MECHANISM AND REGULATION OF YEAST AND HUMAN MITOCHONDRIAL DNA TRANSCRIPTION

By

URMIMALA BASU

A dissertation submitted to the
School of Graduate Studies

Rutgers, The State University of New Jersey

In partial fulfillment of the requirements
For the degree of
Doctor of Philosophy
Graduate Program in Biochemistry and Molecular Biology
Written under the direction of
Smita S. Patel
And approved by

_____

_____

_____

_____

_____

New Brunswick, New Jersey
May, 2019

ABSTRACT OF THE DISSERTATION

Mechanism and Regulation of Yeast and Human Mitochondrial DNA Transcription

By

URMIMALA BASU

Dissertation Director:

Dr. Smita S. Patel

Mitochondria are autonomous double-membrane-bound organelles in eukaryotic cells. Their most important function is to synthesize ATP to meet the energy needs of the organism through oxidative phosphorylation. This cardinal role in energy production makes mitochondria a key player in metabolic, degenerative, and age-related diseases. Dysregulations of mitochondrial energy production in humans affect all organs, but the high energy demanding organs of the body like the heart, brain, kidneys, etc., are the primary targets of mitochondrial malfunctions.

Mitochondria contain their own genome that is replicated and transcribed by enzymes distinct from the nuclear enzymes. Interestingly, the mitochondrial DNA (mtDNA) replication and transcription enzymes are homologous to bacteriophage T7 encoded

enzymes. Thus, yeast and human mtDNA are transcribed by phage T7-like single-subunit RNA polymerases (RNAP) called Rpo41 and POLRMT, respectively. They are structurally homologous to T7 RNAP, but both yeast and human RNAPs require transcription factors to initiate transcription which includes: Mtf1 (in yeast) and mitochondrial transcription factor A /TFAM and mitochondrial transcription factor B2 /TFB2M (in humans). Reliance of the mtRNAP on transcription factors, results in regulation of gene expression that is unprecedented in homologous phage T7. Transcription initiation is a crucial step where gene expression is regulated. However, mtDNA transcription initiation, elongation, and termination mechanisms are understudied. The overarching goals of my graduate research were to investigate the mechanisms of regulation of transcription initiation through biochemical and biophysical characterization.

In the first part of the thesis, I have showed that yeast and human mtRNAPs can initiate transcription with NAD^+/NADH, which results in production of capped RNA transcripts in the mitochondria. Capping with NAD^+/NADH is a new discovery brought to light in the last decade. My research has showed that initiation with NAD^+ or NADH is up to 40% as efficient as initiation with ATP for *S. cerevisiae* mtRNAP and up to 60% as efficient as initiation with ATP for human mtRNAP. Similarly, direct quantitation of NAD^+- and NADH-capped RNA *in vivo* showed up to ~60% NAD^+ and NADH capping of yeast mitochondrial transcripts and up to ~10% NAD^+ capping of human mitochondrial transcripts. Furthermore, both *S. cerevisiae* mtRNAP and human mtRNAP can cap RNA with NAD^+ and NADH more efficiently than bacterial RNAP and eukaryotic nuclear RNAP II. The capping efficiency is higher with promoter derivatives

having R:Y at position -1 than with promoter derivatives having Y:R. The implications of alternatively capping mitochondrial RNAs are not known; however, capping can affect RNA stability, processing, and global gene expression. Since intracellular NAD^+ and NADH levels dictate the efficiency of capping, we propose that mtRNAPs use NAD^+/NADH capping as both a sensor and actuator in coupling cellular metabolism to mitochondrial transcriptional outputs, sensing NAD^+ and NADH levels and adjusting transcriptional outputs accordingly.

In the second part of the thesis, I studied the role of transcription initiation factors, Mtf1 and TFB2M, in transcription initiation. Specifically, I focused on understanding the function of the C-terminal region (C-tail) of Mtf1 and TFB2M. I engineered and purified recombinant proteins with deletions in the C-tail of *S. cerevisiae* Mtf1 and human TFB2M and investigated the effect of C-tail deletion on the various steps in transcription initiation. Using 2-aminopurine fluorescence-based studies, I have showed that the C-tail of Mtf1 is not necessary for promoter melting but the C-tail of TFB2M is essential for promoter melting. Nevertheless, deletion of the C-tail in both systems decreases the affinity for the initiating nucleotide, which indicated that the C-tail is critical for aligning the template strand in the active site. The biochemical phenotypes of C-tail deletion in Mtf1 resemble those of the sigma factor 3.2 region deletion in bacterial RNAP and B-reader loop in the RNA Pol II system. Thus, I have identified the mechanism of template alignment in mtRNAPs that is needed generally to initiate transcription from a specific site on the double-stranded DNA.

ACKNOWLEDGEMENTS

A dissertation is the culmination of knowledge acquisition, on a specific topic, gathered over a span of 5-6 years through focused individual effort. However, having worked in a scientific laboratory for the past six years, I realize how deeply this effort, has been overwhelmingly supported by the people around me both personally and professionally. I would like to thank those people in this section.

I begin this section by thanking my advisor Dr. Smita S. Patel. To me, she has been an epitome of what a scientist should be: smart, curious, persistent, industrious, and relentless. She has been a constant source of inspiration and encouragement in every aspect of my graduate life. I have found her office door always open which I think is quite symbolic of how she is always open to discussions with students. I appreciate the amount of effort she has put in to help me improve my communication skills. Working on grants and fellowships with her, has been particularly instructive. I hope I will be able to take some of this training in the next steps of my professional life.

I would like to thank Dr. Bryce E. Nickels and Dr. Richard H. Ebright for being excellent collaborators and committee members. This collaboration has enabled me to work in a different lab which helped me pick up new skills, get more exposure and forge excellent professional and personal relationships. I would like to thank my other committee members Dr. Mike Hampsey and Dr. Marc Gartenberg for being very supportive and for always being on top of things. I appreciate all your time and effort in providing valuable input to my work. I would like to thank Dr. Hajin Kim and his lab for being such wonderful collaborators.

I would like to thank Carolyn Ambrose, Diane Murano, Tina Marottoli, Perry Dominguez Walter Barnes and Lettice Smith for their invaluable role in administrative support that helped to keep my work running smoothly. Special thanks to Jamie Carr for being so fantastic in her job and helping me smoothly navigate the grant handling process. I would like to thank our janitor Anna for taking care of the lab and for our friendly conversations. I would like to thank Annette for always making sure we had clean glassware and for our early morning chit-chats.

It takes a village to raise a child. My past and present colleagues in the Patel lab have created an excellent lab environment and I am grateful for their support. I would like to thank Dr. Shemaila Sultana, Dr. Aishwarya Deshpande and Dr. Aparna Ramachandran for being my earliest mentors in the lab and training me in protein purification and biochemical techniques. Special thanks to Aparna for always giving me a shoulder to cry on and for being my friend, be it rain or shine. I would like to thank Brandon for being the 'best isle mate ever' and always telling me "You are a superstar" (I still don't want to hear about the Eagles!). I would like to thank Jiayu and Laura for the endless laughs we shared and our deep conversations about life, love, science, career, children, and politics. I'm telling you ladies; our ideas can change the world! I would like to thank Anupam, Mihai, and Candice for always being so kind and friendly to me. Last but not least, I would like to thank Nandini and Mohammed for being the excellent mentees that they are. Training you guys have had a profound impact on me and I have learned from you more than I would have believed to be possible.

I would like to especially thank Dr. Jeremy Bird for being an excellent mentor to me and introducing me to the wonderful world of RNA capping. I have benefitted immensely from your deep understanding of the NAD field, technical acumen, willingness to teach and amazing temperament. I would like to thank Dr. Jared Winkelman for teaching me cross-linking techniques and being my friend. I would like to thank all members of the Nickels lab for the amazing time I have had with you. I would like to thank my friends and classmates for their support in this long and often tiresome journey. I am grateful for the people I have met here at Rutgers and the lasting impression you all have made. I would like to thank all my teachers throughout the years: thank you for guiding me in every step of the way.

I do not believe it is possible to thank your family adequately. But here is my shot at it. I would like to thank my father for getting me interested in Mathematics and Science at an early age. Your insistence of understanding the fundamental concepts of any topic has helped me come a long way. I would like to thank my mother for instilling a sense of discipline and perseverance in me. No other person has influenced me more than you have Maa and thank you for guiding the ship through some very tough waters. I would like to thank my sister for having an enormous influence on me. I admire your calmness and strength, and after Maa, you are the woman I admire the most. I would like to thank my brother-in-law for our cheerful conversations and always assuring me that everything will be fine. A special thank you to my baby niece Titli: you are a piece of my heart, and I cannot wait to see you grow up! I would like to thank my paternal grandparents who were my very first teachers. They nurtured a very curious child and have contributed immensely towards molding my personality. I would like to thank my maternal

grandparents and extended family for their endless love and support. A special thanks to my maternal grandmother for being an endless source of strength for me. Last but not the least, I would like to thank my husband for being my best friend (you have had the job for more than a decade now!), my counselor, my fiercest critic (I really appreciate it although you don't agree!) and most of all for being 'my person'. Surfing the waves has just begun!

DEDICATION

To

Perseverance

&

Patience

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

[A] INTRODUCTION

Mitochondria are at the heart of energy production in eukaryotic cells producing usable energy in the form of ATP through oxidative phosphorylation pathways. The first half of the twentieth century made major advances in identifying ATP as the energy currency and delineating the various steps of aerobic respiration. However, it was only in the latter half of the century that mitochondria were defined as the center for aerobic metabolism. Early electron micrograph studies revealed the iconic double-membrane-bound anatomy of the mitochondria along with the highly convoluted cristae (Palade 1953, Sjostrand 1953). These studies also identified mitochondria as the 'powerhouse' of the cell. It was discovered that enzymes like succinoxidase, cytochrome oxidase, adenylate kinase, glutamate dehydrogenase, transaminases, pyruvate carboxylase were localized exclusively to mitochondria, fatty acid oxidation and the citric acid cycle also occurred exclusively in the mitochondria (Pagliarini and Rutter 2013). These studies laid the foundations of mitochondrial biology. With the debut of the chemiosmotic coupling concept in the 1960s (Mitchell 1966) and subsequent elucidation of the OXPHOS pathways in the 1970s (Boyer, Chance et al. 1977), it was believed that the fundamental discoveries in the field had been already made. Thus, the traditional view of mitochondria became restricted to its function as 'the powerhouse of the cell.'

This monochromatic perspective of the organelle was challenged in the 1990s with the discovery of the role of mitochondria in programmed cell death aka apoptosis (Green and Reed 1998). A profound discovery in the apoptosis field was made: cytochrome c, an electron carrier operating within the OXPHOS system, was released from mitochondria upon apoptotic insult when it initiated caspase activation and the execution of apoptosis

(Liu, Kim et al. 1996, Li, Nijhawan et al. 1997). Work in the 1990s and 2000s also established the role of mitochondrial dysfunction in many rare and common human diseases (Wallace 1999). We now know that these complex and dynamic organelles play astoundingly diverse roles in metabolism, cell signaling, apoptosis etc. (Pagliarini and Rutter 2013) (McBride, Neuspiel et al. 2006, Nunnari and Suomalainen 2012). Their cardinal role in cellular energetics makes mitochondria a key player in metabolic, degenerative and age-related diseases (Shadel 2008, Shutt and Shadel 2010). Dysregulation of mitochondria is ubiquitous in diabetes, obesity, cardiovascular disease, cancer, etc.

To date, there have been several key discoveries in mitochondrial biology that have been awarded the Nobel Prize. Peter Dennis Mitchell was awarded the 1978 Nobel Prize for Chemistry for his discovery of the chemiosmotic mechanism of ATP synthesis. In 1997, Paul D. Boyer and John E. Walker were awarded the Nobel Prize "for their elucidation of the enzymatic mechanism underlying the synthesis of adenosine triphosphate (ATP)." In 2016, the Nobel Prize in Physiology or Medicine was awarded to Yoshinori Ohsumi for "discoveries of the mechanisms for autophagy."

Mitochondrial-related diseases mainly stem from defects in ATP synthesis, which is the major energy source for the heart that continuously pumps blood to all organs of the body to meet their oxygen demands. Mitochondrial DNA codes for some of the proteins of the five complexes that make the ATP synthesis machinery (Fig. D.2). Mitochondrial transcript abundance is higher in energy demanding organs like heart, muscles etc. (heart cells have the highest levels of mitochondrial RNA transcripts accounting for 30% of total RNA in the cell) and lower (~5%) in low energy consuming organs like lung, ovary,

thyroid, prostate etc. (Mercer, Neph et al. 2011). The transcript levels decrease in the senescent/diseased heart (Rosca and Hoppel 2010, Rosca and Hoppel 2013), which causes age/disease related decrease in ATP levels. There is mounting evidence that decrease in mitochondrial energy production is the main cause of heart failure.

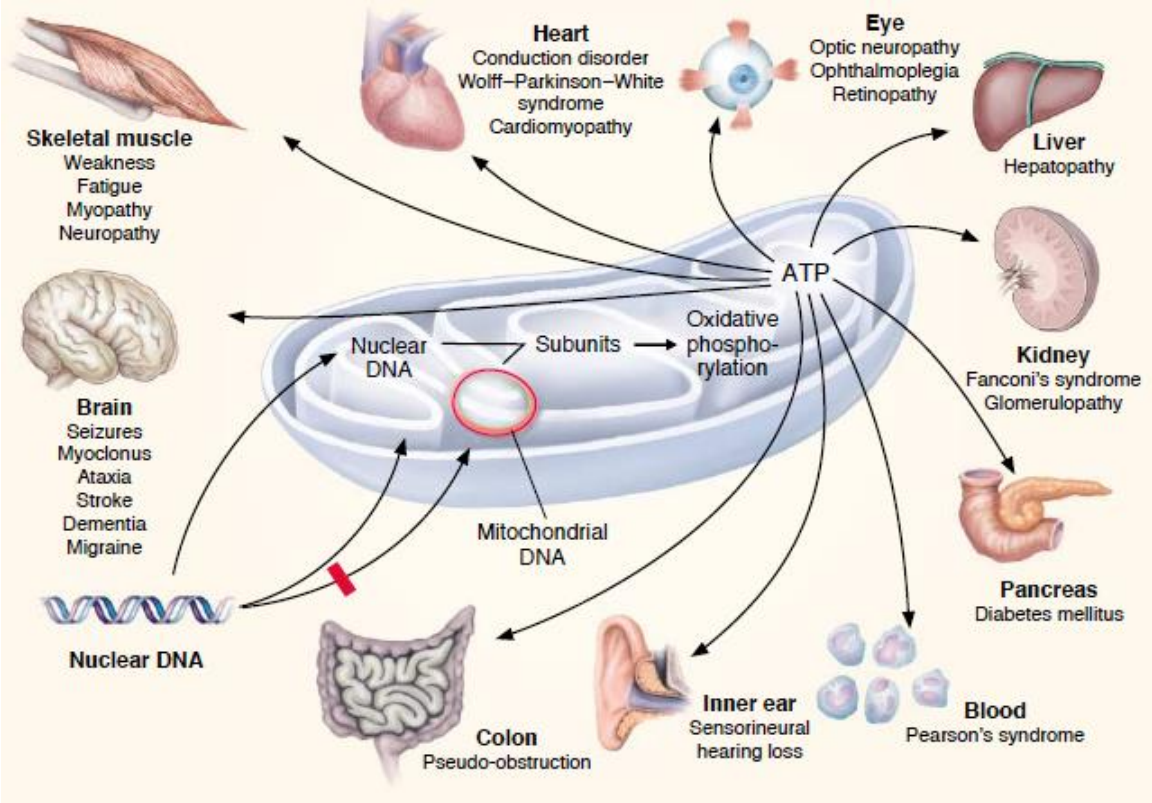


Figure A.1: Role of mitochondria in health and disease. [Adopted from Child Neurology Foundation.]

For instance, in hypertrophic and dilated cardiomyopathies, the mitochondrial morphology changes and this is related to a decrease in ATP synthesis and energy deficit. Similarly, mitochondrial DNA mutations leading to defective transcription cause inherited MELAS (Mitochondrial Myopathy, Encephalopathy Lactic Acidosis, Stroke-like Episodes) (Chomyn, Martinuzzi et al. 1992).

There is a consensus that mitochondria originated from an ancient endosymbiotic event where an Archaea-type host engulfed a ∝-proteobacterium–like ancestor ~1.5 billion years ago. This hypothesis mainly stems from analyses of mitochondrial genes and their genomic organization and distribution (Gray, Burger et al. 1999, Lang, Gray et al. 1999, Dyall, Brown et al. 2004). This endosymbiotic event led to mutual benefits for the host and the ∝-proteobacterium: it equipped the host with 'compartmentalized bioenergetic and biosynthetic factories' (Dyall, Brown et al. 2004) while the endosymbiont acquired access to various metabolites from the host. A crucial step in the transition from autonomous endosymbiont to organelle was genome reduction: many endosymbiont genes have been lost while others have been transferred to the nucleus. Although these gene transfer events have plateaued in most eukaryotes, frequent functional gene transfer events are still seen in angiosperms and yeast (Thorsness and Fox 1990, Adams and Palmer 2003).

A mitochondrion is a world of its own. Mitochondria have their replication, transcription and translation machinery. Transcription initiation is a key event that regulates the amount of RNA made in the mitochondrion. However, the details of transcriptional regulation are not well understood. My studies will address this gap by biochemically characterizing the steps of transcription initiation, comparing and contrasting the yeast

and human mitochondrial systems. My research will provide a fundamental understanding of transcription in yeast and human mitochondria, without which we cannot use knowledge-based methods to design therapies for many mitochondrial diseases.

Yeast and Human mitochondrial genomes and promoters

Saccharomyces cerevisiae mitochondrial genome is an 85 kbp, double-stranded genome that codes for thirteen proteins of the ATP synthesis complex, two ribosomal RNAs (21S and 14S), 24 transfer RNAs, RNA-subunit of mt RNase P, and one ribosomal protein (Biswas 1999). I have studied *Saccharomyces cerevisiae* strain of yeast: thus, any reference of yeast throughout this dissertation means *Saccharomyces cerevisiae.* The yeast mt promoters contain a consensus sequence of 9 base pairs: 5′-(−8) ATATAAGTA (+1) NT sequence (Biswas 1999). This is a remarkably short sequence compared to T7 RNAP (19 base pairs). There are 28 conserved promoters in yeast mt DNA.

The rate of transcription from yeast promoters vary widely with the strong promoters being transcribed 15-20 times faster than the weak promoters (Wettstein-Edwards, Ticho et al. 1986). The +2 position is critical in determining the strength of the promoter: +2 purine and +3 pyrimidine in the non-template strand are seen in strong promoters while +2 pyrimidine and +3 purine are weak promoters (Biswas and Getz 1986, Wettstein-Edwards, Ticho et al. 1986) (Deshpande and Patel 2014). The G:C base pair at -2 position is critical to mt transcription initiation (Biswas and Getz 1986). Though all +1 positions have conserved A:T base pairs, surprisingly, it can be changed to any base pair without affecting transcription (Biswas, Ticho et al. 1987) (Deshpande and Patel 2014).

The human mt genome is a 16.5 kbp, double-stranded circular genome that codes for 13 essential membrane proteins of the oxidative phosphorylation (OXPHOS) pathway. It also codes for two rRNAs and 22 tRNAs required for mt translation (Bonawitz, Clayton et al. 2006). Human mt genome is unique in that there are no introns. This is probably the reason why it is so much smaller than the yeast mt genome (16 kbp vs. 80 kbp). The 37 genes on the human genome are extremely closely spaced. The only non-coding region in this genome is the 1.1 kb displacement-loop control region (D-loop) which is named such due to a stable three-stranded DNA structure found here (Shadel and Clayton 1997). There are only three promoters in the human mitochondrial genome: Heavy Strand Promoters (HSP1 and HSP2) and Light Strand Promoter (LSP). The two strands differ in G+T content and can be separated in alkaline CsCl density sedimentation coefficient method from which the naming of the strands follow (Clayton 1982). Both strands of the genome code for various genes: HSP1 codes for two rRNAs, 12 mRNAs, and 14 tRNAs. LSP codes for one mRNA and eight tRNAs. Transcription of LSP also produces RNA primers required for mtDNA replication initiation (Shadel and Clayton 1997). These promoters are species-specific as it has been shown that recombinant mouse and human transcription machinery are unable to initiate transcription from heterologous LSP promoters (Gaspari, Falkenberg et al. 2004). Transcripts are long polycistronic RNA that is almost as long as the entire genome. This transcript undergoes trimming to excise tRNAs that releases mature mRNAs and rRNAs (Clayton 1982, Chang, Hauswirth et al. 1985, Bonawitz, Clayton et al. 2006, Bestwick and Shadel 2013).

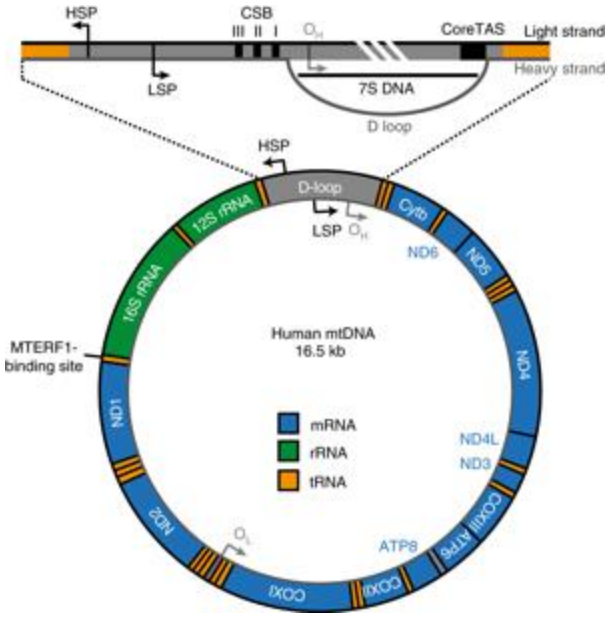


Figure A.2: Human mitochondrial DNA. The transcripts from the three promoters are shown in various colors. LSP produces one mRNA and eight tRNAs; HSP1 (H1) produces two rRNAs and two tRNAs, and HSP2 (H2) produces 12 mRNAs, 12 tRNAs. (Hillen, Temiakov et al. 2018).

Yeast mt RNAP Rpo41 and Human mitochondrial RNAP POLRMT

Sequence analysis first speculated that there might be a structural similarity between T-odd phage RNAPs and Rpo41(Masters, Stohl et al. 1987). The T-odd phages and the mt RNAP is distantly related to the Pol A family of DNAP. The catalytic C-terminal domain resembles a right hand formed by thumb, palm and fingers domain (Sousa, Chung et al. 1993). The palm domain harbors the active site. It utilizes a basic two-metal-dependent

reaction mechanism for nucleic acid polymerization (Steitz, Smerdon et al. 1994, Sosunov, Sosunova et al. 2003). O helix present in the fingers domain is involved in catalysis and strand separation at the downstream edge of the bubble (Tahirov, Temiakov et al. 2002, Yin and Steitz 2002). The NTD which brings about promoter binding and constitutes of two domains: AT-rich recognition loop that recognizes DNA and the intercalating β-hairpin which separates the DNA strands to form the transcription bubble (Cheetham, Jeruzalmi et al. 1999, Cheetham and Steitz 1999). Specificity loop located within the CTD binds to the major grove upstream of the point of strand separation. This guides the template strand to the active site and also stabilizes the initially melted transcription bubble (Cheetham, Jeruzalmi et al. 1999).

The yeast and human mtRNAP, Rpo41 and POLRMT, are both single subunit nuclear-encoded DNA-dependent RNAPs with extensive sequence homology (26% sequence similarity) to T7 bacteriophage RNAP. The classic features of T7 RNAP present in the promoter binding domain (PBD) like AT-rich recognition loop, intercalating β-hairpin, specificity loop are all preserved in mtRNAP. However, the structural similarity between the two remains restricted to the C-terminal domain (CTD). The most striking difference between T7 RNAP and mtRNAPs is that the latter requires the assistance of accessory factors in performing the various steps of transcription. For instance, mtRNAPs alone are incapable of initiating promoter-specific transcription on duplex promoters. This need for additional factors is consistent with the prokaryotic origins of the mitochondria.

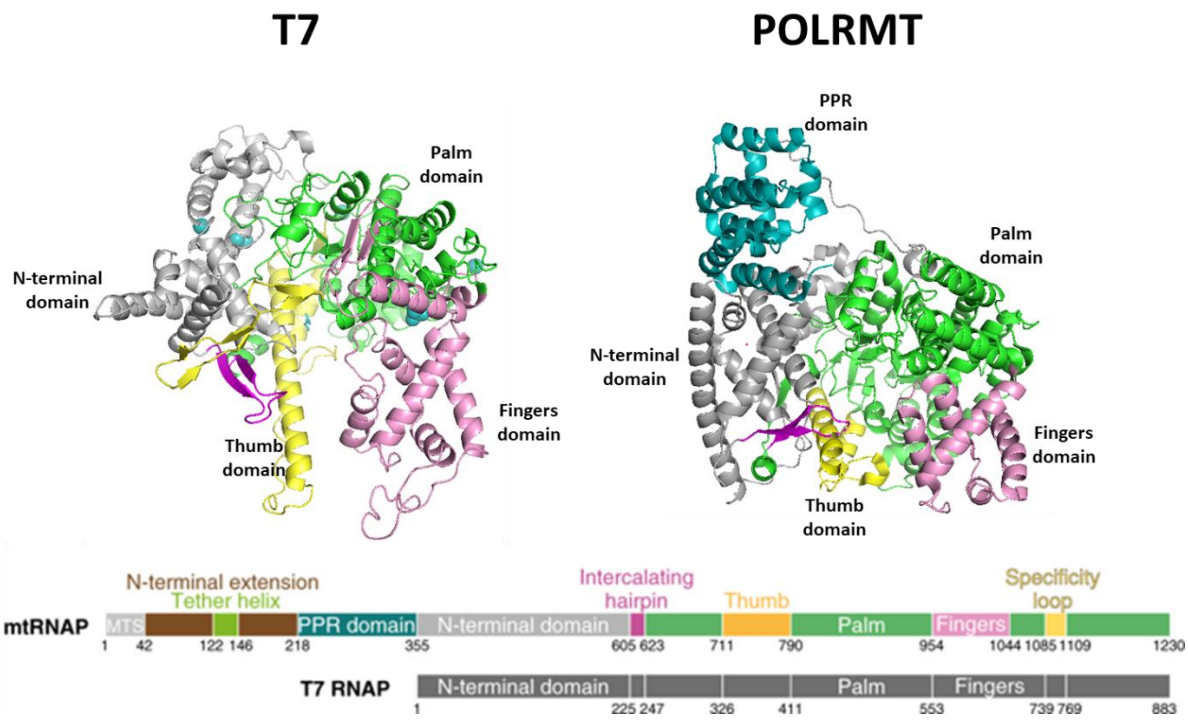


Figure A.3: Comparing and contrasting the structure of T7 RNAP and human mitochondrial RNA polymerase (mtRNAP). The top panel shows ribbon depiction of T7 (PDB 1CEZ) and human mitochondrial RNA polymerase (PDB 3SPA) with domains colored as follows: PPR, teal; NTD: gray; palm: green; fingers: pink; intercalating hairpin: magenta, thumb: yellow. Important structural elements are indicated. The bottom panel shows a schematic depiction of mtRNAP and T7 RNAP domains and important structural elements. Structurally homologous regions between mtRNAP and T7 RNAP are aligned.

In addition to the CTD, POLRMT also has N-terminal domain (NTD) and an N-terminal extension domain (NTE). The NTE is involved in promoter specificity (Posse, Hoberg et al. 2014). The NTD contains all the elements for promoter recognition like AT-rich recognition loop and for promoter melting like the intercalating β-hairpin as seen in T7 RNAP. However, these elements have diverged extensively from T7 RNAP and require the assistance of other accessory factors involved in initiation like TFAM and TFB2M to position them correctly to carry out their functions (Ringel, Sologub et al. 2011). Crystal structure of POLRMT has brought forward the similarities of the catalytic domain with T7 RNAP. Although the NTD has similarities to T7, the AT-rich recognition loop differs substantially. In T7 RNAP, this loop interacts with the -15 to -17 promoter region in the minor groove (Cheetham, Jeruzalmi et al. 1999). However, the AT-rich loop is short, and intraprotein interactions sequester this loop in human POLRMT (Ringel, Sologub et al. 2011). mtRNAP also contains an additional N-terminal pentatricopeptide repeat (PPR) domain and a largely flexible N-terminal extension, which is not resolved in the apo-mtRNAP structure. In the initiation structure (IC), fingers, palm and thumb domains of the CTD adopt a canonical right-hand fold, but the fingers domain appears rotated by roughly 25° approximately around the axis of the O helix, accompanied by a tilting of the O and Y helices into a 'clenched' conformation (Ringel, Sologub et al. 2011). These movements cause a 15 Å translation of the N-terminal part of the Y helix towards the promoter-binding domain (PBD). We speculate that the addition of nucleotides might change to a more 'active' conformation where the active site is more accessible. Another difference between T7 RNAP and POLRMT is that the hairpin connects the conserved helices I and J and has a 14-residue insertion called 'B-loop' or

lever loop that is partly visible in both in the mtRNAP and IC structure. This loop is used to interact with TFB2M as evident from the human mt IC.

Human mt transcription factor A (TFAM)

TFAM, the first transcription factor identified in the human mitochondrial system, belongs to the high mobility group (HMG) family of DNA-binding proteins. Unlike its homolog in yeast Abf2 which has only DNA packaging functions, TFAM is involved in transcription initiation in addition to DNA packaging. TFAM has two HMG boxes separated by a linker region. TFAM can bend and wrap DNA via its characteristic HMG-box DNA binding domains (Fisher, Lisowsky et al. 1992). It also has a C-terminal tail that is missing in Abf2. Deleting this C-tail from TFAM detracts its ability to act as a transcriptional activator (Dairaghi, Shadel et al. 1995). TFAM binds in a sequence-specific manner to certain regions upstream of the promoter and severely bends the DNA by 180º when acting as a transcriptional activator (Fisher, Lisowsky et al. 1992, Dairaghi, Shadel et al. 1995).

It has also been shown that POLRMT and TFAM interact with each other via the tether helix of POLRMT (Morozov, Agaronyan et al. 2014, Morozov, Parshin et al. 2015, Hillen, Morozov et al. 2017). TFAM interaction has been proposed to recruit POLRMT to the promoter initiation site. TFAM binds to the DNA -10 and -35 upstream region from the start site in a sequence-specific manner (Fisher, Topper et al. 1987, Gaspari, Falkenberg et al. 2004). Two independent studies have reported the crystal structure of TFAM in complex with LSP (Ngo, Kaiser et al. 2011, Rubio-Cosials, Sidow et al. 2011). These structures show that the HMG box domains induce a 90° turn in the DNA which

points to how TFAM works to compact mt genome into higher order nucleoids (Alam, Kanki et al. 2003, Kanki, Ohgaki et al. 2004, Kaufman, Durisic et al. 2007, Kukat, Davies et al. 2015). It was also suggested that TFAM binds to HSP1 in reverse orientation as compared to LSP suggesting distinct ICs are formed on the two promoters (Ngo, Lovely et al. 2014). However, only partial sequence of the TFAM binding site was used, and this model is not consistent with various biochemical and cross-linking studies on this subject which point to the similar topology of IC on the two promoters (Morozov and Temiakov 2016).

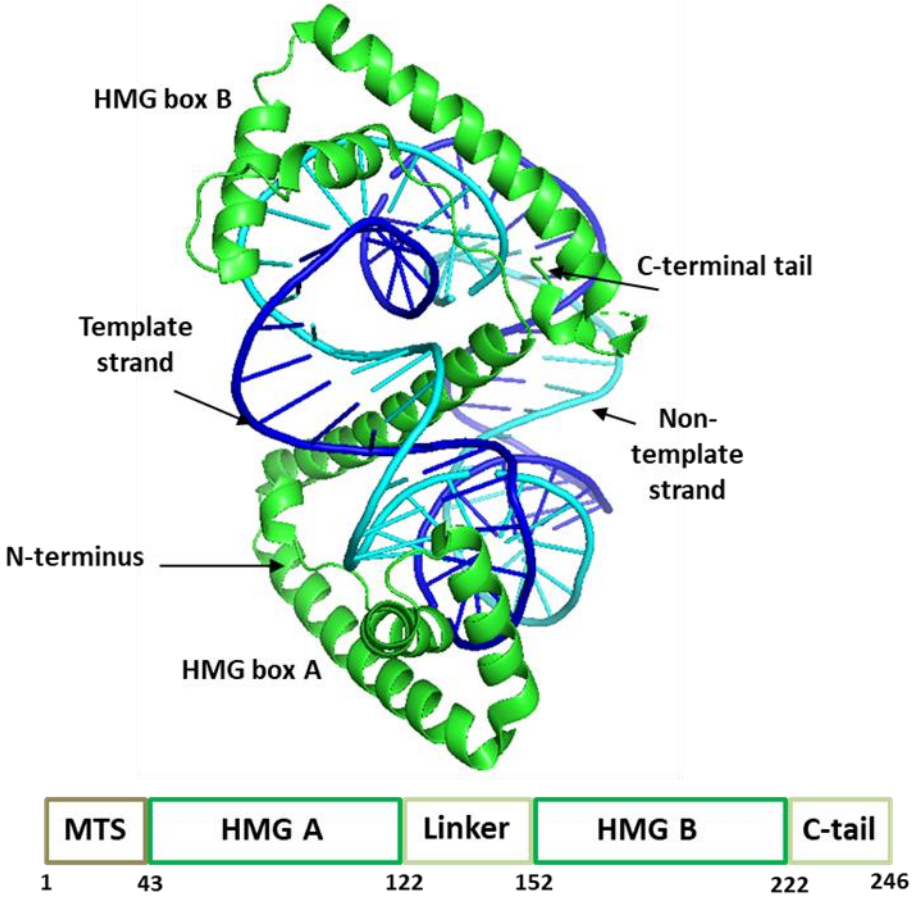


Figure A.4: Structure of human mitochondrial transcription factor TFAM. (Upper panel) Ribbon depiction of human TFAM (PDB 3TQ6) shown in green. The template and

non-template strands are shown in dark blue and cyan respectively. Only one TFAM molecule is shown for simplicity. (Lower panel) Schematic representation of the structural domains of TFAM is shown.

Yeast mt transcription factor Mtf1 & Human transcription factor 2 (TFB2M)

Unlike T7 RNAP, mtRNAPs require one or more transcription factors to catalyze transcription initiation. How and why these transcription factors are required are largely unanswered questions. These factor(s) aid the RNAP in the preinitiation steps of promoter binding, bending, melting, initiating NTPs (iNTPs) binding, abortive RNA synthesis, etc. For the yeast mt system, only one transcription factor Mtf1 is required. Mtf1 helps Rpo41 in binding to promoter DNA, melting of the promoter duplex, and stabilizing the open complex. For the human mt system, two transcription factors TFAM and TFB2M are required. In this study, we have focused on the functions of the C-tail Mtf1and TFB2M.

In 2002, POLRMT, TFAM and a second transcription factor TFB2M were reported to constitute the basal mammalian mitochondrial transcription machinery, and the system was reconstituted *in vitro* (Falkenberg, Gaspari et al. 2002). It was believed that, like its yeast homolog Mtf1, TFB2M is structurally related to bacterial transcription factor sigma (Jang and Jaehning 1991, Cliften, Jang et al. 2000). However, the crystal structure of Mtf1 and the subsequent crystal structure of TFB2M revealed that both these proteins resemble bacterial rRNA methyltransferases (Schubot, Chen et al. 2001, Hillen, Morozov et al. 2017). This raises interesting questions on the evolutionary origins of the mt transcription factors.

Mtf1 and TFB2M are homologous transcription factors in the yeast and human mt systems respectively having 12% sequence identity. Comparing and contrasting the crystal structures of TFB2M and Mtf1 reveal conserved structural features that are different from yeast to humans.

