Structural and Functional Characterization of the Human Enzyme Nocturnin

by

Elizabeth T. Abshire

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Biological Chemistry)
in The University of Michigan
2019

Doctoral Committee:

        Adjunct Associate Professor Aaron Goldstrohm, Co-Chair
        Associate Professor Ray Trievel, Co-Chair
        Professor Ormond MacDougald
        Associate Professor Patrick O'Brien
        Professor Janet Smith

Elizabeth T. Abshire

eabshire@umich.edu

ORCID iD: 0000-0002-2019-3044

Dedication

This work is dedicated to family, and to friends who are family

Acknowledgements

I want to first acknowledge my mentors, Dr. Ray Trievel and Dr. Aaron Goldstrohm. They have provided me not only with guidance but fostered my independence in the lab, scientific exploration, and my future in science. Dr. Peter Freddolino has been instrumental in providing additional guidance on this project. He and Rucheng Diao cannot be thanked enough for their direct contributions to the work presented here. I would also like to thank Dr. Kelsey Hughes, who contributed data and provided invaluable help and scientific discussion.

I would like to thank my committee members, Drs. Ormond MacDougald, Patrick O'Brien, and Janet Smith, who have been immensely supportive in my scientific and career aspirations. Your advice was instrumental to my success. Additional thanks go to my colleagues in the Goldstrohm and Trievel Labs, who contributed to this work both directly and indirectly.

I am grateful to the Biological Chemistry staff, especially Beth Goodwin, Amanda Howard, and Prasanna Baragi helped me navigate graduate school outside of the lab.

The work presented here in this thesis was supported by the support of the American Heart Association Predoctoral Fellowship (16PRE26700002), the National

Table of Contents

List of Figures

List of Tables

CHAPTER 5

List of Abbreviations

2´,5'-A: 2´, 5'- oligoadenylate

3' UTR: 3' untranslated region

3', 5' ADP: 3', 5' adenosine diphosphate

4E-BP: eukaryotic translation initiation factor 4E-binding protein

5' UTR: 5' untranslated region

7mG: 7-methylguanosine cap

ALP: alkaline phosphatase

AMP: adenosine monophosphate

aP2: adipocyte protein 2

βME: β- mercaptoethanol

BMSC: bone marrow mesenchymal stem cells

BSA: bovine serum albumin

cADPR: cyclic ADP ribose

CAF1: see POP2

CCR4: carbon catabolite repression 4

CDS: coding sequence

CEBP: Cytoplasmic Polyadenylation Element Binding Protein

CNOT6: CCR4-NOT transcription complex subunit 6

CNOT6L: CCR4-NOT transcription complex subunit 6-like

CTD: C-terminal domain

DCP1: decapping protein 1

DCP2: decapping protein 2

DCPS: scavenger decapping protein

DHH1: TP-dependent RNA helicase DHH1

DNA: deoxyribonucleic acid

dsDNA: double-stranded deoxyribonucleic acid

DTT: dithiothreitol

EDTA: Ethylenediaminetetraacetic acid

EEP: exonuclease, endonuclease, and phosphatase enzyme superfamily

eIF4A: eukaryotic Initiation Factor 4A

eIF4E: eukaryotic Initiation Factor 4E

eIF4F: eukaryotic Initiation Factor 4F

eIF4G: eukaryotic Initiation Factor 4G

GAPDH: glyceraldehyde-3- phosphate dehydrogenase

GE1: Enhancer of mRNA-decapping protein 4 homolog

GLD2: Poly(A) RNA polymerase GLD2

GO: gene ontology

HEPES: 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid

HFD: high-fat diet

IMPDH1: inosine 5' monophosphate dehydrogenase 1

IP: immunoprecipitation

IPA: Ingenuity Pathway Analysis

IRP: iron response protein

lncRNA: long non-coding ribonucleic acid

LRR: leucine-rich repeat

LSM: U6 snRNA-associated Sm-like protein

MAPK: mitogen-activated protein kinase

miRNA: micro ribonucleic acid

MNase: micrococcal nuclease

mNOCT: mouse Nocturnin

MOPS: 3-N-Morpholino propanesulfonic acid

MPP: mitochondrial processing peptidase

MR: molecular replacement

mRNA: messenger ribonucleic acid

MSA: multiple sequence alignment

MS2-HT-CNOT7: Bacteriophage MS2-HaloTag- CCR4-NOT transcription complex subunit 7 fusion protein

MS2-NOCT$_{64\text{-}431}$: Bacteriophage MS2-Nocturnin $_{64\text{-}431}$

MS2-NOCT$_{1\text{-}431}$: Bacteriophage MS2-Nocturnin $_{1\text{-}431}$

mTOR: mammalian Target of Rapamycin

MTS: mitochondrial targeting sequence

NADP(H): Nicotinamide adenine dinucleotide phosphate

NAADP: nicotinic acid adenine dinucleotide phosphate

NLuc: NanoLuciferase

NLuc 4xMS2BS p(A): NanoLuciferase 4x MS2 Binding Site poly(A)

NLuc 4xMS2BS MALAT1: NanoLuciferase 4x MS2 Binding Site MALAT 1

NOCT: Nocturnin

nt: nucleotide

OC: osteocalcin

OE: overexpression

ORF: open reading frame

PABPC1: Poly(A) Binding Protein Cytoplasm 1

PAIP1: poly(A) binding protein interacting protein 1

PAN2: polyadenosine specific ribonuclease subunit PAN2

PAN3: polyadenosine specific ribonuclease subunit PAN3

PARN: polyadenosine-specific ribonuclease

PAT1: DNA topoisomerase 2-associated protein PAT1

PDB: Protein Data Bank

PDE12: Phosphodiesterase 12

PIC: pre-initiation complex, translation

Poly(A): polyadenosine

Poly(C): polycytosine

Poly(U): polyuridine

POP2: Poly(A) ribonuclease POP2

PPARγ: peroxisome proliferator-activated receptor γ

Pre-mRNA: precursor messenger ribonucleic acid

RBP: ribonucleic acid binding protein

RIP: ribonucleic acid immunoprecipitation

RNA: ribonucleic acid

rRNA: ribosomal ribonucleic acid

RNApol II: ribonucleic acid polymerase 2

ROS: reactive oxygen species

SH2: Src homology 2

shRNA: short hairpin ribonucleic acid

snRNP: small nuclear ribonucleoprotein

dsDNA: single-stranded deoxyribonucleic acid

TBP: TATA binding protein

TIM: Mitochondrial import inner membrane translocase subunit

TOM: Mitochondrial import receptor subunit

TPM: transcripts per million

tRNA: transfer ribonucleic acid

TTP: Tristetraprolin

WT: wild-type

xNOCT: *Xenopus laevis* Nocturnin

XRN1: 5' – 3' Exoribonuclease 1

XRN2: 5' – 3' Exoribonuclease 2

ZT: Zeitgeber time

Abstract

The circadian protein Nocturnin (NOCT) belongs to the exonuclease, endonuclease and phosphatase (EEP) superfamily and is most similar to the CCR4-class of deadenylases that degrade the poly-adenosine poly(A) tails of mRNAs. NOCT-deficient mice are resistant to diet-induced obesity and exhibit increased bone density; however, the mechanisms by which NOCT regulates these processes remain unknown.

Based on homology, we predicted that NOCT adopts the α/β hydrolase fold that is characteristic of EEP deadenylases, has activity against poly(A) RNAs *in vitro* and regulates mRNA expression and half-lives *in vivo*. Since the NOCT null animals display alterations in lipid metabolism, we expected that we would identify target mRNAs of NOCT involved in lipid metabolism and adipogenesis.

To determine the structure of the NOCT catalytic domain, we solved a pair of crystal structures of NOCT, demonstrating that the structure is highly conserved with other EEP deadenylases. Purified recombinant NOCT lacked ribonuclease activity *in vitro*. However, in cell-based assays, we found that NOCT reduces translation and abundance of reporter mRNAs. Incomplete reversal of repression when the NOCT active site is mutated suggests a role for translational repression in addition to ribonuclease activity. NOCT-mediated repression of reporter mRNAs is additionally dependent upon the 3′ end, as reporters terminating with a 3′ MALAT1 structure cannot be repressed by

NOCT. Together, these data suggest a model by which NOCT represses expression from its target mRNAs through RNA decay and translational repression.

We then examined the NOCT protein sequence for motifs indicating subcellular localization. Sequence features of the *NOCT* mRNA suggests that NOCT may be localized to the cytoplasm and the mitochondria. Additionally, the NOCT protein sequence has predicted motifs for mitochondrial localization. Consistent with these observations, we observed that the full-length NOCT is a preprotein that can be processed. Surprisingly, processing of NOCT appears to occur in a tissue-type specific manner, suggesting that NOCT is differentially located to the cytoplasm or mitochondria in different cell types.

To identify potential targets of NOCT we generated HEK293 cells with cytoplasmic overexpression of NOCT and performed RNA-Seq. A dataset of transcripts that change in abundance with NOCT overexpression was identified, demonstrating that NOCT is able to affect steady state levels of mRNAs. In addition to these potential NOCT targets, we further confirmed a report that NOCT has phosphatase activity against NADPH and $NADP^+$, indicating that NOCT can hydrolyze non-mRNA substrates. These data suggest that NOCT may regulate metabolism through modulating the steady state levels of mRNAs as well as NADP(H).

CHAPTER 1

Introduction

1.1 mRNA and the central dogma

The central dogma of molecular biology describes the flow of information from coded instructions in the nucleus to ribonucleic acids that encode temporary instructions for synthesis of proteins that maintain cellular function. This classical view of the central dogma, although intended to be a succinct description of a key biological process, understates the role of ribonucleic acid (RNA). While messenger RNA (mRNA) function is reflected in the transfer of information from DNA to protein, multiple types of RNA are encoded in the genome that function within the processes encompassed by the central dogma. Ribosomal RNAs (rRNAs) compose the catalytic core of the ribosome, which is responsible for protein synthesis in concert with transfer RNAs (tRNAs) that translate the three nucleotide codons of mRNA to a corresponding amino acid of the polypeptide chain. Other layers of gene regulation are performed by yet other RNAs, such as microRNAs (miRNAs), long noncoding RNAs (lncRNAs), and enhancer RNAs. Regulatory mRNAs such as these act on RNAs themselves, as in the case of miRNAs and lncRNAs that bind to specific sequences in mRNAs in order to destabilize them (Bartel, 2018; Yoon et al., 2013). While the remainder of this chapter will focus on regulation of eukaryotic mRNA expression levels and stability, it is worthwhile to consider that this is only one of many important RNAs involved in gene expression.

1.2 Pre-mRNA processing and eukaryotic mRNA maturation

Before an mRNA can be translated, it must first be transcribed, processed, and transported to the cytoplasm. RNA transcribed from DNA is called precursor mRNA (pre-mRNA), which lacks features enabling translation and contains intronic sequences that must be removed for proper expression of the gene product. Therefore, the majority of pre-mRNA must undergo the processes of 5´ capping, splicing, and 3´ cleavage and polyadenylation before the transcripts are exported into the cytoplasm. These processes are closely linked to the transcription of these messages, as the C-terminal domain of RNA polymerase II (RNApol II) has been shown to regulate and facilitate pre-mRNA processing (Hsin and Manley, 2012; Lee and Tarn, 2013).

