University of Wisconsin Milwaukee
UWM Digital Commons

Theses and Dissertations

August 2019

Mechanistic Insights into Translational Silencing of HACL mRNA in Yeast S. Cerevisiae

Leena Sathe
University of Wisconsin-Milwaukee

Follow this and additional works at: https://dc.uwm.edu/etd

Part of the Biology Commons, and the Molecular Biology Commons

Recommended Citation

Sathe, Leena, "Mechanistic Insights into Translational Silencing of HACL mRNA in Yeast S. Cerevisiae" (2019). *Theses and Dissertations*. 2245.
https://dc.uwm.edu/etd/2245

This Dissertation is brought to you for free and open access by UWM Digital Commons. It has been accepted for inclusion in Theses and Dissertations by an authorized administrator of UWM Digital Commons. For more information, please contact open-access@uwm.edu.

MECHANISTIC INSIGHTS INTO TRANSLATIONAL SILENCING OF *HAC1* mRNA IN YEAST *S. cerevisiae*

by

Leena Sathe

A Dissertation Submitted in

Partial Fulfillment of the

Requirements for the Degree of

Doctor of Philosophy

in Biological Sciences

at

The University of Wisconsin-Milwaukee

August 2019

ABSTRACT

MECHANISTIC INSIGHTS INTO TRANSLATIONAL SILENCING OF *HAC1* mRNA IN YEAST *S. cerevisiae*

by

Leena Sathe

University of Wisconsin-Milwaukee, 2019
Under the Supervision of Dr. Madhusudan Dey

Protein synthesis is a fundamental life process. Protein synthesis regulates cellular metabolism, cellular growth, the cell cycle, and cell morphogenesis. Technical advances in molecular biology, advanced high-throughput sequencing technologies and recent developments in crystallographic methodologies have aided in better understanding of the process of protein synthesis. Current knowledge of protein synthesis provides us with an overview of the initiation, elongation, and termination steps, general regulatory mechanisms, and molecular functions of the ribosomal rRNA and proteins. However, even with all this information we are far from understanding the detailed sequence of molecular interactions involved in the process of protein synthesis. We have snapshots of different stages of protein synthesis (initiation, elongation, and termination) but, we are still missing the intricate details responsible for linking these snapshots and creating a complete picture depicting the mechanism of protein synthesis. This thesis is focused on understanding the mechanism of protein synthesis using the yeast *HAC1* mRNA as a model mRNA in yeast *Saccharomyces cerevisiae*.

The messenger RNA (mRNA) bears the genetic information that is decoded in the process of protein synthesis. The fate of a messenger mRNA is decided by regulatory elements present in

the mRNA that are known as *cis*-acting elements. They influence not only translation of the mRNA, but also mRNA splicing, mRNA localization, mRNA processing, and mRNA degradation. One of the *cis*-regulating factors is the RNA secondary structure. The RNA secondary structure controls the translation of an mRNA. The precise mechanism of which is not clearly understood. This thesis is aimed at uncovering the regulation of protein synthesis by the RNA secondary structure.

Proteins in their native conformation either spontaneously attain the folded conformation or they are folded by chaperone proteins in the cytoplasm or in the endoplasmic reticulum (ER). Sometimes the ER experiences an overload of unfolded proteins which results in a condition termed as "ER stress". Under the stress conditions, unfolded protein response (UPR) pathways are activated which result in expression of assorted transcription factors. These transcription factors modulate the cellular transcriptome and proteome to alleviate the ER stress conditions.

Our model mRNA from *Saccharomyces cerevisiae* encodes a transcription factor that is expressed under conditions of ER stress. *HAC1* mRNA contains a cytoplasmic intron (252 nucleotides) that base-pairs with the 5'- untranslated region (5'-UTR) of the mRNA. This base-pairing interaction inhibits the translation of the *HAC1* mRNA under physiological conditions. We use *HAC1* mRNA and the inherent base-pairing interaction to uncover new insights into the mechanism of translational regulation by the RNA secondary structure.

HAC1 pre-mRNA is composed of a 5'-untranslated region {(5'-UTR, 68-nucleotides (nt)}, an exon1 (661-nt), an intron (252-nt), an exon2 (57-nt), and a 3'-UTR (462-nt). Under conditions of ER stress, an endonuclease Ire1 cleaves the intron from the *HAC1* pre-mRNA to relieve the translational block of the mRNA. The two exons are spliced by tRNA ligase and the mature *HAC1* mRNA produces Hac1 protein which is an active transcription factor. We have

shown that the base pairing interaction between 5'-UTR and intron inhibits translation initiation of *HAC1* mRNA [1].

We performed a random genetic screen to identify an intragenic suppressor mutation (s) that would overcome the translational block in *HAC1* mRNA. We identified a point mutation in the base-pairing interaction that relieved the translational block in *HAC1* mRNA. Further mutational analyses of the base-pairing interaction suggested that it is critical for the regulation of *HAC1* mRNA translation. We also showed that insertion of an in-frame AUG start codon upstream of the RNA secondary structure releases the translational block, demonstrating that an elongating ribosome can disrupt the interaction. Moreover, overexpression of translation initiation factor eIF4A, a helicase, enhances production of Hac1 from an mRNA containing an upstream AUG start codon at the beginning of the base-paired region. Together, we showed that the RNA secondary structure regulates translation initiation of *HAC1* mRNA [1].

To dissect the translation initiation block further we shifted the RNA secondary structure from its normal position (which is *cap*-proximal) to one away from the mRNA cap structure (a *cap*-distal position) by inserting an unstructured RNA sequence. We observed that genetically engineered *HAC1* mRNA with the *cap*-distal secondary structure resulted in translation of the Hac1 protein. This result suggested that the cap-proximal RNA secondary structure possibly inhibits the interaction of the pre-initiation complex (PIC) with the *HAC1* mRNA.

Further analyses of the yeast transcriptome and translatome suggest that the *HAC1* gene locus expresses two overlapping transcripts; referred to here as "*HAC1a"* like the one described above, and "*HAC1b", a* newly identified variant. Interestingly, the newly identified *HAC1b* mRNA overlaps with the exon1 of *HAC1a*. Yeast transcriptome analyses show that it is composed of a long 5'-UTR (~400-nt), an open reading frame (693-nt) and a short 3'-UTR (124-nt). We

characterized the role of the *HAC1b* transcript in the context of the ER stress response. We observed that *HAC1b*, like *HAC1a* mRNA, remains translationally silent. However, *HAC1b* can activate the ER stress response as a functional Hac1 protein is synthesized in the absence of "Duh1" protein. "Duh1" is a component of the proteasome complex. These observations are consistent with the previous report that Duh1 targets the protein product from the un-spliced *HAC1* mRNA for degradation.

Taken together, our results provide mechanistic insights into the translational regulation of *HAC1* mRNA. In addition, we provide evidence that the transcript isoform of *HAC1* mRNA might play a role in the ER stress response in yeast, *S. cerevisiae.*

TABLE OF CONTENTS

List of Figures viii

List of Tables x

List of Abbreviations xi

Acknowledgments xii

Chapter 1. Introduction 1

1.1 Mechanism of eukaryotic translation initiation in the yeast system 2

1.2 Regulation of eukaryotic translation initiation 21

1.3 Translational silencing in yeast *HAC1* mRNA 34

Chapter 2. *HAC1* mRNA is translationally repressed at the initiation stage 40

2.1 Introduction 40

2.2 Single base-pair mutation in the intron derepresses translation of unspliced *HAC1* mRNA 41

2.3 Disruption in the 5'-UTR*intron base-pairing by single base pair mutations release the translational block 45

2.4 Insertion of an in-frame start codons in the 5'- UTR of *HAC1* mRNA 48

2.5 Unspliced *HAC1* mRNA is associated with monosomes 53

2.6 Overexpression of helicase eIF4A relieves the translational block in the *HAC1-AUG$_{-42}$,A$_1$G,A$_7$G* allele 55

2.7 Conclusion and future perspectives 57

Chapter 3. *HAC1* mRNA is translationally repressed due to insufficient loading of Initiation complex or helicase 59

3.1 Introduction 59

3.2 Over expression of the cap-binding complex proteins or RNA helicases does not relieve the translation block 60

3.3 Over expression of translation initiation factors and ribosomal proteins does not relieve the translation block 62

3.4 *HAC1* mRNA translation is repressed due to insufficient loading of the 43S PIC 63

3.5 Conclusion and future perspectives 70

Chapter 4. Adaptation to endoplasmic stress through distinct isoform-specific expression of *HAC1* mRNA 87

4.1 Introduction 71

4.2 Two isoforms of *HAC1* mRNA are transcribed from the *HAC1* gene locus 71

4.3 Both the *HAC1* mRNA isoforms: *HAC1a* and *HAC1b* are translationally repressed 74

4.4 *HAC1b* can activate the UPR in the Duh1 protein null strain 78

4.5 *HAC1b* mRNA bearing 3'-UTR of *HAC1a c*an *a*ctivate the ER stress response 82

4.6 *HAC1b* activates only the UPRE1-driven *LacZ* Gene 85

4.7 Conclusion and future perspectives 87

References 89

Materials and Methods 101

Curriculum Vitae 126

LIST OF FIGURES

No.	Caption	Page no.
Chapter 1		
1	Steps of eukaryotic translation initiation	4
2	Cryo-EM structure of yeast 40S ribosome	5
3	The schematics of the eukaryotic elongation cycle	7
4	Ratcheting and nonratcheting states of 80S ribosome	9
5	Crystal structure of the yeast 40S ribosomal subunit	12
6	Ribosomal proteins associated with of the yeast 40S subunit	16
7	X-ray structure of yeast 80S ribosome	16
8	DEAD-box helicase eIF4A	25
9	Translational regulation by IRE and IRP-1 protein	27
10	Cap-binding complex; eIF4F	31
11	Cap-binding complex as a drug target	33
12	Schematic representation of yeast UPR pathway	35
13	Schematic representation of *HAC1* mRNA and Hac1 protein	37
Chapter 2		
14	Single base-pair mutation in the intron derepresses translation of unspliced *HAC1* mRNA	43
15	Disruption in the 5'-UTR*intron base-pairing by single base pair mutations release the translational block	47
16	Insertion of an in-frame start codons in the 5'- UTR of *HAC1* mRNA	50
17	Insertion of an AUG codon at position -60 derepresses translation of unspliced *HAC1* mRNA	52

18 Unspliced HAC1 mRNA is associated with monosomes 54

19 Overexpression of helicase eIF4A relieves the translational block in the *HAC1-AUG$_{-42}$,A$_1$G,A$_7$G* allele 56

Chapter 3

20 Over expression of the cap-binding complex proteins or RNA helicases does not relieve the translation block 61

21 Over expression of translation initiation factors and ribosomal proteins does not relieve the translation block 63

22 Median length of yeast 5'-UTR is 55 nucleotides 65

23 Increasing 5'-UTR length upstream of *HAC1-42AUG* relives the initiation block 67

24 *HAC1* mRNA is translationally silent due to insufficient loading of 43S Pre-initiation complex 69

25 Proposed model showing that a suppressor protein is involved in the translational repression of *HAC1* mRNA 70

Chapter 4

26 Evidence of two isoforms of *HAC1* mRNA 73

27 Both *HAC1a* and *HAC1b* isoforms are translationally repressed 75

28 HAC1b can activate the UPR in the Duh1 protein null strain 80

29 *HAC1b* mRNA Bearing 3'-UTR of *HAC1a* Can Activate the ER stress Response 83

30 *HAC1b* activates only the UPRE1-driven *LacZ* Gene 84

31 Proposed Turnover of *HAC1b* mRNA Isoform 87

LIST OF TABELS

No.	Caption	Page no.
1	Translation initiation factors in yeast	3
2	The translation elongation and termination factors in yeast	8
3	List of ribosomal proteins associated with small ribosomal subunit	15
4	List of ribosomal proteins associated with large ribosomal subunit	19
5	Diseases caused due to mutations in the Kozak sequence	27

LIST OF ABBREVIATIONS

HAC1	
eIF	eukaryotic initiation factor
IRE1	iron requiring element 1
5’-UTR	5’- untranslated region
3’-UTR	3’- untranslated region
43S PIC	43S pre-initiation complex
UPR	unfolded protein response
UPRE	unfolded protein response element
ER	endoplasmic reticulum
uORF	upstream open reading frame
Tm	Tunicamycin
DTT	Dithiothreitol
SD	Synthetic dextrose
Tm^R	Tunicamycin resistance
Tm^S	Tunicamycin sensitive

ACKNOWLEDGMEMTS

PhD dissertation is not just about the research work, but it is also about learning to think and ask questions. Working closely with Dr. Madhusudan Dey was my privilege because I could learn these things from him. I would like to thank Dr. Dey for giving me an opportunity to work with him, for believing in me and for his continuous guidance and support, without that I would not have been able to complete my research work.

I want to thank our collaborates Dr. Cheryl Bolinger and Dr. Amin-ul Mannan who contributed in the paper published in the *Journal of Biological Chemistry*. This would not have been possible without generous help from Dr. Tom E Dever and Dr. Alan Hinnebusch. They not only gave us the helpful insights that shaped this project but also provided us with all the essential plasmids as and when required. I would like to thank all my graduate thesis committee members, Dr. Sergei Kuchin, Dr. Mark McBride, Dr. Steven Forst and Dr. David N Frick. They guided me throughout the PhD program.

I was very fortunate to have colleagues who are understanding and helpful. They made my everyday in the lab memorable and enjoyable. My past lab colleagues, Dr. Amin-ul Mannan and Dr. Ashish Anshu helped me get through first two years of the graduate school. I would like to thank them for all their pointers.

My dearest colleague Chandrima Ghosh was there with me and helped me through all taught times. I would not have finished my work without her helpful insights, troubleshooting tips and constant support. I would also like to thank post-doctoral research scientist Dr. Jagdeesh Kumar Uppala for all his help, support and insightful scientific discussions. I learnt a lot from him just by observing him in the lab. Our newest post-doctoral research scientist Dr. Faiz Ahmad helped me a lot by sharing my work load. I would like to thank him for that.

This thesis would not be in such a good shape if it was not for Anthony Pulvino. I owe a big thanks to Anthony for helping me edit and proof read the thesis. In addition, Anthony made my lab life very smooth by taking care of my lab responsibilities while I was busy writing. I would not have completed and finished writing the thesis without his help.

I want to thank Phylis Shaw and Jenna Rock form making me feel comfortable with all the formalities and paper work in the beginning of my PhD program. My graduate life was made smooth by Chrissy Kozek, Jessica Schultz, Bonnie Murphy and Amber Grupe by answering all the questions and getting all my paperwork aligned through out my PhD.

I am grateful to have found supporting friends who made me laugh on my most depressing lab days. I would like to extend my thanks to my ex and current flat mates Srishti Sahu, Raktima Dasgupta and Sumona Dhara who created a home away from home for me. My dear friends Ankush, Surashree, Unmesh, Snehanshu, Renne, Jecob, Shikha for being there for me whenever I needed support and guidance. And finally, I would also like to thank Dr. Aishwarya Shevade for all her suggestions, help and support in last three years of my PhD.

Lastly, I want to thank my parents Dr. Padmakar Sathe and Dr. Sushama Sathe for encouraging me to pursue higher studies. I would like to thank my sister Anuja who was always available for me and supported me. I am fortunate to have supportive in-laws who encouraged me to aim higher. There are not enough words to thank my husband Dr. Harshad Sahasrabudhe. This work would not have been possible without his constant support and guidance.

1. Introduction

Genetic information is encoded in DNA. DNA undergoes transcription to synthesize RNAs (mRNA, tRNA, rRNA, miRNA, siRNA etc). Messenger RNA (mRNA) acts as the template in the process of protein synthesis *i.e.* translation. Translation is a complex process which requires a messenger RNA (mRNA), transfer RNAs (tRNAs), the ribosome and more than 25 initiation factors (see table 1). It is known to occur in four major steps: Initiation, elongation, termination, and ribosome recycling.

In eukaryotes, translation is known to occur in four steps: 1) Initiation, 2) Elongation, 3) Termination, and 4) Ribosome recycling. The translation initiation step is further divided into three steps: (I) loading of a complex of small ribosomal subunit (40S) and initiation factors, known as the 43S pre-initiation complex (43S PIC), (II) ribosomal scanning by 43S PIC and (III) AUG codon recognition by initiator methionyl tRNA (see the schematic diagram, Fig. 1).

Translation elongation is the iterative step in which amino acids are added to the growing polypeptide chain in the transpeptidation reaction catalyzed by ribosome. The elongation cycle occurs in three steps (I) selection of correct aa-tRNA, (II) peptide bond formation, and (III) translocation. Translation is terminated once the ribosome encounters the stop codon (UAA, UAG and UGA). Then, the newly synthesized polypeptide chain is released from the ribosome. Following release the entire translational machinery is disassembled and is recycled to start a new translation cycle in the last step of ribosomal recycling.

This thesis focuses on understanding the mechanisms of translational regulation using budding yeast, *S. cerevisiae,* as the model organism and yeast *HAC1* mRNA as the model mRNA.

1.1 Mechanism of eukaryotic translation

A. Mechanism of the translation initiation using yeast as model system

Described below is the bird's eye view of the translation initiation process. To begin translation in eukaryotes, translation initiation factor 4E (eIF4E) binds to 7-methylguanosine (m^7G) cap of the mRNA along with the scaffold protein eIF4G, the RNA dependent helicase eIF4A, an activator protein eIF4B, and the poly(A)-binding protein (PABP); a complex known as the cap-binding complex (eIF4F) [1] [2]. The eIF4F; the cap-binding complex, recruits the small ribosomal subunit (40S), a ternary complex (TC) made of eIF2, GTP and initiator met-$tRNA_i^{met}$, and other initiation factors (e.g., eIF1, eIF1A and eIF3, eIF5) to form a 43S-pre-initiation complex (43S-PIC). The 43S-PIC then scans along with the 5'-untranslatred region (5'-UTR) in search of an AUG start codon, and the AUG codon is recognized by an anticodon on the met-$tRNA_i^{met}$ [3]. The codon-anticodon interaction triggers the dissociation of initiation factors, and GTPase eIF5B stimulates joining of the large ribosomal subunit (60S) to form an 80S complex in which the met-$tRNA_i^{met}$ base-paired to the AUG codon. The 80S ribosome then starts the elongation cycle by synthesizing the polypeptide chain [3] [4]. When the 80S ribosome encounters one of the three stop codons (UAA, UAG, UGA), the elongation cycle stops. The newly synthesized polypeptide is then released from the ribosome [5]. Following the release, the 80S ribosome dissociates into the small subunit (40S) and large subunit (60S). Subsequently, both subunits are recycled to begin a new translation cycle.

In eukaryotes the small 40S ribosomal subunit is made of the18S ribosomal RNA (rRNA) and 33 ribosomal proteins. The cryo-EM structure of the 40S subunit reveals that its shape resembles a sitting bird [6] [7] mimicking the shape of its head, beak, neck, and foot (see Fig. 2) . Initiation factors eIF1 and eIF1A cooperatively bind to the 40S ribosome and open the mRNA

entry channel, thus maintaining the 40S ribosome in a scanning competent state, or "open conformation" [8] [9] [10] [11] (see Fig 2A). The mRNA passes through the entry channel and then through the ribosomal P-site. The met-tRNA$_i^{met}$ of the 43S-PIC then starts inspecting the mRNA base-by-base in search of a start codon (AUG) in a process called "ribosomal scanning" [3] [4].

	Factor	Subunit	Gene	Molecular weight (kDa)	Function
1	eIF1		*SUI1*	12.3	AUG recognition, promotes 43S PIC formation
2	eIF1A		*TIF11*	17.4	Promotes 43S PIC formation, ribosomal scanning
3	eIF2	α	*SUI2*	34.7	AUG recognition, global translation regulation.
		β	*SUI3*	31.6	AUG recognition, RNA binding
		γ	*GCD11*	57.9	AUG recognition, GTP binding, interaction with tRNA
4	eIF2B	α	*GCN3*	34.0	regulatory subunit plays role in GEF function
		β	*GCD7*	42.6	regulatory subunit plays role in GEF function
		γ	*GCD1*	65.7	Promotes GEF function of catalytic subunit
		δ	*GCD2*	70.9	Regulatory subunit plays role in GEF function
		ε	*GCD6*	81.2	Catalytic subunit of GEF function
5	eIF3	a	*TIF32*	110.3	Promotes TC and mRNA binding to 40S
		b	*PRT1*	88.1	Contains RRM, promotes TC and mRNA binding to 40S,
		c	*NIP1*	93.2	Contains PIC motif, promotes TC and mRNA binding to 40S, interacts with eIF1,1A and eIF5
		g	*TIF35*	30.5	Contains RRM and Zn domain, binds to RNA and eIF5B
		i	*TIF34*	38.8	Contains 7 WD repeats, binds to eIF3 subunits
		j	*HCR1*	29.6	Non-essential in Yeast, promotes integrity of initiation factor complex
6	eIF4A		*TIF1*	45.1	ATPase, RNA helicase, interacts with eIF4G in cap-binding complex
			TIF2	44.6	ATPase, RNA helicase, , interacts with eIF4G in cap-binding complex
7	eIF4B		*TIF3*	48.5	Forms cap binding complex, promotes eIF4A activity, interaction with Rps20p at the head of the 40S ribosomal subunit alters the structure of the mRNA entry channel
8	eIF4E		*CDC33*	24.3	Binds to m^{7}G cap, forms a cap binding complex with eIF4G
9	eIF4G		*TIF4631*	107.1	Forms cap binding complex by interacting with eIF4E, eIF4A and B, recruits 43S PIC on mRNA by interacting with eIF3
			TIF4632	103.9	Forms cap binding complex by interacting with eIF4E, eIF4A and B, recruits 43S PIC on mRNA by interacting with eIF3
10	eIF5		*TIF5*	45.2	AUG recognition, stimulates eIF2 GTPase by associating with 43SPIC
11	eIF5B		*FUN12*	112.3	GTPase promotes subunit joining

Table 1. Translation initiation factors in yeast: The subunits, gene name, molecular weights, and functions of each factor are given in the table. This table is adapted and modified from the Ref (Genetics. 2016 May;203(1):65-107)

Fig 1: Steps of Eukaryotic translation initiation

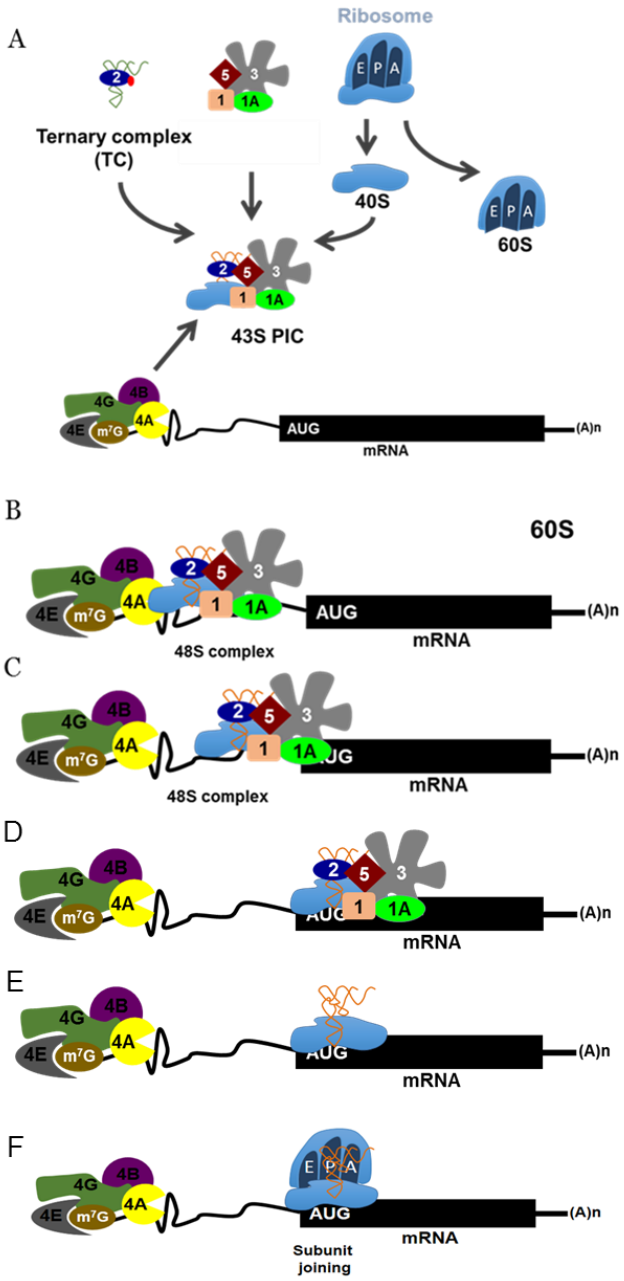

Fig 1 : Steps of eukaryotic translation initiation (A) Cartoon depiction of the eukaryotic mRNA, initiation factors and ribosome. Ternary complex TC is formed by eIF2-GTP-Met $tRNA_i^{met}$. A protein complex consisting of eukaryotic initiation factors (eIFs) eIF1, eIF1A, eIF3 and eIF5 is shown. The 80S ribosome is composed of small subunit (40S) and large subunit (60S). Cap-binding complex (eIF4F) composed of eIF4G, eIF4E, eIF4A and eIF4B. It binds to the m_7G cap of the mRNA. (B) Loading of 43S-pre-initiation complex (PIC) (40S ribosome, TC and protein complex) near the mRNA cap. 43S-PIC together with mRNA is called 48S-PIC (C) Scanning of 43S-PIC along the 5'-UTR of the mRNA in search of the AUG start codon. (D) AUG codon recognition by an anticodon on the initiator tRNA. (E) dissociation of all initiation factors immediately after AUG codon recognition. (F) Large ribosomal subunit (60S) joins small subunit. 80S ribosome starts translation elongation.

During ribosomal scanning, the GTPase eIF5 provides energy by hydrolyzing a GTP molecule to GDP and (Pi). However, P_i is not released from the GDP molecule [3] [12].

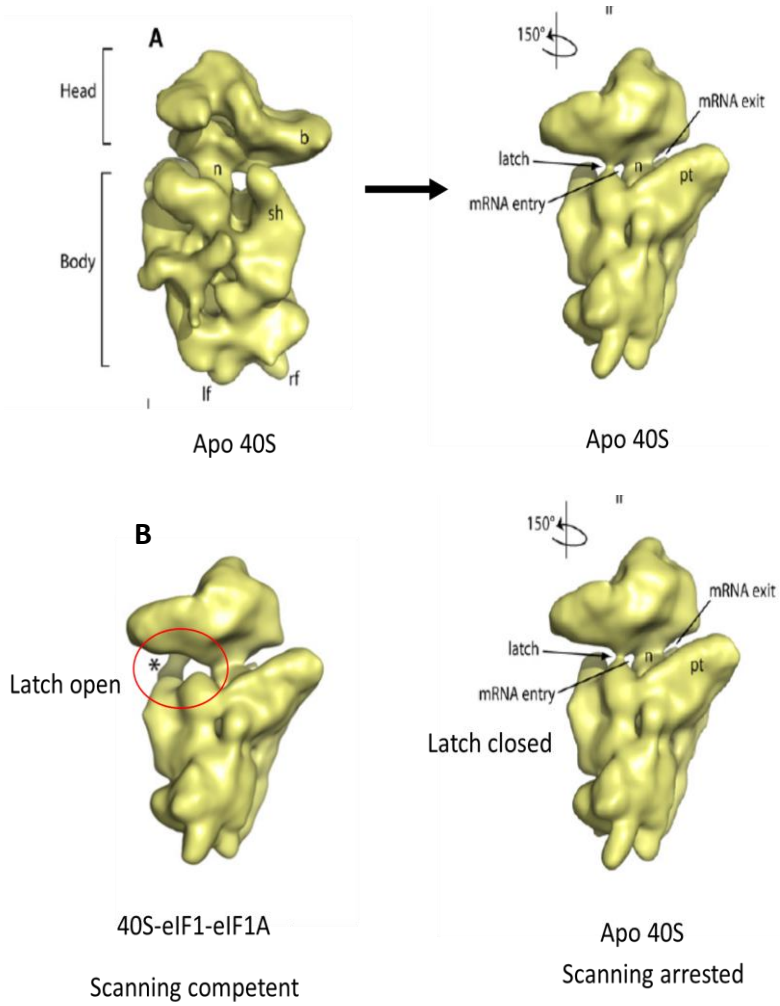

Figure 2. Cryo-EM structure of yeast 40S ribosome: (A) The cryo-EM structure of Apo 40S looks like a sitting bird. Mimicking its h; head, b; beak, n; neck and lf; left foot and rf; right foot. When rotated by 150 ° a latch near mRNA entry channel is observed, which is a non-covalent interaction between body and head. (B) Latch (circled in red) is formed by non-covalent interactions between head and body. It is in open conformation when 40S is bound by eIF1 and eIF1A. This is scanning competent state of 40S. Once the eIF1 and eIF1A dissociate the latch is closed and this is the scanning arrested conformation of 40S. The images are taken from the reference Passmore et al, 2007 (Mol Cell. 2007 Apr 13;26(1):41-50) , edited and labeled.

Interaction with the AUG codon by an anticodon on met-$tRNA_i^{met}$ triggers a conformational change from the scanning competent, open conformation to the scanning arrested or "closed complex"[3] [4] (see Fig 2B) . Such conformation change leads to the release of all initiation factors and previously hydrolyzed P_i [3] [4] [13]. Then, eIF5B stimulates the joining of 60S and 40S subunit to form an 80S complex.

B. Mechanism of translation elongation using yeast model system

Following the identification of the AUG codon, initiation factors must dissociate in order to facilitate the joining of the 60S ribosomal subunit to the 40S. Codon-anticodon interaction results in the release of the GDP bound form of eIF2 which has low affinity for the Met-$tRNA^i_{met}$. The 60S subunit is then recruited by and complexed with the GTPase, eIF5B. Subunit joining is followed by GTP hydrolysis and the release of eIF5B-GDP. eIF1A is subsequently released whereas initiation factors, "eIF1" and "eIF3" remain attached to the 80S ribosome [14] [15]. Multiple questions regarding the complex mechanism behind the subunit joining remain unanswered.

Once the 80S ribosome is formed, it starts the elongation step of translation. The schematic representation of the elongation cycle is shown in Fig. 3 and the yeast elongation factor are shown in Table 2. The mechanism of translation elongation is very much conserved in bacterial and eukaryotic systems as opposed to the process of translation initiation. The elongation factors are abbreviated as eukaryotic elongation factors (eEFs). The eEF1A is orthologous to the bacterial elongation factor EF-Tu which forms a complex with the aminoacyl-tRNA and GTP (eEF1A-GTP-aminoacyl-tRNA). This complex delivers the aminoacyl tRNA to the A (acceptor) t-RNA binding site on the 80S ribosome. The codon-anticodon interaction stimulates hydrolysis of GTP followed by the release of the GDP bound eEF1A.

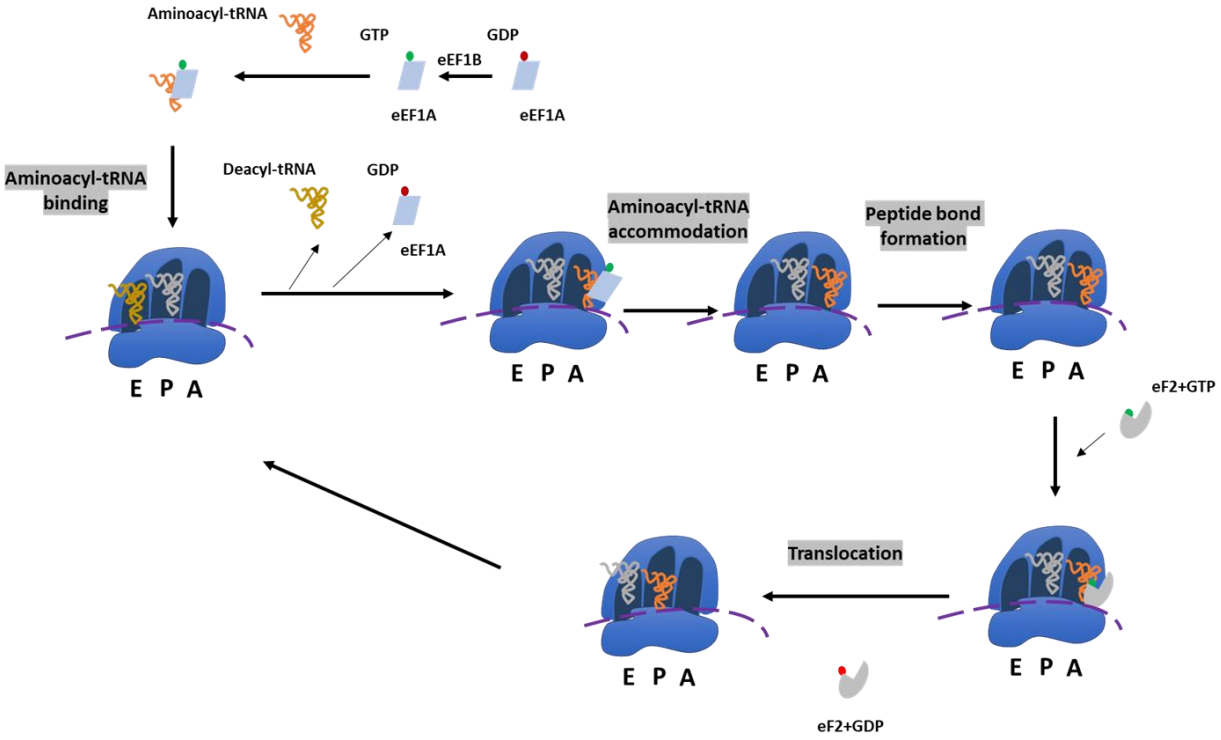

Figure 3. The schematics of the eukaryotic elongation cycle: The 80S ribosome is depicted in blue color with the E, P, and A sites labelled. All other elongation factors, tRNA, and GTP are schematically presented a labelled. The schematic representation does not specifically indicate the exact positions and interaction of the factors and the ribosome. This image is adapted from the Ref [15] (Cold Spring Harb Perspect Biol 2012;4:a013706). eEF1A associated with the GTP binds to the cognate amicoacyl-tRNA and delivers it to the 80S ribosome at the A site. The codon-anticodon interaction at the ribosomal A site results in release of GDP bound form of eEF1A. Peptide bond formation is followed by binding of the eEF2-GTP. This results in translocation of the tRNAs at the P and E sites and release of eEF2-GDP. The deacyl-tRNA is released from the E-site and ribosome is now ready for the next elongation cycle.

The eEF1B recycles the GDP-eEF1A to GTP-eEF1A which binds to the next cognate aminoacyl-tRNA. eEF1B has two subunits, "β" and "γ". The "β" subunit is the catalytic Guanine exchange factor whereas the "γ" is the non-catalytic regulatory subunit.