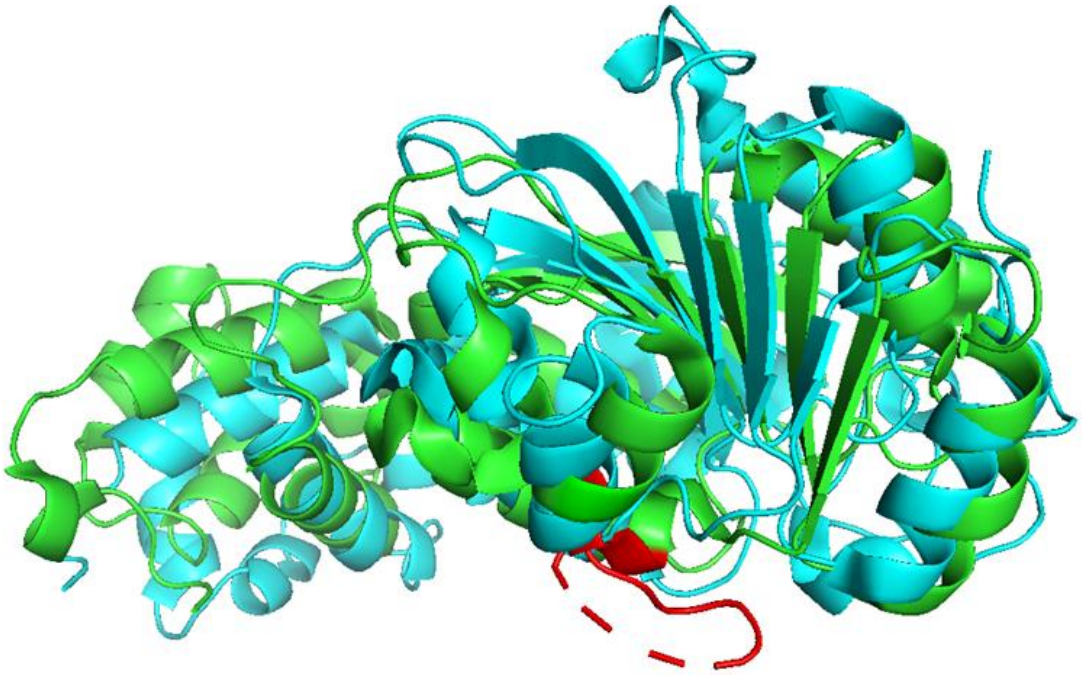


Figure A.5: Structure alignment of yeast Mtf1 and human TFB2M. Ribbon depiction of human TFB2M (PDB 6ERO) in blue, with the structure of the yeast mitochondrial transcription initiation factor Mtf1 (PDB 1I4W) overlaid in green. The extra loop in TFB2M is marked in red.

The N-terminal domain (NTD) assumes a fold similar to S-adenosyl-methionine-dependent methyltransferases with a central seven-stranded β sheet flanked on either side

by three α helices (Schubot, Chen et al. 2001, Martin and McMillan 2002, Hillen, Morozov et al. 2017). A distinct difference between Mtf1 and TFB2M is that the latter has a prominent loop insertion (shown in red in Fig. D4) of unknown function between β3 and α4 regions which may contribute to enhanced domain movements (Moustafa, Uchida et al. 2015). The C-terminal domain of both these homologs consists of four α helices and an extended C-terminal tail. Density for the C-tail was only observed in one of the two copies in the asymmetric unit in the TFB2M structure while the other unit lacks six amino acids (Hillen, Morozov et al. 2017). Sixteen amino acids of the C-tail are missing from the crystal structure of Mtf1 (Schubot, Chen et al. 2001). The lack of the C-tail in the crystal structure indicates the flexibility/ disordered nature of the C-tail in solution. Disordered regions are often crucial for transcription activators (Dyson and Wright 2005).

Similar to methyltransferases, N-terminal domains of both Mtf1 and TFB2M have S-adenosyl-L-methionine (SAM) binding site. But SAM binding is not essential for transcription activity (Schubot, Chen et al. 2001, Cotney and Shadel 2006). TFB2M has been shown to retain some whereas Mtf1 retains none of its methyltransferase activity (McCulloch, Seidel-Rogol et al. 2002).

There is a consensus in the field that the mtRNAP itself is involved in promoter recognition while promoter melting is performed by Mtf1/TFB2M. Mtf1/TFB2M makes specific contacts with the mtRNAP and template DNA to assist the mtRNAP in promoter melting (Gaspari, Falkenberg et al. 2004, Gaspari, Larsson et al. 2004, Sologub, Litonin et al. 2009, Tang, Paratkar et al. 2009, Ringel, Sologub et al. 2011, Hillen, Morozov et al.

2017, Ramachandran, Basu et al. 2017). These proteins also aid in template alignment as shown by our studies (Basu & Patel, unpublished data).

Interactions between Rpo41 and Mtf1

Rpo41 and Mtf1 forms a 1:1 complex both in presence and absence of DNA (Mangus, Jang et al. 1994, Tang, Paratkar et al. 2009, Paratkar, Deshpande et al. 2011). Multiple regions in Mtf1 were shown to interact with Rpo41. The C-tail of Mtf1 was also shown to be involved in interaction with Rpo41 (Mangus, Jang et al. 1994). Two-hybrid assays identified three regions of Mtf1 that interact with Rpo41: Cluster A (Y42C, H44P, L53H), cluster B (V135A, I154T, K157E) and cluster C (S218R, I221K, and D225G). Mapping them on the crystal structure of Mtf1 shows their distribution on different faces of the protein. Mutational studies also indicate that multiple regions of Rpo41 protein are involved in protein-protein interaction with Mtf1. Deletion studies of Rpo41 show that the amino acids in the region 270-380 may mediate interactions with Mtf1 (Paratkar, Deshpande et al. 2011). A genetic suppressor screen using a mutant Mtf1 (V135A) which does not interact with Rpo41, identified three Rpo41 mutants (K1273A, A631V, E1124K) that restored interactions (Cliften, Jang et al. 2000). These mutations were present in the promoter binding elements raises the possibility that this region is responsible for Rpo41-Mtf1 interactions. The A631 lies within the intercalating hairpin loop, E1124 in the promoter specificity loop, and K1273 in the extended foot. It will be interesting to determine whether Mtf1 interacts directly with these promoter binding elements or the mutations reorganize specific regions of Rpo41 that interact with Mtf1.

Interactions between POLRMT and TFB2M

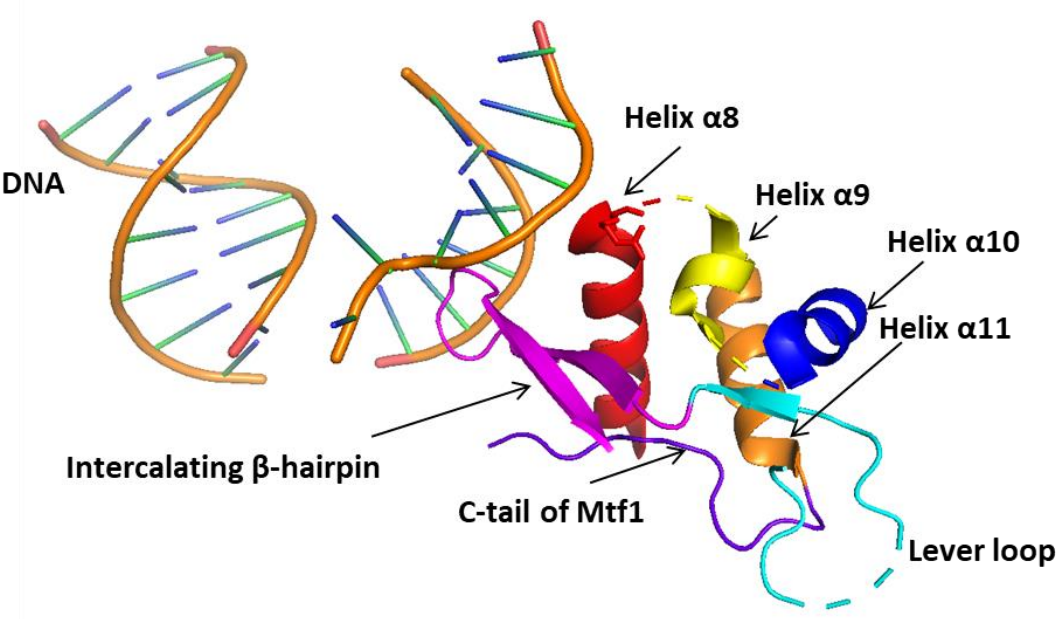


Figure A.6: Interactions of POLRMT and TFB2M. PyMOL image showing the LSP promoter around the transcription start site. The C-tail of Mtf1 is shown in purple, the intercalating β-hairpin of POLRMT is shown in magenta, the lever loop is shown in cyan, the helices α8, α9, α10, α11 are shown in red, yellow, blue and orange respectively.

Cross-linking studies have revealed various interactions between POLRMT and TFB2M which were subsequently confirmed by the crystal structure of the human mt IC. Specific cross-links were detected with the B-loop/lever loop (residues 588-604) and TFB2M. The α-8 helix of TFB2M (residues 315-352) was identified as the region that interacts with lever loop of POLRMT. All these proposed interactions were confirmed by the crystal

structure of the human mt IC (Sologub, Litonin et al. 2009, Morozov, Agaronyan et al. 2014, Morozov, Parshin et al. 2015, Hillen, Morozov et al. 2017). There are additional contacts of the N-terminal region of TFB2M with the specificity loop of POLRMT (Morozov, Agaronyan et al. 2014, Morozov, Parshin et al. 2015).

Interactions between TFB2M and promoter DNA

Cross-links were detected with the -5T position with the N-terminal region of TFB2M. Thus, this region might be functionally important as it is inserted into the RNA–DNA hybrid binding cavity of mtRNAP. Strong TFB2M cross-links were also detected in the +1 to +3 T region of the promoter DNA. Two adjacent regions of TFB2M interact with the DNA around the start site: N-terminus between residues 1 and 42 and another between residues 43 and 59. TFB2M N-terminal region (residues 24-41) was also found to be cross-linked to the priming substrates (Sologub, Litonin et al. 2009, Morozov, Parshin et al. 2015).

Interactions between TFAM and TFB2M

It had been proposed that TFB2M interacts with TFAM via the C-terminal region of the latter (McCulloch and Shadel 2003). This was demonstrated using solid phase protein-protein interaction binding studies. However, later cross-linking studies failed to detect any interaction between TFB2M and TFAM (Morozov, Agaronyan et al. 2014, Morozov, Parshin et al. 2015). This was later confirmed by the crystal structure of the human mt IC which did not detect ant contacts between TFB2M and TFAM (Hillen, Morozov et al. 2017).

[B] SPECIFIC AIMS

The overarching theme of my dissertation is to understand the mechanism of transcription initiation by the yeast and human mtRNAP. Although there has been a recent momentum in research on human mitochondrial biology, we are severely limited in our knowledge of the mechanistic and kinetic aspects of transcription initiation. Before my dissertation project began, we were focusing our studies on the mechanism of T7 bacteriophage and yeast mitochondrial transcription systems. Although the yeast mitochondrial system provides a simplistic platform to understand the more complex human system, I was interested in studying the human system. More importantly, I wanted to compare and contrast the nuances of the regulation of yeast and human transcription initiation. Hence, I initiated studies in the human mitochondrial system. I took a leading role in gearing the lab towards setting up the assays for the human mitochondrial system. Both sections of my dissertation have a comparative view of the yeast and the human mitochondrial systems. The major challenge in studying the human mt initiation was to obtain pure proteins. I had to overcome major hurdles by successfully purifying milligram quantities of recombinant POLRMT, TFB2M, and TFAM, and show that all are active in transcription.

Since the identification of the key components of the human mitochondrial transcription machinery, the last decade has witnessed great strides in research aimed at understanding the mechanisms of transcription initiation. However, many fundamental questions regarding the mechanisms of the key components in the initiation complex remained unanswered. Comprehensive studies aimed at understanding the intricate workings of the proteins in various steps of transcription in my thesis research provided detailed

knowledge of two related transcription initiation factors of the yeast and human mitochondrial transcription systems.

In the first section, I have studied mitochondrial NAD^+ capping by yeast and human mt RNAPs. It has been shown previously that both bacterial and eukaryotic nuclear RNAPs can perform 5' RNA capping by utilizing a plethora of metabolic intermediates called non-canonical initiating nucleotides (NCINs) with structural similarity to ATP. The presence of NAD^+ capped RNA was also detected in the mitochondria. In this section, I have investigated yeast and human mitochondrial RNA capping using the oxidized and reduced forms of the metabolic intermediate nicotinamide adenine dinucleotide NAD^+ and NADH.

Aim 1: To demonstrate NAD^+ capping in mitochondria both *in vitro* and *in vivo*

Aim 1.1: To demonstrate that the yeast and human mtRNAPs can cap RNA with NAD^+ and NADH *in vitro*

Aim 1.2: To determine the efficiency of RNA capping with NAD^+ and NADH by yeast and human mtRNAPs relative to multi-subunit RNAPs

Aim 1.3: To determine the promoter sequence elements that dictate efficiency of mtRNAP capping with NAD^+ and NADH

Aim 1.4: To detect and quantitate NAD^+ and NADH-capped RNA in the yeast and human mitochondria *in vivo*

In the second section, I have studied the role of the C-terminal regions of the yeast mt transcription factor Mtf1 and its human homolog TFB2M in the regulation of transcription initiation. The C-tail of Mtf1 was not resolved in its crystal structure;

however, one of the two molecules in the crystal structure of TFB2M has a resolution for all C-tail residues while the other molecule lacks six amino acids. Previous studies had shown that the C-tail of Mtf1 cross-linked to -3/-4 template strand and that the C-tail of Mtf1 is important for promoter melting (Savkina, Temiakov et al. 2010, Drakulic, Wang et al. 2014). A recent structure of the human mt transcription initiation complex shows that the C-tail of TFB2M is in the vicinity of the intercalating beta hairpin which is involved in open complex formation (Hillen, Morozov et al. 2017). In this section, I have investigated the precise role C-tail of yeast Mtf1 and its human homolog TFB2M in transcription initiation.

Aim 2: To define the roles of the C-terminal regions of yeast mt transcription factor Mtf1 and human mt transcription factor TFB2M in mt transcription initiation

Aim 2.1: To determine the role of the C-tail of Mtf1 and TFB2M in promoter melting

Aim 2.2: To determine the role of the C-tail of Mtf1 and TFB2M in template alignment

Aim 2.3: To determine the role of the C-tail of Mtf1 and TFB2M in catalytic efficiency of transcription initiation

Aim 2.2: To determine the role of the C-tail of Mtf1 and TFB2M in the abortive synthesis

[C] MATERIAL AND METHODS

Protein purifications

Purification of Rpo41 (yeast mtRNAP)

S. cerevisiae mtRNAP was prepared from *E. coli* transformed with pJJ1399 (gift of Judith A. Jaehning). In brief, cells were grown in BL21 codon plus (RIL) cells at 37°C until O.D. reached 0.6-0.8 when they were induced by the addition of 1 mM Isopropyl β-D-1-thiogalactopyranoside (IPTG). Induction was carried out at 16°C for 16 hours. Cells were lysed in the presence of protease inhibitor and lysozyme, followed by polyethyleneimine (PEI) treatment and ammonium sulfate precipitation. This was followed by column chromatographic techniques like Ni-sepharose, DEAE sepharose and Heparin-sepharose chromatography (Tang, Paratkar et al. 2009, Tang, Deshpande et al. 2011).

Cloning, expression, and purification of Mtf1 C-tail deletion mutants

4, 12, 16 and 20 amino acids were systematically deleted from the carboxyl-terminal end of the Mtf1 protein, and the mutants were accordingly named Mtf1-Δ4, Mtf1-Δ12, Mtf1-Δ16, and Mtf1-Δ20. The C-tail deletions were created using Quikchange Site-Directed Mutagenesis Kit from Stratagene. The forward and reverse primers used to create the mutants are listed in Table 1 in the Appendix section. Plasmids of the mutant strains were isolated using QIAGEN mini-prep plasmid purification kit from XL10-Gold ultracompetent cells. The purified plasmids containing each of the mutants were then fully sequenced from Genscript to confirm deletions of respective amino acids. Further confirmation regarding the mutants were obtained from molecular weight determination

of Mtf1-WT and mutants using MALDI-TOF Mass Spectrometry (data not shown) performed at the Center for Integrative Proteomics Research at Rutgers University. Mtf1 was prepared from BL21 codon plus (RIL) cells transformed with pTrcHisC-Mtf1. In brief, cells were grown in at 37°C until O.D. reached 0.6-0.8 when they were induced by the addition of 1 mM Isopropyl β-D-1-thiogalactopyranoside (IPTG). Induction was carried out at 16°C for 16 hours. Cells were lysed in the presence of protease inhibitor and lysozyme, followed by polyethyleneimine treatment and ammonium sulfate precipitation. This was followed by column chromatographic techniques like Ni-sepharose, DEAE sepharose and Heparin-sepharose chromatography as in (Tang, Paratkar et al. 2009, Tang, Deshpande et al. 2011).

Purification of POLRMT (human mtRNAP)

Human mtRNAP was cloned in pPROEXHTb vector by Genscript. POLRMT was prepared from BL21 codon plus (RIL) cells transformed with pPROEXHTb-POLRMT (43-1230)-6xHis. In brief, cells were grown in at 37°C until O.D. reached 0.6-0.8 when they were induced by the addition of 1 mM Isopropyl β-D-1-thiogalactopyranoside (IPTG). Induction was carried out at 16°C for 16 hours. Cells were lysed in the presence of protease inhibitor and lysozyme, followed by polyethyleneimine treatment and ammonium sulfate precipitation. This was followed by column chromatographic techniques like Ni-sepharose, DEAE sepharose and Heparin-sepharose chromatography (Ramachandran, Basu et al. 2017, Bird, Basu et al. 2018).

Purification of TFAM

TFAM was prepared from BL21 codon plus (RIL) cells transformed with pPROEXHTb-TFAM (43-245)-6xHis. In brief, cells were grown in at 37°C until O.D. reached 0.6-0.8

when they were induced by the addition of 1 mM Isopropyl β-D-1-thiogalactopyranoside (IPTG). Induction was carried out at 16°C for 16 hours. Cells were lysed in the presence of protease inhibitor and lysozyme, followed by polyethyleneimine treatment and ammonium sulfate precipitation. This was followed by column chromatographic techniques like Ni-sepharose, DEAE sepharose and Heparin-sepharose chromatography (Ramachandran, Basu et al. 2017, Bird, Basu et al. 2018).

Cloning, expression, and purification of TFB2M C-tail deletion mutants

3, 7, 13 and 17 amino acids were systematically deleted from the carboxyl-terminal end of the TFB2M protein, and the mutants were accordingly named TFB2M-Δ3, TFB2M-Δ7, TFB2M-Δ13, and TFB2M-Δ17. The C-tail deletions were created using Quikchange Site-Directed Mutagenesis Kit from Stratagene. The forward and reverse primers used to create the mutants are listed in Table 1 in the Appendix section. Plasmids of the mutant strains were isolated using QIAGEN mini-prep plasmid purification kit from XL10-Gold ultracompetent cells. The purified mutant plasmids were then fully sequenced from Genscript to confirm deletions of respective amino acids.

Human TFB2M was purified from *E. coli* strain ArticExpress (DE3) (Stratagene) transformed with pT7TEV-HMBP4 (gift of Professor Miguel Garcia-Diaz at the Stony Brook University School of Medicine). Cells were successively grown at 37°C followed by 27°C. Cells were induced by addition of 0.2 mM Isopropyl β-D-1-thiogalactopyranoside (IPTG) and grown overnight at 14°C. Cells were lysed in the presence of protease inhibitor and lysozyme and loaded on Ni-sepharose, heparin-sepharose chromatography, HiTrap SP HP cation exchange column followed by size

exclusion chromatography (Yakubovskaya, Guja et al. 2014, Ramachandran, Basu et al. 2017, Bird, Basu et al. 2018).

The molar concentrations of the proteins were determined from absorbance measurements at 280 nm using a guanidium-HCl buffer and using the respective molar extinction coefficients of the proteins which are compiled in Table 2 in the Appendix section.

Oligonucleotides

All oligonucleotides were custom synthesized from IDT (Integrated DNA Technologies) and were HPLC purified. The molar concentration of purified DNA was determined from its absorbance at 260 nm and the respective extinction coefficients. For transcription assays, the double-stranded DNAs were prepared by annealing complementary non-template and template strands in 1:1 ratio. For fluorescence assays, annealing was carried out by keeping the unlabeled strand in 1.1 fold excess over the fluorescently labeled complementary strand. Sequences of the oligonucleotides used for each experiment are listed with the respective figures.

Promoter melting assays

Promoter melting properties of Mtf1 and TFB2M wild-type and mutants were assayed by conducting steady-state 2-aminopurine (2-AP) fluorescence measurements on a Fluoro-Max-4 spectrofluorometer (Jobin Yvon-Spex Instruments S.A., Inc.). For Mtf1 assays, 200 nM 2-AP incorporated 21S duplex promoter was added at 25°C to the reaction buffer (50 mM Tris-acetate, pH 7.5, 100 mM potassium glutamate and 10 mM magnesium acetate), followed by 400 nM Rpo41 and 400 nM Mtf1-WT or mutants in succession. Increase in fluorescence emission at 380nm (6-nm bandwidth) on these additions were

monitored, when excited at 315nm (2-nm bandwidth). After subtracting the buffer background, the contribution of Mtf1-WT or mutants in promoter melting was depicted as fold change over the basal Rpo41 fluorescence (Tang, Paratkar et al. 2009).

For TFB2M assays, 100 nM 2-AP incorporated duplex LSP promoter was added at 25°C to the reaction buffer (50 mM Tris-acetate, pH 7.5, 50 mM sodium glutamate and 10 mM magnesium acetate, 1 mM DTT), followed by 200 nM POLRMT+TFAM+TFB2M or TFB2M mutants. Fluorescence spectra from 350 to 420 nm (6-nm bandwidth) were collected after excitation at 315 nm (2-nm bandwidth). When used, the concentration of ATP was 500 μM. A parallel experiment was carried out with unmodified LSP DNA to correct for background fluorescence from buffer alone and proteins with unmodified DNA. The 2-AP fluorescence intensities between 360 and 380 nm were integrated and were divided by the fluorescence intensity of free DNA to obtain the fold-increase (Ramachandran, Basu et al. 2017).

Fluorescence assays to test iNTP binding

Binding of +1 and +2 iNTPs was monitored by incorporating 2-AP in the -1T (for yeast) or -1NT (for human) positions. Increase in fluorescence emission at 380nm (6-nm bandwidth) on various additions of was monitored, when excited at 315nm (2-nm bandwidth). 200 nM 2-AP incorporated duplex 21s promoter was added at 25°C in reaction buffer (50 mM Tris-acetate, pH 7.5, 100 mM potassium glutamate and 10 mM magnesium acetate), followed by 400 nM Rpo41 and 400 nM Mtf1-WT or mutants in succession. Then, increasing concentrations of equimolar ATP+GTP were added to determine composite K_d for +1+2 positions. The curves were fitted into a hyperbola equation using Sigmaplot. For human, fluorescence spectra from 360 to 380 nm (6-nm

bandwidth) were collected after excitation at 315 nm (2-nm bandwidth) in samples containing 100 nM 2-AP incorporated duplex LSP promoter, 200 nM each of TFAM, POLRMT and TFB2M and reaction buffer followed by titrating ATP concentrations (Tang, Paratkar et al. 2009, Ramachandran, Basu et al. 2017).

Transcription Assays

Assays performed with *S. cerevisiae* mtRNAP were based on procedures described in (Deshpande and Patel 2014). Assays performed with human mtRNAP were based on procedures described in (Ramachandran, Basu et al. 2017).

Methods for Section [D]: *(Parts of this section have been adapted from Bird* & Basu* et. Al., eLife, 2018)*

For initial RNA product assays in Fig. D.2.1, D.2.2, D.2.6, 1 µM DNA template, 1 µM *S. cerevisiae* mtRNAP, and 1 µM Mtf1 were incubated at 25°C for 10 min in *Sce*-mtRNAP reaction buffer (50 mM Tris-acetate pH 7.5, 100 mM potassium glutamate, 10 mM magnesium acetate, 0.01% Tween-20, 1 mM DTT, and 5% glycerol). A mixture containing the initiating nucleotide (200 µM ATP, 1 mM NAD^+, or 1 mM NADH) and extending nucleotide (10 µM of non-radiolabeled GTP plus 6 mCi of [α^{32}P]-GTP [Perkin Elmer; 3,000 Ci/mmol]) was added, and assays were incubated at 25°C for 30 min.

For assays in Fig. D.2.1, D.2.2 radiolabeled initial products were isolated using a Nanosep 3 kDa cutoff spin concentrator (Pall). For assays in Fig. D.2.6, reactions were stopped with 10 µl RNA loading dye (95% deionized formamide, 18 mM EDTA, 0.25% SDS, xylene cyanol, bromophenol blue, amaranth) and were analyzed by electrophoresis on 7.5 M urea, 1x TBE, 20% polyacrylamide gels (UreaGel System; National

Diagnostics), followed by storage-phosphor imaging (Typhoon 9400 variable-mode imager; GE Life Science).

For the initial RNA product assays in Fig. D.2.1 and D.2.2, 1 µM DNA template, 1 µM human mtRNAP, 1 µM TFAM, and 1 µM TFB2M were incubated at 25°C for 10 min in human-mtRNAP reaction buffer (50 mM Tris-acetate pH 7.5, 50 mM sodium glutamate, 10 mM magnesium acetate, 1 mM DTT, and 0.05% Tween-20). A mixture containing the initiating nucleotide (200 µM ATP, 1 mM NAD^+, or 1 mM NADH) and extending nucleotide (10 µM of non-radiolabeled GTP plus 6 mCi of [α^{32}P]-GTP at [Perkin Elmer; 3,000 Ci/mmol]) was added, and assays were incubated at 25°C for 60 min. Radiolabeled initial RNA products were isolated using a Nanosep 3 kDa cutoff spin concentrator (Pall).

A portion of the recovered initial RNA products was mixed with either 10 U of RppH or 400 nM NudC and incubated at 37°C for 30 min. Reactions were stopped by addition of 10 µl RNA loading dye. Samples were analyzed by electrophoresis on 7.5 M urea, 1x TBE, 20% polyacrylamide gels (UreaGel System; National Diagnostics), followed by storage-phosphor imaging (Typhoon 9400 variable-mode imager; GE Life Science).

For full-length product assays in Fig. D.2.1 and D.2.2 (panel B), 1 µM DNA template, 1 µM *S. cerevisiae* mtRNAP, and 1 µM Mtf1 were incubated at 25°C for 10 min in *Sce*-mtRNAP reaction buffer. A mixture containing the initiating nucleotide (1 mM ATP, 1 mM NAD^+, or 1 mM NADH for Figure D.2.1.C; 200 µM non-radiolabeled ATP plus 10 µCi [γ^{32}P]-ATP [Perkin Elmer; 6,000 Ci/mmol] or 1 mM NAD^+ plus 20 µCi [α^{32}P]-NAD^+ [Perkin Elmer; 800 Ci/mmol] for Fig. D.2.1 and D.2.2 and extending nucleotides (200 µM GTP, 200 µM 3'-deoxy-CTP, 20 µM ATP, 200 µM non-radiolabeled UTP, and

6 mCi of [α^{32}P]-UTP [Perkin Elmer; 3000 Ci/mmol] for Fig. D.2.1; 200 µM GTP, 200 µM 3'-deoxy-CTP, 20 µM ATP, 200 µM UTP for Fig. D.2.1 and D.2.2 was added, and assays were incubated at 25°C for 30 min. Reactions were stopped by addition of stop solution (0.6 M Tris HCl pH 8.0, 18 mM EDTA, 0.1 mg/ml glycogen), samples were extracted with acid phenol:chloroform (5:1, pH 4.5; Thermo Fisher Scientific), and RNA products were recovered by ethanol precipitation and resuspended in NudC reaction buffer (10 mM Tris HCl pH 8.0, 50 mM NaCl, 10 mM $MgCl_2$, 1 mM DTT).

For full-length product assays in Fig. D.2.1 and D.2.2, 1 µM DNA template, 1 µM human mtRNAP, 1 µM TFAM, and 1 µM TFB2M were incubated at 25°C for 10 min in human-mtRNAP reaction buffer. A mixture containing the initiating nucleotide (1 mM ATP, 1 mM NAD^+, or 1 mM NADH for Figure D.2.1.D; 200 µM non-radiolabeled ATP plus 10 µCi [γ^{32}P]-ATP [Perkin Elmer; 6000 Ci/mmol] or 1 mM NAD^+ plus 20 µCi [α^{32}P]-NAD^+ [Perkin Elmer; 800 Ci/mmol] and extending nucleotides (200 µM GTP, 20 µM ATP, 200 µM non-radiolabeled UTP, and 6 mCi of [α^{32}P]-UTP [Perkin Elmer; 3000 Ci/mmol] for Figure D.2.1.D; 200 µM GTP, 20 µM ATP, 200 µM UTP for Fig. D.2.1 and D.2.2 were added, and assays were incubated at 25°C for 60 min. Reactions were stopped by addition of stop solution; samples were extracted with acid phenol: chloroform (5:1) (pH 4.5; Thermo Fisher Scientific), RNA products were recovered by ethanol precipitation and resuspended in NudC reaction buffer.

Full-length RNA products were incubated at 37°C for 30 min with 400 nM NudC alone (Figure D.2.1.C–D and Fig. D.2.1 and D.2.2), 0.25 U FastAP Thermosensitive Alkaline Phosphatase alone (Fig. D.2.1 and D.2.2), or both NudC and FastAP (Fig. D.2.1

and D.2.2). Reactions were stopped by addition of an equal volume of RNA loading dye. Samples were analyzed by electrophoresis on 7.5 M urea, 1x TBE, 20% polyacrylamide gels (UreaGel System; National Diagnostics), followed by storage-phosphor imaging (Typhoon 9400 variable-mode imager; GE Life Science).

Determination of the efficiency of NCIN-mediated initiation vs. ATP-mediated initiation, $(k_{cat}/K_M)_{NCIN}$ / $(k_{cat}/K_M)_{ATP}$ *in vitro*: full-length product assays

For experiments in Fig. D.2.3, 1 µM of template DNA, 1 µM of mtRNAP, and 1 µM transcription factor(s) (Mtf1 for *S. cerevisiae* mtRNAP; TFAM and TFB2M for human mtRNAP) were incubated at 25°C for 10 min in *Sce*-mtRNAP or human reaction buffer. A mixture containing 200 µM ATP, 200 µM UTP, 200 µM non-radiolabeled GTP, and 6 mCi [α^{32}P]-GTP at 3000 Ci/mmol and NCIN (0, 50, 100, 200, 400, 800, 1600, 3200, 6400 µM) was added, and assays were incubated at 25°C for 30 min. Reactions were stopped by addition of an equal volume of RNA loading dye. Samples were analyzed by electrophoresis on 7.5 M urea, 1x TBE, 20% polyacrylamide gels (UreaGel System; National Diagnostics) supplemented with 0.2% 3-acrylamidophenylboronic acid (Boron Molecular), followed by storage-phosphor imaging (Typhoon 9400 variable-mode imager; GE Life Science).

Bands corresponding to uncapped (pppRNA) and NCIN-capped (NCIN-RNA) full-length products were quantified using ImageQuant software. The ratio of NCIN-RNA to total RNA [NCIN-RNA / (pppRNA + NCIN RNA)] was plotted vs. the relative concentrations of NCIN vs. ATP ([NCIN] / [ATP]) on a semi-log plot (SigmaPlot) and non-linear regression was used to fit the data to the equation: $y = (ax) / (b + x)$; where y is [NCIN-

RNA / (pppRNA + NCIN RNA)], x is ([NCIN] / [ATP]), and a and b are regression parameters. The resulting fit yields the value of x for which y = 0.5. The relative efficiency $(k_{cat}/K_M)_{NCIN}$ / $(k_{cat}/K_M)_{ATP}$ is equal to 1/x. Data for determination of relative efficiencies are means of three technical replicates (Bird, Nickels et al. 2017).

Determination of the efficiency of NCIN-mediated initiation vs. ATP-mediated initiation, $(k_{cat}/K_M)_{NCIN}$ / $(k_{cat}/K_M)_{ATP}$, *in vitro*: initial product assays

For experiments in Fig. D.2.6, 1 µM of template DNA, 1 µM of RNAP, and 1 µM transcription factor(s) (Mtf1 for *S. cerevisiae* mtRNAP; TFAM and TFB2M for human mtRNAP; none for T7 RNAP) were incubated at 25°C for 10 min in *Sce*-mtRNAP, human mtRNAP reaction buffer, or T7 RNAP reaction buffer (40 mM Tris HCl pH 7.9, 6 mM $MgCl_2$, 2 mM DTT, 2 mM Spermidine). A mixture containing 1 mM NCIN, ATP (0, 25, 50, 100, 200, 400, 800, 1600 µM), 20 µM non-radiolabeled GTP, and 6 mCi [α^{32}P]-GTP at 3000 Ci/mmol was added, and assays were incubated at 25°C for 30 min. Reactions were stopped by addition of an equal volume of RNA loading dye. Samples were analyzed by electrophoresis on 7.5 M urea, 1x TBE, 20% polyacrylamide gels (UreaGel System; National Diagnostics) supplemented with 0.2% 3-acrylamidophenylboronic acid (Boron Molecular), followed by storage-phosphor imaging (Typhoon 9400 variable-mode imager; GE Life Science).

For experiments in Fig. D.2.5, 200 nM of tailed template and 500 nM RNAP (*S. cerevisiae* mtRNAP, human mtRNAP, *S. cerevisiae* RNA II, T7 RNAP, or *E. coli* RNAP) were incubated at 25°C for 15 min in reaction buffer containing 10 mM Tris pH 8.0, 50 mM potassium glutamate, 10 mM $MgCl_2$, 2 mM DTT, and 50 ug/ml BSA. A

mixture containing NCIN (1 mM NCIN for mtRNAPs, T7 RNAP, and *E. coli* RNAP; 4 mM for *S. cerevisiae* RNAP II), ATP (0, 25, 50, 100, 200, 400, 800, 1600 µM for mtRNAPs and *S. cerevisiae* RNAP II; 0, 6.25, 12.5, 25, 50, 100, 200, 400 µM for *E. coli* RNAP), 10 µM non-radiolabeled CTP, and 6 mCi [α^{32}P]-CTP (Perkin Elmer; 3000 Ci/mmol) was added, and assays were incubated at 25°C for 1 hr. Reactions were stopped by addition of an equal volume of RNA loading dye. Samples were analyzed by electrophoresis on 7.5 M urea, 1x TBE, 20% polyacrylamide gels (UreaGel System; National Diagnostics), followed by storage-phosphor imaging (Typhoon 9400 variable-mode imager; GE Life Science).

Bands corresponding to uncapped (pppApC) and NCIN-capped (NCINpC) initial RNA products were quantified using ImageQuant software. The ratio of NCINpC to total RNA (NCINpC / [pppApC + NCINpC]) was plotted vs. the relative concentrations of NCIN vs. ATP ([NCIN] / [ATP]) on a semi-log plot (SigmaPlot) and non-linear regression was used to fit the data to the equation: $y = (ax) / (b + x)$; where y is [NCINpC / (pppApC + NCINpC)], x is ([NCIN] / [ATP]), and a and b are regression parameters. The resulting fit yields the value of x for which $y = 0.5$. The relative efficiency $(k_{cat}/K_M)_{NCIN}$ / $(k_{cat}/K_M)_{ATP}$ is equal to 1/x. Data for determination of relative efficiencies are means of three technical replicates (Bird, Zhang et al. 2016, Bird, Nickels et al. 2017).

Methods for Section [E]:

Catalytic initiation efficiency assays

The catalytic initiation efficiencies of Mtf1-WT and deletion mutants were determined by measuring the rates of 2-mer RNA synthesis as a function of increasing equimolar concentrations of +1ATP at constant +2GTP (for the +1 position) and +2GTP at constant +1ATP (for the +2 position). In these transcription assays, 1 µM of Rpo41 and 2 µM of Mtf1-WT or deletion mutant was incubated with 2 µM of U20D45 promoter DNA at 25°C, and steady-state concentration of 2-mer RNA was measured by titrating 0-4000 µM of ATP at constant 2 mM +2GTP (+ [α-^{32}P] GTP) for determining the kinetics of the +1 position. Similar kinetics for the +2 position were determined by titrating 0-4000 µM of GTP at constant 2 mM +1 ATP (+ [γ-^{32}P] ATP). The 2-mer RNA was resolved from the free NTP on a 24% sequencing gel containing 4 M urea and quantified using ImageQuant. The rates of 2-mer RNA synthesis were calculated according to equation 1.

$$\text{Rate of 2mer RNA synthesis } (\mu M/s) = \frac{R(2)}{R(2) + A} . \frac{[\text{ATP or GTP}]\ (\mu M)}{t\ (s)} \qquad \text{Eqn. 1}$$

Where R(2) and A are the band intensities of 2-mer RNA and free ATP (or GTP), respectively, [ATP or GTP] is the molar concentration of ATP or GTP added to the reaction, and t is the time of reaction. The rates of 2-mer RNA synthesis were plotted as a function of increasing concentration of +1 or +2 NTPs and fit to the Michaelis-Menten equation to obtain the K_m of the individual +1 and +2 NTPs and the maximum rate of 2-mer formation. The catalytic constant k_{cat} was the same as the maximum rate of 2-mer formation, since Rpo41 was limiting at 1 µM in the Rpo41-Mtf1-DNA complex.

For initiation with dinucleotide primers, 1 µM of Rpo41 and 2 µM of Mtf1-WT or deletion mutant was incubated with 2 µM of 21s U25D20 promoter (45bp) DNA at 25°C, and steady-state concentration of 3-mer RNA was measured by titrating 5-4000 µM of GTP in the presence of 150 µM of -1+1 UpA or CpA primer (spiked with ^{32}P-labeled

UpA or CpA primer) over 5mins reaction time. The 3-mer RNA (-1+1+2 UpApA or CpApA) rates were calculated according to equation 3.

$$\text{Rate of 3mer RNA synthesis } (\mu M/s) = \frac{R(3)}{R(3) + UpA} . \frac{[UpA]\ (\mu M)}{t\ (s)} \qquad \text{Eqn. 3}$$

Where R(3) and UpA are the band intensities of 3-mer RNA and free UpA, respectively, [UpA] is the molar concentration of UpA primer added to the reaction. The rates of 3-mer RNA synthesis were plotted as a function of increasing concentration of +2 ATP and fit to the Michaelis-Menten equation to obtain the K_m of +2 ATP and the catalytic rate of 3-mer (-1+1+2 UpApA) formation.

