSF1) (Biamonti and Vourc'h, 2010; Denegri et al., 2001; Metz, 2004). Some of these factors are recruited through protein-RNA interaction (SF2/ASF) while others such as hnRNP HAP remain at the level of nSBs via a protein-protein interaction. Recently, proteomic studies indicated that more than 150 different proteins can be found at the level of nSBs (Aly et al., 2019; Ninomiya et al., 2020). It was suggested also that following the composition of nSBs they can be

subdivided into two distinct subpopulations characterized by an enrichment of either Scaffold Attachment Factor B (SAFB) or HNRNPM (Aly et al., 2019).

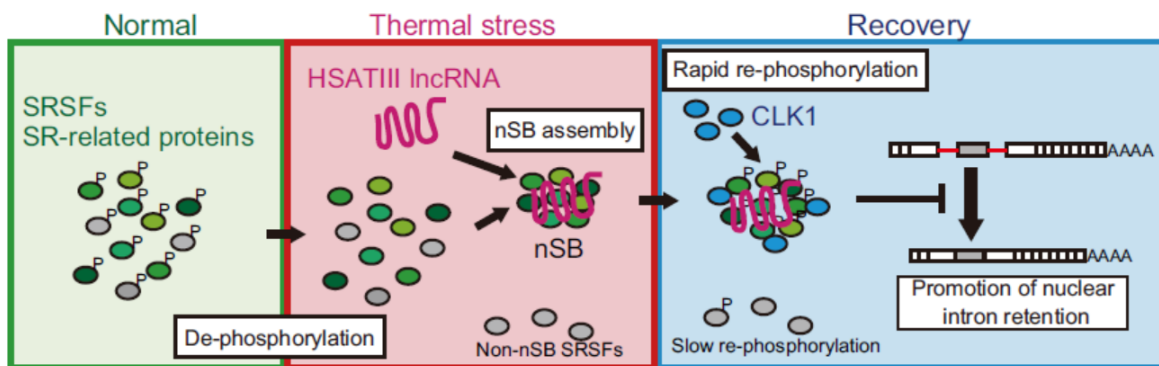
Based on these findings nSBs have been proposed to play the role of molecular traps for various proteins in response to HS. As mentioned previously, the non-coding HSR potentially participates in the global stress-induced downregulation of transcription and RNA-processing through the titration of RNA Pol II, transcription and splicing factors. The acetylation of the pericentromeric regions and transient sequestration of chromatin remodelers at nSBs could also correlate with the global stress-related deacetylation of the transcribed genome.

Another hypothesis on the role of nSBs is their implication in heterochromatin maintenance and reformation. This possibility is tempting since the pericentromeric heterochromatin marks are lost in response to stress and reappear late in the recovery period. *SATIII* lncRNAs are mainly transcribed in one orientation, however anti-sense *SATIII* have been also detected in human cells and may form double stranded RNA (dsRNA) (Jolly et al., 2004; Rizzi et al., 2004; Valgardsdottir et al., 2008). In this context an RNA interference (RNAi) mechanism involving transcriptional gene silencing could be involved in the reformation of heterochromatin following heat shock. Since *SATIII* lncRNAs persist near their transcription site during the recovery after stress, another hypothetical scenario is that they serve as platform for the recruitment of proteins needed for the reformation of heterochromatin.

The HSF1 foci have been a hallmark of the HSR for more than thirty years and several roles have been assigned to these particular nuclear structures. Two very recent studies shed light on the functional implication of nSBs and *SATIII* lncRNAs in post-transcriptional processes.

A transcriptomic analysis in human cells revealed that the depletion of *SATIII*, leading to the elimination of nSBs, promotes the splicing of 533 retained introns particularly during the recovery period after HS (Ninomiya et al., 2020). Ninomiya and colleagues deciphered the underlying mechanisms through the identification of *SATIII* RNA interactome by ChIRP-MS (Chromatin isolation by RNA purification, followed by Mass Spectrometry). This approach demonstrated that *SATIII* lncRNAs interact predominantly with factors involved in pre-mRNA splicing, 3'-end processing and export of mRNAs. Among them most enriched were found to be many of the SR/SRSF splicing factors. Importantly, the phosphorylation state of SR/SRSFs, which

regulates their activity, is dynamically changed in response to environmental stimuli (Ninomiya et al., 2011). It was shown that one of the main SR/SRSF kinases- CLK1 (CDC2-Like Kinase 1) which is autoregulated through splicing events- is also present in the interactome of *SATIII*. Following thermal stress SR/SRSFs are rapidly dephosphorylated leading to modulation of the splicing of wide variety of mRNAs, including *CLK1* mRNA and resulting in the expression of the active form of the kinase. Subsequently active CLK1 rephosphorylates the SR proteins during the recovery period. Once rephosphorylated, SR/SRSF proteins are found to induce mainly intron retention at various mRNAs. The intron retention at *CLK1* mRNA leads then to its inactivation. The study by Ninomiya et al. found that the intron retention at *CLK1* mRNA is inactivated by the KD of the *SATIII* lncRNAs.



**Figure 31 : Model proposed for the role of nSBs in the regulation of splicing events.** Upon HS SRSFs and SR-related proteins are dephosphorylated and are recruited to nSBs. CLK1 is activated through splicing by the SRSF/SR proteins. During the recovery period, CLK1 is also recruited to nSBs and rapidly rephosphorylates and SRSFs which suppress splicing and promote intron retention. *From: Ninomiya et al., EMBO, 2020.*

This finding, together with the analysis of the splicing pattern of hundreds of genes, led to the conclusion that nSBs serve as a platform for the recruitment of SR/SRSFs and CLK1 during the recovery period after HS. Therefore, when *SATIII* sequences are transcribed correctly and nSBs are forming, CLK1 accelerates the rate of rephosphorylation of SR/SRSFs at nSBs and promotes intron retention at diverse mRNAs (Fig.31). Conversely, when *SATIII* lncRNAs are knocked-down and nSBs are not forming, CLK1 is active but SR/SRSFs are not rephosphorylated, thus the massive accumulation of intron retaining transcripts during the recovery period is lost. Hence, this

study identified *SATIII* as an essential element for the intron retention at nearly five hundreds of mRNAs during the recovery period following thermal stress.

Another very recent study unraveled the implication of *SATIII* in chromosomal instability and mitosis (Giordano et al., 2020). HS leads to perturbations of mitosis and to the transient accumulation of cells in pro-metaphase early in the recovery period. It was demonstrated that 48h after the stress induction a significant fraction of mitotic cells present lagging chromosomes and chromosomal bridges bound by nSBs and containing arrays of *satIII* DNA. Later on, after 2 days of recovery, these chromosome segregation defects are resolved potentially in link with the simultaneous processing of *SATIII* lncRNAs into smaller chromatin-associated molecules independently from the DICER and Argonaute 2 machineries. Thus, the study of Giordano and colleagues confirmed that HS and the massive transcription of *SATIII* alter the mitotic behaviors in human cells. Moreover, the role of ncRNAs transcribed from centromeric regions in genome instability has been proposed in *Drosophila* and this role is now also enlarged to Human. These latest results open new perspectives for studies in the field and further demonstrate that *SATIII* may be involved at different levels in the control of gene expression.

Today, the role of nSBs in the regulation of stress-related splicing events and in the control of genomic stability during mitosis have been unraveled. However, perspective studies have to further investigate and decipher the exact mechanisms involved.



In this second chapter of the introduction, the main characteristics of the well conserved stress pathway, activated in response to heat shock- the HSR, have been presented. Over the years, the extensive research on the HSR unraveled the fascinating ability of organisms to adapt to variable external conditions. The robust genome reprogramming, observed in virtually all living organisms following thermal stress, provides numerous opportunities for studies on the cellular mechanisms and on the underlying modulation of gene expression. Still, many aspects of the regulation of the HSR remain elusive. The recent advances in the field of epitranscriptomics demonstrated that RNA modifications, which are today known to control every step of gene expression, could be novel key regulators of diverse stress response pathways. Yet, this emerging field hides a lot of interesting, open questions, needed to be addressed in the context of normal and stress environment. Studies on the role of RNA modifications and their molecular actors in stress biology, as the one which will be presented in this work, are of great significance and can on one hand unravel and decipher the implication of the epitranscriptome in conserved protective cellular programs, and on the other hand can shed light on crucial aspects of stress response mechanisms which have been previously poorly understood.



# **Objectives of the PhD project**



The beginning of my PhD work coincided with the start of a novel research axis in my host team, centered on the nuclear m6A reader, the protein YTHDC1. This axis is dedicated to the nuclear mechanisms of gene expression regulation in response to cellular stress. It is in the continuity of the team's previous work on the roles of the unique fission yeast's potential functional homolog of YTHDC1. The main aims of this research axis are to **unravel new molecular and cellular functions of YTHDC1 and to decipher the involved mechanisms of action, in the context of the cellular response to stress.**

At the beginning of this research work, very little was known regarding the implication of YTHDC1 and the epitranscriptome in the control of gene expression. As highlighted in the introduction section, in the scope of three years, main molecular functions of YTHDC1 have been described and the protein turned out to be an important modulator of fundamental nuclear processes and vital cellular functions. In the search of potential and yet unknown roles of YTHDC1, during my master thesis, we were able to discover that upon heat shock, the protein presents a novel and very specific localization pattern. Indeed, we had observed that YTHDC1 massively relocalizes to nuclear Stress Bodies (nSBs) in response to heat stress. Prior to this observation, an implication of YTHDC1 in cellular stress pathways had not been suggested. The regulation of the HSR and the functions of nSBs in the stress-induced gene expression modulation, which relies on the production of *SATIII* lncRNAs, have been a main interest of research in our team. Therefore, this initial finding immediately caught our interest on multiple intriguing questions regarding the role of YTHDC1 in stress response and the involvement of the epitranscriptome in the conserved Heat Shock Response.

In this context, **the first objective of my PhD project was to characterize the relocalization of YTHDC1 to nuclear Stress Bodies and to unravel the potential role of YTHDC1 in this non-coding response to heat stress.** The obtained results from this axis of my project will be presented in the first chapter of the Results.

Along with the work on the relocalization and function of YTHDC1 at nSBs, we were prompt to address the question of a more global and central role of the protein in the HSR. At this point, few studies had suggested that the RNA modification m6A, via the control of different nuclear processes, could regulate stress responses (Knuckles et al., 2017; Zhou et al., 2015). The

m6A RNA modification had been proposed to control the expression of heat shock proteins. Yet, the molecular actors and the underlying mechanisms remained poorly defined. **The second objective of my research project was to determine the role of YTHDC1 at the level of the protein-coding response to heat stress and, in particular, its implication in the gene expression regulation of heat shock proteins.** The work on this objective will be presented in the second part of the Results.

Even though various functions have been rapidly assigned to YTHDC1, fundamental aspects, such as the extent to which the mode of action of the protein actually relies on the recognition of the m6A modification, remain unclear. For these reasons, **a key aspect of the first two objectives of my PhD project was to evaluate the importance of m6A recognition for YTHDC1 functions in the context of the HSR.** More globally, this question is crucial in order to define the importance of the epitranscriptome in stress biology. It has been studied along with the roles of YTHDC1 at nSBs and for the expression of heat shock proteins, and will be discussed in the first two parts of the Results.

**A third and final objective of my PhD work was to explore the possible role of YTHDC1 for cellular viability and cell cycle progression. This objective was based on initial findings made during my master 2 project.** The results of the preliminary study that I have conducted will be presented in the final, third part of the Results.

# **Materials and methods**





## **I. Cell culture and cell transfections**

### **I.I Cell lines, culture conditions and heat stress treatment**

HEK 293 (Human Embryonic Kidney cells) and HeLa (human cervix carcinoma cells) were purchased via ATCC (American Type Culture Collection, VA, USA). HeLa is an immortal cell line with rapidly dividing cancer cells with variable chromosome number, estimated to be hypertriploid (3n+). HEK 293 are normal cells, immortalized through transfection with Adenovirus DNA and present two or more copies of each chromosome. HeLa stable cell lines expressing WT or mutated YTHDC1 were established in the team via the InGéProt (Ingénierie du Génome et Protéines) platform (see I.III.1). Stable HSF1 Knock Down (KD HSF1) HeLa cells were gently given by Lea Sistonen.

HEK 293, HeLa and HeLa stable inducible cell lines were grown in Dulbecco's Modified Eagle Medium (DMEM) supplemented with 10% (v/v) fetal bovine serum (FBS) and 2% (v/v) L-glutamine under stable physicochemical conditions (37°C, 5% CO<sub>2</sub>, 95% humidity). Stable HSF1 knock-down (KD HSF1) HeLa cells were grown in HeLa wild type medium supplemented with Geneticine antibiotic at 0.4% final concentration.

Heat shock (HS) treatment was performed by immersion of the culture plates or flasks in water bath for 1h at 43°C followed or not by a recovery period at 37°C.

### **I.II Transient transfections**

#### ***I.II.1 siRNA transfection***

Transient knock-down (KD) of gene expression were used for phenotype analyses, functional studies and rescue experiments. The KD assays were performed using specific siRNAs targeting the mRNA(s) of interest and delivered into the cells using a liposomal transfection agent- Lipofectamine RNAi MAX (Thermo Fisher Scientific).

Two protocols are proposed by the manufacturer- a forward protocol and a more rapid reverse protocol. Both protocols were used for KD assays depending on the conditions and time-frame of the experiments with identical final results. Nonspecific siRNA were used as transfection control. The transfection and seeding conditions were the same for HEK, HeLa and HeLa stable inducible cell lines. The specific knockdowns were performed using a combination of two (exclusively for Rescue experiments, see I.II.3 and I.III.2) or three different siRNAs for YTHDC1

and two siRNAs for METTL3 (Table 2). For each well to be transfected in a 6-well plate, 40nM/ml of siRNAs diluted in 250µl of OptiMEM were mixed with 5µl of Lipofectamine RNAi MAX diluted in 250µl of OptiMEM. The mix was incubated for 15min at RT before addition to the cells (forward protocol) or deposition in the well (reverse protocol). For each well to be transfected in a 12-well plate, 20nM/ml of siRNA diluted in 100µl of OptiMEM were mixed with 2µl of Lipofectamine RNAi MAX diluted in 100µl of OptiMEM and incubated for 15min before addition to the cells (forward protocol) or deposition in the well (reverse protocol).

For forward KD the day before transfection with siRNA,  $1.2 \times 10^5$  cells per well in a 6-well plate or  $0.5 \times 10^5$  cells per well in a 12-well plate were seeded in 2ml (6-well plate) or 1ml (12-well plate) of supplemented DMEM and incubated at 37°C for 24h. For reverse KD the day of the transfection with siRNA,  $2.4 \times 10^5$  cells per well in a 6-well plate or  $1 \times 10^5$  cells per well in a 12-well plate were diluted in 2ml (6-well plate) or 1ml (12-well plate) of supplemented DMEM and added to the wells after the addition of the siRNA mix.

### ***I.II.2 DNA plasmid transfection***

Transient DNA plasmid transfections were used for the study of the subcellular localization of YTH-domain proteins and for rescue experiments of YTHDC1 functions.

The DNA was delivered into the cells using a liposomal transfection agent- JetPEI (Polyplus). Plasmid encoding for GFP (pEGFP-N3, Addgene), a plasmid encoding for YTH-domain proteins fused to GFP (Green Fluorescent Protein) - pEGFP-N3-YTH-DF1,DF2,DF3,C1 or C2 (see III.I) or a plasmid encoding for YTHDC1 (mutated or WT) tagged HA (Human influenza hemagglutinin)- pCI-neoYTHDC1-HA (see III.II) were transfected into HeLa cells. The forward transfection protocol was used and  $3 \times 10^5$  cells were seeded one day prior to the transfection. For each well to be transfected 2µl of JetPEI diluted in 100µl of 150mM NaCl (provided with the transfection agent) were incubated for 5min and then mixed with 1µg of plasmid DNA diluted in 100µl of 150mM NaCl. After incubation for 15min at RT, the mix JetPEI / DNA was added to the cells. For immunofluorescence studies, transfections were performed in 12-well plates. For these studies the day prior to transfection  $1.25 \times 10^5$  cells were diluted in 1ml of supplemented medium and were seeded on decontaminated glass coverslips deposited at the bottom of each well. For each well to be transfected 1µl of JePEI diluted in 50µl of 150mM NaCl was incubated for 5min and then mixed with 0.5µg of plasmid DNA diluted in 50µl of 150mM NaCl. After incubation for 15min at RT, the mix JetPEI / DNA was added to the cells. For studies of the localization of the

ectopically expressed YTH-domain proteins fused to GFP cells were incubated for 24h and the medium was not replaced after the transfection. For rescue experiments, medium was replaced 24h after the transfection along with the siRNA transfection and cells were incubated for another 72h before analyses.

### ***I.II.3 Rescue experiments***

Rescue experiments performed with transient transfection were used to study the m6A dependence of YTHDC1 relocalization to nuclear stress bodies and the importance of YTHDC1 binding to m6A for the induction of HSPs. For these studies the plasmid DNA was transfected one day prior to the KD of YTHDC1 in order to obtain effective replacement of the endogenous protein through the expression of the ectopic one.

A batch protocol (trypsinization and transfection of the cells on the same day) was used for the expression of pCI-neo-HA-YTHDC1imm (HA-YTHDC1imm or HA-YTHDC1imm<sup>W428A</sup>, see III.II). For each well to be transfected in a 6-well, 1µg of DNA was diluted in 100µl 150mM NaCl and added to 2µl of JetPEI diluted in 100µl 150mM NaCl. The mix DNA / JetPEI was vortexed briefly, incubated for 15min at RT and added to  $6 \times 10^5$  HeLa cells diluted in 2ml DMEM supplemented medium. Cells together with the DNA / JetPEI mix were deposited in 6-well plates and incubated for 24h. The following day, the transfected cells were harvested after trypsinization and then diluted prior to transfection with siRNAs against YTHDC1. Cells were diluted 5 times to obtain approximately  $2.4 \times 10^5$  cells/ml for seeding in 6-well plates or 12 times to obtain approximately  $1 \times 10^5$  cells/ml for seeding in 12-well plates. The KD of the endogenous YTHDC1, using the two siRNAs against which the ectopically expressed proteins are protected (see III.II), was obtained using the reverse transfection protocol for 6-well or 12-well plates as described previously (see I.II.1). Cells were then incubated for 72h prior to HS treatment.

### **I.III Stable transfection**

#### ***I.III.1 Stable cell lines establishment by CRISPR-Cas9***

Stable HeLa cell lines coding for WT, mutated (W428A) or YTH-domain deleted YTHDC1 were established in the team via the InGéProt (Ingénierie du Génome et Protéines) and were in this study to validate the role of YTHDC1 in the induction of HSPs.

### Donor and Cas9 plasmids

The donor plasmid AAVS1\_Puro\_Tet\_3G\_3×FLAG\_Twin\_Strep-V2 (a gift from Doyon Lab) was used to express YTHDC1 and to confer puromycin resistance to the generated cell line. The donor vector contains 550 bp left and right homology arms matched to the AAVS1 locus in each side, followed by the Tetracycline promoter for the inducible expression of the YTHDC1 gene. YTHDC1 CDS WT, point mutated in the YTH domain or YTH-domain deleted were amplified by PCR from pCI-neo-HA-YTHDC1, pCI-neo-HA-YTHDC1immW428A, and pCI-neo-HA-YTHDC1immΔYTH. The oligonucleotide sequences LT110 and LT112 (Table 2) were used to insert the PCR amplified sequences into the AAVS1 plasmid using Gilson strategy. The used Cas9 plasmid eSpCas9(1.1)\_No\_FLAG\_AAVS1\_T2 (a gift from Doyon Lab) co-expresses the Cas9 and the guides for the AAVS1 locus.

### Cell line generation and individual isolation

HeLa cells were transfected using the JetPRIME (Polyplus transfection) reagent. One day prior to transfection  $1 \times 10^6$  HeLa cells were seeded in 10cm culture plates. On day 1 the transfection was performed following the manufacturer's protocol and using 5µg of the Donor and Cas9 plasmid. Two days later Puromycin (Invitrogen) was added at a final concentration of 2µg/ml in the culture media in order to select transfected cells for 3 additional days. At day 5 post transfection, the puromycin resistant cells were trypsinized and subjected to dilution limit in 96-well plates in the presence of 1µg/ml puromycin. The monoclonal isolation process was performed for 3 weeks. Each time a sole clone was detected in a well it was transferred to a 48-well plates for another 3–4 days in the presence of 1µg/ml of puromycin. When the cells have reached 50 to 70% confluence, they were trypsinized and collected for genotyping, and for transfer into 24-well plates to continue the cell culture in the presence of 1µg/ml of puromycin.

### Characterization of recombinant cell lines by genotyping

The characterization of recombinant cell lines by genotyping was performed using a tail-buffer protocol for DNA extraction. 325µl Tail Buffer (50mM Tris pH 8, 100mM EDTA pH 8, 100mM NaCl, 1% SDS), mixed with 20µl of 10mg/ml Proteinase K (Roche) were added to each cell pellet collected from one well of a 48-well plate. Following an incubation for 1h at 37°C, 125µl 5M NaCl were added and the samples were mixed for 5min at 1000rpm at RT. Next, samples were

centrifuged for 10min at 13 000rpm at RT and 325µl of the supernatant were transferred to new tubes and mixed with 250µl isopropanol. Following a centrifugation 10min at 13 000rpm, RT, the supernatant was eliminated and the pellet was washed in 500µl 70% EtOH and centrifuged for 10min at 13 000rpm, RT. The pellet was then air-dried, dissolved in 50µl elution buffer (10mM Tris pH 7.5, 0.5mM EDTA) and briefly vortexed. The insertion of the exogenous YTHDC1 was examined by PCR using the primers LT44 and LT116 or LT119 and LT120 (Table 2) to detect positive clones. Selected clones were maintained in supplemented DMEM medium or stored at -80°C.

#### *Characterization of recombinant cell lines by Western Blot*

To determine the concentration of Doxycycline needed to induce the expression of the exogenous YTHDC1, one day prior to the induction  $3 \times 10^6$  HeLa cells were seeded in 6-well plates. At day 1 Doxycycline at different concentrations (0.1µg/ml to 2µg/ml) was added to the cell medium. At day 2, the cells were trypsinized, washed with PBS and collected by centrifugation 5min at 2500rpm. The proteins were extracted for Western Blot analyses and 10µg of protein extract from each sample were deposited on polyacrylamide gels. An anti-FLAG antibody was used to detect the signal of the exogenously expressed YTHDC1\_3×FLAG\_ Twin\_Strep-V2 protein.

#### ***I.III.2 Rescue experiments in the stable cell lines***

Rescue experiments in the stable cell lines were performed to confirm the specificity of the impact of YTHDC1 on HSPs expression. The expression of the ectopic WT YTHDC1 was induced using Doxycycline one day prior to the KD of the endogenous YTHDC1 to obtain effective replacement of the endogenous protein through the expression of the ectopic one.

$1.2 \times 10^5$  HeLa stable inducible HeLa cells per well in a 6-well plate or  $0.5 \times 10^5$  cells per well in a 12-well plate were seeded and incubated for 24h in DMEM supplemented medium containing 0.8µg/ml of Doxycycline (concentration defined as described in I.III.2). The following day, the medium was replaced with new one, containing the same concentration of Doxycycline, and a reverse transfection using the two siRNAs against which the ectopically expressed protein is protected (see III.II) was carried out in order to knock-down the endogenous YTHDC1. Cells were incubated for 72h and every 24h the medium was replaced with new DMEM supplemented medium containing 0.8µg/ml of Doxycycline in order to maintain the induction of the ectopic

protein. In the control plates (without addition of doxycycline), the medium was also replaced each 24h.

## II. Subcellular fractionation

Fractionation was performed to obtain cytosolic and nuclear fractions from WT and YTHDC1 KD cells and study the effect of YTHDC1 KD on the export of *HSP* mRNAs.

HeLa cells were seeded in 10cm dishes and transfected transiently with control siRNA or siRNA against YTHDC1 (see I.II.1). 72h after transfection, cells were submitted or not to HS treatment followed or not by a recovery period. Cells together with the dead floating cells were manually collected using a scraper, centrifuged for 5min at 2000rpm and washed in cold 1X PBS (phosphate-buffered saline). On ice 150µl of lysis buffer (10mM Tris pH 7.5, 60mM KCl, 15mM NaCl, 0,34M sucrose, 650µM spermidine, 2mM EDTA, 0,5mM EGTa, 1mM DTT, 0,5mM PMSF, 0.05% (w/v) TritonX-100 + 1µL RNasin) were added gently to the cell pellet (obtained from one 10cm dish) and incubated for 10min. After a 5min centrifugation at 2000rpm, the pellet was kept on ice while the supernatant (cytosolic fraction) was collected and stored on ice. 200µl of washing buffer (10mM Hepes pH 7.5, 10mM KCl, 0,3M sucrose, 1,5mM MgCl<sub>2</sub>, 1mM DTT, 0,5mM PMSF, 1µL RNasin) were added carefully to the pellet and the suspension was centrifuged 5min at 2000rpm. The pellet, containing the washed nuclei, was kept on ice.

For total RNA extraction, 1ml of Trizol Reagent was added to the pellet (nuclear fraction), and 1.5mL to the supernatant. After homogenization, samples were frozen in liquid nitrogen and stored at -80°C. RNA was extracted following the protocol for total RNA extraction (see V.I) using 300µl of chloroform and 750µl of isopropanol for the cytosolic extract and 200µl of chloroform and 500µl of isopropanol for the nuclear extract. RNA concentration and purity were determined using NanoDrop ND-1000 UV spectrophotometer. The RNA integrity was evaluated by electrophoresis on 1% (w/v) agarose gel using the GelGreen Nucleic Acid Stain (Biotium) dye following the manufacturer's instructions.

For Western Blot analyses, the pellet was resuspended into 150µL of nucleus lysis buffer (50mM Tris pH 7.5, 140mM NaCl, 1,5mM MgCl<sub>2</sub>, 0,5% (w/v) NP-40) containing nucleases (benzonase, Merck Millipore) and incubated 20min on ice. The protein concentrations (from the cytosolic and the nuclear fractions) were quantified using Bradford assay (BioRad).

### **III. DNA constructs and mutagenesis**

#### **III.I DNA constructs for expression of YTH-domain GFP-fused proteins**

YTH-domain proteins fused to GFP were used to study their cellular localization in response to HS.

The pEGFP-N3 vector of 4.7kbp, containing the GFP-coding sequence was used for plasmid mammalian expression. The coding sequences for human YTHDF1, YTHDF2, YTHDF3, YTHDC1 and YTHDC2 (obtained from NCBI protein data bank) were amplified from HEK cells cDNA library by PCR (see V.III) using specific primers (Table 2). The PCR products were digested by the restriction enzymes *XhoI-XmaI* for YTHDC1 cDNA and *XhoI-BamHI* for YTHDF1-2-3 and YTHDC2 cDNAs. After purification using the QIAquick PCR Purification kit (Qiagen), the obtained fragments were inserted into the *XhoI-BamHI* or *XhoI-XmaI* digested pEGFP-N3 vector. Positive clones were selected by PCR and validated by sequencing. The selected clones were amplified through bacterial transformation, purified using the NucleoBond Xtra Midi Prep Plus kit (Macherey-Nagel) and stored at -20°C.

#### **III.II Mutagenesis of YTHDC1-coding siRNA non-degradable constructs**

YTHDC1-coding siRNA non-degradable constructs were used to perform rescue experiments (see I.II.3 and I.III.2). Silent mutations were introduced in order to “immunize” the mRNA of the ectopically expressed protein against the siRNAs used for the KD of the endogenous YTHDC1.

First, full-length human YTHDC1 was amplified from HEK cDNA library cells by PCR and the obtained product was cloned into mammalian modified pCI-neo-HA plasmid resulting in HA (Human influenza hemagglutinin)-tagged YTHDC1 (pCI-neo-HA-YTHDC1, previously obtained in the team). Non-degradable plasmids were obtained by site-directed mutagenesis using the Q5 Site-Directed Mutagenesis Kit (BioLabs) and mutagenic primers containing a mismatch in the center of the sequence in order to introduce silent mutations at the site of recognition by two of the specific YTHDC1 siRNAs (primers listed in Table 2). Silent mutations were introduced in two steps- first at the site of recognition by one of the siRNAs, next the obtained construct was again mutated at the site of recognition by the second siRNA.

For each step first a PCR-based amplification using the Q5 Hot Start Fidelity Master mix, the template DNA (20ng) and the mutagenic primers was conducted following the Q5 Site-Directed Mutagenesis Kit instructions. Next the PCR products were added to a Kinase-Ligase-DpnI (KLD) enzyme mix to allow circularization and template removal. NEB 5-alpha Competent *E. coli* were transformed with the resulting constructs. The positive clones were selected by PCR using specific primers unable to amplify the non-mutated cDNA and the selected construct was sequenced to exclude the introduction of non-specific mutations.

The obtained non-degradable plasmid (pCI-neo-HA-YTHDC1imm) was used to establish YTHDC1 point mutants or YTH-domain deletion mutant (delAA354-494) by PCR-based mutagenesis using the Q5 Site-Directed Mutagenesis Kit and specific primers (Table 2). The point mutations were introduced at tryptophan 428 situated in the hydrophobic cage and allowing the recognition of m6A (Xu et al., 2014). The tryptophan residues (W) were substituted by alanine (A). The resulting DNA constructs were named: pCI- pCI-neo-HA-YTHDC1immW428A (referred to in this work as HA-YTHDC1<sup>W428A</sup>) and pCI-neo-HA-YTHDC1immΔYTH. All DNA constructs have been verified by sequencing, produced and purified from bacteria using the NucleoBond Xtra Midi Prep Plus kit (Macherey-Nagel) and stored at -20°C.

## IV. Protein analyses

### IV.I Western blot

#### Protein extraction and sample preparation

For western blot analyses cells were trypsinized and collected together with the dead floating cells. After 5min centrifugation at 2500rpm, cells were washed in 1X PBS and centrifuged again 5min at 2500rpm. Pellets were frozen and stored at -20°C or directly lysed. Depending on the initial quantity of cells 20μl (for one well from a 6-well plate) or 50μl (for two wells from a 6-well plate) of lysis buffer (50mM Tris pH 7.5, 150mM NaCl, 5mM MgCl<sub>2</sub>, 1% (w/v) Triton X-100, 0.5% sodium deoxycholate buffer) were added. Prior to lysis protease inhibitors (LABP, 1/1000, homemade) and benzonase (Merck Milipore, 1/100) were added. Cells were then incubated for 20min on ice and vortexed briefly every 5min. The protein concentration was determined using Bradford assay (BioRad). The proteins were denatured by addition of 5X Laemmli loading buffer



(60mM Tris pH 6.8, 2% (v/v) SDS, 10% (w/v) glycerol, 10% (v/v)  $\beta$ -mercaptoethanol, 0.01% (w/v) bromophenol blue) to final 1X concentration and incubation 5min at 95°C.

#### *Gel electrophoresis, transfer, blocking and detection*

The volumes from each sample corresponding to 10 $\mu$ g of proteins were loaded on 8%, 10% or 12% SDS-polyacrylamide gels. Electrophoresis was performed in 1X Laemmli running buffer (0.1% (v/v) SDS, 25mM Tris, 192mM glycine) at constant voltage. After migration, the gel separated proteins were blotted on nitrocellulose membrane in transfer buffer (25mM Tris, 192mM glycine, 0.1 % (v/v) SDS, 20% (v/v) EtOH) for 1h at 100V. Blocking was performed in 5% non-fat dry milk, TBST solution (Tris Buffered Saline with Tween 20: 50mM Tris pH 7.5, 150mM NaCl, 0.02% (w/v) Tween20) for 45min at RT. The membranes were then incubated with the primary antibody (Table 1) diluted in 1% (w/v) non-fat dry milk, TBST for 1h15 at RT, followed by three washings of 10min in TBST at RT. The secondary antibody coupled to HRP (horseradish peroxidase) (Table 1) was diluted in TBST and membranes were incubated for 45min at RT followed by three washes in TBST buffer at RT. The ECL system (GE Healthcare) was used for immunoblotting detection (the Clarity Western ECL, Biorad was used occasionally for detection of weaker signals).

#### *Images acquisition and signal quantification*

The Western Blot images were acquired using the Fusin FX (Vilber) camera and the automatic exposure option excluding saturated signals. Signals were quantified with the corresponding Vilber software and adjusted to the tubulin quantified levels for comparison between different samples and conditions.

### **IV.II Immunoprecipitation (IP)**

Immunoprecipitation of YTHDC1 was carried out in order to identify the protein molecular partners. These experiments were performed in HEK cell which are often successfully used for the purification of protein complexes and present advantages in terms of protein yield.

### Proteins extraction

Cells from 4 T75 flasks, at 90% of confluence were trypsinized, washed in 1X PBS and lysed 10min on ice in 2ml of lysis buffer (50mM Tris pH 8, 150mM NaCl, 5mM MgCl<sub>2</sub>, 0.5% (w/v) Na-deoxycholate, 1% (w/v) TritonX-100, 10% (w/v) glycerol, 1mM PMSF, 1X LABP and 1mM DTT). The soluble total extract was recovered after centrifugation at 14 000rpm for 10min at 4°C and nucleases (benzonase, Merck Millipore) were added to this extract to degrade nucleic acids. This soluble extract was further used for immunoprecipitation with a specific antibody raised against YTHDC1 and with an irrelevant antibody.

### Preparation of beads-antibody complexes and immunopurification

Magnetic beads coupled to protein A (Dynabeads Protein A, Thermo Fisher) were used for immunoprecipitation of YTHDC1 and the associated proteins. 7µl of beads per IP were washed in 1X PBS, 0.02% (w/v) Tween 20 and incubated with 10µL of the purified polyclonal rabbit antibody raised against the N-terminal part of YTHDC1 (homemade antibody) or with 5µL of an irrelevant polyclonal goat antibody anti-rabbit IgG (control experiment) for 20min at RT under agitation. The Dynabeads/antibody complexes were next equilibrated in the lysis buffer and then incubated with 1 mL of the soluble proteins extract per IP. After incubation for 6h at 4°C on a rotating shaker, the Dynabeads/antibody/antigen complexes were washed four time in 25mM Tris pH 8, 1mM MgCl<sub>2</sub>, 150mM NaCl, 0.01% (w/v) NP40, 1X LABP washing buffer and the antibody-antigen complexes were eluted by addition of 40µl 2X Laemmli loading buffer (60mM Tris-HFCl, pH 6.8, 2% (v/v) SDS, 10% (w/v) glycerol, 10% (v/v) β-mercaptoethanol, 0.01% (w/v) bromophenol blue) and incubated 10min at 70°C.

### Analyses of the purified proteins

The purified proteins were first analyzed by western blot (5µl of each eluted sample were loaded onto the gel) using the antibody raised against YTHDC1. Next, samples were analyzed by SDS-PAGE and silver staining (10µl of each eluted sample were loaded). For silver staining, after migration, the gel was fixed for 30min at RT in 30% (v/v) EtOH, 10% (v/v) acetic acid solution, then activated for 30min at RT in 30% (v/v) EtOH, 0.5% (v/v) acetic acid, 2.5mM sodium thiosulfate, washed three times 15min in H<sub>2</sub>O, stained in the dark for 30min at RT under agitation in 0.1% (w/v) AgNO<sub>3</sub>, 0.03% (v/v) Formaldehyde. After a rapid wash in H<sub>2</sub>O, the gel was

developed in 2.5% (v/v) Na<sub>2</sub>CO<sub>3</sub>, 0.03% (v/v) Formaldehyde solution. The staining was stopped by addition of acetic acid.

YTHDC1 partners were identified using tandem mass spectrometry analyses in collaboration with the EDyP laboratory, CEA Grenoble.

## **V. RNA analyses**

RNA extracted from cells was subjected to different techniques in order to study its abundance and/or splicing and processing characteristics (RT-qPCR, PCR).

### **V.I Total RNA extraction**

Total RNA extraction was performed using the TRIzol Reagent (Invitrogen) protocol. Cells from 6-well plates or 10cm dishes were collected by trypsinization and centrifuged 5min at 2500rpm. For  $1 \times 10^5$ - $10^7$  cells 1ml of Trizol reagent was added and the mix was thoroughly vortexed (at this step occasionally samples were frozen and stored at -80°C). After 5min incubation at RT, 0.2ml of chloroform were added, the mix was vortexed and then incubated for another 2-3min at RT. A centrifugation at 12 000rpm, 15min at 4°C separates the sample into two phases- upper colorless aqueous phase (containing RNA) and lower phenol-chloroform red phase (containing DNA and proteins). The aqueous phase was carefully transferred to a new tube, 0.5ml of isopropanol were added, the mix was vortexed and centrifuged for 10min at 12 000rpm, 4°C. Following this, the pellet was washed in 75% (v/v) EtOH and centrifuged at 7500rpm for 5min, 4°C. At the end the RNA pellet was air-dried, dissolved in 20-50µl RNase-free water, incubated at 60°C for 10min and stored at -20°C. RNA concentration and A260/280 ratio were determined using NanoDrop ND-1000 UV spectrophotometer. The RNA integrity was evaluated by electrophoresis on 1 % agarose gel using the GelGreen Nucleic Acid Stain (Biotium) dye following the manufacturer's instructions.

### **V.II Reverse Transcription (RT) for cDNA synthesis**

RT was carried out in order to synthesize cDNA from RNA. Random hexamer (a range of random sequences) were used to produce cDNA from the whole RNA pool. Specific primers (designed to target a sequence) were used to synthesize cDNA from an RNA of interest.

RT was performed using Transcriptor Reverse Transcriptase (Roche). First RNA is subjected to DNase treatment to eliminate traces of genomic DNA. 1µg of RNA was incubated for 20 min at 37°C in the presence of 5X DNase buffer (100mM Tris pH 8, 10mM MgCl<sub>2</sub>) to final concentration 1X, 100mM DTT and RNase inhibitor (2U/µl). The DNase was inactivated by the addition of 0.7µl 50mM EDTA and incubation for 10min at 70°C.

For synthesis using random hexamer the DNase-treated RNA was hybridized using 2µl of random hexamer (50ng/µl) in the presence of 1µl 50mM MgCl<sub>2</sub> for 10 min at 65°C. Samples were cooled on ice (2min). Next 7µl of the cDNA reaction: 10mM dNTP, RNase Inhibitor, 5X RT buffer and Reverse Transcriptase (Invitrogen), (final concentrations: 1mM dNTP; 2.2U/µl RNase Inhibitor; 1X RT buffer; 10U/µl Reverse Transcriptase) were added and cDNA was synthesized by subsequent incubations for 10min at 25°C and 40min at 55°C. The reverse transcriptase was inactivated by heating 5min at 85°C. A control reaction in the absence of the Reverse Transcriptase was also performed to verify that no DNA was present in the RNA sample and the PCR amplification was specific.

cDNA synthesis was carried out using specific primers for *HSP40* mRNA for the study of the effect of YTHDC1 KD on the transcription termination of *HSP40*. DNase treatment was performed as previously described with 3µg of RNA. Following this RNA was hybridized using a mix of the specific primers at final concentration of 2µM in the presence of 1µl 30mM MgCl<sub>2</sub> and 2µl 10mM dNTPs for 10 min at 65°C. cDNA synthesis was performed using the SuperScript IV (SSIV) Reverse Transcriptase (Invitrogen) (10U/µl final concentration), the corresponding SSIV buffer to 1X final concentration, RNase Inhibitor (2.2U/µl final concentration), 100mM DTT (1mM final concentration) and subsequent incubations for 10min at 23°C and 10min at 55°C. The Reverse transcriptase was inactivated by heating 10min at 80°C. A control reaction in the absence of the SSIV Reverse Transcriptase was also performed to verify that no DNA was present in the RNA sample and the PCR amplification was specific.

### **V.III Polymerase chain reaction (PCR)**

PCR reactions were performed in a total volume of 20µl using the cDNAs produced by RT. 1µl of cDNA was mixed with 1µl of each specific primer at 10µM, 8µl H<sub>2</sub>O and 10µl the 2X BioMix Red (Bioline) or 2X GoTaq G2 Hot Start Master mix (Promega) which contains the polymerase, buffer, dNTPs and MgCl<sub>2</sub>. The applied PCR program consists of a denaturation step

at 95°C, followed by 35 cycles of denaturation 10sec at 95°C, annealing 10sec at the corresponding primers T<sub>m</sub> (melting temperature) and elongation 30sec to 2min (depending on the size of the region of interest) at 72°C and a final step 5min at 72°C. PCR products were migrated on 1%, 1.5% or 2% agarose gels and stained using the GelGreen Nucleic Acid Stain (Biotium) dye following the manufacturer's instructions.

#### **V.IV Quantitative PCR (qPCR)**

Quantitative PCR was performed using the cDNAs obtained after RT (V.II). qPCR reactions were performed in a total volume of 20µl in 96-well plates using the LightCycler 480 PCR platform (Roche). To 4µl of cDNA diluted 5 times were added the specific primers (1.2µl from a mix at 5µM for each primer), 4.8µl of H<sub>2</sub>O and 10µl of the Green Mastermix Plus SYBR (Eurogentec) containing dNTPs, the enzyme and the fluorophore SYBR Green which is incorporated in a non-specific and homogenous way to the synthesized DNA. The qPCR program consists of a denaturation step at 95°C, followed by 40 cycles composed of steps of 15sec each: denaturation at 95°C, hybridization at 60°C and polymerization at 72°C. The final step of the program allows to create the dissociation curves for the primers: 1min 95°C, 30sec 60°C, 30sec 95°C.

The qPCR data was analyzed using the LightCycler 480 software which allows to subtract the value of the linear amplification threshold corresponding to the minimal number of PCR cycles needed to amplify the region of interest in a linear manner. All quantifications were performed on the median of two technical replicates. The RNA levels for each sample were normalized to the RNA levels of a house-keeping gene (tubulin β, actin β or GAPDH). Depending on the experiment the RNA levels were represented as a fold enrichment compared to WT cells or non-heat shocked WT or YTHDC1 KD cells.

### **VI. Microscopy techniques**

#### **VI.I Immunofluorescence**

Immunofluorescence using specific antibodies raised against the proteins of interest was performed to study their subcellular localization in different conditions.

Cells grown on coverslips were fixed in 4% (v/v) formaldehyde, 1X PBS solution for 10min at RT under light agitation, followed by three washes of 5min in 1X PBS. Permeabilization was performed using 0.5% (w/v) Triton X-100, 0.5% (w/v) Saponine, 1X PBS solution for 15min at RT under agitation, followed by three washes of 5min in 1X PBS . Unspecific sites were blocked 30 min at 37°C in 0.3% (w/v) Triton X-100, 10% (v/v) FBS (Fetal Bovine Serum), 1X PBS. Following this, cells were incubated 1 h at 37°C with a primary antibody (Table 1) diluted in 0.3% (w/v) Triton X-100, 2% (v/v) FBS, 1X PBS. Cells were then washed three times 5min in 0.3% (w/v) Triton X-100, 2% (v/v) FBS, 1X PBS solution warmed to 43°C and incubated 30min at 37°C with a secondary antibody (Table 1) diluted in 0.3% (w/v) Triton X-100, 2% (v/v) FBS, 1X PBS. At the end of the incubation cells were washed three times 5min in 0.3% (w/v) Triton X-100, 2% (v/v) FBS, 1X PBS solution warmed to 43°C. DNA was counterstained with 250 ng/mL DAPI diluted in DAKO mounting solution (Agilent).