Hydrolysis of GTP associated with eEF1A is immediately followed by the peptide bond formation at the peptidyl transferase center (PTC), which is located within rRNA elements of the large ribosomal subunit. The crystal structure of the yeast 80S ribosome and the *T. thermophila* 60S ribosome show conservation of the rRNA elements in the PTC suggesting that the mechanism of the peptide bond formation is universal [15]. Peptide bond formation pushes the acceptor end

No.	Factor	Subunits	Gene	Molecular weight (kDa)	Function
1	eEF1A		*TEF1* *TEF2*	50.045	Binds and recruits aminoacyl tRNA to the ribosomal A-site
2	eEF1B	β	*EFB1* *TEF5*	22.603	Guanine exchange factor
3	eEF1B	γ	*TEF4*	46.52	Non-catalytic subunit
4	eEF2		*EFT1* *EFT2*	93.2	Ribosome translocation
5	eEF3		*YEF3* *HEF3*	115.9	Fungal specific elongation factor, ATPase, binds to E near E site on ribosome
6	eIF5A		*HYP1* *ANB1*	17.12	Required for translation elongation
7	eRF1		*SUP 45*	49.8	Identifies stop codon, mimics tRNA
8	eRF2		*SUP35*	76.5	GTPase
9	Rli1		*RLI1*	68.3	APTase

Table 2. The translation elongation and termination factors in yeast: Elongation factors are shown in dark blue color and termination factors are shown in black color. The no of subunits, gene names, molecular weights, and functions are listed for all the factors. The table is adapted and modified from Ref [5] (Genetics, Vol. 203, 65–107 May 2016).

of the tRNA in the ribosomal E site and ribosomal P site whereas the anticodon ends remain at the ribosomal P site and ribosomal A sites, respectively.

This is the hybrid transition stage which requires eEF2-GTP for the translocation of the ribosome. The GTP hydrolysis and release of Pi allows the movement of mRNA and tRNAs with respect to the ribosome. The GTP hydrolysis triggers a conformational change which is thought to unlock the ribosomal subunits allowing the mRNA and tRNA movements and lock the subunits immediately after translocation. eEF2 is an ortholog of bacterial EF-G.

The ribosome performs translational elongation unidirectionally and by rotating its subunits referred to as “ratcheting” of the ribosome [15]. Fig. 4 A and B show the transition from the nonratcheting state (before translocation) to the ratcheting state (immediately after translocation) of the ribosome [16]. In the ratcheted state of the yeast 80S ribosome, the head of the 40S subunit rotates counter-clockwise towards the mRNA exit channel (See Fig. 4 A and B)

[16]. This rotation of the 40S ribosome results in the altering of interactions (bridges) between the 40S and 60S ribosomal subunits. The 5S rRNA and ribosomal proteins enveloping the rRNA also undergo structural rearrangements at this stage to accommodate the alteration of the "bridge" interactions. The proposed model of the "ratcheting" of the ribosome suggests that it is a multistep process and the structural rearrangements are followed by rotation of the head which unlocks the mRNA and tRNA from the 40S ribosomal subunit [7] [16] .

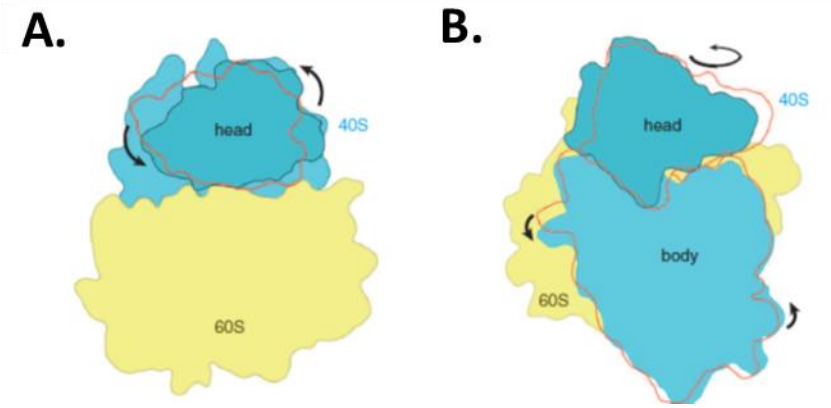

Figure.4 Ratcheting and nonratcheting states of 80S ribosome : (A) Top view of 80S ribosome indicating transition from nonratcheted state to ratcheted state. (B) Solvent view of 80S ribosome indicating transition from nonratcheted state to ratcheted state. This Figure is obtained from Ref (SCIENCE VOL 330 26 NOVEMBER 2010)

Interestingly, eEF3 is a fungal specific elongation factor and has no bacterial or human ortholog. It has an ATPase domain with two ATP binding cassettes (ABC). eEF3 binds to the ribosome near the E site and dissociation of it requires ATP hydrolysis [5][15]. Detailed comparative studies have been performed on bacterial 70S and yeast 80S ribosomal crystal structures to understand the requirement of eEF3 in yeast translation elongation. In the recent work it was observed that "stm1" protein binds to the head of the yeast 80S ribosome which negatively regulates the eEF3 association with the ribosome [15].

C. Mechanism of translation termination using yeast model system

Translation is terminated when one of the stop codons (UAA, UAG, and UGA) is recognized by a termination factor, “eRF1”, at the A site of the ribosome. In termination of translation eRF1 binds to the ribosome along with the GTP-bound form of the eRF2 protein. eRF1 has three functionally important domains; (I) the N-terminal domain which mimics tRNA and contacts the stop codon at the ribosomal A site, (II) central domain which hydrolyses the peptidyl-tRNA bond, and (III) the C-terminal domain which interacts with eRF2 [5] [15].

eRF2 bound GTP is hydrolyzed upon interaction and recognition of a stop codon by eRF1. Hence, the interaction of eRF1 and eRF2 is critical for the translational termination. Upon GTP hydrolysis, eRF2 is released from the ribosome and eRF1 is bound by “Rli1” an ATPase. Binding of Rli1to eRF1 promotes the hydrolysis of the bond between the polypeptide chain and peptidyl-tRNA. The polypeptide is then released followed by dissociation of the 60S ribosomal subunit from the 80S upon ATP hydrolysis by Rli1[5]. Finally, the small ribosomal subunit 40S is dissociated from tRNA and mRNA by recycling factors, “Tma20”, “Tma22”, and “Tma64” the mechanism of which is conserved in yeast and human system.

D. Eukaryotic ribosome

The yeast small ribosomal subunit (40S) is composed of the 18S rRNA and 33 ribosomal proteins [17]. The yeast large ribosomal subunit (60S) is composed of 23S rRNA and 5.8S rRNA with 46 ribosomal proteins. The 18S rRNA forms the structural core of the small subunit 40S and provides attachment sites for the mRNA and tRNA. The tertiary structure of the 40S ribosome resembles a bird, as described above and shown in Fig. 2 and 5.

The secondary structure of 18S rRNA is depicted in Fig. 5A. The structural domains are color coded which include 5' domain, 3' domain, 3' major domain, and central domain (see Fig 3A) [18]. The helices of the 5' domain (h1-h18) form the shoulder of the sitting bird, the right foot, and part of the body. Specifically, h16 and h17 form the right-side border of the body labelled here as "shoulder". The central domain forms the platform having nine helices folded in the shape of a "W". The 3' major domain forms the head and beak part of the sitting bird which contains 15 helices loosely packed with few inter-helical contacts. The 3' minor domain has only two helices h44 and h45 which form the body of the bird and interact with the large ribosomal subunit. The h44 is longest helix and runs from the bottom of the head of the bird to the bottom of the body of the bird [17] [18] [19]. The tertiary structure of this 18S rRNA is conserved with the bacterial 16S, however the 16S rRNA lacks the eukaryotic specific (ES) regions. ES3 A and B are in 5' domain. ES6 and ES7 are in the central domain and ES9 and ES12 are in 3' major and 3' minor domains, respectively. The h16 is located just below the beak and it plays a role in connecting the head to the body when the 40S is bound by eIF1 and eIF1A. This interaction creates the "latch" which forms the mRNA entry channel, depicted in Fig. 2B.

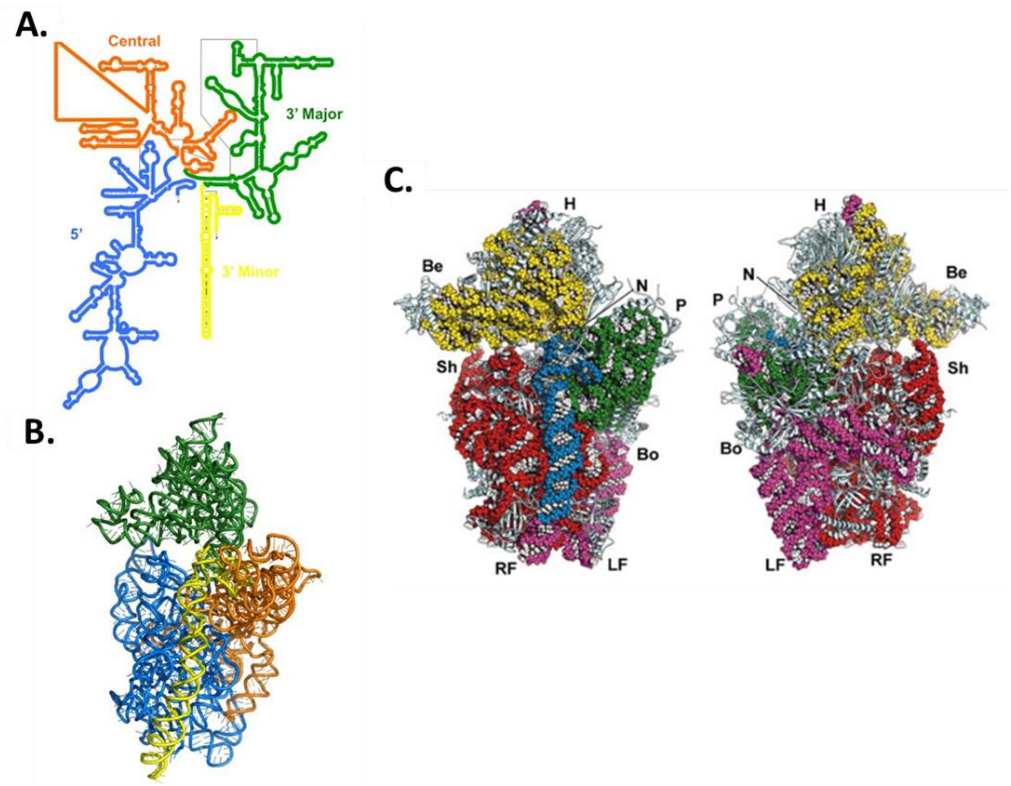

Figure 5. Crystal structure of the yeast 40S ribosomal subunit: (A) Secondary structure of 18S rRNA from *S. cerevisiae*. Adapted and modified from the ref [18] (Genetics, Vol. 195, 643–681). The domains are colored, the central in orange, the 5' domain in blue, 3' major in green, and the 3' minor in yellow color. (B) Ternary structure of 40S ribosome (PDB ID: 2XZM). (C) Front and back view of the 40S ribosome (PDB ID: 2XZM). The proteins are shown in gray color in cartoon representation. The rRNA color scheme is consistent with that of the 18S rRNA domains.

A total of 33 ribosomal proteins are associated with the 40S ribosomal subunit (see Table 3). Ribosomal proteins from the three domains of life had independent nomenclature system. Interestingly it was in 2014 that a universally accepted, systematic nomenclature was developed for all known ribosomal proteins. The universal ribosomal proteins, conserved throughout three domains of life, have the prefix "u". Exclusive bacterial proteins have the prefix "b", and proteins belonging to either eukaryotes, archaea, or both have the prefix "e" [20]. Ribosomal proteins are abbreviated as "rp". The proteins associated with the small ribosomal subunit are donated as "S" *e.g.* rpS, and the proteins associated with the large ribosomal subunit are denoted as "L" *i.e.* rpL. The ribosomal protein genes from different species of the three kingdoms of life are listed in the

Ribosomal Protein Gene Database (RPG database) with their DNA and protein sequences, chromosomal locations, and their corresponding orthologs.

Out of the 33 ribosomal proteins of the 40S ribosome,18 are absent in the 30S (prokaryotic) small ribosomal subunit. In the 30S, the beak of the bird is constituted of solely rRNA however, in eukaryotes the proteins “erpS10” and “erpS12” are bound to the beak of the bird along with h33. [17]. Figure 6 shows the structures of ribosomal proteins of the 40S ribosomal subunit.

Ribosome biogenesis is studied extensively to understand how rRNAs are assembled with ribosomal proteins to form an active ribosomal subunit. The process of ribosomal biogenesis is initiated by transcription of the precursor rRNA (pre-rRNA). The ribosomal proteins are synthesized in the cytoplasm and are transported to the nucleus for subunit assembly. During the assembly of the ribosomal subunit, the pre-rRNA is modified, folded, and the spacer sequences are removed in the presence of the ribosomal proteins. The final assembly of the subunit is completed in the cytoplasm [21]. Extensive deletion and mutational analyses performed in yeast strains have led to the understanding of the pre-rRNA processing required to assemble the 40S subunit. The analyses have revealed that there are at least four rRNA cleavage events; two events occurring at each 5’ and 3’ end of the pre-rRNA. Each cleavage event is coupled with the binding of the ribosomal proteins. Failure in the ribosomal protein association during the assembly results in blocking of the subsequent cleavage steps [21].

Mutations in the ribosomal proteins or the chaperone proteins which assist ribosomal subunit assembly can result in inherited human diseases and disorders referred to as “ribosomopathies” [22] [23]. Specifically, Diamond Blackfan Anemia (DBA), 5q-syndrome, Treacher Collins Syndrome (TCS), and Schwachman Diamond Syndrome are the most well-characterized ribosomopathies [23]. In addition to these disorders, disruptions in the ribosomal

assembly can detrimentally affect the cell cycle progression, growth and can result in malignant transformations.

No.	Proteins	*H. sapiens*	*S. cerevisiae*	*E. coli*
1	RPSA	RPSA	RPS0A RPS0B	rpsB
2	RPS2	RPS2	RPS2	rpsE
3	RPS3	RPS3	RPS3	rpsC
4	RPS3A	RPS3A	RPS1A RPS1B	-
5	RPS4	RPS4X RPS4Y	RPS4A RPS4B	-
6	RPS5	RPS5	RPS5	rpsG
7	RPS6	RPS6	RPS6A RPS6B	-
8	RPS7	RPS7	RPS7A RPS7B	-
9	RPS8	RPS8	RPS8A RPS8B	-
10	RPS9	RPS9	RPS9A RPS9B	rpsD
11	RPS10	RPS10	RPS10A RPS10B	-
12	RPS11	RPS11	RPS11A RPS11B	rpsQ
13	RPS12	RPS12	RPS12	-
14	RPS13	RPS13	RPS13	rpsO
15	RPS14	RPS14	RPS14A RPS14B	rpsK
16	RPS15	RPS15	RPS15	rpsS
17	RPS15A	RPS15A	RPS22A RPS22B	rpsH
18	RPS16	RPS16	RPS16A RPS16B	rpsI
19	RPS17	RPS17	RPS17A RPS17B	-

20	RPS18	RPS18	RPS18A RPS18B	rpsM
21	RPS19	RPS19	RPS19A RPS19B	-
22	RPS20	RPS20	RPS20	rpsJ
23	RPS21	RPS21	RPS21A RPS21B	-
24	RPS23	RPS23	RPS23A RPS23B	rpsL
25	RPS24	RPS24	RPS24A RPS24B	-
26	RPS25	RPS25	RPS25A RPS25B	-
27	RPS26	RPS26	RPS26A RPS26B	-
28	RPS27	RPS27	RPS27A RPS27B	-
29	RPS27A	RPS27A	RPS31	-
30	RPS28	RPS28	RPS28A RPS28B	-
31	RPS29	RPS29	RPS29A RPS29B	rpsN
32	RPS30	RPS30	RPS30A RPS30B	-
33	RACK1		ASC1	-

Table 3. List of ribosomal proteins associated with small ribosomal subunit: The proteins are listed in the second column, the corresponding gene names from human, yeast, and *E. coli* system are listed in the next three columns respectively. This data was obtained from Ribosomal Protein Gene database (http://ribosome.med.miyazaki-u.ac.jp/)

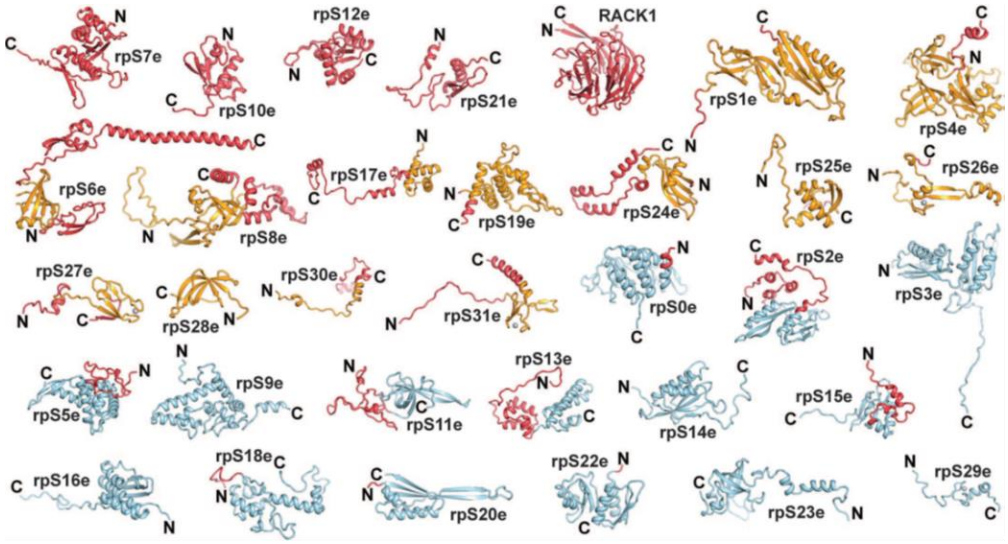

Figure 6. Ribosomal proteins associated with of the yeast 40S subunit: The N and C terminals are indicated for each protein. The proteins that are conserved in all kingdoms are shown in blue, the archaeal protein folds are shown in yellow, and the eukaryotic protein folds are shown in red. The figure is taken from the Ref [17] (*Science* 11 Feb 2011, Vol. 331,

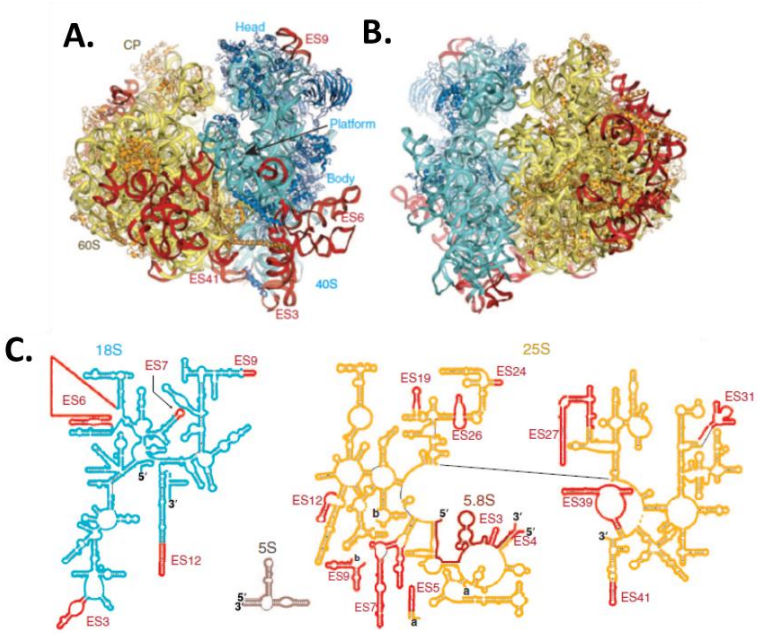

Figure 7. X-ray structure of yeast 80S ribosome: All the images are obtained and modified from the Ref [16] (Science vol 330 26 November 2010). (A) The *S. cerevisiae* 80S ribosomal view from E site (B) The *S. cerevisiae* 80S ribosomal view from A site. The rRNA from 40S is coloured light blue and the rRNA of 60S is coloured light yellow. The 40S ribosomal proteins are coloured dark blue and 60S ribosomal protein are coloured dark yellow. (C) Secondary structures of 18S rRNA, 5S rRNA, 23S rRNA, and 5.8S rRNA. (D) Top view of 80S ribosome indicating transition from nonratcheted state to ratcheted state. E. Solvent view of 80S ribosome indicating transition from nonratcheted state to ratcheted state.

No.		*H. sapiens*	*S. cerevisiae*	*E. coli*
1	RPL3	**RPL3**	**RPL3**	**rplC**
2	RPL4	**RPL4**	**RPL4A** **RPL4B**	
3	RPL5	**RPL5**	**RPL5**	**rplR**
4	RPL6	**RPL6**	**RPL6A** **RPL6B**	
5	RPL7	**RPL7**	**RPL7A** **RPL7B**	**rpmD**
6	RPL7A	**RPL7A**	**RPL8A** **RPL8B**	
7	RPL8	**RPL8**	**RPL2A** **RPL2B**	**rplB**
8	RPL9	**RPL9**	**RPL9A** **RPL9B**	**rplF**
9	RPL10	**RPL10**	**RPL10**	**rplP**
10	RPL10A	**RPL10A**	**RPL1A** **RPL1B**	**rplA**
11	RPL11	**RPL11**	**RPL11A** **RPL11B**	**rplE**
12	RPL12	**RPL12**	**RPL12A** **RPL12B**	**rplK**
13	RPL13	**RPL13**	**RPL13A** **RPL13B**	
14	RPL13A	**RPL13A**	**RPL16A** **RPL16B**	**rplM**
15	RPL14	**RPL14**	**RPL14A** **RPL14B**	
16	RPL15	**RPL15**	**RPL15A** **RPL15B**	
17	RPL17	**RPL17**	**RPL17A** **RPL17B**	**rplV**
18	RPL18	**RPL18**	**RPL18A** **RPL18B**	
19	RPL18A	**RPL18A**	**RPL20A** **RPL20B**	
20	RPL19	**RPL19**	**RPL19A** **RPL19B**	
21	RPL21	**RPL21**	**RPL21A** **RPL21B**	
22	RPL22	**RPL22**	**RPL22A** **RPL22B**	

23	RPL23	**RPL23**	**RPL23A** **RPL23B**	**rplN**
24	RPL23A	**RPL23A**	**RPL25**	**rplW**
25	RPL24	**RPL24**	**RPL24A** **RPL24B**	
26	RPL26	**RPL26**	**RPL26A** **RPL26B**	**rplX**
27	RPL27	**RPL27**	**RPL27A** **RPL27B**	
28	RPL27A	**RPL27A**	**RPL28**	**rplO**
29	RPL28	**RPL28**		
30	RPL29	**RPL29**	**RPL29**	
31	RPL30	**RPL30**	**RPL30**	
32	RPL31	**RPL31**	**RPL31A** **RPL31B**	
33	RPL32	**RPL32**	**RPL32**	
34	RPL34	**RPL34**	**RPL34A** **RPL34B**	
35	RPL35	**RPL35**	**RPL35A** **RPL35B**	**rpmC**
36	RPL35A	**RPL35A**	**RPL33A** **RPL33B**	
37	RPL36	**RPL36**	**RPL36A** **RPL36B**	
38	RPL36A	**RPL36A**	**RPL42A** **RPL42B**	
39	RPL37	**RPL37**	**RPL37A** **RPL37B**	
40	RPL37A	**RPL37A**	**RPL43A** **RPL43B**	
41	RPL38	**RPL38**	**RPL38**	
42	RPL39	**RPL39**	**RPL39**	
43	RPL40	**RPL40**	**RPL40A** **RPL40B**	
44	RPL41	**RPL41**	**RPL41A** **RPL41B**	
45	RPLP0	**RPLP0**	**RPP0**	**rplJ**

46	RPLP1	**RPLP1**	**RPP1A RPP1B**	**rplL**
47	RPLP2	**RPLP2**	**RPP2A**	
48	RPLP3			

Table 4. List of ribosomal proteins associated with large ribosomal subunit: The proteins are listed in the second column, the corresponding gene names from human, yeast, and E. coli system is listed in the next three columns. This data was obtained from Ribosomal Protein Gene database (http://ribosome.med.miyazaki-u.ac.jp/)

X-ray crystal structure of the of the *S. cerevisiae* 80S ribosome at the resolution of 4.15A° is shown in Fig. 7A and 7B. As explained earlier the ribosomal ratcheting movement during translation elongation involves rotation of 40S subunit. This is the most widely accepted mechanism for movement of 80S ribosome along mRNA during translation elongation. This X-ray crystal structure of 80S ribosome is consistent with the cryo-EM structures supporting the ratcheting of the ribosome during translation elongation [16]. The overall assembly and basic architecture of the 80S ribosome is similar to the prokaryotic ribosome and the landmark interactions between the small and large subunits are conserved. However, there are additional eukaryotes specific inter subunit interactions.

The molecular interactions near mRNA entry and exit channel of 80S ribosome highlights eukaryotic specific features [16]. The mRNA entry channel is formed by the interaction of h16 with the head region of the small ribosomal subunit. In prokaryotes the h16 is bent towards the body in contrast the h16 in eukaryotes is extended away from the body assuming an open conformation. In 80S ribosome the h16 is bare, without any interactions with the ribosomal proteins. This observation is consistent with the proposed model of ribosomal scanning where binding of initiation factors eIF1 and eIF1A keeps the mRNA entry channel in an open conformation (see Fig. 2) [16] [8] [24]. Similar interactions are assumed by IRES with the 80S

ribosome in order to initiate the cap-independent translation [25] [26]. Using this X-ray crystal structure eukaryotic mRNA exit channel was modeled. This structure predicted the presence of S28e above the mRNA exit channel in addition to the conserved S5 (S7p) protein. These are a few examples of the molecular interaction details obtained from the crystal structure of the yeast 80S ribosome.

E. The tRNA

There are two distinct tRNAs for amino acid methionine. One is specific for recognizing the start codon 5'-AUG-3' and is known as "Met-$tRNA_i^{met}$". Where the suffix "i" stands for the translation initiation step. The other tRNA recognizes the AUG codons occurring in the ORF. It is known as "Met-$tRNA_e^{met}$", where suffix "e" stands for the translation elongation. Both the tRNAs have an anticodon 5'-CAU-3' however, they differ in the post transcriptional modification. Adenine at the position 64 in the T loop of tRNA is post-transcriptionally modified with an attachment of 2'-O-ribosyl phosphate [27] . The eEF1A contacts the T loop of the tRNA and any modifications at the A64 position in the T loop sterically hinders the interaction of the tRNA with the eEF1A. This steric hinderance prevents the interaction of Met-$tRNA_i^{met}$ with the eEF1A . Rit1 enzyme in *S. cerevisiae* modifies the tRNA.

Four genes encode for the Initiator Met tRNA (*IMT1, IMT2, IMT3,* and *IMT4*) in *S. cerevisiae.* Five genes encode for the elongator Met-tRNA (*EMT1, EMT2, EMT3, EMT4,* and *EMT5*).

Functionally important features of Met-$tRNA_i^{met}$ include base pairing between A1:U72, C3:G70 in the acceptor stem of the tRNA. Yeast cells in which the A1:U72 base pair of Met-$tRNA_i^{met}$ when swapped with G1:C72, exhibit severe growth defects [27]. It is reported that the base

pairing A1:U72 is not only important for the association of Met-tRNA$_i^{met}$ with eIF2 but also for the binding of TC to 40S ribosome [28].

The second most important and conserved interaction is between consecutive bases G29:C41, G30:C40, G31:C39. It is observed that mutations in G31:C39 interaction renders cells with compromised accuracy for start site selection. Similar results were obtained with mutation in C2:G70 base pairing. Both of the base pairings are partially responsible for the TC formation and binding of the TC to 40S ribosome [29].

1.2 Regulation of eukaryotic translation

Translation is an energy-consuming process thus, regulation is exerted at the initiation step of the translation. Regulation of translation initiation is pivotal in cell growth, cell morphogenesis, cell proliferation and cell differentiation. Deregulation of translation initiation results in uncontrolled cell growth; uncontrolled cell growth leads to cancerous growth [30]. Apart from cancer, deregulation of translation is implicated in progression of Alzheimer's, and other inheritable diseases [31] [32] [33]. Cultivating the mechanistic understanding at the molecular level of translation and its regulation helps in identifying drug targets for the treatment of such diseases. Recent advances in the structural biology and crystallographic methodologies enable us to understand the detailed mechanism of translation initiation and its regulation.

Under conditions of cellular stress (oxidative stress, starvation, heat shock stress, and stress due to imbalance in protein homeostasis) global protein synthesis is downregulated but, simultaneously, translation of stress specific transcription factors is upregulated. Global protein synthesis is regulated by controlling formation of the cap-binding complex (eIF4F) and the ternary complex. The mechanisms of which are discussed in subsequent detail.

Transcript specific translation regulation can be attributed to the *cis*-acting elements of the mRNA present in the 5' and 3' untranslated regions (5'-UTR and 3'-UTR) as well as *trans*-acting factors such as transcript specific RNA binding proteins. One of the c*is*-acting factors is the presence of an RNA secondary structure in the 5'-UTR. This thesis is aimed at understanding the mechanisms of translational regulation by an RNA secondary structure. In addition to the RNA secondary structure other *cis*-acting elements such as (I) Nucleotide sequence around AUG codon (Kozak sequence), (II) Length of the 5'-UTR, and (III) upstream open reading frame (uORF) are also known to regulate translation. Regulation of translation by an RNA secondary structure is discussed as follows.

A. Translational regulation by an RNA secondary structure:

mRNA is generally single stranded and often folds on itself. This results in base-pairing of the mRNA with itself; known as mRNA secondary structure. RNA secondary structure present in the 5'-UTR negatively regulates its translation; exact mechanism of which is not clearly understood. Multiple factors decide the mechanism of inhibition of translation by RNA secondary structure including (I) the position of RNA secondary structure with respect to the mRNA cap, (II) the stability of the RNA secondary structure, and (III) the position of the RNA secondary structure with respect to the start codon. These factors are discussed here.

(I) The position of RNA secondary structure with respect to the mRNA cap: RNA secondary structures present within 45 nucleotides from the mRNA cap are classified as cap-proximal secondary structures and if they are present beyond 45 nucleotides they are called as cap-distal [34]. Genome wide analysis of the translational regulation by RNA secondary structure indicates that cap-proximal RNA secondary structures inhibit translation probably by preventing the loading of Pre-initiation complex (PIC) on the mRNA. This result is supported by another

study showing that the PIC occupies ~ 30 nucleotides on the mRNA [35] [36] [37]. Taken together, it can be deduced that cap-proximal RNA secondary structures possibly inhibit loading of PIC due to insufficient room on the 5'-UTR.

As opposed to the cap-proximal, the cap-distal secondary structure inhibits translation probably by impeding the scanning movement of PIC along the 5'-UTR. During the scanning step, initiation factors eIF1 and eIF1A keep the mRNA entry channel in an open conformation also called as scanning competent conformation (see Fig. 2B). Structural studies of the PIC indicate that the diameter of the mRNA entry channel can accommodate only single stranded mRNA [24]. Hence, the RNA secondary structure needs to unwind before entering the channel during the scanning step. Here, the stability of RNA secondary structure comes into play.

(II) The stability of RNA secondary structure: *In vivo* and *in vitro* studies show that, cap-distal RNA secondary structure with free energy of around -30 kcal/mol is melted by scanning PIC [38]. Moreover, the scanning PIC is able to initiate translation at the start codon (AUG) buried in the secondary structure [39]. Further studies show that secondary structures with free energy of around -50 kcal/mol to -60 kcal/mol significantly inhibit translation initiation *in vivo* as well as *in vitro.* Surprisingly, elongating ribosomes can scan pass such secondary structures [38]. Collectively, once attached to the mRNA, the PIC can melt the cap-distal secondary structures depending on their stability. Additionally, the melting of the secondary structure is not solely dependent on the helicase activity of PIC. eIF4A and Ded1 are two RNA helicases present in yeast that play a major role in unwinding the secondary structures.

(III) The position of the RNA secondary structure with respect to the start codon: The secondary structure present 10-12 nucleotides downstream of an AUG codon can positively affect the translation initiation from that start codon. This is because ribosomes spend more time on that

AUG due to stalling by the secondary structure. Such secondary structure is present mostly downstream of AUG not surrounded by the Kozak sequence (weak AUG). This mechanism aids in translation initiation from a weak AUG codon [40] [38]. Another example where RNA secondary structure positively regulates translation is in some viral RNAs.

Some viral RNAs have a stem loop structures in the 5'-UTR region which allows the PIC to directly access the start codon. These elements are called "internal ribosomal entry site" (IRES). IRES elements promote translation initiation in a cap independent manner as opposed to the secondary structures in the 5'-UTR region that inhibits translation initiation [41] [42] [43].

B. Role of RNA helicases in unwinding the RNA secondary structures:

eIF4A is an RNA helicase which forms a cap-binding complex with eIF4G (a scaffold protein) and eIF4E (a cap-binding protein) (see Fig.1 and Table 1). eIF4A exhibits RNA dependent helicase activity in the presence of ATP. Association of eIF4A with the eIF4G protein stimulates the ATPase activity of eIF4A.

eIF4A belongs to the family of DEAD box helicases. It has two RecA-like domains and a conserved Walker B motif (D-E-A-D-). eIF4A is a minimal DEAD box protein with only the conserved core [44]. eIF4A has N terminal and C terminal RecA-like domains called as "NTD and CTD" respectively, they are connected by an 11 amino acid long linker (see Fig. 8B). The ribbon representation of the crystal structure of yeast eIF4A is depicted in Fig 8B. The NTD (yellow), CTD (blue). The NTD has sequence motifs Q, I (walker A), Ia, Ib and II (DEAD). CTD contains sequence motifs III, IV, V, and VI [45]. The sequence and position of motifs in NTD and CTD are listed in Fig 8A.

A.

Motif name	Motif sequence	Amino acid number	Function
Q	GAxxPOxxQ		ATP binding
I/Walker A	ASQSGTGKT	65- 72	ATP binding
Ia	PTRELA	97- 102	RNA binding
Ib	TPGR	145- 148	RNA binding
II/Walker B/ DEAD	DEAD	169- 172	
SAT	SAT	200- 202	
IV	VIFCNTRR	263- 270	RNA binding
V	RGID	321- 325	RNA binding
VI	HRIGRGGR	345- 352	ATP binding

B.