Run-off RNA synthesis assays

Transcription activities of Mtf1-WT and various deletion mutants were determined using the gel-based RNA synthesis assay. Briefly, a complex of 1μM Rpo41 and 2 μM of Mtf1-WT or mutants along with 2μM promoter DNA fragments was incubated at 25°C in the reaction buffer (50 mM Tris-acetate, pH 7.5, 100 mM potassium glutamate, 10 mM magnesium acetate, 0.01% protein-grade Tween 20, 1 mM DTT and 5% glycerol) and the transcription reaction was initiated by adding a mixture of 500μM ATP, UTP, GTP, and 3' dCTP each, spiked with [γ-^{32}P]ATP. For transcription using dinucleotide primers, reaction was initiated by adding a mixture of 150μM of dinucleotide primers, 50μM ATP, and 150μM UTP, GTP, and 3' dCTP each, spiked with ^{32}P-labeled dinucleotide primer.

At particular time intervals, reactions were stopped using 400 mM EDTA and formamide dye (98% formamide, 0.025% bromphenol blue, 10 mM EDTA). Samples were heated to 95°C for 2 min and then chilled on ice, and the RNA products were resolved on 24% sequencing gel containing 4 M urea. The gel was exposed to a phosphor

screen overnight and scanned on a Typhoon 9410 PhosphorImager instrument (Amersham Biosciences). The free ATP and RNA bands were quantified using ImageQuant, and molar amounts of RNA synthesized were calculated according to equation 1.

$$\text{RNA synthesized } (\mu M) = \frac{R}{R + A} . [ATP] \ (\mu M) \qquad \text{Eqn. 1}$$

Where R and A are the band intensities of RNA products and free ATP respectively, and [ATP] is the molar concentration of ATP added to the reaction. The rates of RNA synthesis were expressed in min^{-1} and were obtained by linearly fitting the molar amounts of RNA as a function of time and dividing the slope (μM/min) by the limiting concentration of Rpo41 in the enzyme-DNA complex (1 μM).

Transcript slippage assay

1 μM Rpo41 and 2 μM of Mtf1-WT or Mtf1-Δ20 were preincubated with 2 μM of mutant 15S rRNA promoter duplex having A in -1 non-template position (Table S6). RNA synthesis was initiated by adding 250 μM ATP + [γ-^{32}P] ATP at 25°C for 30 min. Reactions were stopped and ran on sequencing gels as before.

[D] The NAD^+ capping story: Mitochondrial 5'-RNA capping with NAD^+ and NADH by yeast and human mitochondrial RNA polymerase

(Parts of this section have been adapted from Bird & Basu* et al., eLife, 2018)*

NOTE: This paper is the result of collaborative efforts of Dr. Bryce E. Nickels, Dr. Richard H. Ebright, Dr. Smita S. Patel's lab. Hence this section requires acknowledgment of all members involved. Author contributions are listed below.

Jeremy G Bird, David Kuster, Formal analysis, Investigation, Visualization, Methodology, Writing— review and editing; Urmimala Basu, Formal analysis, Funding acquisition, Investigation, Visualization, Methodology, Writing—review and editing; Aparna Ramachandran, Investigation, Methodology, Writing—review and editing; Ewa Grudzien-Nogalska, Atif Towheed, Investigation; Douglas C Wallace, Supervision, Funding acquisition; Megerditch Kiledjian, Supervision, Funding acquisition, Writing—review and editing; Dmitry Temiakov, Conceptualization, Funding acquisition, Methodology, Writing—review and editing; Smita S Patel, Conceptualization, Supervision, Funding acquisition, Visualization, Project administration, Writing—review and editing; Richard H Ebright, Bryce E Nickels, Conceptualization, Supervision, Funding acquisition, Visualization, Writing—original draft, Project administration (Bird, Basu et al. 2018).

D.1 Introduction

Decades of research has brought to light a wealth of information regarding the post-transcriptional modifications of mRNA. Different chemical moieties can be added to mRNA that brings in complex layers of regulation of gene expression. With the discovery of new techniques triggered by technological advances in the form of next-generation sequencing, mass spectrometry, etc., there is renewed interest and deeper understanding of the implications of these modifications. In this section, we focus on 5' capping by the metabolic effector nicotinamide adenine dinucleotide (NAD^+) and its reduced form NADH (Wang and He 2014, Nachtergaele and He 2017, Zhao, Roundtree et al. 2017)..

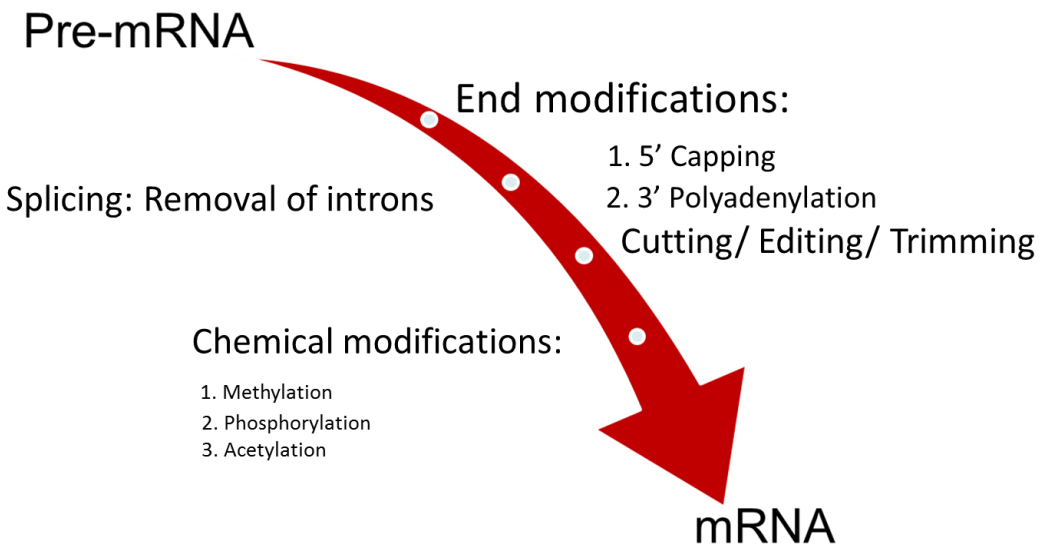


Figure D.1.1: Basic scheme showing major modifications that a pre-mRNA undergoes during its journey toward becoming an mRNA.

There are three main kinds of modifications of mRNA: (i) End modifications like capping and polyadenylation, (ii) Splicing, i.e. removal of exons and joining of introns, (iii) chemical modifications like methylation, phosphorylation, etc. (Fig. D.1.1.). The phenomenon of 5' G-capping was first reported in the 1970s. Seminal work by over the next few decades established the paradigm of m^7G mRNA capping. The 7-methylguanosine or m^7G cap structures (Fig. D.1.2) GpppNmpNm are at the 5' end of nearly all eukaryotic cellular (except those from mitochondria, chloroplast, etc.) and viral RNAs. This cap is added to mRNA precursors and viral transcripts by capping enzymes that replicate in the nucleus.

A single methyl group is present on the terminal G residue at the N-7 position whereas the adjacent nucleotides can be 2'-O-methylated to different degrees. This is the basis of the cap nomenclature: m^7 GpppN (Cap 0), m^7 GpppNm (Cap 1), and m^7 GpppNmpNm (Cap 2). 5' caps are involved in several functions: (i) regulation of nuclear export; (ii) acts as protective group to prevent 5'-3' exonucleolytic degradation; (iii) promotion of translation: facilitates binding of ribosomes to 5' end of mRNA to facilitate cap-dependent protein synthesis; (iv) promotion of 5′ proximal intron excision. Thus chemical modifications of the RNA 5'-end provide a layer of 'epitranscriptomic' regulation (Furuichi, Muthukrishnan et al. 1976, Furuichi and Shatkin 1976, Muthukrishnan, Furuichi et al. 1976, Muthukrishnan, Morgan et al. 1976, Shatkin 1976, Shatkin, Banerjee et al. 1976, Furuichi, LaFiandra et al. 1977, Kozak and Shatkin 1977, Furuichi and Shatkin 2000, Furuichi 2015, Ramanathan, Robb et al. 2016).

m^7G capping is a classic regulatory feature of eukaryotes but has not been reported in prokaryotes (Ramanathan, Robb et al. 2016). However, recently the existence of new

RNA 5'-end cap containing the metabolic effector nicotinamide adenine dinucleotide (NAD^+) has been shown to be added to certain RNAs isolated from bacteria (Chen, Kowtoniuk et al. 2009, Kowtoniuk, Shen et al. 2009, Cahova, Winz et al. 2015). In 2009, chemical screens were developed to identify small molecules coupled to RNA irrespective of the specific type of small-molecule structure or a biological function of the conjugate. These screens identified Coenzyme A (CoA) and several CoA thioesters as covalent conjugates to cellular RNA in *Escherichia coli* and *Streptomyces venezuelae* (Kowtoniuk, Shen et al. 2009). A parallel study used a combination of size exclusion chromatography, enzymatic treatments, and high-resolution LC/MS to report the presence of NAD-linked RNA in *Escherichia coli* and *Streptomyces venezuelae* (Chen, Kowtoniuk et al. 2009, Kowtoniuk, Shen et al. 2009). Later, using chemo-enzymatic capture and next-generation sequencing (NAD capture), specific regulatory small RNAs (sRNAs) and sRNA-like 5'-terminal fragments of certain mRNAs were found to be NAD-capped (Cahova, Winz et al. 2015). All these studies led to further questions about the mechanism of NAD capping. In contrast to an m^7G cap, which is added to nascent RNA by a dedicated capping machinery consisting of several enzymes, it was shown that bacterial RNAP could use NAD or 3′-desphospho-coenzyme A (dpCoA) during transcription initiation as non-canonical initiating nucleotides (NCINs). Thus these metabolites were being added as caps *ab initio* (Bird, Zhang et al. 2016, Bird, Nickels et al. 2017). Thus, in sharp contrast to m^7G capping which can only occur in organisms harboring specialized capping complexes, NAD capping occurs during transcription initiation and hence is more likely to be a universal phenomenon.

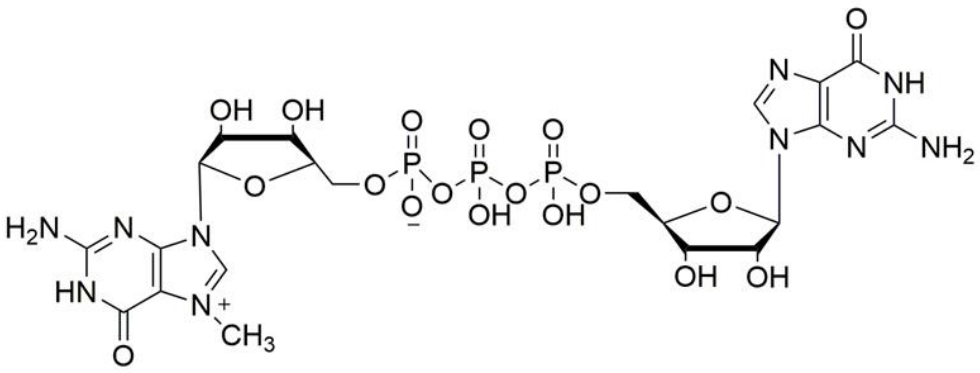


Figure D.1.2: Structure of the m^7G cap.

Using NAD captureSeq techniques, NAD-capped RNA was identified in bacterial cells (Cahova, Winz et al. 2015, Bird, Zhang et al. 2016). The same method was used to detect the presence of NAD-capped RNA in eukaryotic cells: *Saccharomyces cerevisiae* and human cell line HEK293T (Jiao, Doamekpor et al. 2017, Walters, Matheny et al. 2017). In yeast, a wide variety of RNAs were identified as being NAD-capped including RNAs produced by mtRNAP. This formed the premise for our study: Do eukaryotic single-subunit mtRNAPs perform NCIN-mediated capping? In our study, we demonstrate that single-subunit RNAPs like yeast and human mitochondrial RNAPs (also T7 RNAP) can perform NCIN-mediated capping with NAD^+ and NADH *in vitro*. The relative efficiency of NAD versus ATP-mediated initiation is much better than bacterial and eukaryotic multi-subunit RNAPs. We show that capping efficiency is determined by promoter sequence in and around the transcription start site. We also demonstrate very high levels of NAD^+ and NADH are capping up to ~60% of mitochondrial transcripts *in vivo* for yeast and ~15% for humans. The degree of capping *in vivo* and the relative abundance of

NAD^+ capping vs. NADH capping *in vivo* are influenced by intracellular levels of NAD^+ and NADH (Bird, Basu et al. 2018).

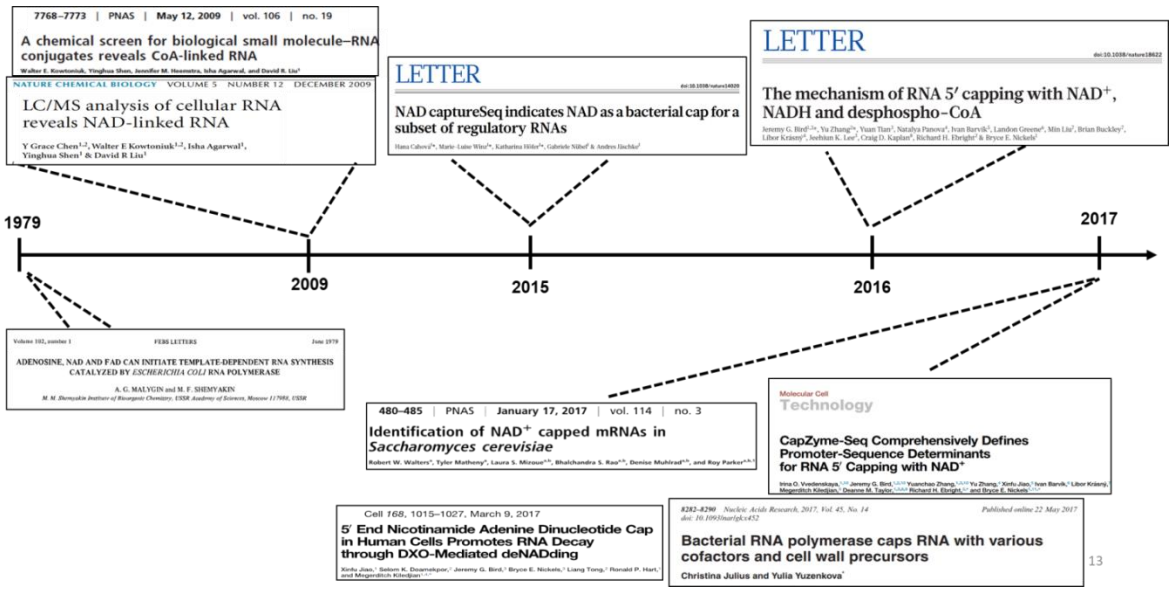


Figure D.1.3: Schematic showing the major discoveries in the NAD-capped RNA field before this work.

D.2 Results

S. cerevisiae and human mtRNAPs cap RNA with NAD^+ and NADH *in vitro*

This study began with a central question: *Do yeast and human mtRNAPs initiate in vitro transcription with NAD^+ and NADH?* To determine this, we performed *in vitro* transcription assays as described before (Figure D.2.1 and D.2.2). To study *in vitro* transcription initiation using NCINs, we first performed transcription initial products assays. We assembled *S. cerevisiae* mitochondrial 21S promoter, Rpo41 and Mtf1 (Deshpande and Patel 2014). In conjunction, we also prepared human mitochondrial mtRNAP light-strand promoter LSP AGU light-strand promoter, LSPAGU (Sologub, Litonin et al. 2009). Reactions were triggered by the addition of either ATP, NAD^+, or NADH as the initiating entity and using [∝-32P]-GTP as the extending nucleotide. For all reactions, we observed the efficient formation of an initial RNA product (Figure D.2.1, middle panel). Next, we confirmed the identities of the initial products using enzymatic treatments. RppH enzyme was previously shown to process 5'-triphosphate RNAs to 5'-monophosphate RNAs while NudC was previously shown to process 5'-NAD^+ and 5'-NADH-capped RNAs to 5'-monophosphate RNAs (Deana, Celesnik et al. 2008, Cahova, Winz et al. 2015, Hofer, Li et al. 2016). We observed that RppH processed the initial RNA products obtained with ATP, but not with NAD^+ or NADH while NudC processed the initial RNA products obtained with NAD^+ or NADH, but not with ATP (Figure D.2.1.B). From these results, we conclude that both *S. cerevisiae* mtRNAP and human mtRNAP can generate initial RNA products using NAD^+ and NADH as NCINs.

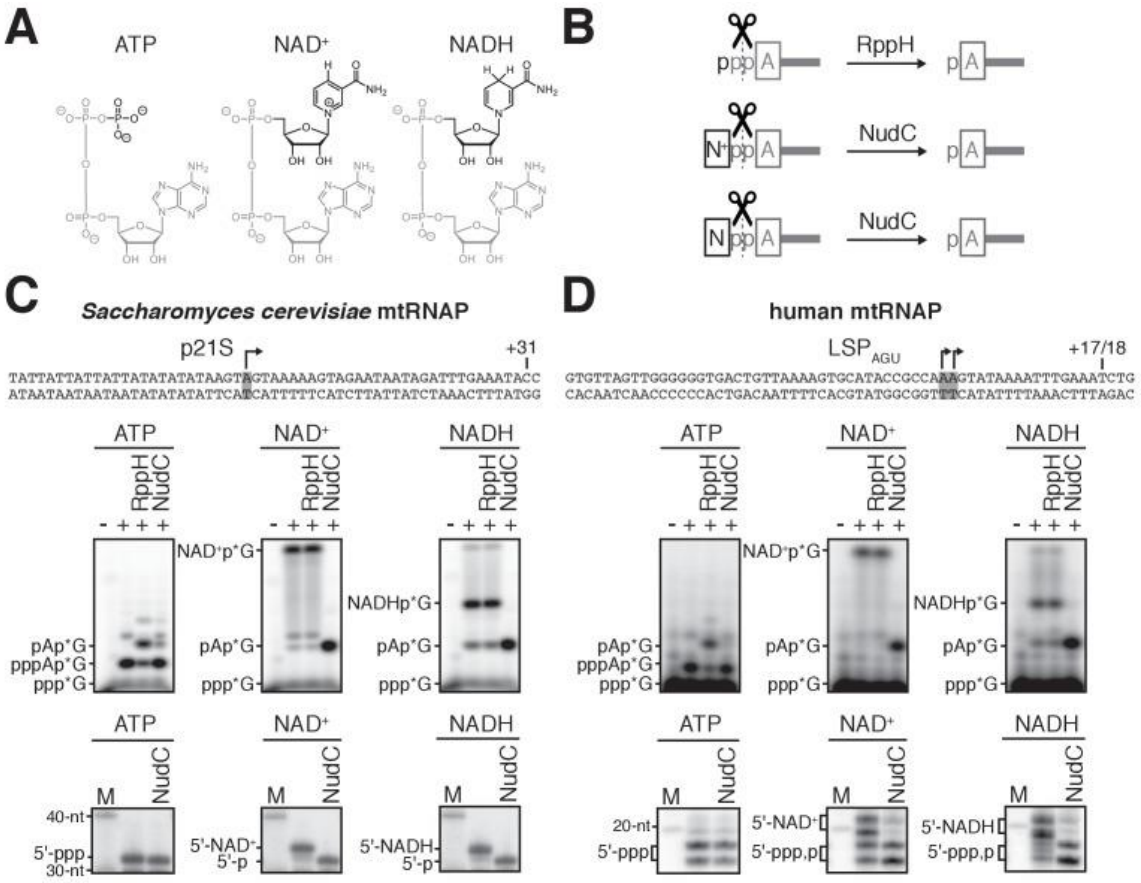


Figure D.2.1: *S. cerevisiae* and human mtRNAPs cap RNA with NAD^+ and NADH *in vitro*. (A) Structures of ATP, NAD^+, and NADH. Grey, identical atoms; black, distinct atoms. (B) Processing of RNA 5' ends by RppH and NudC. A, adenosine; N^+, NAD^+ nicotinamide; N, NADH nicotinamide; p, phosphate. (C and D) NCIN capping with NAD^+ and NADH by *S. cerevisiae* mtRNAP (C) and human mtRNAP (D). Top, promoter derivatives. Middle, initial RNA products of *in vitro* transcription reactions with ATP, NAD^+, or NADH as initiating nucleotide and [α^{32}P]-GTP as extending nucleotide. Bottom, full-length RNA products of *in vitro* transcription reactions with

ATP, NAD^+, or NADH as initiating nucleotide and [α^{32}P]-GTP, ATP, UTP, and 3'-deoxy-CTP (C), or [α^{32}P]-GTP, ATP, and UTP (D) as extending nucleotides. Products were treated with RppH or NudC as indicated. Grey box and arrow, transcription start site (TSS); +31 and+17/18, position of last NTP incorporated into full-length RNA products; M, 10-nt marker. They are adopted from (Bird, Basu et al. 2018).

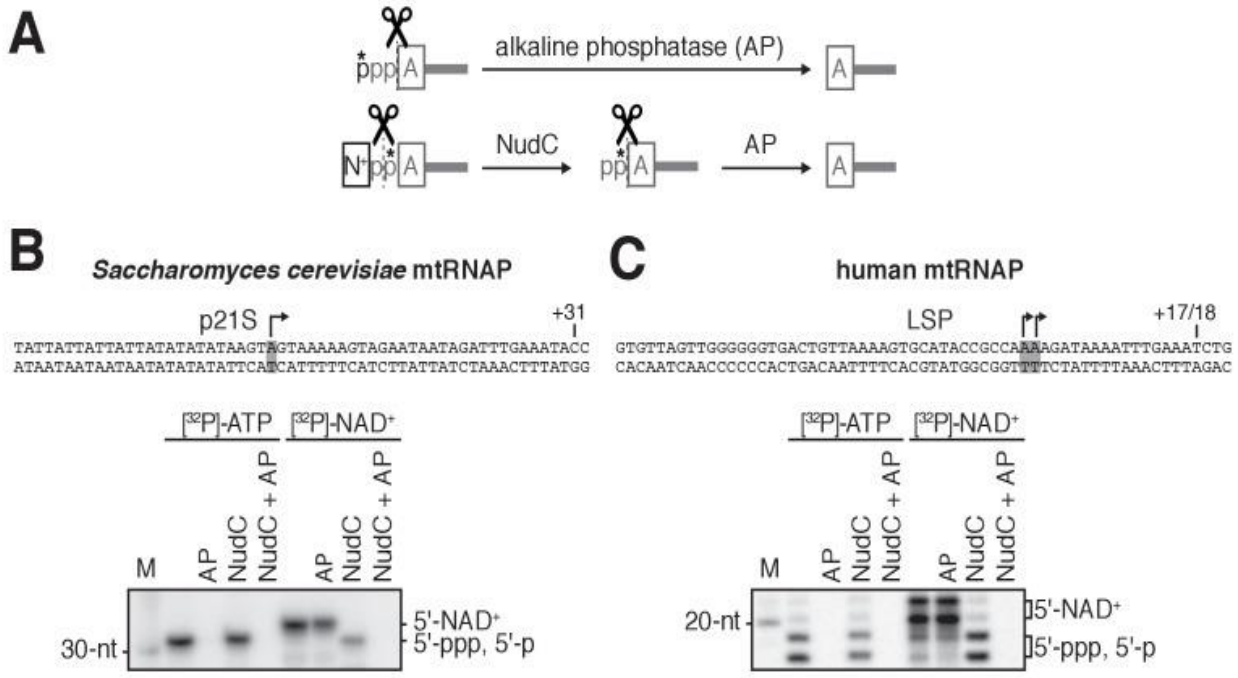


Figure D.2.2: *S. cerevisiae* and human mtRNAPs cap RNA with NAD^+ *in vitro*: additional data. (A) Processing of radiolabeled RNA 5'-ends by alkaline phosphatase (AP) and NudC. A, adenosine; N^+, NAD^+ nicotinamide; p, phosphate. *, radiolabeled phosphate. (B and C). NCIN capping with NAD^+ by *S. cerevisiae* mtRNAP (B) and human mtRNAP (C). Top, promoter derivatives. Bottom, full-length RNA products of *in vitro* transcription reactions with [γ^{32}P]-ATP or [α^{32}P]-NAD^+ as initiating nucleotide and

GTP, ATP, UTP, and 3'-deoxy-CTP (C), or GTP, ATP, and UTP (D) as extending nucleotides. Products were treated with NudC alone, AP alone, or NudC and AP, as indicated. Grey box and arrow, TSS; +31 and +17/18, the position of last NTP incorporated into RNA; M, 10-nt marker. Adapted from (Bird, Basu et al. 2018).

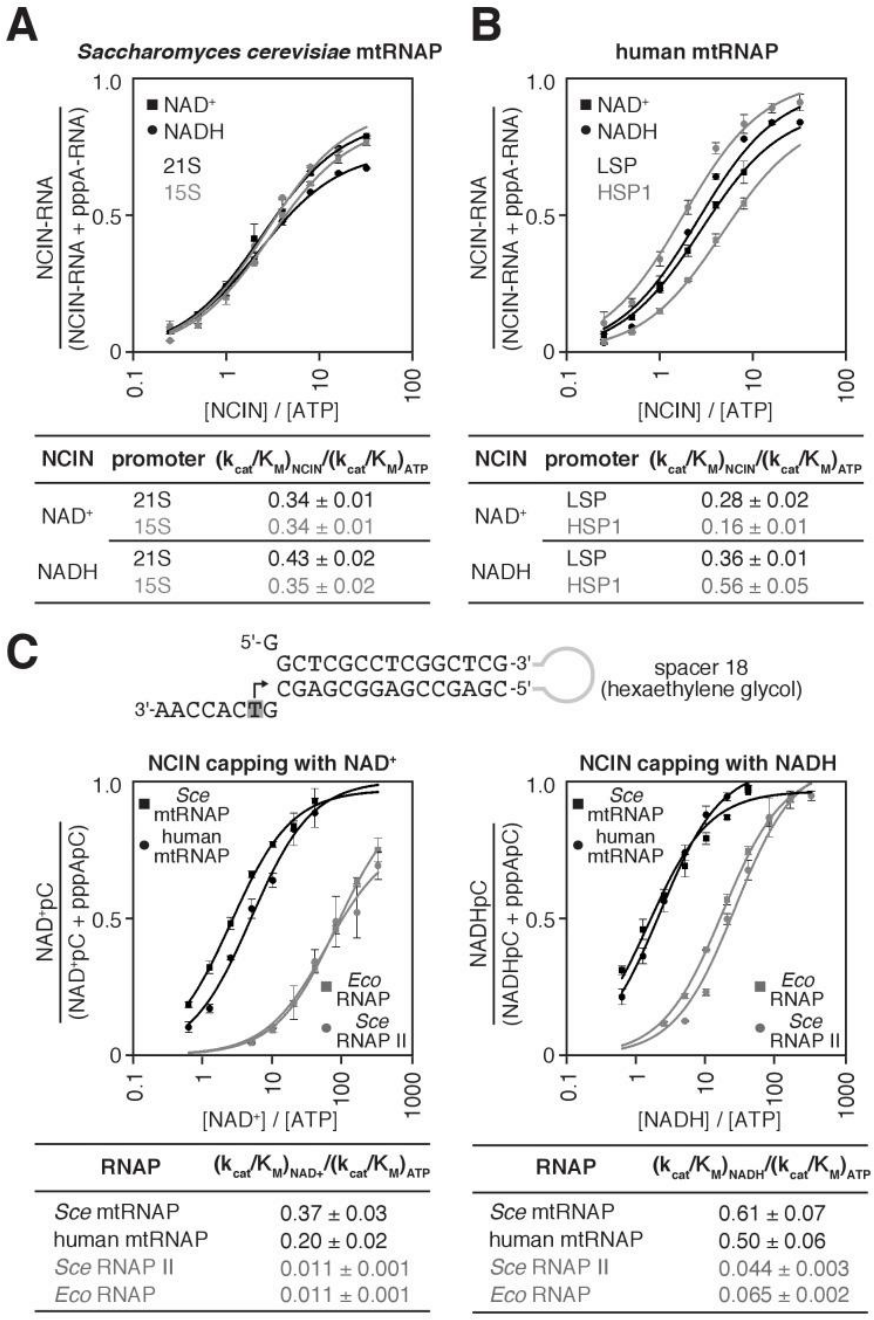
We next asked whether these initial products formed could be extended to form run-off products. We performed parallel transcription experiments using either ATP, NAD^+, or NADH as the initiating entity and using [∝-32P]-GTP, ATP, UTP, and 3' dCTP (Figure D.2.1.C, bottom) or [∝-32P] GTP, ATP, and UTP (Figure D.2.1.D, bottom) as extending nucleotides. We observed the efficient formation of run-off product for all reactions. Only NAD^+ or NADH initiated full-length products but not ATP initiated products were sensitive to NudC treatment (Figure D.2.1.D, bottom). Our results also show the formation of full-length products when [∝-32P]-ATP or [∝-32P]- NAD^+ was used as initiating nucleotides. Full-length products obtained with [∝-32P]- NAD^+ but not with [∝-32P]-ATP were protected against treatment with alkaline phosphatase (5’ phosphate processing enzyme). Full-length products obtained with [∝-32P]- NAD^+ were also treated with NudC which yielded products that now became sensitive to alkaline phosphatase treatment (Figure D.2.1 and D.2.2). From these results, we conclude that both *S. cerevisiae* mtRNAP and human mtRNAP can generate run-off RNA products using NAD+ and NADH as NCINs.

S.cerevisiae and human mtRNAPs cap RNA with NAD^+ and NADH more efficiently than bacterial and nuclear RNAPs

Next, we wanted to know the efficiency of NAD^+ utilization as an initiating nucleotide by mtRNAP as compared to ATP. To address this, we determined the relative efficiencies of NCIN-mediated initiation vs. ATP-mediated initiation, $(k_{cat}/K_M)_{NCIN} / (k_{cat}/K_M)_{ATP}$, for mtRNAPs *in vitro* (Bird, Zhang et al. 2016, Bird, Nickels et al. 2017). We assembled reactions with *S. cerevisiae* mt 21S or 15S promoter and the corresponding transcription machinery: yeast mtRNAP Rpo41 and transcription factor Mtf1. We added ATP+UTP+GTP+3' dCTP and titrated with NAD^+ to compete with the ATP. In parallel, we also assembled reactions with the human mt light-strand promoter (LSP) or heavy-strand promoter (HSP1) and the corresponding transcription machinery: human mtRNAP POLRMT and transcription factors TFB2M and TFAM (Fig. D.2.3 and D.2.4). We added ATP+UTP+GTP (for LSP) and ATP+GTP+CTP (for HSP1) and titrated with NAD^+ to compete with the ATP. We obtained values of $(k_{cat}/K_M)_{NCIN} / (k_{cat}/K_M)_{ATP}$ of ~0.3 for NCIN mediated initiation with NAD^+ by *S. cerevisiae* mtRNAP for both 21S and 15S promoters. For NADH this value varies between ~0.3 to ~0.4 for NCIN mediated initiation (Fig. D.2.3.A). We obtained values of $(k_{cat}/K_M)_{NCIN} / (k_{cat}/K_M)_{ATP}$ of ~0.2 for NCIN mediated initiation with NAD^+ by human mtRNAP for both LSP and HSP1 promoters. For NADH this value varies between ~0.36 to ~0.56 for NCIN mediated initiation (Fig. D.2.3.B). It is curious that for HSP1 promoter, catalytic efficiency seems to be higher for initiation with NADH than NAD^+. These values imply that NCIN-mediated initiation with NAD^+ or NADH is up to 40% as efficient as initiation with ATP for *S. cerevisiae* mtRNAP and up to 60% as efficient as initiation with ATP for human mtRNAP.

It is curious that our observed efficiencies of NCIN-mediated initiation with NAD^+ or NADH by mtRNAPs are substantially higher than reported for NCIN-mediated initiation with NAD^+ or NADH by cellular RNAPs which is repoted to be ~15% (Bird, Zhang et al. 2016, Vvedenskaya, Bird et al. 2018). However, these studies with bacterial RNAPs utilized initial products as the read-out, while in our mtRNAP assays we looked at run-off products. We wanted to design an assay which would let us directly compare cellular and mtRNAPs using the same DNA template with identical reaction conditions and reaction time. We used a hairpin template as shown in Figure D.2.3.C. This template can get around the necessity of sequence-specific RNAP-DNA interactions and transcription-initiation factor-DNA interactions for transcription initiation (Kadesch and Chamberlin 1982, Dedrick and Chamberlin 1985). 200 nM of tailed template and 500 nM RNAP (*S. cerevisiae* mtRNAP, human mtRNAP, *S. cerevisiae* RNAP II, T7 RNAP, or *E. coli* RNAP) were incubated at 25°C for 15 min in reaction buffer (10 mM Tris pH 8.0, 50 mM potassium glutamate, 10 mM MgCl2, 2 mM DTT, and 50 ug/ml BSA). A mixture containing NCIN (1 mM NCIN for mtRNAPs, T7 RNAP, and *E. coli* RNAP; 4 mM for *S. cerevisiae* RNAP II), ATP (0, 25,50, 100, 200, 400, 800, 1600 mM for mtRNAPs and *S. cerevisiae* RNAP II; 0, 6.25, 12.5, 25, 50, 100, 200, 400 mM for *E. coli* RNAP), 10 mM non-radiolabeled CTP, and 6 mCi [a32P]-CTP (Perkin Elmer; 3000 Ci/mmol) was added, and assays were incubated at 25°C for 1 hr. Reactions were quenched as before and loaded on gels. We monitored initial products in this assay as compared to run-off products in Fig. D.2.3.A and D.2.3.B.

Our results show that the efficiencies of NCIN-mediated initiation with NAD^+ and NADH by mtRNAP are about ~10 to ~40 fold higher than efficiencies of NCIN-mediated initiation with NAD^+ and NADH by *E. coli* RNAP and *S. cerevisiae* RNAP II (Figure D.2.3.C, bottom panel). From these results, we conclude that both *S. cerevisiae* mtRNAP and human mtRNAP can cap RNA with NAD^+ and NADH more efficiently than bacterial RNAP and eukaryotic nuclear RNAP II. We further used the same hairpin template and reaction conditions as in assays performed with mtRNAPs to determine the efficiency of NCIN-mediated initiation with NAD^+ and NADH for the single-subunit RNAP of bacteriophage T7 (T7 RNAP) (Figure D.2.5). We observed that the efficiencies of NCIN-mediated initiation with NAD^+ and NADH by T7 RNAP were nearly as high as the efficiencies of NCIN-mediated initiation by mtRNAPs. We conclude that fundamentally the efficiency of NCIN capping by members of the single-subunit RNAP family (T7 RNAP and mtRNAPs) is better than that of the members of the multi-subunit RNAP family (bacterial RNAP and eukaryotic nuclear RNAP II).



NCIN	promoter	$(k_{cat}/K_M)_{NCIN}/(k_{cat}/K_M)_{ATP}$
NAD+	21S	0.34 ± 0.01
	15S	0.34 ± 0.01
NADH	21S	0.43 ± 0.02
	15S	0.35 ± 0.02

NCIN	promoter	$(k_{cat}/K_M)_{NCIN}/(k_{cat}/K_M)_{ATP}$
NAD+	LSP	0.28 ± 0.02
	HSP1	0.16 ± 0.01
NADH	LSP	0.36 ± 0.01
	HSP1	0.56 ± 0.05

RNAP	$(k_{cat}/K_M)_{NAD+}/(k_{cat}/K_M)_{ATP}$
Sce mtRNAP	0.37 ± 0.03
human mtRNAP	0.20 ± 0.02
Sce RNAP II	0.011 ± 0.001
Eco RNAP	0.011 ± 0.001

RNAP	$(k_{cat}/K_M)_{NADH}/(k_{cat}/K_M)_{ATP}$
Sce mtRNAP	0.61 ± 0.07
human mtRNAP	0.50 ± 0.06
Sce RNAP II	0.044 ± 0.003
Eco RNAP	0.065 ± 0.002

Figure D.2.3: *S. cerevisiae* and human mtRNAPs cap RNA with NAD^+ and NADH more efficiently than bacterial and nuclear RNAPs. (A and B) Dependence of NCIN-mediated capping with NAD^+ and NADH on [NCIN] / [ATP] ratio for *S. cerevisiae* mtRNAP (A)

and human mtRNAP (B) (mean ± SD; n = 3). DNA templates and representative data are shown in Fig. D.2.4. (C) Dependence of NCIN-mediated capping with NAD^+ and NADH on [NCIN] / [ATP] ratio for mtRNAPs vs. *E. coli* RNAP and *S. cerevisiae* RNA II. Top, tailed template. Grey box and arrow indicate TSS. Bottom, the dependence of NCIN-mediated capping with NAD^+ and NADH on [NCIN] / [ATP] ratio for *S. cerevisiae* mtRNAP (*Sce* mtRNAP), human mtRNAP, *E. coli* RNAP (*Eco* RNAP) and *S. cerevisiae* RNAP II (*Sce* RNAP II) (mean ± SD; n = 3). Adopted from (Bird, Basu et al. 2018).