## **VI.II Fluorescence in situ hybridization (FISH)**

FISH was performed for the detection of specific RNA or DNA sequences in order to study their presence or abundance, localization or changes in different conditions. For this technique specific probes (fluorescent or labelled through biotinylation) with high complementarity to the target sequence allowed the specific detection and visualization.

### ***VI.II.1 RNA FISH***

#### ***Biotinylation of the probe pH2.3 used for the detection of HSP70 transcripts***

A plasmid containing the genomic probe pH2.3 has been received from Dr R.I. Morimoto's lab. The plasmid contains a sequence of 2.3kb covering the entire coding sequence of *HSP70* mRNA.

The staining of the pH2.3 probe was performed using the kit Bioprime (Life Technologies). The principle of the procedure is to cut the DNA by the Klenow enzyme and randomly incorporate a nucleotide attached to a biotin molecule. 250ng of the plasmid DNA containing the probe were diluted in 20µl of distilled water. On ice 20µl of 2.5X Random Primers were added. The probe was then denatured in boiling water for 5min and placed immediately on ice. 5µl of 10X dNTP mixture, 4µl of distilled water and 1µl Klenow enzyme were added, the mix was briefly centrifuged and

incubated for 1h at 37°C. The reaction was stopped by the addition of 5µl of Stop buffer. The staining increases the size of the probe and was verified by migration of the product on 1% (w/v) agarose gel using the GelGreen Nucleic Acid Stain (Biotium) dye.

#### Precipitation of the probe

The Bioprime kit produces a considerable net DNA synthesis resulting in a 10-40 fold amplification of the probe. For the pH2.3 this amplification was estimated to be 10 times by comparing the signal obtained on agarose gel to the signal of a sample with previously defined concentration. Approximately 200ng of biotinylated probe were used per coverslip. 5µl of biotinylated probe were precipitated in the presence of 1µl of salmon sperm DNA (10µg/µl), 5µl 3M sodium acetate, pH 5.2, 39µl of distilled water and 100µl of 100% EtOH for 30min at -80°C. Following a centrifugation at 12 000rpm for 20min at 4°C, the pellet was washed in 70% (v/v) EtOH and centrifuged for another 10min at 12 000rpm, at 4°C. The pellet was then air-dried (10-15min at RT) and dissolved by pipetting in a mix containing 5µl formamide (UltraPure, Thermo Fisher Scientific) and 5µl hybridization mix (20% dextran sulfate, 4X saline-sodium citrate (SSC)). At the end the probe was denatured at 75°C for 5min and placed immediately on ice.

#### Cells preparation prior to RNA FISH

The cells were fixed and permeabilized as described in the protocol used for immunofluorescence followed by incubation in 20% (w/v) glycerol/ 1X PBS for 20min under agitation at RT. Cells were then washed 5min in 1X PBS and dehydrated by successive incubation for 3min in 70% (v/v) EtOH, 90% (v/v) EtOH, 100% EtOH. Coverslips were air-dried and 5-10µl (depending on the coverslip diameter) of previously biotinylated and precipitated probe were deposited. The cells were incubated overnight at 37°C and rubber cement was used to prevent evaporation. The following day the cells were washed three times for 5min in 60% (v/v) Formamide, 2X SSC solution warmed to 45°C and three times for 5min in 2X SSC at RT under light agitation. 120µl of blocking solution (3% (w/v) BSA, 0.1% Tween 20, 4X SSC) were deposited and the cells were incubated 30min at 37°C. Following this the blocking solution was replaced by the detection solution (3% (w/v) BSA, 0.1% (w/v) Tween 20, 4X SSC) containing the TRITC anti-biotin antibody diluted at 1/200. The cells were incubated for another 30min at 37°C

and washed three times for 5min in 4X SSC, 0.1% Tween 20 solution warmed to 45°C. DNA was counterstained with 250 ng/mL DAPI diluted in DAKO mounting solution.

#### *RNA FISH combined with immunofluorescence*

For studies combining RNA detection using FISH and protein immunostaining best results have been obtained when the IF is performed following the RNA FISH. After the final washing steps of the FISH protocol, , the cells were fixed again in PFA 4% , washed in PBS 1X and then incubated with the blocking solution for IF and following this the other steps of the IF protocol were carried out as usual (see VI.1).

#### ***VI.II.2 DNA FISH***

DNA FISH was performed to study the colocalization of YTHDC1 with *satIII* DNA repetitive sequences.

The cells were fixed and permeabilized as described in the protocol used for immunofluorescence. Cells were then dehydrated by successive incubation for 3min in 70% (v/v) EtOH, 90% (v/v) EtOH, 100% EtOH. Coverslips were air-dried and DNA was denatured 10min at 80°C in 50% Formamide, 1X PBS. Then 5-10µl of the precipitated probe were deposited on each coverslips. The cells were incubated overnight at 37°C. Next, cells were washed three times for 5min in 50% (v/v) Formamide, 2X SSC solution warmed to 45°C and three times for 5min in 15% (v/v) Formamide. For studies combining DNA detection using FISH and protein immunostaining best results have been obtained when the IF is performed following the RNA FISH. After the final washing steps of the FISH protocol, , the cells were fixed again in PFA 4% , washed in 1X PBS and then incubated with the blocking solution for IF and following this the other steps of the IF protocol were carried out as usual (see VI.1).

#### **VI.III Images acquisition and processing**

To have comparative analyses between cell samples submitted to different conditions, images were acquired with the same exposure time using a confocal imaging system (ApoTome-AxioImager Z1 Zeiss) equipped with a monochromatic CCD camera (AxioCam MRm, Zeiss) and navigated by the AxioVision software. Three z-planes were acquired using the 63x oil immersion objective at constant 0.24µm depth. For the acquisition of images at lower magnification, the 40x objective of the same microscope was used without the ApoTome system. Images were converted



to “composite” images containing the channels corresponding to the fluorescent signals kept as 16-bit using the ImageJ software. To arrange the images into panels and create the final figures the FigureJ plugin available with the ImageJ software was used without any signal or size conversions from the original images.

## VII. Chromatin immunoprecipitation

ChIP experiments were carried out to study RNA Pol II-S2P enrichment along *HSP70* and *HSP40* locus in WT and YTHDC1 KD HeLa cells at NHS and HS. YTHDC1 ChIP-seq experiments were carried out in order to identify YTHDC1 targets at NSH and HS in HeLa cells.

ChIP experiments were performed mainly as described previously (Motamedi et al., 2004) with the following modifications. HeLa cells (from two 15 cm dishes at 70-80 % confluency, around  $2 \times 10^7$  cells) were cross-linked with 1% (v/v) formaldehyde (Sigma) at room temperature for 15 min. The reaction was quenched with glycine (Sigma) at a final concentration of 125mM for 5min at RT. Cells were collected, centrifuged for 3min at 2000rpm, and the pellets were washed twice in cold 1X PBS. Cell pellets were resuspended in 7mL of cell lysis buffer (50mM Hepes pH 7.5, 10mM KCl, 0.1% (w/v) NP40, and protease inhibitors) and transferred to a chilled dounce homogenizer on ice and homogenized for 36 strokes. The suspension was then centrifuged for 3 min at 2000rpm at 4°C, and resuspended in 1mL of nuclear lysis buffer (50mM HEPES-KOH pH 7.5, 500mM NaCl, 1mM EDTA pH 8.0, 1% Triton X-100, 0.1% sodium deoxycholate, 0.1% SDS, and protease inhibitors). Lysates were sonicated for 10min (30"ON 30"OFF) at 30% amplitude in a water bath sonicator, centrifuged at 13 000 rpm for 10min and the supernatant was stored at -80°C. To obtain input DNA, 25µL of clarified lysate were used.

Each immunoprecipitation reaction was performed from 50µg of sheared chromatin incubated overnight at 4°C on a wheel with 2µg of specific antibody against RNA polymerase II subunit B1 (phospho CTD Ser-2), against YTHDC1 A305-096A or irrelevant polyclonal goat antibody anti-rabbit IgG (Table 1). 40µl of a 50% slurry of prewashed protA-Sepharose beads (nProteinA Sepharose 4 Fast Flow, GE Health Care) were incubated with each lysate at 4°C for 1h. Beads were washed two times in nuclear lysis buffer, once with 10mM Tris-HCl pH 8.0, 0.25M LiCl, 0.5% NP-40, 0.5% sodium deoxycholate, and 1mM EDTA, and once with TE (10mM Tris-HCl pH 8.0, 1mM EDTA) at RT. Immunopurified material was eluted by incubating with 100µL of 50mM Tris-HCl pH 8.0, 10mM EDTA, and 1% SDS at 65°C for 20min. Eluate was transferred

to a fresh tube and pooled with a final bead wash of 150 $\mu$ L of TE with 0.67% SDS. For input DNA, 200 $\mu$ L of TE with 1% SDS was added to 25 $\mu$ L of lysate. All samples were incubated at 65°C overnight. Then 200 $\mu$ L of TE, 100 $\mu$ g of Proteinase K (Roche) was added to all samples, and incubated at 37°C for 2h. After addition of 55 $\mu$ L of 4M LiCl, DNA samples were extracted once with phenol:chloroform:isoamyl alcohol and once with chloroform and then precipitated (1mL 100% EtOH and 2 $\mu$ l glycogen (20mg/ml) were added to the recovered upper phase, incubated at -80°C for 20min, centrifuged at 4°C for 20min and washed once with EtOH 70%. Precipitated and washed DNA was resuspended in 50 $\mu$ L 10 $\mu$ g/ml RNase A (for the input) or 0.5 $\mu$ g/ml RNase A (for IP) and incubate at 37°C for 1h.

Quantitative PCR was conducted with 2 $\mu$ l of a 1/2 dilution of the IP DNA. Relative enrichment was calculated as the ratio of signal of interest to signal of control IP.

Chromatin immunoprecipitation for whole-genome analysis was performed using the "ChIP-IT High Sensitivity" kit from Active Motif following the manufacturer instructions. Sheared chromatin was obtained using a probe sonicator by 10 cycles, 30"ON 30"OFF at 63W. ChIP were performed using 25  $\mu$ g chromatin with addition of spike-in.

## **VIII. Cell viability and cell cycle assays**

Cell viability and cell cycle analyses were carried out using Fluorescence-activated cell sorting. For cell cycle analyses, prior to incubation with Propidium Iodide (PI) cells were permeabilized. Thus, every cell uptakes the fluorescent intercalating agent and the obtained signal reflects the DNA content of cells.

Cell viability was determined with PI exclusion assay- dead unpermeabilized cells uptake PI which binds to DNA by intercalating between the nucleotides. Living unpermeabilized cells are unable to uptake PI. The number of cells with fluorescent signal corresponding to the wave-length of the PI fluorophore represent the number of dead cells.

Cellular apoptosis was studied through the detection of the cleaved form of caspase 3 via specific antibody.

### **VIII.I Cell growth curves**

Cell growth curves were used in the beginning of the study to estimate the impact of YTHDC1 KD on cellular growth.

To establish cell growth curves the first day KD of YTHDC1 was performed in HeLa or HEK cells in 6-well plates as described in I.II.1. From day 2 until day 4 cells were collected, including dead floating cells in the medium, centrifuged 5min at 1200rpm, 4°C and diluted in 2ml of DMEM supplemented medium. 10µl of the solution were deposited on Malassez counting chamber and the cell counting was performed manually using an upright microscope.

### **VIII.II Propidium iodide staining**

For PI exclusion assay 72h after the induction of the KD of YTHDC1 (I.II.1) HeLa (WT or KD) cells from two wells in a 6-well plate were collected together with dead floating cells in the medium, centrifuged 5min at 1200rpm, 4°C and washed twice in 1X PBS. Cells were then incubated with 500µl of 0.1 mg/ml RNase, 1X PBS for 30min at 37°C. Prior to analyses 10µl of 1mg/ml PI stock solution were added and cells were gently vortexed. The PI up-taking by dead cells was measured using BD Biosciences FACS Accuri- C6.

For cell cycle analyses, prior to incubation with PI HeLa WT or YTHDC1 KD cells were permeabilized. After collection and washing of the cells, as previously described, the pellet was resuspended in 0.5ml 1X PBS and added carefully under agitation drop by drop to ice cold 70% EtOH to allow fixation and permeabilization of the cells. After 30min incubation on ice, cells were centrifuged 5min at 1200rpm, 4°C and washed in 1X PBS. The RNase treatment, PI staining and FACS analyses were performed as described for the PI exclusion assay.

### **VIII.III Assessment of apoptosis**

Cellular apoptosis was studied through the detection of the cleaved form of caspase 3. Caspase 3. This study was conducted using the BD Pharmingen PE Active Caspase-3 Apoptosis Kit. YTHDC1 KD in HeLa cells was performed following the forward protocol (I.II.1). 24h prior to analysis 0.1µM of Actinomycin D were added to WT or KD cells in order to induce apoptosis and serve as a positive control for the assay. The day of analysis, cells from two wells in a 6-well plate were collected together with the dead floating cells in the medium, centrifuged 5min at 1200rpm, 4°C and washed twice in cold 1X PBS. Cells were then fixed and permeabilized using the BD Cytotfix/Cytoperm buffer at a concentration of  $1 \times 10^5$  cells/0.5 ml. After 20min incubation on ice, the cells were centrifuged 5min at 1200rpm, 4°C and washed twice with 0.5ml of 1X BD Perm/Wash buffer (initial 10X concentration, diluted in distilled water). Next, the cell pellet was dissolved in 100µl of 1X BD Perm/Wash buffer containing 20µl of the PE (Phycoerythrin) Rabbit

Anti-Active Caspase 3 antibody and incubated for 30min at RT. At the end cells were washed in 1X BD Perm/Wash buffer and resuspended in 0.5ml of 1X BD Perm/Wash buffer. The levels of activated caspase-3 were assessed by flow cytometry analyses using the BD Biosciences FACS Accuri- C6.

#### **VIII.IV Fluorescence-activated cell sorting (FACS) and data processing**

Fluorescence-activated cell sorting analysis of cell viability, cell cycle and apoptosis were conducted using the BD Biosciences FACS Accuri- C6 flow cytometer. First FSC/SSC graphs were produced with gates selecting live and death cells and eliminating cellular debris. PI and PE caspase-3 channels were set on WT non-treated cells. In order to have comparative analysis between different samples, the same threshold for positive cells was applied to all experimental conditions. All flow cytometry generated data was analyzed and processed into graphs with corresponding experimental statistics charts using the FCS Express Flow 6 De Novo software.

Cell cycle studies were performed using the MultiCycle AV tool available with the software which allows the automatic detection of cycling populations and applies different cell cycle models. The best fitting model was selected manually based on the generated statistical data and the expected (already described in the literature) distribution of WT HeLa and HEK cells into the distinct cell cycle phases. The same modeling was applied to all samples and the automatically generated experimental statistics were used for comparison between different conditions.

## IX. Antibodies and oligonucleotide sequences used in this study

**Table 1: Antibodies used in this study**

Target	Species	Application	Dilution	Reference
YTHDC1	Rabbit	IF WB	1/250 1/1000	Abcam #122340
<i>N</i> -terminal peptide of YTHDC1	Rabbit	WB	1/250	Homemade
YTHDC1	Rabbit	ChIP	1/1000	Bethyl A305-096A
YTHDF2	Rabbit	IF	1/100	Proteintech 24744-1-AP
HSF1	Rabbit	WB IF	1/1000 1/200	Cell Signaling #4356
HSF1	Mouse	IF	1/100	Santa Cruz sc-17757
METTL3	Rabbit	IF WB	1/500 1/1000	Proteintech 15073-1-AP
HA	Mouse	IF WB	1/1000 1/1000	Abcam Cat# ab9110
SC35	Mouse	IF	1/500	Sigma S4045
PABPN1	Rabbit	IF WB	1/500 1/5000	Abclonal A6041
HSP70	Rabbit	WB	1/1000	Stressgen #SPA-812
HSP40 (DNAJB1)	Rabbit	WB	1/1000	Bethyl A305245AM
Active caspase 3 PE-conjugated	Rabbit	FACS	1/5	Biosciences 51-68655X
PARP	Rabbit	WB	1/1000	Abcam #4830
FLAG	Mouse	WB	1/1000	Sigma F3165

<b>Target</b>	<b>Species</b>	<b>Application</b>	<b>Dilution</b>	<b>Reference</b>
Tubulin	Mouse	WB	1/5000	Sigma T5168
GAPDH	Rabbit	WB	1/1000	Proteintech 10494-1-AP
GFP	Chicken	IF	1/500	Abcam 13970
Mouse IgG HRP- conjugated	Sheep	WB	1/5000	Dako Cat# P0447
Rabbit IgG HRP- conjugated	Goat	WB	1/5000	Dako Cat# P0448
Mouse IgG (Alexa 488)	Goat	IF	1/500	Invitrogen A-11029
Rabbit IgG (Alexa 594)	Goat	IF	1/1000	Abcam 150080
Rabbit IgG (Dylight 549)	Goat	IF	1/1000	Vector DI-1549

**Table 2: Oligonucleotide sequences used in this study**

Name	Application	Sequence
siYTHDC1#1	YTHDC1 KD assays	CGAGTACATGATTATGATA
siYTHDC1#2	YTHDC1 KD assays	CAAGATGATGAACTGTATA
siYTHDC1#3	YTHDC1 KD assays	AGTCAGCCACAGAGTATAA
siMETTL3#1	METTL3 KD assays	GCUCUUUCCUUAUACAUGUUUCCA
siMETTL3#2	METTL3 KD assays	CUAAGGAACAACAGAGCAAGAAGGU
siC1_1immF	Immunization of YTHDC1 cDNA against siYTHDC1#1	GACTACGACATGAGGGTGGATGATTTC
siC1_1immR		GTGAACCCTTTTATCTCTGTATCTTGCTTC
siC1_3immF	Immunization of YTHDC1 cDNA against siYTHDC1#3	CCGAATACAAAAATGAGGAATATCAAAGATC
siC1_3immR		TAGCGCTTTTTCTTTTGTACTAACTATTC
LT110	Cloning of YTHDC1 at the NotI site of AAVS1-tet3G plasmid	ggtaccctcagccttaaggcgATGGCGGCTGACAGTCGG
LT112	Cloning of YTHDC1 at the NotI site of AAVS1-tet3G plasmid	tgaaaatacagattctcggcTCTTCTATATCGACCTCTCTCCCC
LT44	Genotyping of CRISPR-Cas9 recombinant cell lines	CGATACCAGGAAGTGGACA
LT116	Genotyping of CRISPR-Cas9 recombinant cell lines	AGTTTGCCCAAGCAGTCACC

Name	Application	Sequence
LT119	Genotyping of CRISPR-Cas9 recombinant cell lines	TCCAGGGGTCCGAGAGCTCAGCTAG
LT120	Genotyping of CRISPR-Cas9 recombinant cell lines	CCAGACAGCCGCGTCAGAGCAGCTC
W377A_F	YTHDC1 Mutagenesis. Substitution of Tryptophan 377 with Alanine	GAAGGGTGTAGCGTCCACGCTC
W377A_R		GCTTTGGCAAGAGACACA
W428A_F	YTHDC1 Mutagenesis. Substitution of Tryptophan 428 with Alanine	TCCTATACACGCGGTGCTTCCAGC
W428A_R		GATCCTCCGTGATGTGATTC
$\Delta$ YTH_F	YTHDC1 Mutagenesis. Deletion of the YTH domain	GATGAAAGTATTGACTTGTATC
$\Delta$ YTH_R		TTGAAGCACATATTTGAGTTTAC
YTHDC1_Xho	Amplification of YTH-domain proteins cDNA from HEK cells cDNA library by RT-PCR	CTATCTCGAGgcATGGCGGCTGACAGTCGG
YTHDC1_Xma		GTTCCCCGGGacTCTTCTATATCGACCTCTCTC
YTHDC2-Xho		CTATCTCGAGgcATGTCCAGGCCGAGCAGC
YTHDC2Bam		GTTCGGATCCATCAGTTGTGTTTTTTTTCTCCC
YTHDF1-Xho		CTATCTCGAGgcATGTCGGCCACCAGCGTGG
YTHDF1-Bam		GTTCGGATCCTTGTGTTTGTTCGACTCTGCCG
YTHDF2-Xho		CTATCTCGAGgcATGTCGGCCAGCAGCCTC
YTHDF2-Bam		GTTCGGATCCTTTCCCACGACCTTGACGTTC
YTHDF3_Xho		CTATCTCGAGgcATGTCAGCCACTAGCGTGG
YTHDF3_Bam		GTTCGGATCCTTGTGTTTCTATTTCTCTCCC
Ex1F	Splicing <i>HSP40</i>	CGTCGGACGAGGAGATCAAG



Name	Application	Sequence
Ex1R	Splicing <i>HSP40</i>	CTTCGGTGGCAGAAATCCCT
Ex2F	Splicing <i>HSP40</i>	CTTCCTATGGGCATGGGTG
Ex3R	Splicing <i>HSP40</i>	CCGAAAGGATTCCCCAG
Int1F	Splicing <i>HSP40</i>	CTTCCAAGCTGTTTCCCCCT
Ex2R	Splicing <i>HSP40</i>	TCTTCTTGGTACAGCCGCTG
Int2R	Splicing <i>HSP40</i>	GTGAACAGGACTGACCAGGG
HSPA1A F	PCR/qPCR <i>HSPA1A</i>	GGGCCTTTCCAAGATTGCTG
HSPA1A R	PCR/qPCR <i>HSPA1A</i>	TGCAAACACAGGAAATTGAGAACT
HSPA1B F	PCR/qPCR <i>HSPA1B</i>	TGTTTGTCTTTGAGGTGGACTGT
HSPA1B R	PCR/qPCR <i>HSPA1B</i>	AAGAAGTGAAGCAGCAAAGAGC
DNAJB1CDS_F	PCR/qPCR <i>DNAJB1</i>	TAACCGACGGAAAGAGC
DNAJB1CDS_R	PCR/qPCR <i>DNAJB1</i>	GGAGGTCGTGTCTCCTTCCT
DNAJB1_F5 DNAJB1_R5	Transcription termination qPCR <i>DNAJB1</i> site <b>A</b>	AGGGAAGGGGAGTTGGAACA ACACTCAAACCTGTCCCCG
DNAJB1_F4 DNAJB1_R4	Transcription termination qPCR <i>DNAJB1</i> site <b>B</b>	CTCCTTCCCTGCTAGAACGG CCCCTGTAGAGAGGACCCA
DNAJB1_F1 DNAJB1_R1	Transcription termination qPCR <i>DNAJB1</i> site <b>C</b>	AAGTATCCCACACGCCAAGG ACATTAAGCCAGGAGACGCC
DNAJB1_F7 DNAJB1_R7	Transcription termination qPCR <i>DNAJB1</i> site <b>D</b>	GGCATTGGCATCACAGA TGGTGAAAAGCCGAGTCCTC
DNAJB1_F10 DNAJB1_R10	Transcription termination qPCR <i>DNAJB1</i> site <b>E</b>	TGAGATGCTGCGGGGAATAC TCCCACCCTACCATCCGATT
HSP1AF1 HSP1AR1	Transcription termination qPCR <i>HSPA1A</i> site <b>A</b>	GGCCTTTCCAAGATTGCTG TGCAAACACAGGAAATTGAGAACT
qP7A_F qP7S_R	Transcription termination qPCR <i>HSPA1A</i> site <b>B</b>	GATGCTGATTGGGTTTAGCCTG GGATGTTTCTGCAAAAGTGGTCA
qP8A_F qP8S_R	Transcription termination qPCR <i>HSPA1A</i> site <b>C</b>	TCCCATGCACTTTCCTTGG CAATCCTTAAAGCTCCCCGGA

Name	Application	Sequence
qP9S_F qP9A_R	Transcription termination qPCR <i>HSPA1A</i> site <b>D</b>	CTCCTCTCCCCTTCAGAGCT GGAAAAGGGTGGAGATGGGG
qP10F qP10R	Transcription termination qPCR <i>HSPA1A</i> site <b>E</b>	CGTTGAGGGAGCTGTGGAAT TGAGGGAGGAAAGTGCAACAG
Actin_F	PCR/qPCR <i>actin</i>	CTTCCAGCCTTCCTTCCTGG
Actin-R	PCR/qPCR <i>actin</i>	AATGCCAGGGTACATGGTGG
Clk1 IRFor (Intron 3)	Splicing <i>CLK1</i>	TGCCGCCCACTTGACGTTTCCAG
Clk1 E6Rev (Exon 6)	Splicing <i>CLK1</i>	TTACTGCTACATGTCTACCTCC
Clk1 E2For (Exon 2)	Splicing <i>CLK1</i>	ATGAGACACTCAAAGAGAACTTACTG
Clk1 E2I4For (Exon 2)	Splicing <i>CLK1</i>	ATCACATAGCAGTGCCCAGG
Clk1 E3For (Exon 3)	Splicing <i>CLK1</i>	CACTCAAGGATGTGTGAACCTGGA
Clk1 E3Rev (Exon 3)	Splicing <i>CLK1</i>	CCATGTGAACGACGATGTGAAG

\* All listed oligonucleotide sequences were ordered at Eurogentec

# Results



# **Chapter I**

## **YTHDC1 relocalization and role at nuclear Stress Bodies (nSBs)**



## **I.I Characterization of YTHDC1 relocalization to nSBs**

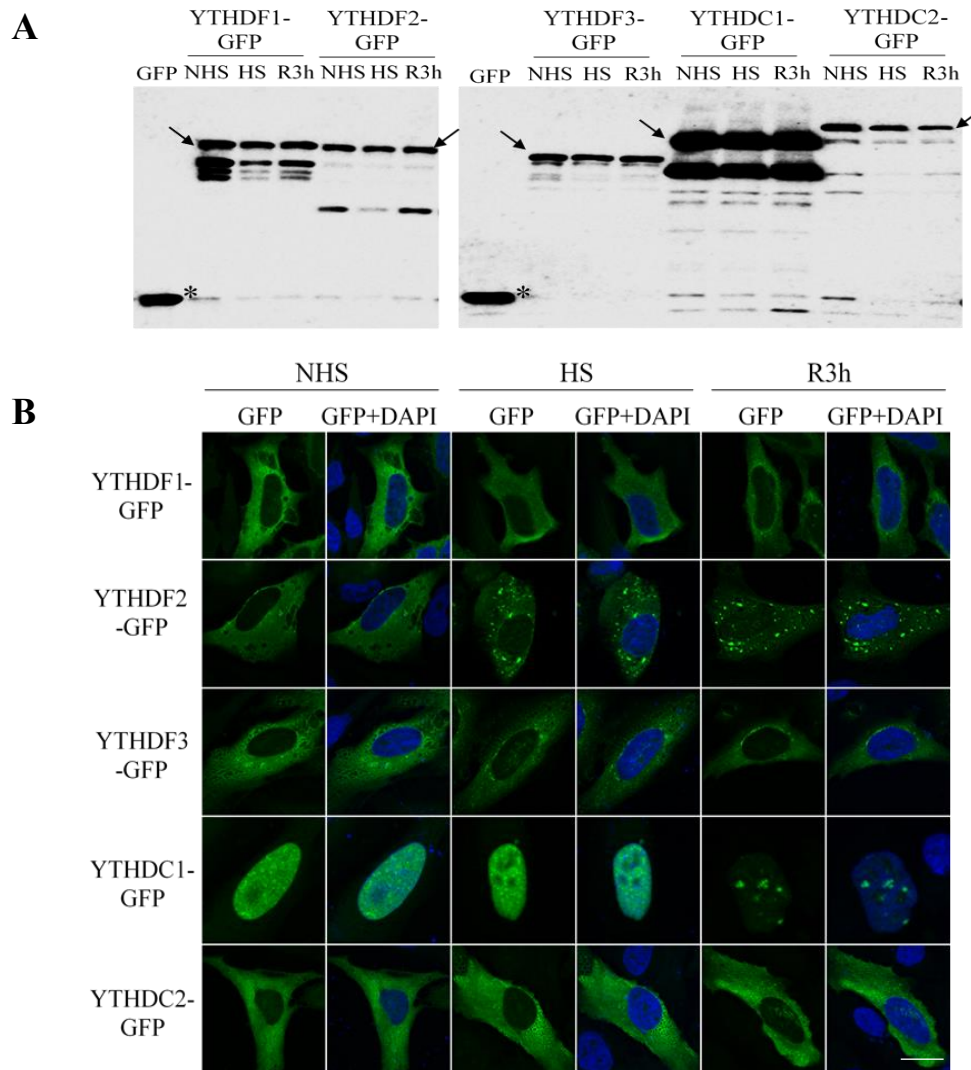
### **I.I.1 Localization of YTH-domain proteins in response to heat shock**

In the beginning of this study, before focusing on the m6A nuclear reader YTHDC1, we were prompt to have a general view of the behavior of YTH-domain proteins in response to heat shock (HS). Very little was known up to this moment regarding a possible role or regulation of these proteins in the context of cellular stress. To address this question, we have studied the subcellular localization of all m6A readers of the YTH-domain family following stress induction.

First, we have constructed plasmids coding for the five members of this family fused to GFP at the C-terminal domain (Materials and Methods III.I). We have defined the correct transfection conditions (Materials and Methods I.II.2) and analyzed by western blot the expression of the proteins in HeLa cells subjected to HS treatment (performed each time in this study for 1h at 43°C) followed or not by a recovery period at 37°C. As shown in the western blot in Fig.32A, all YTH-domain proteins fused to GFP were expressed in HeLa cells in normal conditions (NHS, non-heat shock), after the HS treatment and following the recovery period of 3h at 37°C (R3h). Nevertheless, some protein degradation was observed in all studied time points especially for YTHDF1, YTHDF2 and YTHDC1.

Next, using the same transfection conditions and fluorescent microscopy analysis we have investigated the subcellular localization of these proteins. The obtained transfection rates were variable for the different constructs, ranging from 20 to 30% for the YTHDF proteins and up to 50% for YTHDC1 and YTHDC2. In absence of stress, we observed the subcellular localizations previously described in the literature, with a cytoplasmic distribution of YTHDF1-3 and YTHDC2 and a nuclear localization of YTHDC1 (Fig.32B). Interestingly, following the induction of HS, localization changes were detected for two of the studied proteins- YTHDC1 and YTHDF2.

In the NHS and HS conditions, YTHDC1 presents mainly a diffuse nuclear localization with few YT-bodies reported earlier as specific nuclear structures, formed by the protein and suggested by more recent studies to correspond to nuclear speckles (Nayler et al., 2010; Galganski et al., 2017). The important finding of our study comes from the R3h condition where we observe a remarkable, massive relocalization of YTHDC1 into bright nuclear foci (Fig.32B). No nuclear to cytoplasmic relocation of the protein was detected and nuclear foci were observed in all transfected cells exclusively at the R3h time point.



**Figure 32: Localization of YTH-domain proteins in response to heat shock.** **A.** Western blot analysis of the expression of YTH-domain proteins fused to GFP in HeLa cells transfected with pEGFP-N3-YTH(DF1,DF2,DF3,C1 or C2) and subjected to HS treatment (1h, 43°C), followed or not by 3h recovery period at 37°C (R3h). NHS- non-heat shock. Western blot of whole-cell extract using  $\alpha$ -GFP antibody. Arrows indicate the GFP-fused proteins, \* indicates GFP signal. **B.** Microscopy analysis of the localization of YTH-domain proteins fused to GFP in HeLa cells transfected with the plasmids used at A in NHS, HS and R3h. Cell nuclei are stained with DAPI, scale bar 10 $\mu$ m.

In the case of YTHDF2, immediately after the stress induction the protein presents a granular foci-like cytoplasmic redistribution persisting also later, at R3h. Noticeably, previous studies have suggested a cytoplasmic to nuclear relocalization of this m6A reader in response to stress, however, later on these results have been contested (Zhou et al., 2015; Ries et al., 2019).



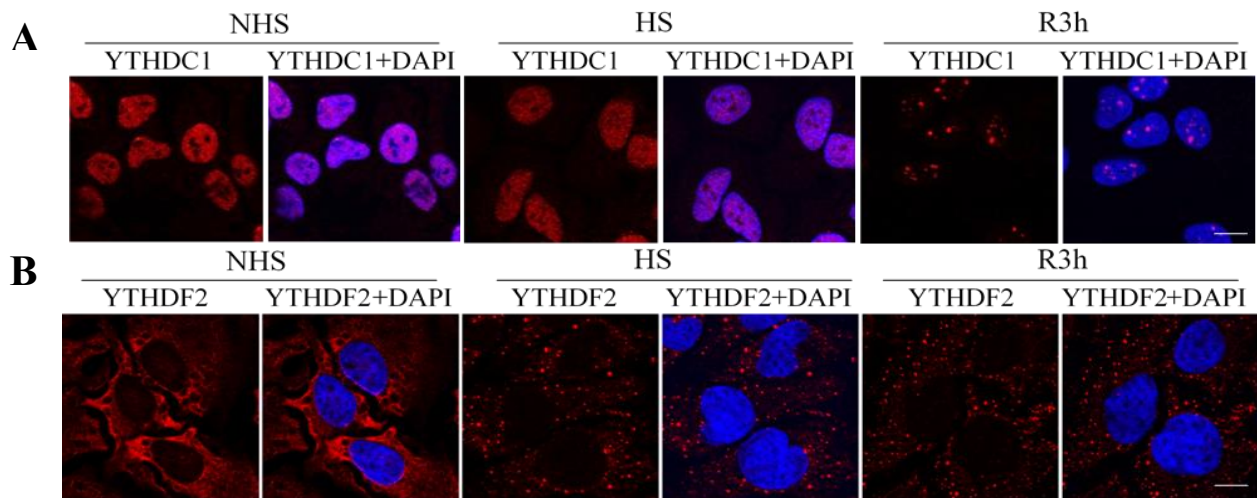
From our initial analysis, we could not exclude a possible interference of the GFP tag with the ability of the protein to relocate. To decipher this we have performed the same experiment this time fusing GFP to the *N*-terminal domain of the protein. Using this construct we once again did not observe a nuclear redistribution of YTHDF2 after HS.

Following on these intriguing findings, we wished to validate the described relocation patterns for YTHDC1 and YTHDF2 through the detection of the endogenous proteins, using specific antibodies.

As shown in Fig.33A in the NHS and HS conditions YTHDC1 has mainly a diffuse, exclusively nuclear localization. However, during the recovery period, the protein forms the previously observed irregular in size and shape nuclear foci, ranging in number from 4 to 7 per cell.

Regarding YTHDF2 (Fig.33B) we detected the diffuse cytoplasmic localization in NHS. After HS and at R3h, the protein concentrates as described earlier into foci-like, cytoplasmic structures whose nature has not been further investigated in this work. Noticeably, these experiments confirmed the absence of nuclear relocation of YTHDF2 in response to stress.

For both YTHDF2 and YTHDC1 this work validated our results obtained using the ectopically expressed, GFP-fused forms.



**Figure 33: Localization of YTHDC1 and YTHDF2 in HeLa cells at NHS, HS and R3h. A.** Immunofluorescence of YTHDC1 (red) in HeLa cells in NHS, HS and R3h. **B.** Immunofluorescence of YTHDF2 (red) in HeLa cells at NHS, HS and R3h. Cell nuclei are stained with DAPI, scale bar 10μm.

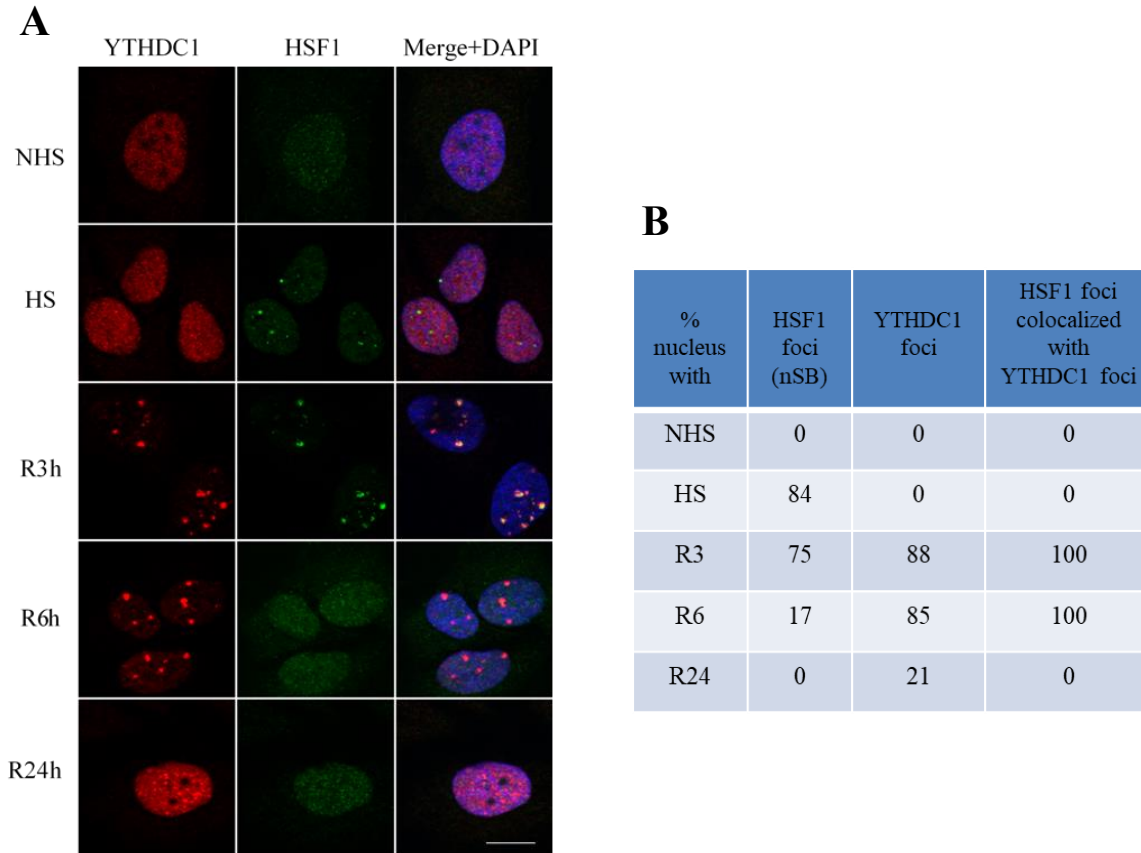
Overall, we have shown for the first time, using two different approaches, a massive nuclear relocation into foci for YTHDC1 during the recovery period after heat stress. Since the main research interest of the team relates to the nuclear processes underlying gene expression regulation, this intriguing discovery prompts us to focus on YTHDC1 and study more in detail its potential implication or regulation in cellular stress response.

### **I.I.2 YTHDC1 relocates to nSBs during the recovery period following heat shock**

Further, in order to better characterize the relocation of YTHDC1, first we wished to investigate the nature of the foci, formed by the protein during the recovery period after HS.

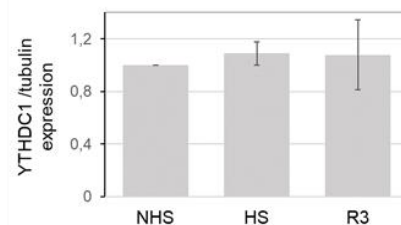
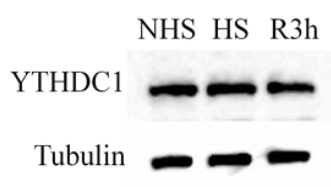
Nuclear Stress Bodies (nSBs), and their implication in the cellular response to heat stress, have been a subject of extensive studies in our team. Upon HS the transcription factor HSF1 (Heat Shock Factor 1) concentrates into nuclear granules and activates the production of the long non-coding RNAs *SATIII* (Introduction II.IV.2). Multiple transcription, processing and splicing factors bind to the *satIII* DNA regions, the lncRNAs *SATIII* and to HSF1 to form the transient structures called nSBs, characterized by a dense nucleoprotein content. The resemblance between nSBs and the foci formed by YTHDC1 led us to address the potential relocation of the nuclear m6A reader to these particular stress-induced granules.

To examine this question we have carried out a co-immunodetection of YTHDC1 and HSF1 in normal conditions and after the stress induction. This study clearly showed a colocalization of the two proteins at R3h (Fig.34A). HSF1 foci are formed immediately after the HS, begin to disassemble at R3h and are no longer detected after R6h. YTHDC1 foci form a bit later, mostly at R3h and persist up to R6h. In Fig.34B are presented the percentage of cells with HSF1 and YTHDC1 foci and the percentage of cells in which the two proteins colocalize at the different time points. Based on this analysis the two proteins colocalize at nSBs in 100% of the cells at R3h. After HS, HSF1 is present at nSBs, while YTHDC1 is not. Conversely, at R6h YTHDC1 is strongly relocated to nSBs while HSF1 is no longer accumulating at these structures. These results indicate that both HSF1 and YTHDC1 foci are forming transiently mostly at the same place in the nucleus, in response to HS but do not present exactly the same kinetics of relocation.



**Figure 34: YTHDC1 relocates to nuclear Stress Bodies (nSBs) in response to heat shock. A.** Immunofluorescence of YTHDC1 (red) and HSF1 (green) in HeLa cells at NHS, HS or after 3, 6 or 24h of recovery (R3h, R6h, R24h). Cell nuclei are stained with DAPI, scale bar 10 $\mu$ m. **B.** Percentage of cells with HSF1 foci, YTHDC1 foci and percentage of the cells in which the two proteins colocalize at NHS, HS, R3h, R6h and R24h. 200 cell nucleus were counted per condition. *Performed by Solenne Dufour.*

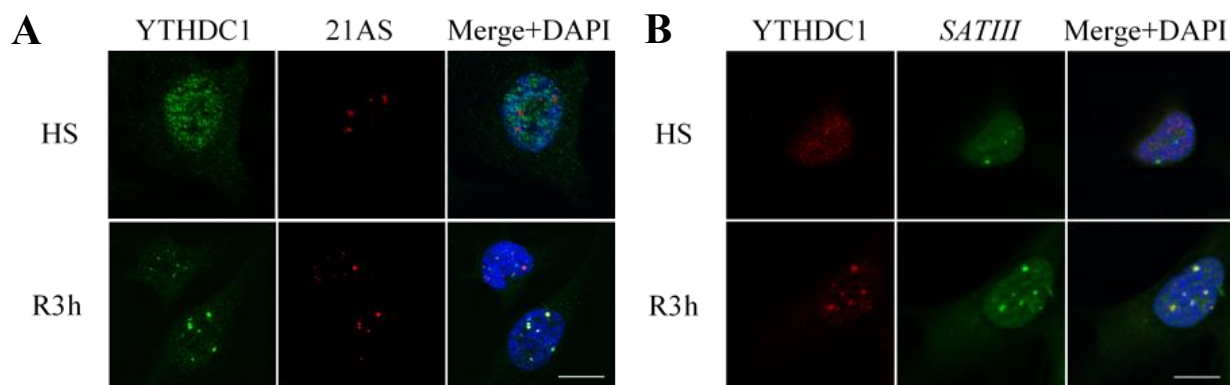
We have next investigated whether the specific concentration of YTHDC1 into granules was associated or not with changes at the protein level. As shown in Fig.35, using western blot analysis we do not detect any modifications of the abundance of this protein, in HS or during the recovery (R3h).