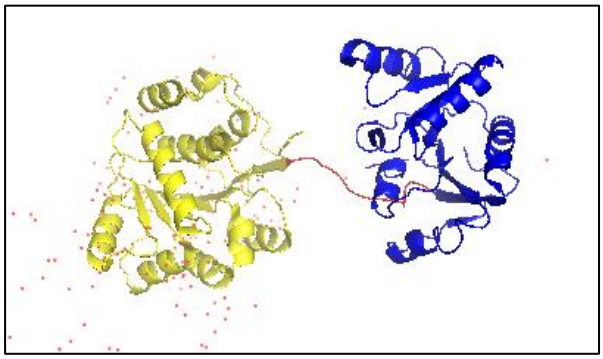

Figure 8. DEAD-box helicase eIF4A: (A) The motifs of DEAD box helicase are listed in the table with corresponding conserved sequence and amino acid positions in the yest eIF4A.The motifs from NTD and CTD are colored yellow and blue respectively. (B) The ribbon representation of crystal structure of the yeast eIF4A (PDB ID: 1FUU). The NTD is colored in yellow, CTD is colored in blue and the linker is colored in red.

eIF4A is a potential target of anti-cancer drugs. Silvestrol and Hippuristanol are two anti-cancer drugs which inhibit eIF4A activity [46] (shown in Fig 11). Pentamine A and FL3 are also effective in inhibiting translation by blocking eIF4A activity.

Ded1 is another DEAD box RNA-dependent helicase involved in the translation initiation of structured 5'-UTR of yeast mRNAs. The RNA unwinding activity of Ded1 is independent of the interaction with eIF4G in contrast to eIF4A [47]. Thus, it is observed that, eIF4A unwinds the cap-proximal RNA secondary structures as opposed to Ded1 which unwinds cap-distal RNA secondary structures [34] [48].

C. Translational regulation by RNA secondary structure bound by a suppressor protein:

In some cases, inhibition of translation is achieved not solely by RNA secondary structure but by a cognate suppressor protein binding to the secondary structure. This interaction of *trans*-acting protein with the RNA secondary structure prevents translation initiation by precluding the loading of PIC on the mRNA.

The iron regulatory protein is an exclusive example of a suppressor protein which binds to the secondary structure of mRNA and precludes the loading of the 43S PIC (see Fig. 9) [49] [50]. Ferritin mRNAs (ferritin H chain and ferritin L chain) have a stem-loop structure referred to as, iron responsive elements (IRE) in the 5'-UTR. Iron regulatory protein-1 (IRP-1) binds to the IRE (secondary structure) and precludes the loading of PIC on the ferritin mRNAs in iron deficient cells. Initial results showed that increasing the distance of the IRE (secondary structure) from the mRNA cap relieves translational inhibition. This result suggests that, increasing the distance of IRE from the mRNA cap provides enough room for the loading of 43S PIC on the mRNA [49]**.** Further experimental evidence clearly showed that binding of IRP-1(suppressor protein) to IRE (secondary structure) prevents recruitment of the PIC [50].

This is the sole, best studied and most well-established example of an RNA secondary structure and a suppressor protein binding together to inhibit translation initiation. However, the precise mechanism of how an RNA secondary structure regulates translation remains elusive. This thesis is aimed at understanding how RNA secondary structure regulates translation using yeast *HAC1* mRNA as a model mRNA. The structure of *HAC1* mRNA is discussed in the following section. Mechanism of translational regulation in *HAC1* mRNA is discussed in chapter 2 and chapter 3.

D. *Cis*-acting factors regulating translation initiation:

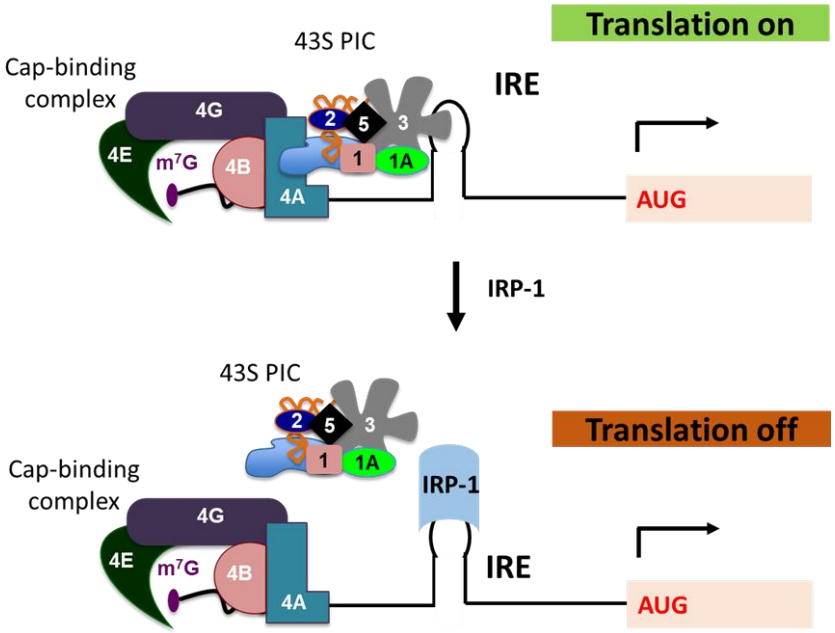

Figure 9. Translational regulation by IRE and IRP-1 protein: Top panel shows recruitment of cap-binding complex and PIC in presence of the IRE (secondary structure). However, bottom panel shows that binding of IRP-1 (suppressor protein) to IRE (secondary structure) does not affect the recruitment of cap-binding complex but precludes PIC loading on the mRNA. This figure is adapted and modified from the reference Mukhentaler et al, 1998 (Molecular Cell, Vol. 2, 383–388, September,1998)

As discussed above, other *cis*-acting elements such as (I) Nucleotide sequence around the AUG codon (Kozak sequence), (II) Length of the 5'-UTR, (III) upstream open reading frame (uORF), and (IV) miRNA binding sites (in human system) are also significant in regulating translation.

No.	Disease	Cause
1	a-Thalassemia	Mutation at position -3 (A ➔C change) of the globin gene
2	Sporadic breast cancer	Mutation at position -3 (G ➔C change) of the BRCA1 gene
3	Androgen insensitivity syndrome	Mutation at position +4 (G ➔A change) of the androgen receptor gene
4	Ataxia with vitamin E deficiency	Mutation at position -1 (C ➔T change) of the a-tocopherol protein gene

Table 5: Diseases caused due to mutations in the Kozak sequence: Table gives a list of diseases with corresponding gene names and point mutations in the Kozak sequence.

(I) Nucleotide sequence around AUG codon (Kozak sequence): The AUG codon with a purine (A/G) at the position -3 and +4 is the preferred translation initiation site identified by Marylin Kozak. Named after her, the "Kozak consensus sequence" is defined as GCC(A/G)CC**AUG**(A/G). The AUG codon with the Kozak consensus sequence is said to be in a good context. During the scanning step, the AUG nearest to the cap is the preferred translation start site unless it is not in a good context [51]. Identification of the correct start codon depends on good context. Mutations around the start codon result in incorrect identification of the start codon. Mutations around the start codon of specific transcripts leads to diseased conditions as a result of said mutations (listed in Table 5) [52] [53]. Hence, while genetically engineering all the constructs expressing various mutant forms of *HAC1* mRNA (discussed throughout the thesis), great precaution was taken to preserve the Kozak sequence around the stat codon AUG.

(II) Length of the 5'-UTR: A long 5'-UTR facilitates binding of multiple PIC at once, thus greatly increasing the rate of translation initiation. PIC occupies ~ 30 nucleotides on the mRNA [35] [36] [37]. Our bioinformatic analysis of yeast transcriptome shows that median length of yeast 5'-UTR is 54 nucleotides (see chapter 3, Fig 19). Our results discussed in chapter 3 shows that increasing 5'-UTR length provides sufficient room for loading of the PIC on the *HAC1* mRNA.

(III) Upstream ORF (uORF): An ORF present in the 5'-UTR and upstream of the authentic start codon is called an upstream ORF (uORF). When an upstream AUG (uAUG) of the uORF is in-frame with the downstream AUG, the ribosome skips the uAUG during scanning. This is referred to as "leaky scanning". Interestingly, mRNA is able to generate two proteins initiated at two different AUGs, one being shorter at the N terminal end [40] [54]. Translating uORFs mostly

inhibit translation of the downstream ORF because, translating ribosomes pauses on the uORF preventing the initiation complex to access the downstream AUG [13].

A unique mechanism of translational regulation is observed on general control non-derepressable 4 (*GCN4*) mRNA (encoding a transcription factor that activates amino acid biosynthetic genes). Four uORFs are present in the 5'-UTR of the *GCN4* mRNA. *GCN4* mRNA is thus, translationally repressed under physiological conditions because ribosomes fail to reinitiate translation after translating uORF-2 to uORF-4. The PIC requires ample ternary complex (TC) to reinitiate translation at uORF-2, uORF-3 and uORF-4. Under amino acid starvation conditions, the kinase "Gcn2" in yeast phosphorylates eIF2α-ser51. The phosphorylated form of eIF2α-ser51 acts as inhibitor of the guanine exchange factor (GEF), responsible for exchanging eIF2-GDP to eIF2-GTP, lowering the overall ternary complex (TC) concentration in the cell. Thus, under amino acid starvation conditions, concentration of TC falls and only 50% of the 43S PIC are bound by TC. The PIC translates the uORF-2, uORF-3 and uORF-4 and falls off before reaching the start codon of *GCN4* transcript. The remaining 50% of PIC, which cannot reinitiate translation due to lack of TC, continues scanning the uORFs until it reaches the start codon of *GCN4* transcript. Once 43S PIC are bound by TC, they start translating the *GCN4* ORF. This is the mechanism of derepression of *GCN4* translation. Low concentration of TC in the cell also shuts off global translation under stress conditions [13].

(IV) miRNA binding sites:

Micro RNAs (miRNA) are small non-coding RNA (ncRNA) which are 19-25 base pairs in length. miRNA regulates gene expression by either modulating mRNA translation or mRNA stability [55]. Mature miRNA attached to the RISC complex binds to the complementary nucleotides (6-8 nucleotides) in the 3'-UTR of the target mRNA. This interaction inhibits mRNA expression [56].

Translational silencing by miRNA is achieved by direct inhibition of the translation initiation or mRNA deadenylation, mRNA decapping, and degradation [55][57]. Recent reports suggest that miRNAs directly inhibit translation by preventing association of RNA helicase eIF4A (eIF4AI, eIF4AII or eIF4AIII) with the target mRNA in the human system [58][59] [60]. Apart from translation initiation, miRNAs are shown to inhibit translation immediately after initiation by rapid deadenylation and degradation in Drosophila and zebrafish systems. It has been proposed that miRNA inhibits subunit joining by interacting with eIF6 factor however, this was only observed in *Caenorhabditis elegans* and was not confirmed in any other system [61].

One miRNA can bind and regulate translation of multiple target mRNAs and target mRNA can have binding sites for more than one miRNA [61]. The recognized miRNAs in human system are rapidly increasing and are classified in miRNA families. miRNAs belonging to a family are simultaneously expressed, colocalized and regulate a set of genes such as immune system regulatory genes or growth and development genes [61].

Growth, development and metabolism requires temporal and tissue specific gene regulation. miRNA mediated regulation is involved at all the stages of growth and development. Deregulation of miRNA mediated regulation results in cancer development and progression [57].

E. Global regulation of translation:

Global regulatory mechanisms are exerted on the two crucial steps of translation initiation: formation of the (I) cap-binding complex (eIF4F), and the (II) Ternary complex. The detailed account of molecular mechanisms accompanied by regulatory mechanisms of the formation of eIF4F and ternary complex are discussed below.

(I)Formation of the cap-binding complex (eIF4F): The cap-binding complex binds to the m^7G cap at the 5'-end of eukaryotic mRNA. The cap-binding complex consists of proteins eIF4E, eIF4G, eIF4A and eIF4B. Out of which eIF4G acts as a scaffold protein on which the cap-binding complex is assembled. eIF4G has binding sites for the proteins: eIF4A, eIF4E, eIF5, Pab1 and the mRNA (Fig 10A). Yeast has two 4G proteins; eIF4G1 encoded by *TIF4631* and eIF4G2 encoded by *TIF4632*. Both are functionally redundant proteins which share only 51% identity [5]. Recent research indicates that a class of eIF4G binding proteins hinder the recruitment of the Preinitiation complex (PIC) on mRNA [62].

eIF4E binds to the m^7G cap of the mRNA. Interaction of eIF4E with eIF4G forms a ternary complex of m^7G-eIF4E-eIF4G. The crystal structure of this ternary complex in ribbon representation is shown in Fig. 10B. The m^7G cap is sequestered by the two tryptophan residues W58 and W104 of eIF4E. eIF4G (amino acids 393-460) (Fig. 10B) forms a bracelet shaped structure around eIF4E. This molecular bracelet is formed by five α-helices, which are conserved in human, rabbit and mouse eIF4G. The helix 4 has the well characterized eIF4E binding motif (YxxxFLL) [63].

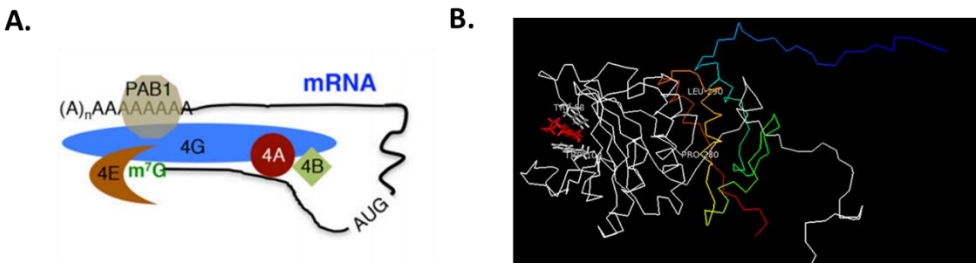

Figure 10. Cap-binding complex; eIF4F: (A) Schematic representation of the cap-binding complex: eIF4F. Image is taken from the review article by Tom E. Dever et al, 2016 (Genetics, Vol. 203, 65–107 May 2016) . (B) The crystal structure of ternary complex formed by m^7G-eIF4E-eIF4G is represented in the ribbon deoiction. The m^7G mRNA cap is shown in red color, the eIF4E is shown in white color and the eIF4G (amino acids 393-460) is shown in the rainbow color.

Binding of eIF4G to eIF4E increases affinity of 4E towards the m^7G cap. Also, the large interaction surface between eIF4G/4E imparts a slow dissociation constant forming a stable solution structure [5] [63]. Experimental evidence suggests that not all the capped mRNAs exhibit identical affinity for eIF4E. Differential affinity towards eIF4E is attributed to the presence of mRNA secondary structure near the mRNA cap [64].

4E binding protein (4E-BP) sequesters the cap binding protein eIF4E. When eIF4E is sequestered by 4E-BP, ribosome recruitment is hindered, without which cellular mRNAs fail to initiate translation [1] [2] [65]. 4E-BP mediated translation is monitored by phosphorylation by mammalian target of rapamycin (mTOR) in human system [66]. Inhibitors have been designed to target eIF4E and mTOR pathway in order to prevent uncontrolled cancerous growth (Fig. 11) [67].

Pab1 binds to the poly A tail of the mRNA. Interaction of eIF4G with Pab1 and eIF4E circularizes the mRNA (Fig. 10A). mRNA circularization enhances the rate of translation by efficient recycling of the dissociated ribosomal subunits after each translation cycle. Yeast eIF4B is also part of eIF4G-eIF4A complex. Studies have revealed that interaction of eIF4G with eIF4A enhanced ATPase activity [68].

Taken together, the association of the cap-binding complex activates the mRNA for initiating translation. Hence, global translational regulatory mechanisms are exerted on the formation of the cap-complex. These mechanisms decrease pan translation of all cellular mRNAs.

(II) Formation and recycling of the Ternary complex (TC): The ternary complex is formed by initiator met-$tRNA_i^{met}$, GTP, and eIF2. The ternary complex delivers the initiator met-$tRNA_i^{met}$ to the 40S ribosome.

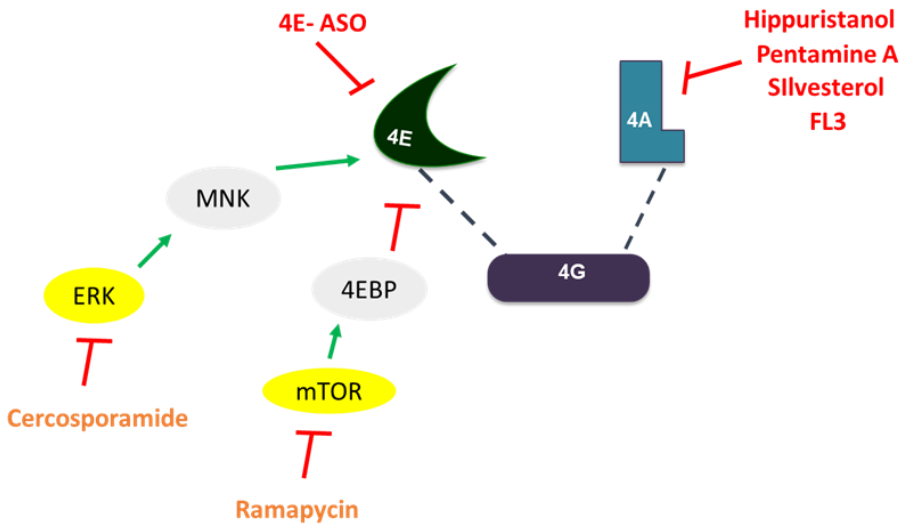

Figure 11. Cap-binding complex as a drug target: The cap-binding complex proteins (eIF4A, eIF4E and eIF4G) are shown at the center. The inhibitors of eIF4E and eIF4A are listed in red color. The primary effector protein is shown in gray oval and secondary effector is shown in fluorescent yellow oval. The inhibitors of the secondary effector proteins are shown in orange color.

eIF2 is a heterotrimeric protein with three subunits; α, β and γ. They are encoded by *SUI2*, *SUI3*, and *GCD11* genes respectively. These genes were discovered in 1998 by Donahue while studying a histidine auxotrophic yeast strain. He replaced the AUG start codon of the *HIS3* gene by AUU and studied the spontaneous mutations. He identified two suppressor mutations in two genes that allowed the translation of *HIS3* mRNA from the third codon; coding for leucine. Those two genes were named as Suppressor of initiator codon (for both *SUI2* and *SUI3*). eIF2γ is the keystone of TC. It binds to GTP and Met-tRNA$_i^{met}$. The TC is recycled after every translation cycle.

The major regulation on translation initiation is exerted by the reversible phosphorylation of initiation factors, which leads to depletion of active initiation factors in the cell. The best studied example is eIF2α-ser51 phosphorylation under stress conditions [2] [69]. Association of the active TC with the small ribosomal subunit forms the pre-initiation complex (PIC) and is an essential step to initiate translation. Phosphorylated eIF2α fails to form a TC. Under low concentrations of

TC, PIC cannot initiate translation. This reduces the overall rate of translation in the yeast cell. In yeast, eIF2α is phosphorylated by a kinase "Gcn2". In humans four kinases: PERK (under ER stress conditions), PKR (upon viral infections), HRI, and Gcn2 (amino acid starvation conditions) phosphorylate eIF2α-ser51 under various stress conditions as previously mentioned. This is the most effective mechanism to dampen global protein synthesis. This mechanism also supports the expression of required transcription factors under stress conditions. One of the most well studied examples is of *GCN4* expression under amino acids starvation conditions.

1.3 Translational silencing in yeast *HAC1* mRNA:

Endoplasmic reticulum (ER) stress conditions arise after accumulation of unfolded proteins in the lumen of the ER. Under conditions of ER stress the unfolded protein response (UPR) pathway is activated which results in translation of *HAC1* mRNA to produce Hac1 protein. The Hac1 protein is required by yeast to survive the ER stress conditions [70]. Hac1 protein is a bZIP transcription factor that binds to the unfolded protein responsive element (UPRE) in the promoters of various UPR responsive genes. Hac1 activates the expression of UPRE under the ER stress conditions [71] [72].

A unique mechanism is involved in regulating the translation of *HAC1* mRNA. Even though *HAC1* mRNA is present in the cells, its expression is regulated post-transcriptionally via an mRNA secondary structure formed by its 5'-UTR and an unconventional cytoplasmic intron [73]. Under conditions of ER stress an active kinase/RNase "Ire1", cleaves the inhibitory intron at two positions [74] [75]. The two exons are then ligated by tRNA ligase. Then mature $HAC1^{s}$ (s stands for spliced) mRNA undergoes translation to produce Hac1 protein. The Hac1 protein enters the nucleus, binds to the UPR elements, and activates gene expression (see Fig. 12).

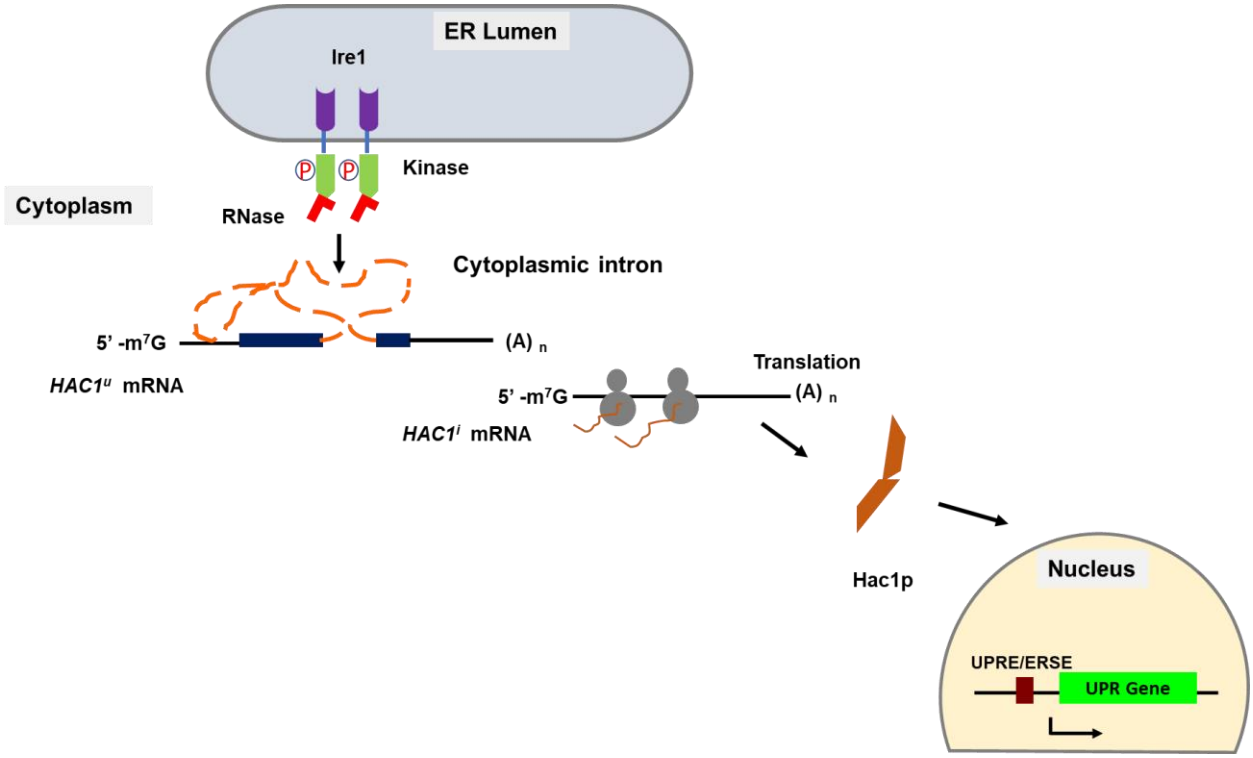

Figure 12. Schematic representation of yeast UPR pathway: Accumulation of unfolded proteins in the lumen of ER activates Ire1. Ire1 is ER resident transmembrane protein. It has luminal domain shown in purple, transmembrane domain shown in blue lines, and cytoplasmic Kinase/RNase domain indicated in green and red figures respectively. Active Ire1 cleaves un spliced *HAC1*^{u} mRNA. The intron is spliced, and mature mRNA (HAC1^{i}) produces Hac1 protein (shown in brown color). Hac1 enters the nucleus and activates UPR target genes whose products help in alleviating ER stress.

It is known that RNA secondary structure in the 5'-UTR of an mRNA negatively regulates translation of that mRNA [51]. However, the precise mechanism of inhibition of translation is not known. The presence of an inherent secondary structure in the 5'-UTR of *HAC1* mRNA makes it a model mRNA to study regulation of protein synthesis by an RNA secondary structure. In this thesis, we uncover the mechanism of translational regulation of *HAC1* mRNA by an RNA secondary structure.

A. Yeast *HAC1* as a model mRNA:

Yeast *HAC1* mRNA is the model mRNA to study translational regulation. The structure of *HAC1* mRNA is shown in Fig 13. *HAC1* mRNA has a 5'-UTR of 69 nucleotides, Exon 1 of 661

nucleotides, Exon 2 of 57 nucleotides, a long intron of 252 nucleotides sandwiched between two exons, and long 3'-UTR of 461 nucleotides. The 17 nucleotides of the intron base-pair with the 5'-UTR to form an RNA secondary structure which keeps the *HAC1* mRNA translationally silent under physiological conditions.

In rare cases, a 230 amino acid long Hac1 protein; known as $Hac1^{u}$ ("u" stands for un-stress conditions) is obtained from exon 1. Under stress condition, the inhibitory intron is removed, and the two exons are ligated to mature *HAC1* mRNA. This mRNA, when translated, produces a 238 amino acid long $Hac1^{i}$ ("i" stands for induced conditions) which will be henceforth referred as, "$Hac1^{i}$ protein". The $Hac1^{i}$ is an active transcription factor and can activate UPRE dependent transcription. The last 8 amino acids encoded by exon 1 are replaced by an 18 amino acid long tail encoded by exon 2 in $Hac1^{i}$ (Fig. 13B) [76]. The N- terminal end of the Hac1 protein contains a DNA binding domain and C- terminal end contains a transcription activation domain.

Both *IRE1* and *HAC1* genes are essential for the yeast to survive under stress conditions (see UPR pathway in Fig 12). Conventionally, the stress conditions are mimicked in cells by treating them with either tunicamycin (Tm) or Dithiothreitol (DTT). DTT reduces the disulfide bonds in folded proteins while tunicamycin inhibits N-linked glycosylation of the proteins. Thus, both treatments likely result in accumulation of unfolded proteins in the lumen of ER mimicking ER stress conditions.

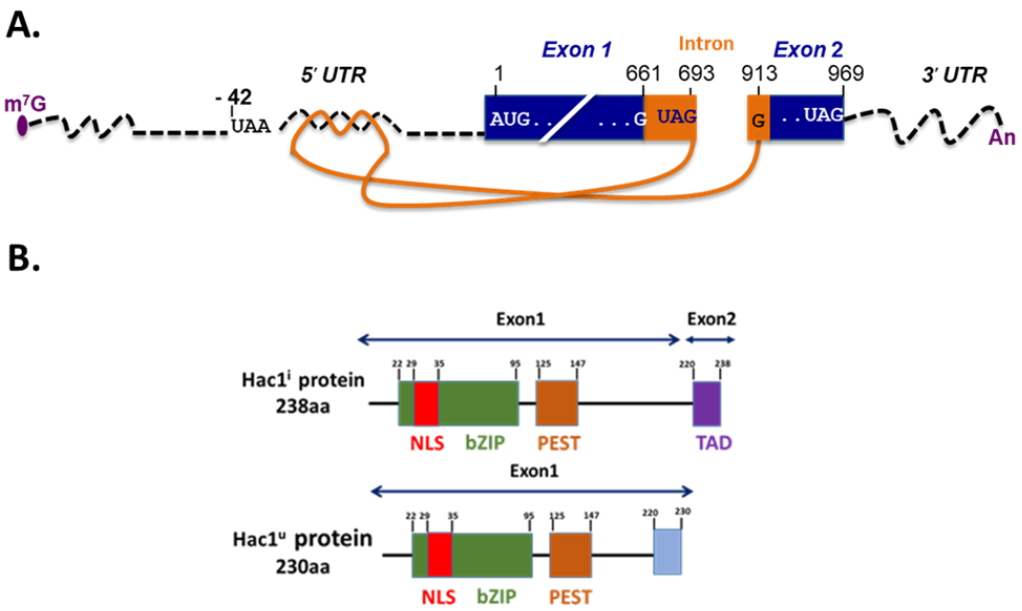

Figure 13. Schematic representation of *HAC1* mRNA and Hac1 protein: (A) Schematic representation of the *HAC1* mRNA. The mRNA cap is shown as m^7G, the 5'-UTR is shown in black dashed lines, exon1 is shown in blue box, intron is shown in orange color, exon 2 is shown is a small blue box followed by 3'-UTR shown in back dashed line. The intron base pairs with the 5'-UTR region forming the RNA secondary structure. (B) Schematic representation of domain organization of Hac1 protein. Domain organization of the $Hac1^i$ protein (238 amino acids) and $Hac1^u$ (230 amino acids). Both the proteins have N-terminal nuclear localization signal; NLS, basic leucine zipper; bZIP and degron sequence called as PEST. $Hac1^u$ lacks the transcription activation domain; TAD encoded by exon2. The domain organization of Hac1 protein is obtained and modified from Ref [76] (Cell. 1996 Nov 1;87(3):391-404)

B. RNA secondary structure inhibits translation initiation in *HAC1* mRNA:

It has been reported that the RNA secondary structure inhibits translation elongation in *HAC1* mRNA [73]. This conclusion was based on a key observation that the *HAC1* mRNA was predominantly associated with the polyribosome. The proposed model suggests that, as soon as the *HAC1* mRNA enters the cytoplasm, it is engaged by ribosomes, which start translating the transcript. When the intron enters the cytoplasm, it base-pairs with the 5'-UTR which traps the translating ribosomes on the mRNA. Thus, *HAC1* mRNA is associated with polyribosomes and the RNA secondary structure inhibits translation elongation [73] .

Using a random mutational screen as well as site-directed mutagenesis, we identified point mutations within the 5'-UTR*intron interaction site, which de-repress translation of the un-spliced *HAC1* mRNA. We also show that insertion of an in-frame AUG start codon upstream of the RNA secondary structure releases the translational block. Moreover, overexpression of translation initiation factor eIF4A, a helicase, enhances production of Hac1 from an mRNA containing an upstream AUG start codon at the beginning of the base-paired region. Thus, our results suggest that translation of *HAC1* mRNA is inhibited at the initiation stage [77]. These results are published, and they are discussed further in chapter 2 of this thesis.

C. RNA secondary structure inhibits either loading of PIC or helicase eIF4A on *HAC1* mRNA:

In chapter 3, we aimed at deciphering the detailed mechanisms of translation initiation block. We asked the following questions: 1) What step of translation initiation is inhibited by RNA secondary structure? 2) Is there any suppressor protein likely involved in translational silencing of *HAC1* mRNA? Our results suggest that the loading of pre-initiation complex (PIC) on the *HAC1* mRNA is inhibited by the RNA secondary structure.

D. *HAC1* mRNA transcript isoform plays a role in adaption to ER stress:

Recent genomic and transcriptomic studies suggest that Hac1 protein is produced from two overlapping mRNA isoforms. One isoform (dubbed here "*HAC1a"*) is the canonical mRNA (Fig 10). The second isoform (dubbed here "*HAC1b"*) is a newly discovered mRNA, containing a long 5'-UTR, exon1, and a short 3'-UTR. As reported earlier, intron of *HAC1a* binds to its 5'-UTR and keeps mRNA translationally silent. Like *HAC1a,* we find that, *HAC1b* remains translationally silent by an unknown mechanism. We find that *HAC1b* could translate an active transcription

factor only under conditions of cellular stress. Moreover, we find that the *HAC1b* mRNA level is substantially low compared to *HAC1a* and the Hac1b protein is stabilized in the absence of Duh1 (a member of a proteasome complex). These results suggest that translational silencing of *HAC1b* mRNA could be because of rapid degradation of both mRNA and its translational product. Taken together, we study the role of a previously undefined *HAC1b* transcript in adaptation to cellular stress; these results are further in chapter 4.

2. *HAC1* mRNA is translationally repressed at the initiation stage

These results are published in JBC with the tittle " Evidence that base-pairing interaction between intron and mRNA leader sequences inhibits initiation of *HAC1* mRNA translation in yeast" (J Biol Chem. 2015 Sep 4;290(36):21821-32).

2.1 Introduction

The regulation of translation of the yeast *HAC1* mRNA involves base-pairing of the 5'-UTR and the intron [73] . The unique intron (spanning nucleotides G661 to G913, with the adenine of AUG start codon assigned as 1) is not spliced in the nucleus by the spliceosome but instead is retained in the mRNA that is exported to the cytoplasm [78]. Previous work established that base-pairing interactions between elements in the intron and the 5'-UTR repress translation of the unspliced *HAC1* mRNA[73]. Their results showed that *HAC1* mRNA translation is repressed at the elongation stage. Under conditions of endoplasmic reticulum (ER) stress, the endoribonuclease Ire1 is activated and cleaves both exon-intron boundaries of the *HAC1* mRNA in the cytoplasm (see Fig. 7). The two exons are then ligated by tRNA ligase [79], resulting in an altered ORF with a new codon starting at nucleotide G661 and a UAG stop codon at nucleotide 963 (see Fig. 11A and Fig. 8B). The Hac1 protein is a transcription factor and primary effector of the unfolded protein response (UPR) [80] [81]. It binds to the UPR elements (UPRE) in the promoters of genes and activates expression of a set of genes involved in maintaining protein homeostasis and alleviating ER stress [75][76] [71][82] (see Fig. 12). The Hac1 proteins produced from both the spliced (238 amino acids) and unspliced (230 amino acids) mRNA are functional transcription factors [83]. Understanding the mechanism of the repression of translation of the unspliced *HAC1* mRNA is key to understanding the UPR pathway in yeast.

To gain mechanistic insights into *HAC1* mRNA translation, a genetic screen was performed to identify mutation(s) that can release the translational block in a splicing-deficient and translationally inert *HAC1* mRNA allele (i.e. *HAC1-G661C*). In this screen we identified a single base mutation (i.e. G771A) in the intron that is predicted to disrupt the base-pairing interaction with the 5'- UTR [73]. This suggests that the mutation G771A in the intron results in the disruption of the 5'- UTR*intron interaction leading to translation of the splicing deficient *HAC1-G661C,G771A* mRNA allele. This suggests that the disruption in 5'-UTR*intron base-pairing interaction relieves the translational block. This result is consistent with the previous observation that the 5'-UTR*intron interaction interferes with translation initiation on the *HAC1* mRNA [73]. Hence, we introduced an upstream in-frame AUG codon before the 5'- UTR*intron base-pairing region which relieved the translational block on the *HAC1* mRNA. These results are discussed in this chapter.