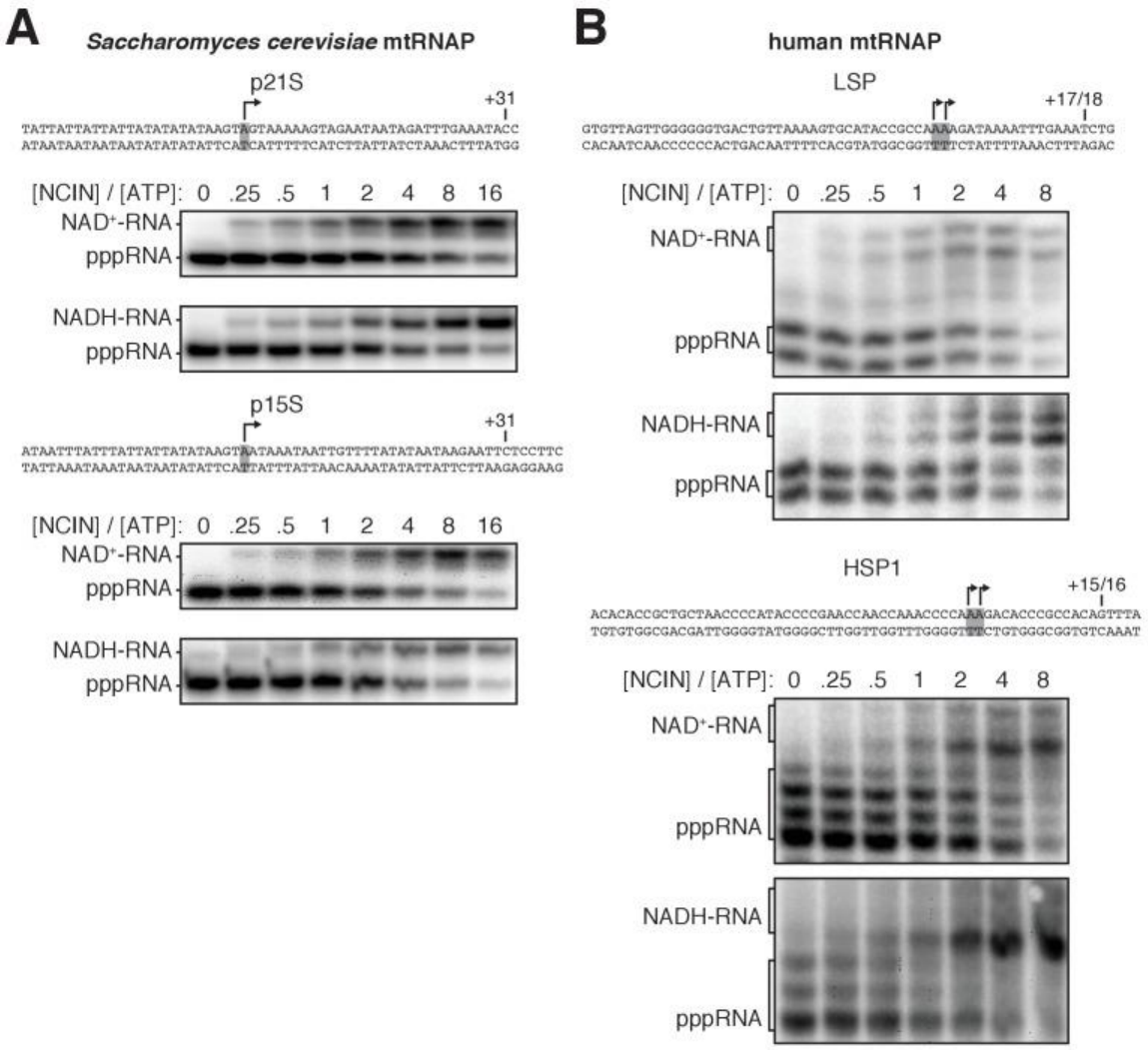


Figure D.2.4.: Dependence of NCIN-mediated capping with NAD^+ and NADH on [NCIN] / [ATP] ratio for mtRNAPs: representative data. (A and B) Panels show DNA templates and full-length RNA products of *in vitro* transcription reactions performed with *S. cerevisiae* mtRNAP (A) and human mtRNAP (B) with the indicated [NCIN] / [ATP] ratio. Grey box and arrow, TSS; +31, +17/18, +15/16, the position of last NTP incorporated into RNA; M, 10-nt marker. Adopted from (Bird, Basu et al. 2018).

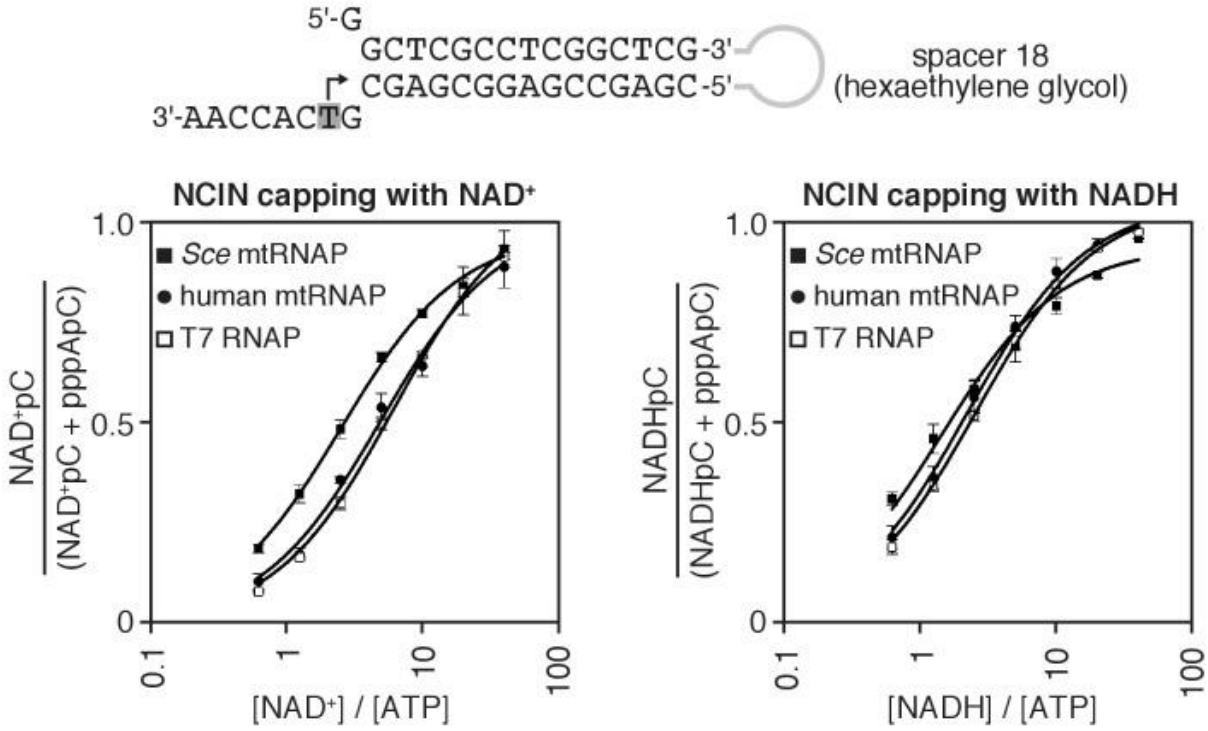
Promoter sequence determines the efficiency of RNA capping by mtRNAP

It has been previously reported that promoter sequence determinants dictate the efficiency of NCIN capping with NAD^+ and NADH for bacterial RNAPs. The sequence at and immediately upstream of the transcription start site determine how efficiently a promoter will be NCIN-capped (Bird, Zhang et al. 2016, Vvedenskaya, Bird et al. 2018). We found that NCIN capping by bacterial RNAP occurs only at promoters where the base pair (NT base:T base) at the transcription start site is an A:T (+1A promoters), and occurs most efficiently at the subset of + 1A promoters where the base pair immediately upstream of the transcription start site is purine:pyrimidine. Sequence determinants for NCIN capping by bacterial RNAP are present in the T strand of promoter DNA (i.e., the strand that templates incoming nucleotide substrates) (Bird, Zhang et al. 2016, Vvedenskaya, Bird et al. 2018). We wanted to ask: *Is the specificity for A:T at the TSS (position + 1) as observed with bacterial RNAPis also observed with mtRNAP?*

To address this, we used *S. cerevisiae* promoter derivatives having A:T or G:C at position + 1 (Figure D.2.6.A–B) to determine specificity at the +1 position. We observed NAD^+ capping in reactions performed using the promoter derivative having A:T at position + 1, but not in reactions performed using the promoter derivative having G:C at position + 1

(Figure D.2.6.B), indicating specificity for A:T at position + 1. Next, we wanted to determine whether like bacterial RNAP system, the specificity resides in the template strand for A:T at position + 1 (Bird, Zhang et al. 2016, Vvedenskaya, Bird et al. 2018). We analyzed NAD^+ capping with *S. cerevisiae* mtRNAP using template derivatives having noncomplementary non-template and template-strand nucleotides (A/C or G/T) at position + 1 (Figure D.2.6.B). We observed NAD^+ capping only with the promoter derivative having T as the template strand base at position + 1, indicating that specificity at position + 1 solely lies in the template strand.

To determine if the identity of the -1 position dictates NAD^+ capping efficiency, we analyzed the relative efficiency of NAD^+ versus ATP-mediated initiation by *S. cerevisiae* mtRNAP using promoter derivatives having either R:Y (A:T or G:C) or Y:R (C:G or T:A) at position -1 (Figure D.2.6.C). We report higher efficiencies of NAD^+ capping with promoter derivatives having R:Y at position -1 than with promoter derivatives having Y:R (Figure D.2.6.C).



RNAP	$(k_{cat}/K_M)_{NAD+}/(k_{cat}/K_M)_{ATP}$
Sce mtRNAP	0.37 ± 0.03
human mtRNAP	0.20 ± 0.02
T7 RNAP	0.18 ± 0.01

RNAP	$(k_{cat}/K_M)_{NADH}/(k_{cat}/K_M)_{ATP}$
Sce mtRNAP	0.61 ± 0.07
human mtRNAP	0.50 ± 0.06
T7 RNAP	0.43 ± 0.02

Figure D.2.5.: *S. cerevisiae* and human mtRNAPs cap RNA with NAD^+ and NADH at least as efficiently as bacteriophage T7 RNAP. Dependence of NCIN-mediated capping with NAD^+ and NADH on [NCIN] / [ATP] ratio for mtRNAPs vs. T7 RNAP. Top, tailed template. Grey box and arrow indicate TSS. Bottom, Dependence of NCIN-mediated capping with NAD^+ and NADH on [NCIN] / [ATP] ratio for *S. cerevisiae* mtRNAP (*Sce* mtRNAP), human mtRNAP, and T7 RNAP (mean ± SD; n = 3). Adopted from (Bird, Basu et al. 2018).

To confirm that the specificity at position -1 resides in the DNA template strand, we performed experiments using promoter derivatives having Y (C or T) or R (A or G) at position -1 of the template strand and having an abasic site (*) on the nontemplate strand (Figure D.2.6.D). We observed higher efficiencies of NAD^+ capping in reactions performed using promoter derivatives having Y at template-strand position -1 than with those having R. We found that the capping efficiencies for promoter derivatives having Y or R at T strand position -1 matched the capping efficiencies for homoduplex promoter derivatives (Figure 3C–D), indicating that the template strand alone dictates NAD^+ capping with *S. cerevisiae* mtRNAP. We conclude that NCIN capping with NAD^+ by mtRNAP is determined by the sequence at, and immediately upstream of, the TSS (positions +1 and -1, respectively). We further conclude that the sequence and strand preferences at positions +1 and -1 for NCIN capping is conserved between mtRNAP and bacterial RNAP (Figure D.2.6.C–E) (Bird, Zhang et al. 2016, Vvedenskaya, Bird et al. 2018). We speculate that that sequence and strand preferences may be universal determinants of NCIN capping with NAD^+ for all RNAPs. This is also found to be true for bacteriophage: sequence preferences for NCIN capping with NAD^+ by T7 RNAP, are similar to the sequence preferences observed for yeast mtRNAP and bacterial RNAP (Figure D.2.6.E and D.2.7). From structural modeling, we hypothesize that the basis for these sequence and strand preferences is universal. T strand + 1T for Watson-Crick base pairing to the NAD^+ adenine moiety is a stringent requirement. Additionally, there is a preference for T strand -1Y for 'pseudo' base pairing to the NAD^+ nicotinamide moiety (Bird, Zhang et al. 2016, Vvedenskaya, Bird et al. 2018).

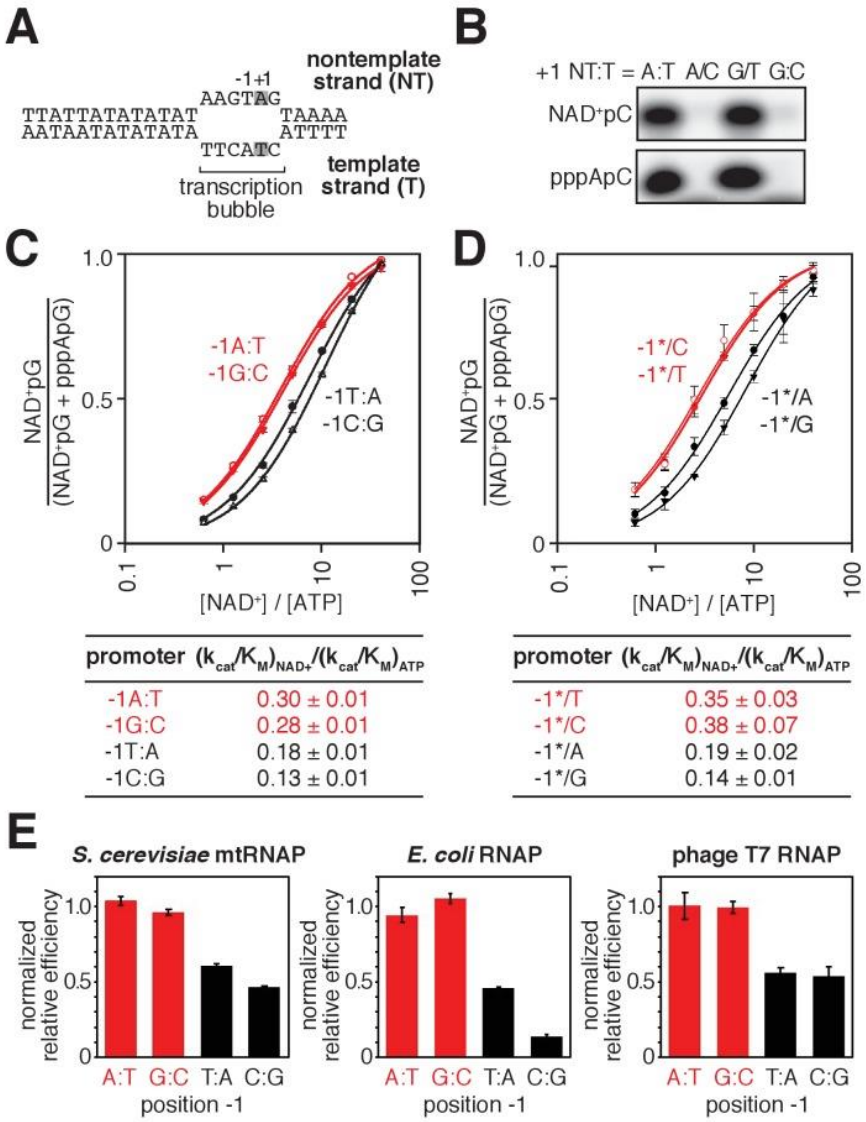
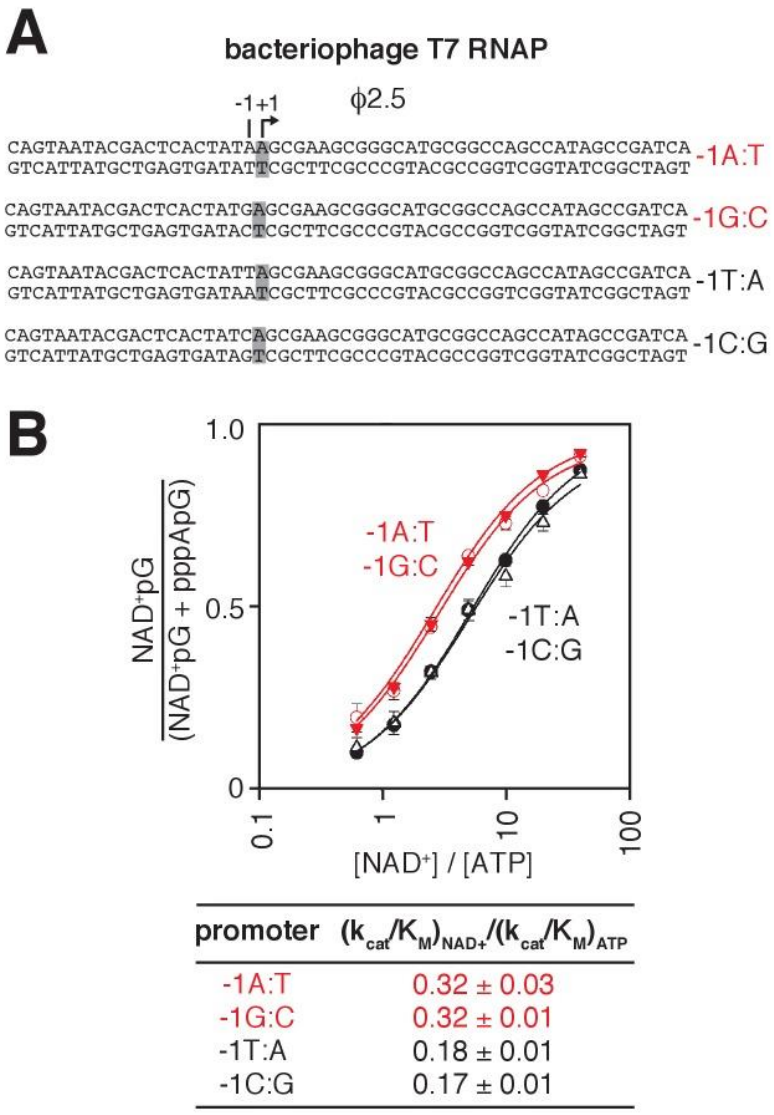


Figure D.2.6.: Promoter sequence determines the efficiency of RNA capping with NAD^+ by mtRNAP. (A) *S. cerevisiae* mitochondrial 21S promoter DNA depicted in the context of the mtRNAP-promoter open complex. DNA nontemplate strand (NT) on top; DNA template strand (T) on bottom; Unwound, non-base-paired DNA region, 'transcription bubble,' indicated by raised and lowered nucleotides; +1 and grey boxes, bases at the TSS; −1, bases immediately upstream of the TSS (the 21S promoter is a −1Y promoter). (B) Products of transcription reactions with NAD^+ as initiating nucleotide and

[α^{32}P]-CTP as extending nucleotide for templates having complementary or non-complementary nucleotides at position +1. (C) Dependence of NAD^+capping on [NAD^+] / [ATP] ratio for homoduplex templates having A:T, G:C, T:A, or C:G at position −1 relative to TSS (mean ± SD; n = 3). Red, −1R promoters; black, −1Y promoters. (D) Dependence of NAD^+capping on [ATP] / [NAD^+] ratio for heteroduplex templates having an abasic site (*) on the DNA nontemplate strand (mean ± SD; n = 3). Red, promoters with a template-strand Y; black, promoters with a template-strand R. (E) Sequence preferences at position −1 for *S. cerevisiae* mtRNAP, *E. coli* RNAP, and T7 RNAP. Graphs show normalized values of $(k_{cat}/K_M)_{NAD+}$ / $(k_{cat}/K_M)_{ATP}$ determined for homoduplex templates having A:T, G:C, T:A, or C:G at position −1 (mean ± SD; n = 3). Normalized values were calculated by dividing the value for each promoter by the average value measured for −1R promoters. Data for *S. cerevisiae* tRNA are from panel C, data for *E. coli* RNAP are from (Vvedenskaya, Bird et al. 2018), and data for T7 RNAP are from Fig. D.2.7. Adopted from (Bird, Basu et al. 2018).



promoter	$(k_{cat}/K_M)_{NAD+}/(k_{cat}/K_M)_{ATP}$
-1A:T	0.32 ± 0.03
-1G:C	0.32 ± 0.01
-1T:A	0.18 ± 0.01
-1C:G	0.17 ± 0.01

Figure D.2.7.: Promoter sequence determines the efficiency of RNA capping with NAD^+: bacteriophage T7 RNAP. (A) Bacteriophage T7 RNAP-dependent promoter derivatives analyzed. Red, −1R promoters; black, −1Y promoters; +1 and grey box, bases at the TSS; −1, bases immediately upstream of the TSS. (B) Dependence of NAD^+ capping on $[NAD^+]$ / [ATP] ratio for homoduplex templates having A:T, G:C, T:A, or C:G at position −1 relative to TSS (mean ± SD; n = 3). Adopted from (Bird, Basu et al. 2018).

Detection and quantitation of NAD^+ - and NADH-capped mitochondrial RNA *in vivo*: boronate affinity electrophoresis with processed RNA and synthetic standards

After establishing that NCIN capping is performed by mtRNAPs *in vitro,* we wanted to investigate if this phenomenon occurs *in vivo*. We decided to use northern blot analysis to detect the presence of NCIN capped RNA in cells. For this, we needed to be able to differentiate capped vs. uncapped RNA *in vitro*. Previously, boronate affinity electrophoretic techniques had been used to resolve capped vs. uncapped RNA (Igloi and Kossel 1985, Igloi and Kossel 1987, Nubel, Sorgenfrei et al. 2017). Different types of capped species like m^7G, $NAD^{+,}$ and NADH capped RNAs that contain vicinal-diol groups can be differentiated using this method. As shown in Fig. D.2.8.B, the vicinal 2 hydroxyl residues react with the phenyl boronic acid and form a large adduct that has differential migration on electrophoretic gel depending on the identity of the cap. However, a major drawback of this technique is that resolution is affected once the length of the RNA goes beyond 200 nucleotides. Additionally, though this technique can differentiate between capped and uncapped species, it cannot differentiate across caps, i.e. between m^7G, $NAD^{+,}$ and NADH, etc.

To overcome these limitations, we used boronate affinity electrophoresis in conjunction with oligodeoxynucleotide-mediated RNA cleavage (Santoro and Joyce 1997, Joyce 2001). Next, we needed standards to identify the different caps based on their mobility. For this, we used *in vitro* generated NCIN-capped RNA. The scheme is shown in Fig. D.2.8.A: we used DNAzyme cleavage to process the 3' end of long RNAs that gives defined and uniform 3' ends. The migration of this RNA will now be dictated by its 5' end. Comparing this RNA with the *in vitro* generated standards enabled us to determine

the identity of the 5' end moiety. First, we treated the *in vitro* standards having sequences identical to RNA of interest with DNAzyme. Similar treatments were given to target RNAs to generate uniform 3' ends. All these RNA fragments were then resolved using boronate affinity electrophoresis. We then detected these DNAzyme-generated 5'-end-containing RNA fragments by hybridization with a radiolabeled oligodeoxyribonucleotide probe. We chose to analyze RNA generated from two previously identified yeast promoters known to contain NAD^+ capping: COX2 and 21S (Walters, Matheny et al. 2017). We isolated total RNA and using our optimized method analyzed COX2 and 21S RNAs. For both promoters, we detected at least one RNA species whose mobility was retarded more than uncapped RNA which points to the presence of capped species in the 5' end. This species was eliminated on treatment with the decapping enzyme NudC which confirmed the presence of a 5' cap. Comparison of the mobility of this species to the *in vitro* standards showed that one of the capped species is NAD^+ capped RNA, and the other capped species, present under these growth conditions only for COX2 RNA, is NADH-capped RNA (Fig. D.2.8.C). So, from this, we see that both COX2 and 21S RNAs are present in NAD^+ -capped forms and that COX2 RNA also is present in an NADH-capped form. From these results, we conclude that *S. cerevisiae* mitochondrial RNAs undergo both NAD^+ and NADH capping in cell and that this capping is at the transcription initiation step. Our previous studies have shown that the levels of NCIN-capping in (exponentially growing) *E. coli* ranges between 1-20%, while NADH capping was not detected (Cahova, Winz et al. 2015, Bird, Zhang et al. 2016, Nubel, Sorgenfrei et al. 2017, Vvedenskaya, Bird et al. 2018). In yeast, we detected surprisingly high levels of NAD^+ and NADH: NAD^+ capped RNA comprises ~50% of

the total COX2 RNA pool and NADH-capped RNA comprises ~10% of the total COX2 RNA pools. For 21S RNA, NAD^+ -capped RNA comprises ~30% of the total 21S RNA pool.

Next, we wanted to conduct similar NCIN capped RNA detection in the human mt system. For this, we performed analogous experiments analyzing RNAs produced by transcription from the human mitochondrial LSP promoter (Figure D.2.9.B, top panel). We isolated and analyzed total RNA from HEK293T cells. We observed an NAD^+ -capped species comprising ~15% of the total LSP-derived RNA pool (Figure D.2.9.B, top panel). The results establish that human mitochondrial RNAs undergo NAD^+ capping in cells.

Mitochondria are the haven for metabolism and energy production in the cell. Both the tricarboxylic acid cycle (TCA) cycle and oxidative phosphorylation occur in the mitochondria. The TCA cycle reduces NAD^+ to NADH, and oxidative phosphorylation oxidizes NADH to NAD^+. Thus, these two processes help maintain the NAD^+/NADH balance in the mitochondria. Thus, we wanted to test whether the observed levels of capping are modulated by the available pools of NAD^+ and NADH in the cell.

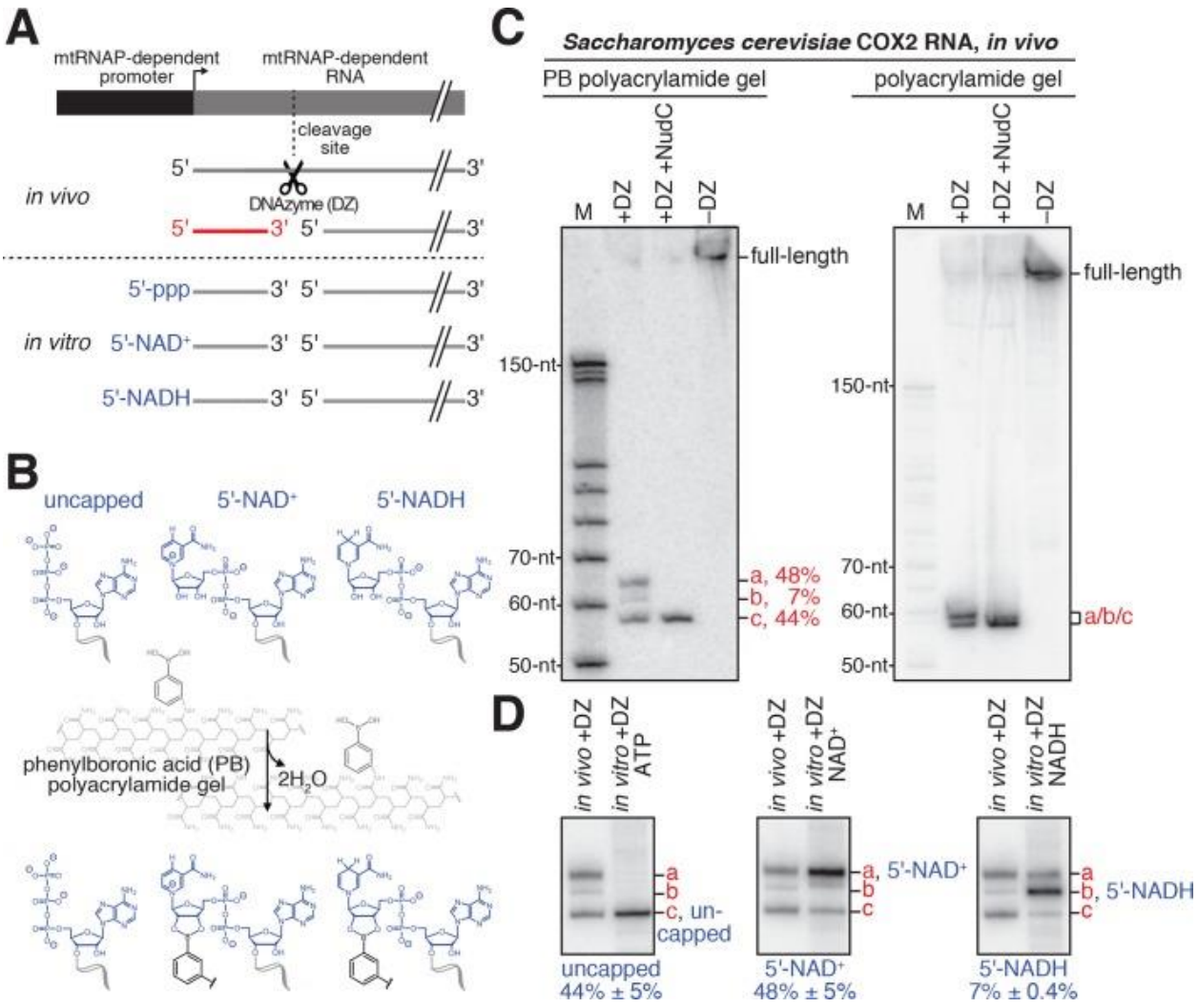


Figure D.2.8.: Detection and quantitation of NAD^+- and NADH-capped mitochondrial RNA *in vivo*: boronate affinity electrophoresis with DNAzyme-cleaved cellular RNA and DNAzyme-cleaved synthetic NCIN-capped RNA standards. (A) Use of DNAzyme (DZ) to process RNA to yield a defined, short 5'-end-containing subfragment, in parallel *in vivo* (red) and *in vitro* (blue). Uncapped, 5'-triphosphate (ppp) end generated using ATP-mediated initiation; 5'-NAD^+, NAD^+-capped end generated using NAD^+-mediated initiation; 5'-NADH, NADH-capped end generated using NADH-mediated initiation. (B)

Use of boronate affinity electrophoresis to resolve 5'-uncapped, 5'-NAD^+, and 5'-NADH containing RNAs. Grey, structure of phenylboronic acid (PB) polyacrylamide gel. (C) PB-polyacrylamide gel (left) and polyacrylamide gel (right) analysis of DNAzyme-generated 5'-end-containing subfragments of *S. cerevisiae* mitochondrial RNA COX2. Red, observed 5'-end-containing RNA subfragments resolved by PB-polyacrylamide-gel (left) or not resolved by polyacrylamide gel (right); identities of these subfragments are defined in Panel D. (D) Comparison of electrophoretic mobilities of observed 5'-end-containing subfragments of COX2 RNA generated *in vivo* to 5'-end-containing subfragments of synthetic RNA standards generated *in vitro*. a, NAD^+-capped RNA; b, NADH-capped RNA; c, uncapped RNA (mean ± SD; n = 3). Adopted from (Bird, Basu et al. 2018).

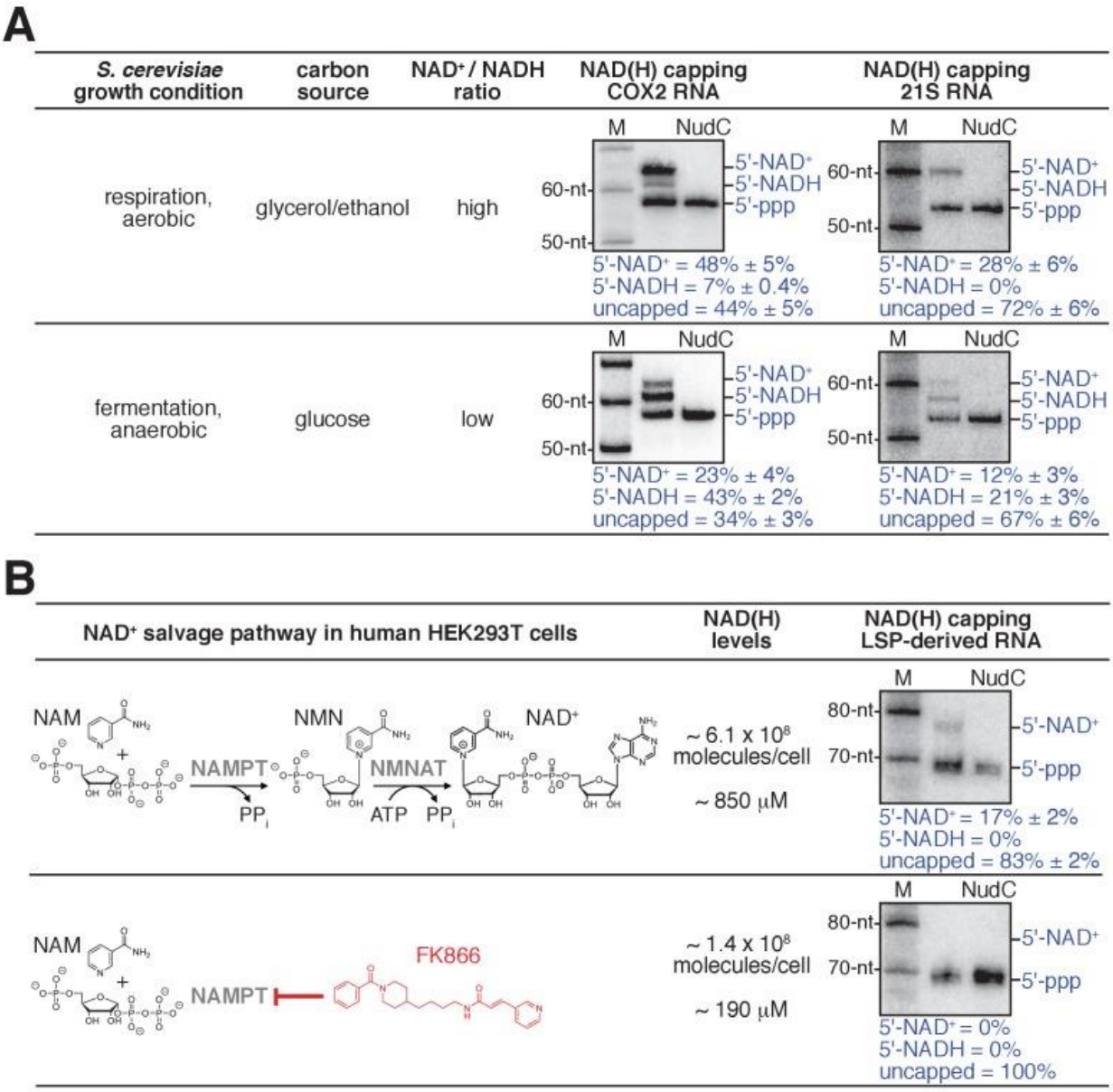


Figure D.2.9.: Detection and quantitation of NAD^+- and NADH-capped mitochondrial RNA *in vivo*: effects of intracellular NAD^+ and NADH levels in *S. cerevisiae* and human cells. (A) Changes in intracellular NAD^+/NADH ratios result in changes in levels of NAD^+- and NADH-capped mitochondrial RNA (mean ± SD; n = 3). Gel images show representative data for *S. cerevisiae* COX2 RNA (left) and 21S RNA (right). Blue annotations as in Fig. D.2.8. (B) Changes in intracellular NAD(H) levels Changes in

intracellular NAD(H) levels result in changes in levels of NAD^+- and NADH-capped mitochondrial RNA (mean ± SD; n = 3). Gel images show representative data for LSP-derived RNAs. Red, NAD(H) biosynthesis inhibitor FK866; NAMPT, Nicotinamide phosphoribosyltransferase; NMNAT, Nicotinamide mononucleotide adenylyltransferase. Adopted from (Bird, Basu et al. 2018).

We assessed whether changing intracellular [NAD^+] / [NADH] ratios affects the NAD^+ and NADH capping of mitochondrial RNAs. We isolated total RNA from *S. cerevisiae* grown under conditions that result either high or low [NAD+] / [NADH] ratios (Canelas, van Gulik et al. 2008, Bekers, Heijnen et al. 2015): that is respiration (glycerol/ethanol media; aerobic) or fermentation (glucose media; anaerobic). We analyzed COX2 and 21S RNAs. For COX2, when cells were grown under a new growth condition that yields a high [NAD+] / [NADH] ratio from a growth condition that yield a low [NAD+] / [NADH] ratio, a sharp decrease in the levels of NAD^+ capping (from ~50% to ~20%) and corresponding increase in the level of NADH capping (from ~10% to ~40%) was observed. However, the total levels of NAD^+ and NADH capping remain constant in both cases indicating the relative number of NCIN capping levels do not change (Fig. D.2.9.A). Similar results were observed for 21S RNA, on changing growth conditions from high to [NAD+] / [NADH] ratio, the level of NAD^+ capping decreases (from ~30% to ~10%), the level of NADH capping increases (from 0% to ~20%), and the total level of NAD^+ and NADH capping remains constant (~30%). Our results show that changing the [NAD+] / [NADH] ratio changes transcriptional outputs of NCIN-capped RNAs *in vivo*.