**Figure 35: YTHDC1 protein levels in response to stress.** Western blot of whole-cell extract of HeLa cells, using  $\alpha$ -YTHDC1. Quantification of YTHDC1 protein levels at NHS, HS and R3h compared to tubulin levels.

Further, using DNA Fluorescence *in situ* hybridization (FISH) combined with immunofluorescence (Materials and methods VI.II.2), we have studied whether YTHDC1 relocates precisely to the genomic sites of formation of nSBs. The number of nSBs correlates with the ploidy of the cells and represents also primary and secondary HSF1 binding sites. Primary HSF1 foci are forming at the 9q12 pericentric repetitive region (three large foci in HeLa cells) and secondary foci are found at the same regions of chromosomes 1, 16, Y and acrocentric chromosomes (13, 14, 15, 21 and 22). These regions encompass the *satIII* DNA repetitive elements, producing the *SATIII* lncRNAs. The specific probe applied in these studies detects the primary as well as the secondary nSBs formed at Y and acrocentric chromosomes. The microscopic analyses indicate that in absence of stress YTHDC1 is not found specifically at the *satIII* repeats. However, we observe that it relocates to pericentric repetitive heterochromatin at R3h when the formation of nSBs is initiated. (Fig.36A). YTHDC1 foci correspond in number and localization to the specifically stained chromosomal sites.

We have next examined the possible colocalization of YTHDC1 with *SATIII*, using a specific probe detecting the lncRNAs. This RNA FISH experiment, combined with immunofluorescence (Materials and methods VI.II.1), demonstrated that in HS *SATIII* are already detected while YTHDC1 is not yet redistributed into foci (Fig.36B). Later, at R3h, we observe the colocalization of the protein with the newly synthesized transcripts. Thus, we conclude from these results that YTHDC1 relocates to heterochromatin at the level of pericentromeric regions formed by *satIII* repeats and colocalizes with the *SATIII* RNAs at R3h.



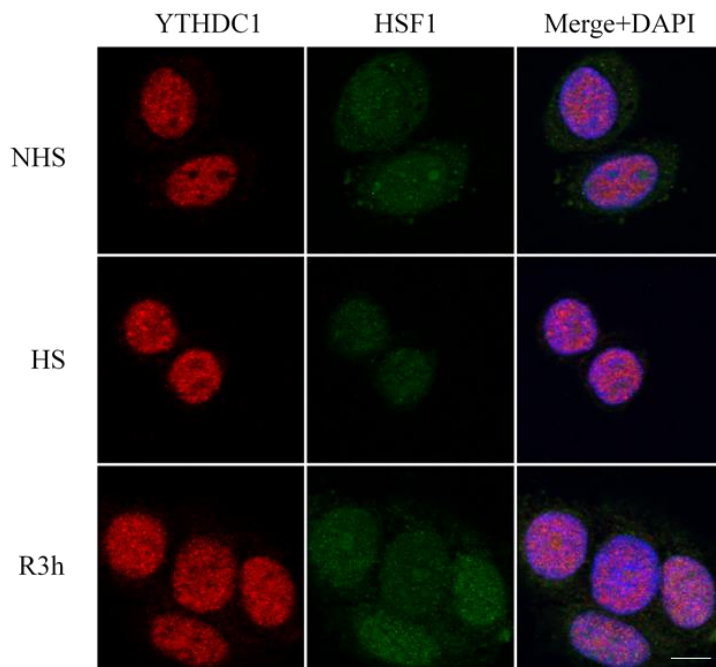
**Figure 36: YTHDC1 relocates to pericentric repetitive heterochromatin and colocalizes with the *SATIII* lncRNAs.** **A.** DNA FISH of the chromosomal (9, Y, 13, 14, 15, 21 and 22) pericentric regions of formation of nSBs (red) combined with YTHDC1 immunofluorescence (green) in HeLa cells at HS and R3h. **B.** RNA FISH of *SATIII* RNAs (green) and immunofluorescence of YTHDC1 (red) in HeLa cells at HS and at R3. Cell nuclei are stained with DAPI, scale bar 10μm. *Performed by Solenne Dufour.*

Altogether, this series of studies identified the nuclear structures at which YTHDC1 is found upon heat stress as nSBs. YTHDC1 relocates to nSBs specifically during the recovery period after HS induction. This is in contrast with HSF1 which is recruited at the onset of the HS response. YTHDC1 persists longer at these structures suggesting that the key element maintaining its relocation is not HSF1.

### I.1.3 HSF1-dependent *SATIII* transcription is required for YTHDC1 relocation

In the search for the determinant of YTHDC1 relocation to nSBs we have investigated on one hand the importance of HSF1 for this process and on the other hand the role of *SATIII*.

First, we have examined YTHDC1 localization in response to stress in HSF1 stable KD cell lines. In these cell lines nSBs are not forming and the *satIII* repeats are not transcribed (Koskas et al., 2017). Importantly, YTHDC1 shows a diffuse nuclear distribution following stress and does not form foci at R3h when HSF1 is not present in the cells (Fig.37). This result highlights the importance of HSF1 and indicates that its presence is required for YTHDC1 relocation in response to stress.

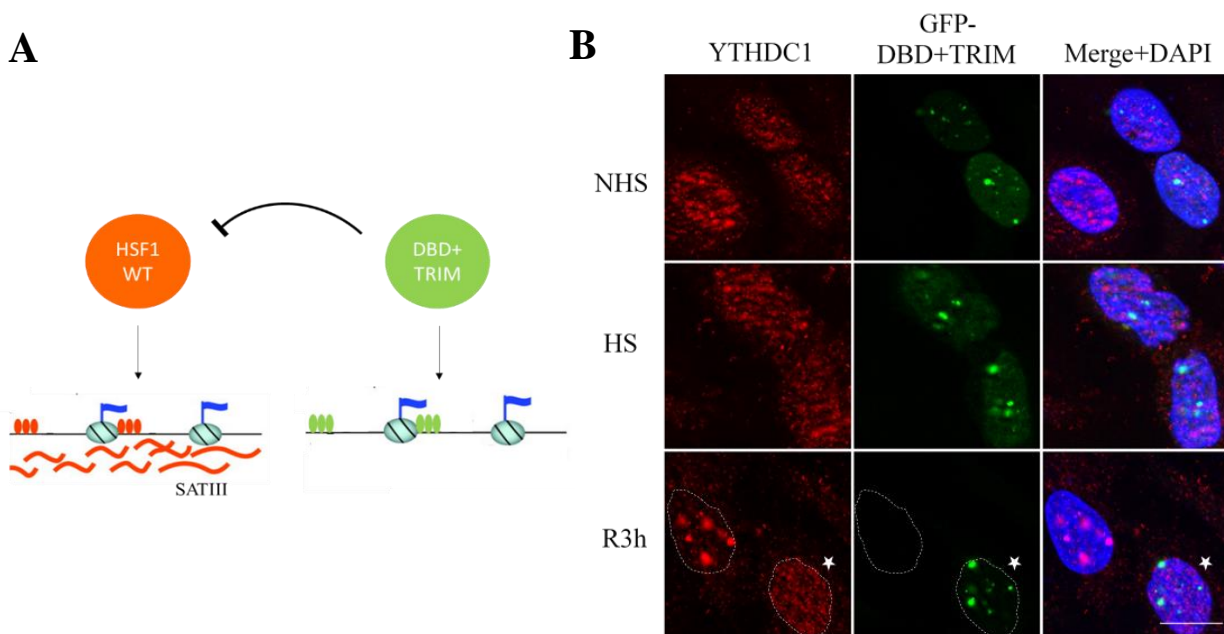


**Figure 37: Active HSF1 is required for the relocation of YTHDC1.** Immunofluorescence of YTHDC1 (red) and HSF1 (green) in HeLa HSF1 stable KD cell lines at NHS, HS and R3h. Cell nuclei are stained with DAPI, scale bar 10µm. *Performed by Solenne Dufour.*

As described previously YTHDC1 foci persist even when HSF1 is no longer found at nSBs at R6h, indicating that if HSF1 is necessary for YTHDC1 relocation to nSBs it is not sufficient to

maintain it there. For this reason, we then examined the importance of *SATIII* RNAs for YTHDC1 relocalization to nSBs. These lncRNAs are transcribed shortly after the HS treatment, tend to accumulate near the transcription site and persist up to 12h of recovery (our unpublished data).

To examine their role in the formation and maintenance of YTHDC1 foci we took advantage of a mutant of HSF1, called HSF1<sup>DBD/TRIM</sup>. This deletion mutant contains only the DNA-binding and trimerization domains of HSF1 and has lost its transactivation domain, which results in the constitutive formation of granules without activation of the transcription of *SATIII* RNAs (Fig.38A). Importantly, HSF1<sup>DBD/TRIM</sup> acts as dominant negative, meaning that in cells transfected with this mutant, endogenous HSF1 is not able to accumulate into stress granules upon HS (Jolly et al., 2002).



**Figure 38: HSF1-dependent *SATIII* transcription is required for YTHDC1 relocalization to nSBs. A.** Schematic representation of WT HSF1 and HSF1<sup>DBD/TRIM</sup> in response to HS. Arrow indicates that the mutant is a dominant negative and inhibits the endogenous HSF1. **B.** Immunofluorescence of YTHDC1 (red) in HeLa cells transfected with GFP- HSF1<sup>DBD/TRIM</sup> at NHS, HS and R3h. Cell nuclei are stained with DAPI, scale bar 10µm. *Performed by Solenne Dufour.*

This experiment indicates that in cells expressing HSF1<sup>DBD/TRIM</sup>, in which HSF1 stress granules are formed by the mutant but the transcription of *SATIII* RNAs is not activated, YTHDC1 does not form nuclear foci at R3h (Fig.38B). In contrast, in the non-transfected cells, where

endogenous HSF1 is active and there is transcription of *SATIII* RNAs, we observed the characteristic redistribution of YTHDC1 into bright foci.

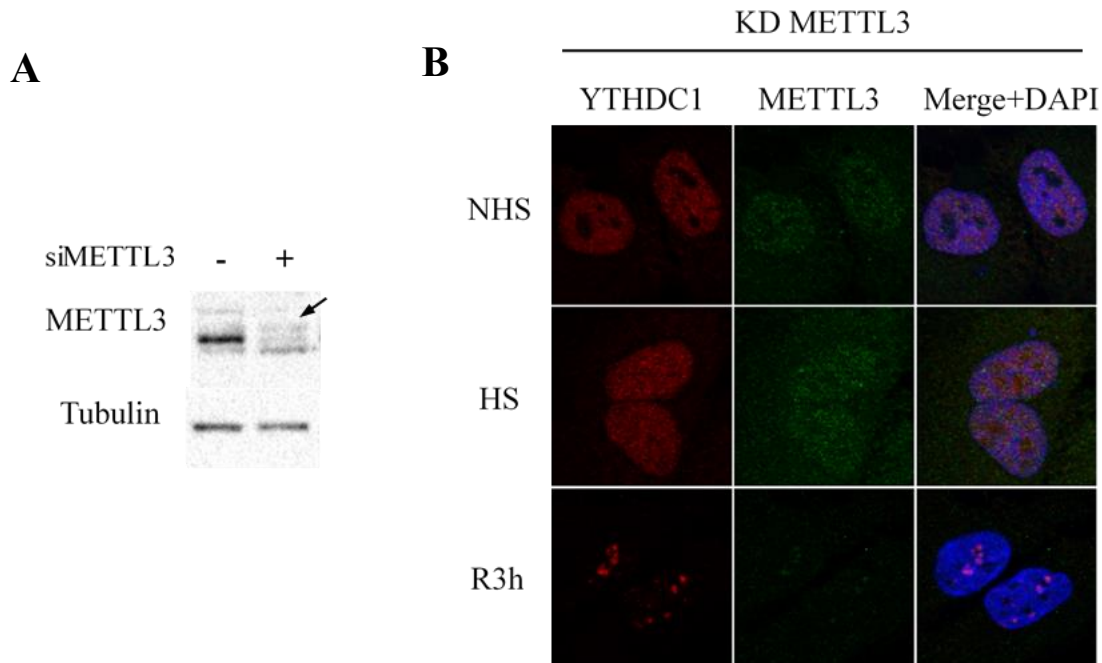
Thanks to the use of this HSF1 mutant, we were able to distinguish the importance of the unique relocation of HSF1 to its genomic targets separately from its capacity to activate in *cis* the transcription of *satIII* repeats. From the obtained results, we can conclude that the accumulation of HSF1 at pericentric regions is necessary but not sufficient for the relocation of YTHDC1. Notably, YTHDC1 relocation to nSBs turns out to be dependent on the production of *SATIII* transcripts. These lncRNAs persist longer than HSF1 at nSBs and possibly maintain YTHDC1 at these sites during the recovery period.

#### **I.I.4 YTHDC1 relocation to nSBs could be m6A-independent**

YTHDC1 is mainly studied as a nuclear reader of m6A and many of its molecular functions have been associated with the presence of this RNA mark. To understand more precisely the molecular mechanisms involved in YTHDC1 relocation to nSBs, we have examined the importance of m6A recognition for this process. Two different approaches have been used to address this question.

In the first approach we have performed a transient knock-down of the main methyltransferase responsible for m6A deposition - METTL3 (Liu J. et al, 2014). To begin we have defined the optimal conditions for the efficient KD. As shown in Fig.39A, we found the KD conditions to induce a strong decrease of the levels of the methyltransferase.

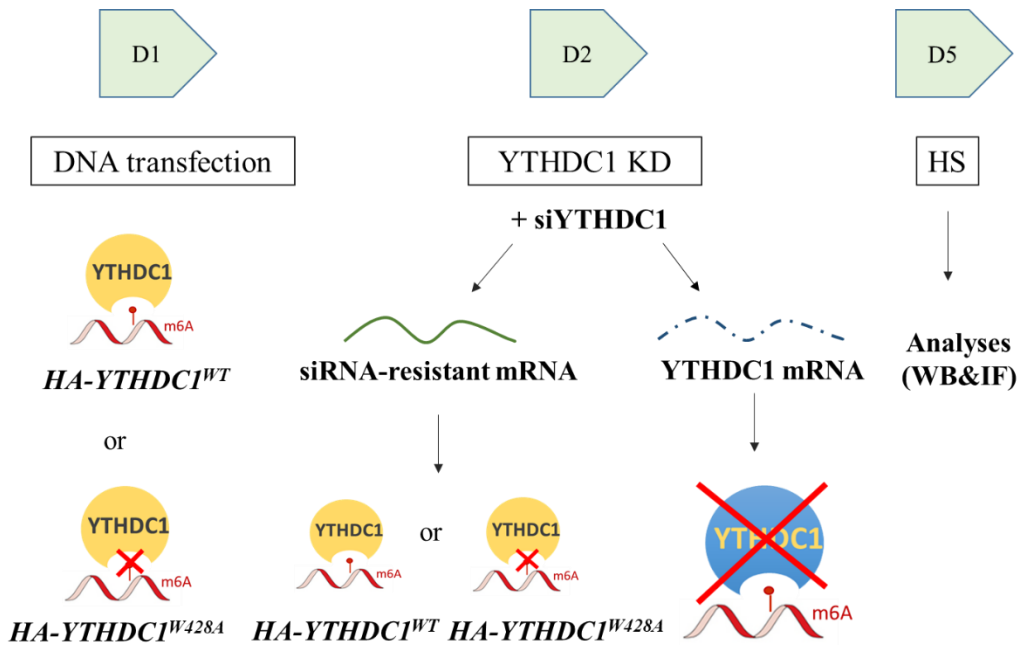
Immunofluorescence of YTHDC1 in METTL3 KD cells following HS treatment demonstrates that the protein relocates to nSBs at R3h (Fig.39B). The decrease of the methyltransferase levels did not correlate with a decrease of the signal of YTHDC1 foci suggesting that YTHDC1 relocation might be an m6A-independent mechanism. Nevertheless, since low levels of METTL3 are still detected by western blot and immunofluorescence we could not completely exclude the presence of residual m6A in these cells.



**Figure 39 : YTHDC1 relocalization could be m6A-independent.** **A.** Western blot analysis of METTL3 on whole-cell extract in HeLa cells transfected (+) or not (-) with specific siRNAs against METTL3. Tubulin is used as loading control. Arrow shows the corresponding METTL3 signal. **B.** Immunofluorescence of YTHDC1 (red) and METTL3 (green) in HeLa METTL3 KD cells at NHS, HS and R3h. Cell nuclei are stained with DAPI, scale bar 10 $\mu$ m.

We have then used a second approach to examine the importance of m6A for YTHDC1 relocalization. This time in YTHDC1 KD cells we have transiently expressed a mutant of the protein which can no longer recognize m6A residues. Importantly, the introduced mutation (substitution of tryptophan 428 to alanine, W428A), does not affect the structure of the protein (Xu et al., 2014). These experiments were conducted in YTHDC1 KD cells based on the fact that previous studies in our team showed that YTHDC1 can dimerize (unpublished data). Therefore, it was important here to observe exclusively the ectopically expressed protein in absence of the endogenous one. The mutant (HA-YTHDC1<sup>imm</sup><sup>W428A</sup>, Materials and Methods III.II) is named here HA-YTHDC1<sup>W428A</sup>. Silent mutations have been introduced in the sequence coding for this mutant in order to protect it against the siRNAs used for the specific KD of the endogenous YTHDC1. As a control of these rescue experiments (Materials and Methods I.II.3) cells were transfected also with HA-YTHDC1<sup>WT</sup> (HA-YTHDC1<sup>imm</sup>), non-degradable by the siRNAs but still binding the m6A mark. The experimental procedure is schematically described in Fig.40.



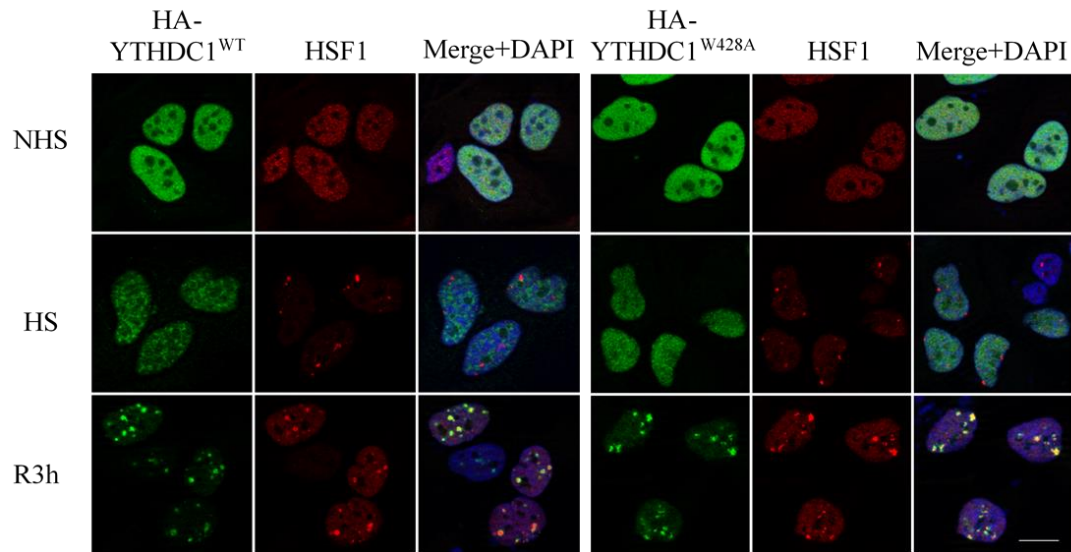


**Figure 40: Schematic representation of the rescue experiments used to study m6A importance for YTHDC1 relocation to nSBs.** On Day 1 (D1) HeLa cells are transfected with HA-YTHDC1<sup>WT</sup> or HA-YTHDC1<sup>W428A</sup> (in red: RNA with the m6A mark). On Day 2 (D2) a transient KD of YTHDC1 is carried out using specific siRNAs. The non-degradeable mRNAs of HA-YTHDC1<sup>WT</sup> and HA-YTHDC1<sup>W428A</sup> are not degraded while YTHDC1 mRNA is targeted and degraded. On Day 5 (D5) HS treatment is performed followed by Western blot (WB) and Immunofluorescence (IF) analyses.

First, using western blot analysis we defined the conditions allowing an effective KD of YTHDC1 and the simultaneous expression of the ectopic proteins.

Next, we have studied specifically the localization of HA-YTHDC1<sup>WT</sup> and HA-YTHDC1<sup>W428A</sup> using immunodetection of the HA tag. As shown in Fig.41 in YTHDC1 KD cells these two ectopically expressed proteins relocate to nSBs in response to stress during the recovery period. The mutation disrupting m6A binding did not affect the formation of YTHDC1 foci at R3h. This result further suggests that m6A is not essential for YTHDC1 localization following HS. However, the KD of the endogenous protein is not absolute and thus we cannot exclude a potential dimerization of YTHDC1 with HA-YTHDC1<sup>WT</sup> or HA-YTHDC1<sup>W428A</sup> which might drive the relocation of the ectopic proteins to nSBs.

Taken together the experiments carried out here using these two different methods support the idea that YTHDC1 relocation to nSBs might be an m6A-independent process.



**Figure 41: YTHDC1 relocation is possibly m6A-independent.** Immunofluorescence of HA (green) and HSF1 (red) at NHS, HS and R3h in YTHDC1 KD HeLa cells transfected with HA-YTHDC1<sup>WT</sup> or HA-YTHDC1<sup>mut</sup>. Cell nuclei are stained with DAPI, scale bar 10 $\mu$ m

More generally with this work we were able to demonstrate that YTHDC1 relocates to nSBs in response to heat stress. Furthermore, we have defined the determinants by which the specific nuclear redistribution of the protein is achieved. We have shown that the process is dependent on the HSF1-activated *SATIII* RNAs transcription and that it is most likely m6A-independent. Following on this initial characterization of YTHDC1 relocation, in the next part of our study we searched whether this event is linked to a specific molecular function of the protein in the context of cellular stress.

## **I.II Role of YTHDC1 at nuclear Stress Bodies**

In order to characterize YTHDC1 molecular and cellular functions one of the strategies initiated in the team, while I was conducting my Master 2 project, was based on the purification of YTHDC1 and the identification of its protein partners by mass spectrometry. This work was mainly conducted by my PhD advisor, Daphné Seigneurin-Berny, and I contributed to the analysis of YTHDC1 interactome. Noticeably, this approach showed a strong physical link between YTHDC1 and the splicing machinery. Later on, a study by Ninomiya and colleagues (Ninomiya et al., 2019) reported a new role for nuclear Stress Bodies in adapting the genome expression through the regulation of the splicing of several hundreds of different mRNAs (discussed in Introduction II.IV.2). Taken together these two series of results led us to realize that YTHDC1 relocation to nSBs might be part of the cell stress response leading to the regulation of splicing of several hundreds of genes upon HS. We then tested this possibility.

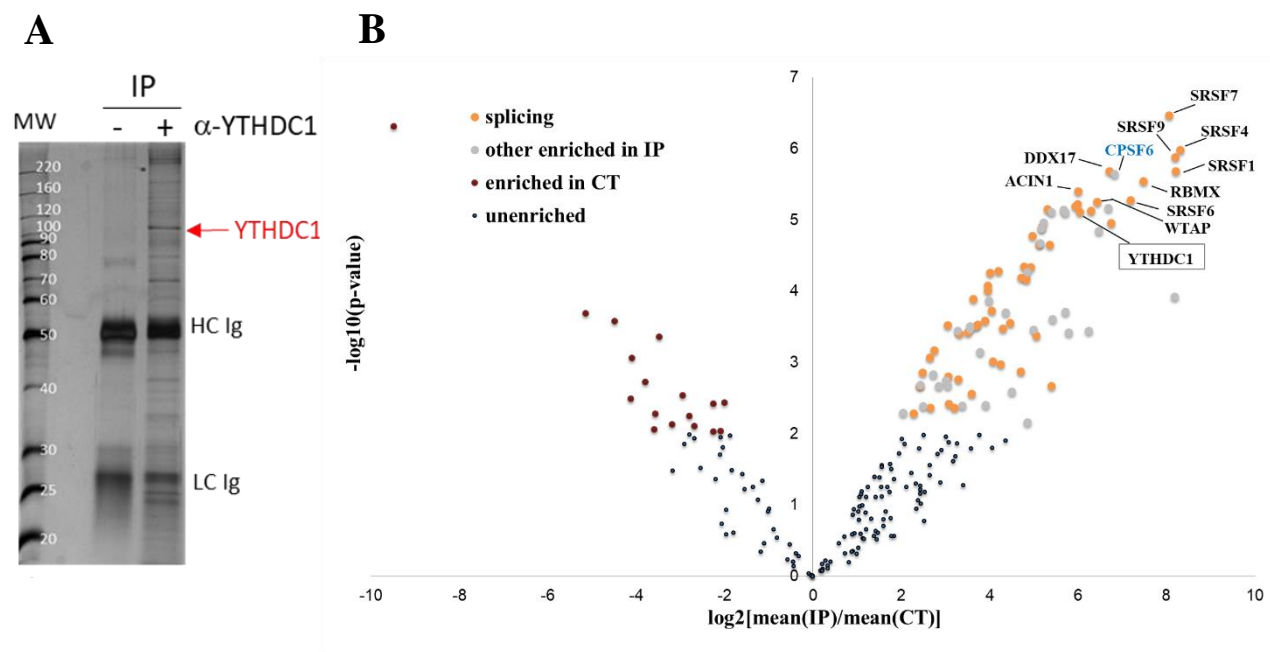
Hence, in this next section, I first present the results of YTHDC1 purification, followed by the comparison of YTHDC1 proteome with the list of proteins found associated with *SATIII* lncRNAs, the core of nSBs formation. I will finish by presenting our findings on YTHDC1 function in regulating splicing in the context of its relocation to nSBs in response to stress.

### **I.II.1 Interactomes of YTHDC1 and *SATIII* lncRNAs**

#### *Interactome of YTHDC1*

In order to identify the protein partners of YTHDC1 we have used a non-targeted approach. An affinity purification of the protein was performed using a specific antibody raised against its *N*-terminal domain. The immunoprecipitation was carried out in triplicate, in HEK cells subjected to normal growth conditions (Materials and Methods IV.II). For each experiment, the soluble proteins extract was used for two purifications conducted in parallel, one realized with the YTHDC1 antibody and the second (the control) realized with an irrelevant antibody against rabbit immunoglobulins. The efficiency of the immunoprecipitation was first estimated by western blot analysis and by silver staining. As shown in Fig.42A, several proteins, including YTHDC1, are clearly enriched from the IP conducted with the specific antibody, compared to the control. These different experimental controls allowed us to conclude that the purification of YTHDC1 and its partners was successful.

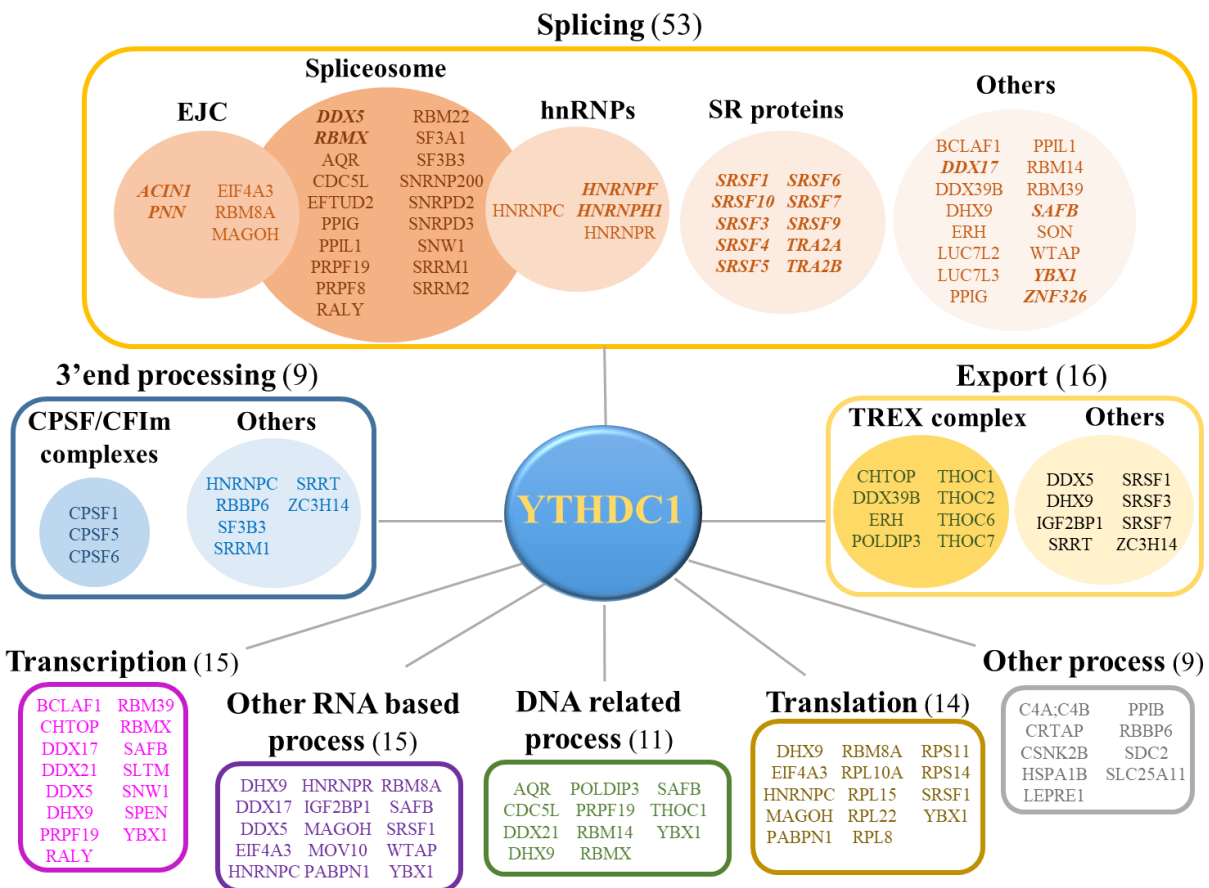
Next, the proteins were identified using Liquid Chromatography coupled with tandem Mass Spectrometry (LC-MS/MS) (collaboration with EDyP, CEA, Grenoble). Noticeably, this analytical technique was chosen since it is qualitative and quantitative. A high stringency *in silico* analysis was conducted on the identified proteins. Exclusively proteins (with at least one unique peptide) found in the three independent experiments and presenting fold enrichment higher than 2 have been considered as putative YTHDC1 partners. The results show that YTHDC1 interactome is composed of a subset of 86 enriched proteins. Almost all of them are involved in nuclear RNA-based processes. In particular, more than 60% of the identified proteins are linked to the splicing process. The volcano plot shown in Fig.42B, represents the distribution of the proteins according to their average enrichment compared to the control and the statistical significance (p-value) of their enrichment.



**Figure 42: Identification of YTHDC1 molecular partners.** **A.** Silver staining of acrylamide gel presenting the eluted fractions using non-relevant antibody anti-rabbit IgG (-) or the antibody against YTHDC1 (+),  $\alpha$ -YTHDC1. HC Ig- heavy chains of immunoglobulins (Ig), LC Ig- light chains of Ig. **B.** Volcano plot of YTHDC1 partners involved in splicing, the other IP-enriched proteins, the CT (control)-enriched proteins and the unenriched proteins identified by LC-MS/MS. Each protein is presented as a dot with corresponding average enrichment compared to the control and a p-value ( $\log_2[\text{mean}(\text{IP})/\text{mean}(\text{CT})]$ ) within the three independent immunoprecipitations. Indicated are the names of the most significantly enriched proteins.

The proteins are classified into specifically enriched in the three immunoprecipitations (gray and orange circles), enriched in the control and unenriched (contaminants). Here, the once implicated in splicing (orange circle) are represented apart from the other potential partners since they are found to be the main YTHDC1 interactors. Remarkably, the ten most significantly enriched proteins are involved in regulating alternative splicing. Their names are also indicated on the plot.

We have further classified the identified proteins according to their molecular functions and when applicable to their corresponding protein complex (Fig.43). It is of note that some of the proteins have multiple roles and therefore may appear in more than one category. We have chosen to represent here the main functional classes of YTHDC1 putative partners.



**Figure 43: Schematic representation of the protein interactome of YTHDC1.** The putative YTHDC1 partners are grouped based on their molecular functions and if applicable to the corresponding protein complex. Represented are the main functions according to analyses performed with PubMed and UniProt databases. The proteins implicated in alternative splicing are indicated in ***bold and italic shrift***. The number of proteins associated with each function is given in brackets.

As mentioned earlier, this study showed that the largest proportion of identified proteins, 53 out of 86, are involved in splicing. Many of these proteins are part of the major splicing complex- the spliceosome or represent spliceosome regulators and adaptors- the Exon-Junction Complex (EJC) proteins, the serine/arginine-rich proteins (SR), the heterogeneous nuclear RiboNucleoProteins (hnRNPs) and others. Although the majority of partners found in this category are essential for constitutive splicing, an important number of them (20/53), particularly the SR proteins and the hnRNPs, are central players of alternative splicing events (cf. Fig.43).

YTHDC1 partners participate also in other mRNA-based processes such as 3'-end processing and mRNA export. For instance we have identified components of the Cleavage and Polyadenylation Specificity Factor complex CPSF, and one of its units- the CFIm complex, responsible for alternative polyadenylation of newly synthesized transcripts. Regarding the export of mRNAs, we detected notably TREX complex constituents (proteins of THO, TREX subcomplex) and several of their assisting partners such as SR proteins. Moreover, some of the identified YTHDC1 interactors are important for translation regulation (RPL proteins and others) or for distinct RNA-based processes as miRNA processing (DHX9, DDX17 and DDX5) or mRNA decay (e.g EIF4A3, RBM8A, MAGOH).

The rest of the functional categories of YTHDC1 partners are implicated in transcription-based regulation (BCLAF, SLTM, SPEN and others) or DNA-related processes (PRPF19, RMB14 and RMBX for example).

Many of the identified proteins correlate with the already described molecular functions of YTHDC1- splicing, mRNA export and 3'-end processing. Moreover, part of them have been previously identified in other studies- SR proteins (SRSF1, SRSF3, SRSF7, SRSF9, and SRSF10), (Xiao et al., 2016) and TREX components (CHTOP TREX subunits and DDX39B) (Lesbirel et al., 2018).

The immunoprecipitation of YTHDC1 and the analysis of its interactome demonstrated that the protein partners are implicated in vast and diverse mechanisms of regulation of RNA biogenesis at multiple levels with a clear emphasis for splicing in general and alternative splicing in particular.

### *SATIII-interacting proteins*

Even though molecular functions have been assigned to the *SATIII* lncRNAs, many of the related mechanisms of action have remained poorly understood for long time. The recent study by Ninomiya et al. (discussed in Introduction II.IV.2) demonstrated that *SATIII* are involved in the intron retention at hundreds of genes during the recovery period following HS. In the search of the molecular mechanisms standing by this process, the authors identified the *SATIII*-interacting proteins. The ChIRP (Chromatin Isolation by RNA Purification) experiments were carried out in HeLa cells during the recovery period after HS treatment. The results identified a protein interactome of 141 proteins susceptible to bind these lncRNAs. A sizeable fraction of *SATIII*-interacting proteins happen to be involved in mRNA processing and in particular in splicing - SR proteins, spliceosome components and splicing cofactors. The description of *SATIII* interactome, together with the studies of the effect of *SATIII* on splicing, shed light for the first time on the molecular function of *SATIII* and their associated nSBs in response to heat stress.

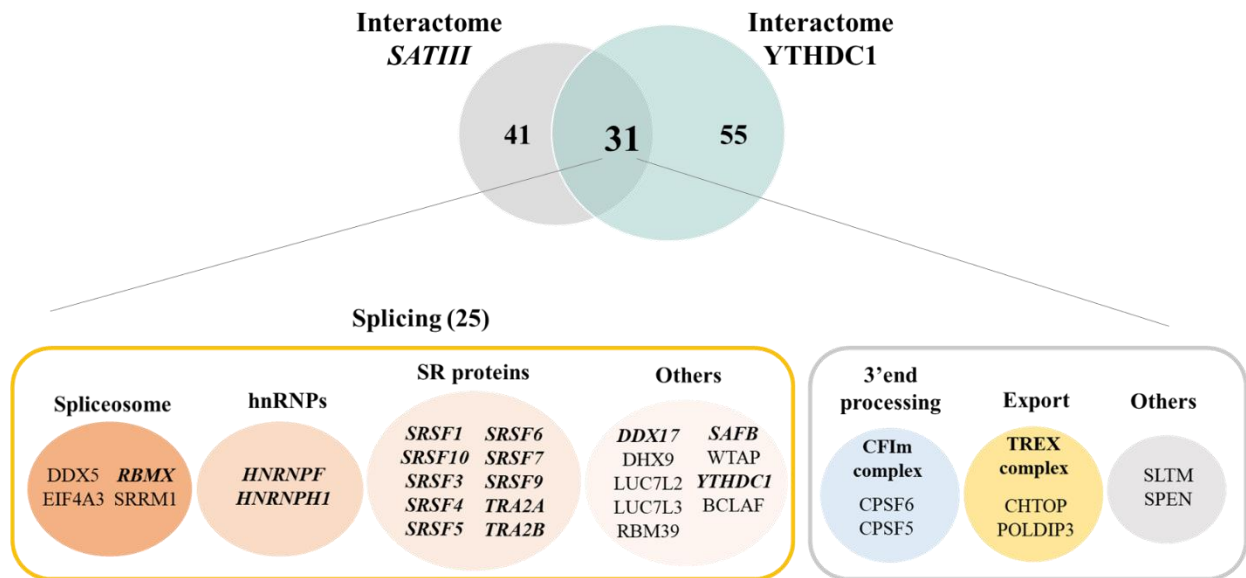
### *YTHDC1 partners and SATIII partners*

We found that following stress YTHDC1 relocalizes to nSBs, the site of transcription of *SATIII* RNAs. The purifications of YTHDC1, on one side, and *SATIII*, on the other side, clearly indicated that both YTHDC1 and the lncRNAs interact predominantly with RNA splicing factors. This led us to the hypothesis that YTHDC1 might be a central actor of the recently described *SATIII*-dependent splicing control taking place upon heat stress.

We then conducted a more in depth comparison of the protein interactomes of *SATIII* and YTHDC1. Based on the selection criteria used for YTHDC1 partners, we have determined from the total of 141 *SATIII*-interacting proteins a subset of 72 partners identified by at least one peptide in the three immunoprecipitation experiments. From the 86 YTHDC1 partners and the 72 *SATIII*-interacting proteins, 31 were found in both lists (Fig.44). This result demonstrates a remarkable resemblance between the two interactomes.

Notably, 25 of the 31 common proteins are major splicing regulators- SR proteins and hnRNPs as well as core spliceosome components and other proteins with diverse roles in splicing (Fig.44). Interestingly, 16 of these 25 splicing actors, are involved also in alternative splicing

processes (cf. Fig.44). Another part of the shared interactors are implicated in 3'end processing (CFIm complex), mRNA export (TREX complex) or other nuclear regulatory mechanisms.



**Figure 44. Schematic representation of YTHDC1 and SATIII interactomes and the functional classes of the common partners.** SATIII-interacting proteins (grey circle), 72 in total. YTHDC1 interacting proteins (green circle), 86 in total. 31 common partners of YTHDC1 and SATIII, most of them (25) are involved in splicing (spliceosome, hnRNPs, SR proteins and others). In **bold and italic** are presented the actors of alternative splicing events. The rest of the common partners participate in 3'end processing, export and other processes.

The important similarity among YTHDC1 and SATIII interactomes, highlighted by this investigation, reinforced the hypothesis that YTHDC1 could be essential player in SATIII-dependent splicing events in the context of the cell response to heat stress.

### I.II.2 YTHDC1 and SATIII-dependent splicing events

The study by Ninomiya et al. showed that during the recovery period after HS a key transcript differentially spliced in presence or absence of SATIII RNAs is *CLK1* (CDC-like Kinase 1) mRNA. CLK1 is a major regulator of alternative splicing, activated in response to stress (Duncan et al., 1997). Following HS induction *CLK1* mRNA is as well subjected to alternative splicing modulated by SR/SRSF proteins, which leads to the expression of its catalytically active form. CLK1 is then recruited to nSBs where it phosphorylates SR/SRSF pre-mRNA splicing factors. In



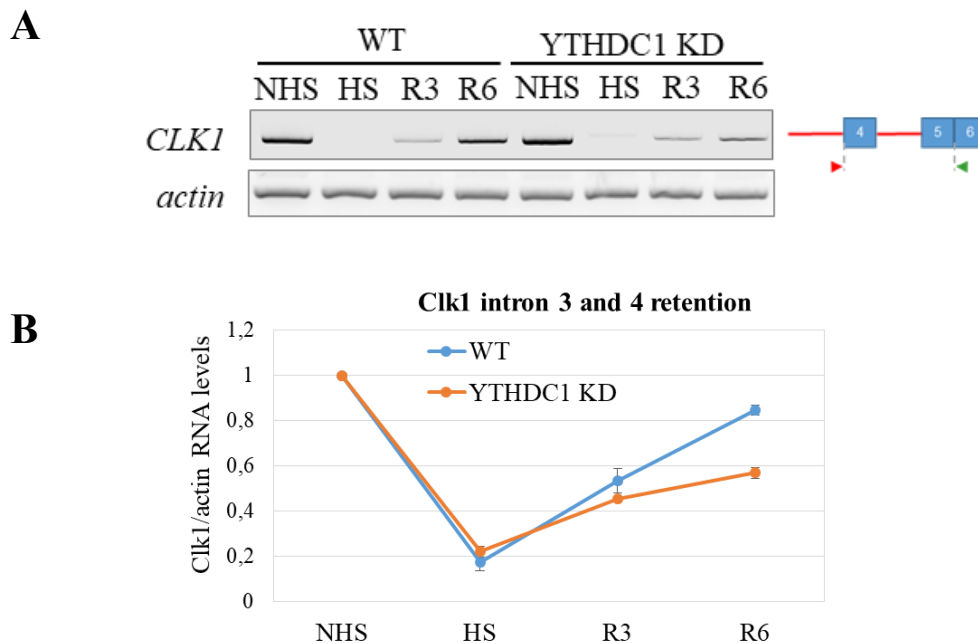
general, the phosphorylation state of SR/SRSFs affects greatly the splicing patterns of approximately 500 hundreds of mRNAs during the recovery period, including *CLK1* mRNA. Therefore the CLK1 autoregulation loop is crucial for the proper expression of a large number of genes in response to stress (cf. Fig. 31, Introduction).