2.2 Single base-pair mutation in the intron derepresses translation of unspliced *HAC1* mRNA

To gain mechanistic insights into the 5'-UTR*intron-mediated translational repression, we screened for intragenic suppressor mutations in a splicing-deficient and translationally inert *HAC1-G661C* allele. We expected that second-site mutation(s) might produce a Hac1 protein from the unspliced mRNA. The suppressor screen was performed using *E. coli XL-1* Red cells from Stratagene. This *E. coli* strain is deficient in efficient DNA repair mechanisms due to mutations in genes *mutS*, *mutD*, and *mutT*. Hence, this strain is suitable for generation of random mutations in the DNA.

The *hac1*Δ yeast strain carrying an empty vector grew well on synthetic dextrose (SD) medium as the same strain expressing *HAC1* from a low copy number plasmid (Fig. 14B, rows 1

and 2) The *hac1*Δ yeast strain required *HAC1* expression to grow on the SD medium containing the ER stress inducer tunicamycin (Fig. 14B, rows 1 and 2). These results are consistent with the previous reports that Hac1 function is essential for the ER stress response [82] [71]. Mutation of the Ire1-cleavage site (i.e. G661C) impaired yeast cell growth on tunicamycin medium (Fig. 14B, row 3). Thus, we assumed that the G661C mutation might reduce the *HAC1* mRNA stability, mRNA splicing, and/or translation.

To determine whether the G661C mutation caused a reduction in *HAC1* mRNA stability and/or mRNA splicing, total RNA was extracted from those cells and then subjected to reverse transcriptase (RT) PCR to amplify the *HAC1* mRNA as well as an endogenous *ACT1* mRNA (housekeeping control). As shown in Fig. 14D, amounts of *ACT1* mRNA were equal in each RNA sample (lanes 1–3). As expected, no *HAC1* mRNA was detected in the *hac1*Δ strain transformed with an empty vector (Fig. 14D, lane 1), whereas equal amounts of *HAC1* mRNA were detected in the strains expressing WT *HAC1* (lane 2) or the *HAC1-G661C* mutant (lane 3). These results suggest that the tunicamycin-sensitive (Tm^S) phenotype of the strain expressing a *HAC1-G661C* mutant was unlikely due to educed *HAC1* mRNA levels. To determine whether the G661C mutation impaired mRNA splicing, RT-PCR was performed using primers that distinguish the un spliced ($HAC1^u$) and spliced ($HAC1^s$) transcripts. Again, as expected, no *HAC1* mRNA was detected in the RNA sample prepared from the *hac1*Δ strain carrying an empty vector (Fig. 14E, upper panel, lane 1). Both the $HAC1^u$ and $HAC1^s$ transcripts were amplified from cells expressing WT *HAC1* (Fig. 11E, upper panel, lane 2) as expected. In cells expressing the *HAC1-G661C* allele, the vast majority of mRNA was un spliced ($HAC1^u$; Fig.14E, lane 3). This result is consistent with the idea that the G661C mutation in the cleavage site impaired *HAC1* mRNA splicing by Ire1 under stress conditions. The defective splicing of the *HAC1-G661C* mRNA is supported by

previous studies that mutation of this guanine nucleotide at the cleavage site impairs Ire1 cleavage of the *HAC1* mRNA both in vivo and in vitro [75] [74].

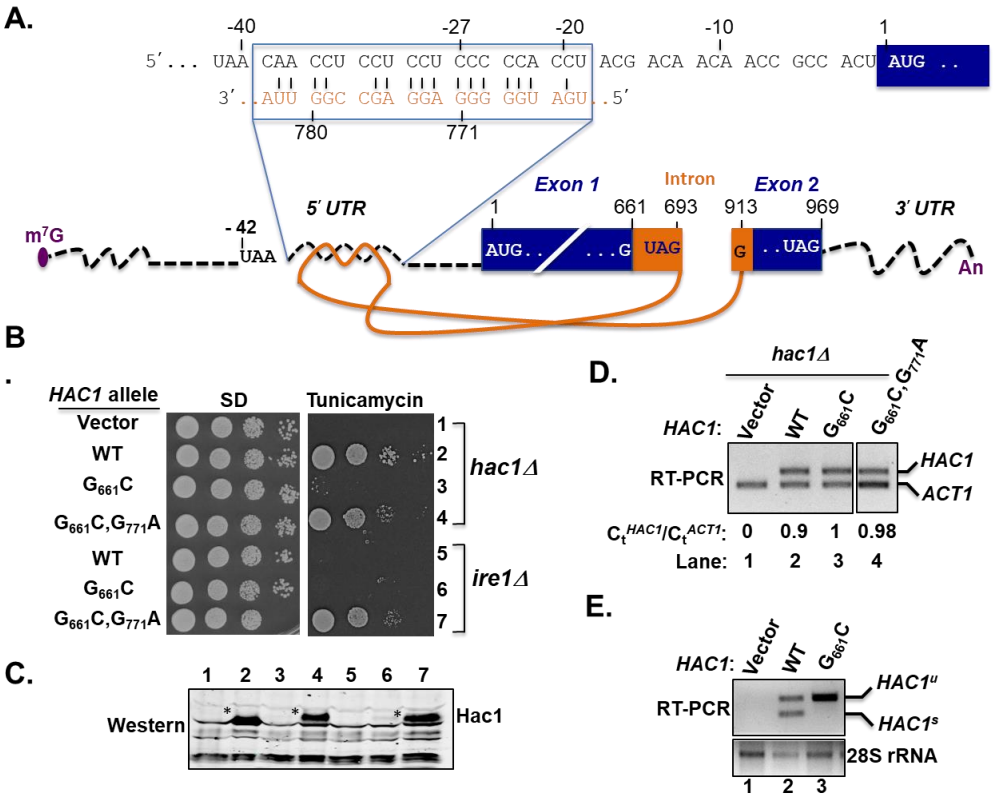

Figure 14. Single base-pair mutation in the intron derepresses translation of unspliced *HAC1* mRNA: (*A*) Schematic diagram of *HAC1* mRNA. The m^7G (7-methylguanosine) cap, 5'- and 3'-UTRs (black dotted lines), exons (dark blue boxes), intron (solid orange line) and polyA tail (An) are shown. The intron spans nucleotides G_{661} to G_{913} and interacts with the 5' UTR. The nucleotide sequence of the 5' UTR–intron interaction is shown at the top. The adenine of the *HAC1* AUG start codon is assigned as position +1 with positive and negative values for the downstream and upstream nucleotides, respectively. Two inframe *HAC1* UAG stop codons are shown: one located in the intron (nucleotide positions +691 to +693) and the second located in exon 2 (nucleotide positions +967 to +969). (*B*) Yeast growth assay. Transformants of *hac1Δ* or *ire1Δ* yeast strains carrying an empty vector or expressing the indicated *HAC1* allele were grown in SD medium to saturation, and 5 µl of serial dilutions (of OD_{600} = 1.0, 0.1, 0.01, and 0.001) were spotted on SD medium or SD medium containing 0.4 µg/ml tunicamycin and incubated 3 d at 30 ˚C.(*C*) Immunoblot analysis of Hac1 protein. Strains described in panel *B* were grown in SD medium, treated with 5 mM DTT to induce ER stress, and then WCEs were prepared and subjected to SDS-PAGE followed by Western blot analysis using a polyclonal antibody raised against recombinant Hac1 protein. Lanes are numbered according to row numbers in panel *B*. (*D*) Analysis of *HAC1* mRNA levels. Total RNA was extracted from the *hac1Δ* strains expressing the indicated *HAC1* alleles and grown under non-stress condition. The RNA was used as a template for reverse transcription (RT) and quantitative (Q)-PCR analyses of *HAC1* and *ACT1* mRNAs as described in Experimental Procedures. The ratio of threshold values (C_t) for *HAC1* and *ACT1* mRNA are shown. (*E*) RT-PCR analysis of *HAC1* mRNA splicing. A *hac1Δ* strain carrying an empty vector or expressing WT *HAC1* or the *HAC1-G661C* mutant was grown in SD medium, treated with 5 mM DTT, and then total RNA was extracted and used as a template for RT-PCR analysis of 28S rRNA or of unspliced (*HAC1*^{u}) and spliced (*HAC1*^{s}) *HAC1* transcripts as previously described (25).

To confirm that the *HAC1-G661C* allele is impaired in Hac1 protein expression, we performed the immunoblot analysis using an antibody raised against the Hac1 protein. The WCEs were prepared from cells exposed to DTT to induce ER stress (as described in materials and methods). The Hac1 protein was detected in the extract obtained from cells expressing the WT *HAC1* allele (Fig. 14C, lane 2). In contrast, no Hac1 protein was detected in cells expressing the *HAC1-G661C* mutant (lane 3). These results are consistent with the model [73] [83] that translation of the unspliced *HAC1* mRNA is blocked.

The splicing- and translation-defective *HAC1-G661C* allele was expressed on plasmid and was subjected to random mutagenesis by amplification in *E. coli* XL-1 red cells (Stratagene) that are deficient in three DNA repair pathways. The mutagenized pool of plasmids was purified and used to transform a *hac1*Δ yeast strain. Transformants were screened for the ability to grow on the medium containing tunicamycin. The *HAC1* plasmid was isolated from the tunicamycin-resistant (Tm^R) colonies and re-tested in a *hac1*Δ strain. The Tm^R phenotype associated with the plasmid was confirmed, and then plasmid was sequenced to detect the mutation. A single second-site suppressor mutation, G771A, was found to restore the ability of the *HAC1-G661C* variant to activate the ER stress response (Fig. 14B, row 4).

The nucleotide G771 is in the region of the intron that is predicted to form base-pairing interactions with the 5'-UTR of the *HAC1* mRNA. In the alignment shown in Fig. 14A, the G771 base in the intron is base-paired with C-27 in the 5'-UTR. We hypothesized that the G771A mutation disrupted the base-pairing interaction with C-27 and thus, the 5'-UTR*intron interaction is disrupted. This leads to the derepression of *HAC1* mRNA translation in the absence of splicing (*HAC1-G661C*).

We tested the possibility that the G771A mutation enabled translation of unspliced *HAC1* mRNA in the *ire1*Δ strain. Plasmids encoding the WT *HAC1*, the splice-site mutant *HAC1-G661C*, and the suppressor mutant *HAC1-G661C, G771A* were introduced into a yeast strain lacking the Ire1 endonuclease. The transformants were then tested for their ability to grow on the medium containing tunicamycin. As shown in the Fig. 14B (rows 5 and 6), the *ire1*Δ strain expressing WT *HAC1* or *HAC1-G661C* failed to grow on the medium containing tunicamycin, consistent with the notion that Ire1-mediated splicing of *HAC1* mRNA is essential for the ER stress response [73] [83] [78]. In contrast, the *HAC1- G661C, G771A* mutant conferred a Tm^R phenotype when expressed in the ire1strain (Fig. 14B, row 7). These results suggest that that the G771A mutation derepressed *HAC1* mRNA translation in the absence of splicing. Consistent with this interpretation, immunoblot analysis revealed substantial $Hac1^u$ protein levels without an increase in the level of the *HAC1* mRNA (Fig. 14D, lane 4) when the *HAC1-G661C, G771A* mutant was expressed in either *hac1*Δ (Fig. 14C, lane 4) or *ire1*Δ (Fig. 14C, lane 7) cells. These data demonstrate that a single base mutation in the intron is sufficient to eliminate translational control of the *HAC1* mRNA.

2.3 Disruption in the 5'-UTR*intron base-pairing by single base pair mutations release the translational block

Previous results suggest that single base-pair mutation in the intron can release the translational block in *HAC1* mRNA [73]. To identify significant single or double base-pairing interactions important for translational repression, we introduced several mutations in the intron or the 5'-UTR of a human influenza hemagglutinin (HA) epitope-tagged version of *HAC1*. The HA-tag was inserted between codons for Ser-10 and Asn-11 of the *HAC1* ORF. Importantly, the HA-tagged *HAC1* allele functioned like untagged WT *HAC1* and complemented the Tm^S

phenotype of a the *hac1*Δ strain (Fig. 15A, row 2) in an Ire1-dependent manner (Fig.15B, row 2). As shown in Fig. 15B, the single C-23G (row 3) and the double C-23G, C-24G (row 4) mutations in the 5'-UTR of the *HA-HAC1* mRNA showed Tm^R phenotype of an *ire1*Δ *hac1*Δ, due to disruption of the base-pairing interaction with the intron (Fig. 14A). As no *HAC1* mRNA splicing takes place in the *ire1*Δ *hac1*Δ strain, these results indicate that, just like the G771A mutation in the intron (Fig. 15B, row 7), the point mutations in the *HAC1* 5'-UTR can release translational block in *HAC1* mRNA. As shown in Fig. 15B, the G767C (row 5) and G767C, G768C (row 6) mutations in other intron residues involved in the base-pairing interactions also led to prominent Tm^R phenotypes in the *ire1*Δ *hac1*Δ strain. This suggests that mutations in either the 5'-UTR or the intron relieved translational block in *HAC1* mRNA.

Immunoblot and Northern analyses were used to assess the impact of the 5'-UTR and intron mutations on Hac1 protein and mRNA levels. As the cells lack Ire1 protein and no splicing of the *HAC1* mRNA will take place, extracts were prepared from cells grown in the absence of ER stress. Whereas no HA-Hac1 was detected in the extracts from cells expressing WT *HA-HAC1* (Fig. 15C, upper panel, lane 2), very low or low levels of HA-Hac1 protein were observed in the extracts prepared from cells expressing the C-23G or C-23G, C-24G mutant alleles, respectively, of *HA-HAC1* (Fig. 15C, upper panel, lanes 3 and 4). Thus, low level expression of HA-Hac1 protein from the *HA-HAC1-C-23G* allele was apparently sufficient to confer a Tm^R phenotype. Consistent with the more pronounced Tm^R phenotype, the HA-Hac1 protein was readily detected in extracts prepared from the *ire1*Δ *hac1*Δ strain expressing the G767C (Fig. 15C, lane 5) and G767C, G768C (lane 6) mutant form of *HA-HAC1*. Northern analyses revealed that the levels of *HA-HAC1* mRNA were similar in cells expressing the WT or various mutant alleles (Fig. 15C, lower panel, lanes 3– 6 versus 2), suggesting that differences in protein expression were not due

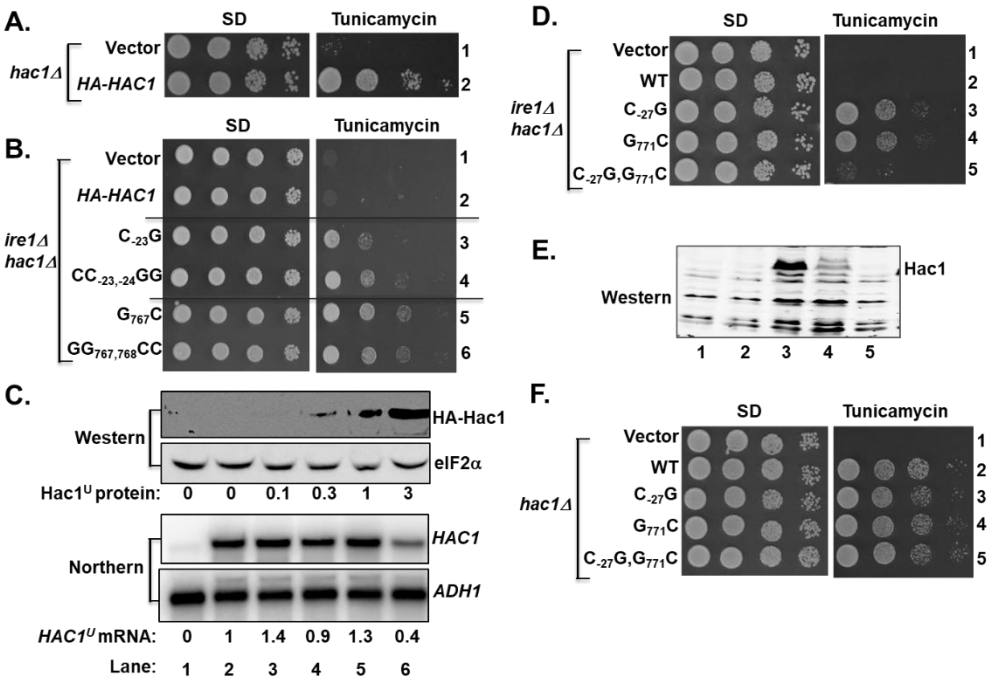

Figure 15. Disruption in the 5'-UTR*intron base-pairing by single base pair mutations release the translational block: (*A& B*) Yeast growth assays. Transformants of *ire1Δ hac1Δ* yeast strain carrying an empty vector or expressing the indicated HA-tagged *HAC1* allele were grown in SD medium to saturation, and 5 µl of serial dilutions (of OD_{600} = 1.0, 0.1, 0.01, and 0.001) were spotted on SD medium or SD medium containing 0.4 µg/ml tunicamycin and incubated 3 d at 30 ˚C. (*C*) Analysis of *HAC1* mRNA and protein expression in transformants of the *ire1Δ hac1Δ* strain. Strains described in panel *A* were grown in SD medium (unstressed conditions), and then (upper two panels) WCEs were prepared and subjected to SDS-PAGE followed by Western analysis using polyclonal antibodies against the HA epitope and translation initiation factor eIF2α, and (lower two panels) total RNA was extracted and subjected to Northern analysis using probes specific for the *HAC1* and *ADH1* mRNAs. Lanes are numbered according to row numbers in panel *A*. The levels of $Hac1^u$ protein were determined by quantitative densitometry and normalized to the levels of eIF2α protein. $HAC1^u$ mRNA levels were normalized to *ADH1* mRNA levels, with the WT ratio (lane 2) set to 1.0 (*D &F*) *Y*east growth assays. Transformants of the *ire1Δ hac1Δ* strain carrying an empty vector or expressing the indicated *HAC1* allele were grown as described in panels *A* and *B*. (*E*) Analysis of Hac1 protein expression. WCEs were prepared and subjected to SDS-PAGE followed by Western analysis using polyclonal anti-Hac1 antibodies. Lanes are numbered according to row numbers in panel *D*.

to altered mRNA levels. These results provide further support for the hypothesis that mutations designed to disrupt the 5'-UTR*intron interaction relieve the translational block in *HAC1* mRNA.

The G771 residue is predicted to form a base-pairing interaction with C-27 (Fig. 14A). To test the importance of this predicted base pair, three additional *HAC1* mutants were generated. The single mutants *HAC1-C-27G* and *HAC1-G771C* were made to destroy the pairing and the double mutant *HAC1-C-27G, G771C* was constructed to restore the pairing in the opposite orientation.

As shown in Fig. 15D, both *HAC1-C-27G* (row 3) and *HAC1-G771C* mutants (row 4) conferred a Tm^R phenotype in the *ire1*Δ *hac1*Δ strain. In contrast, the *HAC1-C-27G, G771C* double mutant (Fig. 15D, row 5) functioned like WT *HAC1* (row 2) and conferred a Tm^S phenotype in the *ire1*Δ *hac1*Δ strain. The Hac1 protein was detected only in the extracts obtained from cells expressing the *HAC1-C-27G* mutant (Fig. 15E, Western, lane 3) and the *HAC1-G771C* mutant (lane 4). The mutual suppression of the *HAC1-C-27G* and *HAC1-G771C* mutants in the *HAC1- C-27G, G771C* double mutant is consistent with the proposed base-pairing interaction between these residues and further demonstrates that repression of *HAC1* expression by the 5'-UTR*intron interaction is sensitive to loss of a single base pair interaction. These three *HAC1* alleles (*HAC1-C-27G*, *HAC1-G771C*, and *HAC1- C-27G,G771C*) when expressed in a *hac1*Δ strain (IRE1 intact in the chromosome) conferred a Tm^R phenotype (Fig. 15F, rows 2–5) just like the WT. Based on the results of these mutational analyses, we conclude that 5'-UTR*intron interaction and, in particular, the C-27/G771 base pair plays an important role in repressing *HAC1* mRNA translation.

2.4 Insertion of in-frame start codons in the 5'- UTR of *HAC1* mRNA

The secondary structure in the 5'-UTR could impair Hac1 synthesis by either blocking the scanning of the initiation complex to the start codon or by impairing translation elongation as has been proposed previously [73]. To differentiate between these models, AUG start codons were inserted at three sites in the 5'- UTR of the *HAC1* gene (Fig. 16A). All three of the introduced AUG codons were in-frame with the *HAC1* ORF and thus would be predicted to encode Hac1 proteins with N-terminal extensions. The inserted AUG codon was made sole initiation codon by mutating the adenine of the authentic *HAC1* start codon to guanine, generating the *HAC1-A1G* allele. Similarly, the adenine of the AUG codon encoding Met3 was likewise altered to guanine to generate the *HAC1-A7G* mutation (Fig. 16A). These mutations ensure that Hac1 protein is

translated from the inserted in-frame AUGs. Consistent with this it was shown in Fig. 16B that *HAC1-A1G,A7G* mutant allele failed to complement the Tm^S phenotype of a *hac1*Δ strain (row 3). The A1G,A7G mutation did not reduce *HAC1* mRNA levels (Fig. 16D, lane 3) consistent with the notion that, even following splicing, scanning ribosomes fail to initiate at the *HAC1* start site (a GUG codon in the A1G mutant), and no Hac1 protein was produced (Fig. 16C, lane 2).

We first characterized the effect of insertion of the AUG immediately upstream of the 5'-UTR*intron interaction (see Figs. 14Aand 16A) at the position -42UAA. This insertion with the mutations of authentic AUG (as described above) generated *HAC1-AUG-42, A1G, A7G* allele. As shown in Fig. 16B, the *HAC1-AUG-42, A1G, A7G* allele conferred a Tm^R phenotype when expressed in a *hac1*Δ strain (row 4), suggesting that a functional Hac1 protein with an extra 14 amino acids at the N terminus was expressed from the AUG-42 codon. Western analyses were performed on WCEs prepared from ER-stressed cells using an antibody prepared against recombinant Hac1 protein. Consistent with the introduction of 14 extra N-terminal residues, the Hac1 protein in the *HAC1-AUG-42, A1G, A7G* cells was of higher molecular weight than the WT Hac1 protein (Fig. 16C, lanes 1 and 3). The above results were obtained using a *hac1*Δ strain in which the chromosomal IRE1 gene was intact. Thus, it is likely that the *HAC1-AUG-42, A1G, A7G* mRNA was spliced in these cells, and we conclude the AUG insertion does not impair Hac1 production or the ability of Hac1 to promote yeast cell growth on medium containing tunicamycin.

We further expressed the *HAC1- AUG-42, A1G, A7G* allele in an *ire1*Δ strain where the mRNA will not be spliced. As shown in Fig. 16B, the *HAC1-AUG-42, A1G, A7G* allele failed to complement the Tm^R phenotype of the *ire1*Δ strain (Fig. 16B, row 5). Western analyses showed that the Hac1 protein was not produced from the inserted AUG without Ire1 mediated splicing (Fig. 16D, lower panel, lane 2). These results indicate that the unspliced *HAC1-AUG-42, A1G,*

A7G mRNA is translationally repressed, and we propose that the presence of the adjacent secondary

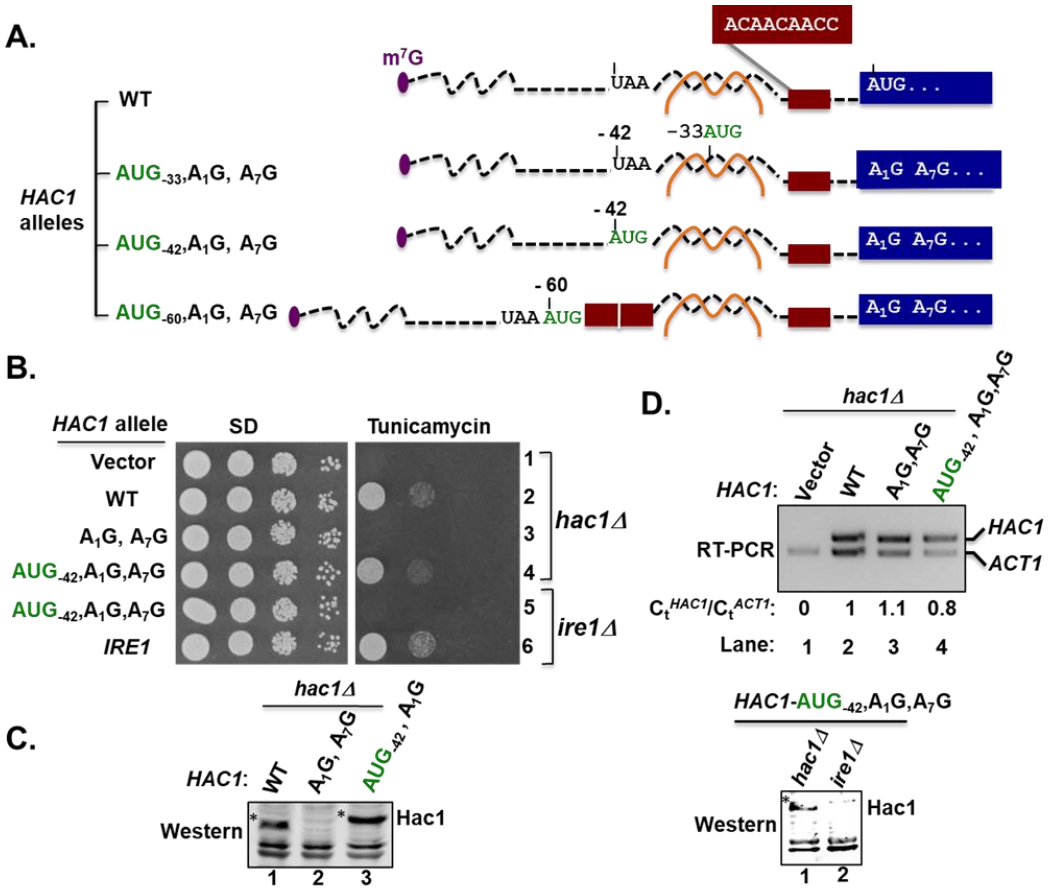

Figure 16. Insertion of an in-frame start codons in the 5'- UTR of *HAC1* mRNA: (*A*) Schematic diagrams showing the 5'-UTRs of *HAC1*, *HAC1-AUG-33,A1G*, *HAC1-AUG-42,A1G* and *HAC1-AUG-60,A1G* alleles. Color schemes are the same as in Fig. 11A. The brown box represents the 9-nucleotide sequence 5'-A-15CAACAACC-7-3'. The A_1G mutation converts the *HAC1* start codon to GUG and includes a second mutation A_7G that alters the Met3 codon. The nucleotides $C_{-33}C_{-32}U_{-31}$ or $U_{-42}A_{-41}A_{-40}$ were replaced by an AUG codon in the *HAC1-AUG-30,A1G* or *HAC1-AUG-42* allele, respectively. In the *HAC1-AUG-60* allele, a 21-nucleotide sequence consisting of an AUG codon plus two repeats of the 9 nucleotides represented by the brown box was inserted upstream of position -39.(*B*) Yeast growth assay. Transformants of *hac1Δ* or *ire1Δ* yeast strains carrying an empty vector or expressing the indicated *HAC1* or *IRE1* allele were grown in SD medium to saturation, and 5 µl of serial dilutions (of OD_{600} = 1.0, 0.1, 0.01, and 0.001) were spotted on SD medium or SD medium containing 0.4 □g/ml tunicamycin and incubated 3 d at 30 ˚C. (*C*) Immunoblot analysis of Hac1 protein. Strains described in panel *B* were grown in SD medium, treated with 5 mM DTT to induce ER stress, and then WCEs were prepared and subjected to SDS-PAGE followed by Western analysis using a polyclonal antibody raised against recombinant Hac1 protein. (*D*) RT-PCR and Western analysis of *HAC1* expression. (Upper panel) Total RNA was extracted from the *hac1Δ* strains expressing the indicated *HAC1* alleles following growth under non-stress conditions. The RNA was used as a template for RT-PCR and Q-PCR analysis of *HAC1* and *ACT1* mRNAs as described in Experimental Procedures. The ratio of threshold values (C_t) from Q-PCR analysis of *HAC1* and *ACT1* mRNAs are shown. (Lower panel) *hac1Δ* and *ire1Δ* strains expressing *HAC1-AUG-42,A1G* were grown in SD medium, treated with 5 mM DTT to induce ER stress, and then WCEs were prepared and subjected to SDS-PAGE followed by Western analysis using polyclonal anti-Hac1 antibodies.

structure in the 5'- UTR prevents the scanning ribosomes from accessing the AUG-42 start codon or may impede formation of a productive 80S ribosome on the AUG-42 codon.

Two additional mutants were constructed by inserting AUG at different positions. First, the nucleotides $C_{-33}C_{-32}U_{-31}$ in the *HAC1-A_1G, A_7G* allele were substituted by an AUG triplet to generate the *HAC1-AUG_{-33}, A_1G, A_7G* allele (Figs. 14A and 16A). This mutation is predicted to disrupt two base-pair interactions between the 5'-UTR and the intron (Fig. 14A). We previously showed that disruption of a single base-pair interaction was sufficient to relive the *HAC1* mRNA translational block (Figs. 14-15). Hence, likewise AUG_{-33} mutation was expected to relieve the translational block. As shown in Figure 17A (row 3), and in contrast to WT *HAC1* (row 1), the *HAC1-AUG_{-33}, A_1G, A_7G* allele conferred a Tm^R phenotype in the *ire1Δ hac1Δ* strain, and Hac1 protein was detected in extracts from cells expressing the mutant, but not the WT protein (Fig. 17B, lane 3 versus 1). These experiments with the *HAC1-AUG_{-33}, A_1G, A_7G* mutant provide further support for the hypothesis that disruption of the secondary structure formed by the base-pairing interaction between the 5'-UTR and intron is sufficient to derepress translation of the unspliced *HAC1* mRNA.

Because the AUG_{-42} insertion may have failed to derepress *HAC1* mRNA translation due to the close proximity of AUG_{-42} to the 5'-UTR*intron secondary structure, the third mutant was constructed by inserting the AUG codon further upstream from the secondary structure. The *HAC1-AUG_{-60},A_1G,A_7G* allele was generated by inserting an AUG codon followed by 18 additional nucleotides at position -39 (Figs. 14A and 16A) to increase the distance between the AUG and the secondary structure. The inserted nucleotides consist of two 9-nucleotide repeats of the sequence normally found between positions -15 and -7 (5'-A_{-15}CAACAACC$_{-7}$-3') of the *HAC1* mRNA which does not form a secondary structure. In this new *HAC1-AUG_{-60},A_1G,A_7G* allele, the

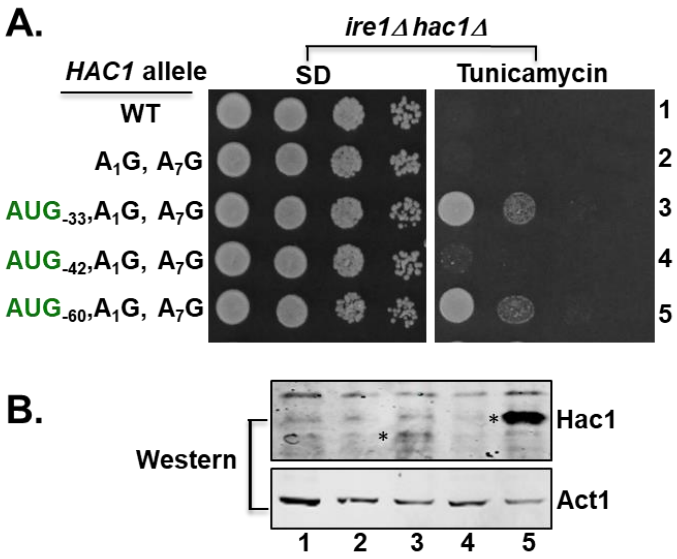

Figure 17. Insertion of an AUG codon at position -60 derepresses translation of unspliced *HAC1* mRNA: (*A*) Yeast growth assays. Transformants of an *ire1Δ hac1Δ* strain carrying an empty vector or expressing the indicated *HAC1* allele were grown in SD medium to saturation, and 5 µl of serial dilutions (of OD_{600} = 1.0, 0.1, 0.01, and 0.001) were spotted on SD medium or SD medium containing 0.4 µg/ml tunicamycin and incubated 3 d at 30 ˚C. (*B*) Immunoblot analysis of Hac1 protein. Strains from panel *A* were grown under non-stressed conditions in SD medium, and then WCEs were prepared and subjected to SDS-PAGE followed by Western analysis using polyclonal antibodies against recombinant Hac1 or Act1 protein. Lanes are numbered according to row numbers in panel *A*, and an asterisk is positioned on the left side of the Hac1 protein in lanes 3 and 5.

inserted AUG_{-60} start codon is located ~26-nts from the 5' cap and ~22-nts upstream of the secondary structure.

Interestingly, when introduced into the *ire1Δ hac1Δ* strain, the *HAC1-AUG$_{-60}$, A_1G,A_7G* allele conferred a Tm^R phenotype (Fig. 17A, row 5). The western analysis of WCEs from the strain grown under non-ER stress conditions revealed a prominent Hac1 protein signal consistent with the growth (Fig. 17B, lane 5). Importantly, the Hac1 protein in the *HAC1-AUG$_{-60}$, A_1G,A_7G* strain migrated more slowly in SDS-PAGE than the Hac1 protein produced in the *HAC1-AUG$_{-33}$,A_1G,A_7G* strain (Fig. 17B, lane 3), consistent with the longer N-terminal extension for the protein initiating at AUG_{-60}. The *HAC1-AUG$_{-60}$, A_1G,A_7G* mRNA allele is unspliced in the *ire1Δ hac1Δ* strain and the 5'-UTR*intron interaction is also intact hence, the ability of AUG_{-60} codon insertion to relieve the translational block in *HAC1* leads to two important conclusions. First, in contrast to

the previously proposed model of a translation elongation block on the unspliced *HAC1* mRNA, elongating ribosomes are able to translate through the 5'-UTR*intron interaction. Second, the data are consistent with a model in which the 5'-UTR*intron interaction impedes translation initiation by blocking ribosomes scanning to the native AUG start codon.

2.4 Unspliced HAC1 mRNA is associated with monosomes

The genetic suppressor and AUG insertion mutant studies support the model that ribosomal scanning on the WT unspliced *HAC1* mRNA is paused at the 5'-UTR*intron interaction site. Accordingly, the point mutations that weaken the 5'-UTR*intron interaction allow scanning ribosomes to traverse the 5'-UTR and initiate translation at the AUG start codon of *HAC1*. If this model is correct, few ribosomes should associate with the unspliced WT *HAC1* mRNA, whereas the *HAC1* mRNA containing the point mutations that disrupt the 5' UTR–intron interaction should be readily translated and thus show a greater association with polysomes.