We wanted to see whether a similar phenomenon could also be observed in human mitochondria. We isolated total RNA from human HEK293T cells grown under conditions yielding either high or low intracellular NAD(H) levels: standard growth media or growth media in the presence of the NAD(H)-biosynthesis inhibitor FK866 (Hasmann and Schemainda 2003, Khan, Tao et al. 2006). On changing from high intracellular NAD(H) levels to low NAD(H) levels by adding FK866, we observe a sharp change in the total level of NCIN capping (from ~15% to 0%) (Fig. D.2.9.B). Thus as in the yeast system, in the human system too, we see that the levels of *in vivo* NCINs dictate the amount of NCIN capped RNA.

Overall from Fig. D.2.9, we conclude that mtRNAP serves as both sensor and actuator in coupling [NAD+] / [NADH] ratios to relative levels of NAD^+ and NADH capped mitochondrial RNAs. mtRNAP is also involved in coupling total NAD(H) levels to total levels of NCIN-capped mitochondrial RNAs. Both these processes couple cellular metabolism to mitochondrial transcription outputs. We propose that mtRNAPs serve as sensors that can elect what residue to use for transcription initiation and thereby function as actuators by incorporating specific residues at the RNA 5' end during transcription initiation.

D.3 Discussion

In this study, we have shown that both *S. cerevisiae* and human mtRNAPs are capable of *in vitro* capping with both NAD^+ and NADH (Fig. D.2.1 and D.2.2). This capping is far more efficient than bacterial and eukaryotic nuclear RNAPs (Fig. D.2.3, D.2.4 and D.2.5). We further show that capping efficiency is determined by sequences at and around the transcription start site (Fig. D.2.6 and D.2.7). Our *in vivo* data shows that a high percentage of transcripts are capped with NAD^+ and NADH and these values are sensitive to the cellular NAD^+ and NADH levels (Fig. D.2.8 and D.2.9).

NCIN capping has been shown to have functional consequences: modulation of RNA stability and translatability (Cahova, Winz et al. 2015, Bird, Zhang et al. 2016, Jiao, Doamekpor et al. 2017). Mitochondrial RNA capping occurs at a higher level than non-mitochondrial RNAs (Fig. D.2.9 vs. previous studies). NCIN capping by single-subunit RNAPs has a higher efficiency than by multi-subunit RNAPs which we speculate might be more important than other kinds of capping in other cellular compartments. The following factors support our argument: (i) mtRNAPs are substantially more efficient at NAD^+ and NADH capping than bacterial and eukaryotic nuclear RNAPs (Fig. D.2.3, D.2.4 and D.2.5); (ii) mitochondria have a high concentration of NAD^+ and NADH relative to ATP than any bacterial or eukaryotic compartments (Cambronne, Stewart et al. 2016, Chen, Freinkman et al. 2016, Park, Rubin et al. 2016); (iii) all yeast and human promoters have +1A. This implies in theory all transcripts can be capped with NAD^+ and NADH (Chang and Clayton 1984, Biswas 1999, Taanman 1999); (iv) as opposed to NAD^+ capping for non-mitochondrial RNAs *in vivo* we observed both NAD^+ and NADH capping for mt RNAs *in vivo* (Fig. D.2.8 and D.2.9). Thus, it is possible that

mitochondria are the only compartment where NAD^+ and NADH capping occurs that affects RNA fate and hence global mt gene expression.

We observed less capping on 21S RNA than COX2 promoters (Fig. D.2.9). This implies that capping for a ribosomal RNA is more than an mRNA. It will be interesting to contrast other ribosomal and mRNA and see whether this is a general trend.

We observed that both in the human and yeast mitochondrial systems, changing the growth conditions (for yeast) or addition of FK866 (for humans) modulate the amount of available NAD^+/NADH. This available level of NCIN in the cell affects the amount of NCIN capping. We speculate that NCIN-capping is an epitranscriptomic modification because the level of specific capping changed as we changed the growth conditions. Thus the NCIN-cap may be acting as an epitranscriptomic switch, possibly affecting RNA function and thus overall gene expression.

D.4 Future Directions

Recent research in the NCIN capping field has opened up interesting but yet unresolved questions. Some of those are listed below.

1. Using NCINs other than NAD and NADH for transcription initiation:
    (i) Several different kinds of metabolic intermediates have been identified that are used during transcription initiation by bacterial RNAP (Julius and Yuzenkova 2017). This indicated there is a wide repertoire of RNA in a cell can be capped at a time with different 5' ends. The affinity of the bacterial RNAP to use these substrates for 5' capping was also found to be variable. Our data suggest that the mtRNAP working as both sensor and actuator caps RNA with NAD^+ and NADH depending on the available cellular pool of NAD^+ and NADH (Julius and Yuzenkova 2017). We hypothesize that the levels of RNA capped by different NCINs might be similarly linked to the available amount of these substrates.
    (ii) The different molecules used for 5' capping impart different identities and possibly different stability and translatability to the RNAs. The functions and implications of these variations in the 5' ends of this RNA remain to be seen.
    (iii) It will be interesting to study whether, like bacterial RNAP, mtRNAPs are also capable of initiating with metabolic intermediated other than NAD^+ and NADH.

2. Role of the mt transcription factors in modulating NCIN capping in mitochondria: In the human mt system, it has been reported that the N-terminal region of the transcription factor TFB2M interacts with the incoming iNTPs. Our data suggest that the C-terminal region of both the yeast and human transcription factors Mtf1 and TFB2M modulated iNTPs binding (Basu & Patel, unpublished data). It will be interesting to study how and at which steps of transcription, these factors modulate NCIN capping in the mitochondria.
3. Single-subunit RNAPs are better at initiating with NCIN than multi-subunit RNAPs: our data suggest that single subunit RNAPs are better at initiating with NCIN than multi-subunit RNAPs (Bird, Basu et al. 2018). Does the transcription start site for single-subunit RNAPs structurally more adept at accommodating NCINs as compared to multi-subunit RNAPs? Our current knowledge is insufficient in this respect as the human mt IC structure does not have a resolution around the transcription start site (Bird, Basu et al. 2018).
4. Other pathways involved in NCIN capping of RNA in mitochondria (other than transcription initiation): A subset of sno/scaRNAs have been reported to contain an NAD^+ cap at their 5' end in human HEK293T cells (Jiao, Doamekpor et al. 2017). Mammalian sno/scaRNAs are mainly intronic and produced after exonucleolytic RNA processing. This would require NAD^+ capping to occur at a later stage than transcription initiation. This makes it possible that other mechanisms in addition to transcription initiation might be involved in 5' NAD^+ capping. These pathways require investigation.

[E] A Tale of Tails: Functions of the C-tail of mitochondrial transcription factors Mtf1 (yeast) and TFB2M (human)

NOTE: This project is the result of the collaborative efforts of Dr. Hajin Kim and Dr. Smita S. Patel's lab. Hence this section requires acknowledgment of all members involved. Author contributions are listed below.

Urmimala Basu: Formal analysis, Funding acquisition, Investigation, Visualization, Methodology, Writing—review and editing; Aishwarya Deshpande: Formal analysis, Investigation, Visualization, Methodology; B.K. Shon: Formal analysis, Investigation, Visualization, Methodology; Jiyau Shen: Formal analysis, Investigation, Visualization, Methodology; Smita S Patel: Conceptualization, Supervision, Funding acquisition, Visualization, Project administration, Writing—review and editing; Hajin Kim: Supervision, Funding acquisition, Visualization, Project administration.

E.1 Introduction

Mitochondria are present in all eukaryotes acting as molecular motors producing usable energy in the form of ATP through oxidative phosphorylation pathways. The traditional view of mitochondria has been restricted to its function as 'the powerhouse of the cell'. This monochromatic perspective of the organelle has been revived over the past couple of decades: we now know that these complex and dynamic organelles play astoundingly diverse roles in metabolism, cell signaling, apoptosis, etc. (Pagliarini and Rutter 2013) (McBride, Neuspiel et al. 2006, Nunnari and Suomalainen 2012). Their cardinal role in cellular energetics makes mitochondria a key player in metabolic, degenerative and age-related diseases (Shadel 2008, Shutt and Shadel 2010). Dysregulation of mitochondria is ubiquitous in diabetes, obesity, cardiovascular disease, cancer, etc.

Mitochondria have endosymbiotic origin: a host single cell organism engulfed an α-proteobacteria giving rise to present day eukaryotes with two genomes, the nuclear DNA and the mitochondrial DNA (mtDNA) (Lang, Gray et al. 1999, Dyall, Brown et al. 2004). The mtDNA is replicated/transcribed by enzymes distinct from the enzymes that catalyze nuclear DNA transcription. Curiously, in spite of its prokaryotic origin, the mtDNA enzymes are homologous to single-subunit phage T7 encoded enzymes (Cermakian, Ikeda et al. 1997). Thus, yeast and human mtDNA are transcribed by phage T7-like single-subunit RNAPs called Rpo41 and POLRMT, respectively. Unlike T7 RNAP that does not require any transcription factors, mtRNAPs require one or more transcription factors to initiate RNA synthesis. Rpo41 depends on one transcription factor Mtf1 for initiation whereas POLRMT depends on two transcription factors, mitochondrial transcription factor B2 (TFB2M) and mitochondrial transcription factor A (TFAM) to

initiate RNA synthesis (Falkenberg, Gaspari et al. 2002, Litonin, Sologub et al. 2010, Shutt, Bestwick et al. 2011). Transcription initiation is the most regulated step of gene expression. Thus, it is crucial to understand the basic mechanisms of how these initiation factors function.

The catalytic function of synthesizing RNA is performed exclusively by Rpo41 or POLRMT. However, Rpo41 by itself binds non-specifically to duplex DNA with a K_d of 58 nM and bends it by 52° (Tang, Deshpande et al. 2011), but it cannot melt the promoter and thus cannot initiate transcription unless the promoter is pre-melted (Matsunaga and Jaehning 2004). Addition of Mtf1 generates a 300-fold tighter complex of Rpo41 with promoter with a K_d of 0.2 nM and induces a sharp 90° DNA bend (Tang, Deshpande et al. 2011), resulting in the formation of an open complex with DNA melted from the -4 to +2 region (Tang, Paratkar et al. 2009). The dynamic process of bending-unbending of DNA was observed by single-molecule studies which showed that Mtf1 facilitates open complex formation by increasing the DNA bending rate and decreasing the unbending rate (Kim, Tang et al. 2012).

Much of the corresponding mechanistic details of the human mtRNAP system remain elusive. TFB2M has been identified as the chief contributor to promoter melting (Gaspari, Larsson et al. 2004, Sologub, Litonin et al. 2009, Morozov, Agaronyan et al. 2014, Hillen, Morozov et al. 2017, Ramachandran, Basu et al. 2017). TFB2M has interactions with POLRMT, with the promoter around the transcription start site, and the initiating NTPs (Sologub, Litonin et al. 2009, Morozov, Agaronyan et al. 2014, Hillen, Morozov et al. 2017). TFAM binds site specifically to the upstream -36 to -16 region of the light strand and heavy strand promoters (LSP and HSP, respectively) inducing U-turn

bend in the DNA. TFAM also interacts with POLRMT on promoter DNA (Morozov, Agaronyan et al. 2014, Yakubovskaya, Guja et al. 2014, Morozov, Parshin et al. 2015). The current model is that the role of TFAM is to recruit POLRMT to the promoter initiation site and TFB2M is required to melt the promoter (Morozov, Agaronyan et al. 2014, Posse, Hoberg et al. 2014).

It has been speculated that Mtf1 and TFB2M are homologous to sigma factors of *Escherichia coli* (Jang and Jaehning 1991, Cliften, Jang et al. 2000). However, the crystal structure of Mtf1 and TFB2M revealed structural homology to ribosomal RNA methyltransferases (Schubot, Chen et al. 2001, Hillen, Morozov et al. 2017). Mtf1 has been shown to retain none, while TFB2M has been shown to retain some of its methyltransferase activity (Cotney and Shadel 2006). Both these homologs have a dumbbell-shaped structure, with a large N-terminal domain (NTD) that adopts a fold similar to S-adenosyl-methionine-dependent methyltransferases with a central seven-stranded β-sheet flanked on either side by three α-helices (Schubot, Chen et al. 2001, Hillen, Morozov et al. 2017). The smaller C-terminal domain consists of four α-helices and an extended C-terminal tail. The 16 amino acids C-terminal region or C-tail of Mtf1 is missing from the crystal structure. For TFB2M, six amino acids were missing from the C-tail in one of the two copies of the asymmetric unit. The absence of the C-tail in the crystal structures points to the flexibility of the tail. Protein-DNA cross-linking studies place Mtf1 in the proximity of the melted strands in the -4 to +2 promoter region (Paratkar and Patel 2010), and its C-terminal 13 amino acids cross-link to -3 and -4 bases of the template strand (Savkina, Temiakov et al. 2010). The *E. coli* sigma 3.2 regions that has been implicated in iNTPs binding, abortive synthesis, template alignment, and

transition into elongation also cross-links to the -3/-4 T strand. The C-tail of Mtf1 also has been implicated in promoter melting (Drakulic, Wang et al. 2014). The IC structure for yeast mt system is not available, but the IC of the human mt system points that the C-tail stabilizes the intercalating hairpin of POLRMT which promotes open complex formation (Hillen, Morozov et al. 2017). Based on this information, we hypothesized that the C-tail of Mtf1 is involved in one or more steps of transcription initiation.

Here we have studied the function of the C-tail of Mtf1. We systematically deleted 4, 12, 16 and 20 amino acids from the C-terminal end of Mtf1 and tested the behavior of the deletion mutants (Mtf1-Δ4, Mtf1-Δ12, Mtf1-Δ16, and Mtf1-Δ20 respectively) during transcription initiation to determine the role of the C-tail. We started by testing the promoter melting ability of these mutants using 2-aminopurine fluorescence assay: all mutants efficiently formed the -4 to +2 open complex, suggesting that the C-tail of Mtf1 is not important for promoter melting. Next, we used single-molecule Fluorescence Resonance Energy Transfer (smFRET) to test the role of C-tail in promoter bending. Our results show that the C-tail helps push the equilibrium towards a highly bent catalytically active state that corresponds to the aligned conformation of the promoter in the IC. We show that C-tail of Mtf1 facilitates template alignment that aids in iNTPs binding. Deletion of Mtf1 C-tail resulted in defective promoter bending with a consequent defect in iNTPs binding as shown by our fluorescent titrations. This is further supported by a reduction in the catalytic efficiency (k_{cat}/K_m) of 2-mer synthesis due to C-tail deletion. The *in vitro* transcriptional profiles showed that all Mtf1 mutants have a defect in synthesizing run-off RNA, although there was a notable decrease in the 2-7-mer abortive products with deletion of the C-tail. Concurrently, there is increased transcript slippage

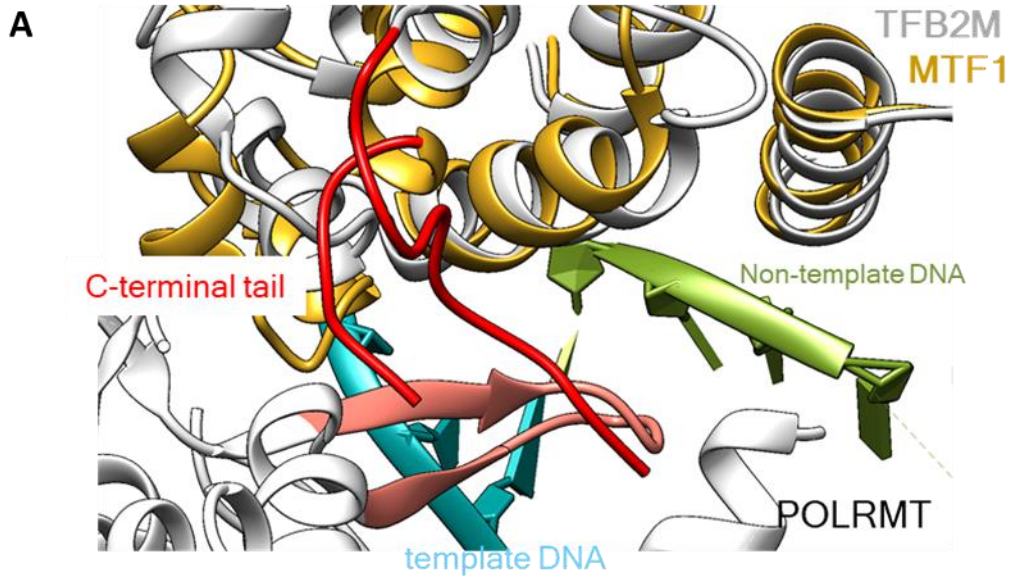
during initiation on deleting the C-tail of Mtf1. We were curious whether the C-tail of TFB2M has similar functions. Our studies show that the template alignment function is conserved between the C-tail of Mtf1 and TFB2M. However, in sharp contrast to Mtf1, the C-tail of TFB2M is needed for open complex formation. The model that emerges from our studies is that the C-tail of Mtf1 plays an essential role in template strand alignment and the C-tail of TFB2M is involved both in promoter melting and template alignment.

E.2 Results

C-tail deletion mutants of Mtf1

To date, there are no crystal structures of the yeast mtRNAP IC. Negative staining and cryo-EM techniques have provided a glimpse into the yeast mtRNAP IC (Drakulic, Wang et al. 2014). However, the recently published crystal structure of the human mtRNAP IC provides a more comprehensive picture of the location of a part of the C-tail (Hillen, Morozov et al. 2017). Overlap of the free structure of TFB2M with the TFB2M in the IC indicates that the C-tail may stabilize the intercalating hairpin that is involved in facilitating open complex formation (Fig. E.2.1.A).

To investigate the role of the C-tails of Mtf1 and TFB2M in transcription initiation, we created a series of Mtf1 and TFB2M deletion mutant proteins from the C-terminal end. Mtf1 with sequential deletions of 4, 12, 16, and 20 amino acids are henceforth referred to as Mtf1-Δ4, Mtf1-Δ12, Mtf1-Δ16, and Mtf1-Δ20, respectively (Fig. E.2.1.B).



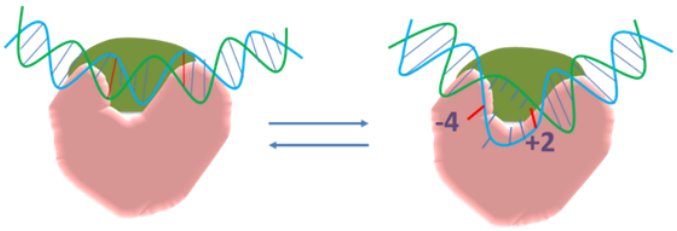
B

Human TFB2M	Δ17 Δ13 Δ7 Δ3 R S K D C A Y K W L Y D E T L E D R
S. cerevisiae MTF1	E W P F K P D I L M D F V D M Y Q T E H S G Δ20 Δ16 Δ12 Δ4

Figure E.2.1: Location of C-tail in the mitochondrial initiation complex and C-tail deletion mutants of Mtf1 and TFB2M. (A) Structure of the human mitochondrial initiation complex (6ero) was overlapped with that of Mtf1 (1I4W). TFAM in 6ero is hidden completely and only parts the template DNA (blue), non-template DNA (green) and POLRMT (white except intercalating β-hairpin which is shown in pink) are shown for clarity. Mtf1 which has the last 16 amino acids missing is shown in yellow. TFB2M which has the last four amino acids missing is depicted in white. The C-terminal region of both is depicted in red. The proximity of the C-tail with the intercalating β-hairpin is evident from the figure. (B) Scheme of TFB2M C-tail deletion: 3, 7, 13, 17 amino acids

were deleted from the C-terminal end of TFB2M. Scheme of Mtf1 C-tail deletion: 4, 12, 16 and 20 amino acids were deleted from the C-terminal end of Mtf1.

A



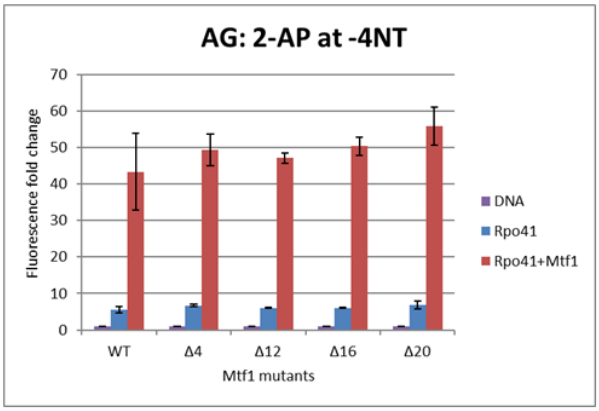
B

21s (−25 TO +20)

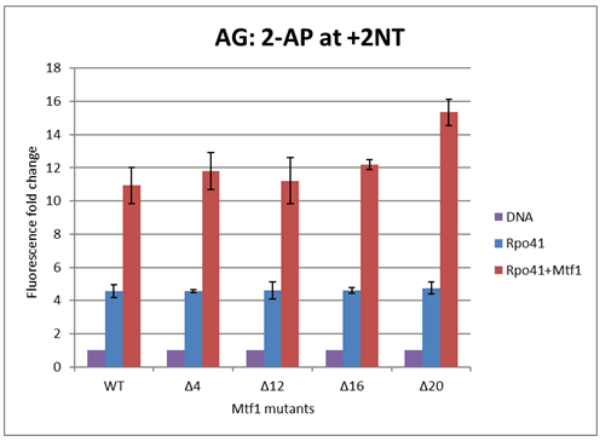
21S_NT: 5′ TATTATTATTATTATATATATAAGTAGTAAAAAGTAGAATAATCT

21S_T: 3′ ATAATAATAATAATATATATATTCATCATTTTTCATCTTATTAGA

C



D



E

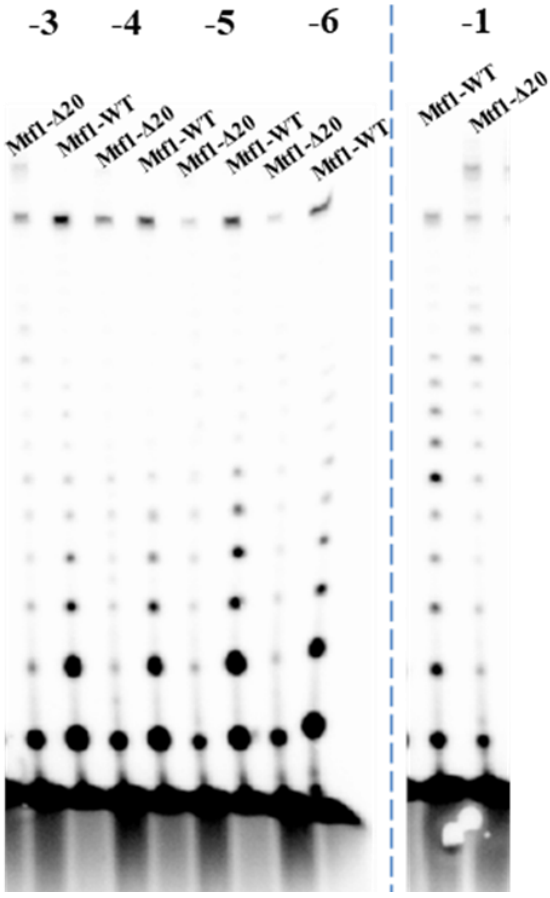


Figure E.2.2: Melting properties of the C-tail deletion mutants of Mtf1. (A) Cartoon showing promoter opening by Rpo41+Mtf1 complex. The -4 and +2 bases are shown in red as they were substituted by 2-aminopurine one at a time. (B) The sequence of the -25 to +20 21S promoter used in promoter melting studies. The -4 and +2 bases are shown in red as they were substituted by 2-aminopurine. The A at the start site is underlined. (C) & (D) Fluorescence intensity of 2-AP in 200 nM duplex DNA modified with 2-AP at −4 non-template (-4NT) and +2 non-template (+2NT). Fluorescence fold change due to DNA addition is normalized as 1. Fluorescence fold change due to the addition of 400 nM Rpo41 (dark blue bar), followed by 400 nM Mtf1-WT or mutants (red bar). Error bars represent the standard deviation from two measurements done on different days. (E) Transcription profile of Mtf1-WtT and Mtf1-Δ20 on 15S promoter with 2-AP substitution. The numbers at the top of the gel indicate the positions which contain the 2-AP substitutions.

Deletion of C-tail of Mtf1 causes no defect in the open complex formation

The fluorescent base analog 2-aminopurine has been used extensively to map the open bubble region of initiation complexes in several RNAPs (Bandwar and Patel 2001, Liu and Martin 2001, Liu and Martin 2002, Tang, Paratkar et al. 2009, Ramachandran, Basu et al. 2017). Previous work from our lab has successfully used this technique to map the extent of the open complex in both yeast and human mtRNAP systems (Fig. E.2.2.A) (Tang, Paratkar et al. 2009, Deshpande, Sultana et al. 2014, Ramachandran, Basu et al. 2017). To test the contribution of the C-tail of Mtf1 in promoter melting, we used the 21s promoter that contains +1+2 AG initiation sequence and introduced a single 2-AP base

either at -4 non-template (NT) or +2 non-template (NT) position (Tang, Paratkar et al. 2009) (Fig. E.2.2.B). The extent of the initiation bubble is from -4 to +2. Hence the two 2AP modifications will monitor upstream and downstream bubble opening steps. The experiments were carried out using 200 nM of 2-AP modified DNA, 400 nM Rpo41, and 400 nM of Mtf1-WT or various Mtf1 C-tail deletion mutants. Fold change in fluorescence intensity of 2AP modified DNA after addition of Rpo41 and Mtf1 was plotted (Fig. E.2.1.C & E.2.1.D). For the -4NT position, the addition of Rpo41 produces a ~5-fold increase in fluorescence as compared to DNA. Subsequent addition of Mtf1-WT produces ~40-fold increase in fluorescence as compared to DNA. All the mutants showed a similar increase in fluorescence as Mtf1-WT (Fig. E.2.1.C). For the +2NT position, the addition of Rpo41 produces a ~5-fold increase in fluorescence as compared to DNA. Subsequent addition of Mtf1-WT produces ~10-fold increase in fluorescence as compared to DNA. All the mutants showed a similar increase in fluorescence as Mtf1-WT (Fig. E.2.1.D). These results indicate that deletion of the C-terminal tail of Mtf1 does not affect promoter melting.

It should be noted that there is a possibility that introducing a mutation such as 2-AP in the DNA may remove the requirement of the C-tail in promoter opening. In the human mtRNAP system, introducing a C:2AP base pair in the -1 position on the Heavy Strand Promoter 1 (HSP1) removes TFAM requirement in transcription initiation (Ramachandran et al., unpublished data). However, the efficiency of transcription of Mtf1-WT and Mtf1-Δ20 remain relatively similar on 2-AP modified vs. unmodified promoters (Fig. Fig. E.2.1.E). Thus, the introduction of the 2-AP mutation does not help

Mtf1-Δ20 overcome the melting defect. We conclude that the C-tail of Mtf1 is not involved in promoter opening.

In the 21s promoter, the +1+2 sequence is AG and incorporating 2-AP in place of the naturally occurring +2G (NT strand) weakens base pairing. The 2-AP:C base pairing is weaker than the 2-AP:T base pairing (Hawkins 2003, Reha-Krantz, Hariharan et al. 2011). To confirm that weakening the +2 base pairing is not affecting our promoter melting results, we conducted similar melting assays with the 15s promoter that contains +1+2AA sequence promoter where the introduction of 2-AP at +2 position forms the usual 2 hydrogen bonds with thymine in the T stand (Hawkins 2003). Again with both the -4NT and +2 NT positions, we observed no defect in promoter melting in any of the Mtf1 C-tail deletion mutants (data not shown). We, therefore, conclude that C-tail of Mtf1 is not involved in promoter melting in either the AA or AG promoter.

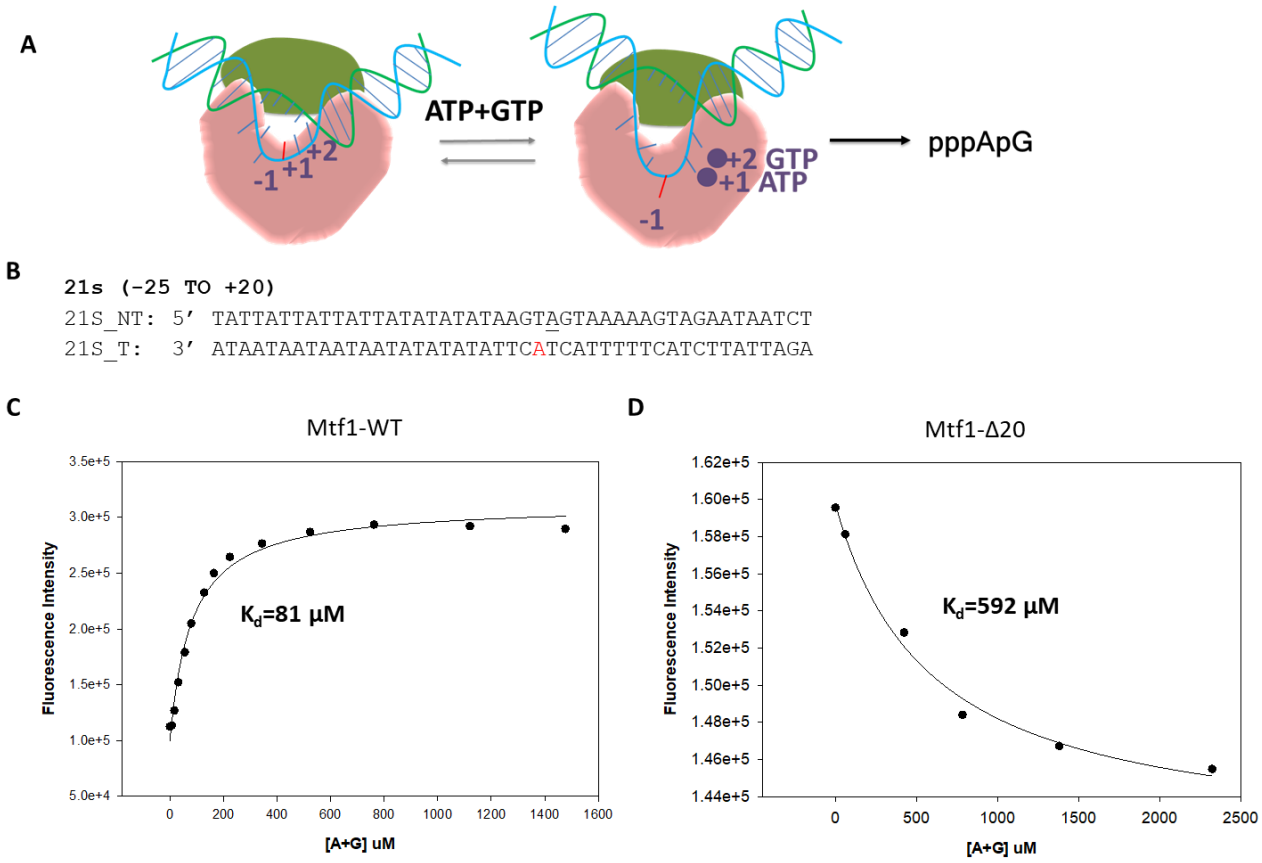


Figure E.2.3: Template alignment properties of Mtf1 C-tail deletion mutant using fluorescence intensity assays. (A) Cartoon showing the scheme of the +1+2 NTP titrations using 2-aminopurine (2-AP) at -1 template as a marker. (B) The sequence of the 21s yeast promoter used for template alignment studies. The -1 base that flips to align template on the addition of +1+2 NTPs was substituted by 2-aminopurine and is marked in red. The A at the start site is underlined. (C) & (D) Determining the K_d values for the +1+2 iNTPs by fluorescence intensity titrations using 2-AP in the -1template position for Mtf1-WT and Mtf1-Δ20. 100 nM duplex promoter DNA was added followed by 200 nM Rpo41+Mtf1. The reaction was titrated with an equimolar mixture of ATP+GTP.

Fluorescence intensity is plotted as a function of increasing +1+2 AG concentration and fit to hyperbola exponential (C) or decay (D) equations obtain the K_d values.

Deletion of C-tail of Mtf1 causes severe defect in iNTPs binding

We were puzzled that C-tail deletions did not affect the promoter melting activity of Mtf1 because a previous report implicated C-tail in promoter melting (Drakulic, Wang et al. 2014). However, this was shown indirectly by comparing iNTPs K_m of Mtf1-WT, and C-tail deleted Mtf1 on duplex vs. bubble promoters. The C-tail deleted Mtf1 had a higher iNTPs K_m on duplex which was rescued by the pre-melted bubble promoter. This result can be explained by a template strand alignment problem which will weaken iNTPs binding affinity. We wanted to directly test the effect of the C-tail deletion on iNTPs binding; hence, we used a fluorescence-based assay to measure the K_d of iNTPs. The position -1 2AP fluorescence is sensitive to ATP binding (Tang, Paratkar et al. 2009) and can be used as a signal to measure iNTPs binding (Fig. E.2.3.A). Experiments were carried out using 200 nM of the 21S promoter with 2-AP at position -1T (Fig. E.2.3.B), 400 nM Rpo41, and 400 nM of various Mtf1-WT or Mtf1-Δ20. To determine the composite K_d of the +1+2 AG iNTPs, we titrated the above mixture with an increasing concentration of an equimolar cocktail of ATP+GTP. Fluorescence intensity increase due to each of these additions was monitored and plotted against ATP+GTP concentration. The curves were fitted using Sigmaplot software into a hyperbolic increase (for Mtf1-WT) or decrease (for Mtf1-Δ20) equations.

Our data show that Mtf1-WT complex binds to +1+2 iNTPs with K_d of 81 μM. The Mtf1-Δ20 complex, on the other hand, has a much weaker binding affinity for iNTPs

with K_d of 592 μM. This indicates an ~8-fold defect in iNTPs binding affinity upon deletion of 20 amino acids of the C-tail of Mtf1 (Fig. E.2.3.C & E.2.3.D).

We conducted similar experiments with the 15s +1+2 AA promoter, and our results show that Mtf1-WT complex has K_d of 65 μM and Mtf1-Δ20 complex has a K_d of 950 μM for ATP (data not shown). Thus, there is a ~15-fold defect in iNTPs binding on the +1+2 AA promoter upon deletion of the C-tail of Mtf1.

Deletion of the C-tail reduces catalytic efficiency of transcription initiation

The above data indicates that the C-tail is involved directly or indirectly in promoting the binding of iNTPs to the Rpo41-Mtf1 complex. A direct role may involve making direct contacts with the iNTPs; however, the crystal structure of the human TFB2M indicates that the C-tail is not likely to be very close to the active site. An indirect role could involve positioning the template strand in the active site of the RNAP, so the +1 and +2 template bases are positioned to bind the iNTPs. The latter is more likely the case.

To test if the iNTPs bound to the complex with Mtf1-Δ20 are catalytically active, we measured the k_{cat} and K_m values, and the catalytic initiation efficiency (k_{cat}/K_m) of 2-mer synthesis. The k_{cat} and K_m measurements were carried out as described in our earlier studies (Deshpande and Patel 2014, Ramachandran, Basu et al. 2017), except here we used the 21s promoter with +1+2 AG initiation sequence, in this way we could measure the catalytic efficiency of ATP and GTP individually. The catalytic efficiency of +1 position ATP was measured by titrating Rpo41-Mtf1-DNA complex with increasing concentrations of +1 ATP at a constant 2 mM +2 GTP. To measure the catalytic

efficiency of +2 position GTP, we titrated with +2 GTP at a constant 2 mM +1 ATP concentration. The resulting 2-mer synthesis data were hyperbolic, and the increase was fit to Michaelis-Menten kinetics.