The kinase is composed of 13 exons and 12 introns (Ninomiya et al., 2011). In normal conditions, the mature RNA retains only introns 3 and 4, which leads to the production of an inactive form of the kinase. In stress conditions, these two introns are spliced to produce the catalytically active form of CLK1. The *CLK1* mRNA with retained introns reappears during the recovery period. In *SATIII* KD cells, Ninomiya et al. observed a delayed retention of introns 3 and 4, which were then detected later during the recovery period compared to the control. This result together with the analysis of the splicing patterns of other transcripts, regulated by CLK1, revealed the role of SATIII, and more generally of nSBs, in promoting widespread intron retention during the recovery period after HS.

In our attempt to uncover the function of YTHDC1 at nSBs, and based on its potential link to *SATIII* functions, we have examined the splicing of *CLK1* in YTHDC1 KD cells. For this WT and YTHDC1 KD HeLa cells were subjected to HS followed by a recovery period of three or six hours at 37°C. We examined the intron retention at *CLK1* pre-mRNA by RT-PCR. The primers were chosen to study the effect of the YTHDC1 KD on *CLK1* pre-mRNA splicing. One of them is situated in intron 3 and allowed us to detect only the transcripts containing this intron while the second one is in exon 6 (Fig.45A, scheme on right). Using these oligonucleotides we are able to estimate the expression of the inactive form, retaining introns 3 and 4, and the splicing of intron 4.

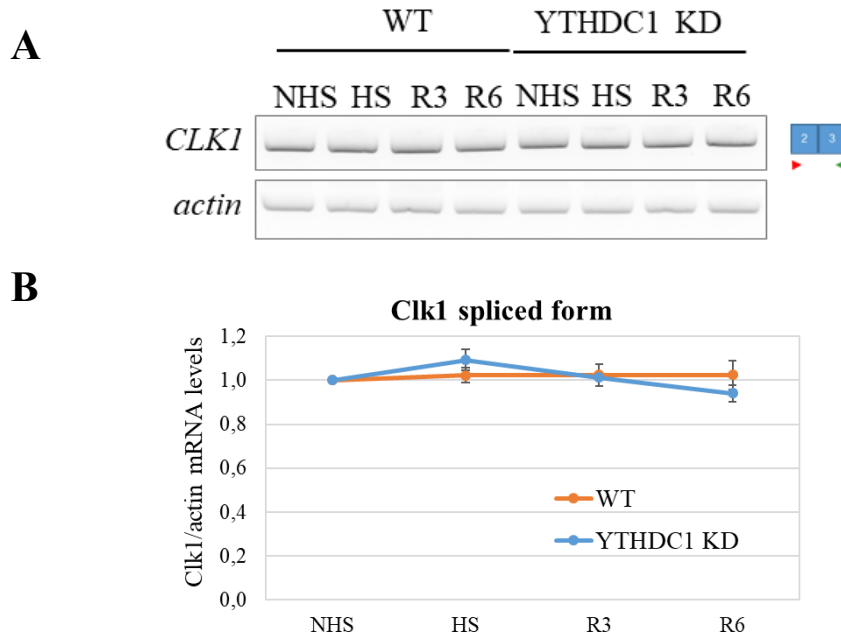
The RT-PCR results are presented in Fig.45A. Here, only the band of interest, corresponding to the form retaining introns 3 and 4 is shown and *actin* was used as loading control. In Fig.45B is shown the quantification of *CLK1* levels compared to *actin*, obtained from three independent experiments. In both WT and YTHDC1 KD HeLa cells we detect the already described alternative splicing of intron 3 and 4 following stress induction. The quantification indicates a strong decrease, of nearly 80%, of the unspliced *CLK1* form immediately after the thermal shock. However, in WT cells after three hours of recovery this unspliced form is restored to the same levels as in NHS. Remarkably, in YTHDC1 KD cells at R3h we notice only a slight increase, of approximately two times, of the unspliced form compared to HS. Even after six hours

of recovery the levels of *CLK1* mRNA with retained introns are not restored as in the control and are still diminished by nearly 50% of the NHS levels. This result uncovers a major splicing defect in YTHDC1 KD cells and suggests that as for the *SATIII* lncRNAs, YTHDC1 promotes intron retention in response to stress.



**Figure 45: YTHDC1 promotes *CLK1* mRNA intron retention during the recovery period following HS.** **A.** Analysis of the splicing pattern of *CLK1* mRNA by semi-quantitative RT-PCR on total RNA extract in HeLa WT and YTHDC1 KD cells at NHS, HS, R3h and R6h. Arrows on the scheme indicate the position of the used primers. Shown is only the band of interest, retaining the two introns. *Actin* is used as loading control. **B.** Quantification of *CLK1* mRNA form with retained introns 3 and 4, normalized to *actin* and to the NHS *CLK1* mRNA levels in WT and YTHDC1 KD HeLa cells at NHS, HS, R3h and R6h. Mean values from three independent experiments.

Additionally we wished to verify that YTHDC1 KD does not alter the levels of *CLK1* mRNA and that YTHDC1 does not participate more generally in the regulation of the constitutive *CLK1* mRNA splicing. Therefore, we have performed RT-PCR experiments, this time targeting a region of the mRNA, subjected to constitutive splicing. This control indicated that YTHDC1 KD does not affect *CLK1* mRNAs levels, compared to the WT, and that YTHDC1 does not participate in the control of the constitutive *CLK1* mRNA splicing (Fig.46).



**Figure 46: YTHDC1 does not affect *CLK1* mRNA levels and constitutive splicing. A.** Analysis of *CLK1* mRNA by semi-quantitative RT-PCR on total RNA extract in HeLa WT and YTHDC1 KD cells at NHS, HS, R3h and R6h. Arrows on the scheme indicate the position of the used primers. *Actin* is used as loading control. **B.** Quantification of *CLK1* E2-E3 spliced mRNA, normalized to *actin* and to the NHS *CLK1* levels in WT and YTHDC1 KD HeLa cells at NHS, HS, R3h and R6h. Mean values from three independent experiments.

Altogether, these results demonstrate that YTHDC1 is essential for the alternative splicing of *CLK1* mRNA in response to HS. The protein promotes intron retention and its KD strongly decreases the restoration of the levels of unspliced *CLK1* at R3h and R6h. Since CLK1 is the key regulator of alternative splicing in the context stress, our results suggest an even more global and central implication of YTHDC1 in the control of splicing events at the level of nSBs. In line with this, we are currently investigating the effect of YTHDC1 KD on the splicing pattern of other transcripts, known to be regulated through the axis *SATIII* (nSBs)-CLK1 during the recovery period after HS.

## **Chapter II**

# **YTHDC1 and gene expression regulation in response to HS**



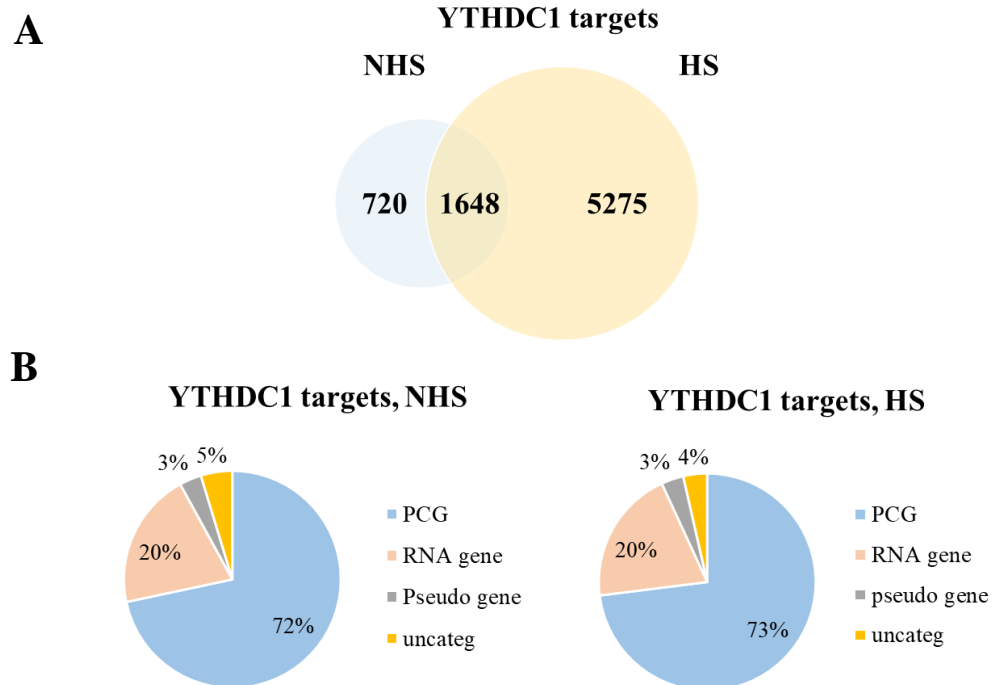
## **II.I HS induces YTHDC1 recruitment to heat responsive genes**

The adaptive cellular stress response is accompanied by a massive genome reprogramming characterized by the activation of a specific set of coding and non-coding genes, and the simultaneous global downregulation of the transcribed genome.

In the first part of the project we have focused on the relocalization and role of YTHDC1 at nSBs, representing part of the non-coding transcriptional response to stress. However, as part of my PhD work, we wished to understand in a more global manner the role of the protein in the cellular response to heat stress.

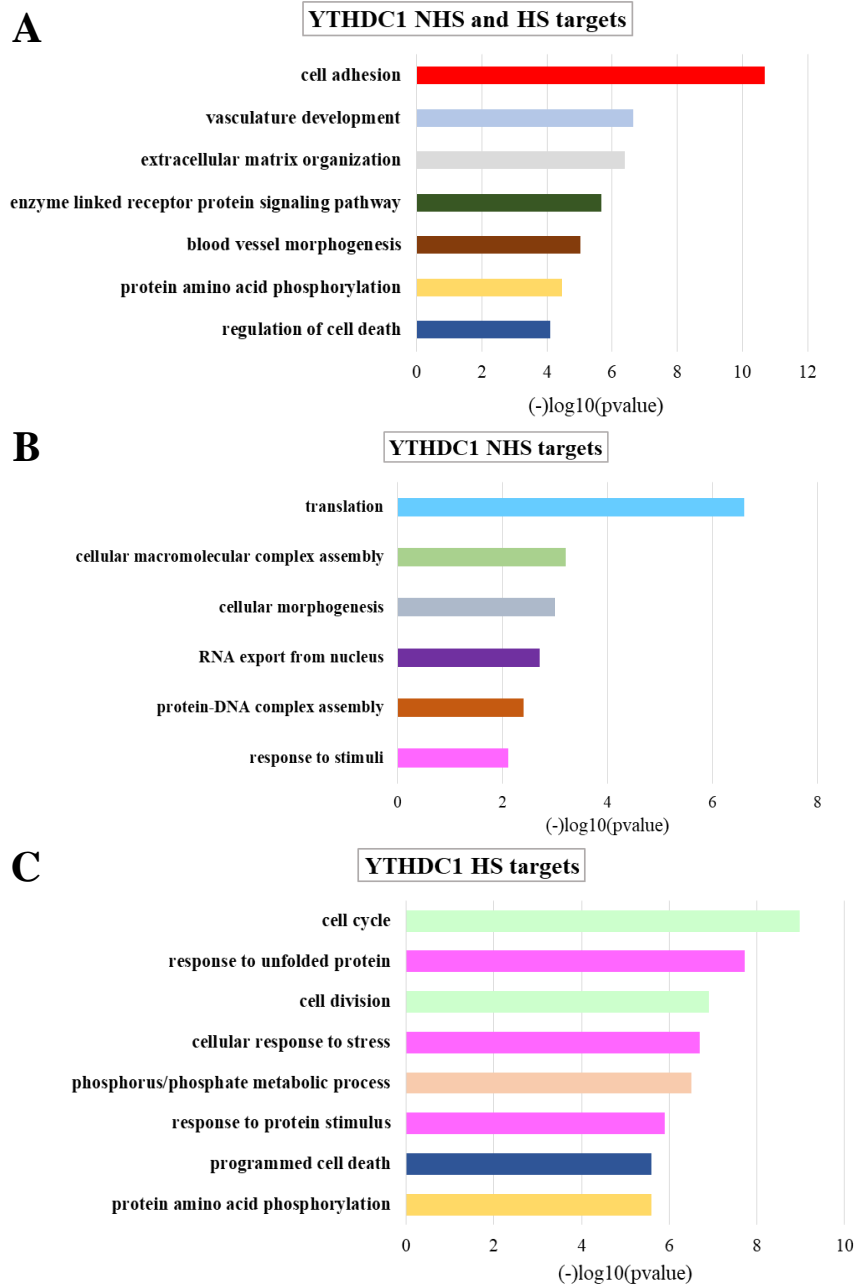
To address this question, and investigate the potential central role of YTHDC1 in the conserved HSR pathway, we have defined its genomic targets in normal and stress conditions by performing genome-wide Chromatin Immunoprecipitation sequencing (ChIP-seq) experiments in collaboration with Rosemary Kiernan's lab (IGH, Montpellier). The ChIP-seq approach was applied for these studies since on one hand YTHDC1 is involved in major co-transcriptional events such as splicing and 3'-end processing and on the other hand the m6A RNA mark is deposited also co-transcriptionally. The chromatin immunoprecipitation was conducted in HeLa cells under normal growth conditions or following HS, using an antibody specific of YTHDC1 (Materials and Methods VII.1).

The analysis of the ChIP-Seq data show that in absence of stress YTHDC1 associates with 2368 different genes (approximate number which can slightly vary depending on the analyses parameters). Remarkably, following HS, YTHDC1 is massively recruited to chromatin and interacts with 6923 different genes, almost three times the number of targets detected in NHS (Fig. 47A). Regarding the nature of the identified targets we observe that the repartition of protein-coding genes (PCG), RNA genes (or non-coding genes), pseudogenes and uncategorized genes does not vary between the two studied conditions (NHS and HS). In both cases ~70% of the total number of YTHDC1 targets are coding genes while 20% are non-coding (Fig. 47B). This initial analyses revealed for the first time the striking recruitment of YTHDC1 to chromatin in response to cellular stress and prompt us to investigate the importance and the potential functional implication of this relocation.



**Figure 47: Heat stress induces massive recruitment of YTHDC1 to chromatin. A.** Number of YTHDC1 targets identified by ChIP-seq analysis in HeLa cells at NHS and HS. YTHDC1 interacts with 2368 genes in NHS, 6923 genes in HS. 1648 of the targets are common between the two conditions. **B.** Repartition of protein-coding genes (PCG), RNA genes, pseudogenes and uncategorized genes (uncatteg) YTHDC1 targets in NHS and HS. Indicated is the percentage of the total number of targets in each condition.

To go further with the study of YTHDC1 targets, we have performed gene ontology analysis of the protein-coding YTHDC1-bound genes in order to determine their corresponding molecular and cellular functions. We have separately analyzed the genes found enriched in both NHS and HS conditions (1648 genes), and the once found enriched specifically in absence of stress (720 genes) and following HS (5275 genes) (Fig.48).



**Figure 48: Biological processes of YTHDC1 targets common between NHS and HS, and specific for NHS and HS.** Biological processes associated with YTHDC1 targets identified in both NHS and HS (A), specifically in NHS (B) and specifically in HS (C). Shown are representative functional categories of genes with highest  $(-\log_{10}(\text{pvalue}))$  for each condition ( $(-\log_{10}(\text{pvalue})) > 4$  for NHS-HS;  $(-\log_{10}(\text{pvalue})) > 2$  for HS and  $(-\log_{10}(\text{pvalue})) > 5$  for HS). The Gene Ontology (GO) analyses are performed using DAVID (Database for Annotation, Visualization and Integrated Discovery).



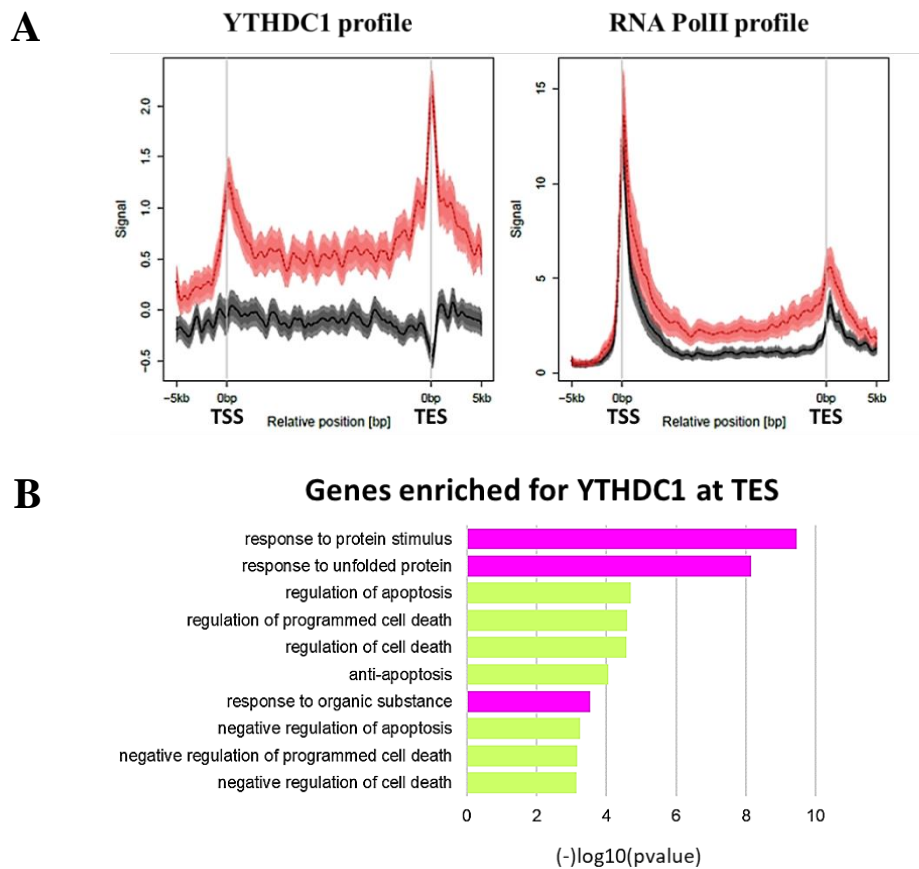
This study demonstrated that the genes enriched in YTHDC1, common between the NHS and HS conditions, are involved in the regulation of cell adhesion (Fig.48A). In this group of targets we find also genes with various roles such as vasculature development, extracellular matrix organization, regulation of cell death and others.

The genes targeted by YTHDC1 only in normal growth conditions are involved in translation regulation, macromolecular complex assembly, cellular morphogenesis, chromatin organization, nuclear export and response to stimulus (Fig.48B). These results point out that in normal conditions the protein could be implicated in the gene expression regulation of transcripts needed for the execution of crucial molecular and cellular processes.

The numerous targets identified in HS could be classified in twenty distinct categories associated with a broad range of functions. Importantly, the most significantly enriched biological processes following stress induction turned out to be cell cycle, cell division and cellular stress response (Fig.48C). Here, we observe that in addition to the targets implicated in stress response, YTHDC1 is recruited particularly to genes related to the response to unfolded proteins and the general cellular protein stimulus. Overall, the performed analyses indicated that in the context of stress YTHDC1 interacts with genes related to various cellular processes, essential for the homeostasis preservation, cell survival and recovery.

Next, we wished to have a more precise and comprehensive view of the enrichment of YTHDC1 along its targets. In order to do so, the YTHDC1-bound genes were classified according to the enrichment of YTHDC1 at the Transcription Start Site (TSS), the Gene Body (GB) and the Transcription End Site (TES) in HS compared to NHS. In each category, the 500 most strongly YTHDC1-enriched genes have been selected and then used to establish the average profile of YTHDC1 and RNA Pol II along them in NHS and HS conditions. The RNA Pol II ChIP-seq experiments in NHS and HS have been performed by our collaborators- Rosemary Kiernan's lab (IGH, Montpellier). The functions and the transcriptional pattern of the genes, bound by YTHDC1 most significantly at the TSS and the GB are currently under investigation and suggest important roles of YTHDC1 in the modulation of the stress-related downregulation of genes. Here, in relation to my PhD work, only the genes enriched for YTHDC1 all along (TES, GB, TSS), but most strongly at the TES region will be discussed.

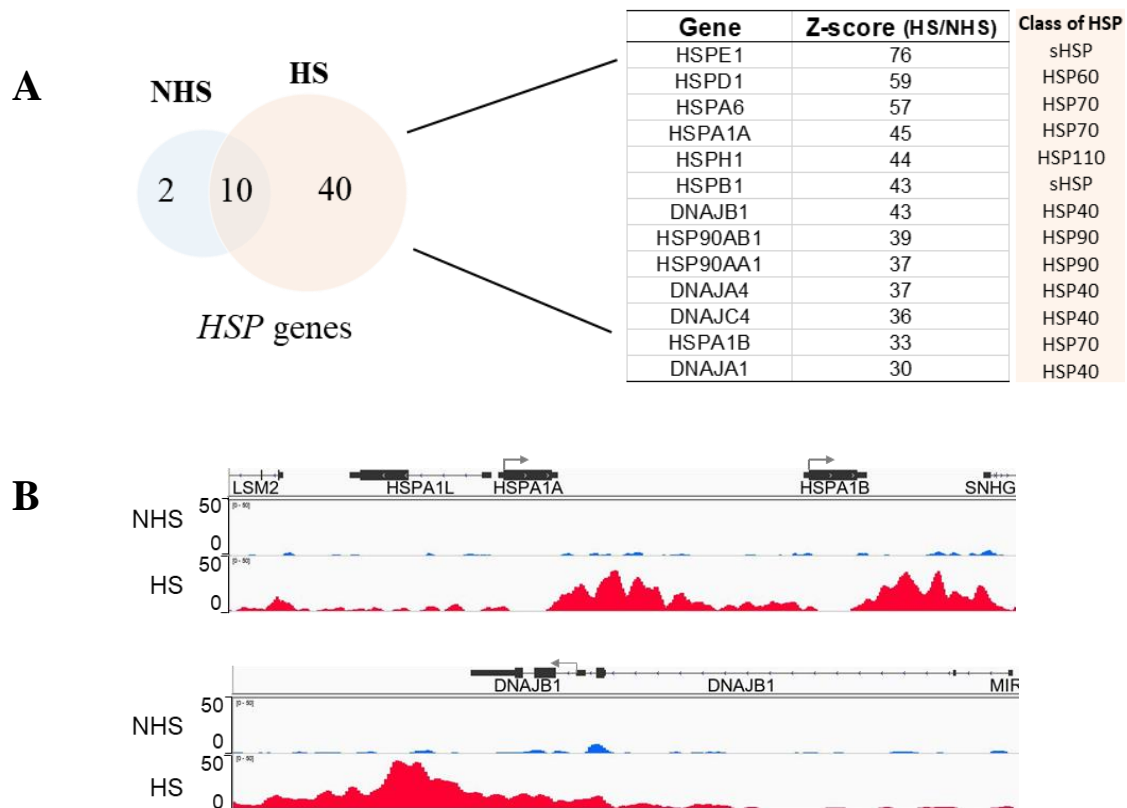
The profiles of YTHDC1 and RNA Pol II at the genes with strong enrichment of YTHDC1 at the TES in response to HS versus NHS are presented in Fig.49A. We notice that in absence of stress YTHDC1 is not specifically enriched at these genes (dark line), while following HS its levels increase mostly at the TES but also at the GB and TSS regions (red line). The strong RNA Pol II signal at the TSS and TES regions of these genes indicates that they are actively transcribed during the stress response (red line).



**Figure 49 : Average profiles of YTHDC1 and RNA Pol II over 500 genes at which YTHDC1 is most strongly enriched at the Transcription End Site during HS compared to NHS. Biological processes linked to these genes. A.** Profiles of the signal of YTHDC1 and RNA Pol II along the 500 genes at which YTHDC1 is most strongly enriched at the Transcription End Site (TES) at HS versus NHS. Black line- enrichment during NHS, red line- enrichment during HS. The Transcription Start Site (TSS) is indicated as well. **B.** Biological processes associated to the 500 most significantly enriched genes among the one with strongest YTHDC1 enrichment at the TES. The categories are presented according to the significance of their enrichment ( $(-\log_{10}(\text{pvalue}))$ ). The Gene Ontology (GO) analyses are performed using DAVID (Database for Annotation, Visualization and Integrated Discovery).

We have next examined the biological processes related to these genes. Importantly, they turned out to be involved mainly in the response to protein stimulus and to unfolded proteins, as well as in the regulation of apoptosis and cell death. Therefore, this study unraveled that following HS the levels of YTHDC1 increase strongly at the TES region of actively transcribed genes, essential for the implementation of the cellular stress response.

The very strong enrichment of genes involved in stress response within YTHDC1 targets prompt our interest and suggested, in line with our investigation, a role of the nuclear m6A reader in the regulation of the heat response pathway. A more detailed study of these targets revealed that many *HSP*-coding genes (Heat Shock Protein-coding genes) are bound by YTHDC1 upon HS. In NHS, YTHDC1 interacts with 12 *HSP* genes, while after HS this number rises to 50 different *HSP*-coding genes, including the 10 targets found in the NHS condition (Fig.50A).



**Figure 50: YTHDC1 is recruited to many *HSP*-coding genes upon HS.** **A.** Repartition of *HSP*-coding genes bound by YTHDC1 in NHS and HS. On the right are presented the *HSP* genes with the highest Z-score (enrichment in HS versus NHS, normalized to the standard deviation) and their class. **B.** IGV (Integrative Genomics Viewer) image showing the enrichment of YTHDC1 at *HSPA1A* and *HSPA1B* (upper panel), and *DNAJB1* (lower panel) in NHS and HS. Arrows indicate the orientation of the genes.

The members of the broadly conserved family of HSPs are classified according to their molecular weight and roles into five categories: HSP110s, HSP90s, HSP70s, HSP40s, HSP60s and small Heat Shock Proteins (sHSPs). We discovered that upon HS YTHDC1 associates with *HSP* genes that are part of all of these main classes. The largest proportion, thirty genes, belong to the HSP70s and their cofactors- the HSP40s. The most strongly enriched *HSP* genes, according to their Z-score (enrichment in HS compared to NHS and normalized to the standard deviation), are represented in Fig.50A along with their corresponding class.

Furthermore, we have examined the enrichment of YTHDC1 along *HSP*-coding genes. As an example, the levels of YTHDC1 at three HSP genes- *HSPA1A* and *HSPA1B* (HSP70s) and *DNAJB1* (HSP40s) are shown in Fig.50B. We observe that the levels of YTHDC1 increase very strongly at these genomic loci under stress compared to normal growth conditions. Notably, we detect as well that the protein associates mainly with the 3'UTR region of these *HSP* genes.

The identification of the genomic targets of YTHDC1 in normal and stress conditions allowed us to discover that the protein is massively recruited to chromatin following HS and associates with genes involved in various functions. Through the study of the enrichment of YTHDC1 along its targets, we found that upon HS the group of genes with strong levels of YTHDC1 at the TES contains many active genes involved in the regulation of stress response. Among them, in particular YTHDC1 turned out to interact with various *HSP*-coding genes.

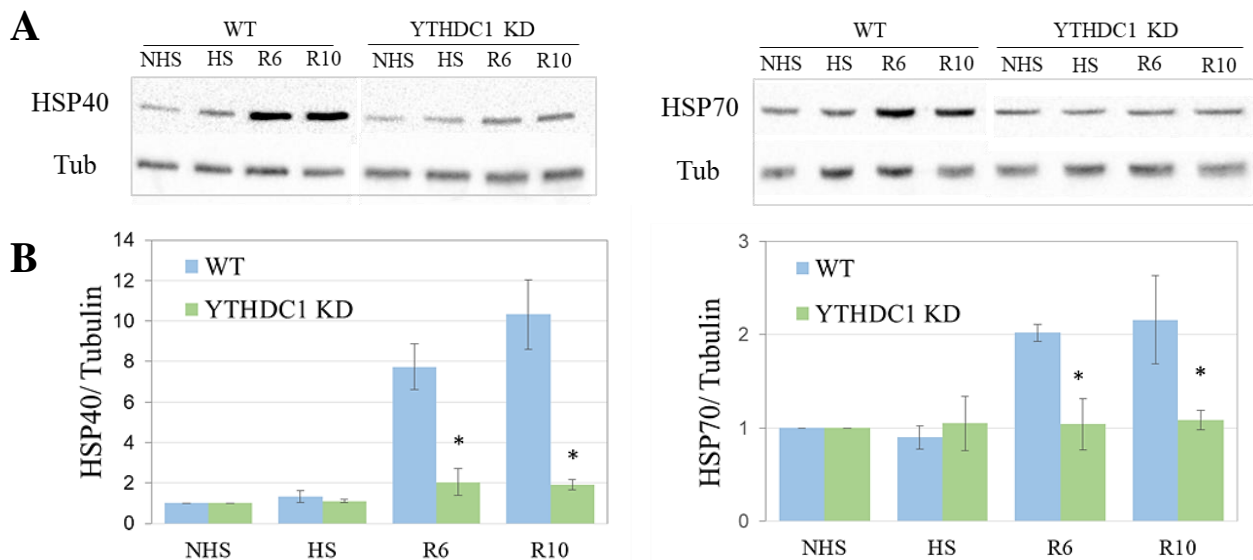
## **II.II YTHDC1 regulates HSPs expression**

The preferential association of YTHDC1 to *HSP* genes caught our interest since recently the distribution and abundance of the RNA modification m6A have been related to the expression of HSPs, in particular HSP70 (discussed in Introduction II.II.2). The levels of m6A were found to increase at the 5'UTR region of heat-responsive transcripts and were proposed to promote their cap-independent translation (Zhou et al., 2015; Meyer et al., 2015). In addition, it was suggested that higher m6A levels at the 3'UTR mark the HS-induced transcripts for degradation following stress (Knuckles et al., 2017). However, the mechanisms and molecular players required for this gene expression regulation involving m6A are still poorly defined. We hypothesized that by recognizing m6A on *HSP* mRNAs, YTHDC1 could be a major molecular actor of m6A-based gene

regulation upon HS. Below, we tested this hypothesis by examining the importance of YTHDC1 on the expression of HSPs.

For our studies we chose to work on three representative *HSP* genes- *HSPA1A*, *HSPA1B* (HSP70) and *DNAJB1* (HSP40), which are enriched in the ChIP-seq analysis of YTHDC1 after HS and present strong levels of YTHDC1 specifically at their 3'UTR following stress (c.f. Fig. 49B). In the next parts, if not indicated otherwise, these HSPs isoforms will be referred to as HSP70 and HSP40. Both of these proteins are inducible. They are expressed at a low levels in normal conditions to satisfy the constitutively low demand in protein folding and are strongly induced after HS in order to prevent protein misfolding and aggregation (Hata et al., 1998).

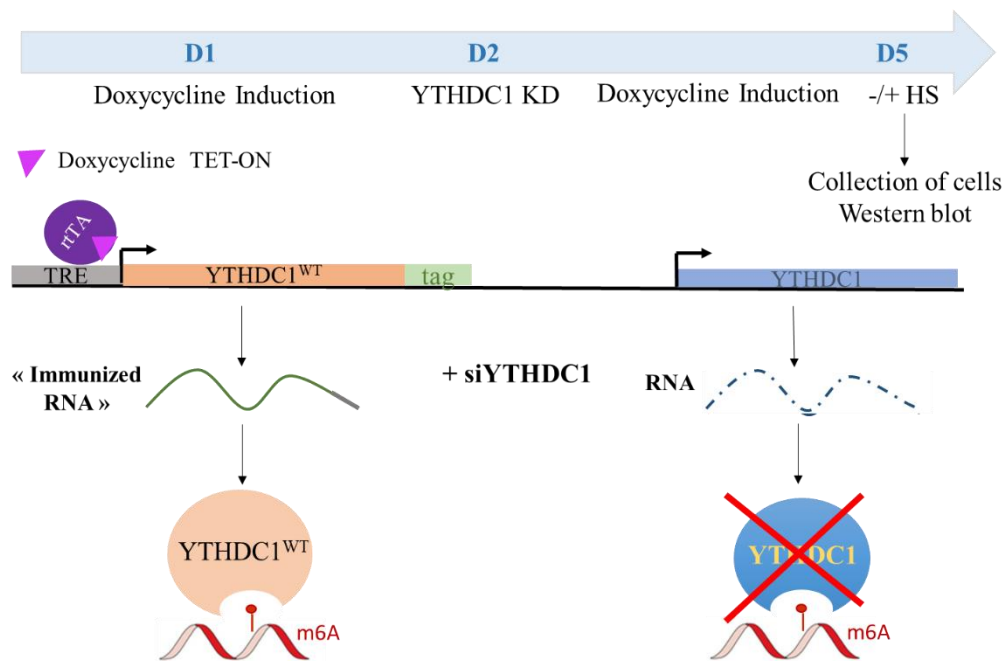
We have first examined the expression of HSP70 and HSP40 in WT and YTHDC1 KD HeLa cells using western blot analysis. Tubulin was used as loading control in these experiments since its expression is not affected by the HS treatment. As shown in Fig. 51A in WT cells, as expected, we detect a strong increase of the protein levels of both HSP40 and HSP70 during the recovery period, at R6h and R10h. In YTHDC1 KD cells at the NHS and HS time points the protein levels stay comparable to the control. Strikingly, in these cells at R6h and R10h we no longer observe the characteristic induction of the expression of the two HSPs.



**Figure 51: YTHDC1 KD impairs HSP40 and HSP70 induction in response to heat stress.** **A.** Western blot analysis of HSP40 and HSP70 on whole-cell extract from WT and YTHDC1 KD HeLa cells at NHS, HS, R6h and R10h. Tubulin (Tub) is used as loading control. **B.** Quantification graphs of the signals of HSP40 and HSP70 compared to tubulin. Mean values from three independent experiments. P-value was calculated using a two-tailed Student's t-test. \*:  $p < 0.05$ .

This complete impairment of HSPs induction after KD of YTHDC1 is further confirmed by the quantification graphs, Fig.51B. HSP40 presents up to 6 times fold-increase at R10h compared to NHS and HSP70 a 2 fold-increase at R6h and R10h compared to the NHS. In YTHDC1 KD cells this increase is abolished and no induction of the protein levels of the two studied HSPs is detected. This result indicates that the protein is essential for the expression of HSPs in response to stress.

Following this finding, we wished to confirm that the observed effect on the induction of HSPs is indeed due to the performed KD of YTHDC1. In order to do this we have carried out rescue experiments in the stable inducible HeLa cell line, coding for the ectopic WT YTHDC1 (Materials and methods I.III.1). The experimental setup is schematically represented in Fig.52.

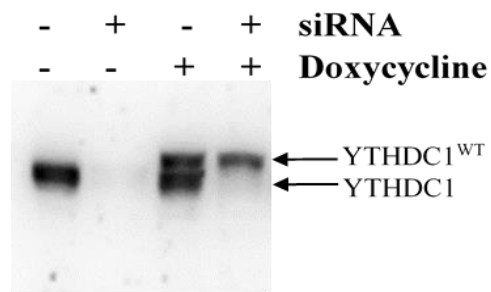


**Figure 52: Rescue experiments in the inducible stable HeLa cell lines.** On day 1 (D1) the expression of the ectopic YTHDC1<sup>WT</sup> is induced using doxycycline. On day 2 (D2) a KD of the endogenous YTHDC1 is performed, the mRNA of YTHDC1<sup>WT</sup> is not degraded and the protein is expressed, while endogenous YTHDC1 mRNA is targeted and degraded. Between D1 and D5 the doxycycline induction is maintained. HS is induced at D5 in order to study HSPs expression.

In this cell line the construct introduced by CRISPR-Cas9 is under the control of a doxycycline responsive promoter, TET-ON system. Briefly, using doxycycline we induce the

expression of the ectopic WT YTHDC1 and perform in parallel a transient KD of the endogenous protein. Silent mutations were introduced in the sequence coding for the ectopic YTHDC1 (YTHDC1imm-Strep-FLAG, for shortly here YTHDC1<sup>WT</sup>) to abolish the recognition by the siRNAs used for the KD of the endogenous YTHDC1 (Materials and Methods III.II). The induction of the ectopic protein is initiated one day prior to the KD to allow its accumulation before the loss of the endogenous protein and is maintained until the endpoint of the experiment. At the last day, we perform a heat shock treatment, collect the cells and analyze the expression of HSPs by western blot.

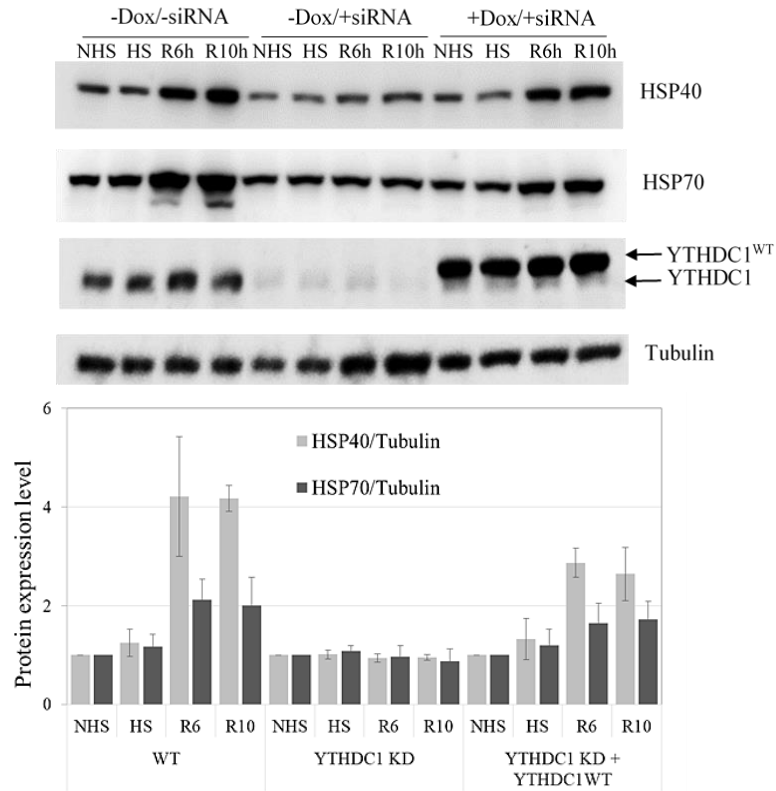
Prior to the studies using this approach, we validated the chosen conditions for KD of the endogenous YTHDC1 and the expression of YTHDC1<sup>WT</sup>. As shown in Fig.52, 72h after the siRNA transfection, the KD results in strong decrease of the endogenous YTHDC1 protein levels compared to the WT. Using western blot we are able to analyze in parallel the two forms of the protein since YTHDC1<sup>WT</sup> contains a C-terminal tag and is slightly longer than YTHDC1 (visible in the condition -siRNA and +doxycycline, Fig.53). With this optimized protocol, at the end of the assay YTHDC1<sup>WT</sup> is expressed in the KD cells at levels comparable to this of the endogenous protein in the WT cells.



**Figure 53: Western blot analysis used to define the conditions for rescue experiments in the stable inducible cell lines.** Immunodetection of YTHDC1 in the stable inducible cell lines in presence (+) or absence (-) of the specific siRNAs against YTHDC1 or Doxycycline. Arrows indicate the two forms of YTHDC1- ectopically expressed or endogenous.

For the study of the expression of HSP40 and HSP70 in these conditions, the HS treatment was followed by 6 and 10 hours of recovery. The western blot analyses of YTHDC1 show that the stress induction does not affect the expression of the ectopic protein in the KD cells (Fig.54A). In this cell line, as previously observed in HeLa cells, the induction of the two HSPs is abolished after the KD of YTHDC1.

Importantly, when YTHDC1<sup>WT</sup> is expressed, this pattern is no longer present and we detect the induction of HSP40 and HSP70 at R6h and R10h (Fig.54A). The quantification graphs in Fig.54B indicate that using these rescue experiments we are able to restore in YTHDC1 KD cells nearly 70% of the expression of HSP40 and HSP70 found in WT cells during the recovery period. Therefore, these results confirmed a specific role of YTHDC1 in the expression of HSPs in response to stress.



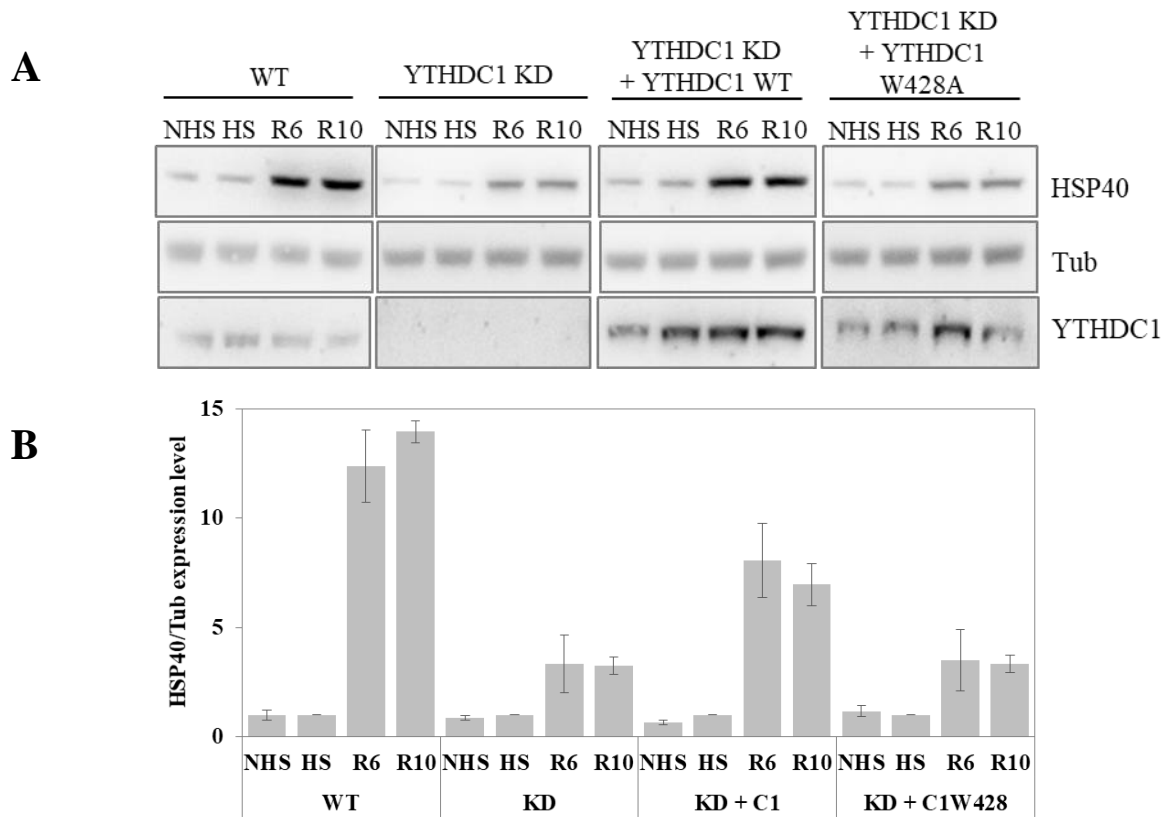
**Figure 54: Rescue experiments of HSP40 and HSP70 in the stable inducible HeLa cell lines coding for WT YTHDC1.** **A.** Western blot analyses of HSP40, HSP70 and YTHDC1 levels in the HeLa stable inducible cell lines coding for WT YTHDC1 in presence (+) or absence (-) of the specific siRNAs against endogenous YTHDC1, or doxycycline (Dox) at NHS, HS, R6h and R10h. Tubulin is used as loading control. **B.** Quantification graphs showing the induction of HSP40 and HSP70 compared to the NHS in the same conditions as in A. Presented are mean values from three independent experiments.