To test this possibility, the WT chromosomal *HAC1* allele (translationally repressed) in the *ire1Δ* strain J751 was replaced with the *HAC1-GG*$_{767,768}$*CC* mutant (translationally active; see Fig. 12B-C, lane 6) to create strain J1167. Next, WCEs from strains J751 and J1167 were subjected to velocity sedimentation in 7 to 47% sucrose gradients, and then the gradients were fractionated while being scanned at 254 nm (see Experimental Procedures) to visualize the free 40S and 60S ribosomal subunits, 80S monosomes, and polyribosomes (Fig. 18).

To monitor the sedimentation of mRNAs in the gradients, RNA was isolated from the gradient fractions and analyzed by Northern blot using probes to detect the *HAC1* mRNA and the *ADH1* mRNA. As shown in Figure 15A, the WT *HAC1* mRNA in unstressed *ire1Δ* cells showed strong association with 40S subunits (lane 2) and only ~53% of the *HAC1* mRNA associated with polysomes. In contrast, the *ADH1* mRNA was primarily (~80%) associated with the polysome

fractions. Disruption of the 5' UTR–intron base-pairing interaction in the *HAC1-GG*$_{767,768}$*CC* mutant resulted in more (~70%) of the *HAC1* mRNA associating with polysomes (Fig. 15B, lanes 5-8).

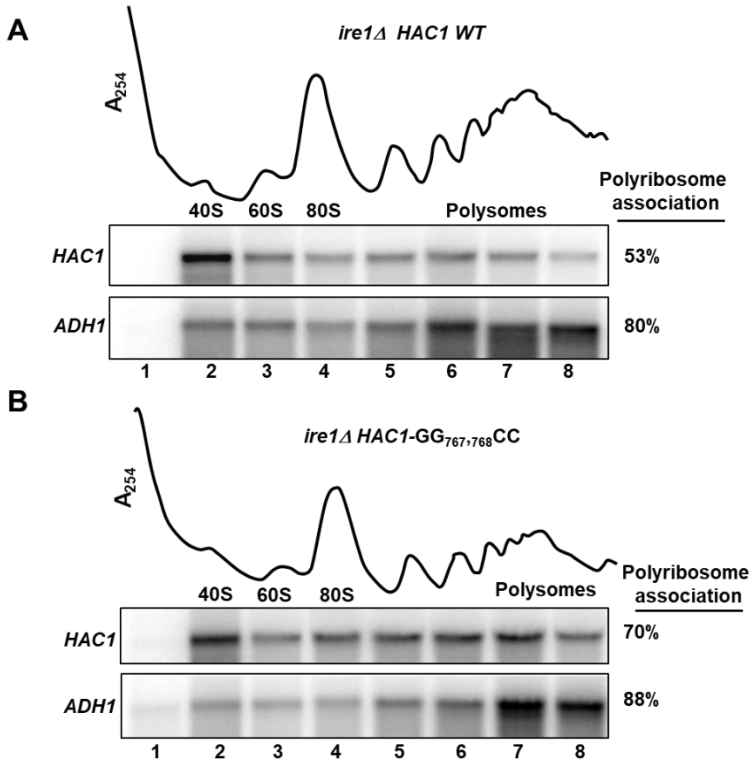

Figure 18. Unspliced HAC1 mRNA is associated with monosomes: WCEs from strains J751 (*HAC1 ire1Δ, panel A*) and J1167 (*HAC1-GG*$_{767,768}$*CC ire1Δ, panel B*) were resolved by velocity sedimentation in 7 to 47% sucrose gradients. Gradients were fractionated while scanning at A_{254} (upper tracings), and the positions of the 40S and 60S subunits, 80S ribosomes and polysomes are indicated. Total RNA was extracted from each fraction and subjected to Northern analysis using probes for *HAC1* and *ADH1*, as indicated. The amounts of *HAC1* mRNA in the gradient fractions were quantified, and the percentage of *HAC1* mRNA associating with polysomes (fractions 4-8) was calculated.

These data provide independent support for the idea that the 5'-UTR–*intron interaction blocks ribosomal scanning, resulting in a substantial fraction of the *HAC1* mRNA associating with a single 40S subunit. Accordingly, disruption of the secondary structure releases the initiation block and enables more ribosomes to translate the *HAC1* mRNA, resulting in more of the mRNA associating with polyribosomes.

2.6 Overexpression of helicase eIF4A relieves the translational block in the HAC1-AUG$_{-42}$,A$_1$G,A$_7$G allele

The insertion of an AUG codon at position -60 far upstream of the secondary structure in the *HAC1* mRNA led to constitutive derepression of Hac1 synthesis (Fig. 17) and the insertion of an AUG codon at position -42 immediately before the secondary structure failed to derepress Hac1 synthesis in the absence of splicing (Fig. 16). As our data indicate that the secondary structure formed by the base-pairing interactions between the intron and 5'-UTR in the *HAC1* mRNA interferes with translation initiation, we reasoned that scanning ribosomes lack sufficient ability to unwind the secondary structure and access the AUG$_{-42}$ start codon.

Translation initiation factor eIF4A is an RNA helicase which melts the secondary structures present in the 5'-UTR specifically cap-proximal secondary structures [84] [34]. Hence, we reasoned that overexpression of the translation initiation factors forming a cap-binding complex (eIF4F) along with eIF4A will melt the secondary structure and facilitate the Hac1 production on the *HAC1-AUG$_{-42}$,A$_1$G,A$_7$G* mRNA. Consistent with this hypothesis, overexpression of helicase eIF4A conferred a tunicamycin-resistant phenotype in an *ire1Δ hac1Δ* strain expressing the *HAC1-AUG$_{-42}$,A$_1$G,A$_7$G* allele (Fig. 19B, right panels, row 2). This derepression of *HAC1* synthesis was dependent on eIF4A helicase activity as overexpression of the eIF4A-D170E mutant that lacks helicase activity failed to support growth on tunicamycin medium (Fig. 19B, right panels, row 5). Moreover, overexpression of eIF4E, the mRNA cap-binding protein, or eIF4B, a factor that promotes eIF4A activity in mRNA remodeling and scanning, failed to promote growth of the strain expressing the *HAC1-AUG$_{-42}$, A$_1$GA$_7$G* allele on tunicamycin medium (Fig. 19B, right panels, rows

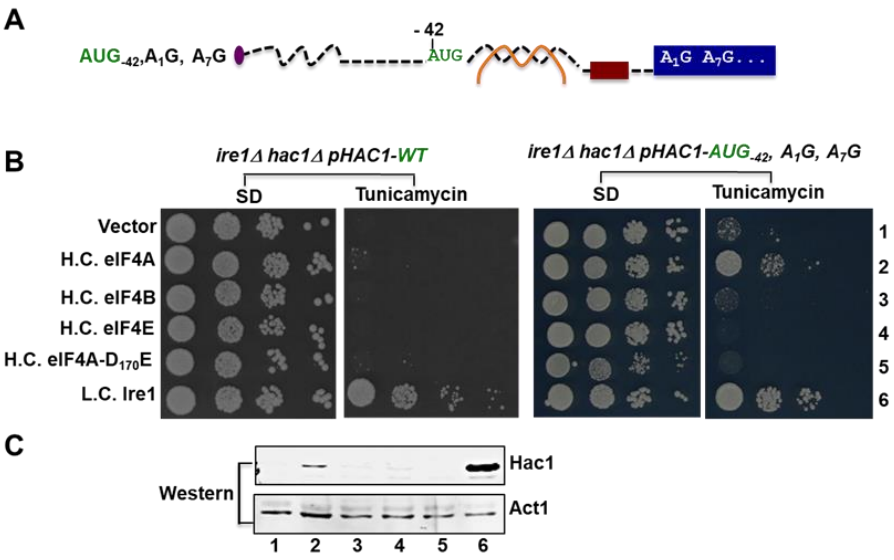

Figure 19. Overexpression of helicase eIF4A relieves the translational block in the *HAC1-AUG-42,A1G,A7G* allele: (*A*) Schematic diagram showing the 5' UTR of the *HAC1-AUG-42,A1G,A7G* allele. Color schemes are the same as in Fig. 11A.(*B*) Yeast growth assays. An *ire1Δ hac1Δ* strain was transformed with a low copy-number *URA3* plasmid expressing *HAC1* (left panels) or the *HAC1-AUG-42,A1G,A7G* allele (right panels). The resulting strains were then transformed with high copy-number (H.C.) *LEU2* plasmids expressing eIF4A, eIF4B, eF4E or eIF4A-D_{170}E or with a low copy-number (L.C.) *LEU2* plasmid encoding *IRE1*. (*C*) Immunoblot analysis of Hac1 protein. Strains from the right panels in *B* were grown under stressed conditions in SD medium, and then WCEs were prepared and subjected to SDS-PAGE followed by Western analysis using a polyclonal antibody raised against recombinant Hac1 protein. Lanes are numbered according to row numbers in panel *B*.

3 and 4). Consistent with these growth phenotypes, Western blot analyses demonstrated that overexpression of eIF4A, but not eIF4E, eIF4B or non-functional eIF4A-D170E, promoted synthesis of Hac1 in from the *HAC1-AUG-42, A1G,A7G* mRNA (Fig. 19C).

These data support the model that increasing the cellular amount of eIF4A enables the scanning ribosome to melt secondary structure on the *HAC1-AUG-42,A1G,A7G* mRNA and gain access to the AUG-42 start codon. In contrast to the ability of eIF4A overexpression to derepress translation of the *HAC1-AUG-42,A1G,A7G* mRNA, overexpression of eIF4A failed to confer a tunicamycin-resistant phenotype in cells expressing the wild type HAC1 mRNA (Fig. 19B, left panel, row 2). Thus, we propose that overexpression of eIF4A can promote the modest level of secondary structure melting required for the scanning ribosome to access AUG-42 but is not

sufficient for the ribosome to traverse through the entire base-paired region and access the wild-type *HAC1* start codon.

2.6 Conclusion and future perspectives

Taken together, these results suggest that intact base-pairing of the 5'-UTR and intron is essential for translational suppression in *HAC1* mRNA which is consistent with previously published results [73]. The previously proposed model predicts that, [73] (see Fig. 9B) translating ribosomes are trapped on the mRNA and spliced *HAC1* mRNA can produce Hac1 protein once the translational block is removed by Ire1 mediated splicing. However, we showed that spliced *HAC1* mRNA with the mutations of the authentic AUGs is unable to translate Hac1 protein, contradicting the previously proposed model.

Our results suggest that *HAC1* mRNA with an insertion of in-frame AUG codon upstream of the base-pairing interaction is able to translate and produce Hac1 protein independent of splicing suggests that base-pairing interaction is blocking the translation initiation in *HAC1* mRNA. This result is further supported by the observation that, the majority of the un spliced *HAC1* mRNA is associated with monosomes (40S ribosomes). Collectively, our results propose a new model where the base-pairing interaction of 5'-UTR and intron inhibits translation initiation and that over expression of the RNA helicase eIF4A relieves the translational block by unwinding of the base-pairing.

With these results, the future studies are targeted at dissecting the translation initiation block. As shown in Fig. 1 and as listed in table1, there are multiple initiation factors that contribute in the translation initiation. Translation initiation can be seen in three steps 1) loading of the Pre-initiation complex (small ribosomal subunit (40S), a ternary complex (TC) made of eIF2, GTP and

initiator met-tRNA$_i^{met}$ and other initiation factors, (eIF1, eIF1A and eIF3, eIF5) on the 5' end of the mRNA. 2) Scanning of PIC on the mRNA and 3) Identification of the start codon by codon-anticodon base-pairing interaction. In the next chapter, we study what step of translation initiation is blocked by the base-pairing interactions between 5'-UTR and intron of the *HAC1* mRNA.

3. *HAC1* mRNA is translationally repressed due to insufficient loading of Initiation complex or helicase

3.1 Introduction

Our previous results suggest that the translation of *HAC1* mRNA is blocked at the initiation stage [77]. Further we wanted to dissect the inner workings of the initiation block. Translation initiation can be seen in three steps; 1) loading of the Pre-initiation complex (PIC) ; 43S PIC {small ribosomal subunit (40S), a ternary complex (TC) made of eIF2, GTP and initiator met-$tRNA_i^{met}$ and other initiation factors, eIF1, eIF1A and eIF3, eIF5} on the 5' end of the mRNA; 2) Scanning of 43S PIC on the mRNA and 3) Identification of the start codon by codon-anticodon base-pairing interaction (see Fig. 1).

Translation initiation requires more than 10 initiation factors (see Table 1). Defects in the recruitment of any of these initiation factors may result in the initiation block. We performed over expression studies with an intent to identify the defect in the translation initiation. Next, we differentiated between the 43S PIC loading block and ribosomal scanning block by modulating the length of the 5'-UTR of the *HAC1* mRNA.

We propose a model to explain the translation initiation block in *HAC1* mRNA. We propose that the 5'-UTR and intron interaction is inhibits loading of either 43S PIC or helicase, (eIF4A) on the *HAC1* mRNA. In this chapter we present results supporting our proposed models.

3.2 Overexpression of the cap-binding complex proteins or RNA helicases does not relieve the translation block

In order to begin translation, mRNA is activated by binding of the cap-binding complex (eIF4F) which is composed of eIF4A, eIF4B, eIF4E, and eIF4G to the mRNA cap [5][14][67]. The eIF4G acts as scaffold protein on which other cap-biding proteins (eIF4A, eIF4B, and eIF4E) are assembled. The eIF4E binds to the mRNA cap. Association of the eIF4G with the Poly A tail binding protein (Pab1 in yeast) circularizes the mRNA which increases the translational efficiency of the mRNA (see Fig. 10). Also, most mRNAs need eIF4G for the recruitment of the 43S PIC. This suggests that impaired recruitment of any of the cap-binding complex components might result in the block of translation on *HAC1* mRNA. We performed overexpression studies to identify if there is any defect in the recruitment of cap-biding complex proteins on the *HAC1* mRNA.

As shown in Fig. 20B, an *ire1*Δ strain was transformed with a high copy plasmid expressing eIF4A, eIF4B, eIF4E or eIF4G. Transformants were then tested for growth on medium containing tunicamycin. As expected, the *ire1*Δ strain containing empty vector showed Tm^S phenotype (Fig. 20B, row 1) and showed Tm^R phenotype (row 2) when the same vector expressed WT *IRE1* gene. Overexpression of any of the cap-binding complex proteins (eIF4A, eIF4B, eIF4E, and eIF4G) failed to show the Tm^R phenotype in the *ire1*Δ strain (Fig. 20B). If recruitment of either eIF4A, eIF4B, eIF4E or eIF4G is defective, then expression of these proteins on a high copy vector in *ire1*Δ strain is expected to overcome the Tm^S phenotype. These results indicated that over expression of the cap-binding proteins could not overcome the translation initiation block.

RNA helicase plays an important role in translation of mRNAs with structured 5'-UTRs [38] [85] [34]. As explained in the introduction, two major RNA helicases, eIF4A and Ded1 are involved in the unwinding of secondary structures located near the cap (cap-proximal) or away from the cap (cap-distal) in yeast, *S. cerevisiae*. We previously have shown that over expression of eIF4A activates *HAC1* expression in *HAC1 -42AUG A1G, A7G* allele (see Fig. 19) [77]. Consistent with this notion we expect that overexpression of the RNA helicases (eIF4A and Ded1) might melt the base-pairing structure and consequently, the ribosome will be able to access the AUG codon and produce Hac1 protein (Fig. 20C). However, we observed that an *ire1*Δ null strain expressing eIF4A, and Ded1 failed to show Tm^R phenotype (Fig. 20D). This result indicates that, over expression of RNA helicase is not sufficient to relive the translational block in *HAC1* mRNA.

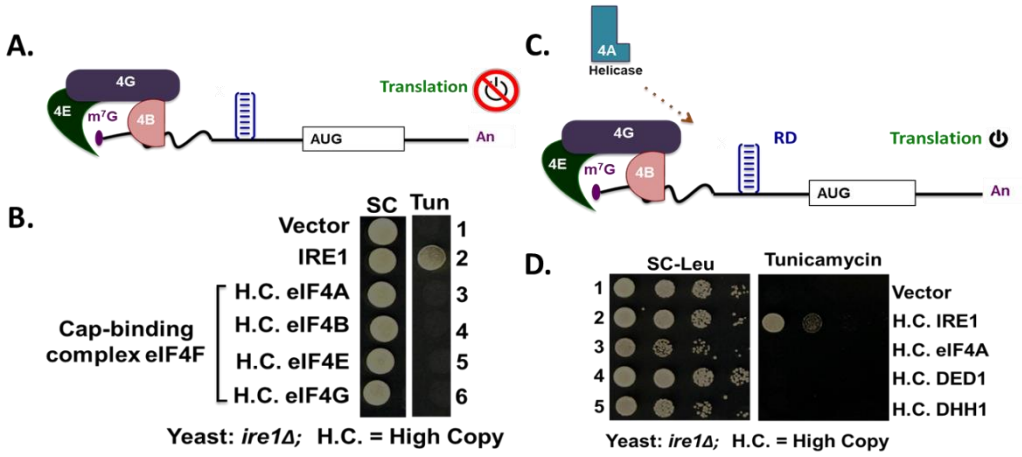

Figure 20. Over expression of the cap-binding complex proteins or RNA helicases does not relieve the translation block: (*A*) Schematic diagram of translationally silent *HAC1* mRNA without eIF4A bound to the mRNA cap. The m^7G (7-methylguanosine) cap, 5'- and 3'-UTRs (black lines), exons (white box), base-pairing interaction is shown the 5'-UTR (blue horizontal lines) and polyA tail (An) are shown. (*B*) and (D) Yeast growth assay. Transformants of *ire1Δ* yeast strains carrying an empty vector or expressing the indicated plasmids were grown in SD medium to saturation, and 5 µl of serial dilutions (of OD_{600} = 1.0, 0.1, 0.01, and 0.001) were spotted on SD medium or SD medium containing 0.4 µg/ml tunicamycin and incubated 3 d at 30 ˚C. (*C*) Schematic diagram of translationally active *HAC1* mRNA with eIF4A recruitment. The color schemes ae the sameas (A).

3.3 Overexpression of translation initiation factors and ribosomal proteins does not relieve the translation block

As indicated by previous results, overexpression of the proteins eIF4A, eIF4B, eIF4E, eIF4G, and the RNA helicase Ded1were not sufficient to relieve the translational block in *HAC1* mRNA. We also over expressed several initiation factors associated with the 43S PIC (see Fig 1 and table 1). These initiation factors are eIF1, eIF1A, eIF3a, eIF5, eIF5B, and Ternary complex; TC (eIF2, GTP and initiator met-$tRNA_i^{met}$). These initiation factors together with the 40S ribosome form the pre-initiation complex (PIC). We purposed that defective loading of any of those initiation factors might be the reason for translational suppression in *HAC1* mRNA. Consequently, over expression of those factors might relieve the translational block. An *ire1*Δ strain was transformed with a high copy plasmids expressing eIF1, eIF1A, eIF3a, eIF5, eIF5B, and Ternary complex. The transformants were then tested for their growth on medium containing tunicamycin. The over expression of the initiation factors eIF1, eIF1A, eIF3a, eIF5, eIF5B, and Ternary complex in *ire1*Δ strain failed to exhibit Tm^R phenotype (Fig. 21B).

The *S. cerevisiae* 40S ribosomal subunit is composed of the 18S rRNA and 33 ribosomal proteins [86]. The crystal structure of the 40S ribosome is shown in Fig. 5 and the list of ribosomal proteins is provided in Table 3. The ribosome filter hypothesis suggests that the small ribosomal subunit (40S) has a regulatory role in addition to translation initiation [87] [88] [89]. Thus, we explored the possibility of involvement of 40S ribosomal proteins in regulating the *HAC1* mRNA translation. We over expressed 40S ribosomal proteins Rps3, Rps4, Rps7A, Rps14, Rps15, Rps19, Rps28, Rps18 and Rps16 in the *ire1*Δ yeast strain. These proteins were chosen because they are present near the mRNA entry channel on the 40S subunit and have been implicated in the helicase activity exhibited by the 40S ribosome [90]. We tested the growth of *ire1*Δ transformants on

medium containing tunicamycin. None of the transformed yeast showed Tm^R phenotype (Fig. 21C). This result suggests that 40S ribosomal proteins might play a role in regulation of

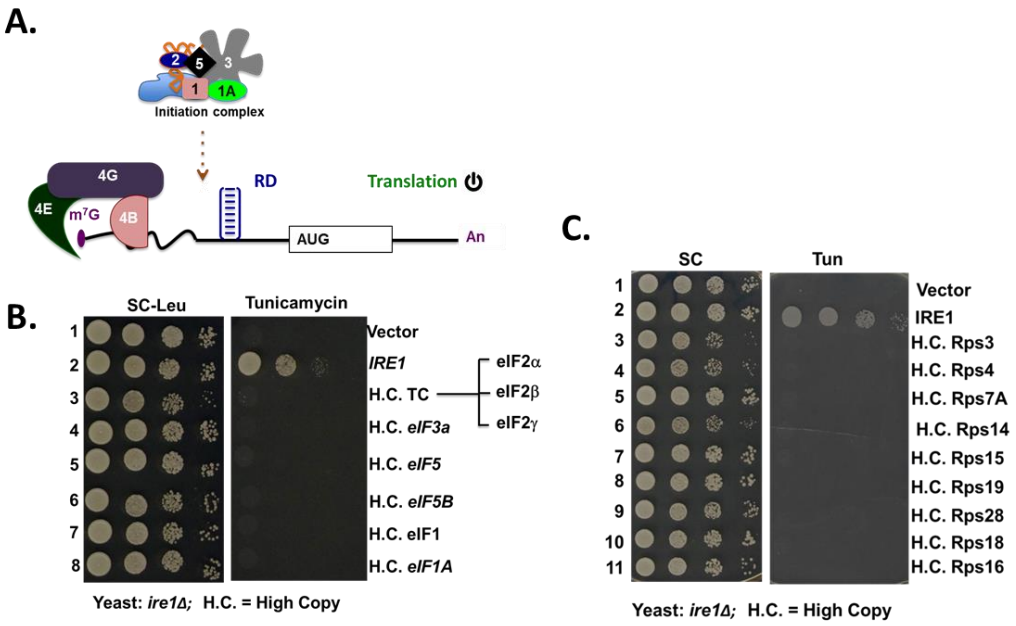

Figure 21. Over expression of translation initiation factors and ribosomal proteins does not relieve the translation block: (*A*) Schematic diagram of *HAC1* mRNA. The color schemes ae the same as Fig. 17A. The schemtic diagram shows translationally active *HAC1* mRNA upon enhanced association with 43S PIC. (*B*) and (C) Yeast growth assay. Transformants of *ire1Δ* yeast strains carrying an empty vector or expressing the indicated plasmids were grown in SD medium to saturation, and 5 µl of serial dilutions (of OD_{600} = 1.0, 0.1, 0.01, and 0.001) were spotted on SD medium or SD medium containing 0.4 µg/ml tunicamycin and incubated 3 d at 30 ˚C.

translation in *HAC1* mRNA however, the overexpression had no effect.

3.4 Translation of *HAC1* mRNA shows no defect in ribosomal scanning

We focused on understanding how the base-pairing interaction inhibits the translation initiation in *HAC1* mRNA. The base-pairing interaction can either inhibit translation initiation by impeding the loading of 43S PIC on the mRNA or by inhibiting the ribosomal scanning for the AUG start codon. The secondary structure in the 5'-UTR, and the length of the 5'-UTR together might regulate the translation of mRNA. As explained in the introduction, a long 5'-UTR enhances

the rate of translation initiation by allowing multiple Pre-initiation (43S PIC) complexes to load on the mRNA [37].

We wanted to understand how the length of the 5'-UTR affects translational silencing in *HAC1* mRNA. The length of a 5'-UTR in *HAC1* mRNA is 69 nucleotides (see Fig.23A) , as reported by the yeast transcriptome analyses [91]. 17 nucleotides of the 5'-UTR are involved in the base-pairing with the intron [73]. The secondary structure lies 30 nucleotides away from the mRNA cap (see Fig 23A). As explained in the Introduction, the base-pairing structures present within the first 45 nucleotides from the mRNA cap are considered as cap-proximal secondary structures [34]. According to this definition, the secondary structure in the *HAC1* mRNA qualifies as a cap-proximal secondary structure.

To understand how the cap-proximal base-pairing in *HAC1* mRNA regulates the translation initiation, we asked two questions: 1) what is the median length of the 5'-UTR in yeast transcriptome? 2) How many nucleotides are occupied by the 43S PIC on the mRNA? To answer the first question we analyzed yeast transcriptome and found that, the length of the 5'-UTR ranges from 5 nucleotides to 450 nucleotides with a median of 55 nucleotides. The frequency distribution is shown in Fig 22. To answer the second question we searched the literature and we found that the 43S PIC occupies ~30 nucleotides on the mRNA [37] [36]. In the context of *HAC1* mRNA the length of 5'-UTR upstream of the base-pairing interaction is 30 nucleotides (See Fig. 23A). This is less than the median 5'-UTR length and is not likely sufficient for the loading of the 43S PIC on the mRNA. Thus, we hypothesized that the 5'-UTR length of the *HAC1* mRNA is insufficient for the loading of the 43S PIC. To test our hypothesis we increased the 5'-UTR length upstream of the base-pairing interaction in the *HAC1 -42 AUG A1G A7G* allele.

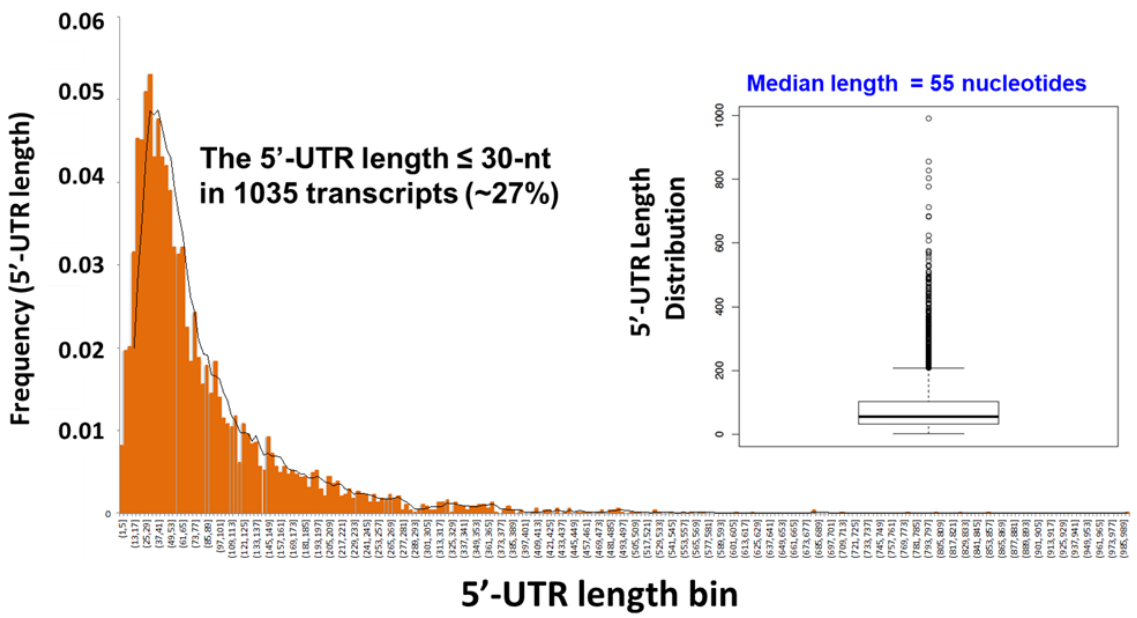

Figure 22. Median length of yeast 5'-UTR is 55 nucleotides: The frequency distribution of the length of the yeast 5'-UTRs. The X-axis shows the length of the 5'-UTRs and Y-axis shows the frequency of the occurrence. The box plot on the right show that the median length of the yeast 5'-UTR is 55 nucleotides.

The *HAC1 -42 AUG A1G, A7G* allele when expressed in the *ire1*Δ strain shows tunciamycin sensitive phenotype (see Fig. 16, row 5) [77]. We reasoned that as the inserted AUG lies very close to the secondary structure and the ribosome is unable to access and initiate translation from that AUG. However, we also showed that upon overexpression of the RNA helicase (eIF4A), the Hac1 protein is generated from the inserted AUG (see Fig. 19, right panel, row 2) [77]. The over expression of eIF4A likely helped in unwinding the secondary structure and hence the AUG codon was accessible to initiate the translation. The 5'-UTR length upstream of the inserted AUG in the *HAC1 -42 AUG A1G, A7G* allele is only 27 nucleotides (see Fig. 23 A). We hypothesized that, if we increase the 5'-UTR length upstream of the inserted AUG, it might provide sufficient extra room for the 43S PIC loading and hence the Hac1 protein would be generated from the inserted AUG. To test our hypothesis, we inserted unstructured nucleotides (5'- ACGACAACAACCACGACAACAACC -3') at position -41 in the 5'-UTR of *HAC1* mRNA.

This 24 nucleotide sequence is obtained from the repeat of the 12 nucleotides (5' -15-ACGACAACAACC -7 3') present in the 5'-UTR of *HAC1* mRNA. Insertion of the 24 nucleotides increased the 5'-UTR length to 50 nucleotides which should provide sufficient space for the loading of the 43S PIC on the mRNA.

As shown in Fig. 23B (right panel) the WT *HAC1* and *HAC1 -42 AUG A1G, A7G alleles* when expressed in *hac1*Δ strain showed Tm^R phenotype consistent with our previous results (Fig. 13) [77]. The Hac1 protein was detected on the Western blot analyses (Fig. 23 B, right panel, Western analyses, lane 2 and 3). The *HAC1 24 –nt RNA -42 AUG A1G, A7G* allele when expressed in the *hac1*Δ strain showed Tm^R phenotype (Fig. 23B, right panel, yeast growth test, lane 4) and Hac1 protein was detected on the Western blot analyses (Fig. 23 B, right panel, Western analyses, lane 4). Both of these results suggest that the insertion of 24 nucleotides in the 5'-UTR did not affect the stability of the *HAC1* mRNA, and Hac1 protein was generated from the inserted AUG codon.

As expected, the WT *HAC1* and *HAC1 -42 AUG A1G, A7G* alleles showed Tm^S phenotype in the *hac1*Δ *ire1*Δ strain (Fig. 23B, left panel, yeast growth test, lanes 2 and 3). However, we observed that, the *HAC1 24 –nt RNA -42 AUG A1G, A7G* allele showed Tm^R phenotype when expressed *hac1*Δ *ire1*Δ strain (Fig. 23B, left panel, yeast growth test, lane 4). We observed no Hac1 protein from WT *HAC1* and *HAC1 -42 AUG A1G, A7G* alleles on the Western blot (Fig. 23B, left panel, Western blot analyses, lane 2 and 3). However, Hac1 protein was produced form the *HAC1 24–nt RNA -42 AUG A1G, A7G* allele (Fig. 23B, left panel, Western blot analyses, lane 4). These results are consistent with our hypothesis that the 5'-UTR length of 50 nucleotides

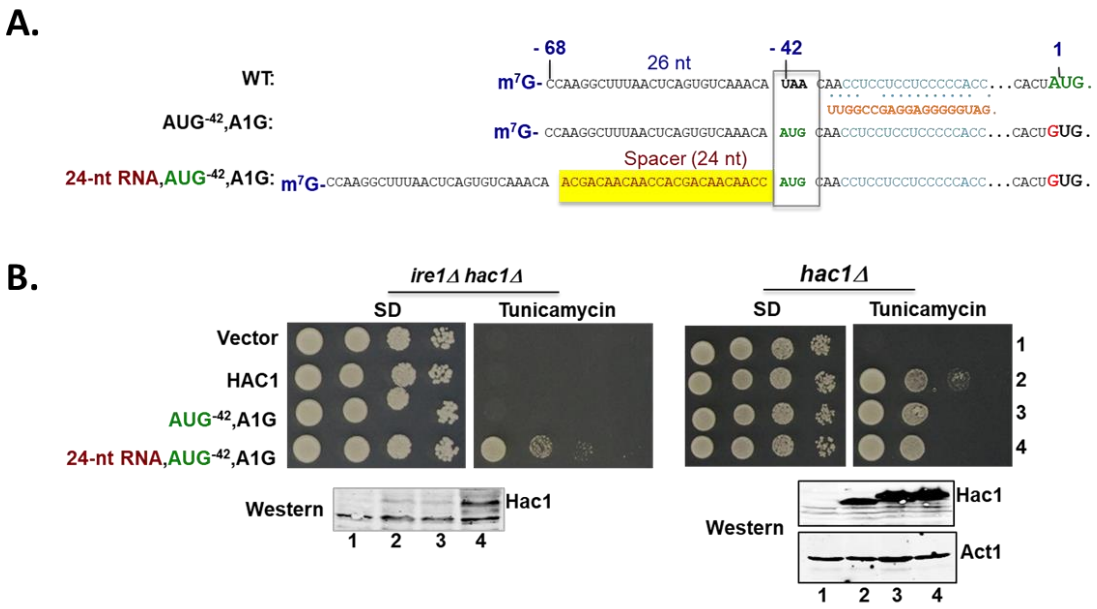

Figure 23. Increasing 5'-UTR length upstream of *HAC1-42AUG* relieves the initiation block: (*A*) Schematic represetation of the nucleotide sequence of the 5'-UTR of *HAC1* mRNA. In the WT HAC1allele the mRNA cap is labeled in blue, the 26 nucleotide sequence upstream of the base-pairing reigion is shown in black color, the -42UAA is indicated followed by the nucleotide sequence of the base-pairing region from the 5'-UTR shown in blue color and nucelotide sequence of the base-pairing region of intron is shown in orange color. The authentic start codon is labeled as 1 and shown in green color. The color scheme is identical for the *HAC1 AUG -42 A1G* and *HAC1 24nt- RN, AUG -42 A1G* allele. The inserted AUG at the position -42 is shown in green and the authentic AUG is mutated to GUG shown in red in the sequecne of *HAC1 AUG -42 A1G* alelle. The spacer sequence (24 nucleotides) is highlighted in yellow, the inserted AUG at the position -42 is shown in green and the authentic AUG is mutated to GUG shown in red in the sequecne of *HAC1 24nt- RN, AUG -42 A1G* allele. (B) Yeast growth assay and Western blot analyses.Yeast growth assay (top panel) The transformants of *hac1Δ ire1Δ* yeast strains carrying an empty vector or expressing the indicated plasmids were grown in SD medium to saturation, and 5 µl of serial dilutions (of OD_{600} = 1.0, 0.1, 0.01, and 0.001) were spotted on SD medium or SD medium containing 0.4 µg/ml tunicamycin and incubated 3 d at 30 ˚C. Western blot analyses (bottom panel) *hac1Δ ire1Δ* and *hac1Δ* strains expressing indicated *HAC1 alleles* were grown in SD medium, treated with 5 mM DTT to induce ER stress, and then WCEs were prepared and subjected to SDS-PAGE followed by Western analysis using polyclonal anti-Hac1 antibodies.

upstream of the base-pairing is relievs the translational block possibly by providing sufficient room for the loading of the 43S PIC.