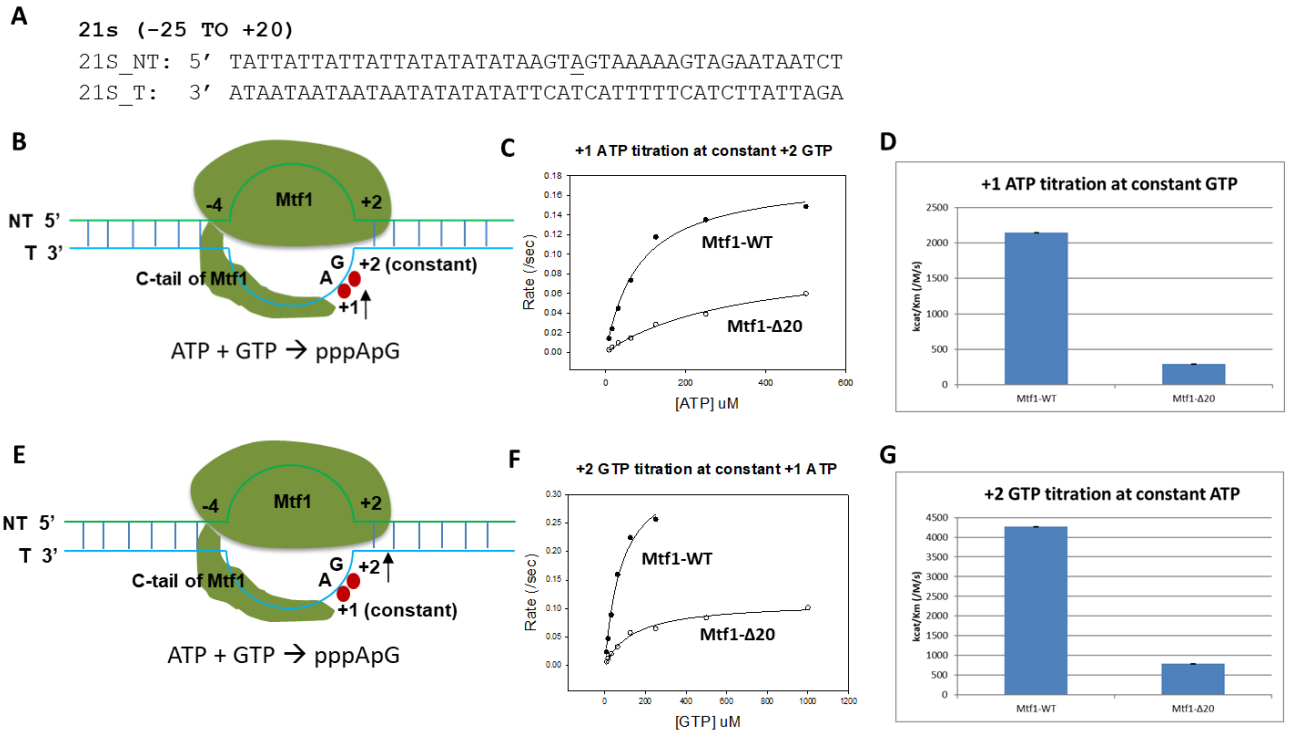


Figure E.2.4: Template alignment properties of Mtf1 C-tail deletion mutants using transcription assays. (A) The sequence of the 21s yeast promoter (-25 to +20) used for template alignment studies. (B) Cartoon is showing a basic scheme of transcription assays to determine the kinetics of +1 position. (C) & (D) Catalytic efficiency (k_{cat}/K_m) values of transcription initiation on the 21S promoter for the Mtf1-WT and Mtf1-Δ20 for +1 position. (E) Cartoon is showing a basic scheme of transcription assays to determine the kinetics of +2 position. (F) & (G) Catalytic efficiency (k_{cat}/K_m) values of transcription initiation on the 21S promoter for the Mtf1-WT and Mtf1-Δ20 for +1 position.

The +1 position ATP binds to the Mtf1-WT complex with K_m of 83 µM and incorporates as a 2-mer with a k_{cat} value of $0.17s^{-1}$, with the resulting catalytic efficiency of 2145 $M^{-1}s^{-1}$. In contrast, Mtf1-Δ20 complex shows +1 ATP K_m of 339 µM and a k_{cat} value of 0.09 s^{-1}, resulting in catalytic efficiency of 290 $M^{-1}s^{-1}$. Thus, there is a ~8-fold defect in catalytic initiation efficiency which stems for a ~4-fold defect in K_m and a ~2-fold defect in k_{cat} (Fig. E.2.4.C, E.2.4.D and Table E.2.1).

The +2 GTP binds to the Mtf1-WT complex with K_m of 83 µM and incorporates with a k_{cat} value of $0.35s^{-1}$, resulting in catalytic efficiency of $4264M^{-1}s^{-1}$. In contrast, Mtf1-Δ20 complex binds to +2 GTP with a K_m value of 142 µM and a k_{cat} value of $0.11s^{-1}$, resulting in catalytic efficiency of $784M^{-1}s^{-1}$. There is a ~6-fold defect in catalytic efficiency, and this results from a ~2-fold defect in K_m and a ~3-fold defect in k_{cat} (Fig. E.2.4.F, E.2.4.G and Table E.2.1).

Deletion of 20 C-terminal amino acids of Mtf1 results in both binding defect of iNTPs at both the +1 and +2 positions and 2-mer synthesis defect. Thus, Mtf1-Δ20 can bind the iNTPs at the catalytic site but greater amounts of iNTPs are needed to catalyze phosphodiester bond formation between the two NTPs. Even when iNTPs are saturating, the lower k_{cat} indicates that the activation energy for bond formation is greater when the C-tail is absent.

The -1+1 matching dinucleotide primer rescues C-tail deletion defect in RNA synthesis

Various RNAPs utilize dinucleotide primers with improved transcription initiation efficiency (Biswas and Getz 1990, Pupov, Kuzin et al. 2014). The yeast mtRNAP can use -1+1 UpA and +1+2 ApA RNA primers for initiation (Biswas and Getz 1990). We hypothesized that the use of -1+1 dinucleotide instead of +1 ATP would rescue the

initiation defect of Mtf1-Δ20 because the extra H-bond with the -1 template base will provide additional binding energy to align the template. We measured the kinetics of transcription initiation for +2 GTP by titrating Rpo41-Mtf1-promoter complex with a mixture of 150 μM of -1+1 UpA (+ γ-[^{32}P] UpA) and increasing GTP. The resulting data were fit to the hyperbolic equation to determine the k_{cat}, K_m and k_{cat}/K_m parameters (Table D.2.1).

The catalytic efficiency of RNA synthesis by the Mtf1-Δ20 complex is 5 times greater than Mtf1-WT when we used UpA as a primer. When we used CpA as a primer that cannot base pair with the -1 template base, the catalytic efficiency of Mtf1-Δ20 decreased and now it was 26 times lower than Mtf1-WT. Based on these results, we conclude that the extra base-pairing energy compensates the defect in transcription initiation by C-tail deletion in Mtf1. This is consistent with the role of C-tail as a placeholder in aligning the template strand in the active site of the RNAP. However, we cannot eliminate the possibility that the C-tail also has direct interactions with the iNTPs.

k_{cat}/K_m ($M^{-1}s^{-1}$)	Mtf1-WT	Mtf1-Δ20
+1A	2150 (±303.8)	291 (±53.9)
+2G	4270 (±2518)	785 (±124)
UpA+G	55 (±40)	274 (±126)
CpA+G	552 ±(±96)	21 (±6.6)
K_m (µM)		
+1A	83.4 (±11.2)	339 (±56.4)
+2G	82.9 (±14)	142 (±20)
UpA+G	54 (±13.2)	1325 (±537.6)
CpA+G	208 (±34.7)	560 (±147.2)
k_{cat} (s^{-1})		
+1A	0.179 (±0.008)	0.0987 (±0.008)
+2G	0.354 (±0.2)	0.112 (±0.008)
UpA+G	0.003 (±0.002)	0.364 (±0.078)
CpA+G	0.115 (±0.006)	0.012 (±0.002)

Table E.2.1: K_m, k_{cat} and k_{cat}/K_m values for 21s promoter for +1 ATP (constant +2 GTP) position, +2 GTP (constant +1 ATP) position, +2 GTP (constant -1+1 matching UA primer) position, +2 GTP (constant -1+1 mismatching CA primer) position.

Mtf1 C-tail deletion mutants make less abortive products

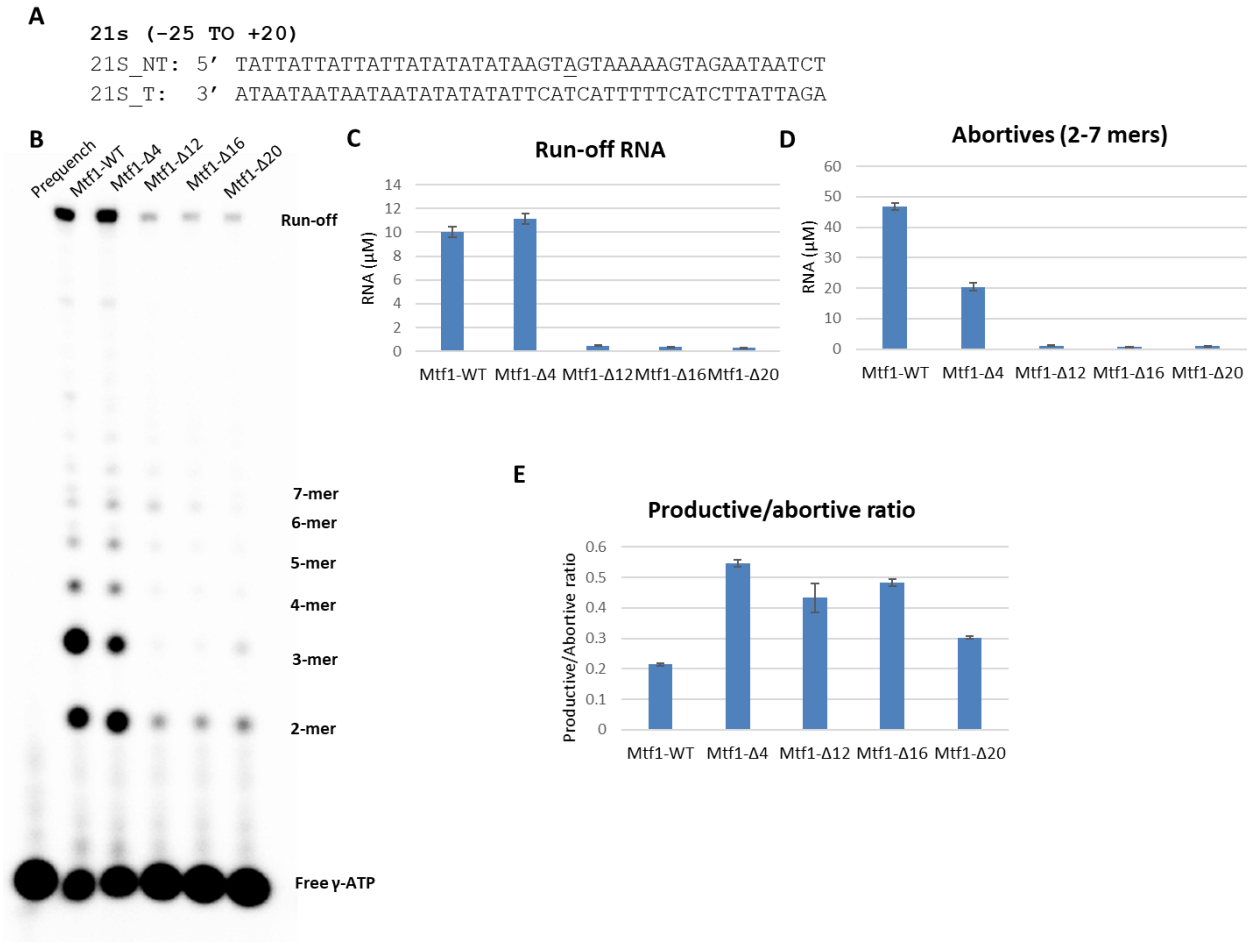


Figure E.2.5: Deletion of the C-tail of Mtf1 causes defect in transcription. (A) The sequence of the 21s yeast promoter used for template alignment studies. (B) Full transcription profile of Mtf1-WT and C-tail mutants showing abortive products and the 19-mer run-off RNA. A pre-formed complex of 2 µM Rpo41, 2 µM Mtf1-WT or Mtf1 C-tail deletion mutants, and 2 µM 21s promoter DNA was pre-incubated and then reacted with 100 µM ATP (+[γ-^{32}P] ATP), UTP, GTP and 3' dCTP at 25°C. Reaction time was

10 mins. (C) (D) & (E) Quantitation of the run-off (C), short abortives (D) and the productive to abortive ratio (E).

Having found a severe defect in catalytic efficiency of 2-mer synthesis on deleting the C-tail, we studied runoff synthesis of Mtf1 mutants using 21s promoter (Fig. E.2.5). In these experiments, 2 µM Rpo41 and 2 µM of Mtf1 or the deletion mutants were pre-incubated with 2 µM promoter DNA (Fig. E.2.5.A), and RNA synthesis was initiated by the addition of 100 µM ATP (+[γ-^{32}P]ATP), UTP, GTP and 3' dCTP at 25°C for 10 minutes reaction times. The inclusion of 3' dCTP helps in preventing non-templated RNA extension and yields a 19-mer run-off RNA product. The RNA products are resolved on 24% sequencing gel to visualize the complete transcription profile (Fig. E.2.5.B).

A typical transcription profile for Rpo41-Mtf1 consists of an initial distributive phase of short abortive products from 2-mer to 7-mer in progressively decreasing amounts, followed by a transition phase of long abortive products from 8-mer to 12-mer, after which it enters into a processive elongation phase that results in the synthesis of the 19-mer run-off RNA product. As we expected, our results show a defect in the formation of run-off products on deleting the C-tail. As shown in Fig. E.2.5.B and E.2.5.C, run-off RNA is 10, 11, 0.4, 0.3, 0.3 µM for Mtf1-WT, Mtf1-Δ4, Mtf1-Δ12, Mtf1-Δ16, and Mtf1-Δ20 respectively. This trend points to a progressive decrease in the amounts of run-off product formed on deleting the C-tail of Mtf1.

We made another interesting observation from the transcription profile: as shown in Figure E.2.5.B and E.2.5.D, we observed that with the deletion of the C-tail there is a decrease in the amount of short 2-7 mer abortives.

The ratio of 19-mer productive run-off RNA to the 2-7 mer abortive denotes the productive to abortive ratio that assesses the efficiency of transition from transcription initiation to elongation. We see that the productive abortive ratio increases with C-tail deletion: all of the C-tail mutants have a higher ratio than Mtf1-WT (Fig. E.2.5.E).

Thus we conclude that the defect in initiation on deleting the C-tail of Mtf1 causes a defect in the formation of run-off products. The unexpected discovery is that there is a significant decrease in the amount of short abortive products on deletion of the C-tail of Mtf1. Interestingly, the decrease in short abortives is reminiscent of the phenotype observed in *E. coli* RNAP upon sigma 3.2 region deletion. Crystal structure of *E. coli* RNAP shows that the Sigma 3.2 loop sterically hinders the growth of the RNA:DNA hybrid and needs to be displaced after 5 to 6 mer RNA synthesis (Hernandez, Hsu et al. 1996, Zhang, Feng et al. 2012). We propose that the C-tail region of Mtf1 plays an equivalent role in the mtRNAP. The placeholder function of C-tail keeps the template aligned for iNTPs binding and 2-mer synthesis, but since the template needs to scrunch from 3-mer onward, the placeholder needs to be disrupted. These opposing needs to keep the template in the +1+2 site and scrunch it for longer RNA synthesis result in abortives. Deletion of the C-tail disrupts the placeholder function but allows the template to scrunch freely and consequently there are less abortive events.

Deletion of the C-tail of Mtf1l increases transcript slippage during initiation

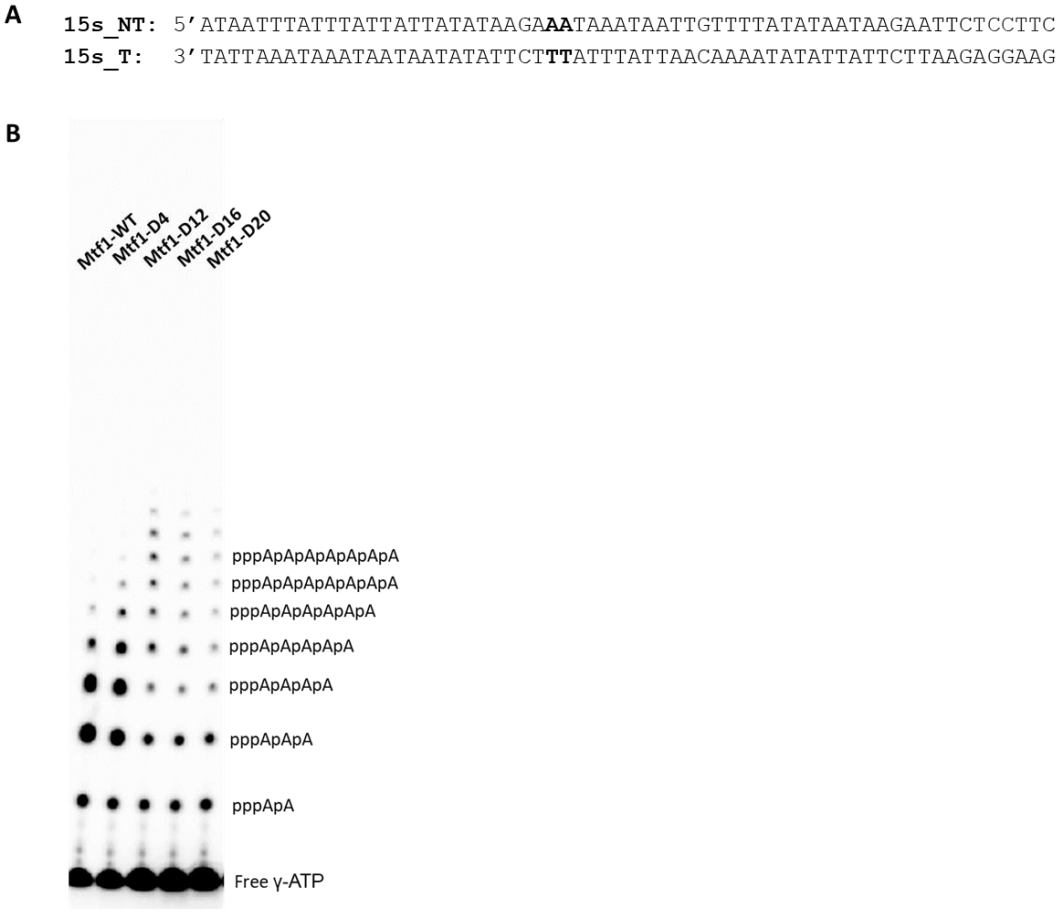


Figure E.2.6: Deletion of the C-tail of Mtf1 more transcript slippage. (A) The sequence of the mutant 15s yeast promoter used for transcript slippage studies. +1+2 bases are in bold. The -1NT position has been mutated to A. (G) Transcript slippage profile of Mtf1-WT and mutants. A pre-formed complex of 1 µM Rpo41, 2 µM Mtf1-WT or Mtf1 C-tail deletion mutants, and 2 µM mutant 15S promoter DNA was pre-incubated and then reacted with 250 µM ATP (+[γ-^{32}P] ATP) at 25°C. Reaction time was 15 mins.

Transcript slippage is observed when the template sequence at the initiation site contains three identical bases in a row. For example, many T7 promoters contain +1GGG, hence T7 RNAP makes a G-ladder as long as 12 nt and longer in the presence of GTP alone (Martin, Muller et al. 1988, Sologub, Litonin et al. 2009). During transcript slippage, the +1 and +2 template bases remain in the active site, and RNA gets longer by slippage. The yeast mt promoters contain +1AA, so we introduced a -1A to create three A's in a row. In this assay, 1 μM Rpo41 and 2 μM of Mtf1 or the deletion mutants were pre-incubated with 2 μM of -1A 15S rRNA promoter duplex (Fig. E.2.6). RNA synthesis was initiated by adding 250 μM ATP + [γ-^{32}P ATP] at 25°C for 30 min. We observed that introduction of -1A in the templating position on the promoter with +1+2 AA resulted in slippage synthesis both in Mtf1-WT and Mtf1 C-tail deletion mutants. However, in comparison to Mtf1-WT, all Mtf1 mutants exhibited a significantly greater tendency for slippage synthesis generating long A-ladders (Fig. E.2.6). This is especially prominent starting from Mtf1-Δ12. Thus, deletion beyond C-terminal 12 amino acids results in enhanced transcript slippage during initiation.

We reasoned that since the Mtf1 C-tail is a placeholder and keeps the template aligned during initiation, it would prevent transcript slippage. On the other hand, C-tail deletion would lead to greater transcript slippage because the growing slipped transcripts are easily accommodated in the active site owing to the missing steric hindrance from the C-tail.

Deletion of the C-tail of TFB2M causes a defect in transcription

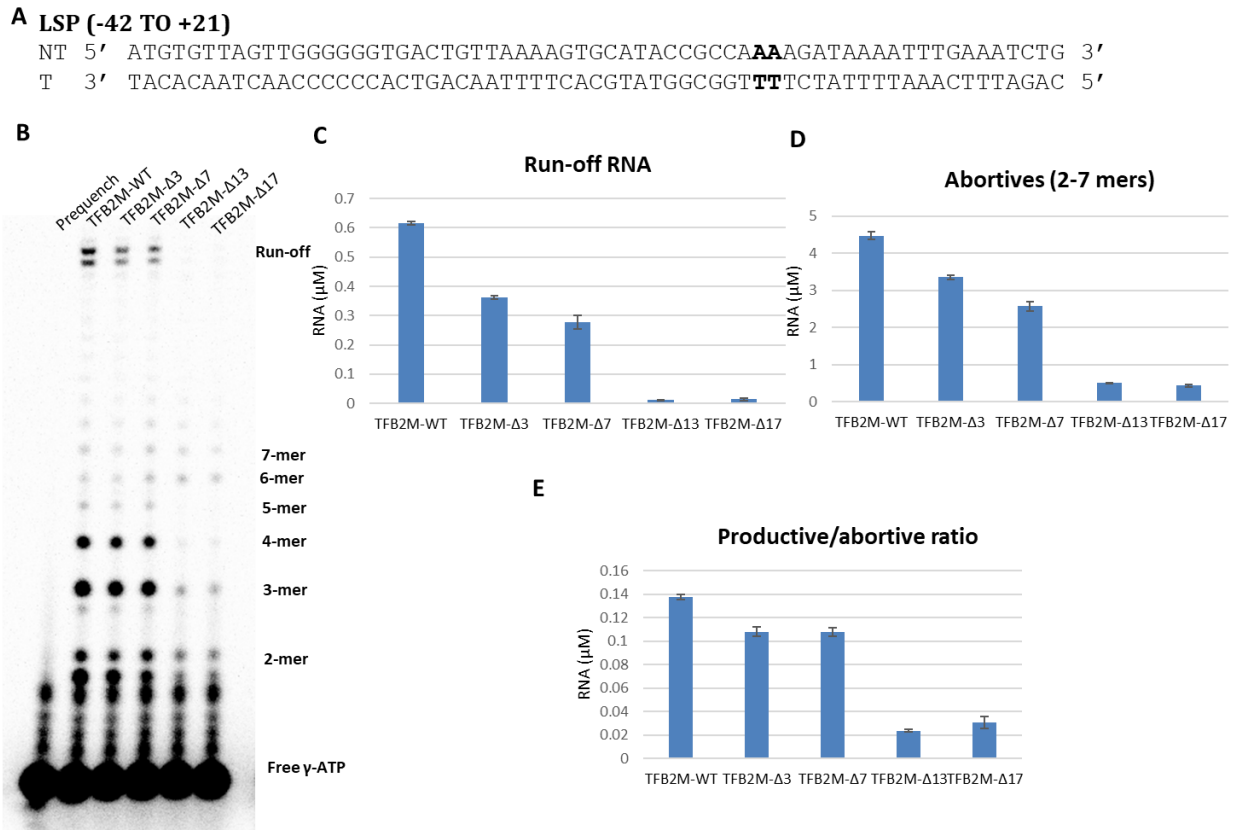


Figure E.2.7: Deletion of the C-tail of TFB2M causes a defect in transcription. (A) The sequence of the LSP human promoter used for transcription assays. (B) Transcription profile of TFB2M-WT and TFB2M C-tail deletion mutants are shown. 1 μM each of POLRMT, TFAM, TFB2M, and LSP DNA was pre-incubated and then reacted with 250 μM ATP (+[γ-^{32}P] ATP), UTP, GTP at 25°C. Reaction time was 10 mins. (C) (D) & (E) Quantitation of the run-off (C), short abortives (D) and the productive to abortive ratio (E).

From our studies of the C-tail deletion mutants of Mtf1, the picture that has emerged so far is that the primary function of the C-tail is a placeholder in facilitating template strand alignment for iNTPs binding. We were curious to find out whether this function is conserved between the C-tails of yeast Mtf1 and human TFB2M. To test this, we made C-tail deletion mutants of TFB2M and named them TFB2M-Δ3, TFB2M-Δ7, TFB2M-Δ13, TFB2M-Δ17 (Fig. E.2.1.B).

We first tested the transcription profile of the TFB2-WT and C-tail deletion mutants. In these experiments, 1 μM each of POLRMT, TFAM, TFB2M or the TFB2M C-tail deletion mutants were pre-incubated with 1 μM LSP DNA (Fig. E.2.4.A), and RNA synthesis was initiated by addition of 250 μM ATP (+[γ-^{32}P] ATP), UTP, GTP at 25°C for 10 minutes reaction time. The RNA products are resolved on 24% sequencing gel to visualize the complete transcription profile (Fig. E.2.7.B).

Our results show a defect in the formation of run-off products on deleting the C-tail of TFB2M. As shown in Fig. E.2.7.B and E.2.7.C, run-off RNA is 0.6, 0.3, 0.2, 0.01, 0.01 μM for TFB2M-WT, TFB2M-Δ3, TFB2M-Δ7, TFB2M-Δ13, TFB2M-Δ17 respectively. This trend points to a progressive decrease in the amounts of run-off product formed on deleting the C-tail of TFB2M very similar to what we saw for Mtf1 C-tail deletion. We observed from Figure E.2.7.B and E.2.7.D, we observed that with the deletion of the C-tail there is a decrease in the amount of short 2-7 mer abortives. However, the productive abortive ratio decreases with C-tail deletion: all of the C-tail mutants have a lower ratio than TFB2M-WT (Fig. E.2.5.E). The latter result is in sharp contrast to Mtf1 where we saw an increase in the productive abortive ratio. Thus, complete deletion of the C-tail of

TFB2M i.e. TFB2M-Δ17 eliminates activity whereas complete deletion of the Mtf1 C-tail i.e. Mtf1-Δ20 does not eliminate transcription initiation activity. We conclude that the C-tail of TFB2M is essential for transcription initiation.

Role of the C-tail of TFB2M in template alignment

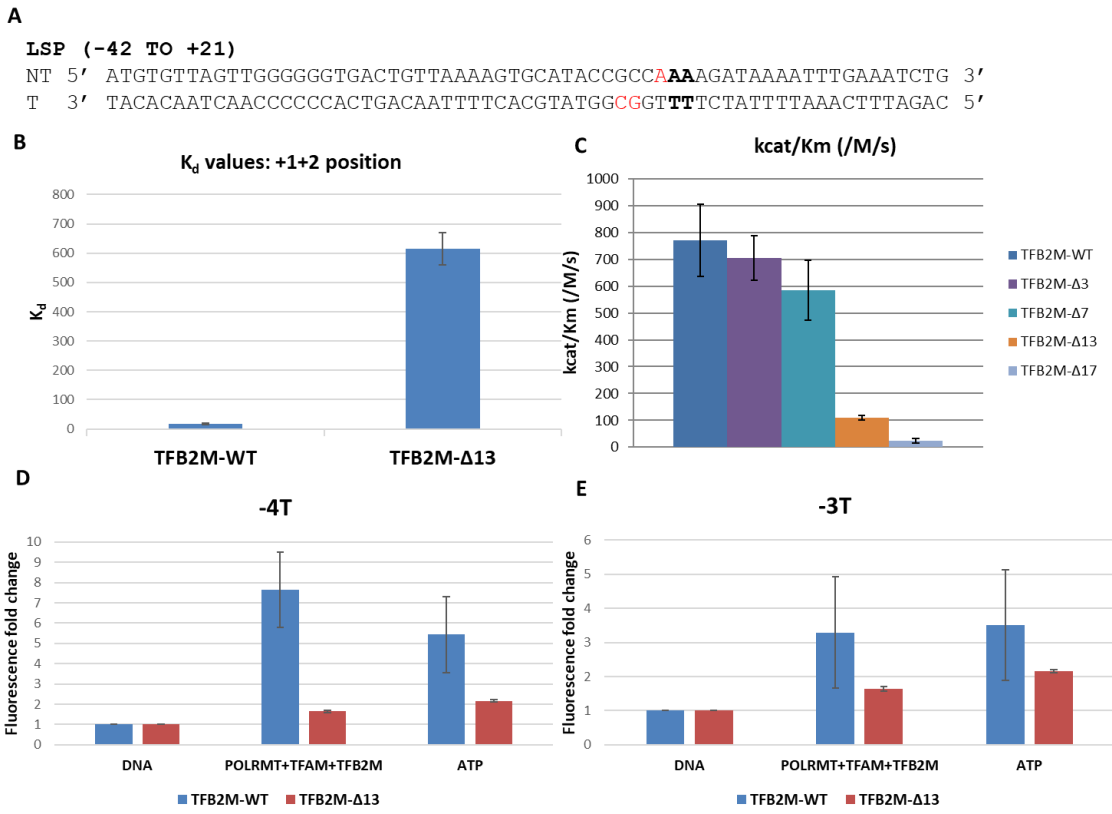


Figure E.2.8: Template alignment and promoter melting properties of the TFB2M deletion mutants. (A) The sequence of the LSP human promoter used for template alignment and melting studies. +1+2 bases are in bold while the -1NT, -3T, -4T positions are marked in red. (B) Determining the composite K_d values for the +1+2 iNTPs by fluorescence titrations using 2-AP in the -1 non-template position in LSP for TFB2M-WT

and TFB2M-Δ13. 100 nM duplex promoter DNA was added followed by 150 nM POLRMT+TFAM+TFB2M or POLRMT+TFAM+TFB2M-Δ13, followed by ATP titration. Average values of the two sets are plotted. (C) Catalytic efficiency (k_{cat}/K_m) values of transcription initiation on the LSP promoter for the TFB2M-WT and TFB2M C-tail deletion mutants. (D) & (E) Fluorescence of 2-AP in 200 nM duplex LSP DNA modified with 2-AP at -4T (D) and (-3T) (E). Fluorescence fold change due to DNA addition is normalized as 1. Fluorescence fold change due to the addition of 200 nM POLRMT+TFAM+TFB2M or POLRMT+TFAM+TFB2M-Δ13 (blue bar), followed by addition of ATP (red bar). Error bars represent the standard deviation from two measurements done on different days.

We measured the iNTPs K_d values to test if the role of C-tail of TFB2M has a similar role as Mtf1 C-tail in template alignment. Our studies have shown previously that, similar to the yeast system, (Tang, Paratkar et al. 2009), 2-AP fluorescence at -1 position in the human system is sensitive to binding of the iNTPs (Ramachandran, Basu et al. 2017) that is reminiscent of yeast mtRNAP system. We replaced the naturally occurring -1 adenine in the NT strand of LSP by 2-AP (Fig. E.2.8.A). We added 100 nM of 2-AP modified DNA to the reaction buffer followed by 200 nM of equimolar POLRMT+TFAM+TFB2M cocktail followed by titration with increasing concentrations of +1+2 ATP. Fluorescence intensity increase due to each of these additions of ATP was monitored and plotted against ATP concentration. The curves were fitted using Sigmaplot software into hyperbola equations. The K_d values for TFB2M-WT and TFB2M-Δ13 are plotted in bar charts (Fig. E.2.8.B). Our data show that the TFB2M-WT complex has K_d of ~15uM for +1+2 ATP. On deleting the C-tail of TFB2M, the K_d of ATP increased to

~576 uM for TFB2M-Δ13. There is a ~40-fold defect in binding of the +1+2 ATP on deletion of the C-tail of TFB2M. These results indicate that the template alignment function is conserved across the C-tails of Mtf1 and TFB2M.

To further validate our results, we measured the catalytic initiation efficiency (k_{cat}/K_m) of initial RNA synthesis in the presence of only +1+2 iNTPs on the LSP promoter. We used 63 bp DNA with consensus LSP -41 to +21 promoter sequence (Fig. E.2.8.A). The initiation sequence contains +1AAA sequence, hence transcription in the presence of ATP results in slippage and poly(A) synthesis. The rate of poly(A) synthesis was measured by titrating a pre-incubated POLRMT+TFAM+TFB2M+DNA complex with increasing concentrations of ATP. The resulting rate versus [ATP] data was fit to the Michaelis-Menten kinetics hyperbola to obtain the kcat/Km.

	k_{cat}/K_m ($M^{-1}s^{-1}$)	**K_m (µM)**	**k_{cat} (/s)**
TFB2M-WT	770 (±133)	318 (±51)	0.24 (±0.01)
TFB2M-Δ3	704 (±82)	455 (±50)	0.32 (±0.01)
TFB2M-Δ7	584 (±110)	1144 (±193)	0.66 (±0.05)
TFB2M-Δ13	109 (±8)	1237 (±82)	0.13 (±0.04)
TFB2M-Δ17	23 (±8)	1716 (±502)	0.04 (±0.0)

Table E.2.2: K_m, k_{cat} and k_{cat}/K_m values for LSP promoter for +1+2 ATP position.

TFB2M-WT shows a catalytic efficiency (k_{cat}/K_m) of 0.77 $M^{-1}s^{-1}$ for pol(A) synthesis with ATP. Successive deletions of the C-tail causes decrease in catalytic efficiency. For TFB2M-Δ3, TFB2M-Δ7, TFB2M-Δ13, TFB2M-Δ17 mutants the catalytic efficiency is 0.7, 0.58, 0.1, 0.02 $M^{-1}s^{-1}$ respectively (Fig. E.2.8.C). This defect in catalytic efficiency mainly stems from a defect in K_m. (Table E.2). These results indicate a similar role of the C-tail of TFB2M towards template alignment as seen for the C-tail of Mtf1.

Role of the C-tail of TFB2M in promoter opening

The deletion of the C-tail of TFB2M has a more severe defect than the deletion of the C-tail of Mtf1. To understand the basis for this result, we measured promoter melting using the 2-AP assay developed earlier (Ramachandran, Basu et al. 2017). To test the promoter melting properties of TFB2M, we introduced 2AP at positions -4T and -3T individually. We added 100 nM of 2-AP modified DNA to the reaction buffer followed by 200 nM of equimolar POLRMT+TFAM+TFB2M cocktail followed by addition of 500 μM ATP. Fluorescence increase due to each addition was monitored. Fluorescence fold change due to DNA addition is normalized as 1. Background fluorescence was corrected as reported before (Ramachandran, Basu et al. 2017).

For -4NT position, the addition of POLRMT+TFAM+TFB2M produces a ~7.5-fold increase in fluorescence as compared to DNA (Fig. E.2.8.D). Addition of POLRMT+TFAM+TFB2M-Δ13 produces a ~1.5-fold increase in fluorescence as compared to DNA. This data points to a ~5-fold defect in promoter melting. Addition of +1+2 ATP does not rescue this defect (Fig. E.2.8.D). For -3NT position, the addition of POLRMT+TFAM+TFB2M produces a ~3-fold increase in fluorescence as compared to DNA (Fig. E.2.8.E). Addition of POLRMT+TFAM+TFB2M-Δ13 produces a ~1.5-fold

increase in fluorescence as compared to DNA. This data points to a ~2-fold defect in promoter melting. Addition of +1+2 ATP does not rescue this defect (Fig. E.2.8.E). These results suggest that deletion of the C-terminal region of TFB2M causes a significant defect in melting the duplex promoter, which contrasts with Mtf1.

E.3 Discussion

Contrasting the structure and function of the yeast and human mt transcription factors Mtf1 and TFB2M

Unlike T7 RNAP, mtRNAPs require one or more transcription factors to catalyze transcription initiation. These factors aid the mtRNAPs in transcription initiation but their role in the preinitiation steps of promoter binding, bending, melting, iNTPs binding, and abortive RNA synthesis is better understood in yeast mtRNAP but not well understood in human mtRNAP. For the yeast mtRNAP system, only one transcription factor Mtf1 is required. Mtf1 helps Rpo41 in binding to promoter DNA, melting of the promoter duplex and stabilizing the open complex. For the human mtRNAP system, two transcription factors TFAM and TFB2M are required, but TFB2M is the homolog of Mtf1 and expected to have similar functions.

Comparing the crystal structures of TFB2M and Mtf1 brings out similarities and differences in the structural features. The N-terminal domains (NTD) of both assume a fold similar to S-adenosyl-methionine-dependent methyltransferases with a central seven-stranded β sheet flanked on either side by three α helices (Schubot, Chen et al. 2001, Martin and McMillan 2002, Hillen, Morozov et al. 2017). One difference is that TFB2M has a prominent loop insertion of unknown function between β3 and α4 regions (Fig. D.5). This extra loop in TFB2M was suggested to enhanced domain movements (Moustafa, Uchida et al. 2015), but since it is in the vicinity of the non-template strand in the initiation bubble region, it could be involved in DNA binding. The C-terminal domain of both these homologs consists of four α helices and an extended C-terminal tail. Density for the C-tail was only observed in one of the two copies in the asymmetric unit

in the TFB2M structure while the other unit lacked six amino acids (Hillen, Morozov et al. 2017). Similarly, 16 amino acids of the C-tail were undefined in the crystal structure of Mtf1 (Schubot, Chen et al. 2001). The lack of the C-tail in the crystal structure indicates flexibility/disordered nature of the C-tail in solution. Disordered regions are often crucial for transcription activators (Dyson and Wright 2005).