Overall, this work tightly links the relocation of YTHDC1 to heat-responsive genes to their induction of expression following stress.



### II.III YTHDC1 regulation of HSPs could be m6A-dependent

We are currently investigating whether the regulation of the expression of HSP40 and HSP70 by YTHDC1 is dependent on the recognition of m6A residues. For this we carry out rescue experiments using transiently transfected YTHDC1 KD HeLa cells, expressing HA-YTHDC1<sup>W428A</sup> (described in Fig.40, Results I.I.1.). The western blot analyses presented in Fig.55A were performed in WT, YTHDC1 KD and YTHDC1 KD cells transfected with HA-YTHDC1<sup>WT</sup> or HA-YTHDC1<sup>W428A</sup> at NHS, HS, R6h and R10h.



**Figure 55: YTHDC1 regulation of HSPs expression could be m6A-dependent.** **A.** Western blot analyses of HSP40 and YTHDC1 from whole-cell extracts in HeLa: WT, YTHDC1 KD or YTHDC1 KD cells transfected with HA-YTHDC1<sup>WT</sup> or HA-YTHDC1<sup>mut</sup> at NHS, HS, R6h, R10h. Tub- tubulin is used loading control. **B.** Quantification graphs of the expression of HSP40 compared to NHS in the same conditions as in A. Levels are normalized to tubulin  $\beta$  signal. Mean values from three independent experiment.

As detected, the expression of the two ectopic proteins was not affected by the siRNAs, nevertheless, the levels of the mutant were weaker than those of the WT protein (Fig.55A). The examination of HSP40 levels indicate that the induction of this chaperone is partially restored in

KD cells expressing HA-YTHDC1<sup>WT</sup> compared to the KD cells. Notably, a rescue was not detected after the reintroduction of the mutated YTHDC1 in KD cells, suggesting an m6A-dependent mechanism of regulation of HSPs induction. However, we cannot exclude the possibility that the protein levels of HSP40 are lower in YTHDC1 KD cells transfected with HA-YTHDC1<sup>W428A</sup> compared to HA-YTHDC1<sup>WT</sup>, due to the lower expression and/or instability of the mutant.

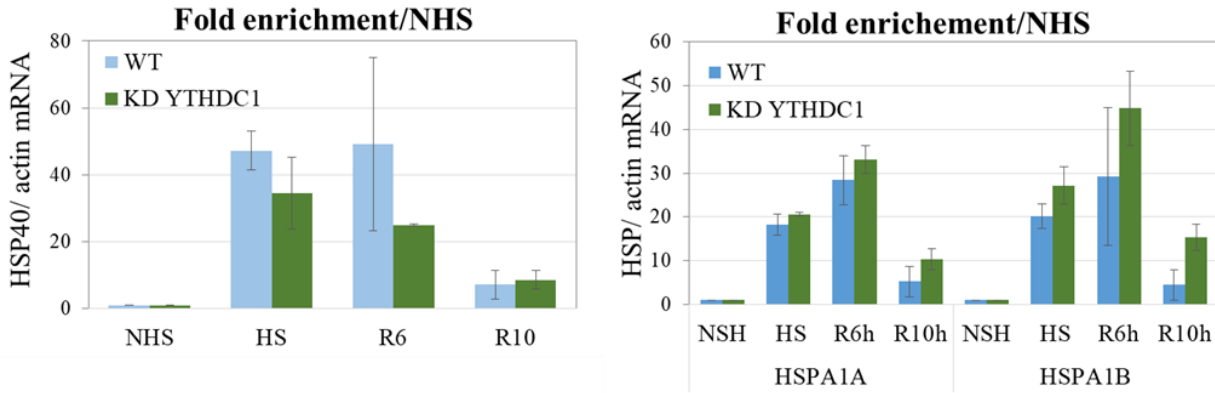
These results indicate that the capacity of YTHDC1 to bind m6A might be necessary for HSPs induction. Therefore the stress-related expression of HSPs is tightly correlated to the presence of this epitranscriptomic mark.

#### **II.IV Role of YTHDC1 in *HSP* mRNA processing and export**

Following on the discovery that YTHDC1 is needed for the stress-induced expression of chaperones we wished to decipher molecular mechanisms involved in this regulation.

##### *Effect of YTHDC1 on the production of HSP mRNAs*

First we wished to determine whether YTHDC1 could interfere directly with the production of *HSP40* and *HSP70* mRNAs during the recovery period. To address this, we conducted RT-qPCR analyses to detect the mRNA levels of the two HSPs in WT and YTHDC1 KD HeLa cells in NHS, HS, R6h and R10h. Two of the major *HSP70* isoforms are *HSPA1A* and *HSPA1B*, which are identical in their sequence at 95%. The antibody used previously for our western blot analyses detects both of them. For our RT-qPCR experiments we used primers specific for the mRNA of one or the other isoform (primers located in the 3'UTR of each mRNA). *Actin* mRNA was used as control since its production is not affected by the HS treatment. For both *HSP* mRNAs the results are presented as fold enrichment compared to the NHS condition in WT or in YTHDC1 KD cells. As shown in Fig.56, in WT cells the mRNA levels of *HSP40* and *HSP70* (*HSPA1A* and *HSPA1B*) increase rapidly after the stress induction, reach a maximum at R6h and decrease almost back to the NHS level at R10h. In YTHDC1 KD cells we do not notice any significant difference of the levels compared to WT cells (Fig.56, green bars). Indeed, following the KD of YTHDC1 the mRNAs of *HSP40* and *HSP70* are strongly induced after the HS and during the first six hours of the recovery period and decrease back to normal at R10h like in the wild-type situation.



**Figure 56: YTHDC1 does not affect the levels of *HSP40* and *HSP70* mRNAs.** RT-qPCR experiments after total RNA extraction on HeLa WT and YTHDC1 KD cells at NHS, HS, R6h and R10h, using specific primers for *HSP40*, *HSPA1A*, *HSPA1B* and *actin* mRNAs. *Actin* is used as non-variable control. Graphs represent fold enrichment of *HSP40* (left), *HSPA1A* and *HSPA1B* (right) for each condition compared to the NHS in the corresponding cells (WT or KD) and normalized to *actin* mRNA levels. Mean values from three independent experiment.

This result indicates that YTHDC1 KD does not affect the increase of *HSP* mRNAs in response to stress. An intriguing outcome of this study is the fact that it highlights the uncoupling between HSPs RNA expression and protein expression. The RNA accumulation was detected immediately after the HS while our western blot analyses indicated that the mRNAs are translated later, between R6h and R10h (Fig.51). The dissociation of the two processes has been already mentioned in the literature, however up to date there is no clear explanation for this differential regulation. The delayed protein production probably reflects a specific control of mRNA-processing steps occurring between the period from HS induction to R6h.

Taken together, these observations led to the hypothesis that rather than the production of *HSP40* and *HSP70* mRNAs, the KD of YTHDC1 might alter one or several steps that further impact on their competency to be translated.

YTHDC1 participates in numerous mRNA-processing steps- transcription termination, alternative polyadenylation, splicing and nuclear export. All of these functions have been investigated to determine whether they are essential for the proper expression of *HSP40* and *HSP70* and the obtained results for each process are described in the next parts of this chapter.

### Effect of YTHDC1 on the transcription termination of HSPs

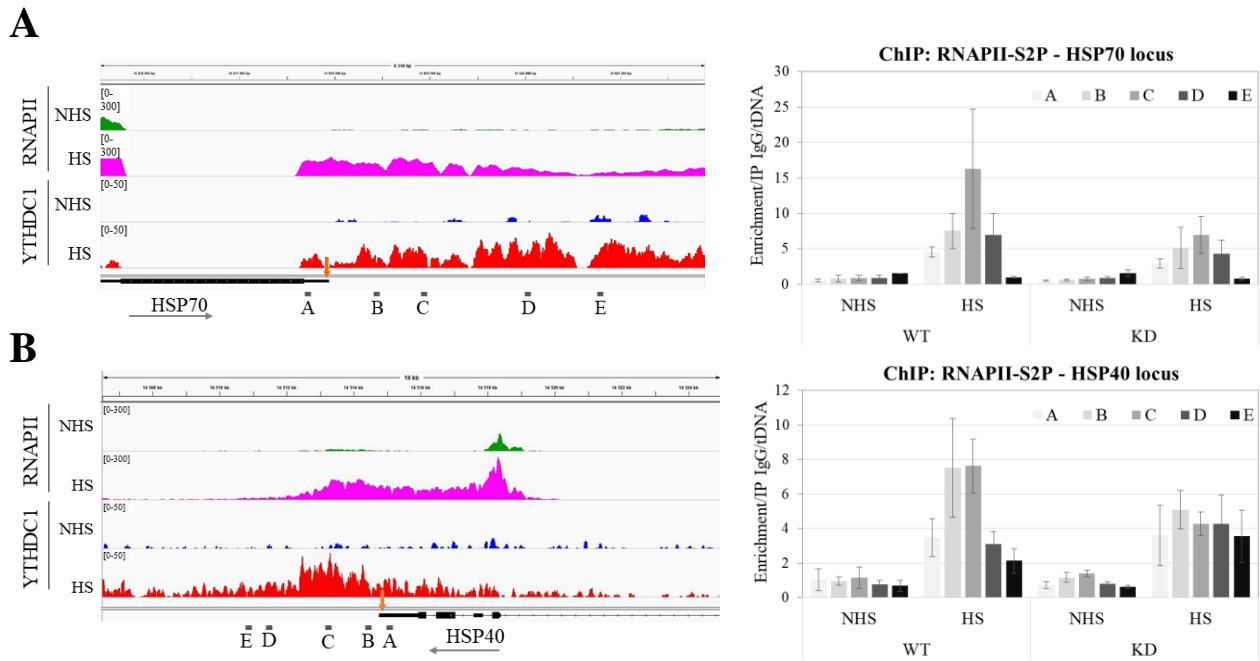
The tightly linked processes of pre-mRNA 3'-end cleavage and polyadenylation are also closely related to the transcriptional termination by the RNA Pol II. A PolyA Signal (PAS) and a correctly functioning cleavage and polyadenylation machinery are required for the 3'-end cleavage which subsequently releases the RNA Pol II and allows transcription termination. If the PAS is altered or incorrectly recognized, this abolishes the cleavage of the pre-mRNA and may lead to defects of the transcription termination which can have severe consequences. For example, in the case of protein-coding genes arranged in tandem failure of transcription termination gives rise to transcription readthrough which may lead to deregulation of gene expression through transcription interference.

When multiple PAS exist Alternative PolyAdenylation (APA) occurs. APA is a mechanism of production of RNAs with different 3'-end lengths. The 3' UTR contains target sites for miRNAs and RNA-binding proteins. As a result lengthening or shortening of 3'UTR may affect greatly the translation efficiency, stability, and localization of mRNA.

It has been described in the literature that YTHDC1 is important for the selection of PAS and is therefore needed for the correct 3'-end processing of mRNAs. The loss of the protein results in APA and modifies the 3'UTR length of hundreds of genes (Kasowitz et al., 2018).

Here we have used two different approaches to investigate if YTHDC1 regulates PAS and transcription termination of *HSP40* and *HSP70*. In the first one we have performed CHIP of the elongating form of RNA Polymerase II, phosphorylated at serine 2 (Pol II-S2P) and looked for its enrichment along the 3'-region of *HSP* genes in WT and YTHDC1 KD cells, in NHS and following HS. When transcription termination is occurring normally Pol II-S2P accumulates at the 3'UTR shortly before it is released. In Fig.57 is presented the previously described HS-induced strong enrichment of YTHDC1 at *HSP40* and *HSP70*, particularly at the 3'UTR region and downstream of it (gene orientation shown by the arrows) as well as the levels of RNA Pol II along these genes.

We have selected specific primers annotated A-E along the 3'UTR of *HSP40* and *HSP70* and downstream from it to perform qPCR analyses of the occupancy of Pol II-S2P at this level (Fig.57).



**Figure 57: Pol II-S2P localization and enrichment at the 3'UTR region of *HSP70* and *HSP40*.** Left panels: enrichment of YTHDC1 and RNA Pol II along *HSP70* (A) and *HSP40* (B) in HeLa cells in NHS and HS (results from YTHDC1 and RNA Pol II genome-wide ChIP-seq experiments). IGV software images. A-E: positions of the primers used for qPCR. Right panels: qPCR experiments performed using the primers A-E along the 3'UTR regions of *HSP70* (A) and *HSP40* (B) following ChIP of Pol II-S2P in WT and YTHDC1 KD cells at NHS and HS. The enrichment of RNA Pol II-S2P in each condition is normalized to an IP with irrelevant antibody and to tDNA content. Mean values from three independent experiments.

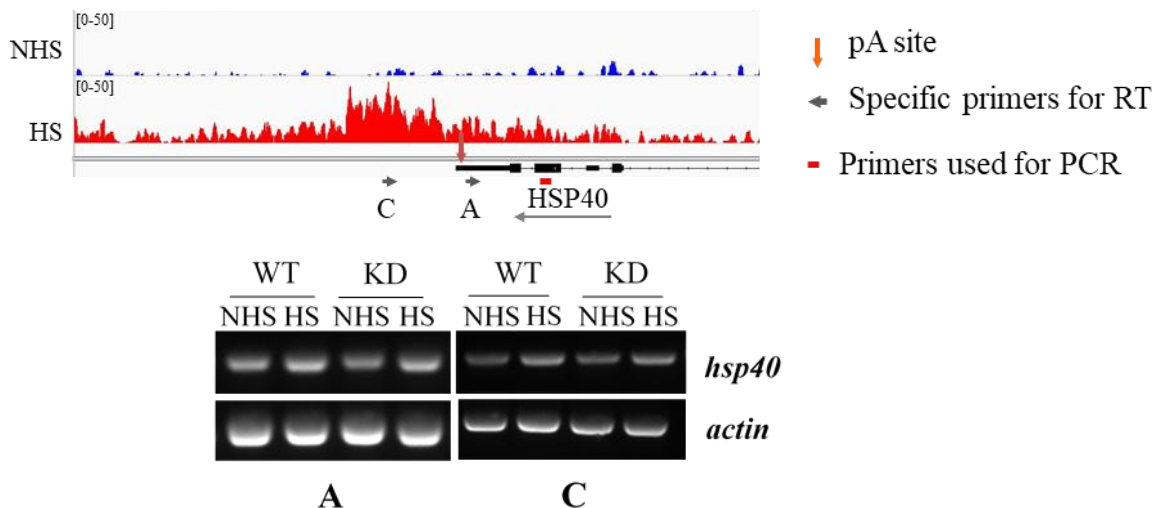
In absence of stress the Pol II-S2P enrichment in WT and YTHDC1 KD cells is low at the studied regions at both *HSP70* and *HSP40*, correlating with their low transcription levels in normal conditions (Fig.57A and B). As expected, following HS the Pol II-S2P is strongly enriched at the 3' end of the two *HSP* genes. In WT cells, at *HSP70*, Pol II-S2P is accumulating mostly around the C region (~1000 nt downstream from the predicted PAS) (Fig.57A, right panel). Further from this region its level drops down to the background level suggesting that an efficient transcription termination is taking place there. At *HSP40* we observe its accumulation around the region amplified with primers B (~500 nt downstream from the predicted PAS) and C (~1600 nt

downstream from the predicted PAS). This region corresponds also to the peak of enrichment of YTHDC1 for *HSP40* (c.f. Fig.57B, left panel).

The polymerase is enriched mostly at ~1000 nt downstream from the predicted PAS (region C) at *HSP70* in both KD and WT cells. Overall, in YTHDC1 KD cells following HS the enrichment of the elongating polymerase is weaker at the 3'UTR of both genes compared to the control. In addition, for *HSP40*, the peak of accumulation of RNA Pol II-S2P (B-C region) present in WT cells is no longer detected after the KD of YTHDC1 (Fig.57B, left panel). Collectively, these results are in agreement with the possibility that YTHDC1 is required for the appropriate transcriptional termination of heat stress responsive genes following HS.

The second approach which we have applied consisted in using RT-PCR to search for the production of readthrough transcripts at the locus of *HSP40* and *HSP70* following KD of YTHDC1. Reverse transcription was carried out with specific primers allowing us to detect if the transcription is not correctly ended and a readthrough is produced.

The positions of the primers used for RT and PCR of *HSP40* are indicated in Fig.58, upper panel. For the RT two primers were used- one ~200 nt upstream from the predicted PAS (A) and one ~1600 nt downstream from it (C), close to the peak of enrichment of YTHDC1.



**Figure 58: The KD of YTHDC1 does not induce production of readthrough transcripts at *HSP40*.**

Upper panel: IGV image of the enrichment of YTHDC1 at the 3'UTR region of *HSP40*, indicated are the positions of the used primers A and C. Lower panel- PCR of *HSP40* following RT performed using the primers A or C after total RNA extraction in WT and YTHDC1 KD HeLa cells at NHS and HS. *Actin* is used as loading control.

On the presented agarose gel we observe that the levels of the transcripts obtained with both RT primers increase upon HS in WT and in YTHDC1 KD cells. The results obtained using primer C indicate that in addition to the proximal PAS, another distal PAS is selected in all of the studied conditions. However, we can conclude that YTHDC1 KD did not affect the PAS selection and transcription termination of *HSP40* in normal, neither in HS conditions since no differential amplification between the WT and the KD cells was obtained.

The same experiments were performed also for *HSP70* and led to the same observation (data not shown). The KD of YTHDC1 did not interfere with the PAS selection and did not increase the formation of readthrough transcripts at *HSP70*.

These results suggest that YTHDC1 is needed for the proper transcription termination of *HSP40* and *HSP70*. However, in the chosen experimental conditions we can conclude that its KD does not produce a significant change in readthrough transcripts at the level of the two *HSP*-coding genes. A possible explanation, in agreement with both the RNA Pol II-S2P ChIP signals and the lack of readthrough transcripts accumulating in YTHDC1 KD cells, would be that YTHDC1 is required for the proper transcription termination of *HSP* genes but not the cleavage and polyadenylation of their corresponding pre-mRNAs. However, further experiments are needed to test this possibility.

#### *Effect of YTHDC1 on the splicing of HSP40 mRNA*

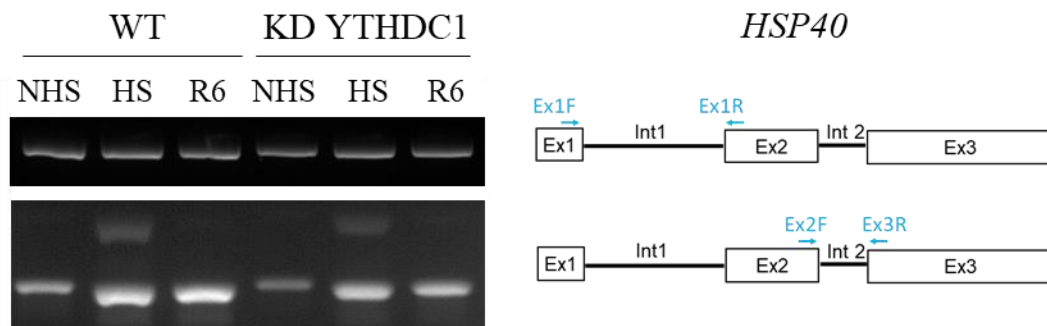
YTHDC1 is described as an important actor in the alternative splicing occurring at several hundreds of genes and some of the related mechanisms have been already elucidated (Xiao, et al.2016). Hence, we investigated whether this molecular function of the protein contributes to the expression control of *HSP* genes upon stress.

Since *HSP70* is an intronless mRNA, here we have focused our study on the splicing of *HSP40* in response to HS in WT and YTHDC1 KD cells. The structure of the transcript is schematically presented in Fig.59, right panel. *HSP40* is composed of 2 introns and 4 main isoforms can be produced- fully spliced, unspliced, one with the second intron and a version with only partially spliced intron 1.

To investigate the role of YTHDC1 in the splicing of *HSP40* we selected couples of primers, situated in the exons adjacent to the two introns (Ex1F and Ex1R; Ex2F and Ex3R), c.f. Fig.59, right panel.

Using these primers we observed that intron 1 seems to be spliced in all examined conditions. This form presents an increase in HS in both WT and YTHDC1 KD cells and no specific splicing changes are observed when YTHDC1 levels are decreased (PCR using Ex1F and Ex1R, first line, Fig.59). We noticed as well that the form containing the second intron (PCR using Ex2F and Ex3R, second line, Fig.59) is usually not present in normal conditions (upper band on the agarose gel). This intron is specifically retained in response to stress and spliced again later during the recovery period, at R6h in WT and YTHDC1 KD cells

Collectively, using this experimental setup, we can conclude that the splicing pattern of *HSP40* is not affected by the decreased YTHDC1 expression levels. Therefore, the proper expression of *HSP40* by YTHDC1 is unlikely to be related to an action of the protein at the level of splicing.



**Figure 59: YTHDC1 KD does not affect the splicing of *HSP40* mRNA.** On the right: schematic representation of *HSP40* mRNA, indicating the exons and introns and the positions of the primers used for PCR experiments. On the left: migration on agarose gel of the PCR products obtained with each couple of primers, after RT on total RNA extracted from HeLa WT or YTHDC1 KD cells in NHS, HS or R6h.

#### Effect of YTHDC1 on the export of *HSP40* and *HSP70* mRNAs

The correct export of mRNAs from nucleus to cytoplasm is not only a mechanism of intracellular transport and relocalization but is also associated to different quality control steps, which allow the mRNA to be efficiently translated in the cytoplasm. Thus, failure of the appropriate



export could result in the retention of mRNA in the nucleus or further on in an inefficient translation in the cytoplasm.

It has been recently described that YTHDC1 is implicated in the export of hundreds of mRNAs (Roundtree et al., 2017). In the search for its role in HSPs induction we investigated the possibility of an impact on the export of *HSP* mRNAs from the nucleus upon stress.

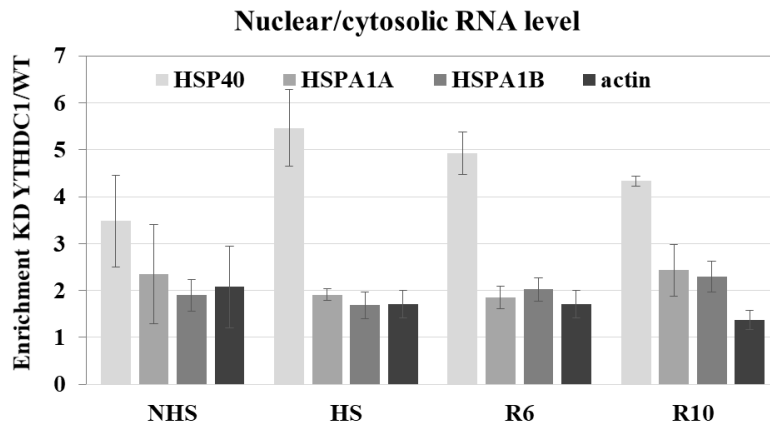
We studied the export of *HSPA1A* and *HSPA1B* (coding for two HSP70 isoforms), of *HSP40* and of *actin* mRNAs after HS and during the recovery period in WT and YTHDC1 KD HeLa cells. For these experiments we carried out subcellular fractionation and extracted separately the RNA from the cytosolic or nuclear fractions (Materials and Methods II). The mean values obtained from three independent experiments are presented as an enrichment of the nuclear to cytosolic mRNA levels of each transcript at NHS, HS, R6h and R10h in YTHDC1 KD cells compared to WT (Fig.60). Thus, here the enrichment indicates the increase in content of nuclear over cytoplasmic mRNA in the KD cells.

For the four examined transcripts we detect that the nuclear/cytoplasmic ratio is more important in the KD YTHDC1 cells than the WT cells in all studied conditions. This result indicates a partial nuclear retention of the mRNAs and corresponds to the known global role of YTHDC1 in export. However, regarding the stress-induced mRNAs some differences are noticeable compared to the levels of the “housekeeping” *actin* mRNA. *HSP40* seems to accumulate most strongly in the nucleus of KD cells even in absence of stress. Later on, after the stress-induced transcriptional activation, this accumulation persists. Notably, in HS the nuclear/cytoplasmic ratio of *HSP40* mRNAs in the KD cells is nearly 2 times more important than the one in NHS following KD. This elevated nuclear content is also observed during the recovery period.

*HSPA1A* and *HSPA1B* present less drastic differences between YTHDC1 KD and WT cells. Their nuclear to cytoplasmic content is comparable to the one of *actin* and is relatively stable in NHS, HS, R6h and R10h.

Altogether these results indicate that YTHDC1 is required for the proper export of the four studied transcripts and its KD leads to their partial nuclear retention independently from the HS treatment. In addition, YTHDC1 is further needed for the export of *HSP40* especially under stress

conditions. Therefore, we cannot assume that YTHDC1 has a general role in the export of stress-responsive mRNAs. Instead, it rather has a distinctive importance for the export of *HSP40* mRNAs especially following stress induction. The distinctive sensitivity to YTHDC1 KD may be linked to another step of *HSP40* mRNAs processing coupled to mRNA export and requiring the activity of YTHDC1. However, this will need to be tested.

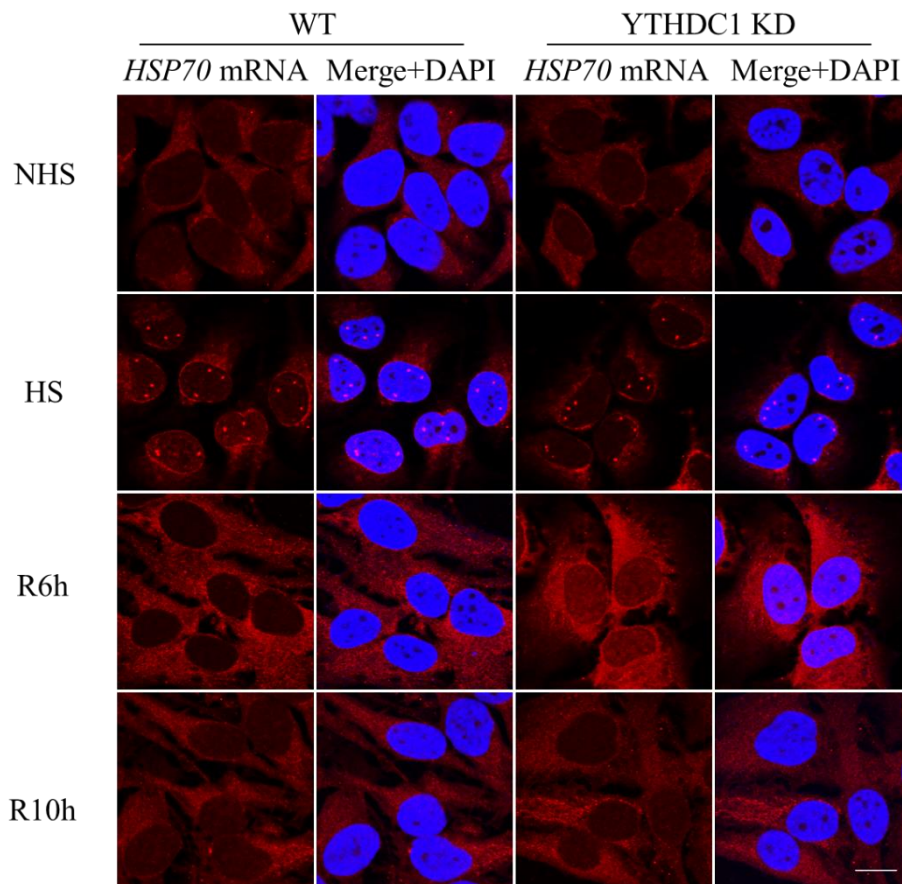


**Figure 60: YTHDC1 KD induces partial nuclear retention of *HSPA1A*, *HSPA1B*, *HSP40* and *actin* mRNAs.** RT-qPCR analyses of the mRNA levels of *HSPA1A*, *HSPA1B*, *HSP40* and *actin* following cell fractioning of HeLa WT or YTHDC1 KD cells in NHS, HS, R6h or R10h. Mean values from three independent experiments are presented as nuclear/cytoplasmic ratio. *Performed by Aude Grézy.*

To evaluate the importance of YTHDC1 for the export of *HSP* mRNAs after HS we also used a microscopy-based approach. RNA Fluorescent *In Situ* Hybridization (FISH) was conducted in WT and YTHDC1 KD cells at NHS, HS, R6h and R10h using a specific probe against *HSP70* mRNA (Materials and Methods VI.2). Importantly, this approach allows us to visualize directly the localization of the transcripts in individual cells and presents a qualitative rather than quantitative assessment.

The microscopic analyses indicated an important heterogeneity of the *HSP70* mRNA signal between cells within each condition, presumably due to the sensitivity of the applied technique and/or resulting from the different cell treatments (+/- HS and +/- siRNA transfection). Regardless, of these variations, in NHS, when *HSP70* mRNAs are present at low levels, no specific signal is detected (Fig.61). After the HS treatment, the expression of *HSP70* is induced and in WT and KD cells we observe the characteristic accumulation of the newly-synthesized mRNAs in the proximity

of their transcription site (three bright nuclear foci in HeLa cells). Importantly, no significant differences of the size of these foci were observed after the KD of YTHDC1. Hence, this observation suggests that the reduced YTHDC1 levels do not lead to increased accumulation of *HSP70* transcripts near their transcription site. Later, after 6h of recovery *HSP70* mRNAs are mainly exported in the cytoplasm. In WT cells at this time point the nuclear signal is strongly decreased while the cytoplasmic one is elevated compared to the HS condition. In YTHDC1 KD cells this cytoplasmic increase is observed but in addition we notice a more important diffuse nuclear signal in approximately 50% of the cells. At R10h the cytoplasmic signal becomes weaker in both WT and KD cells indicating that the mRNAs levels drop.



**Figure 61: YTHDC1 KD induces partial nuclear retention of *HSP70* mRNA during the recovery period after HS.** Fluorescent *in situ* hybridization of *HSP70* mRNAs using the specific pH2.3 genomic probe in WT and YTHDC1 KD HeLa cells in NHS, HS, R6h and R10h. Cell nuclei are stained with DAPI, scale bar- 10 $\mu$ m.

The performed FISH experiments do not allow us to conclude firmly on the effect of YTHDC1 KD on the export of *HSP70* mRNAs. We were able to visualize particularly a nuclear retention of *HSP70* mRNA at R6h in a proportion of YTHDC1 KD. Our previous studies, using cell fractioning, have confirmed a more global role of YTHDC1 in the export of *HSP70* in normal and stress conditions.

Overall, both approaches, especially the cell fractioning experiments, suggest that YTHDC1 is important but not essential for the correct export of *HSP70* and *HSP40*. Thus, we suggest that this molecular function of the protein is involved to some extent in the expression of HSPs.

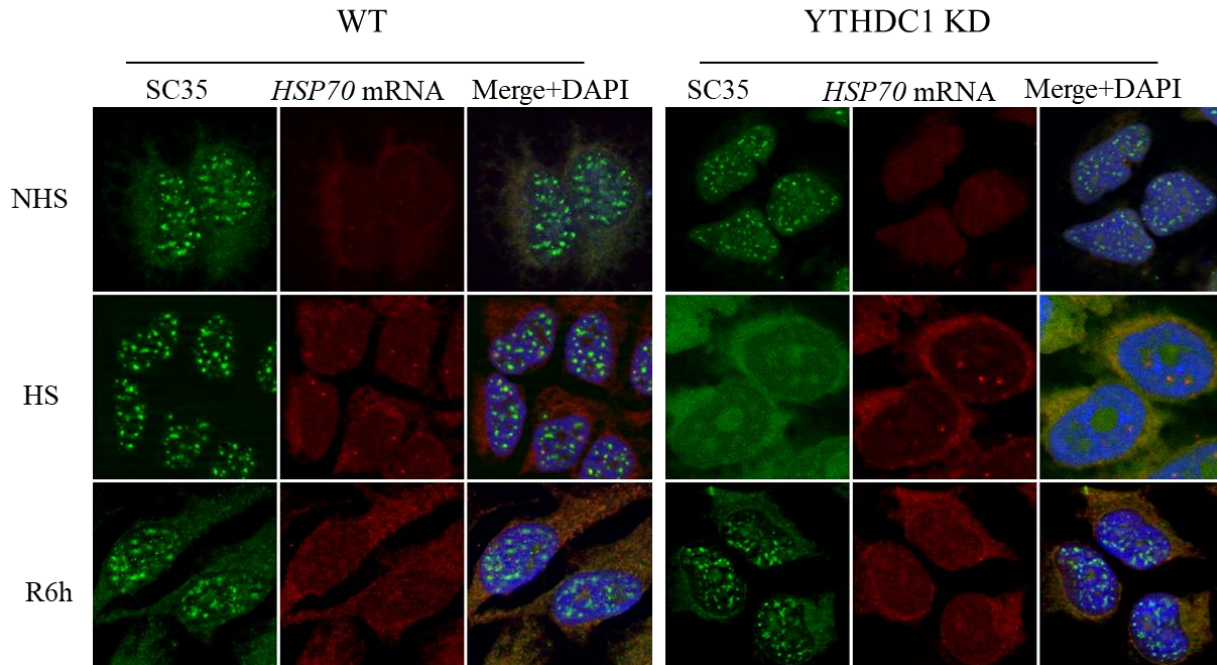
Altogether this work demonstrated that some of the main molecular functions of YTHDC1 are implicated in the induction of chaperones upon heat stress. The protein KD does not seem to affect the transcription of *HSP70* and *HSP40* with the possible exception of transcription termination. Our study also revealed a role of YTHDC1 in the export of *HSP* transcripts.

## **II.V YTHDC1 maintains Nuclear Speckles structure upon HS**

Nuclear Speckles (NSs) are known as important mRNA processing sites, hosting splicing, polyadenylation and 3'-end processing factors as well as components of the export machinery (Galganski et al., 2017). These nuclear structures are needed for the proper mRNA maturation and association with proteins essential for the nuclear export and assisting further the correct translation in the cytoplasm. A recent study described for the first time that intronless mRNAs, especially *HSP70*, need to go through NSs in order to be correctly exported and later translated into proteins (Wang et al., 2018). After our initial results indicating partial nuclear retention of *HSP70* in YTHDC1 KD cells, we were interested in assessing whether the protein is needed for the transit of this intronless mRNA through NSs. We hypothesized that the decreased YTHDC1 levels might impair this process and therefore affect *HSP70* export and translation.

To test this hypothesis, we combined the *HSP70* RNA FISH to the immunofluorescence detection of SC35 (or SRSF2), a well-known marker of NSs (Materials and Methods VI.II). This single cell, double-detection approach showed in WT cells the expected colocalization of *HSP70* transcripts (that localize mainly at their sites of production) and SC35-containing NSs, upon HS induction (Fig.62). The protein SC35 presents a very characteristic granular pattern detected as

multiple irregular foci, distributed all over the nucleus. Surprisingly, in YTHDC1 KD cells, specifically in HS approximately 80% of the cells do not present this particular SC35 staining, but rather a diffuse, homogenous nuclear signal.



**Figure 62: YTHDC1 KD and HS disrupt the foci-like Nuclear Speckle staining.** Fluorescent *in situ* hybridization of *HSP70* mRNA using the specific pH2.3 genomic probe (red) coupled with immunofluorescence of SC35 (green) in WT and YTHDC1 KD HeLa cells in NHS, HS and R6h. DNA is counterstained with DAPI. Scale bar- 10 $\mu$ m.

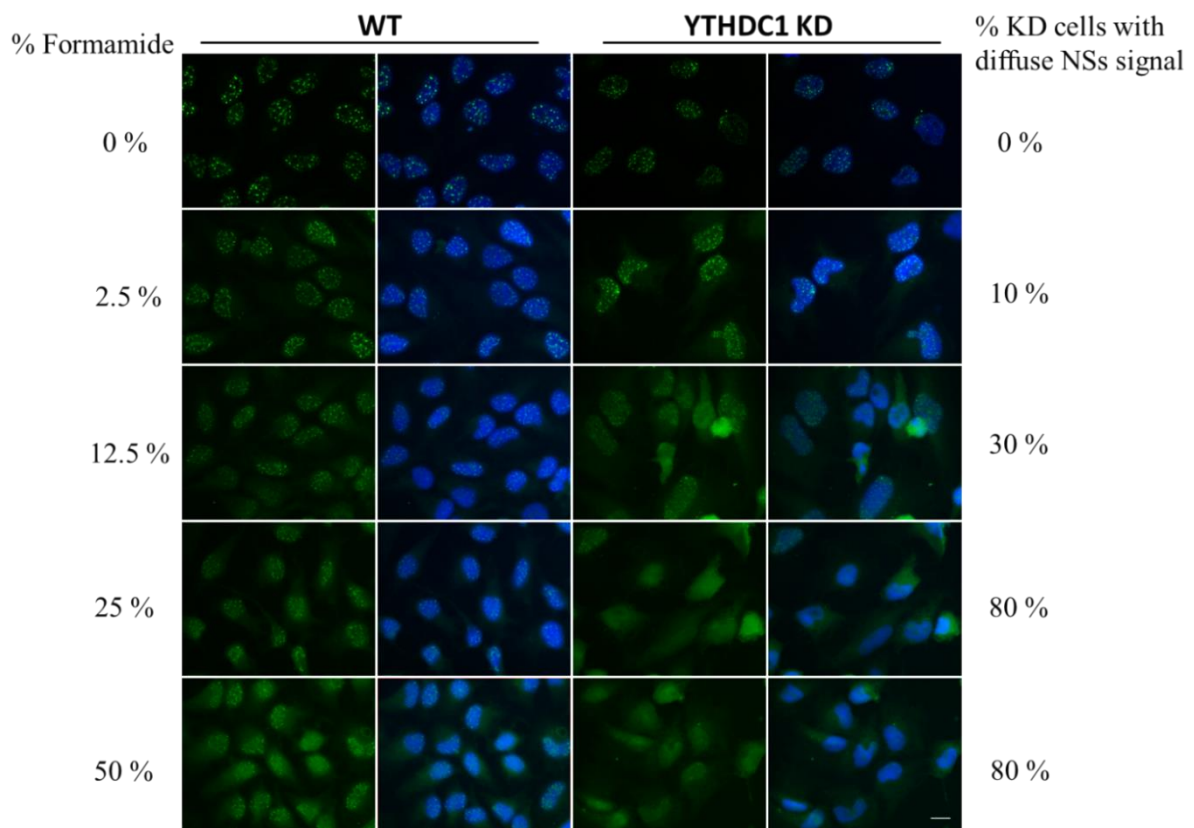
These experiments unraveled an unexpected disruption of the granular foci-like staining of SC35 in YTHDC1 KD cells upon HS. This effect appears to be HS-related since in NHS and after 6 hours of recovery the SC35 protein is detected in KD cells as in the control.

To further explore this intriguing result, we have carried out immunofluorescence of SC35 in the same conditions, without combining it to RNA-FISH. Interestingly, this time we did not detect the loss of the granular nuclear speckles signal in KD cells after HS. Hence, using solely immunofluorescence, we were able to observe a correct nuclear speckles staining in all tested conditions. This curious finding implied that the treatments used for the RNA FISH experiments are possibly contributing to the observed effect.

Following on this hypothesis, we tested the effect of various substances used in the RNA-FISH technique and which may impact on the structure and interaction of proteins or RNAs (saline-sodium citrate buffer, formamide and ethanol). We found that formamide is an essential element for the loss of SC35 foci-like signal in absence of YTHDC1 and after HS.

WT and YTHDC1 KD HeLa cells subjected to HS were incubated with different concentrations of formamide- from 0% to 50% (corresponding to the concentration used for FISH assays) and subjected afterwards to SC35 IF.

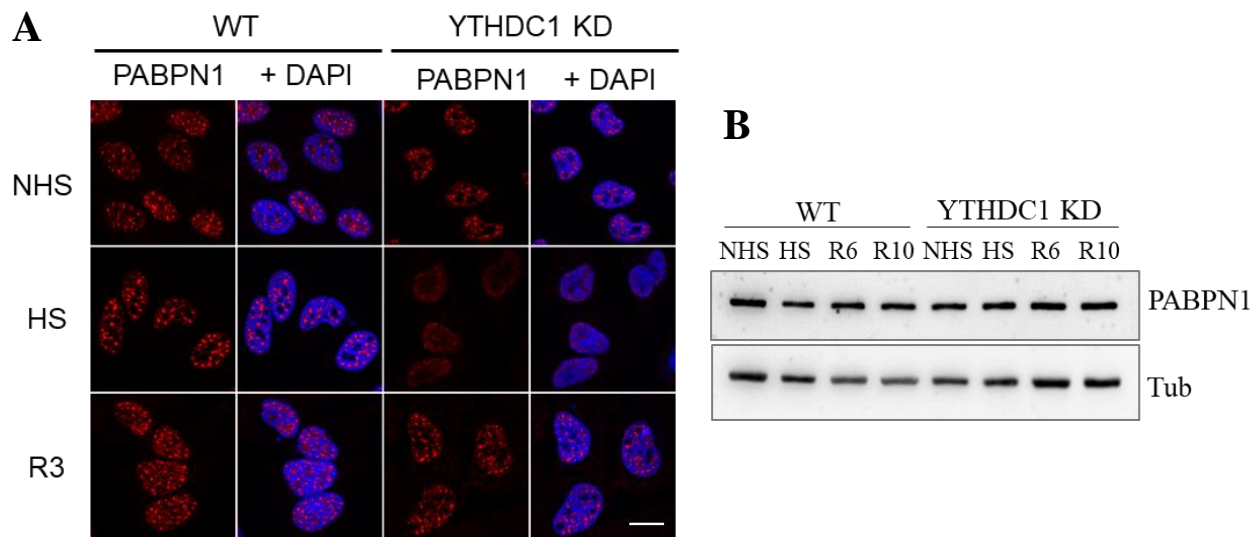
Formamide addition affects the detected fluorescent SC35 signal (Fig.63), which becomes weaker and blurrier as the concentration of formamide increases in both WT and YTHDC1 KD cells. However, the specific staining of nuclear speckles is observed in all conditions in WT cells while in YTHDC1 KD cells the SC35 staining becomes more diffuse and less granular with the gradual increase of the formamide concentration.