We further tested our hypothesis in the context of the *WT HAC1* mRNA. We generated multiple *HAC1* alleles by inserting 12 nucleotides (5'- ACGACAACAACC- 3'), 24 nucleotides (5'- ACGACAACAACCACGACAACAACC -3') and 36 nucleotides (5'- ACGACAACAACCACGACAACAACCACGACAACAACC -3') at the position of -42 in the 5'-UTR of the *HAC1* mRNA (see Fig. 24A). The 24 nucleotides sequence is identical to the spacer

inserted in *24–nt RNA -42 AUG A1G, A7G* allele in the previous experiment which is two consecutive repeats of the 12 nucleotide sequence (5’ -15-ACGACAACAACC -7 3') inherent to the 5’-UTR of *HAC1* mRNA. The 12 nucleotide spacer sequence is obtained by single repeat of the nucleotide sequence mentioned above which is inherent to the 5’-UTR of *HAC1* mRNA. And the 36 nucleotide spacer was obtained by three consecutive repeat of the 12 nucleotide spacer (5’ -15-ACGACAACAACC -7 3') (see Fig. 24A).

Insertion of the 12, 24 and 36 nucleotides spacer increased the 5’-UTR length to 38, 54 and 66 nucleotides upstream of the base-pairing region. Based on the previous result we proposed that increasing the 5’-UTR to at least 50 nucleotides should provide sufficient room for the loading of the 43S PIC on the mRNA. Once the 43S PIC is loaded there are two alternate possible outcomes: 1) the 43S PIC would start scanning the mRNA in search of start codon AUG or unwind base-pairing region, reach the AUG codon and begin translation 2) 43S PIC would be defective in scanning due to the base-pairing region and fail to reach the AUG codon.

To dissect above possibilities, we introduced plasmids expressing *HAC1- 12nt, HAC1-24nt* and *HAC1-36nt* alleles in an *ire1*Δ yeast strain. Transformants were then tested for growth on the tunicamycin containing medium. Our results showed gradual increase in the Tm^R phenotype in *HAC1- 12nt, HAC1-24nt* and *HAC1-36nt* alleles respectively (Fig. 24B, lane 3, 4 ad 5). As expected the *WT HAC1* showed Tm^S phenotype in the absence of IRE1 gene ((Fig. 24B, lane 2) The transformants were then subjected to Western blot analyses. We observed *HAC1-24nt* and *HAC1-36nt* alleles showing Hac1 protein on Western blot analyses (Fig 24C, lane 4 and 5) however no Hac1 protein was obtained from the *WT HAC1* and *HAC1- 12nt alleles* (Fig. 24C, lane 2 and 3).

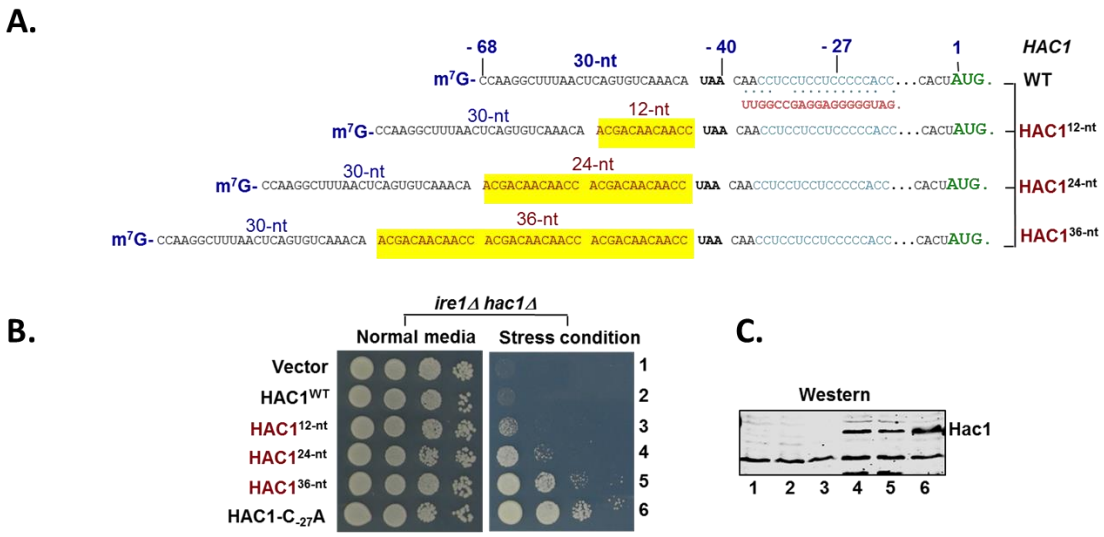

Figure 24. ***HAC1*** **mRNA is translationally silent due to insufficient loading of 43S Pre-initiation complex:** (*A*) Schematic represetation of the nucleotide sequence of the 5'-UTR of *HAC1* mRNA. The color scheme is identical to the Fig 20A. The spacer sequence of 12, 24 and 36 nucleotides is highlighted in yellow. (B) Yeast growth assay. The transformants of *hac1Δ ire1Δ* yeast strains carrying an empty vector or expressing the indicated plasmids were grown in SD medium to saturation, and 5 µl of serial dilutions (of OD_{600} = 1.0, 0.1, 0.01, and 0.001) were spotted on SD medium or SD medium containing 0.4 µg/ml tunicamycin and incubated 3 d at 30 ˚C. (C) Western blot analyses. *hac1Δ ire1Δ* strain expressing indicated *HAC1* alleles were treated with 5 mM DTT to induce ER stress, and then WCEs were prepared and subjected to SDS-PAGE followed by Western analysis using polyclonal anti-Hac1 antibodies.

These results showed that; 1) 38 nucleotide long 5'-UTR of *HAC1- 12nt* allele was not sufficient for the loading of 43S PIC on the mRNA. 2) 54 nucleotide long 5'-UTR provided sufficient room for the loading of 43S PIC on the *HAC1-24nt* allele. Hence, we interpret that, once loaded the 43S PIC shows no scanning defect in *HAC1-24nt* allele (as Hac1 protein is generated from the authentic start codon). Taken together our results provide support for the proposed model where the base-pairing interaction of the 5'-UTR and intron prevents translation initiation by precluding loading of the 43S PIC or eIF4A on the *HAC1* mRNA (see Fig. 25).

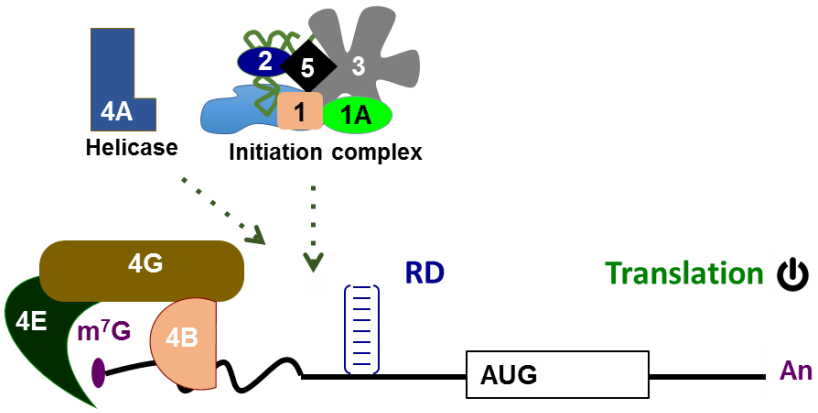

Figure 25. Proposed model showing that the base-pairing interaction (RD) inhibits either loading of 43S PIC (initiation complex) or eIF4A (helicase) on the *HAC1* mRNA.

3.5 Conclusion and future perspectives

The above results suggest that the *HAC1* mRNA is translationally inert due to insufficient loading of the 43S PIC or eIF4A on the mRNA. More experiments such as *in vitro* transcription and translation of *HAC1* mRNA are needed to specifically dissect the 43S PIC or eIF4A loading problem.

4. Adaptation to endoplasmic stress through distinct isoform-specific expression of *HAC1* mRNA

4.1 Introduction

The transcription factor Hac1 in yeast *S. cerevisiae* activates a plethora of genes under conditions of cellular stress. Recent genomic and transcriptomic studies suggest that Hac1 protein is produced from two overlapping mRNA isoforms. One isoform (dubbed here "*HAC1a"*) is the canonical mRNA, containing a short 5'-untranslated region (5'-UTR), two exons (exon1 and exon2) interrupted by an intron and a long 3'-UTR. The second isoform (dubbed here "*HAC1b"*) is a newly discovered mRNA, containing a 5'-UTR, exon1 and a 3'-UTR. As discussed in previous chapters, intron of *HAC1a* binds to its 5'-UTR and keeps mRNA translationally silent. Like *HAC1a, HAC1b,* we found, remains translationally silent by an unknown mechanism. We also found that *HAC1b* can produce an active transcription factor only under conditions of cellular stress. Moreover, we found that the *HAC1b* mRNA level is substantially low compared to *HAC1a* and the Hac1b protein is stabilized in the absence of Duh1, a member of proteasome complex. These results suggest that translational silencing of *HAC1b* mRNA is because of rapid degradation of both mRNA and its translational product. Taken together, we characterized the role of the previously undefined *HAC1b* transcript in adaptation to cellular stress.

4.2 Two isoforms of *HAC1* mRNA are transcribed from the *HAC1* gene locus

We observed that YFL032W gene overlaps with the *HAC1* (YFL031W) gene on chromosome VI. Analyses of previously published yeast transcriptome data showed that two distinct but overlapping transcripts are produced from *HAC1* gene and the overlapping upstream gene [92] [93]. One of the transcripts is the well-studied *HAC1* mRNA having 5'-UTR, exon 1,

cytoplasmic intron, exon 2, and 3'-UTR. This transcript is henceforth referred as "*HAC1a"* (see Fig. 26 A). As reported by previously published transcriptome data, the second transcript has 5'-UTR, exon 1, a 3'-UTR where the nucleotides encoding the exon 1 and the 3'-UTR are shared with *HAC1a* transcript (see Fig. 26). We reviewed literature to look for other evidence supporting this transcriptome analyses.

We identified a microarray analyses studying differential expression of genes under non-stress and ER stress conditions [82]. Microarray data reported the fold change in gene expression under stress conditions as compared to non-stress conditions. We further analyzed the microarray data to identify the genes expressed under ER stress conditions. The data analysis is represented as a volcano plot having transcript length (in Kilo base pairs; kb) on X- axis and fold change in gene expression on Y-axis. The data showed that two distinct transcripts having different lengths were produced from the *HAC1* (YFL031W) and YFL032W gene locus. This result showed that a second transcript, in addition to the "*HAC1a*" transcript was produced. We referred to the second transcript as "*HAC1b*".

To further confirm the presence of *HAC1b* transcript, we performed reverse transcriptase (RT) PCR using wild type yeast *S. cerevisiae* strain. Total RNA was isolated from the yeast strain and RT-PCR was performed as explained in the materials and methods. The primers binding site on the *HAC1* gene locus are shown in Fig. 26. The RT-PCR primers are designed so that primers F1 and R1 bind to only *HAC1b* transcript; primers F2 and R2 bind to both the un spliced and spliced *HAC1a* transcript; and primers F3 and R2 bind to only the spliced *HAC1a* mRNA transcript (Fig. 1C). Amplicons corresponding to the

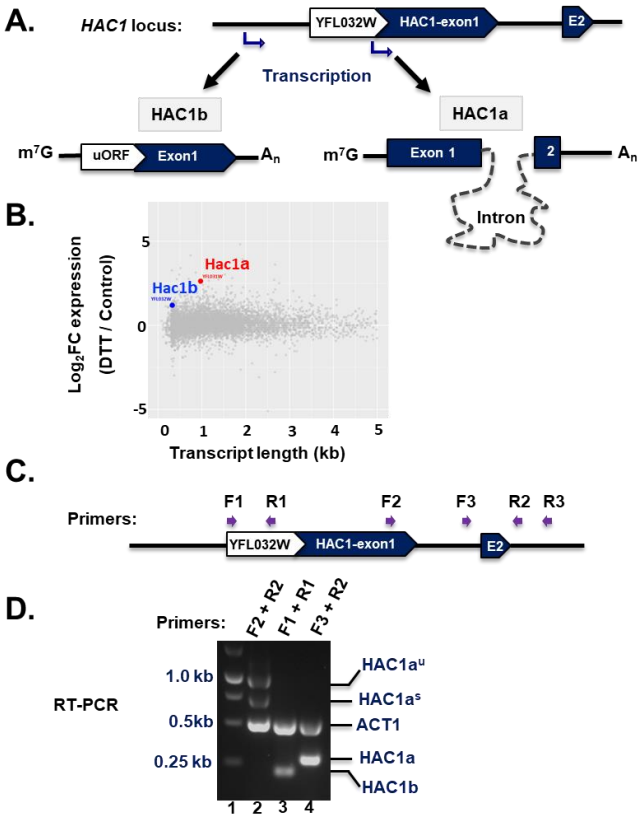

Figure 26. Evidence of two isoforms of *HAC1* mRNA: (A) Schematic representation of *HAC1* gene locus and the two mRNA isoforms produced from the gene locus. The upstream ORF; YFL032W is represented with a white box overlapping with the exon 1 of *HAC1* mRNA shown in blue box, the intron is shown with black line and a shape representing Exon2 labeled as E2. Schematics of *HAC1a* mRNA is shown on right; 5'-end m^7G cap followed by 5'-UTR shown in black line, exon 1 shown in blue box, intron is shown in black dotted line, Exon2 is shown small blue box and 3'-UTR in black line with poly A_n tail. Schematics of *HAC1b* mRNA is shown on left; 5'-end m^7G cap followed by 5'-UTR shown in black line, uORF shown in white box overlapping with the exon 1 shown in blue box followed by 3'-UTR in black line with a poly A_n tail. (B) The volcano plot shows metadata analysis of the microarray data from Travers et al. The volcano plot has transcript length on X-co-ordinate and log_2fold change in gene expression on Y-co-ordinate. The two HAC1 mRNA isoforms; *HAC1a* and *HAC1b* are highlighted in red and blue respectively. (C) Schematic representation of *HAC1* gene locus is as described in (A). The primer binding sites for F1, R1, F2, R2, and F3, R3 are shown. (D) The results Reverse Transcriptase (RT) PCR experiment shown in the agarose gel. The lane 1 has DNA marker with the band corresponding to 0.25kb, 0.5kb and 1.0kb are labeled on left. The primer pairs are shown on top of lane 2,3 and 4. Amplification of *HAC1* mRNA isoforms observed on gel are labeled on right.

un spliced *HAC1a* transcript ($HAC1a^u$) of 1Kb and a spliced *HAC1a* transcript ($HAC1a^s$) of 0.75 kb were observed on agarose gel (Fig.26D, lane 2). An amplicon corresponding to *HAC1b* transcript (~ 0.2 Kbp) was observed upon PCR amplification with primers F1 and R1 (Fig. 26D, lane 3). We also observed PCR amplification of ~0.25 Kbp corresponding to the spliced *HAC1a*

transcript Fig. 26D, lane 4). These results demonstrate the presence of two isoforms, *HAC1a* and *HAC1b* in the wild type cells.

4.3 Both the *HAC1* mRNA isoforms: *HAC1a* and *HAC1b* are translationally repressed

The *HAC1b* transcript consists of a 5'-UTR, a dubious ORF of 321 nucleotides, and 797 nucleotides long 3'-UTR. It is reported that the dubious ORF is unlikely to encode for any functional protein [93]. However, the 797 nucleotide long 3'-UTR encompasses the exon 1 of *HAC1a* transcript, which if translated would produce 230 amino acid long Hac1protein [94]. The 230 amino acid long protein is called as $Hac1^{u}$ ("u" stands for un-stress conditions) which is nonetheless an active transcription factor [94]. If *HAC1b* transcript is able to produce $Hac1^{u}$ protein, it is possible that translation of *HAC1b* transcript is also regulated post transcriptionally like *HAC1a*. To understand the post transcriptional regulation of *HAC1b* transcript, we performed the following experiments.

We generated three Hac1 constructs in single copy number plasmids, which are (1) Hac1a: expressing *HAC1a* mRNA, (2) Hac1b: expressing *HAC1b* transcript and (3) Hac1a-b: expressing both *HAC1a* and *HAC1b* transcripts (see Fig. 26A). The Hac1a construct encodes for the *HAC1a* mRNA consisting of 5'-UTR {69 nucleotides (nt)}, exon 1 (661 nt), intron (252 nt) exon 2 (57 nt). The Hac1b encodes for the *HAC1b* transcript bearing 5'-UTR (341 nt), the dubious ORF (321 nt), and 3'-UTR (797 nt) (see Fig. 26A). The Hac1a-b construct mimics the *HAC1* gene locus encoding both the *HAC1a* and *HAC1b* transcripts. We further transformed these plasmids in *hac1*Δ and *ire1*Δ *hac1*Δ yeast strains. The transformants were further tested for their growth on tunicamycin containing media.

The *hac1*Δ yeast strain expressing Hac1a construct showed tunicamycin resistant (Tm^R) phenotype (Fig. 26B, left panel, lane3) and the *ire1*Δ *hac1*Δ yeast strain expressing the same construct showed tunicamycin sensitive (Tm^S) phenotype (Fig. 26B, right panel, lane 3).

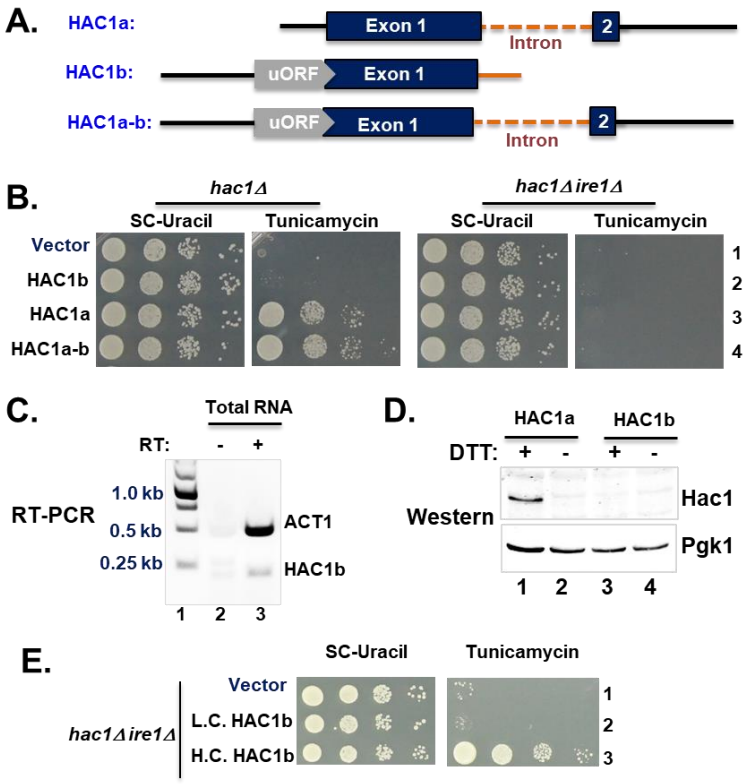

Figure 26. Both *HAC1a* and *HAC1b* isoforms are translationally repressed: (A) Schematic representation of Hac1 constructs: Hac1a, Hac1b and Hac1a-b. The 5'-UTR is represented in black line, uORF represented with a gray box, exon 1 with a blue box intron with orange dotted line and 3'-UTR with black line. (B) Yeast growth test on Sc-Uracil and Sc-Uracil medium containing tunicamycin. *hac1*Δ and *hac1*Δ*ire1*Δ yeast is transformed with constructs with vector, Hac1a, Hac1b and Hac1a-b plasmids. The transformed yeast cells having of A_{600} = 0.4 were serially diluted and spotted on Sc-Uracil and Sc-Uracil medium containing tunicamycin. (C) The results Reverse Transcriptase (RT) PCR experiment shown in the agarose gel. The lane 1 has DNA marker with the band corresponding to 0.25kb, 0.5kb and 1.0kb are labeled on left. The presence of Reverse transcriptase enzyme is labeled on top of the gel. Amplification for *ACT1* and *HAC1b* mRNA is labeled on right. (D) Immunoblot analysis of Hac1 protein: *hac1*Δ yeast strain containing Hac1a and Hac1b were grown in SD medium and treated with 5mM DTT for 2 hours to induce ER stress. The WCEs are prepared and Western blot analysis was performed as explained in materials and methods. The presence and absence of DTT treatment is indicated on top of the gel. (€) Yeast growth test on Sc-Uracil and Sc-Uracil medium containing tunicamycin. *hac1*Δ*ire1*Δ yeast is transformed with vector, low copy Hac1b and high copy Hac1b construct. The transformed yeast cells having of A_{600} = 0.4 were serially diluted and spotted on Sc-Uracil and Sc-Uracil medium containing tunicamycin.

To confirm the Hac1 protein expression from the Hac1a construct, the *hac1*Δ transformants expressing Hac1a construct were grown in the absence and presence of DTT (ER stress inducer) and the WCE was subjected to the immunoblot analysis using an antibody against Hac1 protein (see Materials and methods). As expected, the Hac1 protein was detected only in the DTT treated cell extract (Fig. 26D, lane 1) and was absent in the untreated cell extract (row 2). These results are consistent with the previous reports showing that expression of both Hac1 and Ire1 proteins is required for yeast to survive under the ER stress conditions [94] [71].

Unlike the Hac1a construct, the *hac1*Δ yeast strain expressing Hac1b construct showed Tm^S phenotype (Fig 26B, left panel, lane 2). This phenotype can be attributed to the lack of production of either *HAC1b* mRNA or $Hac1^u$ protein. To determine if *HAC1b* mRNA is synthesized from the Hac1b construct, we performed RT-PCR experiment. The *hac1*Δ yeast cells expressing Hac1b construct was treated with 5mM DTT for two hours and total RNA was isolated as described in the materials and methods. cDNAs were prepared with and without reverse transcriptase enzyme. The *HAC1b* transcript was PCR amplified using the transcript specific primers F1 and R1 as described previously (see Fig. 25). *ACT1* gene specific primers were used to amplify the *ACT1* transcript.

No *ACT1* or *HAC1b* mRNA were detected in the RT-PCR reaction lacking reverse transcriptase enzyme (Fig. 26C, lane 1) confirming the absence of DNA contamination. Very low amount of *HAC1b* mRNA was detected in the RT-PCR reaction with reverse transcriptase enzyme and robust amount of *ACT1* mRNA (Fig. 26C, lane 2). These results showed that very low amount of *HAC1b* mRNA was detected from Hac1b construct and the Tm^S phenotype is not due to the absence of *HAC1b* mRNA. Next, we performed Western blot analysis to test if the *HAC1b* mRNA was translated to produce $Hac1^u$ protein. The *hac1*Δ transformants expressing the Hac1b construct

were grown in absence and presence of DTT (ER stress inducer) and the WCE was subjected to the immunoblot analysis using an antibody against Hac1 protein (see Materials and methods). No Hac1 protein was detected from in both DTT- treated and untreated cells (Fig. 26D, lane 3 and 4). These results were consistent with the growth test (Fig 26B, left panel, lane 2). Collectively, these results suggested that in *hac1*Δ yeast strain, very low amount of *HAC1b* mRNA is detected from the Hac1b construct which (I) might not produce $Hac1^u$ protein (II) might produce $Hac1^u$ protein which is not sufficient to activate UPR and produce Tm^R phenotype (III) might produce $Hac1^u$ protein which is subjected to degradation to prevent activation of UPR genes.

Both the *HAC1a* and *HAC1b* transcripts were generated from the Hac1a-b construct. Like Hac1a construct, Hac1a-b showed Tm^R phenotype in *hac1*Δ yeast strain and Tm^S phenotype in the *ire1*Δ *hac1*Δ yeast strain (see Fig 26B, lane 4). This phenotype was attributed to the expression of *HAC1a* transcript from the Hac1a-b construct, consistent with the above results. These results suggested that both *HAC1a* and *HAC1b* are translationally silent. The translation of *HAC1a* is regulated in Ire1dependent manner however, presence of Ire1 protein has no effect on translational regulation of *HAC1b* mRNA.

It is reported that expression of *HAC1a* transcript from a high copy number plasmid resulted in a Tm^R phenotype in an *ire1*Δ strain [76]. To determine the tunicamycin phenotype upon expression the *HAC1b* from a high copy number plasmid in an *ire1*Δ *hac1*Δ yeast strain we performed the following experiment. We generated a high copy number construct expressing *HAC1b*. This construct was transformed into the in *ire1*Δ *hac1*Δ yeast strain and the transformants were subjected to the growth test on the tunicamycin containing media. We observed that expression of *HAC1b* from high copy number plasmid resulted in Tm^R phenotype in *ire1*Δ *hac1*Δ yeast strain. (Fig 26E, lane 3). The Tm^R phenotype confirms the production of $Hac1^u$ protein from

the *HAC1b* mRNA. Thus, *HAC1b* mRNA is translated to produce the $Hac1^{u}$ protein but when *HAC1b* is expressed from a single copy number plasmid, $Hac1^{u}$ protein produced from the low amount of the *HAC1b* mRNA is not sufficient to show Tm^{R} phenotype. When expressed on a high copy number plasmid, increased amount of *HAC1b* transcript accumulates enough $Hac1^{u}$ protein which results in the Tm^{R} phenotype. Taken together our results suggest that *HAC1b* mRNA is translationally silent and might be regulated by maintaining the transcript abundance.

4.4 *HAC1b* can activate unfolded protein response (UPR) in the absence of Duh1 protein mediated degradation

$Hac1^{u}$ protein is produced by leaky translation of *HAC1a* mRNA under rare conditions [73] [77]. The $Hac1^{u}$ protein is an active transcription factor and thus might activate the target genes in the absence of ER stress conditions. To prevent this aberrant activation of genes by $Hac1^{u}$ under non-stress conditions, a fail-safe mechanism depletes the $Hac1^{u}$ protein by targeting it for degradation [95]. According to a recent report, $Hac1^{u}$ is targeted for ubiquitination by Duh1(Duh1: degrader of un-spliced Hac1) protein and is further degraded by proteasome mediated degradation [95]. Duh1 mediated $Hac1^{u}$ degradation requires a degron sequence (AVITMTRKLQ) encoded by the intronic region of the *HAC1a* mRNA [95]. Sequence analysis of the *HAC1b* mRNA showed that $Hac1^{u}$ produced from *HAC1b* also contains this degron sequence. To determine if $Hac1^{u}$ protein produced by *HAC1b* is a target of Duh1 we performed the following experiments.

We generated a *hac1*Δ *ire1*Δ *duh1*Δ yeast strain as described in the materials and methods. We transformed the *hac1*Δ *ire1*Δ *duh1*Δ yeast with single copy plasmid expressing *HAC1b*. The transformants were tested for the growth on tunicamycin medium. We observed that, *HAC1b* showed Tm^{R} phenotype in the *hac1*Δ *ire1*Δ *duh1*Δ yeast strain (Fig 28A, lane 2) and as expected the *hac1*Δ *ire1*Δ *duh1*Δ yeast strain expressing empty vector showed Tm^{S} phenotype. These results

suggested that in the absence of Duh1 protein, enough Hac1^{u} protein is present to produce Tm^{R} phenotype. Hence, Duh1 protein might have a role in degradation of Hac1^{u} protein produced from *HAC1b* transcript.

To confirm that the above Tm^{R} phenotype is indeed due to accumulation of the Hac1^{u} protein produced from *HAC1b* mRNA, we performed the following experiment. We generated *HAC1b A1G A7G* construct where the adenine of the authentic start codon of exon1 is mutated to guanine and adenine of the AUG codon encoding the third amino acid is also mutated to guanine. The plasmid was transformed in *hac1*Δ *ire1*Δ *duh1*Δ yeast strain and the transformants were tested for growth on tunicamycin medium. The *HAC1b A1G A7G* construct showed Tm^{S} phenotype in *hac1*Δ *ire1*Δ *duh1*Δ yeast strain (Fig 28A, lane 3). These results confirmed that the Hac1^{u} protein produced from *HAC1b* mRNA is encoded by the exon1 and might be targeted for degradation by Duh1 protein.

The Hac1^{u} protein accumulated due to absence of Duh1 protein in *hac1*Δ *ire1*Δ *duh1*Δ yeast strain could be detected on Western blot analyses. The *hac1*Δ *ire1*Δ *duh1*Δ yeast strains expressing a vector, *HAC1b,* and *HAC1b A1G A7G* allele were grown in the presence and absence of DTT. The WCEs were subjected to the immunoblot analyses using an antibody against Hac1 protein (see Materials and methods). Very low levels of Hac1^{u} protein was detected in the DTT-treated *hac1*Δ *ire1*Δ *duh1*Δ yeast cells expressing the *HAC1b* allele (Fig 28B, right panel, lane 2) and no Hac1^{u} protein was detected from the untreated cells (Fig 28B, left panel, lane 2). No Hac1^{u} protein was detected in *hac1*Δ *ire1*Δ *duh1*Δ yeast cells expressing the *HAC1b A1G A7G* allele in both untreated and treated samples (Fig 28B, right panel and right panel, lane 3). Our western blot analyses suggested that (I) In *hac1*Δ *ire1*Δ *duh1*Δ yeast cells Hac1^{u} protein is detected only under

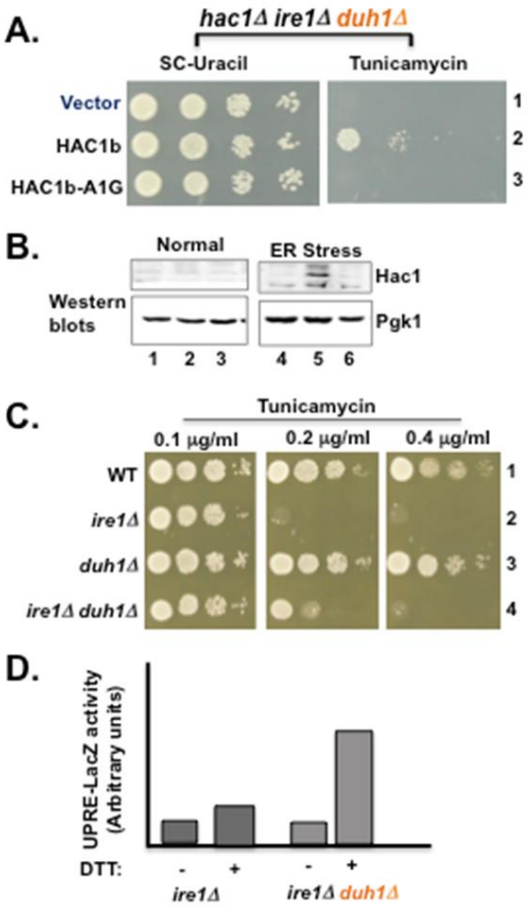

Figure 28. *HAC1b* can activate the UPR in the Duh1 protein null strain: (A) Yeast growth test on Sc-Uracil and Sc-Uracil medium containing tunicamycin. *hac1*Δ*ire1*Δ *duh1*Δ yeast strain is transformed with constructs vector, Hac1b, Hac1b A1G,A7G and Hac1a^{intronless} plasmids. The transformed yeast cells having of A_{600} = 0.4 were serially diluted and spotted on Sc-Uracil and Sc-Uracil medium containing tunicamycin. (B) Immunoblot analysis of Hac1 protein: Left panel- strains from panel A were grown in SD medium. The WCEs are prepared and Western blot analysis was performed as explained in materials and methods. Right panel- strains 1, 2 and 3 from panel A were grown in SD medium and treated with 5mM DTT for 6 hours. The WCEs are prepared and Western blot analysis was performed as explained in materials and methods. Western blot analysis was performed with polyclonal antibody raised against recombinant Hac1 protein. (C) Yeast growth test on YEPD and YEPD medium containing tunicamycin. WT cells, *ire1*Δ, *duh1*Δ *and ire1*Δ *duh1*Δ strains having of A_{600} = 0.4 were serially diluted and spotted. (D) The UPRE-LacZ assay: *ire1*Δ and *ire1*Δ *duh1*Δ strains were transformed with UPRE-1-LacZ plasmid. The cells were grown in YEPD medium with and without DTT treatment. The activity of *ire1*Δ and *ire1*Δ *duh1*Δ strains is represented in bar diagram indicated in arbitrary units.

ER stress conditions mimicked by DTT treatment. (II) The low levels of Hac1^{u} protein detected on Western blot in *hac1*Δ *ire1*Δ *duh1*Δ yeast cells is enough to activate UPR and exhibit Tm^{R} phenotype. Even in the absence of Duh1 mediated protein degradation, very low amount of Hac1^{u}

was detected on the Western blot in *hac1*Δ *ire1*Δ *duh1*Δ yeast cells suggesting the involvement of an additional post-transcriptional regulatory mechanism that might operate by destabilizing the *HAC1b* transcript.

The Hac1^{u} protein produced by *HAC1b* has an active DNA binding domain at the N-terminus. However, it lacks the 18 amino acid tail encoded by exon 2 of the *HAC1a* mRNA, which enables the robust transcriptional activation of UPR elements (UPRE) [76]. Previous reports showed that Hac1^{u} protein expressed on the plasmid is able to activate the transcription of UPRE-driven *LacZ* reporter gene [76]. To determine if the endogenously produced Hac1^{u} protein encoded on *HAC1b* mRNA can activate UPRE-driven *LacZ* genes we used *ire1*Δ *duh1*Δ yeast strain.

We first tested the growth of the WT, *ire1*Δ, *duh1*Δ, and *ire1*Δ *duh1*Δ yeast strains on media containing tunicamycin. Consistent with the previous results the WT yeast strain showed Tm^{R} phenotype whereas the *ire1*Δ yeast showed Tm^{S} phenotype (Fig. 28C, lane 1 and 2) [94]. The *duh1*Δ yeast strain exhibited Tm^{R} phenotype just like WT yeast strain and the *ire1*Δ *duh1*Δ yeast strain showed weak Tm^{R} phenotype as expected (Fig. 28C, lane 4) [95]. Previous paper showed that the Tm^{R} of the *ire1*Δ *duh1*Δ yeast strain which was due to production of Hac1^{u} protein from *HAC1a* transcript under rare conditions. Here we suggest that the *ire1*Δ *duh1*Δ yeast strain would express both *HAC1a* and *HAC1b* transcripts where, *HAC1a* transcript will not produce in Hac1 protein due to absence of *IRE1* gene and absence of *DUH1* gene would accumulate Hac1^{u} protein produced from *HAC1b* transcript. Thus, we propose that Tm^{R} phenotype of *ire1*Δ *duh1*Δ yeast is due to endogenous Hac1^{u} protein produced from *HAC1b* transcript which can activate the UPR genes.