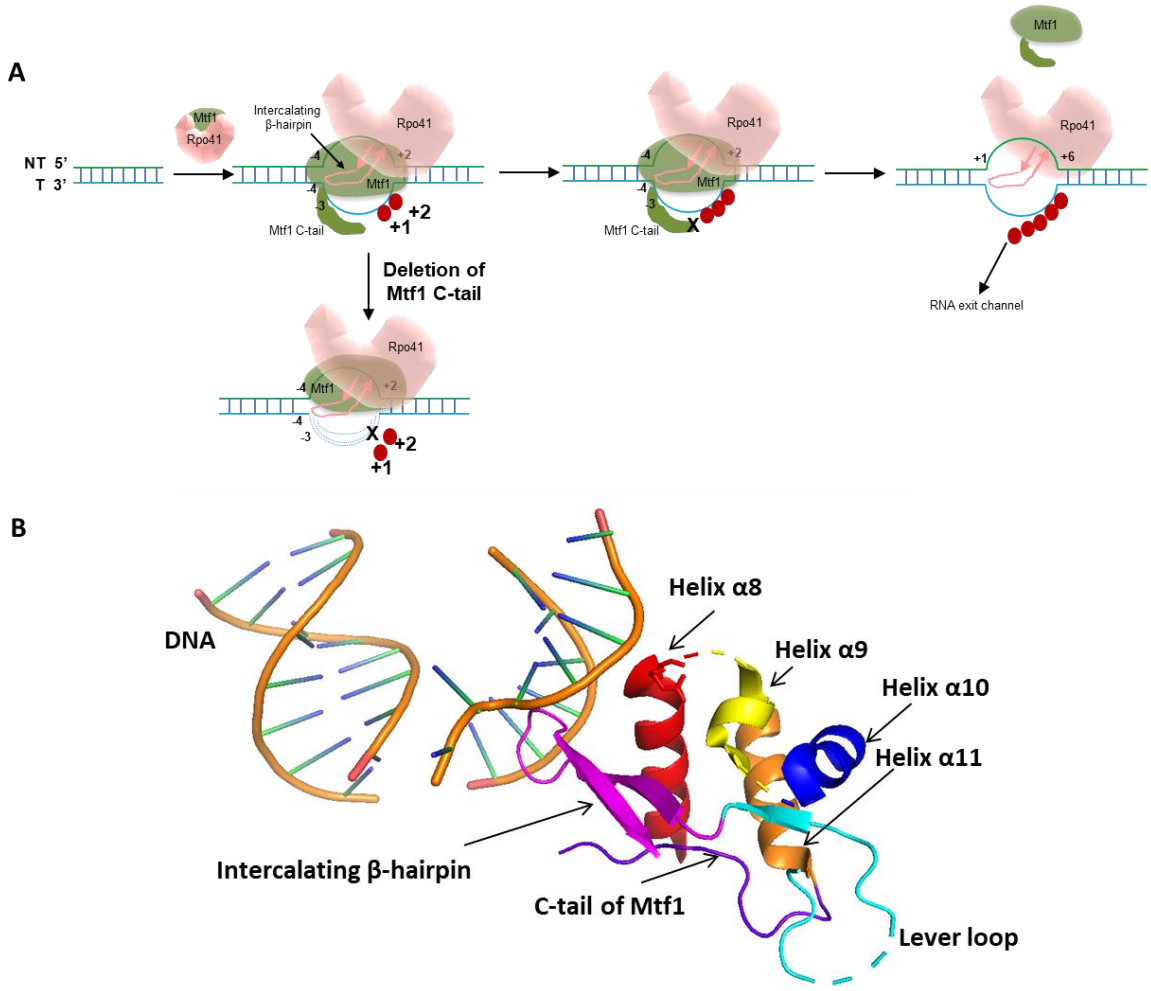


Figure E.3.1: Role of the Mtf1 C-terminal tail in various stages of transcription. (A) Cartoon showing the various events that occur during transcription initiation. The addition of the Rpo41+Mtf1 complex causes melting of the duplex promoter. The intercalating β-hairpin wedges open the duplex while the C-tail of Mtf1 establishes contact in the -3/-4 template region that acts as an anchor to align the template to

facilitate iNTPs binding. Lack of the C-tail of Mtf1 causes misalignment of the template strand that hampers iNTPs binding. The growing RNA chain causes a collision with the Mtf1 C-tail that destabilizes the C-tail interactions with the -3/-4 template region. Loss of these stabilizing interactions probably causes repositioning or ejection of Mtf1 from the initiation complex and subsequent transition into elongation. (B) Crystal structure is showing the LSP promoter around the transcription start site. The C-tail of Mtf1 is shown in purple, the intercalating β-hairpin of POLRMT is shown in magenta, the lever loop is shown in cyan, the helices α8, α9, α10, α11 are shown in red, yellow, blue and orange respectively.

The role of the C-tail in promoter melting, bending and template alignment

In this project, my goal was to investigate the role of the C-tail of Mtf1 and TFB2M. When I began these studies, a report had suggested that the C-tail of Mtf1 is important for promoter melting (Drakulic, Wang et al. 2014). However, this was shown indirectly by comparing activities of Mtf1-WT and C-tail deleted Mtf1 on duplex vs. bubble promoters. Using direct promoter melting assays based on 2AP fluorescence change, we found that the C-tail of Mtf1 is not needed for promoter melting. On the other hand, the C-tail of TFB2M has a significant contribution to promoter melting.

The structural basis for the different roles of the C-tails in promoter melting is not known. The human mt RNAP IC structure indicates that helix α8 of TFB2M positions the β-hairpin between the DNA strands to induce promoter melting. The C-tail of TFB2M also moves relative to its conformation without POLRMT to a new position to stabilize the β-hairpin (Hillen, Morozov et al. 2017), which may explain the role of C-tail of TFB2M in promoter melting. The structure of yeast mtRNAP IC is not known; however, assuming a

similar mechanism, it appears that the helix α8 of Mtf1 is enough to buttress the intercalating β-hairpin and catalyze promoter melting without assistance from the C-tail.

The need for a separate/dedicated element for template alignment

Melting of the promoter DNA strands close to the catalytic active site and accurate positioning of the single-stranded template at positions +1 and +2 for base pairing with incoming ribonucleoside triphosphates are the key steps in transcription initiation. Different strategies have been adopted across RNAPs for precise execution of these steps.

T7 RNAP is studied as a prototype of all single-subunit RNAPs. Several elements in the N-terminal domain (NTD) and fingers domain of T7 RNAP bring about promoter melting and subsequent template alignment. T7 RNAP also contains the intercalating β-hairpin that inserts a Val residue, replacing the position -4 base pair, to stack on base pair -5 and stabilize the upstream end of the bubble in the initiation complex. The melted promoter is further supported by sequence-independent interactions between the phosphodiester backbone of the single-stranded T strand and residues from the specificity loop (from the fingers domain) (Cheetham and Steitz 1999). Template alignment is mainly brought about by the palm domain: the aromatic side chain of a Trp residue stacks on the -1 base which induces a sharp bend in the T strand to align the +1 and +2 bases for catalysis. Four protein side chains of the specificity loop also interact with the upstream -7, -8, -9 base pairs to anchor the promoter thereby contributing to a template alignment. Thus, we see there is no single dedicated element but several elements contributing to template alignment in this system. An extensive promoter sequence and consequent extensive base-specific interactions with the RNAP is instrumental in promoter melting and alignment.

We looked into the structures of POLRMT IC to determine why transcription factors are needed for the function of melting and template alignment. In the apo structure of POLRMT, the fingers domain is rotated by 25° around the O helix, the latter also being tilted. These deviations cause a 15Å translation of the Y helix which makes the active site inaccessible by the NTPs make catalysis difficult (Hillen, Cheetham Nature). This clenched conformation has the intercalating β-hairpin (in NTD) in a position that is incompatible towards DNA melting. TFB2M binds around the -7 NT strand (similar upstream interactions as provided by the specificity loop of T7 RNAP) which triggers a conformational change in the form of rotation of PBD of POLRMT which repositions intercalating β-hairpin between the DNA strands triggering promoter melting. The intercalating β-hairpin of TFB2M is also supported by the helix α8 and C-tail. As observed from the human mt IC structure and also our alignment of T7 RNAP, Rpo41, and POLRMT, there is an insertion of ~15 amino acids ahead of the intercalating β-hairpin called lever loop. This loop interacts with the helix α9 and α10 of TFB2M further supporting the rotated POLRMT (Morozov, Agaronyan et al. 2014, Hillen, Morozov et al. 2017).

Unlike T7 promoter, the human promoters LSP and HSP1 do not have any sequence conservation in the promoter region. Thus we see that the transcription factor TFB2M is required to provide the upstream interactions that probably maintain the upstream edge of the bubble open. The stabilizing interactions by TFB2M CTD and C-tail with the added lever loop and the intercalating β-hairpin allow POLRMT to transition from the apo to IC structure. For yeast, there is a conserved -8 to +1 nona-nucleotide sequence. Thus, it is possible that similar to T7, the specificity loop of Rpo41 is involved in base-specific

interactions with the conserved -7/-8 region, which stabilizes the CTD interactions with the lever loop and beta hairpin and C-tail is not needed for this purpose.

However, the template alignment function of the C-tail is conserved between Mtf1 and TFB2M. For Mtf1, the C-tail was shown to be cross-linked to the -3/-4 T strand. Thus the C-tail through these and possibly more interactions act as a place-holder assuring proper template alignment that facilitates iNTPs binding. From the human mtRNAP IC, we do not have a resolution of the T strand downstream of the -2 position and the C-tail of TFB2M for the last 4 amino acids. But, similar to Mtf1, the C-tail of TFB2M can act as a placeholder for template alignment.

A conserved element involved in template alignment across RNAPs

Template alignment is a key step in transcription initiation both for efficient iNTPs binding and for correct start site. Therefore, an element dedicated to template alignment can be identified by examining other RNAPs. The *E. coli* RNAP is extensively studied both biochemically and structurally with the availability of several different crystal structures that provide valuable snapshots of various steps during transcription initiation (Murakami, Masuda et al. 2002, Zhang, Feng et al. 2012, Feng, Zhang et al. 2016, Boyaci, Chen et al. 2019). A conserved hairpin-like loop region called Sigma 3.2 has been shown to protrude into the active center of RNAP in the IC. In the initiation complex, it occupies part of the region that is occupied by the RNA in the elongation complex. It has base edge-edge interactions with the -4T strand and H-bonding interactions with the -3/-4 T strand. These interactions between the Sigma 3.2 and the template strand positions the latter deep into the active site of the RNAP that facilitates iNTPs binding. This structural prediction has been corroborated by biochemical studies

that show that deletion of sigma 3.2 region causes an increase in K_m of iNTPs (Murakami, Masuda et al. 2002, Kulbachinskiy and Mustaev 2006, Pupov, Kuzin et al. 2014). These observations point to a role in template alignment. The structure also predicts that the interactions between the Sigma 3.2 and the template strand have to be disrupted and the Sigma 3.2 region has to be displaced to generate longer RNA products pointing to the role of sigma 3.2 regions in abortive initiation. This was confirmed by biochemical studies where deletion of sigma 3.2 region decreases abortive products (Hernandez, Hsu et al. 1996, Murakami, Masuda et al. 2002, Kulbachinskiy and Mustaev 2006, Pupov, Kuzin et al. 2014). This loss in interaction prepares the RNAP for subsequent steps of sigma ejection and transition to elongation (Basu, Warner et al. 2014). However contrary to some biochemical data that show cross-linking between the Sigma 3.2 region and the iNTPs (Kulbachinskiy and Mustaev 2006), from the structure it is apparent that the Sigma 3.2 region does not have any direct interactions with the iNTPs (Zhang, Feng et al. 2012). A functionally homologous region to sigma 3.2 in the RNA Pol II system is the B-reader loop. This region is also involved in template alignment and abortive initiation (Kostrewa, Zeller et al. 2009, Sainsbury, Niesser et al. 2013).

Final model: Melting, Template alignment & abortive initiation

From all our biochemical data and comparative structural analysis with T7 RNAP, the model for C-tail function that emerges is as follows: The C-tail of TFB2M is required for promoter melting, but the C-tail of Mtf1 is not needed for promoter melting. The interactions of the C-tail with the β-hairpin that keep the initiation bubble open and other unknown elements around the transcription start site anchor the template strand and align to position it deep into the active site for iNTPs binding. As transcription initiation

progresses, the two processes of nascent RNA growth and template alignment by C-tail interactions oppose each other. The growing RNA chain requires realignment of the template which is opposed by the C-tail template interactions, and this conflict results in abortive events during initial transcription. We speculate that like the sigma 3.2 region, the growing RNA eventually pushes the C-tail out of the active site to trigger the subsequent steps of Mtf1 ejection or its repositioning during the transition into elongation (Fig. E.2.6).

E.4 Future Directions

Our studies along with the crystal structure of the human mt IC (Hillen, Morozov et al. 2017) have brought to light the possible roles of the C-tail of Mtf1/TFB2M in transcription initiation. Several interesting and as yet unanswered questions arise from these studies. Some of them are listed below.

1. Function of the N-tail of TFB2M: Alignment of the sequences of Mtf1 and TFB2M shows the presence of an additional ~60 amino acids that are present in the N-terminal end of TFB2M but not seen in Mtf1 (Schubot, Chen et al. 2001, Sologub, Litonin et al. 2009, Hillen, Morozov et al. 2017). Cross-linking studies have shown that this N-tail of TFB2M interacts with +1+2 iNTPs and also with the promoter DNA from the +1 to +3 regions (Sologub, Litonin et al. 2009). This gives rise to the possibility that the N-tail modulates the iNTP binding, a function we found to be executed by the C-tail. Now it is possible that either of these N or C-tail directly interacts with the iNTPs while the other has an allosteric effect. It would be interesting to address the functions of the N-tail of TFB2M.

2. The crystal structure of the human mt IC suggests that the C-tail of TFB2M promotes promoter melting (Hillen, Morozov et al. 2017). Our 2-AP assays confirm this data. Using stop-flow studies, it would be interesting to study the rates of opening to further study the role of the C-tail of TFB2M in the opening.

3. Comparison of the crystal structures of TFB2M in the apo form and bound to POLRMT and DNA in the human mt IC suggests that C-tail of TFB2M must be

repositioned to avoid a clash with the NT stand of the promoter DNA. From this observation, we hypothesize that the C-tail might be binding to the NT DNA binding sites of TFB2M. Thus, binding of the C-tail and NT strand to this cleft might be a mutually exclusive process. DNA binding properties of Mtf1 and TFB2M wild-type and C-tail deletion mutants need to be tested to test this 'auto-inhibition hypothesis.'

[I] SUMMARY

The central theme of my thesis is the biochemical and biophysical characterization of the mechanisms of mt transcription initiation. Although there has been a recent interest in research on human mitochondrial biology, we are severely limited in our knowledge of the mechanistic aspects of transcription initiation. My thesis has aimed to fill some of this gap in the field.

In the first section of the thesis, I have studied mt 5' NAD^+ capping. In this study, we have shown that both *S. cerevisiae* and human mtRNAPs are capable of *in vitro* capping with both NAD^+ and NADH. This capping is far more efficient than bacterial and eukaryotic nuclear RNAPs. We further show that capping efficiency is determined by sequences at and around the transcription start site. Our *in vivo* data shows that a high percentage of transcripts are capped with NAD^+ and NADH and these values are sensitive to the cellular NAD^+ and NADH levels.

In the second section of the thesis, I have studied the functions of the unresolved C-terminal regions of the yeast mt transcription factor Mtf1 and its human homolog TFB2M in transcription initiation. We identified a function of the C-tail that is conserved across yeast and human: template alignment. The C-tail acts as a place-holder and stabilizes the promoter DNA that facilitates binding of the iNTPs. We also identified a function of the C-tail that is not conserved across yeast and human: promoter melting. The C-tail of Mtf1 is not necessary for promoter melting while the C-tail of TFB2M has significant contributions to promoter melting. Our results point to the C-terminal region of Mtf1and TFB2M being involved in regulating transcription initiation in yeast mitochondria.

[J] APPENDIX

I. Expression of Bpa labeled Mtf1

Expression of Mtf1 containing unnatural amino acid unnatural amino acid p-benzoyl-L-phenylalanine (pBPA), involved the following sequence of experiments.

1. Site-directed mutagenesis of Mtf1, Plasmid purification, and Sequencing:

   As described in the Methods section, the QuikChange® Site-Directed Mutagenesis Kit was used to introduce TAG stop codon at the terminal 24 codons of the C-tail region of Mtf1. Plasmids were purified and sequenced to confirm the introduction of the TAG stop codon at correct positions.

2. Preparation of C321ΔA competent cells:

   A single colony of *E. Coli* (all known TAG stop codons in Escherichia coli MG1655 were replaced with synonymous TAA codons, which permitted the deletion of release factor 1 and reassignment of UAG translation function. The strain was purchased by the Nickels lab from Addgene (Bacterial strain #49018) was picked from a plate and inoculated in a 100 mL culture overnight at 30°C. The next morning, a secondary culture with a 1:100 dilution was started. The cells were grown until O.D. 0.4, chilled for 15 minutes, and then spun down. The pellets were suspended in 50 mL of autoclaved water and then stored in 4°C overnight. The next day, the cells were spun down, resuspended in 50 mL of 10% glycerol solution, and stored overnight at 4°C. The next day, the cells were again spun down, suspended in 15 mL 10% glycerol solution, and stored overnight at

4°C. The next morning, cells were centrifuged and resuspended in 10% glycerol and then aliquoted and stored at -80°C.

3. Co-transformation of mutant plasmid and pBPA plasmid:

   C321ΔA cells were co-transformed with ptrc-His plasmid containing mutant Mtf1 gene (ampicillin selection) and pEVOL plasmid containing the genes for the aminoacyl-tRNA synthetase and tRNA pair which are required for incorporation of pBPA (chloramphenicol selection). The cell line requires electroporation for the uptake of the plasmid. Cells were pulsed for 2-3 seconds and then immediately recovered with LB, and then set to shaking at 37°C and 250 rpm. Each mutant was plated on 3 ampicillin + chloramphenicol plates and grown overnight at 37°C.

4. Cell growth and Lysis:

   0.1 M pBPA solution (Bachem Product #4017646.0005) was added to a total volume of LB, and the resulting solution was titrated to a pH of 7.3 using concentrated HCl. Chloramphenicol and ampicillin were added to the media, which was aliquoted into separate flasks for each mutant. Plates of bacteria containing different mutants were suspended in media and transferred into separate falcons. Each culture was therefore started at an O.D. of 0.3 and simultaneously induced with 1 mM arabinose to amplify the amount of tRNA synthetase/tRNA in each cell. At an O.D. of 0.6-0.8, 1 mM IPTG was added to each culture to induce Mtf1 gene. Cells were grown overnight at 16°C. On the next day, cells were spun down at 5000 rpm for 30 minutes. Induction was checked to confirm protein expression. Each cell pellet was lysed with 5 mL B-

PER, 50 µl of DnaseI, and lysozyme (1mg/mL) for 30 minutes at room temperature. Then it was dissolved in 20 mL wash buffer (50 mM NaP buffer (pH 8), 300 mM NaCl, 10 % glycerol, 20 mM imidazole, ddH2O). After incubation, 20 mL of wash buffer was added to lysate before spinning down at 16.5K rpm for 45 minutes at 4°C.

5. Nickel-NTA Chromatography:

   Each column was loaded with nickel ion beads and equilibrated with water and then wash buffer (50 mM NaP buffer (pH 8), 300 mM NaCl, 10 % glycerol, 20 mM imidazole, ddH_2O). The lysate was filtered to remove large impurities and was subsequently loaded on the column. After lysate flowed through, the column was washed with wash buffer. Protein was eluted with elution buffer containing 500 mM imidazole. The concentration of each protein was determined as described in the Methods section.

II. Detailed purification of POLRMT (human mtRNAP)

Human mtRNAP (43-1230) was cloned in pPROEXHTb vector by Genscript. POLRMT was prepared from BL21 codon plus (RIL) cells transformed with pPROEXHTb-POLRMT (43-1230)-6xHis. Single colonies were picked and were grown at 37°C in 5mL culture with the addition of ampicillin for ~6 hours. 200 µL culture was transferred to 200 mL media containing ampicillin and grown overnight at 30°C. 10mL inoculum per liter LB media was added, and cells were grown in at 37°C until O.D. reached 0.6-0.8, when they were induced by addition of 1 mM Isopropyl β-D-1-thiogalactopyranoside (IPTG). Induction was carried

out at 16°C for 16 hours. Cells were lysed in the presence of protease inhibitor and lysozyme, followed by polyethyleneimine treatment and ammonium sulfate precipitation.

DEAE sepharose column was attached in tandem with Ni-sepharose column, and the lysate was loaded overnight. The DEAE sepharose column was detached, and the Ni-sepharose column was washed (20 mM imidazole in wash buffer). Elution was done over 200 mL gradient between 0-50% (20 mM imidazole in wash buffer and 500 mM imidazole in elution buffer). The eluent was loaded into Heparin-sepharose column overnight. The Heparin-sepharose column was washed (150 mM NaCl in wash buffer). Elution was done over 100 mL gradient between 0-50% (150 mM NaCl in wash buffer and 1 M NaCl in elution buffer). The eluent was concentrated using amicon filters, dialyzed and stored at -80°C.

III. Detailed purification of TFAM

TFAM was prepared from BL21 codon plus (RIL) cells transformed with pPROEXHTb-TFAM (43-245)-6xHis. Separate colonies were picked and added to 24 culture tubes containing 10 ml of LB media. The cells were grown at 37°C until O.D. reached 0.5 to 0.8, before transferring the culture from 2 tubes into 1 L LB, for a total of 12 flasks containing cell culture. The 1 L cultures were grown to an OD of 0.6-0.8 before 1-ml pre-induction samples were taken and stored in -20°C. The induced cells grew for another ~4 hours before post-induction samples were taken. 4-20% gel was run at this point to check the induction of the TFAM protein. The cells were spun at 5000 rpm for 20 minutes. Cells were lysed in the

presence of protease inhibitor and lysozyme, followed by polyethyleneimine treatment and ammonium sulfate precipitation. The cells were then sonicated for 3 minutes with 30 seconds on/off pulses (4 pulses that were 2 minutes apart). The lysate was then centrifuged at 16500 rpm for 1.5 hours and filtered.

The lysate was loaded overnight on Ni-sepharose column. The Ni-sepharose column was washed (20 mM imidazole in wash buffer). Elution was done over 140 mL gradient between 0-60% (20 mM imidazole in wash buffer and 500 mM imidazole in elution buffer). Protein started eluting at approximately 10-25% (68-140 mM imidazole), with a peak at about 15%. The eluent was loaded into Heparin-sepharose column overnight. The Heparin-sepharose column was washed (150 mM NaCl in wash buffer). Elution was done over 100 mL gradient between 0-50% (150 mM NaCl in wash buffer and 1 M NaCl in elution buffer). TEV protease was added to the eluent in 1:30 ratio and the sample was added to a SnakeSkin dialysis bag and submerged in dialysis buffer for 9-10 hours. The solution was collected and loaded on 5ml reverse nickel ion column and 5ml heparin column arranged in tandem. The Ni-sepharose column was removed, and the heparin column was washed (200 mM NaCl in wash buffer). Elution was done over 80 mL gradient between 0-50% (200 mM NaCl in wash buffer and 1.5 M NaCl in elution buffer). Protein started eluting in various fractions (30%, 40%, and 60%). The three peaks were pooled separately, buffer exchanged and concentrated using amicon filters, dialyzed and stored at -80°C.

IV. Detailed purification of TFB2M

Human TFB2M (21-396) was purified from *E. coli* strain ArticExpress (DE3) (Stratagene) transformed with pT7TEV-HMBP4 (gift of Professor Miguel Garcia-Diaz at the Stony Brook University School of Medicine). Single colonies were picked and added to 200 mL primary culture containing 50ug/ml Kan were grown overnight at 37°C. 16ml primary culture was added to per 1L secondary culture (without antibiotic) and grown at 27°C till O.D. reaches 0.7-0.8. Once in the range, cells were induced with 0.2 mM IPTG for 20 hrs at 14°C.

Cells were lysed in the presence of high salt (1 M KCl), protease inhibitor and lysozyme and loaded on Ni-sepharose column. The column was washed (20 mM imidazole in wash buffer). Elution was done over 140 mL gradient between 0-60% (20 mM imidazole in wash buffer and 500 mM imidazole in elution buffer). TEV protease was added to the eluent in 1:50 ratio and the sample was added to a beaker and left ~15 hours. Proteolysis was checked on a gel. The eluent was loaded into Heparin-sepharose column. The Heparin-sepharose column was washed (150 mM KCl in wash buffer). Elution was done over 80 mL gradient between 0-50% (200 mM KCl in wash buffer and 1.5 M KCl in elution buffer). The eluent was loaded into the SP column. The SP column was washed (150 mM KCl in wash buffer). Elution was done over 30 mL gradient between 0-75% (200 mM KCl in wash buffer and 1.5 M KCl in elution buffer). The eluent was

V. Sequence alignment of T7 RNAP, Rpo41, POLRMT

```
CLUSTAL format alignment by MAFFT FFT-NS-2 (v6.864b): T7/Human/Yeast
sp|P00573|RPOL_ ------------------------------------------------------------
sp|O00411|RPOM_ MSALCWGRGAAGLKRALRPCGRPGLPGKEGTAGGVCGPRRSSSASPQEQDQDRRKDWGHV
sp|P13433|RPOM_ ---------------MLRP----------------------AYKSLVKTSLLQRR------


sp|P00573|RPOL_ -----------------MNTINIAKNDFS-----------------------------
sp|O00411|RPOM_ ELLEVLQARVRQLQAESVSEVVVNRVDVARLPECGSGDGSLQPPRKVQMGAKDATPVPCG
sp|P13433|RPOM_ -LISSKGSKLFKPSPDSTSTILISEDPLV----TGS-----SPTSSTTSGIISSNDFP--
                  . : : .  .

sp|P00573|RPOL_ ------------------------------------------------------------
sp|O00411|RPOM_ RWAKILEKDKRTQQMRMQRLKAKLQMPFQSGEFKALTRRLQVEPRLLSKQMAGCLEDCTR
sp|P13433|RPOM_ ----LFNKNRK---------DAKSSISYQ-----------------------------


sp|P00573|RPOL_ ----------------------DIELAAIPFNT--------------------------
sp|O00411|RPOM_ QAPESPWEEQLARLLQEAPGKLSLDVEQAP-SGQHSQAQLSGQQQR----LLAFFKCCLL
sp|P13433|RPOM_ ------WKN---------PSELEFD----PFNKSHASAVTSMTRTRDVMQLWSLLEACLQ
                        .::    * .

sp|P00573|RPOL_ --------------LADHYGERLAREQLALE-HESYEMGEAR----FR-----------
sp|O00411|RPOM_ TDQLPLAHHL-----LVVHHGQRQKRKLLTLDMYNAVMLGWARQGA-FKELVYVLFMVKD
sp|P13433|RPOM_ SNLMKRAFSILESLYLVPEHKQR------FIEDYNMYLNSFSKNDPNFP----ILKM--N
                  *. .: :*      :: ::    . ::    *

sp|P00573|RPOL_ -------------------------------KMFERQLKAGEVADNAAAKPLITTLLPK
sp|O00411|RPOM_ AGLTPDLLSYAAALQCMGRQDQDAGTIERCLEQMSQEGLKLQALFTAVLLSEEDRATVLK
sp|P13433|RPOM_ EKLTNDLET----------SFKDVNYNDKTLAIMIHHALNFHSTTSSMLLKPIISAYLKM
                                  * .. *:         .    : :

sp|P00573|RPOL_ MIARINDWFE-------------------------------------------------
sp|O00411|RPOM_ AVHKVKPTFS--------------------LPPQL--------------PPPVNT-----
sp|P13433|RPOM_ SVNGIREIFSCLDILTISDLNILMNDLKVITPSQLPNSVRPILESLTLSPTPVNNIENEE
           : :.  *.

sp|P00573|RPOL_ ---------------EVKAKRG-------------------------------------
sp|O00411|RPOM_ ----------SKLLRDVYAKDGRVSYPKLHLPLKTLQCLFEKQLHMELASRVCVVSV---
sp|P13433|RPOM_ GLNKVEAENDSKLHKASNASSDSIKKPSLD-PLREVSFHGSTEVLSKDAEKLIAVDTIGM
                    *. .

sp|P00573|RPOL_ ---------------------------------KRPT----------------------
sp|O00411|RPOM_ ---------------------------------EKPT----------------LPSKEV
sp|P13433|RPOM_ RVIRHTLLGLSLTPEQKEQISKFKFDANDNVLKMKPTKNDDNNNSINFFEIYNSLPTLEE
                                  :**

sp|P00573|RPOL_ --AFQ-------------------------------------------------------
sp|O00411|RPOM_ KHA----------------RKTLKTLRDQW---------------EKAL----------
sp|P13433|RPOM_ KKAFESALNIFNQDRQKVLENRATEAARERWKHDFEEAKARGDISIEKNLNVKLWKWYNE
           *

sp|P00573|RPOL_ -------------------------------------FLQEIKPEAVAYITIKTTLACL
sp|O00411|RPOM_ -----------CRALRETK---NRLEREVYEGRFSLYPFLCLLDEREVVRMLLQVLQALP
sp|P13433|RPOM_ MLPLVKEEINHCRSLLSEKLSDKKGLNKVDTNRLGYGPYLTLIDPGKMCVITILELLKLN
                                  :* :.  : : :

sp|P00573|RPOL_ TSA---DNTTVQAVASAIGRAIEDEARFGRIRDLEAKHFKKNVEEQLNKRVGHVYKKAFM
sp|O00411|RPOM_ AQG---ESFTTLARELS-ARTFSRHVVQRQRVSGQVQALQN-----------HYRKYLC
sp|P13433|RPOM_ STGGVIEGMRTARAVISVGKAIEMEFRSEQVLKSESQAFRD-----------VNKK---
           : .  :. .    : .:::. .   : . : : ::.           :*

sp|P00573|RPOL_ QVVEADMLSKGLLGGEAWSS--------------------WHKEDSIHVGVRCIEMLIES
sp|O00411|RPOM_ LLASDAEVPEPCLPRQYWEEL--GAPEALR----EQP---WPLPVQMELGKLLAEMLVQA
sp|P13433|RPOM_ -------SPE-------FKKLVQNAKSVFRSSQIEQSKILWPQSIRARIGSVLISMLIQV
                 .:     :..               *     .:*    .**::
230-240 hairpin : these extra 14 aa what interactions with b2?                 273
sp|P00573|RPOL_ TGM----------VSLHRQNAGVVGQDSETIELAPEYAEAIATRAGALAGI---------
```

```
sp|O00411|RPOM_ TQM------PCSLDKPHRSS-----------RLVPVLYHVYSFRNVQQIGILKPHPAYVQ
sp|P13433|RPOM_ AKVSVQGVDPVTKAKVHGEA----------PAFAHGYQYHNGSKLGVLKIHKTLIR
                : :        . * .          *  .  :   *:
                                                                        333
sp|P00573|RPOL_ ----------SPMF------QPCVVPPKPWTGITGGGYWANGRRPLALVRTHSKKALMR
sp|O00411|RPOM_ LLE-----KAAEPTLTFEAVDVPMLCPPLPWTSPHSGAFLLS---PTKLMRTVEGATQHQ
sp|P13433|RPOM_ QLNGERLIASVQPQL------LPMLVEPKPWVNWRSGGYHYT---QSTLLRTKDSPEQVA
                     .* :     * : * **..  .*.: .    *:** .
                                                                        393
sp|P00573|RPOL_ Y-----EDVYMPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPM
sp|O00411|RPOM_ ELLETCPPTALHGALDALTQLGNCAWRVNGRVLDLVLQLFQAKGCPQLGVPA-PPSEAPQ
sp|P13433|RPOM_ YLKAASDNGDIDRVYDGLNVLGRTPWTVNRKVFDVVSQVWN-KGEGFLDIPG-AQDEMVL
                     : . ..:.  . .* :* :*: :.  : : *    .:*.  .*
Palm(386-448)                                                           433
sp|P00573|RPOL_ KPEDI---------DMNPEALTAWKRAAAAVYRKDKARKSRRISLEFMLEQANKFANHKA
sp|O00411|RPOM_ PPEAHLPHSAAPARKAE---LRRELAHCQKVAREMHSLRA---EALYRLSLAQHLRDRV-
sp|P13433|RPOM_ PP--------APPKNSDPSILRAWKLQVKTIANKFSSDRSNRCDTNYKLEIARAFLGEK-
                *          . :  *       : .:  : ::   .  : *. *. : ..
W is polar on surface of protein; F is hydrophobic; 1cez
sp|P00573|RPOL_ IWFPYNMDWRGRVYAVS-MFNPQGNDMTKGLLTLAKGKPIGKEGYYWLKIHGANCAGVDK
sp|O00411|RPOM_ FWLPHNMDFRGRTYPCPPHFNHLGSDVARALLEFAQGRPLGPHGLDWLKIHLVNLTGLKK
sp|P13433|RPOM_ LYFPHNLDFRGRAYPLSPHFNHLGNDMSRGLLIFWHGKKLGPSGLKWLKIHLSNLFGFDK
                :::*:*:*:***.*. .  **  *.*:::.** : :*: :*  *  *****  *  *..*

sp|P00573|RPOL_ V-PFPERIKFIEENHENIMACAKSPLE-NTWWAEQDSPFCFLAFCFEYAGVQHHG--LSY
sp|O00411|RPOM_ REPLRKRLAFAEEVMDDILDSADQPLTGRKWWMGAEEPWQTLACCMEVANAVRASDPAAY
sp|P13433|RPOM_ L-PLKDRVAFTESHLQDIKDSAENPLTGDRWWTTADKPWQALATCFELNEVMKMDNPEEF
                *: .*: * *.  ::*  .*..**    **  :.*:  ** *:*   . : .   :

sp|P00573|RPOL_ NCSLPLAFDGSCSGIQHFSAMLRDEVGGRAVNLLPSETVQDIYGIVAKKVNEILQADAIN
sp|O00411|RPOM_ VSHLPVHQDGSCNGLQHYAALGRDSVGAASVNLEPSDVPQDVYSGVAAQVEVFRRQDAQR
sp|P13433|RPOM_ ISHQPVHQDGTCNGLQHYAALGGDVEGATQVNLVPSDKPQDVYAHVARLVQKRLEIAAEK
                . *:  **:*.*:**::*:  *  *.  *** **:  **:*. **  *:  . * .

sp|P00573|RPOL_ GTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYGVTRSVTKRSVMTLAYGSKEFGFRQ
sp|O00411|RPOM_ GMRVAQV-------------------------LEGFITRKVVKQTVMTVVYGVTRYGGRL
sp|P13433|RPOM_ GDENAKI-------------------------LKDKITRKVVKQTVMTNVYGVTYVGATF
                *    :                   *  .:**.*.*::*** .** .  *

sp|P00573|RPOL_ QVLEDTIQPAIDSGKGLMFTQPNQAAGYMAKLIWESVSVTVVAAVEAMNWLKSAAKLLAA
sp|O00411|RPOM_ QI-EKRLRELSDFPQEFVW----EASHYLVRQVFKSLQEMFSGTRAIQHWLTESARLISH
sp|P13433|RPOM_ QI-AKQLSPIFDDRKESL-----DFSKYLTKHVFSAIRELFHSAHLIQDWLGESAKRISK
                *: . :   *  : :    : : *:.: ::.::   . .:   .** .:*: ::

sp|P00573|RPOL_ EVKDKKTGEILR-------KRCAVHWVTPDGFPVWQEYK-----KPIQTRLNLMFL---G
sp|O00411|RPOM_ -------------------MGSVVEWVTPLGVPVIQPYRLDSKVKQIGGGIQSITYTHNG
sp|P13433|RPOM_ SIRLDVDEKSFKNGNKPDFMSSVI-WTTPLGLPIVQPYREESK-KQVETNLQTVFIS--D
                      ..: *.** *.*: * *:    * :  :: :    .
T760, N762 missing in Rpo41 & POLRMT Q744 which is polar and forms H bonds is replaced by G (hydrophobic, buried inside protein core) & E (charged) none of this will stabilize DNA through backbone interaction;770=D; look at POLRMT same region template interactions
sp|P00573|RPOL_ QFRLQPTINTNKDSEIDAHKQESGIAPNFVHSQDGSHLRKTVVWAHEKYGIESFALIHDS
sp|O00411|RPOM_ DISRKP----------NTRKQKNGFPPNFIHSLDSSHMMLTALHCYRK-GL-TFVSVHDC
sp|P13433|RPOM_ PFAVNP---------VNARRQKAGLPPNFIHSLDASHMLLSAAECGKQ-GL-DFASVHDS
                : :*       ::::*: *:.***:** *.**:  :.  . .: *:  *. :**.

sp|P00573|RPOL_ FGTIPADAANLFKAVRETMVDTYESC---------------------------------D
sp|O00411|RPOM_ YWTHAADVSVMNQVCREQFVRLHSEP-------------------------------
sp|P13433|RPOM_ YWTHASDIDTMNVVLREQFIKLHEVDLVLRLKEEFDQRYKNYVKIGKLKRSTDLAQKIIR
                : * .:*   :  . ** ::  :.

sp|P00573|RPOL_ VLAD----------------FYDQFADQLHESQLD-------------------------
sp|O00411|RPOM_ ILQDLSRFLVKR--------FCSEPQKILEASQLKE------------------------
sp|P13433|RPOM_ IRKDLSRKLGRSTTLADEIYFEKKRQELLNSPLIEDRNVGEKMVTTVSLFEDITDLDALE
                : *            * .:  . *. . :.

sp|P00573|RPOL_ -----------------KMPALPAKGNLNLRDILESDFAFA
sp|O00411|RPOM_ -----------------TLQAVPKPGAFDLEQVKRSTYFFS
sp|P13433|RPOM_ LENGGDENSGMSVLLPLRLPEIPPKGDFDVTVLRNSQYFFS
                      : :*  * :::  : .* : *:
```

VI. Tables

Mtf1 mutant	**Primers used in mutagenesis**
Mtf1-Δ4	Forward primer: 5' GTTGATATGTATCAAACATAGCACTCTGGTTG 3' Reverse primer: 5' CAACCAGAGTGCTATGTTTGATACATATCAAC 3'
Mtf1-Δ12	Forward primer: 5' GGCCTTTCAAACCAGATATTTTAATGTAATTTGTTGATATG 3' Reverse primer: 5' CATATCAACAAATTACATTAAAATATCTGGTTTGAAAGGCC 3'
Mtf1-Δ16	Forward primer: 5' GGAGTGGCCTTTCAAACCATAAATTTTAATGGATTTTGTTG 3' Reverse primer: 5' CAACAAAATCCATTAAAATTTATGGTTTGAAAGGCCACTCC 3'
Mtf1-Δ20	Forward primer: 5' CGAAATTATTCATGGAGTGGTAGTTCAAACCAGATATT 3' Reverse primer: 5' AATATCTGGTTTGAACTACCACTCCATGAATAATTTCG 3'
TFB2M-Δ3	Forward primer: 5' CGGCCGCAAGCTTCTACAGGGTTTCATCATAC 3' Reverse primer: 5' GTATGATGAAACCCTGTAGAAGCTTGCGGCCG 3'
TFB2M-Δ7	Forward primer: 5' TGCTTATAAATGGCTGTATTAGAAGCTTGCGGCCGC 3' Reverse primer: 5' GCGGCCGCAAGCTTCTAATACAGCCATTTATAAGCA 3'
TFB2M-Δ13	Forward primer: 5' CGGCCGCAAGCTTCTAACAATCTTTGGAACGC 3' Reverse primer: 5' GCGTTCCAAAGATTGTTAGAAGCTTGCGGCCG 3'
TFB2M-Δ17	Forward primer: 5' GCGGCCGCAAGCTTCTAACGCTCTATAGTTTCAA 3' Reverse primer: 5' TTGAAACTATAGAGCGTTAGAAGCTTGCGGCCGC 3'

TABLE 1: List of forward and reverse primers used to create Mtf1 and TFB2M C-terminal deletion mutants using site-directed mutagenesis technique.