**Figure 63: Formamide treatment leads to diffuse Nuclear Speckles signal in HS, following KD of YTHDC1.** Immunofluorescence of SC35 (green) in WT and YTHDC1 KD HeLa cells at HS. Prior to the IF, the cells were incubated or not (0%) with increasing formamide concentrations (% indicated on the left). The percentage of KD cells with diffuse NSs signal is indicated on the right. Cell nuclei are stained using DAPI, scale bar 10µm.

At 0% of formamide, all KD cells show strong regular SC35 signal, at 12.5% formamide around 30% of the cells lose the punctual signal and at 25% and 50% of formamide nearly 80% of them present only diffuse nuclear speckles staining and no foci distribution is visible (Fig.63). These analyses confirmed that formamide specifically impacts the foci-like distribution of SC35 in HS, following KD of YTHDC1.

In order to verify whether the combined YTHDC1 KD-HS-formamide effect is specific for SC35 or more globally for NSs, the same experiments were next redone with a second marker of NSs- the protein PABPN1 (for Poly(A) Binding Protein Nuclear 1). In these conditions, once again we observed the loss of signal of the protein in nearly 80% of the KD cells in HS (Fig.64A). Moreover, western blot analysis indicated that this loss of punctual signal upon HS does not correspond to decreased PABPN1 protein levels in YTHDC1 KD cells following stress (Fig.64B). The levels of PABPN1 stay also comparable to the once in WT cells during the recovery period (R6h and R10h, Fig.64B). Therefore, these experiments confirmed a global impact on nuclear speckles upon reduced levels of YTHDC1, HS and formamide treatment, which was not associated with the expression of NSs components.



**Figure 64: Formamide treatment allows us to visualize an effect of YTHDC1 KD on the structure of nuclear speckles upon HS.** **A.** Immunofluorescence of PABPN1 (red) in WT and YTHDC1 KD HeLa cells at NHS, HS and R3h. Prior to the IF, the cells were incubated 50% formamide solution. Cell nuclei are stained using DAPI, scale bar 10µm. **B.** Western blot analysis of PABPN1 on whole-cell extract of WT and YTHDC1 KD HeLa cells at NHS, HS, R6h and R10h. Tubuline (Tub) is used as loading control.

Taken together the studies presented in this section led us to the conclusion that formamide disorganizes nuclear speckles which are weakened upon HS in absence of YTHDC1. Our discovery indicates that YTHDC1 is maintaining nuclear speckles formation, structure or stability specifically in stress conditions. This function of YTHDC1 might be important for the proper export and translation of numerous different mRNAs in the context of stress.

Altogether the analyses presented in these two parts of the results highlight the importance of YTHDC1 for the regulation of the HSR. The protein is involved at multiple levels of the pathway and regulates diverse processes- alternative splicing at nSBs, stress-related induction of HSPs, with a particular relevance for the transcription termination and export of *HSP* mRNAs, and has as well as a less investigated role in the maintain of NSs upon HS.





## **Chapter III**

# **New cellular functions of YTHDC1 outside the stress context**

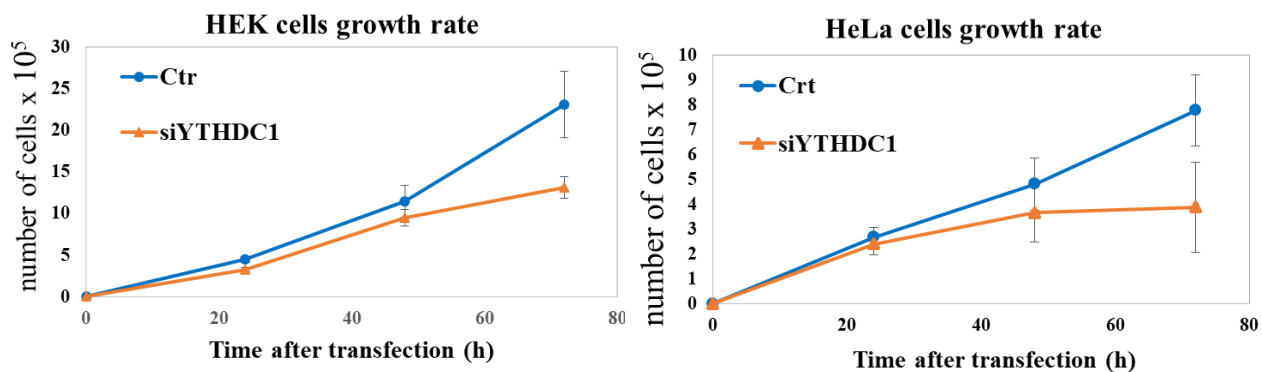


### **III.I YTHDC1 KD induces cell mortality through apoptosis**

Along with the recent extensive studies of the epitranscriptome, some of the main functions of the nuclear m6A reader, YTHDC1, have been also described. Our research contributed and completed this description with the discovery of the key role of YTHDC1 in cellular stress response. Nevertheless, a lot remains unknown and findings suggests that various important molecular functions of the protein might still be unknown.

A major interest of our work on YTHDC1 is to unravel and characterize these potential new functions. For these reasons during my PhD I have searched for the implication and role of the protein in the regulation of fundamental cellular processes such as cell proliferation. The RNA mark m6A has been already described as essential for cell proliferation. Particularly the genomic deletion of m6A writers (METTL3, METTL14 and WTAP) has been associated with decreased cellular growth in the case of human and mouse embryonic stem cells, and some specific cancer cell lines (Batista et al., 2014; Geula et al., 2015; Shuibin Lin et al., 2016). However, very little was known in regard to the molecular actors of m6A involved in the regulation of cellular proliferation. Therefore, we have addressed the question of the potential implication of YTHDC1 in this key cellular process. The preliminary study of the changes of cellular growth, viability and cell cycle when YTHDC1 expression levels are strongly decreased in cells will be presented in this last chapter of the results.

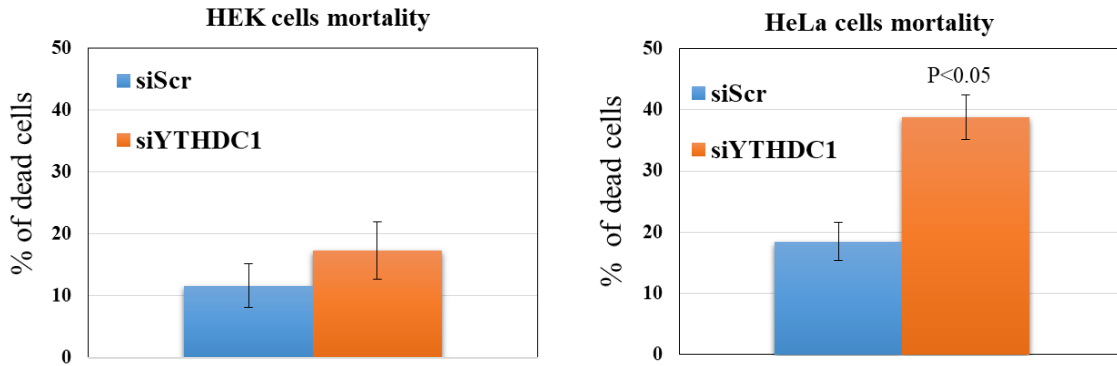
To begin we have established growth curves of WT and YTHDC1 KD HeLa and HEK cells. The transient KD was performed as described previously (Material and methods I.II.1) and was maintained for 72h following the cell transfection. The number of cells was counted from the day of transfection till 72h later and compared to the number of non-transfected, WT cells. As shown in Fig.64 both cell lines presented a reduced cell growth when YTHDC1 levels were diminished compared to the control. The effect was more pronounced toward the end of the assay-between 48h and 72h when a maximum KD efficiency is achieved. The cell growth rates were overall decreased in a similar way in both cells lines- 2 times in HEK cells and 2.5 times in HeLa cells. These initial findings suggest that YTHDC1 could have an important role in cell proliferation and prompt us to decipher this possible function of the protein



**Figure 65: YTHDC1 KD decreases the growth rates of HeLa and HEK cells.** Growth rates of HEK and HeLa cells at 24h, 48h and 72h after transfection with siRNAs against YTHDC1. Ctr-WT cells. Graphs represent mean values from three independent experiments.

The observed effect of YTHDC1 KD could reflect cell death and/or could result from defects at the cell cycle level. In order to examine these possibilities, we have first studied the cellular viability using propidium iodide exclusion assay and flow cytometry analyses. Propidium iodide is a fluorescent intercalating agent which does not cross the cell membrane and is excluded from viable cells. It is used to specifically stain the DNA content of dead cells with fragmented membrane. The levels of cell mortality were assessed 72h after the siRNA transfection in both HEK and HeLa cells. For each flow cytometry experiment an important number of events was analyzed (60 000 events after debris elimination/ condition). To rule out possible cell death due to the transfection, cells transfected with non-specific siRNAs (siScramble, siScr) were used as a control. In HEK cells, we noticed only a slight increase of cell mortality of about 5%, 72h after the transfection (Fig.66). In HeLa cells, on the other hand, we detected a 2.2-fold increase of mortality in YTHDC1 KD cells compared to the control. At 72h nearly 40% of HeLa cells were dead and presented strong propidium iodide staining (Fig.66).

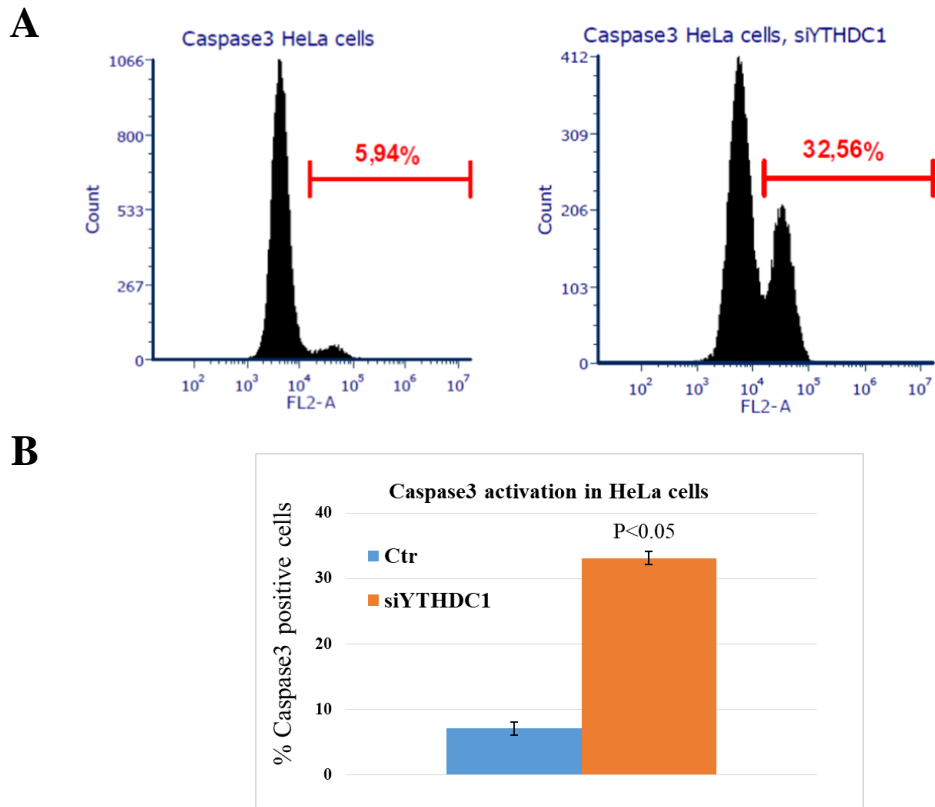
These interesting results reveal a differential effect of the KD of YTHDC1 in HEK and HeLa cells and suggest distinct mechanisms of action of the protein in the two cell lines.



**Figure 66: The KD of YTHDC1 induces increased mortality in HeLa cells.** Mortality of HEK and HeLa cells, 72h after the KD of YTHDC1 assessed using propidium iodide exclusion assay and flow cytometry. Cells transfected with siScramble (siScr, blue) are used as control. Histograms represent mean percentage of dead cells from three independent experiments. P-value was calculated using a two-tailed Student's t-test. \*:  $p < 0.05$ .

To discern the role of YTHDC1 in cellular proliferation, first we have focused on its impact on cell viability in HeLa cells. To begin we wished to characterize the observed cell mortality and understand whether it is due to the activation of apoptotic pathways. We have used two approaches— we have studied by flow cytometry the levels of activated caspase 3 and by western blot analysis the activation of PARP (Poly (ADP-Ribose) Polymerase). These two proteins are main actors of the apoptotic signaling and are rapidly affected by external and internal stress stimuli and/or DNA damage. To be activated they need to be cleaved and the use of specific antibodies against the active forms allows us to conclude whether apoptosis is taking place or not (Materials and Methods VIII.3). As shown in Fig.67A, in HeLa WT cells, treated with the transfection agent used for the KD assays, only 6% of the cells present positive caspase 3 staining after 72h of incubation. In YTHDC1 KD cells, on the other hand, approximately 33% of the cells present positive staining for the activated form of caspase 3. The results indicate that following KD of YTHDC1 the levels of active caspase 3 increase 5.5 times compared to the control (Fig.67B).

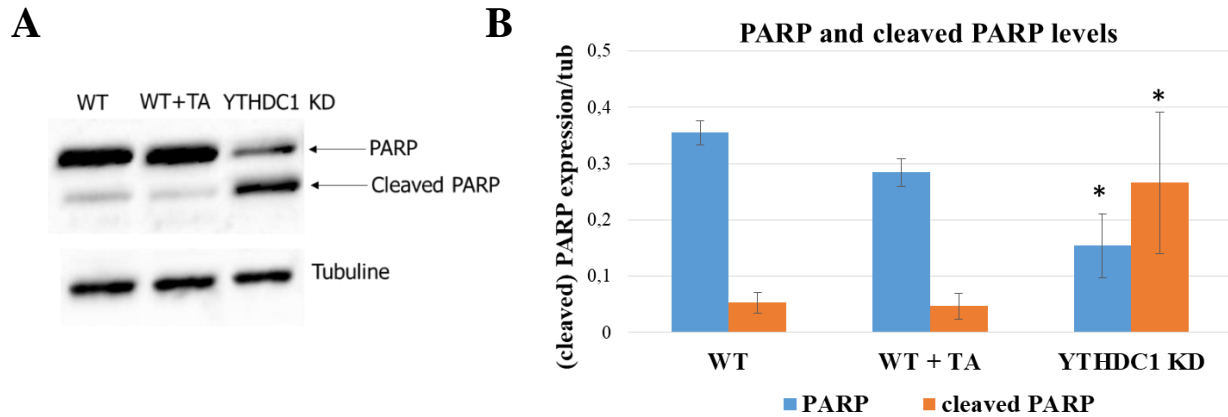
These findings were the first indication that in YTHDC1 KD cells apoptotic pathways are being activated.



**Figure 67: YTHDC1 KD induces activation of caspase 3 in HeLa cells.** **A.** Flow cytometry analyses graphs of the number of cells (count) and the fluorescent intensity of cleaved caspase 3 signal (FL2-A). Analyses on HeLa cells 72h after transfection with specific siRNAs or treatment with the transfection agent (control, Ctr). Active Caspase 3 is detected with PE Rabbit Anti-Active Caspase 3 antibody. **B.** Histogram of the percentage of cells with active caspase 3 staining in control (Ctr) and YTHDC1 KD HeLa cells. Mean values from three independent experiments. P-value was calculated using a two-tailed Student's t-test. \*:  $p < 0.05$ .

Next, we have analyzed the cleavage of PARP in the same experimental conditions. For the western blot study of PARP we have used a specific antibody detecting the full-length and the cleaved protein.

As shown in Fig.68A in WT cells, treated or not with the transfection agent, we detect low levels of the cleaved PARP form. In YTHDC1 KD cells the levels of PARP decrease while those of cleaved PARP increase drastically. The average signals from three independent experiments indicate a 2.2-fold decrease of PARP and a 5.2-fold increase of cleaved PARP signals in YTHDC1 KD cells. Thus, this second approach demonstrates a significant activation of PARP when YTHDC1 levels are strongly reduced.



**Figure 68: YTHDC1 KD induces activation of PARP in HeLa cells.** **A.** Western blot analysis of full-length (PARP) and cleaved PARP in WT cells, treated or not with the transfection agent (TA) used for YTHDC1 KD assays, and in YTHDC1 KD cells. Tubulin is used as loading control. **B.** Levels of PARP and cleaved PARP obtained from three independent experiments in the same conditions as in A. Tubulin is used as loading control. P-value was calculated using a two-tailed Student's t-test. \*:  $p < 0.05$ .

Overall, these analyses revealed a new molecular function of YTHDC1 in cellular proliferation. In response to the KD of YTHDC1, HeLa cells activate apoptotic pathways which result in the increased cell mortality and reduced cell growth rates. When HEK cells present diminished YTHDC1 levels, their growth is greatly affected. However, in this cell lineage viability does not seem to be altered.

### III.II YTHDC1 KD affects cell cycle progression

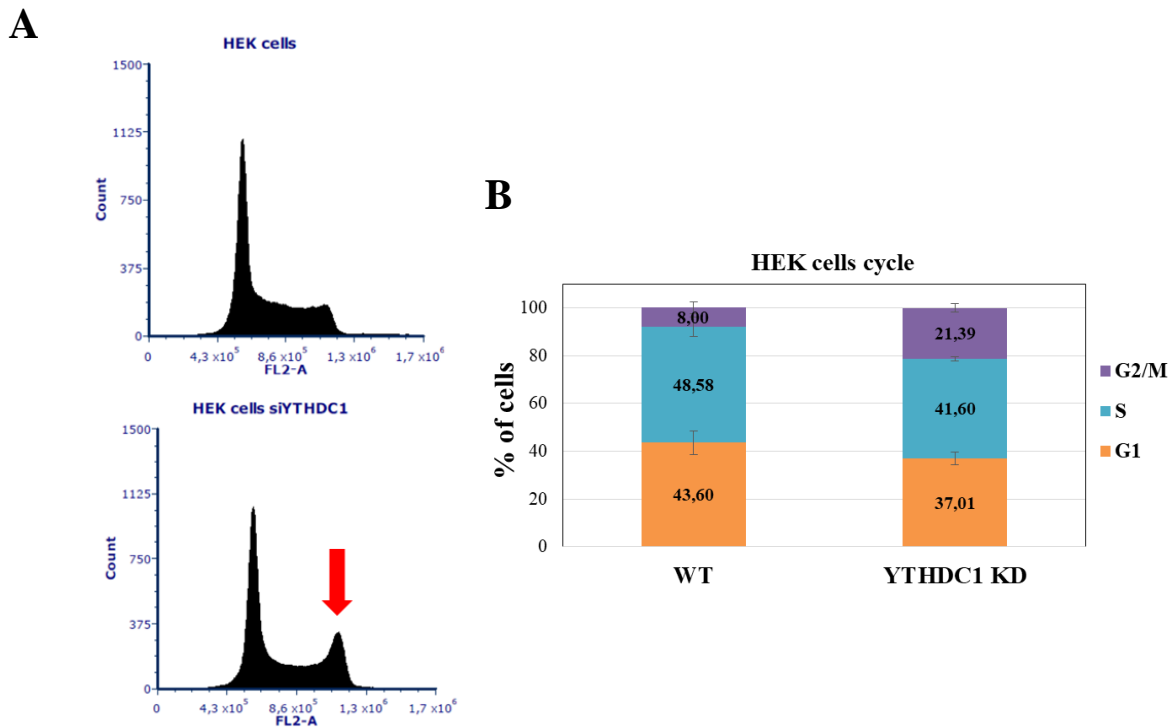
Apoptosis and cell cycle progression are tightly linked. Apoptotic stimuli often lead to cell cycle defects and conversely improper cell cycle progression can cause activation of apoptotic pathways in order to eliminate the damaged cells. After the description of the effect of YTHDC1 KD on cell viability through apoptosis we wished to analyze its impact on the cell cycle of HEK and HeLa cells.

For these experiment, the cells were permeabilized prior to the incubation with propidium iodide. Therefore, the intensity of the fluorescent dye reflects directly the DNA quantity of all cells. In this analysis 60 000 events were processed after the elimination of cellular debris. The distribution of cells in each phase of the cell cycle was determined using a specific software. The



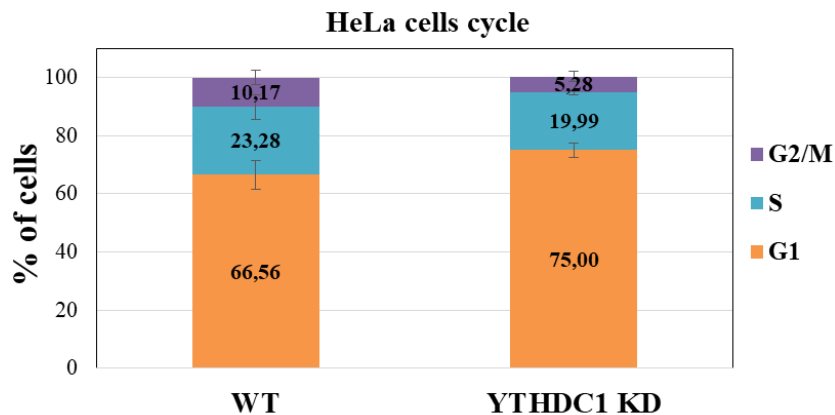
proposed cell cycle models were first applied to WT cells and afterwards to YTHDC1 KD cells allowing comparative analyses.

In WT HEK cells we detect a regular cell cycle distribution with important number of dividing cells at the time point of the experiment. Following KD of YTHDC1 in this cell line a strong increase of the G2/M peak is observed (Fig.69A, lower panel). The obtained data indicates that when the levels of the protein are reduced the percentage of cells in the G2/M phase increases nearly 2.8 times compared to the control (Fig.69B). In WT cells 8% of the cells are in G2/M phase and in YTHDC1 KD cells approximately 22% are present in this phase of the cell cycle. Meanwhile the numbers in G1 and S phase are slightly reduced compared to the WT cells. This interesting result shows a specific blockage of HEK cells in G2/M phase following a strong expression decrease of YTHDC1.



**Figure 69: YTHDC1 KD leads to G2/M phase blockage of HEK cells.** **A.** Flow cytometry analyses of WT and YTHDC1 KD (siYTHDC1) HEK cells. Cell cycle assessed using propidium iodide staining of permeabilized cells. Count- number of cells, FL2-A- intensity of the propidium iodide signal. Arrow indicates the increased G2/M peak. **B.** Histogram representation of the mean percentage of cells in each phase of the cell cycle in WT and YTHDC1 KD HEK cells. Represented are mean values from three independent experiments.

The same experiments were performed also in HeLa cells. The highest proportion of WT cells were found in G1 phase (66%), around 25% of the cells were in S-phase and 10% in the G2/M phase at the time of the analysis. No differences were detected in the overlay distribution of WT and YTHDC1 KD cells. However, as shown in Fig.70 the percentage of HeLa cells in G2/M phase is reduced two times in the YTHDC1 KD cells while the G1 peak slightly increases. This finding indicates a possible function for YTHDC1 in the regulation of cell cycle progression in HeLa cells.



**Figure 70: YTHDC1 KD decreases the proportion of HeLa cells in G2/M phase.** Histogram representation of the mean percentage of cells in each phase of the cell cycle in WT and YTHDC1 KD HeLa cells. Represented are mean values from three independent experiments.

Taken together, these results point again to a differential effect of the KD of YTHDC1 in HEK and HeLa cell lines. YTHDC1 turns out to be an important regulator of HEK cells proliferation and its decreased levels lead to a significant cell blockage in G2/M phase. On the other hand, in HeLa cells the protein KD affects slightly the cell cycle and results in reduced number of cells in G2/M phase. These contrasted outcomes lead to the hypothesis that distinct pathways are being affected by the reduced levels of the protein.

In general, our preliminary study on the role of YTHDC1 in cell proliferation unraveled a novel, yet relatively less understood function of the protein which might also correlate with the implication of m6A in cell viability.



# **Discussion and perspectives**



In this last section of my manuscript the results obtained during my PhD thesis are discussed in the light of other studies carried out in the fields of stress biology and the epitranscriptome. Ideas for additional analysis as well some of the many opportunities for future investigations will be presented here. At the end a general conclusion, including the main results will be given.

## **I. Regulation of cellular stress responses through the epitranscriptome?**

The conserved cellular Heat Stress Response (HSR) is one of the most extensively studied survival pathways activated in the case of increased temperature and deregulated protein and cellular homeostasis (Morimoto, 1998). Even though many of the molecular mechanisms, linked to the proper execution of the HSR, have been revealed over the past two decades, still many aspects of its regulation remain elusive.

For instance, the transcription factor HSF1 was believed for long time to be the main orchestrator of the HSR, acting at the level of both coding and non-coding genes. However, recent genome-wide studies revealed that HSF1 has a rather limited role in the control of the heat shock-induced transcriptomic landscape (Mahat et al., 2016). Therefore, these latest findings highlight the need to search for novel key regulators of the fascinating genome reprogramming observed in response to thermal stress.

Another intriguing question in the domain is the potential control of the HSR through RNA modifications. Recently a central role for the emerging field of the epitranscriptome in the implementation of stress programs was proposed and presents many perspectives for studies on the involved regulatory mechanisms.

In particular, regarding the execution of the HSR, several studies revealed that m6A is required for the control of the transcriptional coding response to thermal stress (Knuckles et al., 2017; Meyer et al., 2015; Zhou et al., 2015). Nevertheless, the precise mechanisms and the involved m6A molecular players were not fully elucidated.

In this work we have examined the implication of the nuclear m6A reader, the protein YTHDC1 in the regulation of the cellular response to heat stress. Along with the rapid progress in the field of the epitranscriptome, many of the functions of YTHDC1, which have been unknown at the start of my PhD project, were discovered. Today, we know that the protein is essential for gene expression and regulates key steps of the process such as alternative splicing, alternative

polyadenylation and mRNA export. Using different approaches we have demonstrated the central role of YTHDC1 in the control of the HSR and the importance of some of its molecular roles for the expression of stress-responsive genes. We have observed that YTHDC1 functions and genomic targets are diverse following the distinct phases of the HSR- the immediate response upon heat stress and the recovery period following thermal shock. These results and the observed effects of YTHDC1 upon HS and during recovery will be discussed in the next parts and have been summarized in a model presented in Fig.71.

## **II. Heat shock recovery: YTHDC1 at nuclear Stress Bodies**

### **II.I YTHDC1 relocalization to nSBs**

In the beginning of this work, we have investigated the subcellular localization of YTH-domain proteins in HeLa cells, in response to heat stress. Our results indicated that YTHDC1 presents a particular foci-like staining during the recovery period after HS (Fig.33 and 34A). This finding, indicated for the first time that the protein might be regulated and/or present a specific role in the context of increased growth temperature.

Regarding the rest of the known m6A readers, we found that YTHDF2 forms cytoplasmic foci immediately after the HS, persisting also during the recovery period (Fig.33 and 34B). The subcellular localization of YTHDF2 following heat stress was often discussed in the literature during the past few years. It was proposed that YTHDF2 relocalizes to the cell nucleus upon HS (Zhou et al., 2015), while others indicated that after heat stress the protein is exclusively cytoplasmic in different cell lineages (Ries et al., 2019). Therefore, it is worth mentioning that our results go along with the study of Ries et al. and show no cytoplasmic to nuclear relocation of the protein in response to stress. Several recent publications uncovered the phase separation capacities of YTHDF proteins and their implication in the formation of stress granules (SGs) (Fu and Zhuang, 2020; Ries et al., 2019b). Thus, the observed relocation of YTHDF2 into cytoplasmic foci, might be linked to the formation of SGs upon HS, however this question has not been examined in this work.

Following the discovery that YTHDC1 localization changes during the recovery period after heat treatment, we have investigated the nature of the detected foci-like nuclear structures. Importantly, our analyses revealed that YTHDC1 relocates to the transiently formed characteristic

stress compartments called nuclear Stress Bodies (nSBs) at the level of *satIII* pericentromeric heterochromatin (Fig.71). The protein is found at nSBs mainly at three and six hour of recovery following HS and redistributes homogenously in the nucleus 24h after the stress induction (Fig.34).

We found YTHDC1 to colocalize with HSF1 at R3h and with the core component of nSBs- the lncRNAs *SATIII* at R3h and R6h (Fig.36). Further additional analysis, using HSF1 KD stable cell lines and a particular HSF1 mutant, revealed that the HSF1-dependent *SATIII* transcription is essential for the relocation of YTHDC1 to nSBs (Fig.37 and 38). In line with this we suggest that *SATIII*, which persist longer than HSF1 at nSBs, are required to maintain YTHDC1 at the level of these nuclear structures during the recovery period.

Our finding of YTHDC1 relocation to nSBs prompt us to study the importance of m6A for this process. The two applied methods- KD of METTL3 and rescue experiments with mutated YTHDC1, not able to bind m6A, pointed to an m6A-independent mechanism of YTHDC1 relocation to nSBs. However, both approaches present several limitations and we can only suggest that the process is m6A-independent. To have a more accurate vision on this question, YTHDC1 localization after HS can be studied following the concomitant deletion of METTL3 and METTL14 which was found to induce the loss of nearly 99% of m6A residues in polyadenylated RNAs in human cells (Geula et al., 2015; Lin et al., 2017).

This first part of the project unraveled YTHDC1 relocation to nSBs during the recovery period following HS and demonstrated the importance of *SATIII* lncRNAs for this process. However, the precise molecular mechanism responsible for YTHDC1 relocation was not fully elucidated in these studies. YTHDC1 could relocate to nSBs through an interaction with RNAs found at nSBs, factors recruited to nSBs in a *SATIII*-dependent manner or with both.

Based on a performed ChIP of YTHDC1 at the level of *satIII* DNA repetitive sequences (data not shown), we propose that YTHDC1 does not interact directly with the *SATIII* lncRNAs. Up to date the RNA content of nSBs has not been fully examined. Hence, apart from their core components- the *SATIII*, it is unknown whether other RNAs can be found at nSBs. Nevertheless, if YTHDC1 relocation is driven via an interaction with RNAs, it seems not to be dependent on the presence of m6A.



The comparison of the interactomes of YTHDC1 and *SATIII* (discussed in the next part) showed that in normal growing conditions YTHDC1 already interacts with many proteins found later at the level of nSBs. This strongly suggests that YTHDC1 relocation can rely on the interaction with factors found also to relocalize to nSBs in a *SATIII*-dependent manner. Comparative studies of YTHDC1 partners at NSH, HS and during the recovery period, together with the additional analysis of their importance for YTHDC1 relocalization may shed light onto this question.

We have observed as well that the foci-like concentration of the protein does not correlate with changes of its expression levels upon HS (Fig.35). Interestingly, early studies evoked that the phosphorylation of YTHDC1 regulates its subcellular localization and its role in splicing (Rafalska et al., 2004). To go further, we have performed western blot analysis of the protein upon HS and during the recovery period, this time in presence of phosphatase inhibitors (data not shown). The obtained results did not indicate the presence of potential YTHDC1 PostTranslational Modifications (PTMs) in these conditions. In perspective, it will be interesting to apply mass spectrometry analysis in order to investigate more precisely if YTHDC1 is subjected to PTMs in response to stress and subsequently address their importance for its relocation to nSBs.

## **II.II YTHDC1 and the stress-induced regulation of alternative splicing**

Over the years many processing and in particular splicing factors were found to relocate to nSBs in response to HS. Therefore, it was believed that *SATIII* participate to the regulation of stress-induced splicing events via the titration of diverse splicing regulators at nSBs (Biamonti and Vourc'h, 2010; Denegri et al., 2001; Metz, 2004). The precise molecular mechanisms involving *SATIII* in the control of alternative splicing have been demonstrated very recently (Ninomiya et al., 2020). The study of Ninomiya and indicated that *SATIII* lncRNAs promote intron retention during the recovery period following HS. In particular, nSBs turned out to serve as platform for the recruitment of SR proteins and their main regulator- the kinase CLK1 which is subjected itself to alternative stress-related splicing by SR proteins (Long et al., 2019). The axis *SATIII* (nSBs)-CLK1-SR proteins was found to modulate the splicing of hundreds of mRNAs during the recovery period following stress (Introduction II.IV.2).

In our attempt to decipher the potential functional significance of YTHDC1 relocalization to nSBs, we hypothesized that the protein might participate to the specific modulation of alternative

splicing in the context of stress. To investigate this, we have compared the interactomes of YTHDC1 and *SATIII*. The immunoprecipitation of YTHDC1 was carried out earlier, in the course of studies on the molecular roles of the protein, and was performed in HEK cells, in normal growth conditions. Meanwhile the ChIRP-seq analysis of *SATIII* partners was conducted in HeLa cells, during the recovery period after HS induction (Ninomiya et al., 2020). Regardless of the different experimental conditions, we were able to find out that YTHDC1 and *SATIII* share 31 common partners among which 25 are main splicing regulators- SR proteins, hnRNPs, spliceosome components and others (Fig.44).

We have further investigated the effect of YTHDC1 KD on the splicing of *Clk1* mRNA. Our results indicated that as the *SATIII* lncRNAs, YTHDC1 promotes intron retention at *Clk1* (Fig.71). We discovered that YTHDC1 KD does not affect *CLK1* mRNA splicing upon HS but leads specifically to a strong decrease of the intron-retaining form of *CLK1* mRNA at R3h and R6h (Fig.45). This decrease was not associated with reduced *CLK1* mRNA levels and was assigned specifically to the stress-related splicing modulation of *CLK1* mRNA.

Importantly, since CLK1 is a master regulator of alternative splicing, it was revealed that *SATIII* KD affects the splicing of nearly five hundred mRNAs. Hence, we suggest that YTHDC1 at the level of nSBs could have a key and central role for the alternative splicing and the stress-related expression of hundreds of mRNAs. However, this aspect is currently under investigation and needs to be further confirmed.

During the recovery period, the intron retention at *CLK1* mRNA is induced by phosphorylated SR proteins. Since nSBs serve as platform for the interaction of CLK1 with SR proteins, the KD of *SATIII* which disrupts nSBs, prevents the phosphorylation of SR proteins by CLK1 and subsequently affects the splicing of the kinase itself and many other mRNAs. The results of Ninomiya and colleagues identified particularly the phosphorylation state of SRSF9 to be involved in the intron retention during recovery.

In our study we have observed that nSBs are forming in YTHDC1 KD cells as in WT cells (data not shown). Thus, we can hypothesize that even if nSBs are forming correctly, YTHDC1 KD might alter the interaction of CLK1 with SR proteins. Importantly, SRSF9 is found to be significantly enriched in the interactome of YTHDC1. Therefore, YTHDC1 might be essential for the localization of SRSF9 (and potentially other SR proteins) to nSBs. Indeed, it was observed that

in normal conditions YTHDC1 is required for the Nuclear Speckles (NSs) localization of several SR proteins (Xiao et al., 2016). However, the mechanisms involved in the maintain of SR proteins at NSs and at nSBs are still poorly understood. Additional studies on the localization of CLK1 and SR proteins in YTHDC1 KD cells could further decipher the mechanisms behind the role of YTHDC1 in alternative splicing at nSBs.

Another possible explanation for the observed effects is that in YTHDC1 KD cells CLK1 and SR proteins interact at nSBs, however the phosphorylated SR proteins are not successfully recruited to their targets and subsequently the intron retention is abolished. In fact it has been previously shown that YTHDC1 is required for the recruitment of SRSF3 (Srp20) to its targeted mRNAs (Xiao et al., 2016). The study by Ninomiya and colleagues proposed that intron retention occurs outside of nSBs since the intron-retaining mRNAs were not detected in the *SATIII* ChIRP-seq fraction. Thus, upon stress YTHDC1 might shuttle between nSBs and nucleoplasm in order to bring SR proteins to their targets. To investigate this, the association of SR proteins and targeted mRNAs can be studied in WT and YTHDC1 KD cells during the recovery period following HS.

In this axis of the project it is important in addition to examine the relevance of m6A for the regulation of splicing events via YTHDC1 at nSBs. Our initial analyses on the relocalization of the protein to nSBs suggest that the process might be m6A-independent. However, if YTHDC1 is needed to recruit the phosphorylated SR proteins, this process may require m6A binding at the level of the SR-targeted mRNAs. Such m6A-dependent mechanism of recruitment of SRSF3 by YTHDC1 was previously reported (Xiao et al., 2016).

### **III. YTHDC1 and the expression of stress-responsive genes**

#### **III.I YTHDC1 genomic targets following heat stress**

In the framework of our study, in order to have a more comprehensive view of YTHDC1 behavior in response to stress we have performed genome-wide ChIP-seq analysis of its targets in absence of stress and upon HS. Previously, YTHDC1 targets have been investigated in normal conditions using RIP-seq (Roundtree et al., 2017) or PAR-CLIP analysis (Xiao et al., 2016). Here, we have selected the ChIP-seq approach based on the fact that YTHDC1 is essential for co-transcriptional RNA processing events (alternative splicing and alternative polyadenylation) and

moreover the m6A mark, potentially needed for YTHDC1 functions, was recently found to be deposited in a co-transcriptional manner (Ke et al., 2017).

Importantly, the performed ChIP-seq analysis of YTHDC1 at NHS and HS were pivotal and revealed the broad variety of coding genes which are bound and possibly regulated by the protein in both normal and stress condition (Fig.48). Thus, this genome-wide study presents countless opportunities for future investigations of the molecular and cellular functions of YTHDC1.

The striking observation that the number of YTHDC1 targets increases nearly three times following stress demonstrated that the protein is recruited to new chromatin sites in response to HS. This result revealed for the first time on the genome level the dynamic stress-related behavior of YTHDC1. Regarding, the role of YTHDC1 in the control of the HSR, the comparative study of the profiles of YTHDC1 and RNA Pol II suggests a role of YTHDC1 not only at the level of stress-induced genes but maybe also at the level of the genes found to be downregulated upon stress. This aspect is currently under investigation and could unravel the association and possible importance of YTHDC1 for the stress-related massive shut-down of thousands of active genes observed rapidly after HS.

It is of note that we have performed the YTHDC1 ChIP-seq study also during the recovery period after HS, at R6h. The exciting preliminary results from this analysis suggest that during the HSR YTHDC1 shuttles between different genomic sites in order to participate on a large scale to the stress-related genome reprogramming. Thus, the final results from this investigation could further highlight the central role of YTHDC1 in the implementation of the conserved stress response to HS.

Within the various interesting results obtained from the YTHDC1 ChIP-seq analysis at NHS and HS, our attention was caught by the fact that upon stress many of YTHDC1 targets are involved especially in the regulation of cellular stress response. Additional analysis showed that YTHDC1 associates with 50 different HSPs upon HS (Fig.50). The levels of YTHDC1 were found to increase mostly at the 3'UTR region of diverse *HSP*-coding genes, part of all main classes of chaperones (Fig.71).

Recently, stress-induced increase of m6A at the 5'UTR and the 3'UTR was proposed to regulate the expression of HSPs (Knuckles et al., 2017; Zhou et al., 2015). Based on our finding that YTHDC1 is recruited to various *HSP* genes following HS and the suggested role for m6A, we have examined the potential implication of YTHDC1 in the expression of HSPs in stress conditions.

### **III.II YTHDC1 is essential for the induction of HSPs upon HS**

We have studied the role of YTHDC1 by investigating specifically its function on the expression of two HSPs- HSP70 (*HSPA1A* and/or *HSPA1B*) and HSP40 (*DNAJB1*). Both of these HSPs are inducible and their expression, and cellular roles have been extensively studied in the literature (Michels et al., 1997; Ohtsuka and Hata, 2000). HSP40 are cofactors of HSP70 and the two chaperones are central players in the preservation and restoration of the proteome, the cellular homeostasis and viability in stress environment.

One of the most striking results in this work was that YTHDC1 is essential for the induction of HSPs during the recovery period after HS (Fig.51). Using rescue experiments in our stable cell line model, which will be further discussed, we have validated the specific relevance of YTHDC1 for the expression of HSPs (Fig.54). Importantly, the abolished induction after KD of YTHDC1 was not associated with decreased *HSP* mRNA levels (Fig.56).

It is of note, that such loss of HSP70 induction has been previously described following KD of YTHDF2 while an increased induction of HSP70 was observed in FTO KD cells (Meyer et al.2015; Zhou et al., 2015). As mentioned earlier, YTHDF2 was proposed to relocalize to the nucleus and to protect m6A-modified transcripts from demethylation by FTO (one of the two known m6A demethylases). It was reported that the loss of induction (YTHDF2 KD) or the increased induction (FTO KD) does not correlate with modifications of *HSP70* mRNA levels, compared to WT cells. Altogether, these two studies proposed that m6A is needed for the increased expression of HSP70 and in particular for the cap-independent translation of *HSP70* mRNAs following heat stress.

On the other hand, increased m6A levels at the 3'UTR were associated with the degradation of stress-induced transcripts by the RNAi machinery (Knuckles et al., 2017). In line with this, the

KD of METTL3 (the main m6A methyltransferase) was found to correlate with increased induction of HSP70 following HS in opposition to the described effect of FTO KD.

In the course of our study, we have observed that YTHDF2 KD abolishes HSP70 induction as previously described in the literature. However, our results indicated that METTL3 KD leads as well to loss of stress-related HSP70 induction rather than increased expression as reported by Knuckles et al. Therefore, collectively our analysis support the proposed essential role of m6A for the induction of HSPs.

Related to these intriguing recent findings, we have investigated if YTHDC1 capacity to recognize m6A is needed for the stress-related expression of HSP proteins. The results from the performed rescue experiments indicate that YTHDC1 requires m6A in order to regulate the induction of HSP40 (Fig.55). These analyses allowed us to determine the importance of m6A for YTHDC1-dependent expression of HSPs and to further validate the key role of this RNA mark for the induction of HSPs upon stress.

YTHDC1 is required for the regulation of gene expression and is implicated in mechanisms which are essential for mRNA metabolism and decisive for RNA fate (3'-end processing and transcription termination, alternative splicing, mRNA export). Therefore, we have suggested that YTHDC1 KD might alter crucial steps of the processing of *HSP* mRNAs which would later on hinder their translation.

Using different and complementary approaches we have studied the importance of YTHDC1 for the transcription termination of *HSP70* and *HSP40*. We observed that at both *HSP70* and *HSP40* 3'UTR regions the levels of RNA Pol II-S2P are decreased in YTHDC1 KD cells compared to the enrichment of the Pol II-S2P at the same regions in WT cells (Fig.57). These results imply that YTHDC1 might be required for the appropriate transcriptional termination of heat stress-responsive genes (Fig.71). Nevertheless, additional controls such as analysis of the levels of total RNA Pol II are needed in parallel to these experiments in order to have a solid conclusion on the impact of YTHDC1 KD.