Next to quantify the activation of UPRE by endogenous Hac1^{u} protein we performed the UPRE-LacZ assay. For the assay, the *LacZ* gene is expressed under the promoter with UPR element. The *ire1*Δ and *ire1*Δ *duh1*Δ yeast strains expressing the UPRE-*LacZ* construct were grown in the presence and absence of DTT and the total proteins were subjected to LacZ assay as described before and in the materials and methods [96]. We observed 0.75 times more activation of UPRE- driven *LacZ* gene in the *ire1*Δ *duh1*Δ yeast strain as compared to the *ire1*Δ strain (Fig 28D). These results suggested that endogenously produced Hac1^{u} protein from *HAC1b* transcript can activate the UPR genes under ER stress conditions.

4.5 *HAC1b* bearing long 3'-UTR of *HAC1a* activates UPR in the presence of Duh1

Our previous result showed that very low amount of *HAC1b* transcript is detected in *hac1*Δ yeast strain expressing Hac1b construct. Our next result showed even in the absence of Duh1 mediated Hac1^{u} degradation, Western blot analyses failed to detect robust Hac1^{u} protein in *hac1*Δ *ire1*Δ *duh1*Δ yeast cells expressing the *HAC1b* allele. These results suggested that *HAC1b* transcript might be unstable. We wanted to investigate the mechanism resulting in instability of the *HAC1b* transcript. The 3'-UTR of *HAC1b* mRNA has the exon1 of *HAC1a* followed by a short 130 nucleotide intronic sequence. *HAC1a* transcript has a long 3'-UTR of 521 nucleotides which is lacking in *HAC1b* transcript. To determine the role of short intronic region in the instability of *HAC1b* transcript we performed the following experiments.

We generated a construct Hac1b-TER$_{Hac1a}$ in which we replaced the 130 nucleotide intronic tail of *HAC1b* with the 521 nucleotides long 3'-UTR of *HAC1a* (Fig 29A). We transformed *ire1*Δ *hac1*Δ yeast strain with this construct and the transformants were tested for the growth on tunicamycin containing media. We observed that the *ire1*Δ *hac1*Δ yeast strain expressing Hac1b-

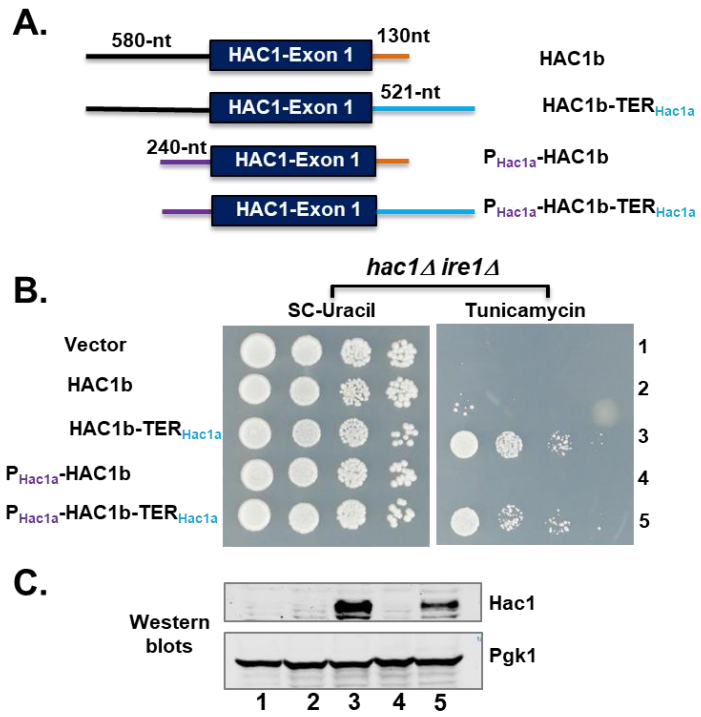

Figure 29. *HAC1b* mRNA Bearing 3'-UTR of *HAC1a* Can Activate the ER stress Response: Schematic representation of Hac1 constructs: Hac1b, Hac1b-$TER_{Hac1a,}$ P_{Hac1b}-Hac1b and P_{Hac1a}-Hac1b-TER_{Hac1a}. (B) Yeast growth test on Sc-Uracil and Sc-Uracil medium containing tunicamycin. *hac1Δire1Δ* yeast strain is transformed with vector, construct expressing Hac1b, Hac1b-$TER_{Hac1a,}$ P_{Hac1b}-Hac1b and P_{Hac1a}-Hac1b-TER_{Hac1a}. The transformed yeast cells having of A_{600} = 0.4 were serially diluted and spotted on Sc-Uracil and Sc-Uracil medium containing tunicamycin. (C) Immunoblot analysis of Hac1 protein: strains from panel A were grown in SD medium. The WCEs are prepared and Western blot analysis was performed as explained in materials and methods. Western blot analysis was performed with polyclonal antibody raised against recombinant Hac1 protein.

TER_{Hac1a} showed Tm^R phenotype (Fig 29B, lane 3) in contrast to the Tm^S phenotype exhibited by *ire1*Δ *hac1*Δ yeast expressing the Hac1b construct (Fig 29B, lane 2). We performed Western blot analyses to detect the $Hac1^u$ protein. Consistent with the growth test, robust Hac1 protein was detected in the *ire1*Δ *hac1*Δ yeast cells expressing Hac1b-TER_{Hac1a} Western blot analyses (Fig 29C, lane 3) and no Hac1 protein was detected in the *ire1*Δ *hac1*Δ yeast cells expressing Hac1b construct (Fig 29C, lane 2). These results suggested that long 3'-UTR of *HAC1a* transcript may serve two purposes. (I) It might improve the stability of *HAC1b* transcript; (II) It does not have the degron sequence (AVITMTRKLQ) encoded by the intron region which targets the $Hac1^u$

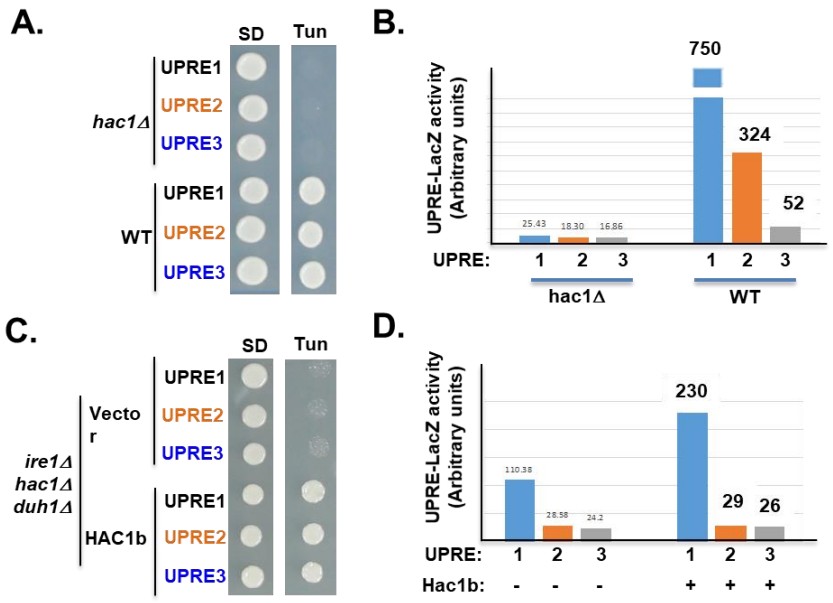

Figure 30. Hac1b activated only UPRE1 driven *LacZ* gene: (A) Yeast growth test on SD and SD medium containing tunicamycin. The *hac1*Δ strain and WT cells transformed with UPRE-1-LacZ plasmid, UPRE-2-LacZ plasmid and UPRE-3-LacZ plasmid. The transformed yeast cells having of A_{600} = 0.4 were serially diluted and spotted on SD and SD medium containing tunicamycin. (B) The UPRE-LacZ activity of the corresponding strains from (A) is represented in bar diagram indicated in arbitrary units. (C) Yeast growth test on SD and SD medium containing tunicamycin. The *ire1*Δ *hac1*Δ *duh*1Δ strain is co-transformed with empty vector and UPRE-1-LacZ plasmid, UPRE-2-LacZ plasmid and UPRE-3-LacZ plasmid. And *ire1*Δ *hac1*Δ *duh*1Δ strain is also co-transformed with vector expressing single copy of Hac1b and UPRE-1-LacZ plasmid, UPRE-2-LacZ plasmid and UPRE-3-LacZ plasmid. The transformed yeast cells having of A_{600} = 0.4 were serially diluted and spotted on SD and SD medium containing tunicamycin. (D) The UPRE-LacZ activity of the corresponding strains from panel C is represented in bar diagram indicated in arbitrary units.

protein for degradation by Duh1. Thus, long 3'-UTR of *HAC1a* transcript also stabilizes with $Hac1^{u}$ protein as observed on Western blot analysis. These results indicate that the 3'-UTR intronic sequence might be responsible for instability of the *HAC1b* transcript.

To determine the role of 5'-UTR region in stability of the *HAC1b* transcript, we performed following experiment. We generated two constructs: (I) P_{HAC1a}HAC1bTER_{HAC1a}; where both the promoter and 3'-UTR sequence of *HAC1b* is replaced by that of *HAC1a* , (II) P_{HAC1a}HAC1b; where only the 5'-UTR of *HAC1b* is replaced by the short promoter sequence of *HAC1a* gene. We transformed both these plasmids in *ire1*Δ *hac1*Δ yeast cells. The transformants were tested for

growth on tunicamycin containing media. We observed that the *ire1*Δ *hac1*Δ yeast expressing P_{HAC1a}HAC1bTER$_{HAC1a}$ showed Tm^R phenotype (Fig 29B, lane 5) whereas *ire1*Δ *hac1*Δ yeast expressing P_{HAC1a}HAC1b showed Tm^S phenotype (Fig 29B, lane 4) much like *HAC1b* construct. (Fig 29B, lane 2). Consistent with the growth test Hac1 protein was detected on the Western blot for P_{HAC1a}HAC1bTER$_{HAC1a}$ construct and was not detected for P_{HAC1a}HAC1b construct (Fig 29C, lane 4 and 5 respectively). Taken together, these results suggested that the long 3'-UTR of *HAC1a* rescues the instability of transcript and protein independent of the 5'-UTR sequence and dubious ORF sequence of *HAC1b* transcript. Promoter of *HAC1a* fails to rescue the instability of *HAC1b* transcript suggesting that 130 nucleotide intronic sequence imparts instability in the transcript as well as $Hac1^u$ protein.

4.6 *HAC1b* activates UPRE-1 driven *LacZ* gene expression

Hac1 is a bZip transcription factor that binds to the unfolded protein response element (UPRE) in the promoters to activate gene expression under conditions of ER stress [76] [70]. Detailed promoter analysis of the UPR responsive genes identified three prominent UPRE sequences named as UPRE-1, UPRE-2 and UPRE-3 [82]. We wanted to know if *HAC1b* can activate UPR genes driven by activating any specific UPRE sequence? We performed the UPRE-LacZ assay to answer this question.

Three plasmids expressing LacZ gene driven by UPRE-1, UPRE-2 and UPRE-3 sequences were transformed in *hac1*Δ and WT yeast cells. The transformants were tested for growth on the tunicamycin containing media (Fig 30A). As expected all the *hac1*Δ transformants showed Tm^S phenotype and WT transformants showed Tm^R phenotype (Fig. 30A left panel). The transformed *hac1*Δ and WT yeast were grown in presence and absence of DTT. The total proteins were isolated and LacZ assay was performed as described in the materials and methods [97]. Our results show

that in the WT yeast stain, the Hac1a protein showed highest LacZ activity in UPRE-1-driven LacZ followed by UPRE-2-driven LacZ and lowest LacZ activity was obtained in UPRE-3-driven LacZ (Fig 30B, WT bars). As compared to the *hac1*Δ strain where no significant LacZ activity was observed (Fig. 30B, *hac1*Δ bars). Our results are inconsistent with the previously published results where UPRE-2-driven LacZ shows maximum activity [82]. After sequencing the UPRE-1 plasmid we observe an additional UPRE-1 sequence element which was absent in the previous studies [82]. The discrepancy in the results is attributed to the presence of the additional UPRE-1 sequence.

Next, we co-transformed the *ire1*Δ *hac1*Δ *duh1*Δ strain with vector expressing *HAC1b* and UPRE-1, UPRE-2 and UPRE-3 plasmid respectively. Similarly, we also co-transformed UPRE-1, UPRE-2 and UPRE-3 plasmids with empty vector as control. The transformed yeast cells were subjected to the growth test on medium containing tunicamycin. As expected the *ire1*Δ *hac1*Δ *duh1*Δ yeast strain expressing *HAC1b* co-transformed with UPRE-1, UPRE-2 and UPRE-3 plasmids showed Tm^R phenotype (Fig 30C, lane 4, 5 and 6) and *ire1*Δ *hac1*Δ *duh1*Δ yeast strain expressing empty vector co-transformed with UPRE-1, UPRE-2 and UPRE-3 plasmids showed Tm^S phenotype (Fig 30C, lane 1, 2 and 3). The transformed yeast strains were grown in the presence and absence of DTT. The total proteins were isolated and LacZ assay was performed as described in the materials and methods [97]. Our LacZ assay results showed that *HAC1b* can activate UPRE-1-driven LacZ expression (Fig 30D). LacZ activation driven by UPRE-2 and UPRE-3 is comparable to that of the vector control (Fig 30D). Thus, our results suggest that *HAC1b* can produce an active transcription factor that is able to activate UPRE-1 driven *LacZ* gene expression.

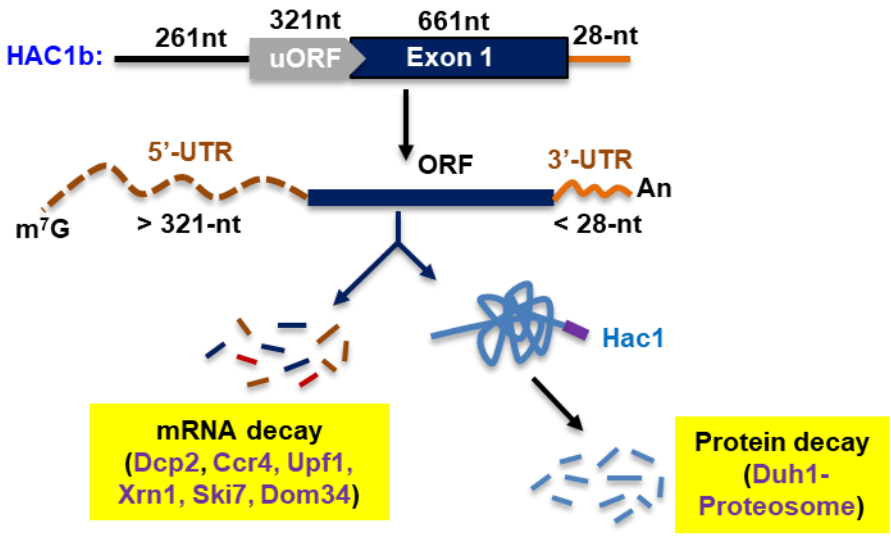

Figure 31. Proposed Turnover of Hac1b mRNA Isoform: The schematic representation of *HAC1b* gene locus shows 5'-UTR indicated in black line, dubious open reading frame shown in gray box labeled as uORF, exon1 indicated in blue box followed by small intronic sequence shown in orange. The schematic of *HAC1b* mRNA has 5'-UTR shown in orange dashed line, followed by ORF encoding the exon one in blue and short 3'-UTR composed of intronic sequence in orange. Most of the mRNA is degraded and only small portion of mRNA undergoes translation to produce Hac1 protein indicated in blue line with degron sequence shown in purple. The Hac1 protein is targeted for degradation by Duh1 dependent proteasome mediated degradation

Taken together our results suggest that two *HAC1* mRNA isoforms; *HAC1a* and *HAC1b* are expressed in yeast cells under endoplasmic reticulum stress conditions. Both these transcripts are translationally silent under normal physiological conditions. *HAC1b* transcript produces $Hac1^{u}$ protein only in the absence of Duh1 protein which targets $Hac1^{u}$ protein for degradation. The endogenous $Hac1^{u}$ protein activates UPR, preferentially by activating UPRE1 driven gene expression. The instability of *HAC1b* transcript is attributed to the short intronic sequence. Based on these results we propose a model shown in Fig. 31 Our model suggests that most of the *HAC1b* transcript is degraded by one or more mRNA decay pathways and the $Hac1^{U}$ protein produced is rapidly degraded by Duh1 mediated proteasome dependent degradation.

4.7 Conclusion and future perspectives

We characterized the role of a previously identified *HAC1* mRNA isoform; *HAC1b* in ER stress conditions. Here we show that the expression of *HAC1b* transcript is tightly regulated by low transcript levels, rapid degradation of the transcript, and proteasome mediated $Hac1^{u}$ protein degradation. Further, our results show that, $Hac1^{u}$ protein obtained from the *HAC1b* transcript is an active transcription factor and able to activate the UPRE-driven gene expression. Collectively, our data shows that *HAC1* gene locus produces two transcripts, both encoding active Hac1 transcription factor however, different mechanisms are exerted to regulate the transcript expression.

Bibliography

[1] N. Sonenberg and A. G. Hinnebusch, “Regulation of Translation Initiation in Eukaryotes: Mechanisms and Biological Targets,” *Cell*, vol. 136, no. 4, pp. 731–745, 2009.

[2] F. Gebauer and M. W. Hentze, “Molecular mechanisms of translational control.,” *Nat Rev Mol Cell Biol*, vol. 5, no. 10, pp. 827–35, 2004.

[3] A. G. Hinnebusch, “The Scanning Mechanism of Eukaryotic Translation Initiation,” *Annu. Rev. Biochem.*, vol. 83, no. 1, pp. 779–812, Jun. 2014.

[4] A. G. Hinnebusch, “Molecular mechanism of scanning and start codon selection in eukaryotes.,” *Microbiol. Mol. Biol. Rev.*, vol. 75, no. 3, p. 434–67, first page of table of contents, 2011.

[5] T. E. Dever, T. G. Kinzy, and G. D. Pavitt, “Mechanism and regulation of protein synthesis in Saccharomyces cerevisiae,” *Genetics*, vol. 203, no. 1, pp. 65–107, 2016.

[6] L. A. Passmore *et al.*, “The Eukaryotic Translation Initiation Factors eIF1 and eIF1A Induce an Open Conformation of the 40S Ribosome,” *Mol. Cell*, vol. 26, no. 1, pp. 41–50, Apr. 2007.

[7] D. N. Wilson and J. H. Doudna Cate, “The structure and function of the eukaryotic ribosome.,” *Cold Spring Harb. Perspect. Biol.*, vol. 4, no. 5, p. a011536, May 2012.

[8] L. A. Passmore *et al.*, “The eukaryotic translation initiation factors eIF1 and eIF1A induce an open conformation of the 40S ribosome.,” *Mol. Cell*, vol. 26, no. 1, pp. 41–50, Apr. 2007.

[9] A. K. Saini, J. S. Nanda, J. R. Lorsch, and A. G. Hinnebusch, “Regulatory elements in

eIF1A control the fidelity of start codon selection by modulating tRNA(i)(Met) binding to the ribosome.," *Genes Dev.*, vol. 24, no. 1, pp. 97–110, Jan. 2010.

[10] Y. Yu *et al.*, "Position of eukaryotic translation initiation factor eIF1A on the 40S ribosomal subunit mapped by directed hydroxyl radical probing.," *Nucleic Acids Res.*, vol. 37, no. 15, pp. 5167–82, Aug. 2009.

[11] J. L. Battiste, T. V Pestova, C. U. Hellen, and G. Wagner, "The eIF1A solution structure reveals a large RNA-binding surface important for scanning function.," *Mol. Cell*, vol. 5, no. 1, pp. 109–19, Jan. 2000.

[12] A. G. Hinnebusch and J. R. Lorsch, "The mechanism of eukaryotic translation initiation - new insights and challenges.pdf," *Cold Spring Harb. Perspect. Biol.*, vol. 4, pp. 1–25, 2012.

[13] A. G. Hinnebusch, "TRANSLATIONAL REGULATION OF *GCN4* AND THE GENERAL AMINO ACID CONTROL OF YEAST*," *Annu. Rev. Microbiol.*, vol. 59, no. 1, pp. 407–450, 2005.

[14] A. G. Hinnebusch and J. R. Lorsch, "The mechanism of eukaryotic translation initiation: New insights and challenges," *Cold Spring Harb. Perspect. Biol.*, vol. 4, no. 10, 2012.

[15] T. E. Dever and R. Green, "The Elongation, Termination, and Recycling Phases of Translation in Eukaryotes."

[16] L. Jenner *et al.*, "Crystal structure of the 80S yeast ribosome," *Curr. Opin. Struct. Biol.*, vol. 22, pp. 759–767, 2012.

[17] J. Rabl, M. Leibundgut, S. F. Ataide, A. Haag, and N. Ban, "Crystal structure of the

eukaryotic 40S ribosomal subunit in complex with initiation factor 1.," *Science*, vol. 331, no. 6018, pp. 730–6, Feb. 2011.

[18] J. L. * Woolford and S. J. Baserga, "Ribosome Biogenesis in the Yeast Saccharomyces cerevisiae," 2013.

[19] B. T. Wimberly *et al.*, "Structure of the 30S ribosomal subunit," pp. 327–339, 2000.

[20] N. Ban *et al.*, "A new system for naming ribosomal proteins HHS Public Access," *Curr Opin Struct Biol*, vol. 24, pp. 165–169, 2014.

[21] J. de la Cruz, K. Karbstein, and J. L. Woolford Jr, "Functions of Ribosomal Proteins in Assembly of Eukaryotic Ribosomes In Vivo."

[22] E. F. Freed, F. Bleichert, L. M. Dutca, and S. J. Baserga, "When ribosomes go bad: diseases of ribosome biogenesis," 2010.

[23] A. Narla and B. L. Ebert, "Ribosomopathies: human disorders of ribosome dysfunction," 2010.

[24] J. L. Llácer *et al.*, "Conformational Differences between Open and Closed States of the Eukaryotic Translation Initiation Complex," *Mol. Cell*, vol. 59, no. 3, pp. 399–412, Aug. 2015.

[25] C. M. Spahn *et al.*, "Hepatitis C virus IRES RNA-induced changes in the conformation of the 40s ribosomal subunit.," *Science*, vol. 291, no. 5510, pp. 1959–62, Mar. 2001.

[26] "Cryo-EM Visualization of a Viral Internal Ribosome Entry Site Bound to Human Ribosomes: The IRES Functions as an RNA-Based Translation Factor."

[27] S. U. Åström, U. von Pawel-Rammingen, and A. S. Byström, "The Yeast Initiator

tRNAMet can Act as an Elongator tRNAMetIn Vivo," *J. Mol. Biol.*, vol. 233, no. 1, pp. 43–58, Sep. 1993.

[28] L. D. Kapp, S. E. Kolitz, and J. R. Lorsch, "Yeast initiator tRNA identity elements cooperate to influence multiple steps of translation initiation."

[29] T. Hussain *et al.*, "Structural Changes Enable Start Codon Recognition by the Eukaryotic Translation Initiation Complex," 2014.

[30] R. Spilka, C. Ernst, A. K. Mehta, and J. Haybaeck, "Eukaryotic translation initiation factors in cancer development and progression," *Cancer Lett.*, vol. 340, no. 1, pp. 9–21, 2013.

[31] P. J. Hernandez and T. Abel, "The role of protein synthesis in memory consolidation: Progress amid decades of debate."

[32] G. C. Scheper, M. S. van der Knaap, and C. G. Proud, "Translation matters: protein synthesis defects in inherited disease," *Nat. Rev. Genet.*, vol. 8, no. 9, pp. 711–723, Sep. 2007.

[33] M. Schröder and R. J. Kaufman, "ER stress and the unfolded protein response," *Mutat. Res. Mol. Mech. Mutagen.*, vol. 569, no. 1–2, pp. 29–63, Jan. 2005.

[34] N. D. Sen, F. Zhou, N. T. Ingolia, and A. G. Hinnebusch, "Genome-wide analysis of translational efficiency reveals distinct but overlapping functions of yeast DEAD-box RNA helicases Ded1 and eIF4A," *Genome Res.*, vol. 25, pp. 1196–1205, 2015.

[35] A. V Pisarev, V. G. Kolupaeva, M. M. Yusupov, C. U. T. Hellen, and T. V Pestova, "Ribosomal position and contacts of mRNA in eukaryotic translation initiation

complexes.," *EMBO J.*, vol. 27, no. 11, pp. 1609–1621, 2008.

[36] L. A. Passmore *et al.*, "EIF5 and eIF5B together stimulate 48S initiation complex formation during ribosomal scanning," *Cell*, vol. 42, no. 3, pp. 12052–12069, 2013.

[37] M. Kozak, "Structural features in eukaryotic mRNAs that modulate the initiation of translation," *J. Biol. Chem.*, vol. 266, no. 30, pp. 19867–19870, 1991.

[38] M. Kozak, "Circumstances and mechanisms of inhibition of translation by secondary structure in eukaryotic mRNAs.," *Mol Cell Biol*, vol. 9, no. 11, pp. 5134–5142, 1989.

[39] M. Kozak, "Influences of mRNA secondary structure on initiation by eukaryotic ribosomes (translational control/insertion mutagenesl/scanning model/RNA stem-and-loop structures/preproinsulin)," *Proc. Nati. Acad. Sci. USA*, vol. 83, no. May, pp. 2850–2854, 1986.

[40] C. E. Aitken *et al.*, "Eukaryotic translation initiation factor 3 plays distinct roles at the mRNA entry and exit channels of the ribosomal preinitiation complex.," *Elife*, vol. 5, Oct. 2016.

[41] A. A. Komar and M. Hatzoglou, "Cellular IRES-mediated translation: The war of ITAFs in pathophysiological states," *Cell Cycle*, vol. 10, no. 2, pp. 229–240, 2011.

[42] I. K. Ali, L. Mckendrick, S. J. Morley, and R. J. Jackson, "Activity of the Hepatitis A Virus IRES Requires Association between the Cap-Binding Translation Initiation Factor (eIF4E) and eIF4G," *J. Virol.*, vol. 75, no. 17, pp. 7854–7863, 2001.

[43] J. S. Kieft, "Viral IRES RNA structures and ribosome interactions Cap-dependent and cap-independent translation in eukaryotes," *Trends Biochem Sci*, vol. 33, no. 6, pp. 274–

283, 2008.

[44] A. Z. Andreou and D. Klostermeier, “The DEAD-box helicase eIF4A: paradigm or the odd one out?,” *RNA Biol*, vol. 10, no. 1, pp. 19–32, 2013.

[45] J. M. Caruthers, E. R. Johnson, D. B. Mckay, and M. F. Morales, “Crystal structure of yeast initiation factor 4A, a DEAD-box RNA helicase.”

[46] R. Cencic and J. Pelletier, “Hippuristanol - A potent steroid inhibitor of eukaryotic initiation factor 4A,” 2016.

[47] Q. Yang and E. Jankowsky, “ATP-and ADP-Dependent Modulation of RNA Unwinding and Strand Annealing Activities by the DEAD-Box Protein DED1 †,” *Biochemistry*, vol. 44, pp. 13591–13601, 2005.

[48] J. M. Caruthers, E. R. Johnson, and D. B. Mckay, “Crystal structure of yeast initiation factor 4A, a DEAD-box RNA helicase.”

[49] B. Goossen and M. W. Hentze, “Position is the critical determinant for function of iron-responsive elements as translational regulators.,” *Mol. Cell. Biol.*, vol. 12, no. 5, pp. 1959–66, 1992.

[50] N. K. Gray and M. W. Hentze, “Iron regulatory protein prevents binding of the 43S translation pre-initiation complex to ferritin and eALAS mRNAs.,” *EMBO J.*, vol. 13, no. 16, pp. 3882–91, 1994.

[51] M. Kozak, “Pushing the limits of the scanning mechanism for initiation of translation,” *Gene*, vol. 299, no. 1–2. pp. 1–34, 2002.

[52] F. Usuki and K. Maruyama, “Ataxia caused by mutations in the-tocopherol transfer

protein gene," 2000.

[53] M. Pirastus, G. Sagliop, J. C. Changll, A. Caos, and Y. W. Kanllll, "THE JOURNAL OF BIOLOGICAL CHEMISTRY Initiation Codon Mutation as a Cause of cu Thalassemia*," 1984.

[54] M. Kozak, "Regulation of translation via mRNA structure in prokaryotes and eukaryotes," *Gene*, vol. 361, no. 1–2, pp. 13–37, 2005.

[55] T. Fukaya and Y. Tomari, "MicroRNAs Mediate Gene Silencing via Multiple Different Pathways in Drosophila," *Mol. Cell*, vol. 48, pp. 825–836, 2012.

[56] T. Dalmay, "Mechanism of miRNA-mediated repression of mRNA translation," *Essays Biochem*, vol. 54, pp. 29–38, 2013.

[57] S. Oliveto, M. Mancino, N. Manfrini, and S. Biffo, "Role of microRNAs in translation regulation and cancer World Journal of Biological Chemistry," *World J Biol Chem*, vol. 8, no. 1, pp. 45–56, 2017.

[58] A. Fukao *et al.*, "MicroRNAs Trigger Dissociation of eIF4AI and eIF4AII from Target mRNAs in Humans," *Mol. Cell*, vol. 56, pp. 79–89, 2014.

[59] E. Izaurralde, "A role for eIF4AII in microRNA–mediated mRNA silencing," *Nat. Struct. Mol. Biol.*, vol. 20, no. 5, pp. 543–545, May 2013.

[60] G. Mathonnet *et al.*, "MicroRNA inhibition of translation initiation in vitro by targeting the cap-binding complex eIF4F.," *Science*, vol. 317, no. 5845, pp. 1764–7, Sep. 2007.

[61] S. Oliveto, M. Mancino, N. Manfrini, and S. Biffo, "Role of microRNAs in translation regulation and cancer World Journal of Biological Chemistry," *World J Biol Chem Febr.*

World J Biol Chem, vol. 26, no. 81, pp. 45–56, 2017.

[62] P. Rajyaguru, M. She, and R. Parker, "Scd6 Targets eIF4G to Repress Translation: RGG Motif Proteins as a Class of eIF4G-Binding Proteins," *Mol. Cell*, vol. 45, no. 2, pp. 244–254, 2012.

[63] J. D. Gross *et al.*, "Ribosome loading onto the mRNA cap is driven by conformational coupling between eIF4G and eIF4E," *Cell*, vol. 115, no. 6, pp. 739–750, 2003.

[64] S. E. O'leary, A. Petrov, J. Chen, and J. D. Puglisi, "Dynamic Recognition of the mRNA Cap by Saccharomyces cerevisiae eIF4E," 2013.

[65] A. Gingras, B. Raught, and N. Sonenberg, "e IF4 I NITIATION F ACTORS : E FFECTORS OF m RNA R ECRUITMENT OF T RANSLATION," pp. 913–963, 1999.

[66] N. Sonenberg and T. E. Dever, "Eukaryotic translation initiation factors and regulators," *Curr. Opin. Struct. Biol.*, vol. 13, no. 1, pp. 56–63, 2003.

[67] H. El Ene Malka-Mahieu, M. Newman, L. D. Esaubry, C. Robert, and S. Ephan Vagner, "Molecular Pathways Molecular Pathways: The eIF4F Translation Initiation Complex-New Opportunities for Cancer Treatment," *Clin Cancer Res*, vol. 23, no. 1, 2017.

[68] A. Z. Andreou, U. Harms, and D. Klostermeier, "eIF4B stimulates eIF4A ATPase and unwinding activities by direct interaction through its 7-repeats region," *RNA Biol.*, vol. 14, 2017.

[69] R. J. Jackson, C. U. T. Hellen, and T. V Pestova, "THE MECHANISM OF EUKARYOTIC TRANSLATION INITIATION AND PRINCIPLES OF ITS REGULATION."

[70] P. M. Fordyce *et al.*, “Basic leucine zipper transcription factor Hac1 binds DNA in two distinct modes as revealed by microfluidic analyses.”

[71] K. Mori, T. Kawahara, H. Yoshida, H. Yanagi, and T. Yura, “Signalling from endoplasmic reticulum to nucleus: Transcription factor with a basic-leucine zipper motif is required for the unfolded protein-response pathway,” *Genes to Cells*, vol. 1, no. 9, pp. 803–817, Sep. 1996.

[72] K. Mori, “The unfolded protein response: the dawn of a new field.”

[73] U. Rüegsegger, J. H. Leber, and P. Walter, “Block of HAC1 mRNA translation by long-range base pairing is released by cytoplasmic splicing upon induction of the unfolded protein response,” *Cell*, vol. 107, no. 1, pp. 103–114, 2001.

[74] J. S. Cox, C. E. Shamu, and P. Walter, “Transcriptional induction of genes encoding endoplasmic reticulum resident proteins requires a transmembrane protein kinase.,” *Cell*, vol. 73, no. 6, pp. 1197–206, Jun. 1993.

[75] C. Sidrauski and P. Walter, “The Transmembrane Kinase Ire1p Is a Site-Specific Endonuclease That Initiates mRNA Splicing in the Unfolded Protein Response,” *Cell*, vol. 90, pp. 1031–1039, 1997.

[76] J. S. Cox and P. Walter, “A novel mechanism for regulating activity of a transcription factor that controls the unfolded protein response.,” *Cell*, vol. 87, no. 3, pp. 391–404, 1996.

[77] L. Sathe, C. Bolinger, M. Amin-Ul Mannan, T. E. Dever, and M. Dey, “Evidence that base-pairing interaction between intron and mRNA leader sequences inhibits initiation of

HAC1 mRNA translation in yeast," *J. Biol. Chem.*, vol. 290, no. 36, pp. 21821–21832, 2015.

[78] J. S. Cox, R. E. Chapman, and P. Walter, "The unfolded protein response coordinates the production of endoplasmic reticulum protein and endoplasmic reticulum membrane.," *Mol. Biol. Cell*, vol. 8, no. 9, pp. 1805–1814, 1997.

[79] C. Sidrauski, J. S. Cox, and P. Walter, "tRNA Ligase Is Required for Regulated mRNA Splicing in the Unfolded Protein Response," *Cell*, vol. 87, pp. 405–413, 1996.

[80] D. Ron and P. Walter, "Signal integration in the endoplasmic reticulum unfolded protein response," *Nat. Rev. Mol. Cell Biol.*, vol. 8, no. 7, pp. 519–529, Jul. 2007.

[81] P. Walter and D. Ron, "The Unfolded Protein Response: From Stress Pathway to Homeostatic Regulation."

[82] K. J. Travers, C. K. Patil, L. Wodicka, D. J. Lockhart, J. S. Weissman, and P. Walter, "Functional and Genomic Analyses Reveal an Essential Coordination between the Unfolded Protein Response and ER-Associated Degradation," *Cell*, vol. 101, pp. 249–258, 2000.