Capping occurs at the 5´ end of the nascent message, appending a 7-methylguanosine cap (7mG) structure to the pre-mRNA in three enzymatic steps (Ramanathan et al., 2016). First, the 5´ γ phosphate is removed by an RNA 5´ triphosphatase before a guanosine monophosphate is transferred to the 5´ end by a guanylyl transferase to form a unique 5´-5´ triphosphate linkage. Finally, the guanosine N7 position is methylated. The 5´ cap plays several roles, including protecting the mRNA from exonucleolytic decay, recruiting splicing factors to nascent pre-mRNA, and facilitating nuclear export of matured mRNAs (Ramanathan et al., 2016).

In addition to capping, the pre-mRNA undergoes splicing. This is performed by the spliceosome complex, a large ribonucleoprotein complex that assembles on splice sites as they are transcribed. These splice sites delineate the 5´ and 3´ boundaries of the introns, which are removed through the activity of the spliceosome catalytic RNAs (Herzel et al., 2017; Shi, 2017a, b). The small nuclear ribonucleoproteins (snRNPs) of the

spliceosome assemble on the 5′ and 3′ splice sites and a site close to the 3′ end called a branch point. The branch point is ligated to the 5′ end of the intron after cleavage of the 5′ splice site. At this point the intron is removed at the 3′ end and released as the 5′ and 3′ exons are ligated together (Herzel et al., 2017; Shi, 2017a). Many transcripts are alternatively spliced, creating mRNA isoforms with varied properties and functionality. Regulation of this process includes differential expression of specific splicing regulators and post translational modification of general splicing factors (Chen and Manley, 2009). Proper splicing of pre-mRNAs is critical for generating an open reading frame (ORF) that will encode the correct gene product upon translation in the cytoplasm.

Further maturation of the pre-mRNA is carried out by the pre-mRNA 3′ end processing complex (Hirose and Manley, 1998). Cleavage by the multi-subunit 3′ end processing machinery occurs at sites defined by a CA or UA dinucleotide and upstream AAUAAA as well as a downstream U or GU-rich sequence element (Neve et al., 2017). After 3′ cleavage, the poly(A) polymerase subunit adds the 3′ polyadenosine (poly(A)) tail in a non-templated polymerase reaction (Elkon et al., 2013; Neve et al., 2017). On average, poly(A) polymerase adds approximately 200-250 adenosines to the ends of mRNAs, though this can be highly variable (Chang et al., 2014; Jalkanen et al., 2014; Kuhn et al., 2009). Alternative cleavage and polyadenylation signals contribute to the diversity of isoforms that can be created from a given pre-mRNA (Elkon et al., 2013; Neve et al., 2017; Tian and Manley, 2013). These alternative signals can be in different exons, functionally linking alternative polyadenylation to alternative splicing (Kyburz et al., 2006; Tian and Manley, 2013).

Export of mRNAs from the nucleus is highly dependent on completion of pre-mRNA processing and mRNA quality control surveillance pathways (Soheilypour and Mofrad, 2018). During pre-mRNA maturation, a large number of RNA binding proteins (RBPs) and RNA modifying proteins are recruited to the transcript, some of which are utilized in subsequent export pathways. The diversity of length, sequence, and structure found among mRNAs results in variation in what proteins are bound to the maturing transcripts, which necessitates the use of a number of different export receptors. Once an mRNA is paired with its appropriate export receptor, it is targeted to the nuclear pore complex and translocated into the cytoplasm. The mature mRNA is then released from the export machinery through an ATP-dependent mechanism, where it is then able to be translated in the cytoplasm (Carmody and Wente, 2009; Kohler and Hurt, 2007; Mandon et al., 2013; Stewart, 2019).

Mature mRNA transcripts, with few exceptions, are made up of a coding sequence that provides sequence information for the encoded protein in the open reading frame (ORF) of the mRNA (Figure 1.1). The ORF is defined by specific codons that dictate where translation should start and stop (Figure 1.1). Flanking the ORF on either side is a 5´ untranslated region (5´ UTR) and 3´ untranslated region (3´ UTR) that contain a series
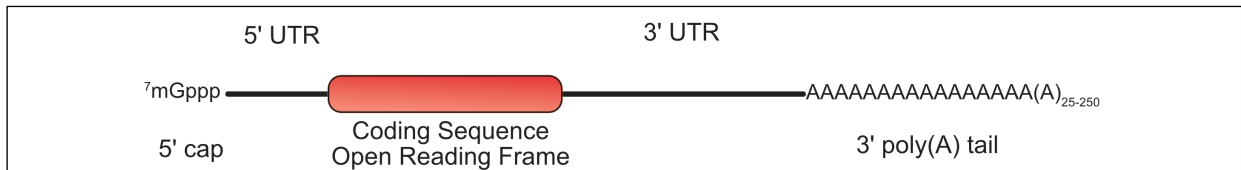
Figure 1.1: mRNAs are comprised of an open reading frame as well as regulatory elements. The 5´ end of the transcript contains the 7 methylguanosine cap attached to the transcript through a triphosphate bond. The 5´ UTR houses regulatory elements that effect expression of the transcript's gene product from the ORF. The ribosome translates the protein from this sequence. The 3´ UTR follows the CDS and also contains regulatory elements, many of which effect transcript translation and stability. The 3´ end of the transcript is polyadenylated, and this poly(A) tail ranges in size from approximately 25 to 250 nucleotides in length. The poly(A) tail is critical for translation and the removal of this tail affects translation and stability of the transcript.

of regulatory elements that control the efficiency of translation and the stability of the mRNA. Regulatory elements can also be found in the ORF itself (Figure 1.1).

Pre-mRNA processing of nuclear-encoded mRNAs is an important process by which the RNA transcribed from its gene matures into a form with the elements needed for translation as well as the correct flanking UTR sequences, which are important for downstream regulation of a transcript in the cytoplasm. For example, alternative polyadenylation can create transcript isoforms with different 3′ UTRs containing variable regulatory sequences that affect half-life, translation efficiency, and localization of mRNAs (Zheng et al., 2018). Differential polyadenylation can be regulated by extracellular signaling pathways such as the mitogen-activated protein kinase (MAPK) pathway, which responds to a variety extracellular signals and propagates them through multiple signaling cascades. In one example, MAPK signaling influences 3′ end processing of the thrombin mRNA during cellular stress responses (Cargnello and Roux, 2011; Danckwardt et al., 2011). Additionally, the adiponectin receptor gene expresses multiple mRNA splice isoforms that are differentially translated, as one isoform contains a short upstream ORF that represses translation of the downstream ORF coding for the functional gene product (Ashwal et al., 2011). Thus, the role of pre-mRNA processing cannot be overlooked within the larger context of mRNA regulation. Mechanisms of crosstalk between mRNA metabolism in the cytoplasm and nucleus have been reported, including nuclear RBPs traveling with mRNAs into the cytoplasm and signal transduction pathways regulating pre-mRNA splicing (Ashton-Beaucage et al., 2014; Kataoka et al., 2000; Moore and Proudfoot, 2009). Control of co-transcriptional processes from outside the nucleus is still

incompletely understood but may have large effects on differential pre-mRNA processing and downstream post-transcriptional and translational control.

1.3 Regulation of translation initiation of mRNAs

Translation by the ribosome occurs once an mRNA enters the cytoplasm. First, the mRNA is activated, bringing the 5′ and 3′ ends together to form a closed loop structure through interaction of proteins binding the poly(A) tail and 7mG cap (Figure 1.2). Poly(A) Binding Protein Isoform Cytoplasmic 1 (PABPC1) bound to the poly(A) tail interacts with eukaryotic Initiation Factor 4G (eIF4G) of the 7mG cap- binding complex (Figure 1.2) (Borman et al., 2000; Gallie, 1991; Michel et al., 2000). Another initiation factor, eIF4E, is bound to eIF4G and directly interacts with the 7mG cap. eIF4Ea and eIF4G, as well as the subunits eIF4B and eIF4A, make up the eIF4F translation initiation complex (Aitken and Lorsch, 2012; Ramanathan et al., 2016). The activated mRNA closed loop structure then stimulates translation initiation (Borman et al., 2000; Gallie, 1991; Michel et al., 2000).

During initiation, the translation pre-initiation complex (PIC) is assembled, which requires the concerted action of ribosomal subunits and initiation factors (Figure 1.2). Formation of the PIC starts with association of
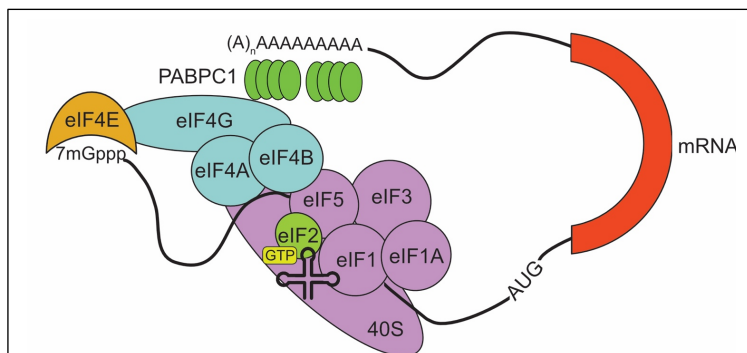
Figure 1.2: A schematic representation of the translation pre-initiation complex. The PIC is composed of the 7mG cap binding protein eIF4E, which associates with PABPC1 on the poly(A) tail through interaction with eIF4G to form the closed loop structure. In addition to the other initiation factors, the PIC is made up of the 40S small ribosomal subunit, and initiator Met-tRNA. The PIC scans for the AUG start codon, is joined by the 60S ribosome, and then commences elongation. Figure adapted from Arvola et al, 2016.

the small ribosomal subunit with initiation factors eIF1, eIF1A, eIF3, eIF5, and eiF2 bound

to GTP and initiator Met-tRNA. The 43S PIC then associates with the activated mRNA through contacts between the PIC, PABPC1, the mRNA, and eIF4F (Aitken and Lorsch, 2012; Andreev et al., 2017; Jackson et al., 2010). The resulting 48S PIC then scans through the 5´ UTR to identify the AUG start codon while the initiation factors stabilize the open conformation of the ribosome and the RNA helicase eIF4A unwinds secondary structures in the 5´ UTR (Marintchev, 2013; Parsyan et al., 2011). The AUG start codon and surrounding consensus sequence (Kozak sequence) indicates the location for translation initiation (Kozak, 1991a, b). In the case where the Kozak sequence context has deviated from the consensus, downstream AUGs can be used for initiation in a process known as leaky scanning (Hinnebusch, 2014, 2017).