It is well described in the literature that at the 3'UTR of protein-coding genes Pol II is capable of sensing its passage across a functional poly-A site (PAS) where it pauses and interacts with factors from the cleavage-polyadenylation complex (CPA). However, depletion of key

components of the CPA could reduce this pausing effect. It has been reported that YTHDC1 depletion leads to extensive alternative polyadenylation (APA) in mouse oocytes and results in the shortening or lengthening of the 3'UTR of nearly thousand genes (Kazowitz et al., 2018). This function of the protein was proposed to rely on its interaction with several factors of the 3'-end processing machinery- CPSF6 (CFIm complex), SRSF3 and SRSF7. In the YTHDC1 immunoprecipitation analysis which we have carried out, we also identified 9 different 3'-end processing factors among which CPSF1, CPSF5 and CPSF6 (Fig.43). Therefore, we can suggest that YTHDC1 KD might affect the recruitment of CPA components to the PAS which leads to reduced Pol II-S2P pausing at the 3'UTR and alters the transcription termination.

We have examined as well the production of readthrough transcripts at the level of *HSP70* and *HSP40* in YTHDC1 KD cells in normal and stress conditions. We did not detect differences of readthrough formation between WT and YTHDC1 KD cells at neither of the two genes (Fig.58). In the same way it has been recently reported that hTHO is required for the transcription termination at *HSPA1A*, however its depletion did not correlate with an increased formation of readthrough transcripts at this locus (Katahira et al., 2019). Importantly, hTHO is also found significantly enriched in YTHDC1 interactome, thus the two proteins could be involved in the same mechanisms of regulation of transcription termination at *HSP* genes.

Another possible effect of YTHDC1 on the expression of HSPs could be linked to its function in splicing. In response to stress splicing is one of the many gene expression regulatory mechanisms which is affected. In the case of *HSP*-coding genes, which are mainly intronless, their differential splicing upon HS has been relatively poorly studied. Yet, the question on how some HSP isoforms resulting from alternative splicing are produced following stress is elusive. In our study, we have observed that even though *HSP40* mRNA is differentially spliced upon HS, YTHDC1 is not required for the involved mechanisms (Fig.59).

We have studied as well the impact of YTHDC1 KD on the export of *HSP40* and *HSP70* mRNAs. We found that YTHDC1 KD leads to partial nuclear retention of *HSP40*, *HSPA1A*, *HSPA1B* and *actin* in both normal and stress conditions (Fig.60). Such global role of YTHDC1 in the export of mRNAs has been already described in the literature (Roundtree et al., 2017). Interestingly, we observed a stronger impact on the export of *HSP40* which was accentuated

following stress. Altogether, these analysis indicated that YTHDC1 is not particularly required for the export of stress-responsive transcripts but is needed especially for the export of *HSP40* mRNAs.

It was shown that YTHDC1 through an interaction with SRSF3 participates in the recruitment of the export adaptor NXF1 (Roundtree et al., 2017). Furthermore, the interaction of SR/SRSF proteins is required for the loading of TREX components and the subsequent nuclear export of mRNAs (Viphakone et al., 2019). Since *HSP40* is subjected to alternative splicing which is tightly correlated to the export of mRNAs we can imagine that even if YTHDC1 is not involved precisely in the splicing of *HSP40* it might be needed for the loading of the export machinery onto this transcript.

Additionally, we have carried out RNA FISH experiments to detect *HSP70* mRNAs in WT and YTHDC1 KD cells after HS and during the recovery period. Due to the technical limitations of this approach and the variables introduced by the different cell treatments we noticed only a partial nuclear retention of *HSP70* mRNAs in a proportion of YTHDC1 KD cells and especially at R6h. Nevertheless, this analysis led to the conclusion that YTHDC1 KD does not induce accumulation of *HSP70* mRNAs near their transcription sites. Such accumulation and increase of the size of the transcription foci of *HSP70* has been demonstrated in the case of abolished nuclear export following KD of Thoc5 or Aly/Ref (Katahira et al., 2009). The role of YTHDC1 in the export of mRNAs was linked to its interaction with TREX components and export adaptors (Lesbirel et al., 2018). The correlation between the mechanisms of loading of TREX components, SRF3 and NXF1 involving YTHDC1 as mediator is yet unclear. We can hypothesize that YTHDC1 affects one of this steps, which however does not disrupt the release of mRNAs from their transcription site but later on affects at least partially their export from nucleus to cytoplasm.

We propose that YTHDC1 is involved in HSPs induction possibly via its capacity to bind m6A. We have observed that YTHDC1 is mostly enriched at the 3'UTR of *HSP*-coding genes which could strongly correlate with its role in the transcription termination and export of *HSP* mRNAs. However, at certain *HSP* genes YTHDC1 was found in addition at the 5'UTR. YTHDF2 was suggested to protect 5'UTR m6A, still since its relocalization to the nucleus in response to stress remains elusive, this mechanism of protection of m6A might rely on YTHDC1. Thus, in addition to its role in the processing of *HSP* mRNAs, YTHDC1 could be indirectly required for their translation. In the same context, it is tempting to address the question of the levels of m6A at



*HSP* mRNAs in YTHDC1 KD cells. Another possible future study would be to examine the targets of a mutated YTHDC1, not binding m6A, in NHS and HS and study globally the importance of m6A for the recruitment of YTHDC1 to chromatin, and more specifically to *HSP* genes.

Importantly, the induction of HSPs is crucial for the preservation of cell viability in the context of stress. HSPs are involved in the regulation of cell cycle and apoptosis following HS (Antonova et al., 2019; Truman et al., 2012; Wang et al., 2014). Our ChIP-seq analysis indicated that some of YTHDC1 strongly enriched targets in HS are implicated in cell cycle, cell division and cell death. Taken together these observations suggest that YTHDC1 might be needed for the preservation of cell viability and for the proper cell recovery following stress. This question presents an interesting opportunity for future studies.

Moreover, the expression of HSPs is often found deregulated and increased in diverse chronic diseases (Calderwood and Gong, 2016). Since recent studies often link the development of various pathological conditions to different m6A-dependent signaling pathways, we could imagine that YTHDC1 role in the regulation of increased HSPs expression extends beyond the stress context to the level of pathological onset and progression.

### **III.III YTHDC1 and the stability of Nuclear Speckles following stress**

Along with the study on the function of YTHDC1 related to the induction of HSPs, we have discovered a potential novel role for the nuclear m6A reader in the maintain of the structure of Nuclear Speckles (NSs) in stress conditions (Fig.71). The granular and foci-like staining of two NSs markers- SC35 and PABPN1 were lost following KD of YTHDC1 and especially upon HS. Interestingly, diffuse NSs signal was detected in these conditions only after incubation of the cells with formamide (Fig.63). Few studies have shown that PABPN1 and its cytoplasmic analog PABP1 are sensitive to detergent treatment which can lead to their aggregation and subsequently affect their localization. Here, we observed a particular effect of formamide on the nuclear distribution of NSs components, which was never reported before. We propose that this chemical agent affects protein-protein or protein-RNA interactions and therefore disorders the association between NSs constituents. Since we detect disruption of NSs specifically in YTHDC1 KD and after HS we imagine that the combined effect of YTHDC1 KD and heat stress weakens the structure of NSs which further disassemble after addition of formamide. This effect is observed solely upon

HS and is transient since during the recovery period NSs are again forming correctly in YTHDC1 KD cells (Fig.62).

Notably, YTHDC1 KD did not impact the expression of PABPN1 following stress. This additionally suggests that upon HS YTHDC1 is required to sustain the structure of NSs rather than to regulate the expression of NSs components.

NSs are essential for the orchestration of mRNA processing events. Interestingly, the transit of *HSP70* mRNAs through NSs was shown to be required for their nuclear export and translation (Fig.71) (Kim et al., 2019; Wang et al., 2018). The decreased stability of NSs following YTHDC1 KD and HS could be potentially linked to the partial nuclear retention of *HSP70* in these conditions. Furthermore, YTHDC1 could impact the loading of essential export adaptors needed not only for the translocation of mRNAs from nucleus to cytoplasm but later on for their proper translation. In the case of *HSP* mRNAs this mechanisms could be required for their induction upon HS. More globally, via the maintain of NSs structure YTHDC1 might be implicated in vast processes of gene expression regulation after HS. Nevertheless, this aspect hides numerous open questions.

There is still no clear evidence if YTHDC1 is found at NSs in normal and stress conditions. It was proposed that, based on their molecular functions, YTHDC1 and other players of m6A signaling could localize to NSs (Galganski et al., 2017). However, future proteomic studies of NSs are needed to decipher this hypothesis.

Indeed, our immunoprecipitation studies show that YTHDC1 interacts with many proteins, known from the literature to localize to NSs and implicated especially in the splicing and export of mRNAs. This observation further supports the idea that YTHDC1 localizes, at least partially, to NSs.

Previous studies indicated that in normal conditions YTHDC1 is required, through yet unknown mechanisms, for the localization of SRSF3 and SRSF10 at NSs (Xiao et., 2016) However, we observe a more global impact of YTHDC1 KD on the structure of NSs, exclusively in specific conditions- following stress and formamide treatment. The involved mechanisms might be common or at least related. We can hypothesize that at NSs the interaction of YTHDC1 with RNAs and/or other proteins could be particularly important for the stability of these nuclear

compartments upon HS. Further, in YTHDC1 KD cells after HS, formamide leads to the disassembly of NSs due to their decreased stability.

To continue this investigation, it is important first to validate the specificity of YTHDC1 for the observed effect. For this, the same experiments need to be redone in a rescue model.

The structure of NSs has been relatively less studied under stress conditions compared to normal conditions and similar effects, as the one we detect, have not been described for other proteins. It is interesting to compare the impact of formamide with other substances which disrupt specifically protein-protein or protein-RNA interactions. This can further point to the mechanisms through which YTHDC1 can participate in the stabilization of NSs structure. If the mechanism is protein-dependent, a key study would be to search among YTHDC1 partners known to localize to NSs for the one (or more) needed as well to maintain the structure of NSs upon HS. The identification of such partner(s) could indicate which interaction of YTHDC1 with another molecular actor is potentially essential for the stability of NSs upon HS.

Importantly, if YTHDC1 role in NSs structure/stability relies on an interaction with RNA, this mechanisms could further rely on m6A recognition. It has been shown that *MALAT1* lncRNA, which is essential for the structure of NSs is m6A-modified (Zhou et al., 2016). Therefore, an investigation of the interaction of YTHDC1 with *MALAT1* is also a potential interesting study which could point to mechanisms of NSs stability linked to both m6A prevalence and YTHDC1 action.

#### **IV. YTHDC1 as regulator of fundamental cellular processes**

In the course of this work, we carried out a preliminary study on the role of YTHDC1 on cell proliferation and cell cycle progression. We have discovered that in HeLa cells YTHDC1 KD leads to strong cell mortality (Fig.66). Further, we have associated this mortality to the activation of apoptotic pathways in YTHDC1 KD cells (Fig.67, Fig. 68). On the other hand, in HEK cells YTHDC1 KD led also to decreased cellular growth. In this cell line, we observed that YTHDC1 is required especially for cell cycle progression since its KD led to an arrest in the G2/M phase (Fig.69).

These results unravel a novel cellular function of the protein in the regulation of apoptosis and cell cycle progression. Interestingly, we notice differential effects in HeLa and HEK cells. We

can imagine that the signaling pathways regulating cell death and cell cycle are not executed in the same way in these two lineages and thus, they might be differentially affected by the KD of YTHDC1.

Similar results have been observed in studies on the effect of m6A on fundamental cellular processes. The levels of m6A writers, readers and erasers have been very often correlated to deregulated cellular proliferation and defective cell cycle progression (Cai et al., 2019; Batista et al., 2014; Geula et al., 2015; Lin et al., 2016). However, many of the mechanisms are still not fully understood since often the impact of m6A and its signaling network turns out to be controversial depending on the examined cell line and the underlying signaling pathways. However, the significance of m6A recognition for YTHDC1 functions in cell proliferation and cell cycle, and possibly for its differential effects in distinct cell lines, need to be further investigated.

Previously, the deletion of YTHDC1 was found to induce early embryonic lethality in mice which was proposed to result from severe splicing defects and modulated APA events (Kasowitz et al., 2018). Our study reports on one hand activation of apoptosis and on the other hand cell cycle arrest in respectively HeLa and HEK YTHDC1 KD cells. These effects could be as well linked to defective mRNA processing after KD of YTHDC1, especially of transcripts coding for apoptotic or cell cycle regulators.

This preliminary investigation opens a lot of opportunities for future studies. The comparative analysis of YTHDC1 partners and targets in HeLa and HEK cells can reveal distinct mechanisms possibly regulated by the protein in these cell lines. Notably, our ChIP-seq study showed that in both normal and stress conditions YTHDC1 associates with genes coding for cell adhesion. Cell viability and proper cell cycle progression are dependent on the cellular capacities to adhere and establish contacts with the extracellular matrix (Jones et al., 2019; Suzanne and Steller, 2009). YTHDC1 through its various molecular functions might be on the crossroad of these fundamental cellular processes.

We noticed that particularly after HS, some of YTHDC1 most significantly enriched targets code for regulators of cell cycle, cell division and cell death. Therefore, the role of YTHDC1 in the regulation of cell viability and cell cycle might be even more important in the context of hostile environment and threatened cellular homeostasis.

## V. m6A-dependent YTHDC1 functions?

In this work we have addressed the relevance of m6A for the functions of YTHDC1. This question was investigated based on several important observations.

First, earlier studies in our team and other research groups revealed that Mmi1, the homolog of YTHDC1 in the fission yeast *S.Pombe*, shares many of its molecular functions with YTHDC1- RNA 3' end processing, splicing, decay and export (Chen et al., 2011; Kilchert et al., 2015; Touat-Todeschini et al., 2017). Importantly, in *S. Pombe* m6A is not present since the m6A writers homologous to human METTL3 and METTL14 proteins are absent in this organism. Therefore, Mmi1 functions rely on the binding of RNA via its YTH domain but independently from m6A. This suggests that its human homolog might also act via mechanisms beyond m6A recognition.

Even though diverse molecular roles have been assigned to YTHDC1, the mode of action of the protein is yet poorly understood. Based on its capacity to bind m6A, YTHDC1 functions have been continuously linked to the epitranscriptomic network. However, only few studies examined directly, through the use of mutated YTHDC1, the significance of this RNA modification for the functions of the protein. For the moment the role of YTHDC1 in alternative splicing is the sole to be described as m6A-dependent (Xiao et al., 2016; Kasowitz et al., 2018).

Further, previously published YTHDC1 PAR-CLIP data demonstrated that only 39% of YTHDC1 targets are methylated (Xiao et al., 2016). This observation also implies a potential m6A-independent mechanisms of action of YTHDC1.

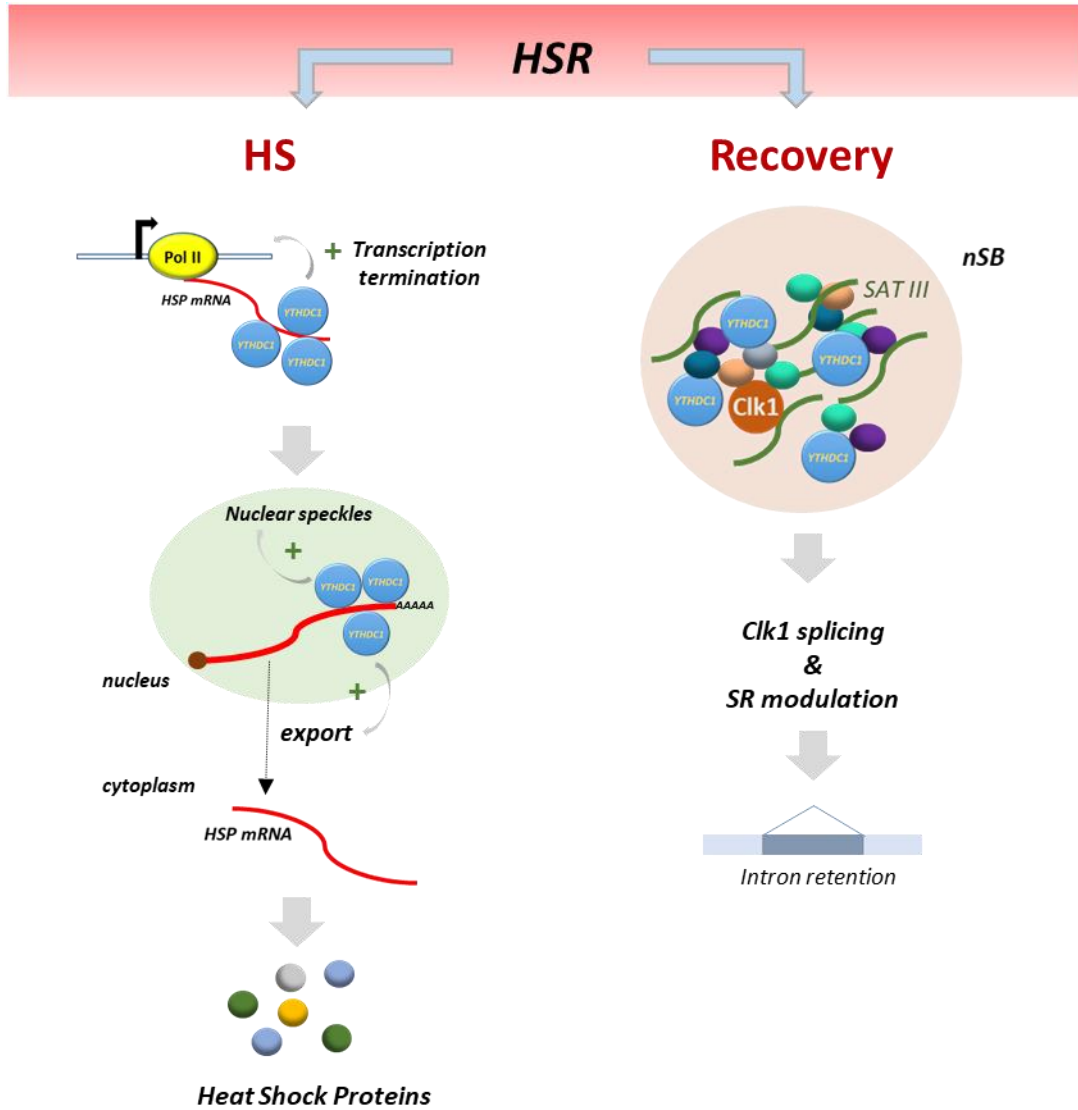
In our study we have developed a particular model of genetically modified HeLa inducible cell lines, coding for siRNA-resistant WT YTHDC1 or mutated YTHDC1, unable to recognize m6A. Notably, this model was selected since YTHDC1 is essential and a knock-out of the protein was not achievable. However, the induction of YTHDC1 (WT or mutated), combined with its very effective transient KD, will allow us to investigate the m6A-dependency of some of YTHDC1 functions. It is of note that compared to the rescue experiments performed using transient transfection, the stable cell line rescue setup presents several advantages. In this model the expression levels can be easily adjusted to be close to the expression levels of the endogenous protein. Moreover, this model allows to overcome problems such as low transfection rates and variable expression levels between cells which can be often observed after transient transfection.

This stable inducible cell line model was not applied in all of the studies performed during my PhD project. However, it can be used in the future to decipher the mechanisms standing by YTHDC1 molecular and cellular functions, and to study the significance of m6A for diverse processes. More importantly, using different mutation constructs, the established model can be employed to investigate numerous aspects of YTHDC1 functioning.

# Conclusion

In this work we have discovered a novel role of YTHDC1 in the regulation of the cellular heat stress response (HSR). The protein may act at multiple levels of this conserved pathway. We found that upon HS YTHDC1 is recruited to chromatin where it associates with numerous protein-coding genes, essential for the implementation of the HSR. Among them, many *HSP*-coding genes were found to be targets of YTHDC1. Furthermore, we demonstrated that YTHDC1, potentially through binding m6A, is essential for the stress-related induction of HSPs. Additional analyses indicated that YTHDC1 may be required for the transcription termination of *HSPs* and the subsequent nuclear export of their mRNAs. Moreover, following HS YTHDC1 somehow maintains the structure of nuclear speckles, which are key actors in gene expression. Later on, during the recovery period after HS, we showed that YTHDC1 relocates to nuclear Stress Bodies (nSBs), formed at the level of pericentromeric heterochromatin. We characterized this relocalization as dependent on the HSF1-induced transcription of *SATIII* lncRNAs and as possibly m6A-independent. Our studies revealed that at the level of nSBs YTHDC1 may participate in the stress-related regulation of the alternative splicing of potentially hundreds of mRNAs. Finally, our preliminary analyses indicated as well that in normal conditions YTHDC1 is required for cell viability and cell cycle progression in two different cell lineages.

Taken together the results presented in this work associate through complex molecular mechanisms the fields of stress biology and the epitranscriptome. The highlighted role of YTHDC1, and by extension the m6A signaling, in the implementation of stress programs and in the regulation of fundamental cellular processes outside the stress context provides opportunities for numerous perspective studies.



**Figure 71: Model for the role of YTHDC1 in the regulation of the cellular Heat Stress Response.** YTHDC1 acts at multiple levels of the HSR- upon HS and during the recovery period. After HS YTHDC1 is recruited mainly to the 3'UTR of *HSP*-coding genes and is essential for their induction. YTHDC1 might be required for the transcription termination and nuclear export of *HSP* mRNAs. The protein maintains the structure of nuclear speckles following stress. During the recovery period YTHDC1 relocates to nuclear Stress Bodies (nSBs) where together with *SATIII* lncRNAs, CLK1 and SR proteins it participates in the stress-related modulation of alternative splicing events.



# Bibliography

- Abravaya, K., Phillips, B., and Morimoto, R.I. (1991). Attenuation of the heat shock response in HeLa cells is mediated by the release of bound heat shock transcription factor and is modulated by changes in growth and in heat shock temperatures. *12*.
- Adelman, K., and Lis, J.T. (2012). Promoter-proximal pausing of RNA polymerase II: emerging roles in metazoans. *Nat Rev Genet 13*, 720–731.
- Åkerfelt, M., Morimoto, R.I., and Sistonen, L. (2010). Heat shock factors: integrators of cell stress, development and lifespan. *Nat Rev Mol Cell Biol 11*, 545–555.
- Alarcón, C.R., Goodarzi, H., Lee, H., Liu, X., Tavazoie, S., and Tavazoie, S.F. (2015). HNRNPA2B1 Is a Mediator of m<sup>6</sup>A-Dependent Nuclear RNA Processing Events. *Cell 162*, 1299–1308.
- Allen, T.A., Von Kaenel, S., Goodrich, J.A., and Kugel, J.F. (2004). The SINE-encoded mouse B2 RNA represses mRNA transcription in response to heat shock. *Nat Struct Mol Biol 11*, 816–821.
- Aly, M.K., Ninomiya, K., Adachi, S., Natsume, T., and Hirose, T. (2019). Two distinct nuclear stress bodies containing different sets of RNA-binding proteins are formed with HSATIII architectural noncoding RNAs upon thermal stress exposure. *Biochemical and Biophysical Research Communications 516*, 419–423.
- Amor, D.J., Kalitsis, P., Sumer, H., and Andy Choo, K.H. (2004). Building the centromere: from foundation proteins to 3D organization. *Trends in Cell Biology 14*, 359–368.
- Anckar, J., and Sistonen, L. (2011). Regulation of HSF 1 Function in the Heat Stress Response: Implications in Aging and Disease. *Annu. Rev. Biochem. 80*, 1089–1115.
- Anders, M., Chelysheva, I., Goebel, I., Trenkner, T., Zhou, J., Mao, Y., Verzini, S., Qian, S.-B., and Ignatova, Z. (2018). Dynamic m<sup>6</sup>A methylation facilitates mRNA triaging to stress granules. *Life Sci. Alliance 1*, e201800113.
- Andersson, R., Refsing Andersen, P., Valen, E., Core, L.J., Bornholdt, J., Boyd, M., Heick Jensen, T., and Sandelin, A. (2014). Nuclear stability and transcriptional directionality separate functionally distinct RNA species. *Nat Commun 5*, 5336.
- Antonova, A., Hummel, B., Khavaran, A., Redhaber, D.M., Aprile-Garcia, F., Rawat, P., Gundel, K., Schneck, M., Hansen, E.C., Mitschke, J., et al. (2019). Heat-Shock Protein 90 Controls the Expression of Cell-Cycle Genes by Stabilizing Metazoan-Specific Host-Cell Factor HCFC1. *Cell Reports 29*, 1645-1659.e9.
- Ashworth, W., Stoney, P.N., and Yamamoto, T. (2019). States of decay: The systems biology of mRNA stability. *Current Opinion in Systems Biology 15*, 48–57.

- Baler, R., Welch, W., and Voellmy, R. (1992). Heat shock gene regulation by nascent polypeptides and denatured proteins: hsp70 as a potential autoregulatory factor. *The Journal of Cell Biology* *117*, 1151–1159.
- Barakat, T.S., and Gribnau, J. (2012). X chromosome inactivation in the cycle of life. *Development* *139*, 2085–2089.
- Batista, P.J., Molinie, B., Wang, J., Qu, K., Zhang, J., Li, L., Bouley, D.M., Lujan, E., Haddad, B., Daneshvar, K., et al. (2014). m6A RNA Modification Controls Cell Fate Transition in Mammalian Embryonic Stem Cells. *Cell Stem Cell* *15*, 707–719.
- Biamonti, G. (2004). Nuclear stress bodies: a heterochromatin affair? *Nat Rev Mol Cell Biol* *5*, 493–498.
- Biamonti, G., and Vourc'h, C. (2010). Nuclear Stress Bodies. *Cold Spring Harbor Perspectives in Biology* *2*, a000695–a000695.
- Boccaletto, P., Machnicka, M.A., Purta, E., Piątkowski, P., Bagiński, B., Wirecki, T.K., de Crécy-Lagard, V., Ross, R., Limbach, P.A., Kotter, A., et al. (2018). MODOMICS: a database of RNA modification pathways. 2017 update. *Nucleic Acids Research* *46*, D303–D307.
- Bonasio, R., Tu, S., and Reinberg, D. (2010). Molecular Signals of Epigenetic States. *Science* *330*, 612–616.
- Buchan, J.R., and Parker, R. (2009). Eukaryotic Stress Granules: The Ins and Outs of Translation. *Molecular Cell* *36*, 932–941.
- Budzyński, M.A., Puustinen, M.C., Joutsen, J., and Sistonen, L. (2015). Uncoupling Stress-Inducible Phosphorylation of Heat Shock Factor 1 from Its Activation. *Mol. Cell. Biol.* *35*, 2530–2540.
- Cai, J., Yang, F., Zhan, H., Situ, J., Li, W., Mao, Y., and Luo, Y. (2019). RNA m<sup>6</sup>A Methyltransferase METTL3 Promotes The Growth Of Prostate Cancer By Regulating Hedgehog Pathway. *OTT Volume 12*, 9143–9152.
- Cairns, B.R. (2009). The logic of chromatin architecture and remodelling at promoters. *Nature* *461*, 193–198.
- Calderwood, S.K., and Gong, J. (2016). Heat Shock Proteins Promote Cancer: It's a Protection Racket. *Trends in Biochemical Sciences* *41*, 311–323.
- Cardiello, J.F., Goodrich, J.A., and Kugel, J.F. (2018). Heat Shock Causes a Reversible Increase in RNA Polymerase II Occupancy Downstream of mRNA Genes, Consistent with a Global Loss in Transcriptional Termination. *Mol Cell Biol* *38*, e00181-18, /mcb/38/18/e00181-18.atom.
- Carrillo Oesterreich, F., Preibisch, S., and Neugebauer, K.M. (2010). Global Analysis of Nascent RNA Reveals Transcriptional Pausing in Terminal Exons. *Molecular Cell* *40*, 571–581.

- Cartegni, L., Chew, S.L., and Krainer, A.R. (2002). Listening to silence and understanding nonsense: exonic mutations that affect splicing. *Nat Rev Genet* 3, 285–298.
- Chen, F.X., Smith, E.R., and Shilatifard, A. (2018). Born to run: control of transcription elongation by RNA polymerase II. *Nat Rev Mol Cell Biol* 19, 464–478.
- Chen, H.-M., Fitcher, B., and Leatherwood, J. (2011). The Fission Yeast RNA Binding Protein Mmi1 Regulates Meiotic Genes by Controlling Intron Specific Splicing and Polyadenylation Coupled RNA Turnover. *PLoS ONE* 6, e26804.
- Christians, E.S., Zhou, Q., Renard, J., and Benjamin, I.J. (2003). Heat shock proteins in mammalian development. *Seminars in Cell & Developmental Biology* 14, 283–290.
- Clapier, C.R., and Cairns, B.R. (2009). The Biology of Chromatin Remodeling Complexes. *Annu. Rev. Biochem.* 78, 273–304.
- Coker, H., Wei, G., and Brockdorff, N. (2019). m6A modification of non-coding RNA and the control of mammalian gene expression. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms* 1862, 310–318.
- Col, E., Houghoughi, N., Dufour, S., Penin, J., Koskas, S., Faure, V., Ouzounova, M., Hernandez-Vargash, H., Reynoird, N., Daujat, S., et al. (2017). Bromodomain factors of BET family are new essential actors of pericentric heterochromatin transcriptional activation in response to heat shock. *Sci Rep* 7, 5418.
- Colgan, D.F., and Manley, J.L. (1997). Mechanism and regulation of mRNA polyadenylation. *Genes & Development* 11, 2755–2766.
- Cotto, J.J., Fox, S.G., and Morimoto, R.I. (1997). HSF1 granules: a novel stress-induced nuclear compartment of human cells. 10.
- Crisp, P.A., Smith, A.B., Ganguly, D.R., Murray, K.D., Eichten, S.R., Millar, A.A., and Pogson, B.J. (2018). RNA Polymerase II Read-Through Promotes Expression of Neighboring Genes in SAL1-PAP-XRN Retrograde Signaling. *Plant Physiol.* 178, 1614–1630.
- Cui, Q., Shi, H., Ye, P., Li, L., Qu, Q., Sun, G., Sun, G., Lu, Z., Huang, Y., Yang, C.-G., et al. (2017). m<sup>6</sup>A RNA Methylation Regulates the Self-Renewal and Tumorigenesis of Glioblastoma Stem Cells. *Cell Reports* 18, 2622–2634.
- Dai, C., Whitesell, L., Rogers, A.B., and Lindquist, S. (2007). Heat Shock Factor 1 Is a Powerful Multifaceted Modifier of Carcinogenesis. *Cell* 130, 1005–1018.
- Denegri, M., Chiodi, I., Corioni, M., Cobianchi, F., Riva, S., and Biamonti, G. (2001). Stress-induced Nuclear Bodies Are Sites of Accumulation of Pre-mRNA Processing Factors. *MBoC* 12, 3502–3514.
- Deng, X., Chen, K., Luo, G.-Z., Weng, X., Ji, Q., Zhou, T., and He, C. (2015). Widespread occurrence of N<sup>6</sup>-methyladenosine in bacterial mRNA. *Nucleic Acids Res* 43, 6557–6567.

Desrosiers, R., Friderici, K., and Rottman, F. (1974). Identification of Methylated Nucleosides in Messenger RNA from Novikoff Hepatoma Cells. *Proceedings of the National Academy of Sciences* 71, 3971–3975.

Di Timoteo, G., Dattilo, D., Centrón-Broco, A., Colantoni, A., Guarnacci, M., Rossi, F., Incarnato, D., Oliviero, S., Fatica, A., Morlando, M., et al. (2020). Modulation of circRNA Metabolism by m6A Modification. *Cell Reports* 31, 107641.

Di Giammartino, D.C., Shi, Y., and Manley, J.L. (2013). PARP1 Represses PAP and Inhibits Polyadenylation during Heat Shock. *Molecular Cell* 49, 7–17.

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

Douglas, R.J., and Martin, K.A.C. (2009). Inhibition in cortical circuits. *Current Biology* 19, R398–R402.

Dow, E.C., Liu, H., and Rice, A.P. (2010). T-loop phosphorylated Cdk9 localizes to nuclear speckle domains which may serve as sites of active P-TEFb function and exchange between the Brd4 and 7SK/HEXIM1 regulatory complexes. *J. Cell. Physiol.* n/a-n/a.

Du, H., Zhao, Y., He, J., Zhang, Y., Xi, H., Liu, M., Ma, J., and Wu, L. (2016). YTHDF2 destabilizes m6A-containing RNA through direct recruitment of the CCR4–NOT deadenylase complex. *Nat Commun* 7, 12626.

Dubois, M. (1999). Heat shock of HeLa cells inactivates a nuclear protein phosphatase specific for dephosphorylation of the C-terminal domain of RNA polymerase II. *Nucleic Acids Research* 27, 1338–1344.

Eaton, J.D., Francis, L., Davidson, L., and West, S. (2020). A unified allosteric/torpedo mechanism for transcriptional termination on human protein-coding genes. *Genes Dev.* 34, 132–145.

Engel, M., Eggert, C., Kaplick, P.M., Eder, M., Röh, S., Tietze, L., Namendorf, C., Arloth, J., Weber, P., Rex-Haffner, M., et al. (2018). The Role of m6A/m-RNA Methylation in Stress Response Regulation. *Neuron* 99, 389-403.e9.

Epshtein, V., Cardinale, C.J., Ruckenstein, A.E., Borukhov, S., and Nudler, E. (2007). An Allosteric Path to Transcription Termination. *Molecular Cell* 28, 991–1001.

Eymery, A., Callanan, M., and Vourc'h, C. (2009). The secret message of heterochromatin: new insights into the mechanisms and function of centromeric and pericentric repeat sequence transcription. *Int. J. Dev. Biol.* 53, 259–268.

Fan, J., Kuai, B., Wu, G., Wu, X., Chi, B., Wang, L., Wang, K., Shi, Z., Zhang, H., Chen, S., et al. (2017). Exosome cofactor HMTR 4 competes with export adaptor ALYREF to ensure balanced nuclear RNA pools for degradation and export. *EMBO J* 36, 2870–2886.

- Fernandes, J., Acuña, S., Aoki, J., Floeter-Winter, L., and Muxel, S. (2019). Long Non-Coding RNAs in the Regulation of Gene Expression: Physiology and Disease. *NcRNA* 5, 17.
- Fica, S.M., and Nagai, K. (2017). Cryo-electron microscopy snapshots of the spliceosome: structural insights into a dynamic ribonucleoprotein machine. *Nat Struct Mol Biol* 24, 791–799.
- Fray, R.G., and Simpson, G.G. (2015). The Arabidopsis epitranscriptome. *Current Opinion in Plant Biology* 27, 17–21.
- Fritah, S., Col, E., Boyault, C., Govin, J., Sadoul, K., Chiocca, S., Christians, E., Khochbin, S., Jolly, C., and Vourc'h, C. (2009). Heat-Shock Factor 1 Controls Genome-wide Acetylation in Heat-shocked Cells. *MBoC* 20, 4976–4984.
- Fu, Y., and Zhuang, X. (2020). m6A-binding YTHDF proteins promote stress granule formation. *Nat Chem Biol*.
- Fu, Y., Jia, G., Pang, X., Wang, R.N., Wang, X., Li, C.J., Smemo, S., Dai, Q., Bailey, K.A., Nobrega, M.A., et al. (2013). FTO-mediated formation of N6-hydroxymethyladenosine and N6-formyladenosine in mammalian RNA. *Nat Commun* 4, 1798.
- Fujimoto, M., Takaki, E., Takii, R., Tan, K., Prakasam, R., Hayashida, N., Iemura, S., Natsume, T., and Nakai, A. (2012). RPA Assists HSF1 Access to Nucleosomal DNA by Recruiting Histone Chaperone FACT. *Molecular Cell* 48, 182–194.
- Fustin, J.-M., Doi, M., Yamaguchi, Y., Hida, H., Nishimura, S., Yoshida, M., Isagawa, T., Morioka, M.S., Kakeya, H., Manabe, I., et al. (2013). RNA-Methylation-Dependent RNA Processing Controls the Speed of the Circadian Clock. *Cell* 155, 793–806.
- Galganski, L., Urbanek, M.O., and Krzyzosiak, W.J. (2017). Nuclear speckles: molecular organization, biological function and role in disease. *Nucleic Acids Research* 45, 10350–10368.
- Gebert, L.F.R., and MacRae, I.J. (2019). Regulation of microRNA function in animals. *Nat Rev Mol Cell Biol* 20, 21–37.
- Geula, S., Moshitch-Moshkovitz, S., Dominissini, D., Mansour, A.A., Kol, N., Salmon-Divon, M., Hershkovitz, V., Peer, E., Mor, N., Manor, Y.S., et al. (2015). m<sup>6</sup>A mRNA methylation facilitates
- Goenka, A., Sengupta, S., Pandey, R., Parihar, R., Mohanta, G.C., Mukerji, M., and Ganesh, S. (2016a). Human satellite-III non-coding RNAs modulate heat-shock-induced transcriptional repression. *J Cell Sci* 129, 3541–3552.
- Goenka, A., Sengupta, S., Pandey, R., Parihar, R., Mohanta, G.C., Mukerji, M., and Ganesh, S. (2016). Human satellite-III non-coding RNAs modulate heat-shock-induced transcriptional repression. *J Cell Sci* 129, 3541–3552.
- Gokhale, N.S., McIntyre, A.B.R., McFadden, M.J., Roder, A.E., Kennedy, E.M., Gandara, J.A., Hopcraft, S.E., Quicke, K.M., Vazquez, C., Willer, J., et al. (2016). N<sup>6</sup>-Methyladenosine in Flaviviridae Viral RNA Genomes Regulates Infection. *Cell Host & Microbe* 20, 654–665.

- Gonatopoulos-Pournatzis, T., and Cowling, V.H. (2014). Cap-binding complex (CBC). *Biochemical Journal* 457, 231–242.
- Gonzalez, I., Munita, R., Agirre, E., Dittmer, T.A., Gysling, K., Misteli, T., and Luco, R.F. (2015). A lncRNA regulates alternative splicing via establishment of a splicing-specific chromatin signature. *Nat Struct Mol Biol* 22, 370–376.
- Gothard, L.Q., Ruffner, M.E., Woodward, J.G., Park-Sarge, O.-K., and Sarge, K.D. (2003). Lowered Temperature Set Point for Activation of the Cellular Stress Response in T-lymphocytes. *J. Biol. Chem.* 278, 9322–9326.
- Grewal, S.I.S., and Jia, S. (2007). Heterochromatin revisited. *Nat Rev Genet* 8, 35–46.
- Gurova, K., Chang, H.-W., Valieva, M.E., Sandlesh, P., and Studitsky, V.M. (2018). Structure and function of the histone chaperone FACT – Resolving FACTual issues. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms* 1861, 892–904.
- Hata, M., Twai, T., Hayashi, Y., Ueda, M., and Ohtsuka, K. (1998). Stresses and Tissue Specific Expression. 9.
- Hattendorf, D.A. (2002). Cooperative kinetics of both Hsp104 ATPase domains and interdomain communication revealed by AAA sensor-1 mutants. *The EMBO Journal* 21, 12–21.
- Hilleren, P., McCarthy, T., Rosbash, M., Parker, R., and Jensen, T.H. (2001). Quality control of mRNA 3'-end processing is linked to the nuclear exosome. *Nature* 413, 538–542.
- Holcik, M., and Sonenberg, N. (2005). Translational control in stress and apoptosis. *Nat Rev Mol Cell Biol* 6, 318–327.
- Horiuchi, K., Kawamura, T., Iwanari, H., Ohashi, R., Naito, M., Kodama, T., and Hamakubo, T. (2013). Identification of Wilms' Tumor 1-associating Protein Complex and Its Role in Alternative Splicing and the Cell Cycle. *J. Biol. Chem.* 288, 33292–33302.
- Huang, H., Weng, H., Sun, W., Qin, X., Shi, H., Wu, H., Zhao, B.S., Mesquita, A., Liu, C., Yuan, C.L., et al. (2018). Recognition of RNA N6-methyladenosine by IGF2BP proteins enhances mRNA stability and translation. *Nat Cell Biol* 20, 285–295.
- Huang, Y., Gattoni, R., Stévenin, J., and Steitz, J.A. (2003). SR Splicing Factors Serve as Adapter Proteins for TAP-Dependent mRNA Export. *Molecular Cell* 11, 837–843.
- Hussong, M., Kaehler, C., Kerick, M., Grimm, C., Franz, A., Timmermann, B., Welzel, F., Isensee, J., Hucho, T., Krobitsch, S., et al. (2017). The bromodomain protein BRD4 regulates splicing during heat shock. *Nucleic Acids Res* 45, 382–394.
- Jantschitsch, C., and Trautinger, F. (2003). Heat shock and UV-B-induced DNA damage and mutagenesis in skin. *Photochem. Photobiol. Sci.* 2, 899.

- Jia, G., Fu, Y., Zhao, X., Dai, Q., Zheng, G., Yang, Y., Yi, C., Lindahl, T., Pan, T., Yang, Y.-G., et al. (2011). N6-Methyladenosine in nuclear RNA is a major substrate of the obesity-associated FTO. *Nat Chem Biol* 7, 885–887.
- Jolly, C., Morimoto, R.I., Robert-Nicoud, M., and Vourc'h, C. (1997). HSF1 transcription factor concentrates in nuclear foci during heat shock: relationship with transcription sites. 7.
- Jolly, C., Konecny, L., Grady, D.L., Kutsikova, Y.A., Cotto, J.J., Morimoto, R.I., and Vourc'h, C. (2002). In vivo binding of active heat shock transcription factor 1 to human chromosome 9 heterochromatin during stress. *Journal of Cell Biology* 156, 775–781.
- Jolly, C., Metz, A., Govin, J., Vigneron, M., Turner, B.M., Khochbin, S., and Vourc'h, C. (2004). Stress-induced transcription of satellite III repeats. *Journal of Cell Biology* 164, 25–33.
- Jolly & Morimoto, C. (2000). Role of the Heat Shock Response and Molecular Chaperones in Oncogenesis and Cell Death. *Journal of the National Cancer Institute* 92, 1564–1572.
- Kachaev, Z.M., Lebedeva, L.A., Kozlov, E.N., and Shidlovskii, Y.V. (2020). Interplay of mRNA capping and transcription machineries. *Bioscience Reports* 40, BSR20192825.
- Kadumuri, R.V., and Janga, S.C. (2018). Epitranscriptomic Code and Its Alterations in Human Disease. *Trends in Molecular Medicine* 24, 886–903.
- Kamieniarz-Gdula, K., and Proudfoot, N.J. (2019). Transcriptional Control by Premature Termination: A Forgotten Mechanism. *Trends in Genetics* 35, 553–564.
- Kampinga, H.H., and Craig, E.A. (2010). The HSP70 chaperone machinery: J proteins as drivers of functional specificity. *Nat Rev Mol Cell Biol* 11, 579–592.
- Kan, L., Grozhik, A.V., Vedanayagam, J., Patil, D.P., Pang, N., Lim, K.-S., Huang, Y.-C., Joseph, B., Lin, C.-J., Despic, V., et al. (2017). The m6A pathway facilitates sex determination in *Drosophila*. *Nat Commun* 8, 15737.
- Kasowitz, S.D., Ma, J., Anderson, S.J., Leu, N.A., Xu, Y., Gregory, B.D., Schultz, R.M., and Wang, P.J. (2018). Nuclear m6A reader YTHDC1 regulates alternative polyadenylation and splicing during mouse oocyte development. *PLoS Genet* 14, e1007412.
- Katahira, J. (2012). mRNA export and the TREX complex. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms* 1819, 507–513.
- Katahira, J., Inoue, H., Hurt, E., and Yoneda, Y. (2009). Adaptor Aly and co-adaptor Thoc5 function in the Tap-p15-mediated nuclear export of HSP70 mRNA. *EMBO J* 28, 556–567.
- Katahira, J., Ishikawa, H., Tsujimura, K., Kurono, S., and Hieda, M. (2019). Human THO coordinates transcription termination and subsequent transcript release from the *HSP70* locus. *Genes Cells* 24, 272–283.