[83] T. Kawahara, H. Yanagi, T. Yura, and K. Mori, "Endoplasmic Reticulum Stress-induced mRNA Splicing Permits Synthesis of Transcription Factor Hac1p/Ern4p That Activates the Unfolded Protein Response," *Mol. Biol. Cell*, vol. 8, pp. 1845–1862, 1997.

[84] M. Sokabe and C. S. Fraser, "A helicase-independent activity of eIF4A in promoting mRNA recruitment to the human ribosome."

[85] A. PAUSE *et al.*, "The requirement for eukaryotic initiation factor 4A (eIF4A) in

translation is in direct proportion to the degree of mRNA 5′ secondary structure," *Rna*, vol. 7, no. 3, pp. 382–394, 2002.

[86] M. Weisser, F. Voigts-Hoffmann, J. Rabl, M. Leibundgut, and N. Ban, "The crystal structure of the eukaryotic 40S ribosomal subunit in complex with eIF1 and eIF1A," *Nat. Struct. Mol. Biol.*, vol. 20, no. 8, pp. 1015–1017, Jul. 2013.

[87] V. P. Mauro and G. M. Edelman, "The ribosome filter hypothesis.," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 99, no. 19, pp. 12031–6, Sep. 2002.

[88] V. P. Mauro and G. M. Edelman, "The Ribosome Filter Redux."

[89] V. P. Mauro and D. Matsuda, "Translation regulation by ribosomes: Increased complexity and expanded scope," 2016.

[90] C. Chen *et al.*, "Dynamics of translation by single ribosomes through mRNA secondary structures," 2013.

[91] U. Nagalakshmi *et al.*, "The Transcriptional Landscape of the Yeast Genome Defined by RNA Sequencing," *Science (80-.).*, vol. 320, no. 5881, pp. 1344–1349, 2008.

[92] U. Nagalakshmi *et al.*, "The Transcriptional Landscape of the Yeast Genome Defined by RNA Sequencing."

[93] D. G. Fisk *et al.*, "Saccharomyces cerevisiae S288C genome annotation: a working hypothesis," 2006.

[94] J. S. Cox and P. Walter, "A Novel Mechanism for Regulating Activity of a Transcription Factor That Controls the Unfolded Protein Response," 1996.

[95] R. Di Santo, S. Aboulhouda, D. E. Weinberg, and R. Green, "The fail-safe mechanism of

post-transcriptional silencing of unspliced HAC1 mRNA," *Elife*, vol. 5, no. OCTOBER2016, pp. 1–23, 2016.

[96] M. A. Mannan *et al.*, "An Ire1–Phk1 Chimera Reveals a Dispensable Role of Autokinase Activity in Endoplasmic Reticulum Stress Response," *J. Mol. Biol.*, vol. 425, no. 12, pp. 2083–2099, Jun. 2013.

[97] M. A. Mannan *et al.*, "An Ire1–Phk1 Chimera Reveals a Dispensable Role of Autokinase Activity in Endoplasmic Reticulum Stress Response," *J. Mol. Biol.*, vol. 425, no. 12, pp. 2083–2099, Jun. 2013.

MATERIAL AND METHODS

A. Yeast strains

X1	F862- *ire1*Δ (ire1::KanMX)	Research genetics
X2	F1764- *hac1*Δ (hac1::KanMX)	Research genetics
X3	F1772- *ire1*Δ *hac1*Δ (ire1::kanMX hac1::NATMX)	This study
X1511	*ire1*Δ *duh1*Δ (ire1::KanMX duh1::hpt I)	This study
X1289	*ire1*Δ *hac1*Δ*duh1*Δ (ire1::KanMX hac1Δ::NatMX duh1::hpt I)	This study

B. Plasmids

Glycerol stock no	Description	Source
D1	lc-pRS313-His3	p700
D2	lc-pRS314-Trp1	p701
D3	lc-pRS315-Leu2	p702
D4	lc-pRS316-Ura3	p703
D5	hc-pRS423, 2 u	p1376
D6	hc-pRS424, 2 u	p1377
D7	hc-pRS425, 2 u	p1378
D8	hc-pRS426, 2 u	p1379
D1038	TIF-2 (eIF4A) B3349-hc-Ura3	Alan Hinnebusch collection
D1039	TIF-11 (eIF1A) B3603-hc-Ura3	Alan Hinnebusch collection
D1040	SUI-1 (eIF1) B4387-hc-Ura3	Alan Hinnebusch collection
D1045	TIF-2 (eIF4A) B3349-hc-Ura3	Alan Hinnebusch collection

D1116	Hac1(A1G, A7G)-D4	This study
D1117	Hac1(-33AUG)-A1G,A7G-D4	This study
D1118	Hac1(-42AUG)-A1G,A7G-D4	This study
D1119	Hac1(-60AUG)-A1G,A7G-D4	This study
D1193	B3898, HC Leu2, elF3a	Alan Hinnebusch collection
D1194	B4629, HC Leu2, TC	Alan Hinnebusch collection
D1195	B3262, HC Leu2, elf2	Alan Hinnebusch collection
D1196	B1873, HC Leu2, GCDG/11 (8,E)	Alan Hinnebusch collection
D1197	B3355, HC Leu2, 4E	Alan Hinnebusch collection
D1198	B3356, HC Leu2, 4G	Alan Hinnebusch collection
D1199	PC 111, HC, α, ß, 8 (elf 2)	Alan Hinnebusch collection
D1200	B3994, HC, Leu2, elf4A	Alan Hinnebusch collection
D1201	PC478, HC, Leu2, TIF5	Alan Hinnebusch collection
D1202	B5300, HC, Leu2,TIF5	Alan Hinnebusch collection
D1203	B3995, HC, Leu2, elf4A-D170E	Alan Hinnebusch collection
D1204	B3998, HC, Leu2, elf4A-R347I	Alan Hinnebusch collection
D1205	B4277, LC, Leu2, Ded1	Tom E. Dever collection
D1278	HAC-1 C-27G in D4	This study
D1279	HAC-1 G771C in D4	This study
D1280	HAC-1 C-27G, G771C in D4	This study
D1296	RPS3 in D7	This study
D1297	RPS4 in D7	This study
D1298	RPS7A in D7	This study
D1299	DPS14 in D7	This study
D1300	RPS15A in D7	This study

D1301	RPS15 in D7	This study
D1302	RPS18 in D7	This study
D1303	RPS19 in D7	This study
D1304	RPS28 in D7	This study
D1390	Hac1 WT spacer (24 nt) in D4	This study
D1435	Hac1 spacer (36nt) WT in D4 (Ura3)	This study
D1436	Hac1 spacer (24nt) biotin WT in D4 (Ura3)	This study
D1825	Hac1b clone in D4 (KpnI/HindIII)	This study
D1911	Hac1 promoter with AUG mutation (A1G, A7G) [Kpn1- Hac1 promoter - AUG mutations -Spe1]	This study
D1951	Hac1β- long 5' UTR- (281nt) long 3' UTR (52nt)	This study
D1997	pPW344 (pJC104) UPRE1-LacZ	Peter walter
D1998	pPW666 UPRE2-LacZ	Peter walter
D1999	pPW667 UPRE3-LacZ	Peter walter
D2022	Hac1a-WT clone in p702 (leu2)	This study
D1988	Hac1b- in Leu2 vector (p702)	This study

C. Media, Buffers and Stocks

C1. Media

Solid media is prepared by adding 2% agar in all the media. All the media is autoclaved at 15psi for 20 minutes.

Luria broth

1 % (W/V) Bactopeptone

0.5 % (W/V) Yeast extract

1 % (W/V) NaCl

Yeast Extract Peptone Dextrose (YEPD)

2 % (W/V) Bactopeptone

1 % (W/V) Yeast extract

2 % (W/V) Dextrose

Yeast minimal media (SD)

0.142 % (W/V) Yeast Nitrogen Base

0.5 % (W/V) Ammonium Sulphate

0.2 % (W/V) Amino acid mixture

2% (W/V) Dextrose

Super Optimal Broth (SOC)

2 % (W/V) Bactopeptone

0.5 % (W/V) Yeast extract

0.05 % (W/V) NaCl

2% (W/V) Dextrose

250 nM KCl = 10 ml

2 M $MgCl_2$ = 5 ml

C2. Buffers

10X Tris-Glycine

Tris-Cl	30.0 g
Glycine	144 g
Water	1000 ml

10X Tris-Glycine SDS

Tris-Cl	30.0 g
Glycine	144 g
Sodium dodecyl sulphate (SDS)	10 g
Water	1000 ml

50X Tris-Borate EDTA

Tris base	540 g
Boric acid	275 g
0.5 M EDTA (pH=8.0)	200 ml
Water	1000 ml

1M Tris-Cl (pH= 6.8)

The pH= 6.8 is adjusted by using concentrate HCl

Tris base	121.1 g
Water	1000 ml

1M Tris-Cl (pH= 8.8)

The pH= 8.8 is adjusted by using concentrate HCl

Tris base	121.1 g
Water	1000 ml

1M Tris-Cl (pH= 7.0)

The pH= 7.0 is adjusted by using concentrate HCl

Tris base	121.1 g
Water	1000 ml

0.5M EDTA (pH= 8.0)

The pH= 8.0 is adjusted using NaOH pallets

EDTA	186.1 g
Water	1000 ml

5X SDS loading dye

5 % β-Mercaptoethanol

0.02 % Bromophenol Blue

30% Glycerol

10 % Sodium Dodecyl Sulphate (SDS)

250nM Tris-cl (pH= 6.8)

5X native loading dye

0.02 % Bromophenol Blue

30% Glycerol

250nM Tris-cl (pH= 6.8)

Poncaeu dye

0.1 % (W/V) Ponceau S dye

1 % Glacial acetic acid

10X TBS (pH= 7.6)

Adjust the pH with 1 M HCl to 7.6.

Tris-Cl	6.05 g

NaCl	8.76 g
Water	1000 ml

Add 0.001 % of Tween 20 in the buffer after preparation.

Polyacrylamide gel composition

Running gel	4 gels	Stacking gel	4 gels
water	15.2ml	water	16.8 ml
Tris-Cl (1.5M) pH= 8.8	10.4ml	Tris-Cl (1M) pH= 6.8	2.4 ml
Acrylamide (30%)	13.6ml	Acrylamide (30%)	4 ml
10% SDS	400µl	10% SDS	240 µl
TEMED	20µl	TEMED	20 µl
10% APS	400µl	10% APS	240 µl

5 % Bovine Serum Albumin (BSA)

1) Dissolve 5 g of BSA in 100 ml of 1X TBS.
2) Filter sterile and store at 4 °C.

10X MOPS (pH = 7.0)

Adjust the pH = 7.0 using 10M NaOH.

3-(N-morpholino) propanesulfonic acid (MOPS)	41.9 g
Sodium acetate	8.2 g
EDTA	3.72 g
Water	1000 ml

C3. Stocks

Ampicillin: 100mg/ ml

Kanamycin: 50mg/ ml

Tunicamycin: 1g/ ml

Dithreothreitol (DTT): 1M (0.154 g in 1 ml water)

Neurothreocine:

Hygromycine:

Kanamycin:

D. Lab protocols

D1. Plasmid DNA isolation

The Biobasic plasmid DNA isolation kit was used for plasmid DNA preparation.

1) Grow single bacterial colony bearing the plasmid in 10 ml of Luria broth overnight.
2) Centrifuge the tube containing overnight culture at 2,900 rpm for 10 minutes. Drain the liquid media by inverting tube on a paper towel for 5-10 minutes.
3) Add 100 µl of Solution I (containing RNaseI) to the pellet, mix well, and keep for 1 minute. Transfer the solution in an Eppendorf tube.
4) Add 200 µl of Solution II to the mixture and mix gently by inverting the tube 4-6 times and keep at room temperature for 1 minute.
5) Add 350 µl of Solution III and mix gently. Incubate at room temperature for 1 minute.
6) Centrifuge at 12,000 rpm for 7 minutes.
7) Transfer the supernatant to the EZ-10 column. Centrifuge at 10,000rpm for 2 minutes.
8) Discard the flow-through in the tube. Add 500 µl of Wash Solution to the column, and centrifuge at 10,000 rpm for 2 minutes.
9) Spin the column at 12,000 rpm for 5 minutes. Discard the flow-through in the collection tube.
10) Transfer the column to a clean 1.5ml microfuge tube. Add 100 µl of TE into the center part of the column and incubate at room temperature for 2 minutes. Centrifuge at 10,000 rpm for 2 minutes.
11) Store purified DNA at -20°C

D2. Yeast genomic DNA isolation

Reagents:

1) Yeast breaking buffer
2) Tris saturated Phenol
3) Chloroform: Isoamyl alcohol (24:1)

Protocol:

1) Take 200 µl of yeast breaking buffer in an Eppendorf tube and resuspend one yeast colony in the buffer.
2) Add about 100 µl of zirconium beads to the tube. Break the cells at 4 °C by vortexing for 10 minutes.
3) Add 200 µl of Tris saturated phenol and invert the tube 2-3 times.
4) Spin the tube at 12,000 rpm for 10 minutes at 4 °C. Transfer the aqueous layer into a new tube.
5) Add 200 µl of chloroform: isoamyl alcohol and spin at 12,000 rpm for 5 minutes.
6) Transfer the aqueous layer in a tube with 500 µl of binding buffer. Mix it and transfer it to the EZ- spin column. Spin at 10,000 rpm for 2 minutes.
7) Discard the flow through. Add 500 µl of wash buffer. Spin at 10,000 rpm for 2 minutes. Discard the flow through and spin at 10, 000 rpm for 5 minutes.
8) Transfer the column in a 1.5 ml Eppendorf tube. Add 35 µl of TE and incubate it in column for 2 minutes.
9) Spin at 10,000 rpm for 2 minutes. Store the genomic DNA at -20 °C.

D3. Chemical competent cells preparation

Reagents:

1) Autoclaved distilled water
2) 10 mM Calcium chloride
3) 10 % Glycerol

Protocol:

Day1:

1) Streak *E.coli* DH5 alpha cells on LB agar plate with and without Ampicillin.

Day2:

1) Autoclave 6 flasks containing 750 ml of Luria broth and one flask with 50 ml of Luria broth. Autoclave 10% Glycerol and centrifuge bottles.

2) Inoculate single colony of *E.coli* DH5 alpha from LB plate without ampicillin in 50 ml of Luria broth. Grow it overnight.

Day 3:

1) Record A_{600} and inoculate A_{600} = 0.05 cells in 750 ml of broth.

2) Grow till A_{600} reaches 0.3- 0.5. Centrifuge the culture at 4,500 rpm for 5 minutes. Collect the cell pallet in one centrifuge bottle.

3) Wash the pellet with cold autoclaved water twice.

4) Resuspend the pallet in 5-8 ml of calcium chloride and glycerol (10nM calcium chloride in 10% glycerol).

5) Aliquot 50 µl of cells in each eppendorf tube in dry ice. Store at -80 °C. Use the cells for chemical transformation.

D4. Electro competent cells preparation

Reagents:

1) Autoclaved distilled water
2) 10 % Glycerol

Day1:

1) Streak *E.coli* DH5 alpha electrocompetent cells on LB agar plate with and without Ampicillin.

Day2:

1) Autoclave 6 flasks containing 750 ml of Luria broth and one flask with 50 ml of Luria broth. Autoclave 10% Glycerol and centrifuge bottles.

2) Inoculate single colony of *E.coli* DH5 alpha from LB plate without ampicillin in 50 ml of Luria broth. Grow it overnight.

Day3:

1) Record A_{600} and inoculate A_{600} = 0.05 cells in 750 ml of broth.

2) Grow till A_{600} reaches 0.3- 0.5. Centrifuge the culture at 4,500 rpm for 5 minutes. Collect the cell pallet in one centrifuge bottle.

3) Wash the pellet with cold autoclaved water twice.

4) Resuspend the pallet in 5-8 ml of 10% Glycerol.

5) Aliquot 50 µl of cells in each eppendorf tube in dry ice. Store at -80 °C. Use the cells for electro transformation.

D5. Bacterial transformation

Reagents:

1) SOC medium
2) Dry bath

Protocol:

1) Thaw 1 vial of 50 µl chemical competent cells on ice.
2) Add 50 ng to 100 ng of Plasmid DNA to the cells.
3) Incubate it on ice for 2 minutes.
4) Set up the dry bath at 42 °C and incubate the tube on dry bath for 90 seconds.
5) Put the tube on ice for 2 minutes.
6) Add 500 µl of SOC in the tube and incubate it on shaker at 37 °C for 1 hour.
7) Transfer the mixture on LB plate with appropriate antibiotic (Ampicillin/ kanamycin).
8) Incubate the plate at 37 °C overnight.

D6. Yeast transformation

Reagents:

1) TE
2) 0.1 M Lithium acetate in TE
3) PEG
4) Calf thymus DNA (CT DNA)

Preparation of reagents:

1) Prepare 1 M Lithium acetate stock solution by dissolving 6.6 g of Lithium acetate in water and filter sterilize. Add 1 ml of 1M lithium acetate in 9 ml of TE to make 0.1 M Lithium acetate in TE.

2) Dissolve 225 g of PEG-3250 from Sigma in 250 ml of TE. Autoclave it and add 25 ml of 1M Lithium acetate.

Protocol:

Day1:

1) Inoculate yeast colony in 5 ml of YEPD or in appropriate minimal medium. Grow overnight.

Day2:

1) Inoculate 1 ml primary culture in 50 ml YEPD.
2) Grow cells at 30 °C for 3.5 – 4 hours.
3) Spin the yeast cells at 2,900 rpm for 10 minutes. Wash the cell pellet with TE and transfer the suspension in an eppendorf tube.
4) Wash the pellet with 0.1 M Lithium acetate in TE.
5) Resuspend the pellet in 750 µl of 0.1 M Lithium acetate in TE. Incubate the tube on shaker at 30 °C for one hour.
6) In an eppendorf add 5 µl of CT DNA and 1 milligram of plasmid DNA with 100 µl of yeast cells from the incubator. Incubate it for 45 minutes.
7) Add 750 µl of PEG in Lithium acetate in each tube. Incubate for 45 minutes more at 30 °C.
8) Incubate the tube at 42 °C for 8 minutes. And spin the tube at 10,000 rpm for 2 minutes.
9) Throw the liquid and resuspend the cells in 400 µl of sterile TE.
10) Spread the supernatant on appropriate plate. Incubate the plate at 30 °C for 48 hours.

D7. Total protein isolation from yeast

Reagents:

1) 100% TCA (Sigma)
2) 20% TCA (diluted in distill water from 100% TCA)
3) 5% TCA (diluted in distill water from 100% TCA)
4) 0.5 mm zirconium beads
5) 1 M Tris-Cl (pH not adjusted)

Protocol:

Day1:

1) Inoculate one yeast colony in 5-10 ml of appropriate liquid medium. Grow it overnight.

Day2:

1) Inoculate the 0.5 ml of overnight culture in 25 ml of liquid medium. Grow cells till the A_{600} reaches 0.6-0.8.
2) Treat the cells with 5 mM DTT for 2 hours for the detection of Hac1 protein expression on Western blot.
3) Harvest the cells. Throw away all the media and transfer the pellet in an eppendorf tube. Wash the pellet with TE and throw away all the supernatant.
4) Add 200 µl of 100 % TCA and incubate overnight at 4 °C.

Day3:

1) Centrifuge for 4 mins at 4000rpm at RT. Pellets contain protein sample.
2) Discard the supernatant and re-suspend in 200 µl 20% TCA.
3) Add zirconium beads and break the cells in cold room with a bead-beater: 2 cycles of 1min 30seconds each.
4) Collect 200 µl of cell suspension in a fresh tube.
5) Add another 200 µl 5% TCA and vortex.
6) Again add 200µl 5% TCA and vortex. Total collection is now 600µl.
7) Centrifuge at 3000 rpm for 10 mins at RT.
8) Throw the supernatant. Proteins are in the pellet.
9) For western blot, add 50 µl 2X LSB. The color is supposed to turn yellow from the acid.
10) Titrate with 50µl 1M Trizma-base (pH not adjusted). Make sure the color changes to a steady blue.
11) Incubate at 95°C and tap in-between to dissolve the cell pellets.
12) Centrifuge at 3000rpm for 10 mins at RT.
13) Load the protein sample on protein gel.

D8. Western blot analysis

Reagents:

Transfer buffer

10X Tris Glycine	50 ml
Ethanol	150 ml
Cold water	350 ml

Day1:

Polyacrylamide gel electrophoresis (PAGE)

1) Load 15 µl of protein sample isolated using TCA method in the 10 % PAGE gel. Run for 65 minutes on 150 Volts. Use 4µl protein ladder.

Transfer

1) Set up transfer using PVDF membrane.
2) Set up overnight transfer at 4 °C using transfer buffer.

Day2:

1) Remove the transfer and stain membrane with Ponceau dye.
2) Wash off the ponceau dye with 1X TBS.
3) Block membrane with 5% BSA for 1 hour at Room temperature.
4) Add primary antibody. Leave overnight at 4 °C.

Day3:

1) Wash the membrane with 1X TBS 5 times (each wash should be for 10 mins).
2) Incubate the membrane with secondary antibody (0.5µl in 5ml BSA).
3) Wash again with 1X TBS 5 times (each wash should be for 10 mins).

D9. Total RNA isolation from yeast

Reagents:

1) Elution buffer

TE	300 µl
RNase free water	700 µl

2) 0.1% DEPC water

Take one-liter MiliQ water in a bottle. Add 0.5 ml of DEPC (from Sigma), shake it gently and keep on table o/n (DEPC is carcinogenic handle with care, wear gloves, open DEPC bottle in fume hood.)

3) Formaldehyde RNA running gel

Dissolve agarose in the TBE then add formaldehyde in the fume hood followed by EtBr.

Agarose	0.7 g
10X MOPS	5 ml
Formaldehyde	750 µl
Ethidium bromide	1.0 µl
TBE	45 ml

Protocol:

1) Add 10 µl of β-ME (Sigma) in 1.0 ml lysis buffer keep it in ice.
2) Add 600 µll of lysis buffer in all the eppendorf tube containing yeast cells and allow cells to thaw on ice.
3) Add Zirconium beads approximate equivalent volume of 100 µl (Sigma or Zirconium).
4) Transfer 600 µl solution in glass bead tubes.
5) Break the cells by vortex in the cold room for 10 minutes (Disruptor genie from Scientific Industries)
6) Centrifuge for 12K for 5 minutes.
7) Collect supernatant 600 µl in a tube containing 600 µl of 70% ethanol, thus total volume 1200 µl.
8) Add 600 µl on a column from Ambion kit, centrifuge (10 K for 2 min), throw liquid from collector.
9) Reload the remaining 600 µl on the same column repeat step 8.
10) Add 500 µl of wash buffer on the column, wash twice.
11) Dry spin for 12K for 5 minutes.
12) Elute RNA with 70 µl of elution buffer. Store at -80 °C.
13) Examine RNA quality by RNA gel electrophoresis. Check the concentration of RNA using a Nanodrop.

D10. Reverse Transcriptase PCR

Reagents:

1) 10mM dNTP – NEB Biolabs dNTP mix- #N0447S (make aliquots of 20 □l and store)
2) 100uM random primer – Biolabs #1254S (make aliquots of 10 □M and store)
3) DEPC-treated water – prepared on your own
4) 5X first strand buffer – Invitrogen- P/N y02321 (make aliquots of 20 □l and store)
5) 0.1M DTT – Invitrogen -- P/N y00147 (make aliquots of 20 □l and store)
6) RNAase out – Invitrogen- P/N 100000840
7) Reverse Transcriptase III - Invitrogen 18080-093

Protocol:

1) First strand cDNA synthesis reaction. Make a primer-dNTP mixture

	For 1 reaction	For 5 reaction
10 mM dNTP	1.0 µl	5.0 µl
10 µM Random primer	0.5 µl	2.5 µl
DEPC water	8.5 µl	32.5 µl
Total	9.5 µl	47 µl

2) Dispense 9.5 µl in 5 eppendorf tubes
3) Add 2000ng of RNA, keep the volume 5 µl
4) Heat at 70 ºC for 3 minutes and transfer to ice immediately (this is to break secondary structures of RNA)
5) Make a reaction mixture as follows:

	For 1 reaction	For 5 reaction
5X First strand buffer	4.0 µl	20 µl
0.1 M DTT	1 µl	5 µl
RNase out	0.2 µl	1 µl

Reverse transcriptase	0.2 µl	1 µl

6) Add 5.4ul in each tube containing RNA mix.

7) Incubate the reaction mixture in a PCR tube with a program for cDNA synthesis:

Temperature	Time
50 ºC	45 minutes
65 ºC	10 minutes
Hold on 10 ºC	

8) Store the synthetic cDNA in -20ºC.

D11. Detection of *HAC1* mRNA splice variants

1) PCR for testing the *HAC1* mRNA splicing

PCR super mix (Invitrogen)	44 µl
10uM Primer omd 1225	2 µl
10uM Primer omd 1226	2 µl
Synthetic cDNA	2 µl

Run program as follows:

Temperature	Time
94 ºC	4 minutes
94 ºC	1 minute
55 ºC	45 Sec
72 ºC	30 Sec
Go to step 2 repeat	20 times
72 ºC	5 minutes
4 ºC	hold

2) Cast a 1.5% agarose gel to run the amplified PCR product

Load 10 µl of PCR product

D12. β -galactose assay

Reagents:

1) 1x Z-buffer

Adjust the pH to 7.0 with NaOH or HCl. Prepare aliquots and Store at -20 °C.

$Na_2HPO_4.7H_2O$	16.1 g
KCl	0.75 g
$MgSO_4 . 7H_2O$	0.246 g
$NaH_2PO_4.H_2O$	5.5 g
12 M β- Mercaptoethanol	2.7 ml
Water	1000 ml

2) ONPG (o-nitrophenyl-β-D-galactoside) stock solution.

Prepare aliquots and Store at -20 °C.

ONPG	4 g
Z- buffer	1000 ml

3) β-galactosidase breaking buffer

Tris-Cl (pH= 8.0)	100 mM
DTT	1 mM
Glycerol	10 %

4) 1M sodium carbonate

Sodium carbonate	5.3 g
Water	50 ml

Protocol:

1) Grow desired yeast cells till A_{600} reaches 0.6. Treat cells with 5mM DTT for 30 minutes to 2 hours as per required.
2) Spin down the DTT treated cells. Throw the media. Wash the pellet with TE and transfer the pellet in eppendorf tubes. Store the pellet at -20°C.
3) Without thawing the pellet add 200 µl of β-galactosidase breaking buffer in cell pellet (add 5 µl β-Mercaptoethanol in 10 ml lysis buffer) on ice.
4) Add about 100 µl of zirconium beads, vortex in cold room for 10 min, spun down at 4 °C for 10000 rpm 5 min.
5) Transfer all supernatant in another clean tube.
6) Thaw ONPG and Z-buffer aliquots at 28°C for at least 1 hour.
7) Add 2 µl of protein sample in 1 ml of Bradford reagent. Record the A_{600} for each sample.
8) Place 1ml Z-buffer in 5 ml glass test tube and equilibrate at 28°C chamber. (Three tubes per sample)
9) Then add 25µl protein supernatant, briefly vortex to mix.
10) Add 200 µl ONPG and vortex keep difference of about 20 sec in each of samples. Record the time.
11) Wait for 5-30 minutes for the development of yellow color. Stop the reaction after yellow color saturates by adding 500µl of sodium carbonate.
12) Record the A_{405}.
13) The beta galactosidase units are calculated by using the following formula.

$$\text{Beta galactosidase Units} = \frac{(OD420 \times 1000)/\text{Concentration} \times 1.7}{0.0045 \times \text{time (min)} \times \text{volume (µl)}}$$

OD420 = Take 1 ml of assay reaction mixture to read the OD420

Concentration = 1 µl /protein concentration, µg/ml

Time (min) = Incubation time (20 min)

Volume (µl) = 25 µl

D15. Flag tagged protein purification

Reagents:

1) Flag purification buffer

1M Tris-Cl (pH= 8.0)	400 µl
5M NaCl	1 ml
Triton X100 (25 %)	40 µl
1M PMSF	34 mg (dissolve in ethanol)
NaF	Pinch
β- Glycerol phosphate	Pinch
Protease tablet	1 mini tablet

Protocol:

1) Total protein isolation: Put the cell pallet on the ice. Add 750µl of buffer. Thaw cells on ice
2) Add glass beads 200µl. Break the cells in cold room for 10 minutes. Spin at 12K for 10 minutes.
3) Take supernatant in fresh tube. Add 500µl of buffer. Break for 10 minutes in cold room
4) Spin at 12K for 10 minutes. Total supernatant 1250µl
5) Bradford assay for protein quantification: Dilute Bradford reagent- 5X to 1X (2ml reagent + 8ml H_20). Dispense 1ml in 8 tubes. Add 2 µl to Bradford reagent. Take OD_{600}. OD_{600} = 0.1 is equivalent to 1µg of protein. Take around 100mg of protein (too much protein is bad for beads)
6) Protein Isolation: Slurry 100µl, add 500µl buffer. Spin 4ºC 3K for 1 minute. Throw supernatant. Wash once more. Mutator shaker 4ºC for 1 hour 30 minutes. Spin at 3k for 1 minute 4 ºC . Throw supernatant
7) wash before elution- to remove unbound protein: Add 750µl flag purification buffer in each tube. Minimum 5 wash. Do Bradford of 5µl wash buffer to check unbound protein concentration
8) Elution peptide preparation- preparation of flag peptide: Elution buffer (50mM Tris pH= 8.00, 10% glycerol) is already prepared]. Weigh 0.5mg of Flag peptide. Prepare 1 M DTT in elution buffer

9) Elution: 45µl of elution buffer with Flag peptide+ 50µl DTT. Add 100-150 µl of Flag peptide in each tube. Tap the tube and wait for 10 minutes. Spin at 10K for 1 minute at 4 ºC. Using 5µl of sample do Bradford

10) PAGE gel: Load around 0.5 µg protein on gel. Use SDS sample buffer (pink color- Biolabs)

    3X loading dye -30µl

    1M DTT -5µl

    Around 10µl sample + 5µl dye

    Heat it

    Load on gel

CURRICULUM VITAE

PERSONNEL

Name: Leena Sathe
Address: Department of Biological Sciences
University of Wisconsin-Milwaukee
3209 North Maryland ave, Milwaukee, WI- 53211

EDUCATION

09/2013 – 05/2019	**University of Wisconsin-Milwaukee,** USA Degree: PhD in Biological Sciences Concentration: Protein Synthesis and Protein Folding Advisor: Dr. Madhusudan Dey Overall GPA: 3.60 /4.0
07/2010 – 05/2012	**Maharaja Sayajirao University**, Baroda, India Degrees: M.Sc. in Microbiology Overall GPA: 8.2/10.00
07/2007 – 05/2010	**University of Mumbai**, Bombay, India Degrees: B.Sc. in Microbiology Overall GPA: 7.20/10.00

PUBLICATIONS

Leena Sathe, Cheryl Bolinger, M. Amin-ul-Mannan, Thomas E. Dever, Madhusudan Dey "Evidence that Base-pairing Interaction Between Intron and mRNA Leader Sequences Inhibits Initiation *HAC1* mRNA translation in Yeast" *Journal of Biological Chemistry* Published 14th July, 2015

Jagadeesh Uppala, Chandrima Ghosh , **Leena Sathe**, Madhusudan Dey "Phosphorylation of Translation Initiation Factor eIF2α at Ser51 Depends on Site- and Context-specific Information" *FEBS Letters* Published 1th August, 2018

Chandrima Ghosh , **Leena Sathe**, Joel David Paprocki , Valerica Raicu and Madhusudan Dey "Adaptation to Endoplasmic Reticulum Stress Requires Transphosphorylation within the Activation Loop of Protein Kinases Kin1 and Kin2, Orthologs of Human Microtubule Affinity-regulating Kinase" *Journal of Molecular and Cell Biology* Published 13th November, 2018

ABSTRACTS AND PRESENTATIONS

Leena Sathe, Chandrima Ghosh, Madhusudan Dey "Mechanistic insights into translational silencing of *HAC1* mRNA" Cold Spring Harbor Laboratory meeting – September 2018, Translational Control, ***Poster***

Leena Sathe, Madhusudan Dey "An RNA duplex inhibits initiation of *HAC1* mRNA translation by precluding pre-initiation complex formation" Graduate Research Symposium- April 2017, University of Wisconsin – Milwaukee ***Talk***

Leena Sathe, Cheryl Bolinger, M. Amin-ul-Mannan, Thomas E. Dever, Madhusudan Dey "Secondary structure formation between 5'-UTR and intron in *HAC1* pre- mRNA impairs translation initiation**"** Cold Spring Harbor Laboratory meeting – September 2016, Translational Control, ***Poster***

Leena Sathe, Madhusudan Dey "An RNA duplex inhibits initiation of *HAC1* mRNA translation by precluding pre-initiation complex formation" Graduate Research Symposium- October 2016, University of Wisconsin – Milwaukee ***Talk***

Leena Sathe, Cheryl Bolinger, M. Amin-ul-Mannan, Thomas E. Dever, Madhusudan Dey "Intron Inhibits Initiation of *HAC1* mRNA Translation" RNA-May 2015, Annual Conference, ***Poster***

Leena Sathe, Cheryl Bolinger, M. Amin-ul-Mannan, Thomas E. Dever, Madhusudan Dey "Intron Inhibits Initiation of *HAC1* mRNA Translation**"** Graduate Research Symposium- April 2015, University of Wisconsin – Milwaukee ***Talk***

Leena Sathe, "Translational control in *HAC1* mRNA" Interdepartmental Sciences Breakfast Talk-August 2014, UWM Research Foundation, ***Talk***

Leena Sathe, Cheryl Bolinger, M. Amin-ul-Mannan, Thomas E. Dever, Madhusudan Dey "Mechanism of regulation of *HAC1* mRNA translation**"** Graduate Research Symposium- April 2014, University of Wisconsin – Milwaukee ***Poster***

AWARDS AND HONORS

Ruth Walker Grant in Aid award- UWM, April 2015

Ruth Walker Grant in Aid award- UWM, April 2016

Ruth Walker Grant in Aid award- UWM, April 2017

Best Microbiology Poster Award - Biological Sciences Symposium- UWM, April 2014

Best Microbiology Talk award- Biological Sciences Symposium- UWM, April 2018

ADDITIONAL ACTIVITIES

Secretary of Graduate Organization of Biological Sciences (GOBS), Department of Biological Sciences, University of Wisconsin-Milwaukee, September 2015- April 2016.

Symposium coordinator for the annual research symposium, Department of Biological Sciences, University of Wisconsin-Milwaukee, September 2017- April 2018