The AUG start codon pairs with the Met-tRNA anticodon, at which point the PIC adopts a closed conformation. Initiator Met-tRNA is now poised in the ribosome P-site and eIF1 is released. eIF2 then hydrolyzes its bound GTP and is released along with the other initiation factors, with the exception of eIF1A (Hinnebusch, 2014). The 80S ribosome, bound to eIF5B, binds the PIC before eIF5B cleaves its bound GTP and is released with eIF1A (Aitken and Lorsch, 2012; Jackson et al., 2010). The ribosome then proceeds to the elongation step and then translation is terminated at the first in-frame stop codon (UGA, UAG, or UAA) (Andreev et al., 2017; Dever and Green, 2012).

Translation initiation is highly regulated, which includes global measures, as is observed with regulation by the mammalian Target of Rapamycin (mTOR) pathway. mTOR promotes PIC assembly on capped mRNAs by phosphorylating and inhibiting the action of eIF4E binding proteins (4E-BPs) that prevent association with the PIC (Gingras et al., 1998; Nandagopal and Roux, 2015). Regulation of translation can also be specific

to a subset of mRNAs sharing cis-acting sequence elements that recruit trans-factors. This is the case for miRNAs, which are highly sequence specific and function through base pairing to a complementary seed sequence to promote translational repression and/or mRNA decay (Bartel, 2009). Proposed mechanisms of this repression include promoting dissociation of PABPC1 from the poly(A) tail, disrupting PABPC1 interaction with eIF4G, or causing dissociation of eIF4A (Fukao et al., 2014; Iwakawa and Tomari, 2015; Tritschler et al., 2010). These mechanisms commonly work by disrupting interactions stabilizing the mRNA closed loop structure.

Translational repression can also occur by leaving the mRNA vulnerable to decay pathways (Kawahara et al., 2008; Zekri et al., 2013). For example, Pumilio binding to its consensus sequence disrupts binding of Pabp (the *Drosophila* homolog of human PABPC1), increasing deadenylation by the Ccr4-NOT complex (Weidmann et al., 2014). Tristetraprolin (TTP) interacts with AU-rich elements in the 3′UTR of mRNAs and recruits decapping factors (Fenger-Gron et al., 2005; Lykke-Andersen and Wagner, 2005).

Positive regulation of translation occurs when factors stabilize the bridging interactions between the 5′ and 3′ ends of the mRNA and their associated factors. Poly(A) Interacting Protein 1 (PAIP1) interacts with PABPC1, eIF4A, and the initiation factor eIF3 to enhance translation (Craig et al., 1998; Martineau et al., 2008; Roy et al., 2002). The Cytoplasmic Polyadenylation Element Binding Protein (CPEB) can promote reversal of deadenylation by recruiting the poly(A) polymerase GLD2 to lengthen the tails of a deadenylated transcript, allowing for PABPC1 binding and formation of the activated closed-loop structure (Ivshina et al., 2014; Kim and Richter, 2008).

Cis-acting regulatory elements, most commonly found in the UTRs of mRNAs, bind RBPs that have effects ranging from promoting translation to affecting localization of mRNAs to regulate transcript expression (Abaza and Gebauer, 2008; Buxbaum et al., 2015; Jackson et al., 2010). Examples include inhibition of PIC formation by the Iron Response Protein (IRP), which binds to a 5´ UTR RNA stem loop structure in the *Ferritin* mRNA when intracellular iron is low (Anderson et al., 2012; Muckenthaler et al., 1998; Thomson et al., 1999). CPEB is a bifunctional RBP; in contrast to the previous example of CPEB-mediated activation it can also bind to U-rich elements in 3´ UTRs during oogenesis and recruit 4E-BP to reducing expression of the bound mRNA transcript (Groisman et al., 2000; Stebbins-Boaz et al., 1999).

This overview highlights the prevalence and diversity of post-transcriptional regulatory pathways. Further examples reviewed elsewhere, and surveys of RBPs have identified additional regulatory mechanisms (Dominguez et al., 2018; Gehring et al., 2017; Harvey et al., 2018; Ryder, 2016). RBP-mediated regulation shares common themes of targeting closed loop formation through disruption of protein-protein interactions or removal of the 5´ cap or poly(A) tail. These structures bind proteins involved in translation, but also protect the transcript from mRNA decay pathways. In this way, regulation of translation and mRNA decay are inexorably linked to one another.

1.4 Canonical Pathways of mRNA Decay

In general, mRNAs undergo spatially and temporally regulated decay to control levels of protein expression. Intrinsic or extrinsic signals alter decay rates through the processes of deadenylation, decapping, and exonucleolytic decay. RNA decay pathways are initiated from either the 3´ or 5´ ends of the transcript, or by endonucleolytic cleavage,

and can act redundantly to efficiently degrade an mRNA (Figure 1.3) (Garneau et al.,

2007).

Figure 1.3: Canonical pathways of mRNA decay. (A) The PAN2/PAN3 complex begins deadenylation by shortening the poly(A) tail, which dissociates PABPC1. **(B)** The CCR4-NOT complex completes deadenylation through the action of two deadenylase subunits, CCR4, which can degrade the poly(A) tail bound by PABPC1, and **(C)** CAF1, which degrades the poly(A) tail sequences that are exposed when PABPC1 disassociates. The rest of the transcript is degraded by one of two alternative mRNA decay pathways. **(D)** The 5′ → 3′ decay pathway involves removal of the 7mG cap by the decapping complex of DCP1/DCP2, followed by **(E)** 5′ to 3′ directed decay of the remainder of the transcript by the XRN1 exonuclease. **(F)** Alternatively, the 3′ → 5′ decay pathway commences with the 3′ to 5′ exonucleolytic decay of the mRNA via the exosome, **(G)** leaving the 5′ 7mG cap which is processed by DCPS, the scavenging decapping enzyme that recovers free cap for eventual reuse. Figure adapted from Arvola et al, 2016.

1.4.I Deadenylation of mRNAs

Decay of cytoplasmic mRNAs frequently initiates through the action of deadenylases, magnesium-dependent exonucleases that degrade the poly(A) tail from the 3´ end, releasing 5´ AMP as a product (Goldstrohm and Wickens, 2008). Deadenylases generally fall into of two subfamilies. The DEDD domain deadenylases, named for the conserved glutamic and aspartic acid side chains in the active site, include PAN2, PARN, and CAF1 nucleases. The other deadenylases are members of the exonuclease, endonuclease, and phosphatase (EEP) superfamily, which include CCR4 (and its human homologs CNOT6 and CNOT6L), Nocturnin (NOCT), and Phosphodiesterase 12 (PDE12) based on the sequence conservation (Goldstrohm and Wickens, 2008). Much of the work characterizing mRNA decay has been completed in yeast; however, mRNA decay is highly conserved and much of this data can be generally extrapolated to mammalian systems (Garneau et al., 2007; Siwaszek et al., 2014).

Deadenylation is typically the first and rate-limiting step of mRNA decay and is thought to occur in two phases. Initial shortening of the poly(A) tail is catalyzed by the PAN2-PAN3 complex, shortening the tail down from ~250 to 60-80 nucleotides. The PAN2 deadenylase subunit requires the PAN3 dimer for activity and is stimulated by PABPC1 (Boeck et al., 1996; Brown et al., 1996; Yamashita et al., 2005; Yi et al., 2018). The second phase of deadenylation completes the degradation of the remaining poly(A) tail by the CCR4-NOT deadenylase complex (Yamashita et al., 2005; Yi et al., 2018). The CCR4-NOT complex is associated with two separate deadenylase subunits, Carbon catabolite repression 4 (yeast Ccr4; human CNOT6 and CNOT6L) and Caf1 (POP2 in *Drosophila*, CNOT7 and CNOT8 in humans) along with seven other components with

various functions outside of deadenylation (Tucker et al., 2001). CCR4-NOT is recruited to transcripts through interactions with RBPs and miRNA-mediated decay pathways. Recently, CC4 and CAF1 have been described as catalyzing alternating deadenylation, based on the differential ability of these subunits to deadenylate the poly(A) sequence associated with PABPC1. While CCR4 is able to degrade PABPC1-bound RNA, leading to the dissociation of PABPC1 RRM domains from the RNA, CAF1 is only able to degrade RNA in which PABPC1 is not bound, and has activity against the RNA that has been cleared during CCR4-mediated deadenylation (Webster et al., 2018; Yi et al., 2018).

In general, deadenylases function in multiprotein complexes to remove the poly(A) tails of mRNAs, effecting stability and translation of their target transcripts (Goldstrohm and Wickens, 2008). The presence of multiple deadenylases with apparent redundancies in function belies the importance of this process in gene regulation as well as some of the important remaining questions in the field. It is probable that deadenylases have both overlapping and non-overlapping sets of substrates, and that their activities are differentially modulated by both extrinsic and intrinsic signals or expression of different sets of RBPs that can preferentially recruit different deadenylase enzymes. Ultimately, these regulatory processes may change which of the deadenylases are dominant in regulation of mRNAs involved in specific biological processes.

1.4.II Decapping of mRNA transcripts

Deadenylation is followed by one of two pathways that enzymatically degrade the remainder of the mRNA transcript. The 5´ to 3´ directed pathway begins with removal of the 5´ 7mG cap via the action of the decapping complex, a heterodimer of Decapping Protein 1 (DCP1) and Decapping Protein 2 (DCP2) (Grudzien-Nogalska and Kiledjian,

2017). Decapping factors are recruited by the deadenylase machinery, though deadenylase-independent decapping also occurs (Alhusaini and Coller, 2016; Badis et al., 2004; Behm-Ansmant et al., 2006; Yamashita et al., 2005). Within the DCP1/DCP2 heterodimer, DCP2 is responsible for enzymatic decapping (Wang et al., 2002). Additional decapping factors such as DXO enzymes remove caps from mRNAs that are missing N7 methylation modifications, leading to their rapid destruction (Grudzien-Nogalska and Kiledjian, 2017). The removal of the 5´ 7mG cap is highly regulated, with multiple protein partners controlling efficiency (Jonas and Izaurralde, 2013; Piccirillo et al., 2003; She et al., 2008). This process is activated by RNA helicases such as DHH1, the GE1 decapping factor, and other factors such as PAT1, which recruits the LSM complex to the 3´ deadenylated end to further promote decapping (Ling et al., 2011; She et al., 2008; Tharun and Parker, 2001; Yu et al., 2005). While deadenylation and decapping are reversible, decapping is frequently the step in which the fate of the mRNA is determined and the message is destroyed (Trotman and Schoenberg, 2019).

1.4.III 5´ to 3´ directed mRNA decay

In the 5´ to 3´ decay pathway, decapping is followed by exoribonucleolytic decay by 5´ – 3´ Exoribonuclease 1 (XRN1) and other 5´ exoribonucleases (Nagarajan et al., 2013). XRN1 is recruited by decapping factors and recognizes the 5´ monophosphate of the RNA product post-decapping, rapidly and processively degrading the mRNA (Fischer and Weis, 2002; Nagarajan et al., 2013; Sinturel et al., 2012). While XRN1 is the major 5´ exoribonuclease in the cytoplasm, other homologs of this protein, such as XRN2, function in the nucleus and are involved in maturation of rRNAs and small nucleolar RNAs

(snoRNAs), termination of RNA polymerase II transcription, and destruction of aberrant mRNAs (Das et al., 2003).