Ke, S., Alemu, E.A., Mertens, C., Gantman, E.C., Fak, J.J., Mele, A., Haripal, B., Zucker-Scharff, I., Moore, M.J., Park, C.Y., et al. (2015). A majority of m<sup>6</sup>A residues are in the last exons, allowing the potential for 3' UTR regulation. *Genes Dev.* *29*, 2037–2053.

Ke, S., Pandya-Jones, A., Saito, Y., Fak, J.J., Vågbø, C.B., Geula, S., Hanna, J.H., Black, D.L., Darnell, J.E., and Darnell, R.B. (2017). m<sup>6</sup>A mRNA modifications are deposited in nascent pre-mRNA and are not required for splicing but do specify cytoplasmic turnover. *Genes Dev.* *31*, 990–1006.

Kelley, P.M., and Schlesinger, M.J. (1978). The effect of amino acid analogues and heat shock on gene expression in chicken embryo fibroblasts. *Cell* *15*, 1277–1286.

Khalil, A.M., Guttman, M., Huarte, M., Garber, M., Raj, A., Rivea Morales, D., Thomas, K., Presser, A., Bernstein, B.E., van Oudenaarden, A., et al. (2009). Many human large intergenic noncoding RNAs associate with chromatin-modifying complexes and affect gene expression. *PNAS* *106*, 11667–11672.

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

Kilchert, C., Wittmann, S., Passoni, M., Shah, S., Granneman, S., and Vasiljeva, L. (2015). Regulation of mRNA Levels by Decay-Promoting Introns that Recruit the Exosome Specificity Factor Mmi1. *Cell Reports* *13*, 2504–2515.

Kim, Y.K., and Jang, S.K. (2002). Continuous heat shock enhances translational initiation directed by internal ribosomal entry site. *Biochemical and Biophysical Research Communications* *297*, 224–231.

Kim, J., Han, K.Y., Khanna, N., Ha, T., and Belmont, A.S. (2019). Nuclear speckle fusion via long-range directional motion regulates speckle morphology after transcriptional inhibition. *J Cell Sci* *132*, jcs226563.

Kim, K., Heo, D., Kim, I., Suh, J.-Y., and Kim, M. (2016). Exosome Cofactors Connect Transcription Termination to RNA Processing by Guiding Terminated Transcripts to the Appropriate Exonuclease within the Nuclear Exosome. *J. Biol. Chem.* *291*, 13229–13242.

Kim, Y.E., Hipp, M.S., Bracher, A., Hayer-Hartl, M., and Ulrich Hartl, F. (2013). Molecular Chaperone Functions in Protein Folding and Proteostasis. *Annu. Rev. Biochem.* *82*, 323–355.

Knuckles, P., Carl, S.H., Musheev, M., Niehrs, C., Wenger, A., and Bühler, M. (2017). RNA fate determination through cotranscriptional adenosine methylation and microprocessor binding. *Nat Struct Mol Biol* *24*, 561–569.

Knuckles, P., Lence, T., Haussmann, I.U., Jacob, D., Kreim, N., Carl, S.H., Masiello, I., Hares, T., Villaseñor, R., Hess, D., et al. (2018). Zc3h13/Flacc is required for adenosine methylation by bridging the mRNA-binding factor Rbm15/Spenito to the m<sup>6</sup>A machinery component Wtap/Fl(2)d. *Genes Dev.* *32*, 415–429.



- Konkel, M.K., and Batzer, M.A. (2010). A mobile threat to genome stability: The impact of non-LTR retrotransposons upon the human genome. *Seminars in Cancer Biology* 20, 211–221.
- Kornberg, R.D. (1977). *Structure of Chromatin*. 26.
- Kouzarides, T. (2007). Chromatin Modifications and Their Function. *Cell* 128, 693–705.
- Kraynik, S.M., Gabanic, A., Anthony, S.R., Kelley, M., Paulding, W.R., Roessler, A., McGuinness, M., and Tranter, M. (2015). The stress-induced heat shock protein 70.3 expression is regulated by a dual-component mechanism involving alternative polyadenylation and HuR. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms* 1849, 688–696.
- Kültz, D. (2005). MOLECULAR AND EVOLUTIONARY BASIS OF THE CELLULAR STRESS RESPONSE. *Annu. Rev. Physiol.* 67, 225–257.
- LaCava, J., Houseley, J., Saveanu, C., Petfalski, E., Thompson, E., Jacquier, A., and Tollervey, D. (2005). RNA Degradation by the Exosome Is Promoted by a Nuclear Polyadenylation Complex. *Cell* 121, 713–724.
- Lama, D., and Verma, C.S. (2020). Deciphering the mechanistic effects of eIF4E phosphorylation on mRNA-cap recognition. *Protein Science* 29, 1373–1386.
- Latham, J.A., and Dent, S.Y.R. (2007). Cross-regulation of histone modifications. *Nat Struct Mol Biol* 14, 1017–1024.
- Lemaux, P.G., Herendeen, S.L., Bloch, P.L., and Neidhardt, F.C. (1978). Transient rates of synthesis of individual polypeptides in *E. coli* following temperature shifts. *Cell* 13, 427–434.
- Leppek, K., Das, R., and Barna, M. (2018). Functional 5' UTR mRNA structures in eukaryotic translation regulation and how to find them. *Nat Rev Mol Cell Biol* 19, 158–174.
- Lesbirel, S., Viphakone, N., Parker, M., Parker, J., Heath, C., Sudbery, I., and Wilson, S.A. (2018). The m6A-methylase complex recruits TREX and regulates mRNA export. *Sci Rep* 8, 13827.
- Li, G., and Reinberg, D. (2011). Chromatin higher-order structures and gene regulation. *Current Opinion in Genetics & Development* 21, 175–186.
- Li, X., Xu, B., Li, X., Wang, D., An, Z., Jia, Y., Niu, J., Gao, J., Zhang, M.Q., Li, F., et al. (2019). Widespread transcriptional responses to the thermal stresses are prewired in human 3D genome (Genomics).
- Li, Z., Huang, C., Bao, C., Chen, L., Lin, M., Wang, X., Zhong, G., Yu, B., Hu, W., Dai, L., et al. (2015). Exon-intron circular RNAs regulate transcription in the nucleus. *Nat Struct Mol Biol* 22, 256–264.
- Liao, S., Sun, H., and Xu, C. (2018). YTH Domain: A Family of N<sup>6</sup>-methyladenosine (m<sup>6</sup>A) Readers. *Genomics, Proteomics & Bioinformatics* 16, 99–107.

- Liebelt, F., Sebastian, R.M., Moore, C.L., Mulder, M.P.C., Ovaa, H., Shoulders, M.D., and Vertegaal, A.C.O. (2019). SUMOylation and the HSF1-Regulated Chaperone Network Converge to Promote Proteostasis in Response to Heat Shock. *Cell Reports* 26, 236-249.e4.
- Lin, S., Choe, J., Du, P., Triboulet, R., and Gregory, R.I. (2016). The m<sup>6</sup>A Methyltransferase METTL3 Promotes Translation in Human Cancer Cells. *Molecular Cell* 62, 335–345.
- Lin, Z., Hsu, P.J., Xing, X., Fang, J., Lu, Z., Zou, Q., Zhang, K.-J., Zhang, X., Zhou, Y., Zhang, T., et al. (2017). Mettl3-/Mettl14-mediated mRNA N<sup>6</sup>-methyladenosine modulates murine spermatogenesis. *Cell Res* 27, 1216–1230.
- Linder, B., Grozhik, A.V., Olarerin-George, A.O., Meydan, C., Mason, C.E., and Jaffrey, S.R. (2015). Single-nucleotide-resolution mapping of m<sup>6</sup>A and m<sup>6</sup>Am throughout the transcriptome. *Nat Methods* 12, 767–772.
- Lindquist, S. (1986). THE HEAT-SHOCK RESPONSE. 44.
- Liu, Z., and Zhang, J. (2018). Most m<sup>6</sup>A RNA Modifications in Protein-Coding Regions Are Evolutionarily Unconserved and Likely Nonfunctional. *Molecular Biology and Evolution* 35, 666–675.
- Liu, J., Yue, Y., Han, D., Wang, X., Fu, Y., Zhang, L., Jia, G., Yu, M., Lu, Z., Deng, X., et al. (2014). A METTL3–METTL14 complex mediates mammalian nuclear RNA N<sup>6</sup>-adenosine methylation. *Nat Chem Biol* 10, 93–95.
- Liu, J., Dou, X., Chen, C., Chen, C., Liu, C., Xu, M.M., Zhao, S., Shen, B., Gao, Y., Han, D., et al. (2020). N<sup>6</sup>-methyladenosine of chromosome-associated regulatory RNA regulates chromatin state and transcription. *Science* 367, 580–586.
- Liu, N., Dai, Q., Zheng, G., He, C., Parisien, M., and Pan, T. (2015). N<sup>6</sup>-methyladenosine-dependent RNA structural switches regulate RNA–protein interactions. *Nature* 518, 560–564.
- Liu, Z.-X., Li, L.-M., Sun, H.-L., and Liu, S.-M. (2018). Link Between m<sup>6</sup>A Modification and Cancers. *Front. Bioeng. Biotechnol.* 6, 89.
- Long, Y., Wang, X., Youmans, D.T., and Cech, T.R. (2017). How do lncRNAs regulate transcription? *Sci. Adv.* 3, eaao2110.
- Long, Y., Sou, W.H., Yung, K.W.Y., Liu, H., Wan, S.W.C., Li, Q., Zeng, C., Law, C.O.K., Chan, G.H.C., Lau, T.C.K., et al. (2019). Distinct mechanisms govern the phosphorylation of different SR protein splicing factors. *J. Biol. Chem.* 294, 1312–1327.
- Louis, I.V.-St., and Sagarsky, C. (2018). Mammalian Cis-Acting RNA Sequence Elements. In *Gene Expression and Regulation in Mammalian Cells - Transcription From General Aspects*, F. Uchiumi, ed. (InTech), p.

- Ma, H., Wang, X., Cai, J., Dai, Q., Natchiar, S.K., Lv, R., Chen, K., Lu, Z., Chen, H., Shi, Y.G., et al. (2019). N6-Methyladenosine methyltransferase ZCCHC4 mediates ribosomal RNA methylation. *Nat Chem Biol* *15*, 88–94.
- Ma, L., Zhao, B., Chen, K., Thomas, A., Tuteja, J.H., He, X., He, C., and White, K.P. (2017). Evolution of transcript modification by N<sup>6</sup>-methyladenosine in primates. *Genome Res.* *27*, 385–392.
- Mahat, D.B., Salamanca, H.H., Duarte, F.M., Danko, C.G., and Lis, J.T. (2016). Mammalian Heat Shock Response and Mechanisms Underlying Its Genome-wide Transcriptional Regulation. *Molecular Cell* *62*, 63–78.
- Malmendal, A., Overgaard, J., Bundy, J.G., Sørensen, J.G., Nielsen, N.Ch., Loeschcke, V., and Holmstrup, M. (2006). Metabolomic profiling of heat stress: hardening and recovery of homeostasis in *Drosophila*. *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology* *291*, R205–R212.
- Mao, Y., Dong, L., Liu, X.-M., Guo, J., Ma, H., Shen, B., and Qian, S.-B. (2019). m6A in mRNA coding regions promotes translation via the RNA helicase-containing YTHDC2. *Nat Commun* *10*, 5332.
- Marianna Penzo, Ania Guerrieri, Federico Zacchini, Davide Treré, and Lorenzo Montanaro (2017). RNA Pseudouridylation in Physiology and Medicine: For Better and for Worse. *Genes* *8*, 301.
- Mariner, P.D., Walters, R.D., Espinoza, C.A., Drullinger, L.F., Wagner, S.D., Kugel, J.F., and Goodrich, J.A. (2008). Human Alu RNA Is a Modular Transacting Repressor of mRNA Transcription during Heat Shock. *Molecular Cell* *29*, 499–509.
- Martin, K.C., and Ephrussi, A. (2010). mRNA Localization: Gene Expression in the Spatial Dimension. *21*.
- Mauer, J., Luo, X., Blanjoie, A., Jiao, X., Grozhik, A.V., Patil, D.P., Linder, B., Pickering, B.F., Vasseur, J.-J., Chen, Q., et al. (2017). Reversible methylation of m6Am in the 5' cap controls mRNA stability. *Nature* *541*, 371–375.
- Mayer, M.P. (2010). Gymnastics of Molecular Chaperones. *Molecular Cell* *39*, 321–331.
- Mendillo, M.L., Santagata, S., Koeva, M., Bell, G.W., Hu, R., Tamimi, R.M., Fraenkel, E., Ince, T.A., Whitesell, L., and Lindquist, S. (2012). HSF1 Drives a Transcriptional Program Distinct from Heat Shock to Support Highly Malignant Human Cancers. *Cell* *150*, 549–562.
- Metz, A. (2004). A key role for stress-induced satellite III transcripts in the relocalization of splicing factors into nuclear stress granules. *Journal of Cell Science* *117*, 4551–4558.
- Meyer, K.D., Saletore, Y., Zumbo, P., Elemento, O., Mason, C.E., and Jaffrey, S.R. (2012). Comprehensive Analysis of mRNA Methylation Reveals Enrichment in 3' UTRs and near Stop Codons. *Cell* *149*, 1635–1646.

- Meyer, K.D., Patil, D.P., Zhou, J., Zinoviev, A., Skabkin, M.A., Elemento, O., Pestova, T.V., Qian, S.-B., and Jaffrey, S.R. (2015). 5' UTR m6A Promotes Cap-Independent Translation. *Cell* *163*, 999–1010.
- Michels, A.A., Kanon, B., Konings, A.W.T., Ohtsuka, K., Bensaude, O., and Kampinga, H.H. (1997). Hsp70 and Hsp40 Chaperone Activities in the Cytoplasm and the Nucleus of Mammalian Cells. *J. Biol. Chem.* *272*, 33283–33289.
- Miozzo, F., Sabéran-Djoneidi, D., and Mezger, V. (2015). HSFs, Stress Sensors and Sculptors of Transcription Compartments and Epigenetic Landscapes. *Journal of Molecular Biology* *427*, 3793–3816.
- Moran, L.A. (1986). Molecular Cloning and Analysis of DNA Complementary to Three Mouse M, = 68,000 Heat Shock Protein mRNAs. *11*.
- Morimoto, R.I. (1998). Regulation of the heat shock transcriptional response: cross talk between a family of heat shock factors, molecular chaperones, and negative regulators. *Genes & Development* *12*, 3788–3796.
- Motamedi, M.R., Verdel, A., Colmenares, S.U., Gerber, S.A., Gygi, S.P., and Moazed, D. (2004). Two RNAi Complexes, RITS and RDRC, Physically Interact and Localize to Noncoding Centromeric RNAs. *Cell* *119*, 789–802.
- Mozzetta, C., Pontis, J., and Ait-Si-Ali, S. (2015). Functional Crosstalk Between Lysine Methyltransferases on Histone Substrates: The Case of G9A/GLP and Polycomb Repressive Complex 2. *Antioxidants & Redox Signaling* *22*, 1365–1381.
- Muchowski, P.J., and Wacker, J.L. (2005). Modulation of neurodegeneration by molecular chaperones. *Nat Rev Neurosci* *6*, 11–22.
- Murray, J.I., Whitfield, M.L., Trinklein, N.D., Myers, R.M., Brown, P.O., and Botstein, D. (2004). Diverse and Specific Gene Expression Responses to Stresses in Cultured Human Cells. *Molecular Biology of the Cell* *15*, 14.
- Nayler, O., Hartmann, A.M., and Stamm, S. (2000). The ER Repeat Protein Yt521-B Localizes to a Novel Subnuclear Compartment. *Journal of Cell Biology* *150*, 949–962.
- Nesse, R.M., Bhatnagar, S., and Ellis, B. (2016). Evolutionary Origins and Functions of the Stress Response System. In *Stress: Concepts, Cognition, Emotion, and Behavior*, (Elsevier), pp. 95–101.
- Newton, E.M., Knauf, U., Green, M., and Kingston, R.E. (1996). The regulatory domain of human heat shock factor 1 is sufficient to sense heat stress. *Mol. Cell. Biol.* *16*, 839–846.
- Ninomiya, K., Kataoka, N., and Hagiwara, M. (2011). Stress-responsive maturation of Clk1/4 pre-mRNAs promotes phosphorylation of SR splicing factor. *The Journal of Cell Biology* *195*, 27–40.

- Ninomiya, K., Adachi, S., Natsume, T., Iwakiri, J., Terai, G., Asai, K., and Hirose, T. (2020). Lnc RNA -dependent nuclear stress bodies promote intron retention through SR protein phosphorylation. *EMBO J* 39.
- Ohtsuka, K., and Hata, M. (2000). Molecular chaperone function of mammalian Hsp70 and Hsp40—a review. *International Journal of Hyperthermia* 16, 231–245.
- Olins, D.E., and Olins, A.L. (2003). Chromatin history: our view from the bridge. *Nat Rev Mol Cell Biol* 4, 809–814.
- Oren, Y.S., McClure, M.L., Rowe, S.M., Sorscher, E.J., Bester, A.C., Manor, M., Kerem, E., Rivlin, J., Zahdeh, F., Mann, M., et al. (2014). The unfolded protein response affects readthrough of premature termination codons. *EMBO Mol Med* 6, 685–701.
- Pandit, S., Wang, D., and Fu, X.-D. (2008). Functional integration of transcriptional and RNA processing machineries. *Current Opinion in Cell Biology* 20, 260–265.
- Panniers, R. (1994). Translational control during heat shock. *Biochimie* 76, 737–747.
- Paparidis, N.F. dos S., Durvale, M.C., and Canduri, F. (2017). The emerging picture of CDK9/P-TEFb: more than 20 years of advances since PITALRE. *Mol. BioSyst.* 13, 246–276.
- Patil, D.P., Chen, C.-K., Pickering, B.F., Chow, A., Jackson, C., Guttman, M., and Jaffrey, S.R. (2016). m6A RNA methylation promotes XIST-mediated transcriptional repression. *Nature* 537, 369–373.
- Patriarca & Maresca, E. (1990). Thermotolerance following Heat Shock Prevents impairment of Mitochondrial Alleviated Temperatures in Saccharom. 8.
- Pauli, D., Arrigo, A.-P., and Tissières, A. (1992). Heat shock response in *Drosophila*. *Experientia* 48, 623–629.
- Pelham, H.R.B. (1982). A regulatory upstream promoter element in the *Drosophila* Hsp 70 heat-shock gene. *Cell* 30, 517–528.
- Pendleton, K.E., Chen, B., Liu, K., Hunter, O.V., Xie, Y., Tu, B.P., and Conrad, N.K. (2017). The U6 snRNA m6A Methyltransferase METTL16 Regulates SAM Synthetase Intron Retention. *Cell* 169, 824–835.e14.
- Pernet, L., Faure, V., Gilquin, B., Dufour-Guérin, S., Khochbin, S., and Vourc'h, C. (2014). HDAC6–ubiquitin interaction controls the duration of HSF1 activation after heat shock. *MBoC* 25, 4187–4194.
- Peterlin, B.M., Brogie, J.E., and Price, D.H. (2012). 7SK snRNA: a noncoding RNA that plays a major role in regulating eukaryotic transcription: 7SK: a regulatory snRNA. *WIREs RNA* 3, 92–103.

- Poljšak, B., and Milisav, I. (2012). Clinical implications of cellular stress responses. *Bosn J of Basic Med Sci* *12*, 122.
- Proudfoot, N.J. (2016). Transcriptional termination in mammals: Stopping the RNA polymerase II juggernaut. *Science* *352*, aad9926–aad9926.
- Rafalska, I., Zhang, Z., Benderska, N., Wolff, H., Hartmann, A.M., Brack-Werner, R., and Stamm, S. (2004). The intranuclear localization and function of YT521-B is regulated by tyrosine phosphorylation. *Human Molecular Genetics* *13*, 1535–1549.
- Rashid, F., Shah, A., and Shan, G. (2016). Long Non-coding RNAs in the Cytoplasm. *Genomics, Proteomics & Bioinformatics* *14*, 73–80.
- Rhoads, R.E., and Lamphear, B.J. (1995). Cap-Independent Translation of Heat Shock Messenger RNAs. In *Cap-Independent Translation*, P. Sarnow, ed. (Berlin, Heidelberg: Springer Berlin Heidelberg), pp. 131–153.
- Richter, K., Haslbeck, M., and Buchner, J. (2010). The Heat Shock Response: Life on the Verge of Death. *Molecular Cell* *40*, 253–266.
- Ries, R.J., Zaccara, S., Klein, P., Olarerin-George, A., Namkoong, S., Pickering, B.F., Patil, D.P., Kwak, H., Lee, J.H., and Jaffrey, S.R. (2019). m6A enhances the phase separation potential of mRNA. *Nature* *571*, 424–428.
- Rinn, J.L., and Chang, H.Y. (2012). Genome Regulation by Long Noncoding RNAs. *Annu. Rev. Biochem.* *81*, 145–166.
- Ritossa, F. (1962). A new puffing pattern induced by temperature shock and DNP in drosophila. *Experientia* *18*, 571–573.
- Rizzi, N., Denegri, M., Chiodi, I., Corioni, M., Valgardsdottir, R., Cobianchi, F., Riva, S., and Biamonti, G. (2004). Transcriptional Activation of a Constitutive Heterochromatic Domain of the Human Genome in Response to Heat Shock□D. *Molecular Biology of the Cell* *15*, 9.
- Rothrock, C.R., House, A.E., and Lynch, K.W. (2005). HnRNP L represses exon splicing via a regulated exonic splicing silencer. *EMBO J* *24*, 2792–2802.
- Rougvie, A.E., and Lis, J.T. (1991). The RNA Polymerase II Molecule at the 5' End of the Uninduced hspTOGene of *D. melanogaster* Is Transcriptionally Engaged. *10*.
- Roundtree, I.A., Luo, G.-Z., Zhang, Z., Wang, X., Zhou, T., Cui, Y., Sha, J., Huang, X., Guerrero, L., Xie, P., et al. (2017). YTHDC1 mediates nuclear export of N6-methyladenosine methylated mRNAs. *ELife* *6*, e31311.
- Roy, B., and Jacobson, A. (2013). The intimate relationships of mRNA decay and translation. *Trends in Genetics* *29*, 691–699.

Salzman, J., Gawad, C., Wang, P.L., Lacayo, N., and Brown, P.O. (2012). Circular RNAs Are the Predominant Transcript Isoform from Hundreds of Human Genes in Diverse Cell Types. *PLoS ONE* 7, e30733.

Sánchez-Vásquez, E., Alata Jimenez, N., Vázquez, N.A., and Strobl-Mazzulla, P.H. (2018). Emerging role of dynamic RNA modifications during animal development. *Mechanisms of Development* 154, 24–32.

Sarge, K.D., Murphy, S.P., and Morimoto, R.I. (1993). Activation of Heat Shock Gene Transcription by Heat Shock Factor 1 Involves Oligomerization, Acquisition of DNA-Binding Activity, and Nuclear Localization and Can Occur in the Absence of Stress. *MOL. CELL. BIOL.* 13, 16.

Schwartz, S., Agarwala, S.D., Mumbach, M.R., Jovanovic, M., Mertins, P., Shishkin, A., Tabach, Y., Mikkelsen, T.S., Satija, R., Ruvkun, G., et al. (2013). High-Resolution Mapping Reveals a Conserved, Widespread, Dynamic mRNA Methylation Program in Yeast Meiosis. *Cell* 155, 1409–1421.

Schwartz, S., Mumbach, M.R., Jovanovic, M., Wang, T., Maciag, K., Bushkin, G.G., Mertins, P., Ter-Ovanesyan, D., Habib, N., Cacchiarelli, D., et al. (2014). Perturbation of m6A Writers Reveals Two Distinct Classes of mRNA Methylation at Internal and 5' Sites. *Cell Reports* 8, 284–296.

Shalgi, R., Hurt, J.A., Lindquist, S., and Burge, C.B. (2014). Widespread Inhibition of Posttranscriptional Splicing Shapes the Cellular Transcriptome following Heat Shock. *Cell Reports* 7, 1362–1370.

Shamovsky, I., and Nudler, E. (2009). Isolation and Characterization of the Heat Shock RNA 1. In *Riboswitches*, A. Serganov, ed. (Totowa, NJ: Humana Press), pp. 265–279.

Shi, H., Wang, X., Lu, Z., Zhao, B.S., Ma, H., Hsu, P.J., Liu, C., and He, C. (2017). YTHDF3 facilitates translation and decay of N6-methyladenosine-modified RNA. *Cell Res* 27, 315–328.

Shi, Y., Mosser, D.D., and Morimoto, R.I. (1998). Molecular chaperones as HSF1-specific transcriptional repressors. *Genes & Development* 12, 654–666.

Shima, H., Matsumoto, M., Ishigami, Y., Ebina, M., Muto, A., Sato, Y., Kumagai, S., Ochiai, K., Suzuki, T., and Igarashi, K. (2017). S-Adenosylmethionine Synthesis Is Regulated by Selective N6-Adenosine Methylation and mRNA Degradation Involving METTL16 and YTHDC1. *Cell Reports* 21, 3354–3363.

Sikorski, T.W., and Buratowski, S. (2009). The basal initiation machinery: beyond the general transcription factors. *Current Opinion in Cell Biology* 21, 344–351.

Skaggs, H.S., Xing, H., Wilkerson, D.C., Murphy, L.A., Hong, Y., Mayhew, C.N., and Sarge, K.D. (2007). HSF1-TPR Interaction Facilitates Export of Stress-induced HSP70 mRNA. *Journal of Biological Chemistry* 282, 33902–33907.

Srijyothi, L., Ponne, S., Prathama, T., Ashok, C., and Baluchamy, S. (2018). Roles of Non-Coding RNAs in Transcriptional Regulation. In *Transcriptional and Post-Transcriptional Regulation*, K. Ghedira, ed. (InTech), p.

St. Laurent, G., Wahlestedt, C., and Kapranov, P. (2015). The Landscape of long noncoding RNA classification. *Trends in Genetics* *31*, 239–251.

Strahl, B.D., and Allis, C.D. (2000). The language of covalent histone modifications. *403*, 5.

Strålfors, A., and Ekwall, K. (2011). Heterochromatin and Euchromatin-Organization, Boundaries, and Gene Regulation. In *Encyclopedia of Molecular Cell Biology and Molecular Medicine*, R.A. Meyers, ed. (Weinheim, Germany: Wiley-VCH Verlag GmbH & Co. KGaA), p. mcb.200400018.pub2.

Sullivan, E.K., Weirich, C.S., Guyon, J.R., Sif, S., and Kingston, R.E. (2001). Transcriptional Activation Domains of Human Heat Shock Factor 1 Recruit Human SWI/SNF. *Mol. Cell. Biol.* *21*, 5826–5837.

Tabuchi, Y., Takasaki, I., Wada, S., Zhao, Q.-L., Hori, T., Nomura, T., Ohtsuka, K., and Kondo, T. (2008). Genes and genetic networks responsive to mild hyperthermia in human lymphoma U937 cells. *International Journal of Hyperthermia* *24*, 613–622.

Tan, M., Luo, H., Lee, S., Jin, F., Yang, J.S., Montellier, E., Buchou, T., Cheng, Z., Rousseaux, S., Rajagopal, N., et al. (2011). Identification of 67 Histone Marks and Histone Lysine Crotonylation as a New Type of Histone Modification. *Cell* *146*, 1016–1028.

Tang, C., Klukovich, R., Peng, H., Wang, Z., Yu, T., Zhang, Y., Zheng, H., Klungland, A., and Yan, W. (2018). ALKBH5-dependent m6A demethylation controls splicing and stability of long 3'-UTR mRNAs in male germ cells. *Proc Natl Acad Sci USA* *115*, E325–E333.

Taniguchi, Y. (2016). The Bromodomain and Extra-Terminal Domain (BET) Family: Functional Anatomy of BET Paralogous Proteins. *IJMS* *17*, 1849.

Tellier, M., Maudlin, I., and Murphy, S. (2020). Transcription and splicing: A two-way street. *WIREs RNA*.

Tian, B., and Manley, J.L. (2017). Alternative polyadenylation of mRNA precursors. *Nat Rev Mol Cell Biol* *18*, 18–30.

Toivola, D.M., Strnad, P., Habtezion, A., and Omary, M.B. (2010). Intermediate filaments take the heat as stress proteins. *Trends in Cell Biology* *20*, 79–91.

Touat-Todeschini, L., Shichino, Y., Dangin, M., Thierry-Mieg, N., Gilquin, B., Hiriart, E., Sachidanandam, R., Lambert, E., Brettschneider, J., Reuter, M., et al. (2017). Selective termination of lnc RNA transcription promotes heterochromatin silencing and cell differentiation. *EMBO J* *36*, 2626–2641.



Tranter, M., Helsley, R.N., Paulding, W.R., McGuinness, M., Brokamp, C., Haar, L., Liu, Y., Ren, X., and Jones, W.K. (2011). Coordinated Post-transcriptional Regulation of Hsp70.3 Gene Expression by MicroRNA and Alternative Polyadenylation. *J. Biol. Chem.* *286*, 29828–29837.

Trinklein, N.D., Murray, J.I., Hartman, S.J., Botstein, D., and Myers, R.M. (2004). The Role of Heat Shock Transcription Factor 1 in the Genome-wide Regulation of the Mammalian Heat Shock Response. *Molecular Biology of the Cell* *15*, 8.

Trixl, L., and Lusser, A. (2019). The dynamic RNA modification 5-methylcytosine and its emerging role as an epitranscriptomic mark. *WIREs RNA* *10*, e1510.

Truman, A.W., Kristjansdottir, K., Wolfgeher, D., Hasin, N., Polier, S., Zhang, H., Perrett, S., Prodromou, C., Jones, G.W., and Kron, S.J. (2012). CDK-Dependent Hsp70 Phosphorylation Controls G1 Cyclin Abundance and Cell-Cycle Progression. *Cell* *151*, 1308–1318.

Tuck, A.C., Rankova, A., Arpat, A.B., Liechti, L.A., Hess, D., Iesmantavicius, V., Castelo-Szekely, V., Gatfield, D., and Bühler, M. (2020). Mammalian RNA Decay Pathways Are Highly Specialized and Widely Linked to Translation. *Molecular Cell* *77*, 1222-1236.e13.

Tutar, Y., Song, Y., and Masison, D.C. (2006). Primate Chaperones Hsc70 (Constitutive) and Hsp70 (Induced) Differ Functionally in Supporting Growth and Prion Propagation in *Saccharomyces cerevisiae*. *Genetics* *172*, 851–861.

Valgardsdottir, R., Chiodi, I., Giordano, M., Rossi, A., Bazzini, S., Ghigna, C., Riva, S., and Biamonti, G. (2008). Transcription of Satellite III non-coding RNAs is a general stress response in human cells. *Nucleic Acids Research* *36*, 423–434.

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., et al. (2019). The human 18S rRNA m6A methyltransferase METTL5 is stabilized by TRMT112. *Nucleic Acids Research* *47*, 7719–7733.

Vazquez, J. (1991). Response to heat shock of gene 1, a *Drosophila melanogaster* small heat shock gene, is developmentally regulated. *Molec. Gen. Genet.* *226*, 393–400.

Vigh, L., Nakamoto, H., Landry, J., Gomez-Munoz, A., Harwood, J.L., and Horvath, I. (2007). Membrane Regulation of the Stress Response from Prokaryotic Models to Mammalian Cells. *Annals of the New York Academy of Sciences* *1113*, 40–51.

Vihervaara, A., Mahat, D.B., Guertin, M.J., Chu, T., Danko, C.G., Lis, J.T., and Sistonon, L. (2017). Transcriptional response to stress is pre-wired by promoter and enhancer architecture. *Nat Commun* *8*, 255.

Vihervaara, A., Duarte, F.M., and Lis, J.T. (2018). Molecular mechanisms driving transcriptional stress responses. *Nat Rev Genet* *19*, 385–397.

Viphakone, N., Sudbery, I., Griffith, L., Heath, C.G., Sims, D., and Wilson, S.A. (2019). Co-transcriptional Loading of RNA Export Factors Shapes the Human Transcriptome. *Molecular Cell* *75*, 310-323.e8.

- Wang, E.T., Sandberg, R., Luo, S., Khrebtkova, I., Zhang, L., Mayr, C., Kingsmore, S.F., Schroth, G.P., and Burge, C.B. (2008). Alternative isoform regulation in human tissue transcriptomes. *Nature* 456, 470–476.
- Wang, H., Xu, B., and Shi, J. (2020). N6-methyladenosine METTL3 promotes the breast cancer progression via targeting Bcl-2. *Gene* 722, 144076.
- Wang, K., Wang, L., Wang, J., Chen, S., Shi, M., and Cheng, H. (2018). Intronless mRNAs transit through nuclear speckles to gain export competence. *Journal of Cell Biology* 217, 3912–3929.
- Wang, P., Doxtader, K.A., and Nam, Y. (2016). Structural Basis for Cooperative Function of Mettl3 and Mettl14 Methyltransferases. *Molecular Cell* 63, 306–317.
- Wang, X., Chen, M., Zhou, J., and Zhang, X. (2014). HSP27, 70 and 90, anti-apoptotic proteins, in clinical cancer therapy. *International Journal of Oncology* 45, 18–30.
- Wang, X., Zhao, B.S., Roundtree, I.A., Lu, Z., Han, D., Ma, H., Weng, X., Chen, K., Shi, H., and He, C. (2015). N6-methyladenosine Modulates Messenger RNA Translation Efficiency. *Cell* 161, 1388–1399.
- Warda, A.S., Kretschmer, J., Hackert, P., Lenz, C., Urlaub, H., Höbartner, C., Sloan, K.E., and Bohnsack, M.T. (2017). Human METTL16 is a N<sup>6</sup>-methyladenosine (m<sup>6</sup>A) methyltransferase that targets pre-mRNAs and various non-coding RNAs. *EMBO Rep* 18, 2004–2014.
- Welch, W.J., and Suhan, J.P. (1985). Morphological study of the mammalian stress response: characterization of changes in cytoplasmic organelles, cytoskeleton, and nucleoli, and appearance of intranuclear actin filaments in rat fibroblasts after heat-shock treatment. *The Journal of Cell Biology* 101, 1198–1211.
- Welch, W.J., and Suhan, J.P. (1986). Cellular and biochemical events in mammalian cells during and after recovery from physiological stress. *The Journal of Cell Biology* 103, 2035–2052.
- Westerheide, S.D., Anckar, J., Stevens, S.M., Sistonen, L., and Morimoto, R.I. (2009). Stress-Inducible Regulation of Heat Shock Factor 1 by the Deacetylase SIRT1. *Science* 323, 1063–1066.
- Whitesell, L., and Lindquist, S. (2009). Inhibiting the transcription factor HSF1 as an anticancer strategy. *Expert Opinion on Therapeutic Targets* 13, 469–478.
- Wojtas, M.N., Pandey, R.R., Mendel, M., Homolka, D., Sachidanandam, R., and Pillai, R.S. (2017). Regulation of m6A Transcripts by the 3'→5' RNA Helicase YTHDC2 Is Essential for a Successful Meiotic Program in the Mammalian Germline. *Molecular Cell* 68, 374–387.e12.
- van Wolfswinkel, J.C., and Ketting, R.F. (2010). The role of small non-coding RNAs in genome stability and chromatin organization. *Journal of Cell Science* 123, 1825–1839.
- Wu, C. (1995). Heat Shock Transcription Factors: Structure and Regulation. 29.

Wu, K., Stull, F., Lee, C., and Bardwell, J.C.A. (2019). Protein folding while chaperone bound is dependent on weak interactions. *Nat Commun* 10, 4833.

Xiang, Y., Laurent, B., Hsu, C.-H., Nachtergaele, S., Lu, Z., Sheng, W., Xu, C., Chen, H., Ouyang, J., Wang, S., et al. (2017). RNA m6A methylation regulates the ultraviolet-induced DNA damage response. *Nature* 543, 573–576.

Xiao, X. (1999). HSF1 is required for extra-embryonic development, postnatal growth and protection during inflammatory responses in mice. *The EMBO Journal* 18, 5943–5952.

Xiao, H., and Lis, J. (1988). Germline transformation used to define key features of heat-shock response elements. *Science* 239, 1139–1142.

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

Xu, C., Wang, X., Liu, K., Roundtree, I.A., Tempel, W., Li, Y., Lu, Z., He, C., and Min, J. (2014). Structural basis for selective binding of m<sup>6</sup>A RNA by the YTHDC1 YTH domain. *Nat Chem Biol* 10, 927–929.

Xu, C., Liu, K., Ahmed, H., Loppnau, P., Schapira, M., and Min, J. (2015). Structural Basis for the Discriminative Recognition of *N*<sup>6</sup>-Methyladenosine RNA by the Human YT521-B Homology Domain Family of Proteins. *J. Biol. Chem.* 290, 24902–24913.

Youn, J.-Y., Dunham, W.H., Hong, S.J., Knight, J.D.R., Bashkurov, M., Chen, G.I., Bagci, H., Rathod, B., MacLeod, G., Eng, S.W.M., et al. (2018). High-Density Proximity Mapping Reveals the Subcellular Organization of mRNA-Associated Granules and Bodies. *Molecular Cell* 69, 517–532.e11.

Young, J.C., Barral, J.M., and Ulrich Hartl, F. (2003). More than folding: localized functions of cytosolic chaperones. *Trends in Biochemical Sciences* 28, 541–547.

Yue, Y., Liu, J., Cui, X., Cao, J., Luo, G., Zhang, Z., Cheng, T., Gao, M., Shu, X., Ma, H., et al. (2018). VIRMA mediates preferential m<sup>6</sup>A mRNA methylation in 3'UTR and near stop codon and associates with alternative polyadenylation. *Cell Discov* 4, 10.

Zaccara, S., and Jaffrey, S.R. (2020). A Unified Model for the Function of YTHDF Proteins in Regulating m<sup>6</sup>A-Modified mRNA. *Cell* 181, 1582–1595.e18.

Zaccara, S., Ries, R.J., and Jaffrey, S.R. (2019). Reading, writing and erasing mRNA methylation. *Nat Rev Mol Cell Biol* 20, 608–624.

Zander, G., Hackmann, A., Bender, L., Becker, D., Lingner, T., Salinas, G., and Krebber, H. (2016). mRNA quality control is bypassed for immediate export of stress-responsive transcripts. *Nature* 540, 593–596.

Zhang, C., and Jia, G. (2018). Reversible RNA Modification N<sup>1</sup>-methyladenosine (m<sup>1</sup>A) in mRNA and tRNA. *Genomics, Proteomics & Bioinformatics* 16, 155–161.

Zhang, C., Chen, L., Peng, D., Jiang, A., He, Y., Zeng, Y., Xie, C., Zhou, H., Luo, X., Liu, H., et al. (2020a). METTL3 and N<sup>6</sup>-Methyladenosine Promote Homologous Recombination-Mediated Repair of DSBs by Modulating DNA-RNA Hybrid Accumulation. *Molecular Cell* S1097276520304020.

Zhang, F., Kang, Y., Wang, M., Li, Y., Xu, T., Yang, W., Song, H., Wu, H., Shu, Q., and Jin, P. (2018). Fragile X mental retardation protein modulates the stability of its m<sup>6</sup>A-marked messenger RNA targets. *Human Molecular Genetics*.

Zhang, H., Shi, X., Huang, T., Zhao, X., Chen, W., Gu, N., and Zhang, R. (2020). Dynamic landscape and evolution of m<sup>6</sup>A methylation in human. *Nucleic Acids Research* gkaa347.

Zhao, B.S., Roundtree, I.A., and He, C. (2017). Post-transcriptional gene regulation by mRNA modifications. *Nat Rev Mol Cell Biol* 18, 31–42.

Zhao, X., Yang, Y., Sun, B.-F., Shi, Y., Yang, X., Xiao, W., Hao, Y.-J., Ping, X.-L., Chen, Y.-S., Wang, W.-J., et al. (2014). FTO-dependent demethylation of N<sup>6</sup>-methyladenosine regulates mRNA splicing and is required for adipogenesis. *Cell Res* 24, 1403–1419.

Zhao, Y.L., Liu, Y.H., Wu, R.F., Bi, Z., Yao, Y.X., Liu, Q., Wang, Y.Z., and Wang, X.X. (2019). Understanding m<sup>6</sup>A Function Through Uncovering the Diversity Roles of YTH Domain-Containing Proteins. *Mol Biotechnol* 61, 355–364.

Zheng, S. (2016). Alternative splicing and nonsense-mediated mRNA decay enforce neural specific gene expression. *Int. j. Dev. Neurosci.* 55, 102–108.

Zheng, G., Dahl, J.A., Niu, Y., Fedorcsak, P., Huang, C.-M., Li, C.J., Vågø, C.B., Shi, Y., Wang, W.-L., Song, S.-H., et al. (2013). ALKBH5 Is a Mammalian RNA Demethylase that Impacts RNA Metabolism and Mouse Fertility. *Molecular Cell* 49, 18–29.

Zhong, X.-Y., Ding, J.-H., Adams, J.A., Ghosh, G., and Fu, X.-D. (2009). Regulation of SR protein phosphorylation and alternative splicing by modulating kinetic interactions of SRPK1 with molecular chaperones. *Genes & Development* 23, 482–495.

Zhou, Z., and Fu, X.-D. (2013). Regulation of splicing by SR proteins and SR protein-specific kinases. *Chromosoma* 122, 191–207.

Zhou, J., Wan, J., Gao, X., Zhang, X., Jaffrey, S.R., and Qian, S.-B. (2015). Dynamic m<sup>6</sup>A mRNA methylation directs translational control of heat shock response. *Nature* 526, 591–594.

Zhu, Z.-M., Huo, F.-C., and Pei, D.-S. (2020). Function and evolution of RNA N<sup>6</sup>-methyladenosine modification. *Int. J. Biol. Sci.* 16, 1929–1940.

Zinder, J.C., and Lima, C.D. (2017). Targeting RNA for processing or destruction by the eukaryotic RNA exosome and its cofactors. *Genes Dev.* 31, 88–100.