Protein	**ε ($M^{-1}cm^{-1}$)**
Rpo41	156650
Mtf1-WT	73590
Mtf1-Δ4	73590
Mtf1-Δ12	72100
Mtf1-Δ16	72100
Mtf1-Δ20	72100
POLRMT	139490
TFAM	41495
TFB2M-WT	49890
TFB2M-Δ3	49890
TFB2M-Δ7	49890
TFB2M-Δ13	41410
TFB2M-Δ17	41285

TABLE 2: List of molar extinction coefficient values of all proteins used.

ABBREVIATIONS

2-AP	2-Aminopurine
A	Adenine
ATP	Adenine triphosphate
3'-dATP	3'-deoxy-ATP
3'-dCTP	3'-deoxy-CTP
3'-dGTP	3'-deoxy-GTP
aa	amino acid
bp	Base-pair
C	Cytosine
CTD	C-terminal domain
CTP	Cytidine triphosphate
DMS	Dimethyl suberimidate
DNA	Deoxyribonucleic acid
DNAzyme	DNA enzyme
DTT	Dithiothreitol
DEPC	Diethyl pyrocarbonate
ds	Double stranded
FL	Full-length
FRET	Förster Resonance Energy Transfer
G	Guanine
GTP	Guanosine triphosphate
HSP	Heavy strand promoter
k_{cat}	Catalytic rate
k_{cat}/K_m	Catalytic efficiency
K_d	Equilibrium dissociation constant
K_m	Michaelis-Menten constant
LSP	Light strand promoter
mt	Mitochondrial
mtRNAP	Mitochondrial RNA polymerase
m^7G	7-methylguanylate

NAD	Nicotinamide Adenine Dinucleotide
NADH	Nicotinamide Adenine Dinucleotide (reduced)
NCIN	Non-canonical initiating nucleotides
nt	Nucleotides
NTD	N-terminal domain
NTP	Nucleoside triphosphate
RNA	Ribonucleic acid
RNAP	RNA polymerase
ss	Single stranded
T	Thymine
TCA	Tricarboxylic acid cycle
TSS	Transcription Start Site
UTP	Uridine triphosphate

[K] REFERENCES

Adams, K. L. and J. D. Palmer (2003). "Evolution of mitochondrial gene content: gene loss and transfer to the nucleus." Mol Phylogenet Evol **29**(3): 380-395.

Alam, T. I., T. Kanki, T. Muta, K. Ukaji, Y. Abe, H. Nakayama, K. Takio, N. Hamasaki and D. Kang (2003). "Human mitochondrial DNA is packaged with TFAM." Nucleic Acids Res **31**(6): 1640-1645.

Bandwar, R. P. and S. S. Patel (2001). "Peculiar 2-aminopurine fluorescence monitors the dynamics of open complex formation by bacteriophage T7 RNA polymerase." J Biol Chem **276**(17): 14075-14082.

Basu, R. S., B. A. Warner, V. Molodtsov, D. Pupov, D. Esyunina, C. Fernandez-Tornero, A. Kulbachinskiy and K. S. Murakami (2014). "Structural basis of transcription initiation by bacterial RNA polymerase holoenzyme." J Biol Chem **289**(35): 24549-24559.

Bekers, K. M., J. J. Heijnen and W. M. van Gulik (2015). "Determination of the in vivo NAD:NADH ratio in Saccharomyces cerevisiae under anaerobic conditions, using alcohol dehydrogenase as sensor reaction." Yeast **32**(8): 541-557.

Bestwick, M. L. and G. S. Shadel (2013). "Accessorizing the human mitochondrial transcription machinery." Trends Biochem Sci **38**(6): 283-291.

Bird, J. G., U. Basu, D. Kuster, A. Ramachandran, E. Grudzien-Nogalska, A. Towheed, D. C. Wallace, M. Kiledjian, D. Temiakov, S. S. Patel, R. H. Ebright and B. E. Nickels (2018). "Highly efficient 5' capping of mitochondrial RNA with NAD(+) and NADH by yeast and human mitochondrial RNA polymerase." Elife **7**.

Bird, J. G., B. E. Nickels and R. H. Ebright (2017). "RNA Capping by Transcription Initiation with Non-canonical Initiating Nucleotides (NCINs): Determination of Relative Efficiencies of Transcription Initiation with NCINs and NTPs." Bio Protoc **7**(12).

Bird, J. G., Y. Zhang, Y. Tian, N. Panova, I. Barvik, L. Greene, M. Liu, B. Buckley, L. Krasny, J. K. Lee, C. D. Kaplan, R. H. Ebright and B. E. Nickels (2016). "The mechanism of RNA 5' capping with NAD+, NADH and desphospho-CoA." Nature **535**(7612): 444-447.

Biswas, T. K. (1999). "Nucleotide sequences surrounding the nonanucleotide promoter motif influence the activity of yeast mitochondrial promoter." Biochemistry **38**(30): 9693-9703.

Biswas, T. K. and G. S. Getz (1986). "A critical base in the yeast mitochondrial nonanucleotide promoter. Abolition of promoter activity by mutation at the -2 position." J Biol Chem **261**(9): 3927-3930.

Biswas, T. K. and G. S. Getz (1986). "Nucleotides flanking the promoter sequence influence the transcription of the yeast mitochondrial gene coding for ATPase subunit 9." Proc Natl Acad Sci U S A **83**(2): 270-274.

Biswas, T. K. and G. S. Getz (1990). "Regulation of transcriptional initiation in yeast mitochondria." J Biol Chem **265**(31): 19053-19059.

Biswas, T. K., B. Ticho and G. S. Getz (1987). "In vitro characterization of the yeast mitochondrial promoter using single-base substitution mutants." J Biol Chem **262**(28): 13690-13696.

Bonawitz, N. D., D. A. Clayton and G. S. Shadel (2006). "Initiation and beyond: multiple functions of the human mitochondrial transcription machinery." Mol Cell **24**(6): 813-825.

Boyaci, H., J. Chen, R. Jansen, S. A. Darst and E. A. Campbell (2019). "Structures of an RNA polymerase promoter melting intermediate elucidate DNA unwinding." Nature **565**(7739): 382-385.

Boyer, P. D., B. Chance, L. Ernster, P. Mitchell, E. Racker and E. C. Slater (1977). "Oxidative phosphorylation and photophosphorylation." Annu Rev Biochem **46**: 955-966.

Cahova, H., M. L. Winz, K. Hofer, G. Nubel and A. Jaschke (2015). "NAD captureSeq indicates NAD as a bacterial cap for a subset of regulatory RNAs." Nature **519**(7543): 374-377.
Cambronne, X. A., M. L. Stewart, D. Kim, A. M. Jones-Brunette, R. K. Morgan, D. L. Farrens, M. S. Cohen and R. H. Goodman (2016). "Biosensor reveals multiple sources for mitochondrial NAD(+)." Science **352**(6292): 1474-1477.
Canelas, A. B., W. M. van Gulik and J. J. Heijnen (2008). "Determination of the cytosolic free NAD/NADH ratio in Saccharomyces cerevisiae under steady-state and highly dynamic conditions." Biotechnol Bioeng **100**(4): 734-743.
Cermakian, N., T. M. Ikeda, P. Miramontes, B. F. Lang, M. W. Gray and R. Cedergren (1997). "On the evolution of the single-subunit RNA polymerases." J Mol Evol **45**(6): 671-681.
Chang, D. D. and D. A. Clayton (1984). "Precise identification of individual promoters for transcription of each strand of human mitochondrial DNA." Cell **36**(3): 635-643.
Chang, D. D., W. W. Hauswirth and D. A. Clayton (1985). "Replication priming and transcription initiate from precisely the same site in mouse mitochondrial DNA." EMBO J **4**(6): 1559-1567.
Cheetham, G. M., D. Jeruzalmi and T. A. Steitz (1999). "Structural basis for initiation of transcription from an RNA polymerase-promoter complex." Nature **399**(6731): 80-83.
Cheetham, G. M. and T. A. Steitz (1999). "Structure of a transcribing T7 RNA polymerase initiation complex." Science **286**(5448): 2305-2309.
Chen, W. W., E. Freinkman, T. Wang, K. Birsoy and D. M. Sabatini (2016). "Absolute Quantification of Matrix Metabolites Reveals the Dynamics of Mitochondrial Metabolism." Cell **166**(5): 1324-1337 e1311.
Chen, Y. G., W. E. Kowtoniuk, I. Agarwal, Y. Shen and D. R. Liu (2009). "LC/MS analysis of cellular RNA reveals NAD-linked RNA." Nat Chem Biol **5**(12): 879-881.
Chomyn, A., A. Martinuzzi, M. Yoneda, A. Daga, O. Hurko, D. Johns, S. T. Lai, I. Nonaka, C. Angelini and G. Attardi (1992). "MELAS mutation in mtDNA binding site for transcription termination factor causes defects in protein synthesis and in respiration but no change in levels of upstream and downstream mature transcripts." Proc Natl Acad Sci U S A **89**(10): 4221-4225.
Clayton, D. A. (1982). "Replication of animal mitochondrial DNA." Cell **28**(4): 693-705.
Cliften, P. F., S. H. Jang and J. A. Jaehning (2000). "Identifying a core RNA polymerase surface critical for interactions with a sigma-like specificity factor." Mol Cell Biol **20**(18): 7013-7023.
Cotney, J. and G. S. Shadel (2006). "Evidence for an early gene duplication event in the evolution of the mitochondrial transcription factor B family and maintenance of rRNA methyltransferase activity in human mtTFB1 and mtTFB2." J Mol Evol **63**(5): 707-717.
Dairaghi, D. J., G. S. Shadel and D. A. Clayton (1995). "Addition of a 29 residue carboxyl-terminal tail converts a simple HMG box-containing protein into a transcriptional activator." J Mol Biol **249**(1): 11-28.
Deana, A., H. Celesnik and J. G. Belasco (2008). "The bacterial enzyme RppH triggers messenger RNA degradation by 5' pyrophosphate removal." Nature **451**(7176): 355-358.
Dedrick, R. L. and M. J. Chamberlin (1985). "Studies on transcription of 3'-extended templates by mammalian RNA polymerase II. Parameters that affect the initiation and elongation reactions." Biochemistry **24**(9): 2245-2253.
Deshpande, A. P. and S. S. Patel (2014). "Interactions of the yeast mitochondrial RNA polymerase with the +1 and +2 promoter bases dictate transcription initiation efficiency." Nucleic Acids Res **42**(18): 11721-11732.
Deshpande, A. P., S. Sultana and S. S. Patel (2014). "Fluorescent methods to study transcription initiation and transition into elongation." Exp Suppl **105**: 105-130.
Drakulic, S., L. Wang, J. Cuellar, Q. Guo, G. Velazquez, J. Martin-Benito, R. Sousa and J. M. Valpuesta (2014). "Yeast mitochondrial RNAP conformational changes are regulated by

interactions with the mitochondrial transcription factor." Nucleic Acids Res **42**(17): 11246-11260.
Dyall, S. D., M. T. Brown and P. J. Johnson (2004). "Ancient invasions: from endosymbionts to organelles." Science **304**(5668): 253-257.
Dyson, H. J. and P. E. Wright (2005). "Intrinsically unstructured proteins and their functions." Nat Rev Mol Cell Biol **6**(3): 197-208.
Falkenberg, M., M. Gaspari, A. Rantanen, A. Trifunovic, N. G. Larsson and C. M. Gustafsson (2002). "Mitochondrial transcription factors B1 and B2 activate transcription of human mtDNA." Nat Genet **31**(3): 289-294.
Feng, Y., Y. Zhang and R. H. Ebright (2016). "Structural basis of transcription activation." Science **352**(6291): 1330-1333.
Fisher, R. P., T. Lisowsky, M. A. Parisi and D. A. Clayton (1992). "DNA wrapping and bending by a mitochondrial high mobility group-like transcriptional activator protein." J Biol Chem **267**(5): 3358-3367.
Fisher, R. P., J. N. Topper and D. A. Clayton (1987). "Promoter selection in human mitochondria involves binding of a transcription factor to orientation-independent upstream regulatory elements." Cell **50**(2): 247-258.
Furuichi, Y. (2015). "Discovery of m(7)G-cap in eukaryotic mRNAs." Proc Jpn Acad Ser B Phys Biol Sci **91**(8): 394-409.
Furuichi, Y., A. LaFiandra and A. J. Shatkin (1977). "5'-Terminal structure and mRNA stability." Nature **266**(5599): 235-239.
Furuichi, Y., S. Muthukrishnan, J. Tomasz and A. J. Shatkin (1976). "Mechanism of formation of reovirus mRNA 5'-terminal blocked and methylated sequence, m7GpppGmpC." J Biol Chem **251**(16): 5043-5053.
Furuichi, Y. and A. J. Shatkin (1976). "Differential synthesis of blocked and unblocked 5'-termini in reovirus mRNA: effect of pyrophosphate and pyrophosphatase." Proc Natl Acad Sci U S A **73**(10): 3448-3452.
Furuichi, Y. and A. J. Shatkin (2000). "Viral and cellular mRNA capping: past and prospects." Adv Virus Res **55**: 135-184.
Gaspari, M., M. Falkenberg, N. G. Larsson and C. M. Gustafsson (2004). "The mitochondrial RNA polymerase contributes critically to promoter specificity in mammalian cells." EMBO J **23**(23): 4606-4614.
Gaspari, M., N. G. Larsson and C. M. Gustafsson (2004). "The transcription machinery in mammalian mitochondria." Biochim Biophys Acta **1659**(2-3): 148-152.
Gray, M. W., G. Burger and B. F. Lang (1999). "Mitochondrial evolution." Science **283**(5407): 1476-1481.
Green, D. R. and J. C. Reed (1998). "Mitochondria and apoptosis." Science **281**(5381): 1309-1312.
Hasmann, M. and I. Schemainda (2003). "FK866, a highly specific noncompetitive inhibitor of nicotinamide phosphoribosyltransferase, represents a novel mechanism for induction of tumor cell apoptosis." Cancer Res **63**(21): 7436-7442.
Hawkins, M. E. (2003). Fluorescent Nucleoside Analogues as DNA Probes. Topics in Fluorescence Spectroscopy, Volume 7: DNA Technology. **Volume 7:** 151-175.
Hernandez, V. J., L. M. Hsu and M. Cashel (1996). "Conserved region 3 of Escherichia coli final sigma70 is implicated in the process of abortive transcription." J Biol Chem **271**(31): 18775-18779.
Hillen, H. S., Y. I. Morozov, A. Sarfallah, D. Temiakov and P. Cramer (2017). "Structural Basis of Mitochondrial Transcription Initiation." Cell **171**(5): 1072-1081 e1010.

Hillen, H. S., D. Temiakov and P. Cramer (2018). "Structural basis of mitochondrial transcription." Nat Struct Mol Biol **25**(9): 754-765.
Hofer, K., S. Li, F. Abele, J. Frindert, J. Schlotthauer, J. Grawenhoff, J. Du, D. J. Patel and A. Jaschke (2016). "Structure and function of the bacterial decapping enzyme NudC." Nat Chem Biol **12**(9): 730-734.
Igloi, G. L. and H. Kossel (1985). "Affinity electrophoresis for monitoring terminal phosphorylation and the presence of queuosine in RNA. Application of polyacrylamide containing a covalently bound boronic acid." Nucleic Acids Res **13**(19): 6881-6898.
Igloi, G. L. and H. Kossel (1987). "Use of boronate-containing gels for electrophoretic analysis of both ends of RNA molecules." Methods Enzymol **155**: 433-448.
Jang, S. H. and J. A. Jaehning (1991). "The yeast mitochondrial RNA polymerase specificity factor, MTF1, is similar to bacterial sigma factors." J Biol Chem **266**(33): 22671-22677.
Jiao, X., S. K. Doamekpor, J. G. Bird, B. E. Nickels, L. Tong, R. P. Hart and M. Kiledjian (2017). "5' End Nicotinamide Adenine Dinucleotide Cap in Human Cells Promotes RNA Decay through DXO-Mediated deNADding." Cell **168**(6): 1015-1027 e1010.
Joyce, G. F. (2001). "RNA cleavage by the 10-23 DNA enzyme." Methods Enzymol **341**: 503-517.
Julius, C. and Y. Yuzenkova (2017). "Bacterial RNA polymerase caps RNA with various cofactors and cell wall precursors." Nucleic Acids Res **45**(14): 8282-8290.
Kadesch, T. R. and M. J. Chamberlin (1982). "Studies of in vitro transcription by calf thymus RNA polymerase II using a novel duplex DNA template." J Biol Chem **257**(9): 5286-5295.
Kanki, T., K. Ohgaki, M. Gaspari, C. M. Gustafsson, A. Fukuoh, N. Sasaki, N. Hamasaki and D. Kang (2004). "Architectural role of mitochondrial transcription factor A in maintenance of human mitochondrial DNA." Mol Cell Biol **24**(22): 9823-9834.
Kaufman, B. A., N. Durisic, J. M. Mativetsky, S. Costantino, M. A. Hancock, P. Grutter and E. A. Shoubridge (2007). "The mitochondrial transcription factor TFAM coordinates the assembly of multiple DNA molecules into nucleoid-like structures." Mol Biol Cell **18**(9): 3225-3236.
Khan, J. A., X. Tao and L. Tong (2006). "Molecular basis for the inhibition of human NMPRTase, a novel target for anticancer agents." Nat Struct Mol Biol **13**(7): 582-588.
Kim, H., G. Q. Tang, S. S. Patel and T. Ha (2012). "Opening-closing dynamics of the mitochondrial transcription pre-initiation complex." Nucleic Acids Res **40**(1): 371-380.
Kostrewa, D., M. E. Zeller, K. J. Armache, M. Seizl, K. Leike, M. Thomm and P. Cramer (2009). "RNA polymerase II-TFIIB structure and mechanism of transcription initiation." Nature **462**(7271): 323-330.
Kowtoniuk, W. E., Y. Shen, J. M. Heemstra, I. Agarwal and D. R. Liu (2009). "A chemical screen for biological small molecule-RNA conjugates reveals CoA-linked RNA." Proc Natl Acad Sci U S A **106**(19): 7768-7773.
Kozak, M. and A. J. Shatkin (1977). "Sequences of two 5'-terminal ribosome-protected fragments from reovirus messenger RNAs." J Mol Biol **112**(1): 75-96.
Kukat, C., K. M. Davies, C. A. Wurm, H. Spahr, N. A. Bonekamp, I. Kuhl, F. Joos, P. L. Polosa, C. B. Park, V. Posse, M. Falkenberg, S. Jakobs, W. Kuhlbrandt and N. G. Larsson (2015). "Cross-strand binding of TFAM to a single mtDNA molecule forms the mitochondrial nucleoid." Proc Natl Acad Sci U S A **112**(36): 11288-11293.
Kulbachinskiy, A. and A. Mustaev (2006). "Region 3.2 of the sigma subunit contributes to the binding of the 3'-initiating nucleotide in the RNA polymerase active center and facilitates promoter clearance during initiation." J Biol Chem **281**(27): 18273-18276.
Lang, B. F., M. W. Gray and G. Burger (1999). "Mitochondrial genome evolution and the origin of eukaryotes." Annu Rev Genet **33**: 351-397.

Li, P., D. Nijhawan, I. Budihardjo, S. M. Srinivasula, M. Ahmad, E. S. Alnemri and X. Wang (1997). "Cytochrome c and dATP-dependent formation of Apaf-1/caspase-9 complex initiates an apoptotic protease cascade." Cell **91**(4): 479-489.
Litonin, D., M. Sologub, Y. Shi, M. Savkina, M. Anikin, M. Falkenberg, C. M. Gustafsson and D. Temiakov (2010). "Human mitochondrial transcription revisited: only TFAM and TFB2M are required for transcription of the mitochondrial genes in vitro." J Biol Chem **285**(24): 18129-18133.
Liu, C. and C. T. Martin (2001). "Fluorescence characterization of the transcription bubble in elongation complexes of T7 RNA polymerase." J Mol Biol **308**(3): 465-475.
Liu, C. and C. T. Martin (2002). "Promoter clearance by T7 RNA polymerase. Initial bubble collapse and transcript dissociation monitored by base analog fluorescence." J Biol Chem **277**(4): 2725-2731.
Liu, X., C. N. Kim, J. Yang, R. Jemmerson and X. Wang (1996). "Induction of apoptotic program in cell-free extracts: requirement for dATP and cytochrome c." Cell **86**(1): 147-157.
Mangus, D. A., S. H. Jang and J. A. Jaehning (1994). "Release of the yeast mitochondrial RNA polymerase specificity factor from transcription complexes." J Biol Chem **269**(42): 26568-26574.
Martin, C. T., D. K. Muller and J. E. Coleman (1988). "Processivity in early stages of transcription by T7 RNA polymerase." Biochemistry **27**(11): 3966-3974.
Martin, J. L. and F. M. McMillan (2002). "SAM (dependent) I AM: the S-adenosylmethionine-dependent methyltransferase fold." Curr Opin Struct Biol **12**(6): 783-793.
Masters, B. S., L. L. Stohl and D. A. Clayton (1987). "Yeast mitochondrial RNA polymerase is homologous to those encoded by bacteriophages T3 and T7." Cell **51**(1): 89-99.
Matsunaga, M. and J. A. Jaehning (2004). "Intrinsic promoter recognition by a "core" RNA polymerase." J Biol Chem **279**(43): 44239-44242.
McBride, H. M., M. Neuspiel and S. Wasiak (2006). "Mitochondria: more than just a powerhouse." Curr Biol **16**(14): R551-560.
McCulloch, V., B. L. Seidel-Rogol and G. S. Shadel (2002). "A human mitochondrial transcription factor is related to RNA adenine methyltransferases and binds S-adenosylmethionine." Mol Cell Biol **22**(4): 1116-1125.
McCulloch, V. and G. S. Shadel (2003). "Human mitochondrial transcription factor B1 interacts with the C-terminal activation region of h-mtTFA and stimulates transcription independently of its RNA methyltransferase activity." Mol Cell Biol **23**(16): 5816-5824.
Mercer, T. R., S. Neph, M. E. Dinger, J. Crawford, M. A. Smith, A. M. Shearwood, E. Haugen, C. P. Bracken, O. Rackham, J. A. Stamatoyannopoulos, A. Filipovska and J. S. Mattick (2011). "The human mitochondrial transcriptome." Cell **146**(4): 645-658.
Mitchell, P. (1966). "Chemiosmotic coupling in oxidative and photosynthetic phosphorylation." Biol Rev Camb Philos Soc **41**(3): 445-502.
Morozov, Y. I., K. Agaronyan, A. C. Cheung, M. Anikin, P. Cramer and D. Temiakov (2014). "A novel intermediate in transcription initiation by human mitochondrial RNA polymerase." Nucleic Acids Res **42**(6): 3884-3893.
Morozov, Y. I., A. V. Parshin, K. Agaronyan, A. C. Cheung, M. Anikin, P. Cramer and D. Temiakov (2015). "A model for transcription initiation in human mitochondria." Nucleic Acids Res **43**(7): 3726-3735.
Morozov, Y. I. and D. Temiakov (2016). "Human Mitochondrial Transcription Initiation Complexes Have Similar Topology on the Light and Heavy Strand Promoters." J Biol Chem **291**(26): 13432-13435.
Moustafa, I. M., A. Uchida, Y. Wang, N. Yennawar and C. E. Cameron (2015). "Structural models of mammalian mitochondrial transcription factor B2." Biochim Biophys Acta **1849**(8): 987-1002.

Murakami, K. S., S. Masuda and S. A. Darst (2002). "Structural basis of transcription initiation: RNA polymerase holoenzyme at 4 A resolution." Science **296**(5571): 1280-1284.
Muthukrishnan, S., Y. Furuichi, G. W. Both and A. J. Shatkin (1976). "Effect of 5'-terminal structures on the binding of ribopolymers to eukaryotic ribosomes." Prog Nucleic Acid Res Mol Biol **19**: 473-476.
Muthukrishnan, S., M. Morgan, A. K. Banerjee and A. J. Shatkin (1976). "Influence of 5'-terminal m7G and 2'--O-methylated residues on messenger ribonucleic acid binding to ribosomes." Biochemistry **15**(26): 5761-5768.
Nachtergaele, S. and C. He (2017). "The emerging biology of RNA post-transcriptional modifications." RNA Biol **14**(2): 156-163.
Ngo, H. B., J. T. Kaiser and D. C. Chan (2011). "The mitochondrial transcription and packaging factor Tfam imposes a U-turn on mitochondrial DNA." Nat Struct Mol Biol **18**(11): 1290-1296.
Ngo, H. B., G. A. Lovely, R. Phillips and D. C. Chan (2014). "Distinct structural features of TFAM drive mitochondrial DNA packaging versus transcriptional activation." Nat Commun **5**: 3077.
Nubel, G., F. A. Sorgenfrei and A. Jaschke (2017). "Boronate affinity electrophoresis for the purification and analysis of cofactor-modified RNAs." Methods **117**: 14-20.
Nunnari, J. and A. Suomalainen (2012). "Mitochondria: in sickness and in health." Cell **148**(6): 1145-1159.
Pagliarini, D. J. and J. Rutter (2013). "Hallmarks of a new era in mitochondrial biochemistry." Genes Dev **27**(24): 2615-2627.
Palade, G. E. (1953). "An electron microscope study of the mitochondrial structure." J Histochem Cytochem **1**(4): 188-211.
Paratkar, S., A. P. Deshpande, G. Q. Tang and S. S. Patel (2011). "The N-terminal domain of the yeast mitochondrial RNA polymerase regulates multiple steps of transcription." J Biol Chem **286**(18): 16109-16120.
Paratkar, S. and S. S. Patel (2010). "Mitochondrial transcription factor Mtf1 traps the unwound non-template strand to facilitate open complex formation." J Biol Chem **285**(6): 3949-3956.
Park, J. O., S. A. Rubin, Y. F. Xu, D. Amador-Noguez, J. Fan, T. Shlomi and J. D. Rabinowitz (2016). "Metabolite concentrations, fluxes and free energies imply efficient enzyme usage." Nat Chem Biol **12**(7): 482-489.
Posse, V., E. Hoberg, A. Dierckx, S. Shahzad, C. Koolmeister, N. G. Larsson, L. M. Wilhelmsson, B. M. Hallberg and C. M. Gustafsson (2014). "The amino terminal extension of mammalian mitochondrial RNA polymerase ensures promoter specific transcription initiation." Nucleic Acids Res **42**(6): 3638-3647.
Pupov, D., I. Kuzin, I. Bass and A. Kulbachinskiy (2014). "Distinct functions of the RNA polymerase sigma subunit region 3.2 in RNA priming and promoter escape." Nucleic Acids Res **42**(7): 4494-4504.
Ramachandran, A., U. Basu, S. Sultana, D. Nandakumar and S. S. Patel (2017). "Human mitochondrial transcription factors TFAM and TFB2M work synergistically in promoter melting during transcription initiation." Nucleic Acids Res **45**(2): 861-874.
Ramanathan, A., G. B. Robb and S. H. Chan (2016). "mRNA capping: biological functions and applications." Nucleic Acids Res **44**(16): 7511-7526.
Reha-Krantz, L. J., C. Hariharan, U. Subuddhi, S. Xia, C. Zhao, J. Beckman, T. Christian and W. Konigsberg (2011). "Structure of the 2-aminopurine-cytosine base pair formed in the polymerase active site of the RB69 Y567A-DNA polymerase." Biochemistry **50**(46): 10136-10149.
Ringel, R., M. Sologub, Y. I. Morozov, D. Litonin, P. Cramer and D. Temiakov (2011). "Structure of human mitochondrial RNA polymerase." Nature **478**(7368): 269-273.

Rosca, M. G. and C. L. Hoppel (2010). "Mitochondria in heart failure." Cardiovasc Res **88**(1): 40-50.
Rosca, M. G. and C. L. Hoppel (2013). "Mitochondrial dysfunction in heart failure." Heart Fail Rev **18**(5): 607-622.
Rubio-Cosials, A., J. F. Sidow, N. Jimenez-Menendez, P. Fernandez-Millan, J. Montoya, H. T. Jacobs, M. Coll, P. Bernado and M. Sola (2011). "Human mitochondrial transcription factor A induces a U-turn structure in the light strand promoter." Nat Struct Mol Biol **18**(11): 1281-1289.
Sainsbury, S., J. Niesser and P. Cramer (2013). "Structure and function of the initially transcribing RNA polymerase II-TFIIB complex." Nature **493**(7432): 437-440.
Santoro, S. W. and G. F. Joyce (1997). "A general purpose RNA-cleaving DNA enzyme." Proc Natl Acad Sci U S A **94**(9): 4262-4266.
Savkina, M., D. Temiakov, W. T. McAllister and M. Anikin (2010). "Multiple functions of yeast mitochondrial transcription factor Mtf1p during initiation." J Biol Chem **285**(6): 3957-3964.
Schubot, F. D., C. J. Chen, J. P. Rose, T. A. Dailey, H. A. Dailey and B. C. Wang (2001). "Crystal structure of the transcription factor sc-mtTFB offers insights into mitochondrial transcription." Protein Sci **10**(10): 1980-1988.
Shadel, G. S. (2008). "Expression and maintenance of mitochondrial DNA: new insights into human disease pathology." Am J Pathol **172**(6): 1445-1456.
Shadel, G. S. and D. A. Clayton (1997). "Mitochondrial DNA maintenance in vertebrates." Annu Rev Biochem **66**: 409-435.
Shatkin, A. J. (1976). "Capping of eucaryotic mRNAs." Cell **9**(4 PT 2): 645-653.
Shatkin, A. J., A. K. Banerjee, G. W. Both, Y. Furuichi and S. Muthukrishnan (1976). "Dependence of translation on 5'-terminal methylation of mRNA." Fed Proc **35**(11): 2214-2217.
Shutt, T. E., M. Bestwick and G. S. Shadel (2011). "The core human mitochondrial transcription initiation complex: It only takes two to tango." Transcription **2**(2): 55-59.
Shutt, T. E. and G. S. Shadel (2010). "A compendium of human mitochondrial gene expression machinery with links to disease." Environ Mol Mutagen **51**(5): 360-379.
Sjostrand, F. S. (1953). "Electron microscopy of mitochondria and cytoplasmic double membranes." Nature **171**(4340): 30-32.
Sologub, M., D. Litonin, M. Anikin, A. Mustaev and D. Temiakov (2009). "TFB2 is a transient component of the catalytic site of the human mitochondrial RNA polymerase." Cell **139**(5): 934-944.
Sosunov, V., E. Sosunova, A. Mustaev, I. Bass, V. Nikiforov and A. Goldfarb (2003). "Unified two-metal mechanism of RNA synthesis and degradation by RNA polymerase." EMBO J **22**(9): 2234-2244.
Sousa, R., Y. J. Chung, J. P. Rose and B. C. Wang (1993). "Crystal structure of bacteriophage T7 RNA polymerase at 3.3 A resolution." Nature **364**(6438): 593-599.
Steitz, T. A., S. J. Smerdon, J. Jager and C. M. Joyce (1994). "A unified polymerase mechanism for nonhomologous DNA and RNA polymerases." Science **266**(5193): 2022-2025.
Taanman, J. W. (1999). "The mitochondrial genome: structure, transcription, translation and replication." Biochim Biophys Acta **1410**(2): 103-123.
Tahirov, T. H., D. Temiakov, M. Anikin, V. Patlan, W. T. McAllister, D. G. Vassylyev and S. Yokoyama (2002). "Structure of a T7 RNA polymerase elongation complex at 2.9 A resolution." Nature **420**(6911): 43-50.
Tang, G. Q., A. P. Deshpande and S. S. Patel (2011). "Transcription factor-dependent DNA bending governs promoter recognition by the mitochondrial RNA polymerase." J Biol Chem **286**(44): 38805-38813.

Tang, G. Q., S. Paratkar and S. S. Patel (2009). "Fluorescence mapping of the open complex of yeast mitochondrial RNA polymerase." J Biol Chem **284**(9): 5514-5522.
Thorsness, P. E. and T. D. Fox (1990). "Escape of DNA from mitochondria to the nucleus in Saccharomyces cerevisiae." Nature **346**(6282): 376-379.
Vvedenskaya, I. O., J. G. Bird, Y. Zhang, Y. Zhang, X. Jiao, I. Barvik, L. Krasny, M. Kiledjian, D. M. Taylor, R. H. Ebright and B. E. Nickels (2018). "CapZyme-Seq Comprehensively Defines Promoter-Sequence Determinants for RNA 5' Capping with NAD<sup/>." Mol Cell **70**(3): 553-564 e559.
Wallace, D. C. (1999). "Mitochondrial diseases in man and mouse." Science **283**(5407): 1482-1488.
Walters, R. W., T. Matheny, L. S. Mizoue, B. S. Rao, D. Muhlrad and R. Parker (2017). "Identification of NAD+ capped mRNAs in Saccharomyces cerevisiae." Proc Natl Acad Sci U S A **114**(3): 480-485.
Wang, X. and C. He (2014). "Dynamic RNA modifications in posttranscriptional regulation." Mol Cell **56**(1): 5-12.
Wettstein-Edwards, J., B. S. Ticho, N. C. Martin, D. Najarian and G. S. Getz (1986). "In vitro transcription and promoter strength analysis of five mitochondrial tRNA promoters in yeast." J Biol Chem **261**(6): 2905-2911.
Yakubovskaya, E., K. E. Guja, E. T. Eng, W. S. Choi, E. Mejia, D. Beglov, M. Lukin, D. Kozakov and M. Garcia-Diaz (2014). "Organization of the human mitochondrial transcription initiation complex." Nucleic Acids Res **42**(6): 4100-4112.
Yin, Y. W. and T. A. Steitz (2002). "Structural basis for the transition from initiation to elongation transcription in T7 RNA polymerase." Science **298**(5597): 1387-1395.
Zhang, Y., Y. Feng, S. Chatterjee, S. Tuske, M. X. Ho, E. Arnold and R. H. Ebright (2012). "Structural basis of transcription initiation." Science **338**(6110): 1076-1080.
Zhao, B. S., I. A. Roundtree and C. He (2017). "Post-transcriptional gene regulation by mRNA modifications." Nat Rev Mol Cell Biol **18**(1): 31-42.