1.4.IV 3′ to 5′ directed mRNA decay

The exosome complex, comprised of 10-12 subunits, degrades deadenylated mRNAs via the 3′ exoribonuclease activity of the Dis3 subunit, leaving a 5′ capped polynucleotide (10$\geq$ nt) (Liu et al., 2006; Reis et al., 2013; Siwaszek et al., 2014). A second exoribonuclease subunit, RRP6, is also associated with the nuclear exosome complex (Wasmuth and Lima, 2017). Nuclear and cytoplasmic forms of the exosome have been identified, and the complex functions in other processes beyond mRNA decay (LaCava et al., 2005; Zinder and Lima, 2017). After the mRNA is degraded, the scavenger decapping enzyme (DCPS) hydrolyzes the cap into 7mG monophosphate and a nucleotide diphosphate (Chen et al., 2005; Liu and Kiledjian, 2005).

Individual mRNA transcripts can be degraded through the action of the 3′ to 5′ pathway, the 5′ to 3′ pathway, or through both pathways. What determines pathway preference for a certain mRNA transcript is not well understood. In addition to the nucleases described in canonical mRNA decay pathways, enzymes with very similar or identical activities could feasibly also function in mRNA decay. For example, the deadenylases, PAN2, CCR4, and CAF1, are well-described in mRNA decay pathways. However, other deadenylases including PARN, NOCT, ANGEL1 and ANGEL2, and PDE12 could also feasibly play a role (Goldstrohm and Wickens, 2008; Pearce et al., 2017; Rorbach et al., 2011; Wang et al., 2010). These other proteins likely function in independent mRNA decay pathways, including alternative pathways of mRNA quality control. It is also possible that other deadenylases feed mRNAs into the 3′ to 5′ and 5′ to

3´ mRNA decay pathways and that alternative deadenylases regulate the decay rates of specific subsets of transcripts. These questions are incompletely understood in the mRNA decay field and remain important avenues of study in learning the precise mechanisms of gene control at the mRNA level.

1.5 Molecular Functions of Human Deadenylases
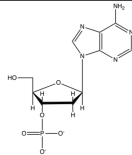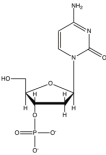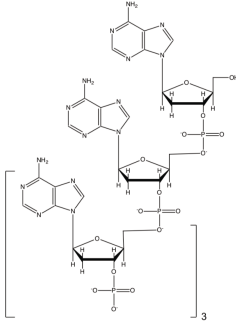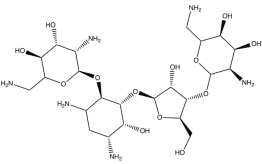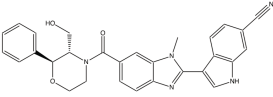
Yeast Ccr4 (which has human homologs CNOT6 and CNOT6L) is the founding member of the EEP deadenylase subfamily (Doidge et al., 2012; Maryati et al., 2015; Tucker et al., 2001). As a subunit of the CCR4-NOT complex, Ccr4 promotes mRNA decay through its deadenylase activity in canonical mRNA decay pathways (Figure 1.3) (Doidge et al., 2012; Maryati et al., 2015; Temme et al., 2010; Tucker et al., 2002). While these deadenylase enzymes were first described as poly(A)-specific 3´ exoribonucleases, non-mRNA substrates have been identified in the case of PDE12, which is active against 2´, 5´ oligoadenylate (2´,5´-A) and tRNAs with erroneous polyadenylation at the 3´ end (Pearce et al., 2017; Wood et al., 2015).

NOCT was first classified as a member of the EEP deadenylase subfamily based on its sequence homology to Ccr4 (Baggs and Green, 2003). As biochemical characterization of NOCT had been limited prior to the studies discussed in Chapter 3, comparisons between NOCT and better characterized deadenylases can give insight into funcntion. The validity of this approach is highlighted by sequence conservation of the catalytic EEP domain among the 6 mammalian EEP deadenylases NOCT, CNOT6, CNOT6L, PDE12, ANGEL1, and ANGEL2 (Figure 1.4) (Ashkenazy et al., 2016; Ashkenazy et al., 2010; Celniker G., 2013; Glaser et al., 2003; Landau et al., 2005). This sequence conservation is reflected in the structural conservation between EEP family

15

members and highlights the functional and/or structural importance of the conserved residues. Although structures of these enzymes in complex with RNA substrates have yet to be reported, both PDE12 and CNOT6L have been crystallized in complex with inhibitors and other nucleotides, offering potential insight into the molecular mechanism of substrate binding (Wang et al., 2010; Wood et al., 2015; Zhang et al., 2016). These structures will be described here and summarized in the context of what may be informative to the function of NOCT.

Multiple structures of the catalytic (EEP) domain of human CNOT6L have been reported, including structures of the apoenzyme (Protein Data Bank (PDB) accession code 3NGQ) as well as in complex with multiple ligands, including AMP (3NGN) and poly(A) DNA (3NGO) (Table 1.1) (Wang et al., 2010). These structures illustrate the expected Mg^{2+}- bound hydrolase fold observed in EEP-deadenylases. Mutation of the metal-binding residues Glu240 and Asp489 to alanine ablates enzyme activity completely *in vitro*. CNOT6L has the highest activity against poly(A) RNA substrates compared to other RNA homopolymers (e.g., poly-uridine or poly-cytosine) and prefers RNA over DNA substrates, all which support the described function of CNOT6L as a classical CCR4-type deadenylase (Wang et al., 2010).

The CNOT6L catalytic domain has been co-crystallized individually with either AMP or a poly(A) DNA 5-mer to approximate the interactions between the enzyme and bound nucleotide substrates (Figure 1.5). In the CNOT6L-DNA co-crystal, the deoxyribose of A_1 and the complete A_2 and A_3 nucleotides are resolved, occupying the binding cleft and active site with the A_2 phosphate bond poised for hydrolysis. The A_2 phosphate is coordinated by the two active-site Mg^{2+} ions as well as by Asn412. This

Enzyme	PDB	Ligand Name	Ligand structure
CNOT6L	3NGN	Adenosine 5' Monophosphate	
CNOT6L	5DV2	Cytidine 5' Monophosphate	
CNOT6L	3NGO	poly(A)$_5$ DNA	
CNOT6L	5DV4	Neomycin	
PDE12	4Z2B	GSK3036342A	

Table 1.1: Chemical structures of ligands co-crystalized with human deadenylases. In addition to nucleotide ligands, PDE12 and CNOT6L have been crystallized in the presence of the inhibitors neomycin and GSK3036342A, respectively (Wang et al., 2010; Wood et al., 2015; Zhang et al., 2016).

substrate conformation would release a dinucleotide product instead of AMP, suggesting that the DNA, a non-native substrate of CNOT6L, may bind in the active site in a manner distinct from the endogenous RNA substrate. The rings of the adenine DNA base A_2 interact through base-stacking interactions with Phe484, an interaction which is preserved in the CNOT6L-AMP co-crystal, suggesting that Asn412 and Phe484 could be important for binding of poly(A) substrates. While this interaction was observed for both poly(A)

DNA and AMP co-crystals, there are no Mg^{2+} ions resolved in the active site of the AMP

co-crystal, which are predicted to be important for coordinating the phosphate groups of

CNOT6L substrates.

Figure 1.4: Conservation of the catalytic domain in mammalian EEP deadenylases. The ConSurf server was used to generate a color-coded conservation alignment mapped to the NOCT amino acid sequence. The multiple sequence alignment (MSA) was generated from 489 different sequences representing mammalian NOCT, CNOT6, CNOT6L, PDE12, ANGEL1 and ANGEL2, which were curated from UniProt and aligned using Clustal W. The MSA was submitted to ConSurf for analysis and the results are represented here. The NOCT catalytic domain is underlined in orange, the conserved active site residues are demoted by a yellow star, and the NOCT basic patch is denoted by blue triangles.

Though these interactions have been hypothesized to be responsible for poly(A)

specificity, the interactions identified through the use of non-native substrates must be

validated using biochemical approaches and RNA substrates to fully elucidate the

mechanisms of binding and catalysis. Co-crystals of CNOT6L with CMP (5DV2) and

neomycin (5DV4) suggest that the CNOT6L active site displays a certain degree of

plasticity, enabling the binding of nucleotides other than adenine as well as non-nucleotides (Figure 1.5 and Table 1.1) (Ren et al., 2002; Zhang et al., 2016).

Figure 1.5: CNOT6L structure active sites with bound ligands. CNOT6L structures for the apoenzyme (3NGQ), or bound to AMP (3NGN), poly(A) DNA (3NGO), CMP (5DV2) or neomycin (5DV4) are shown. CNOT6L is depicted in yellow, Mg^{2+} ions are depicted as grey spheres, and ligands are depicted in cyan. Interactions between Mg^{2+} and between active site residues are shown as dashed lines. Hydrogen bond cutoffs are set to 3.5 Å, where applicable.

PDE12 (also known as 2´5´-PDE or 2´PDE) was first described as a member of the type-1 interferon (IFN) pathway, which plays a role in resisting infection by viruses with positive-strand RNA genomes (Ezelle and Hassel, 2012; Li et al., 2011; Silverman, 2007). PDE12 degrades 2´,5´-A, an oligoadenylate nucleotide activator of RNAse L, which degrades viral RNA (Wood et al., 2015). Later, PDE12 was described to localize to the mitochondrial matrix in human cells and was originally believed to be active against mitochondrial mRNAs (mt-mRNAs) (Pearce et al., 2017; Rorbach et al., 2011). Because PDE12 overexpression resulted in shorter mt-mRNA tail lengths, it was predicted that

PDE12 was a mitochondrial deadenylase (Poulsen et al., 2011; Rorbach et al., 2011). However, mt-mRNA tail length and stabilitiy is not affected by PDE12 knockout, suggesting previously observed reductions in mt-mRNA tail length could be explained by PDE12 overexpression. Further studies demonstrated that the likely cellular function of PDE12 is in mitochondrial tRNA quality control. Deletion of PDE12 in HEK293 cells significantly impaired mitochondrial translation, as a subset of mt-tRNAs had erroneous 3´ polyadenylate sequences that impaired charging with amino acids and increased mitoribosome stalling, reducing translational efficiency (Pearce et al., 2017). The deadenylase activity of PDE12 maintains the quality of mt-tRNAs by deadenylating erroneous 3´ poly(A) extensions (Pearce et al., 2017).

The crystal structure of PDE12 shows that this enzyme adopts an α/β hydrolase fold with a single bound Mg^{2+}ion in the active site, which is coordinated Glu351 (Figure 1.6). This residue is highly conserved in EEP enzymes and corresponds to the Mg^{2+}-binding residues in NOCT (Glu195) and in CNOT6L (Glu240). The viral infection-resistant phenotype of PDE12 knockout cells suggests that this enzyme may be a potential antiviral drug target (Kubota et al., 2004; Wood et al., 2015). A screen to identify potential lead compounds led to identification of inhibitor, GSK3036342A, that was ultimately co-crystallized with PDE12. These structures allow for comparison of ligand binding between CNOT6L and PDE12 (Wang et al., 2010; Wood et al., 2015). A comparison of this structure with the nucleotide-bound structures of CNOT6L shows that the inhibitor is bound in a similar orientation to the poly(A) DNA ligand in CNOT6L, though the inhibitor structure is largely different than that of a nucleotide (Figure 1.6 and Table 1.1). The inhibitor bound structures of PDE12 and CNOT6L illustrate that common residues are

involved in ligand binding, for example, Pro461 and Phe556 in PDE12 interact in base stacking interactions and are analogous to interaction in bound ligands and CNOT6L residues Pro365 and Phe484. However, these interactions have only been validated in biochemical assays for CNOT6L (Figures 2.2 and 2.3) (Wang et al., 2010; Wood et al., 2015). Steady state kinetic studies were utilized to confirm that the PDE12 inhibitors were acting as substrate competitors, which, along with the co-crystal structure, indicates that the interactions between ligands and active site are flexible and can accommodate the binding of non-native ligands through the apparent plasticity of the enzyme active sites. While both CNOT6L and PDE12 have activity against 2´,5´-A, PDE12 is significantly more active against this substrate and PDE12 inhibitors do not affect CNOT6L activity (Wood et al., 2015).

Figure 1.6: PDE12 structure active sites with bound ligands. PDE12 structures for the apoenzyme (4Z0V) or bound to the GSK3036342A inhibitor (4Z2B) are shown. PDE12 is depicted in green, Mg^{2+} ions are depicted as grey spheres, and ligands are depicted in cyan. Interactions between Mg^{2+} and between active site residues are shown as dashed lines. Hydrogen bond cutoffs are set to 3.5 Å, where applicable.

These data indicate that there are differences in how the substrates and ligands are recognized by the different EEP deadenylases, although the mechanism of this is not fully understood and likely lay outside of the active site. The apparent differences in EEP deadenylase function can be addressed in part through structural studies that can

highlight differences between the human EEP deadenylases and can yield insights into the relative specificities and unique features within this family of enzymes. We next review the observed phenotypes of NOCT and the biological processes in which NOCT is proposed to be involved, which indicate that NOCT has some specificity for mRNAs involved in lipid metabolism and storage.

1.6 Biological Roles of NOCT

NOCT is most closely related to the CCR4 deadenylase subunit of the CCR4-NOT complex by sequence homology, and thus, when first reported in the literature, was described to be an mRNA deadenylase. Initially, NOCT was described in a study that identified circadian transcripts within the *Xenopus* retina, which reported that NOCT has deadenylase activity (Green and Besharse, 1996a, b, 1997). This activity was later reported in the murine homolog as well (Garbarino-Pico et al., 2007).

The effect of depletion of NOCT has been examined at the cellular level and in mice. These studies determined that depletion of NOCT led to resistance to obesity on a high fat diet (HFD) and defects in lipid absorption, trafficking, and adipogenesis (Douris et al., 2011; Green et al., 2007; Kawai et al., 2010c). These observations suggest that the deadenylase activity of NOCT regulates the expression of mRNA transcripts associated with lipid metabolism. Here, the existing studies on NOCT function at the organismal and cellular levels are reviewed and commented on in the context of putative NOCT enzymatic activity (Figure 1.4).

1.6.I Phenotypes related to the biological functions of NOCT

The phenotypes associated with NOCT suggest ribonuclease activity against a more specialized subset of mRNAs in contrast to the global regulation observed for CCR4 deadenylase orthologs CNOT6 and CNOT6L. NOCT knockout (KO) mice are viable and showed little difference from the wild type (WT) animals on a standard chow diet. When fed on an HFD, the WT mice become obese whereas the NOCT KO mice remain lean. This resistance to diet-induced obesity manifests in the reduced size of visceral adipose tissue in the KO versus WT animals. In addition, the KO mice accumulate fewer lipid deposits in their livers while on a HFD, the tissue where *Noct* mRNA expression is highest in mice (Douris et al., 2011; Green et al., 2007). These changes in body composition and the lack of weight gain seen in the NOCT KO animals on an HFD suggest that NOCT may function by regulating transcripts related to lipid metabolism, transport, and storage (Figure 1.7).

NOCT appears to play a role in regulating absorption of dietary lipids in the small intestine, providing a potential explanation for the resistance to weight gain on an HFD. NOCT expression in the proximal small intestine is congruent with a putative function in known circadian-controlled processes involved in digestion and nutrient absorption (Hoogerwerf et al., 2007; Pan and Hussain, 2007, 2009). *Noct* mRNA expression fluctuates by approximately 5-fold in the small intestine, increasing in expression between Zeitgeber time (ZT) 0 (light onset) and ZT12 (dark onset) (Douris et al., 2011). This increase in *Noct* expression coincides with the expected feeding time for mice (Kavaliers et al., 1985).

Figure 1.7: Biological roles of NOCT have been observed at multiple levels of complexity, from molecular to physiological. At the molecular level, NOCT is observed to take part in regulation of translational control and mRNA metabolism, which appears to effect cellular differentiation and metabolism. These functions ultimately modulate body mass, development, and lipid metabolism. Figure by E. Abshire and taken from Hughes, K.L., Abshire, E.T., and Goldstrohm, A.C., 2018.

NOCT KO mice were then treated with dietary fats to determine if the lean phenotype on an HFD could be explained by defects in lipid transport. NOCT-deficient mice fed an olive oil gavage had significantly lower plasma triglyceride and cholesterol levels versus WT mice, indicating that they may have defects in dietary lipid absorption. The proximal small intestines of NOCT KO mice retained significantly more lipids, indicating that decreased plasma lipid levels of KO mice may be due to defects in lipid trafficking from the small intestine. Consistent with this observation, isolated intestinal enterocytes retained these lipids in intracellular droplets instead of secreting them for

24

transport to other tissues (Douris et al., 2011). Deficits in lipid absorption and transport provide a reasonable explanation for resistance to weight gain on HFD, though the mechanism by which NOCT promotes absorption and trafficking of ingested lipids is undefined. Furthermore, the unabsorbed dietary lipids were not reported to be found in fecal matter, leaving the fate of these lipids ambiguous (Douris et al., 2011).

1.6.II Proposed function of NOCT in negative regulation of osteogenesis

The effects of NOCT are not limited to lipid absorption and storage, as the NOCT KO animals have increased bone mass relative to WT animals (Green et al., 2007; Kawai et al., 2010c). These observations suggest that NOCT may play a role in fate determination of a subset of progenitor mesenchymal stem cells (MSCs) that are differentiated into either adipose or osteoblasts (Chen et al., 2016; James, 2013; Yuan et al., 2016). Imbalances in this regulation cause diseases such as osteoporosis which is marked by increased bone adiposity (Choi et al., 2017; You et al., 2016). NOCT is negatively associated with osteogenesis in MC3T3-E1 pre-osteoblast cells and primary calvarial osteoblasts (COBs), as levels of NOCT mRNA and protein decrease during COB differentiation. Short hairpin (shRNA)-mediated knockdown of NOCT in MC3T3-E1 cells increased levels of mineralization, alkaline phosphatase activity, and expression of osteoblastic mRNAs in cultures treated with osteoblastogenic factors. In contrast, overexpression of NOCT in MC3T3-E1 cells had the opposite effect (Kawai et al., 2010c).

In COBs lacking NOCT, an increase in *Runx2* mRNA, which expresses a master positive regulator of osteoblastogenesis, was reported to increase by less than two-fold (Kawai et al., 2010c). Transcript levels of other important osteoblastogenic transcripts, *Osterix*, *Osteocalcin*, and *Atf4* were examined in MC3T3-E1 cells overexpressing NOCT

and undergoing osteoblastogenesis. In all cases, the transcript levels decreased with NOCT overexpression though changes were also modest in these experiments with the exception of *Osteocalcin* mRNA, which decreased 4 to 5-fold (Kawai et al., 2010c). None of these transcripts have been examined for direct association with NOCT so it in unknown if this effect is due to NOCT ribonuclease activity. In an additional study, NOCT was reported to interact with *Igf1* mRNA, a transcript that is associated with accumulation of bone mass and skeletal tissue maintenance (Kawai et al., 2010a). *Noct* mRNA was shown to be expressed in mouse femur in a circadian manner antiphase to *Igf1* mRNA, and overexpression of NOCT in MC3T3-E1 cells reduced *Igf1* expression. Furthermore, NOCT was shown to be directly associated with *Igf1* mRNA. However, regulation of the *Igf1* 3′UTR in reporter assays was dependent on which mouse strain the sequence of the *Igf1* 3′ UTR was derived from, which indicates that there is natural variation in what transcripts are regulated by NOCT (Kawai et al., 2010a). As such, the mechanism of NOCT-mediated regulation of the adipogenesis/osteoblastogenesis differentiation axis has yet to be determined.

1.6.III Proposed function of NOCT in positive regulation of adipogenesis

The obesity-resistant phenotypes of the NOCT KO mice indicate that NOCT is likely a positive regulator of adipogenesis, (Green et al., 2007). Subsequent studies in 3T3-L1 pre-adipocytes demonstrated that differentiation of these cells into adipocytes increased the level of NOCT mRNA and protein expression (Kawai et al., 2010c). mRNA levels of *peroxisome proliferator-activated receptor γ* (*Pparγ2*), a well-described proadipogenic factor, increases coincidently with *Noct* expression during adipogenesis of

3T3-L1 preadipocytes, whereas knockdown of NOCT reduced *Pparγ2* expression (Kawai et al., 2010c; Kawai and Rosen, 2010).

Overexpression of NOCT in differentiating 3T3-L1 cells led to increased accumulation of intercellular lipids relative to controls, a marker for adipogenesis. mRNA levels of the adipogenic differentiation marker *Adipocyte protein 2* (*aP2*) were elevated in the NOCT overexpression (OE) cells. The opposite effect was observed in differentiated 3T3-L1 cells wherein NOCT was depleted, with reduced lipid staining and decreased expression of adipogenic mRNAs (Kawai and Rosen, 2010). It was posited that a physical association between PPARγ and NOCT enhanced nuclear translocation of PPARγ, increasing its transcriptional activity. However, the exact mechanism of translocation remains unclear (Kawai and Rosen, 2010). Although NOCT loss of function models have impaired adipogenesis and overexpression models have increased markers of differentiation, the mechanism by which NOCT promotes adipogenesis has still not been established. The proposed exoribonuclease function of NOCT would predict that NOCT-mediated RNA decay would target transcripts associated with negative regulation of adipogenesis. These observations necessitate further study into the relationship between NOCT targets and fate determination of adipogenic precursors.

1.6.IV Circadian regulation of NOCT expression

Early studies of NOCT function focused on the potential role of the enzyme in circadian function. NOCT was first identified in a screen for genes regulated by the circadian clock in the retina of *Xenopus laevis*, but the majority of characterization of the circadian regulation of NOCT expression has been performed in mice (Baggs and Green, 2003, 2006; Green and Besharse, 1996a, b, 1997; Green et al., 2007; Wang et al., 2001).

In mice and in *Xenopus*, the zenith of NOCT expression occurs just after the onset of the dark phase in a 12:12 light-dark cycle. *Noct* mRNA levels fluctuated most significantly in the liver (~30-fold), where the expression of *Noct* is generally the highest versus other tissues. Smaller fluctuations were reported in heart and kidney (Wang et al., 2001). NOCT KO mice do not display defects in circadian rhythms or behaviors, which indicates that it is downstream of circadian effectors (Green et al., 2007).

The NOCT promoter in mice and humans have conserved E-box elements that bind the core clock regulators BMAL1 and CLOCK (Li et al., 2008). However, the time-dependence of these interactions varies by species—in mice, BMAL1 association with E-boxes varies over the 24-hour cycle, while in humans, CLOCK and BMAL1 occupancy at this site is independent of the time of day. Binding of RevERBα, another core clock transcription factor, has also been described in mice and humans (Li et al., 2008). Transcriptional regulation of NOCT does not appear to be wholly controlled by the global molecular clock, as mice deficient in the global circadian clock via loss of CLOCK continue to rhythmically express NOCT in the liver, suggesting mechanisms of local control (Kornmann et al., 2007; Oishi et al., 2003). Overall, studies of NOCT transcriptional regulation remain limited, and the consequences of disrupting binding of CLOCK, BMAL1, or RevERBα in the NOCT promoter remain untested.

Studies of circadian regulation of NOCT remain largely limited to measurements of mRNA abundance. To date, only one study has measured fluctuations in NOCT protein expression in mouse livers over a 24-hour time period, and the magnitude of these changes was less than two-fold (Kojima et al., 2010). In contrast, *NOCT* mRNA levels fluctuate in human Huh7 hepatocarcinoma cells and mouse liver tissues around 30-fold

(Li et al., 2008). As the protein fluctuation in mouse liver is comparably minor compared to that of the mRNA, the significance of circadian rhythmicity of the mRNA remains unclear. Circadian fluctuation of *NOCT* also occurs in human oral mucosa, further suggesting that rhythmic expression occurs inn peripheral tissues; however, only mRNA levels were examined, and the magnitude of change between the light and dark phases was small (Zieker et al., 2010). An additional transcriptome analysis also provides evidence for rhythmic expression in humans in skeletal muscle (Perrin et al., 2018). The resistance of NOCT KO mice to diet-induced obesity suggests that NOCT functions in metabolic regulation, which is also tightly controlled by the circadian clock. Circadian regulation of NOCT may be relevant to the reported phenotypes of NOCT null mice, as livers of KO mice show changes in circadian expression of metabolic mRNAs and circadian fluctuations in levels of circulating lipids and bile acids (Stubblefield et al., 2018).

1.6.V Regulation of NOCT expression in response to external stimuli

NOCT expression is regulated by extracellular stimuli, including nutritional states, growth factors, and mitogens (Douris et al., 2011; Garbarino-Pico et al., 2007; Gilbert et al., 2011). This finding is consistent with the well-described relationship between metabolism and circadian regulation (Eckel-Mahan and Sassone-Corsi, 2013; Mayeuf-Louchart et al., 2017). *Noct* mRNA expression was induced in NIH 3T3 cells by serum shock and also by the phorbal ester 12-O-Tetradecanoylphorbol-13-acetate (TPA), leading to the interpretation that NOCT expression is also modulated by external cues (Garbarino-Pico et al., 2007). It should be noted that TPA treatment has been shown to reset the circadian clock, which would indicate that in this case, external stimulation of NOCT may also affect its circadian expression (Akashi and Nishida, 2000). A similar

29

effect on the circadian clock has also been shown for serum shock (Balsalobre et al., 1998). These responses cannot fully be attributed to clock resetting, as *Noct* mRNA increases to its maximal expression in the first two hours after TPA or serum treatment. This is a more rapid response than what is observed over circadian time, in which *Noct* reaches maximal expression over six hours (Garbarino-Pico et al., 2007; Wang et al., 2001). The promoter of the mouse *Noct* gene contains binding sites for PPARγ, CRE, and STAT3, while the human *NOCT* promoter has an observed FoxO binding site (Bourillot et al., 2009; Hamza et al., 2009; Kawai et al., 2010b; Paik et al., 2007). These transcription factor binding sites are indicative of the potential for multi-level control of *NOCT* transcription in response to nutrient states or other external cues.

Evidence of NOCT induction by nutritional stimuli exists at the organismal level. In mice, an olive oil gavage induces expression of NOCT at ZT3 in the small intestine, when NOCT expression is typically low based on patterns of circadian expression (Douris et al., 2011). This is especially pronounced in the proximal small intestine, where the majority of dietary fat absorption takes place during digestion.

Overall, NOCT expression appears to be jointly regulated by the circadian clock and external stimuli, especially signals that are related to nutritional state. These regulatory processes modulate the activity and expression of gene products related to metabolism. Therefore, NOCT regulation cannot necessarily be cleanly separated into regulation by the clock versus external signals. It should be noted that the mechanisms of direct regulation remain unexplored. For example, stimulation of NOCT expression by TPA and serum shock has been observed but the signaling pathways that exert these changes have not been identified. Describing how NOCT is modulated will be important

in understanding how perturbations in expression can affect body mass and lipid metabolism, and how these regulatory mechanisms are translated into post-transcriptional regulation downstream of NOCT.

1.6.VI Global analysis of NOCT-regulated transcripts

The molecular mechanism by which NOCT regulates body mass and lipid metabolism has been a persistent question in the field of NOCT biology. Identification of endogenous mRNA targets of NOCT has been pursued but still remains elusive. Determination of these targets is critical for further characterization of the NOCT molecular mechanism, as *in vitro* assays are limited by the lack of knowledge of natural substrates. While this lack of ribonuclease activity may also be a function of missing protein partners in *in vitro* experiments (discussed in Chapters 2 and 6), it is also possible that the substrates tested *in vitro* lack sequence or structural elements that are required for NOCT function. Such sequence elements may be able to be identified from datasets of NOCT-regulated transcripts. Several studies have reported efforts to identify NOCT-regulated transcripts by looking at global changes in gene expression (Hee et al., 2012; Kojima et al., 2015; Stubblefield et al., 2018).

Based on its relationship to exoribonucleases and its ability to repress mRNAs when directed to a reporter transcript, we would predict that NOCT targets will increase in abundance when NOCT is depleted. A study by Hee *et al.*, assessed changes in transcript levels upon treatment of differentiated 3T3-L1 cells with NOCT shRNAs. 89 upregulated genes and 184 downregulated genes were reported, which may include both direct targets of NOCT-mediated regulation as well as genes that are indirectly affected.

A)

Category	Term	Gene Count	P-Value
GOTERM_BP_DIRECT	brown fat cell differentiation	4	0.00032
GOTERM_BP_DIRECT	oxidation-reduction process	10	0.0023
GOTERM_BP_DIRECT	glucose transport	3	0.005
GOTERM_BP_DIRECT	developmental process involved in reproduction	2	0.0019
GOTERM_BP_DIRECT	lipid metabolic process	6	0.0034
GOTERM_BP_DIRECT	hepatocyte apoptotic process	2	0.0054
GOTERM_BP_DIRECT	negative regulation of lipid catabolic process	2	0.0061
GOTERM_BP_DIRECT	phospholipid efflux	2	0.0061
GOTERM_BP_DIRECT	cellular response to hypoxia	3	0.0062
GOTERM_BP_DIRECT	immune system process	5	0.0064

B)

Gene	Fold Increase	Gene Ontology
Slc16a1	9.8	Mitochondria
Cox6a2	7.9	Oxioreductase, Mitochondira, Obesity resistance
Ero1l	6.8	ER Stress and Apoptosis
Paqr5	6.7	Fat metabolism
Elovl3	6.7	Obesity resistance
Akp2	6.4	Bone mineralization
Cideda	6.2	Lipid storage, Mitochondria, Obesity resistance
Cyp2c44	5.9	Oxioreductase, Fat metabolism
Ccl24	5.3	Immune response
Apoc1	5.2	Lipid transport and metabolism

Table 1.2 Analysis of differentially regulated genes in 3T3-L1 cells with depleted NOCT. (A) DAVID GO Term analysis of the 89 upregulated genes identified by microarray in Hee et al. **(B)** Top ten upregulated genes identified by microarray in Hee et al.

Genes that exhibited more than two-fold differences in expression were analyzed using Ingenuity Pathway Analysis (IPA), which identified cellular proliferation, movement, and development as the main pathways enriched in these gene sets (Hee et al., 2012). We took the 89 upregulated genes from this dataset and reanalyzed them using DAVID, an open-source platform for gene ontology analysis, which identified the enrichment of genes involved in lipid metabolism and mitochondrial function (Table 1.2) (Huang da et al., 2009a, b). Different databases are used for DAVID and IPA, which may explain the differences in the identified pathways. Overall, the DAVID analysis provides terms that are more consistent with the described NOCT phenotypes related to lipid metabolism and

transport (Douris et al., 2011; Green et al., 2007; Kawai et al., 2010c). It remains unknown whether or not any of the upregulated transcripts physically associate with NOCT.

Other global analyses of NOCT-mediated gene expression have elected to use poly(A) tail length and circadian expression patterns as criteria for identification of putative NOCT-regulated transcripts. Circadian time points, measured by Zeitgeber Time (ZT), standardize light onset to ZT0 and dark onset as ZT12. *Noct* mRNA reaches maximal expression after dark onset to ZT16 and has minimal expression at ZT4 on the other end of the light:dark cycle (Wang et al., 2001). Kojima *et al*. performed an analysis of the changes in global poly(A) tail lengths of mRNAs purified from the livers of WT and NOCT KO mice at ZT4 and ZT16 (Kojima et al., 2015). RNAs were separated into long and short poly(A) tail populations based on relative binding affinities for oligo (dT) resin. One limitation of this approach is negative bias against recovery of mRNAs with very short tails. Though this was controlled for by measuring correlation of between fractionated and unfractionated samples, this bias may be reflected in the higher variability of long to short tail length ratios in the KO vs WT animals (Kojima et al., 2015). The assumption that NOCT targets would show significant differences in poly(A) tail lengths may not be accurate, as mRNA decay intermediates often do not accumulate in cells (Figure 4.7) (Abshire et al., 2018). While it was assumed that NOCT-mediated tail-length shortening would correlate to the rhythmic expression of *Noct* mRNA, only one study has shown that NOCT protein levels are also circadian and the timing of peak protein expression has not been assessed (Kojima et al., 2010; Kojima et al., 2015). In total, only 10 transcripts were identified by that had increased tail lengths at ZT16 in the NOCT KO samples and only 213 transcripts overall that had longer poly(A) tails in the NOCT KO mice at both ZT16

and ZT4. The combined ZT16 and ZT4 datasets were enriched for transcripts involved in oxidative phosphorylation, suggesting a role for NOCT in mitochondrial biology. None of the 10 transcripts with longer tails at ZT16 in the NOCT KO livers are involved in metabolic function (Kojima et al., 2015). Ultimately, it is not known if these mRNAs represent *bona fide* NOCT targets. The approach used in this study limits identification of targets from the outset by assuming poly(A) tail shortening and circadian expression of these transcripts (Kojima et al., 2015). Therefore, the pool of identified potential targets is restricted and unlikely to capture the full effect of NOCT regulation.

A third study analyzed circadian changes in mRNA levels in livers from WT and NOCT KO mice using RNA-Seq. This analysis similarly assumed that *bona fide* NOCT targets would display cyclic behavior for both poly(A) tail lengths and mRNA abundance (Stubblefield et al., 2018). Transcripts that had increased amplitudes of circadian expression in KO mice were identified and included mRNAs involved in acetyl CoA, triglyceride, and cholesterol production (Stubblefield et al., 2018). To assess potential differences in lipid metabolism between the KO and WT mice, the authors measured levels of circulating and hepatic triglycerides and cholesterol, as well as the levels of circulating lipoproteins in the plasma (Stubblefield et al., 2018). Most of these changes in lipid and lipoprotein concentrations between the WT and KO animals were modest so it is unclear that NOCT KO has a large effect on circulating and hepatic lipids. As homeostasis of lipid metabolism is tightly controlled, small changes are not necessarily unexpected. A complex endocrine signaling network coordinates metabolism in different tissues to respond to changes in energy states, which may allow for compensatory mechanisms that minimize the effects of NOCT KO (Hardie, 2012). The most notable

changes observed in the NOCT KO were increased gallbladder volumes during the light phase relative to WT mice (Stubblefield et al., 2018). This observation indicates that NOCT plays a role in bile acid production, which is vital for digestion of dietary fats (Qi et al., 2015; Stubblefield et al., 2018). Finally, the poly(A) tail lengths for a subset of transcripts identified in this study were analyzed for changes in the distribution of tail lengths between WT and NOCT KO mice. Very little difference was observed in the tail lengths of these transcripts between WT and NOCT KO mice, though this does not rule out differential regulation of transcripts by NOCT through methods of translational control (discussed in Chapter 4).

Past studies analyzing NOCT-mediated changes in gene expression used a series of assumptions about the theoretical behavior of putative NOCT transcripts to determine which transcripts were considered significant (Kojima et al., 2015; Stubblefield et al., 2018). First, it was assumed that the poly(A) tails of target transcripts will be significantly longer in the absence of NOCT. In the tethered function assays discussed in Chapter 4, mutagenesis of the NOCT active site did not eliminate reporter repression, suggesting that other methods of translational control may play a role. If this is true, then using abundance or tail-length changes as criteria to identify targets may exclude transcripts that do not change significantly in these assays but are repressed through association with NOCT. Second, it has been assumed that circadian NOCT expression will enforce rhythmic expression patterns on its target mRNAs. This assumption is not necessarily inaccurate for these studies as they utilized liver tissue, where the amplitude of circadian expression of *Noct* is most significant (Wang et al., 2001). Validation of rhythmic expression at the protein level would strengthen this observation and indicate that

changes in the amplitudes of circadian mRNA abundances translate into functional differences (Kojima et al., 2015; Stubblefield et al., 2018; Wang et al., 2001). While paradigms of circadian regulation may be an important consideration, the assumption that NOCT targets should follow this expression pattern may again exclude potential NOCT targets. As *Noct* expression is also regulated by extracellular stimuli, its targets may be regulated by NOCT in other contexts (Douris et al., 2011; Garbarino-Pico et al., 2007).

1.7 Remaining Questions in NOCT Biology

Previous observations about NOCT functions fall into two categories: *in vitro* biochemical studies that describe NOCT as a deadenylase, cellular and genetic studies that describe the biological phenotypes of NOCT function. The human enzyme had not yet been characterized so we therefore addressed questions of *in vitro* and *in vivo* function of human NOCT by solving the first structure of NOCT from any species and measuring the activity of the human enzyme *in vitro*. We also determined whether NOCT could repression mRNA expression and reduce mRNA steady state levels in cells, which would be consistent with its predicted role as a deadenylase. This work aims to build a better understanding of relationship between NOCT phenotypes and mechanisms of mRNA decay and translational control.

1.8 References

Abaza, I., and Gebauer, F. (2008). Trading translation with RNA-binding proteins. Rna *14*, 404-409.

Abshire, E.T., Chasseur, J., Bohn, J.A., Del Rizzo, P.A., Freddolino, P.L., Goldstrohm, A.C., and Trievel, R.C. (2018). The structure of human Nocturnin reveals a conserved ribonuclease domain that represses target transcript translation and abundance in cells. Nucleic acids research *46*, 6257-6270.

Aitken, C.E., and Lorsch, J.R. (2012). A mechanistic overview of translation initiation in eukaryotes. Nature structural & molecular biology *19*, 568-576.

Akashi, M., and Nishida, E. (2000). Involvement of the MAP kinase cascade in resetting of the mammalian circadian clock. Genes Dev *14*, 645-649.

Alhusaini, N., and Coller, J. (2016). The deadenylase components Not2p, Not3p, and Not5p promote mRNA decapping. Rna *22*, 709-721.

Anderson, C.P., Shen, M., Eisenstein, R.S., and Leibold, E.A. (2012). Mammalian iron metabolism and its control by iron regulatory proteins. Biochimica et biophysica acta *1823*, 1468-1483.

Andreev, D.E., O'Connor, P.B., Loughran, G., Dmitriev, S.E., Baranov, P.V., and Shatsky, I.N. (2017). Insights into the mechanisms of eukaryotic translation gained with ribosome profiling. Nucleic acids research *45*, 513-526.

Ashkenazy, H., Abadi, S., Martz, E., Chay, O., Mayrose, I., Pupko, T., and Ben-Tal, N. (2016). ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules. Nucleic acids research *44*, W344-350.

Ashkenazy, H., Erez, E., Martz, E., Pupko, T., and Ben-Tal, N. (2010). ConSurf 2010: calculating evolutionary conservation in sequence and structure of proteins and nucleic acids. Nucleic acids research *38*, W529-533.

Ashton-Beaucage, D., Udell, C.M., Gendron, P., Sahmi, M., Lefrancois, M., Baril, C., Guenier, A.S., Duchaine, J., Lamarre, D., Lemieux, S.*, et al.* (2014). A functional screen reveals an extensive layer of transcriptional and splicing control underlying RAS/MAPK signaling in Drosophila. PLoS Biol *12*, e1001809.

Ashwal, R., Hemi, R., Tirosh, A., Gordin, R., Yissachar, E., Cohen-Dayag, A., Rosenberg, A., Karasik, A., Bluher, M., and Kanety, H. (2011). Differential expression of novel adiponectin receptor-1 transcripts in skeletal muscle of subjects with normal glucose tolerance and type 2 diabetes. Diabetes *60*, 936-946.

Badis, G., Saveanu, C., Fromont-Racine, M., and Jacquier, A. (2004). Targeted mRNA degradation by deadenylation-independent decapping. Mol Cell *15*, 5-15.

Baggs, J.E., and Green, C.B. (2003). Nocturnin, a deadenylase in Xenopus laevis retina: a mechanism for posttranscriptional control of circadian-related mRNA. Current biology : CB *13*, 189-198.

Baggs, J.E., and Green, C.B. (2006). Functional analysis of nocturnin: a circadian clock-regulated gene identified by differential display. Methods in molecular biology *317*, 243-254.

Balsalobre, A., Damiola, F., and Schibler, U. (1998). A serum shock induces circadian gene expression in mammalian tissue culture cells. Cell *93*, 929-937.

Bartel, D.P. (2009). MicroRNAs: target recognition and regulatory functions. Cell *136*, 215-233.

Bartel, D.P. (2018). Metazoan MicroRNAs. Cell *173*, 20-51.

Behm-Ansmant, I., Rehwinkel, J., Doerks, T., Stark, A., Bork, P., and Izaurralde, E. (2006). mRNA degradation by miRNAs and GW182 requires both CCR4:NOT deadenylase and DCP1:DCP2 decapping complexes. Genes Dev *20*, 1885-1898.

Boeck, R., Tarun, S., Jr., Rieger, M., Deardorff, J.A., Muller-Auer, S., and Sachs, A.B. (1996). The yeast Pan2 protein is required for poly(A)-binding protein-stimulated poly(A)-nuclease activity. The Journal of biological chemistry *271*, 432-438.

Borman, A.M., Michel, Y.M., and Kean, K.M. (2000). Biochemical characterisation of cap-poly(A) synergy in rabbit reticulocyte lysates: the eIF4G-PABP interaction increases the functional affinity of eIF4E for the capped mRNA 5'-end. Nucleic acids research *28*, 4068-4075.

Bourillot, P.Y., Aksoy, I., Schreiber, V., Wianny, F., Schulz, H., Hummel, O., Hubner, N., and Savatier, P. (2009). Novel STAT3 target genes exert distinct roles in the inhibition of mesoderm and endoderm differentiation in cooperation with Nanog. Stem Cells *27*, 1760-1771.

Brown, C.E., Tarun, S.Z., Jr., Boeck, R., and Sachs, A.B. (1996). PAN3 encodes a subunit of the Pab1p-dependent poly(A) nuclease in Saccharomyces cerevisiae. Molecular and cellular biology *16*, 5744-5753.

Buxbaum, A.R., Haimovich, G., and Singer, R.H. (2015). In the right place at the right time: visualizing and understanding mRNA localization. Nature reviews Molecular cell biology *16*, 95-109.

Cargnello, M., and Roux, P.P. (2011). Activation and function of the MAPKs and their substrates, the MAPK-activated protein kinases. Microbiology and molecular biology reviews : MMBR *75*, 50-83.

Carmody, S.R., and Wente, S.R. (2009). mRNA nuclear export at a glance. J Cell Sci *122*, 1933-1937.

Celniker G., N.G., Ashkenazy H., Glaser F., Martz E., Mayrose I., Pupko T., and Ben-Tal N. (2013). ConSurf: Using Evolutionary Data to Raise Testable Hypotheses about Protein Function. Israel Journal of Chemistry.

Chang, H., Lim, J., Ha, M., and Kim, V.N. (2014). TAIL-seq: genome-wide determination of poly(A) tail length and 3' end modifications. Mol Cell *53*, 1044-1052.

Chen, M., and Manley, J.L. (2009). Mechanisms of alternative splicing regulation: insights from molecular and genomics approaches. Nature reviews Molecular cell biology *10*, 741-754.

Chen, N., Walsh, M.A., Liu, Y., Parker, R., and Song, H. (2005). Crystal structures of human DcpS in ligand-free and m7GDP-bound forms suggest a dynamic mechanism for scavenger mRNA decapping. J Mol Biol *347*, 707-718.

Chen, Q., Shou, P., Zheng, C., Jiang, M., Cao, G., Yang, Q., Cao, J., Xie, N., Velletri, T., Zhang, X.*, et al.* (2016). Fate decision of mesenchymal stem cells: adipocytes or osteoblasts? Cell Death Differ *23*, 1128-1139.

Choi, Y.J., Song, I., Jin, Y., Jin, H.S., Ji, H.M., Jeong, S.Y., Won, Y.Y., and Chung, Y.S. (2017). Transcriptional profiling of human femoral mesenchymal stem cells in osteoporosis and its association with adipogenesis. Gene *632*, 7-15.

Craig, A.W., Haghighat, A., Yu, A.T., and Sonenberg, N. (1998). Interaction of polyadenylate-binding protein with the eIF4G homologue PAIP enhances translation. Nature *392*, 520-523.

Danckwardt, S., Gantzert, A.S., Macher-Goeppinger, S., Probst, H.C., Gentzel, M., Wilm, M., Grone, H.J., Schirmacher, P., Hentze, M.W., and Kulozik, A.E. (2011). p38 MAPK controls prothrombin expression by regulated RNA 3' end processing. Mol Cell *41*, 298-310.

Das, B., Butler, J.S., and Sherman, F. (2003). Degradation of normal mRNA in the nucleus of Saccharomyces cerevisiae. Molecular and cellular biology *23*, 5502-5515.

Dever, T.E., and Green, R. (2012). The elongation, termination, and recycling phases of translation in eukaryotes. Cold Spring Harb Perspect Biol *4*, a013706.

Doidge, R., Mittal, S., Aslam, A., and Winkler, G.S. (2012). Deadenylation of cytoplasmic mRNA by the mammalian Ccr4-Not complex. Biochem Soc Trans *40*, 896-901.

Dominguez, D., Freese, P., Alexis, M.S., Su, A., Hochman, M., Palden, T., Bazile, C., Lambert, N.J., Van Nostrand, E.L., Pratt, G.A.*, et al.* (2018). Sequence, Structure, and Context Preferences of Human RNA Binding Proteins. Mol Cell *70*, 854-867 e859.

Douris, N., Kojima, S., Pan, X., Lerch-Gaggl, A.F., Duong, S.Q., Hussain, M.M., and Green, C.B. (2011). Nocturnin regulates circadian trafficking of dietary lipid in intestinal enterocytes. Current biology : CB *21*, 1347-1355.

Eckel-Mahan, K., and Sassone-Corsi, P. (2013). Metabolism and the circadian clock converge. Physiol Rev *93*, 107-135.

Elkon, R., Ugalde, A.P., and Agami, R. (2013). Alternative cleavage and polyadenylation: extent, regulation and function. Nat Rev Genet *14*, 496-506.

Ezelle, H.J., and Hassel, B.A. (2012). Pathologic effects of RNase-L dysregulation in immunity and proliferative control. Front Biosci (Schol Ed) *4*, 767-786.

Fenger-Gron, M., Fillman, C., Norrild, B., and Lykke-Andersen, J. (2005). Multiple processing body factors and the ARE binding protein TTP activate mRNA decapping. Mol Cell *20*, 905-915.

Fischer, N., and Weis, K. (2002). The DEAD box protein Dhh1 stimulates the decapping enzyme Dcp1. The EMBO journal *21*, 2788-2797.

Fukao, A., Mishima, Y., Takizawa, N., Oka, S., Imataka, H., Pelletier, J., Sonenberg, N., Thoma, C., and Fujiwara, T. (2014). MicroRNAs trigger dissociation of eIF4AI and eIF4AII from target mRNAs in humans. Mol Cell *56*, 79-89.

Gallie, D.R. (1991). The cap and poly(A) tail function synergistically to regulate mRNA translational efficiency. Genes Dev *5*, 2108-2116.

Garbarino-Pico, E., Niu, S., Rollag, M.D., Strayer, C.A., Besharse, J.C., and Green, C.B. (2007). Immediate early response of the circadian polyA ribonuclease nocturnin to two extracellular stimuli. Rna *13*, 745-755.

Garneau, N.L., Wilusz, J., and Wilusz, C.J. (2007). The highways and byways of mRNA decay. Nature reviews Molecular cell biology *8*, 113-126.

Gehring, N.H., Wahle, E., and Fischer, U. (2017). Deciphering the mRNP Code: RNA-Bound Determinants of Post-Transcriptional Gene Regulation. Trends Biochem Sci *42*, 369-382.

Gilbert, M.R., Douris, N., Tongjai, S., and Green, C.B. (2011). Nocturnin expression is induced by fasting in the white adipose tissue of restricted fed mice. PloS one *6*, e17051.

Gingras, A.C., Kennedy, S.G., O'Leary, M.A., Sonenberg, N., and Hay, N. (1998). 4E-BP1, a repressor of mRNA translation, is phosphorylated and inactivated by the Akt(PKB) signaling pathway. Genes Dev *12*, 502-513.

Glaser, F., Pupko, T., Paz, I., Bell, R.E., Bechor-Shental, D., Martz, E., and Ben-Tal, N. (2003). ConSurf: identification of functional regions in proteins by surface-mapping of phylogenetic information. Bioinformatics *19*, 163-164.

Goldstrohm, A.C., and Wickens, M. (2008). Multifunctional deadenylase complexes diversify mRNA control. Nature reviews Molecular cell biology *9*, 337-344.

Green, C.B., and Besharse, J.C. (1996a). Identification of a novel vertebrate circadian clock-regulated gene encoding the protein nocturnin. Proceedings of the National Academy of Sciences of the United States of America *93*, 14884-14888.

Green, C.B., and Besharse, J.C. (1996b). Use of a high stringency differential display screen for identification of retinal mRNAs that are regulated by a circadian clock. Brain Res Mol Brain Res *37*, 157-165.

Green, C.B., and Besharse, J.C. (1997). Identification of vertebrate circadian clock-regulated genes by differential display. Methods in molecular biology *85*, 219-230.

Green, C.B., Douris, N., Kojima, S., Strayer, C.A., Fogerty, J., Lourim, D., Keller, S.R., and Besharse, J.C. (2007). Loss of Nocturnin, a circadian deadenylase, confers resistance to hepatic steatosis and diet-induced obesity. Proceedings of the National Academy of Sciences of the United States of America *104*, 9888-9893.

Groisman, I., Huang, Y.S., Mendez, R., Cao, Q., Theurkauf, W., and Richter, J.D. (2000). CPEB, maskin, and cyclin B1 mRNA at the mitotic apparatus: implications for local translational control of cell division. Cell *103*, 435-447.

Grudzien-Nogalska, E., and Kiledjian, M. (2017). New insights into decapping enzymes and selective mRNA decay. Wiley interdisciplinary reviews RNA *8*.

Hamza, M.S., Pott, S., Vega, V.B., Thomsen, J.S., Kandhadayar, G.S., Ng, P.W., Chiu, K.P., Pettersson, S., Wei, C.L., Ruan, Y.*, et al.* (2009). De-novo identification of PPARgamma/RXR binding sites and direct targets during adipogenesis. PloS one *4*, e4907.

Hardie, D.G. (2012). Organismal carbohydrate and lipid homeostasis. Cold Spring Harb Perspect Biol *4*.

Harvey, R.F., Smith, T.S., Mulroney, T., Queiroz, R.M.L., Pizzinga, M., Dezi, V., Villenueva, E., Ramakrishna, M., Lilley, K.S., and Willis, A.E. (2018). Trans-acting translational regulatory RNA binding proteins. Wiley interdisciplinary reviews RNA *9*, e1465.

Hee, S.W., Tsai, S.H., Chang, Y.C., Chang, C.J., Yu, I.S., Lee, P.C., Lee, W.J., Yun-Chia Chang, E., and Chuang, L.M. (2012). The role of nocturnin in early adipogenesis and modulation of systemic insulin resistance in human. Obesity *20*, 1558-1565.

Herzel, L., Ottoz, D.S.M., Alpert, T., and Neugebauer, K.M. (2017). Splicing and transcription touch base: co-transcriptional spliceosome assembly and function. Nature reviews Molecular cell biology *18*, 637-650.

Hinnebusch, A.G. (2014). The scanning mechanism of eukaryotic translation initiation. Annu Rev Biochem *83*, 779-812.

Hinnebusch, A.G. (2017). Structural Insights into the Mechanism of Scanning and Start Codon Recognition in Eukaryotic Translation Initiation. Trends Biochem Sci *42*, 589-611.

Hirose, Y., and Manley, J.L. (1998). RNA polymerase II is an essential mRNA polyadenylation factor. Nature *395*, 93-96.

Hoogerwerf, W.A., Hellmich, H.L., Cornelissen, G., Halberg, F., Shahinian, V.B., Bostwick, J., Savidge, T.C., and Cassone, V.M. (2007). Clock gene expression in the

murine gastrointestinal tract: endogenous rhythmicity and effects of a feeding regimen. Gastroenterology *133*, 1250-1260.

Hsin, J.P., and Manley, J.L. (2012). The RNA polymerase II CTD coordinates transcription and RNA processing. Genes Dev *26*, 2119-2137.

Huang da, W., Sherman, B.T., and Lempicki, R.A. (2009a). Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Nucleic acids research *37*, 1-13.

Huang da, W., Sherman, B.T., and Lempicki, R.A. (2009b). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc *4*, 44-57.

Ivshina, M., Lasko, P., and Richter, J.D. (2014). Cytoplasmic polyadenylation element binding proteins in development, health, and disease. Annu Rev Cell Dev Biol *30*, 393-415.

Iwakawa, H.O., and Tomari, Y. (2015). The Functions of MicroRNAs: mRNA Decay and Translational Repression. Trends Cell Biol *25*, 651-665.

Jackson, R.J., Hellen, C.U., and Pestova, T.V. (2010). The mechanism of eukaryotic translation initiation and principles of its regulation. Nature reviews Molecular cell biology *11*, 113-127.

Jalkanen, A.L., Coleman, S.J., and Wilusz, J. (2014). Determinants and implications of mRNA poly(A) tail size--does this protein make my tail look big? Semin Cell Dev Biol *34*, 24-32.

James, A.W. (2013). Review of Signaling Pathways Governing MSC Osteogenic and Adipogenic Differentiation. Scientifica (Cairo) *2013*, 684736.

Jonas, S., and Izaurralde, E. (2013). The role of disordered protein regions in the assembly of decapping complexes and RNP granules. Genes Dev *27*, 2628-2641.

Kataoka, N., Yong, J., Kim, V.N., Velazquez, F., Perkinson, R.A., Wang, F., and Dreyfuss, G. (2000). Pre-mRNA splicing imprints mRNA in the nucleus with a novel RNA-binding protein that persists in the cytoplasm. Mol Cell *6*, 673-682.

Kavaliers, M., Hirst, M., and Teskey, G.C. (1985). Nocturnal feeding in the mouse--opiate and pineal influences. Life Sci *36*, 973-980.

Kawahara, H., Imai, T., Imataka, H., Tsujimoto, M., Matsumoto, K., and Okano, H. (2008). Neural RNA-binding protein Musashi1 inhibits translation initiation by competing with eIF4G for PABP. The Journal of cell biology *181*, 639-653.

Kawai, M., Delany, A.M., Green, C.B., Adamo, M.L., and Rosen, C.J. (2010a). Nocturnin suppresses igf1 expression in bone by targeting the 3' untranslated region of igf1 mRNA. Endocrinology *151*, 4861-4870.

Kawai, M., Green, C.B., Horowitz, M., Ackert-Bicknell, C., Lecka-Czernik, B., and Rosen, C.J. (2010b). Nocturnin: a circadian target of Pparg-induced adipogenesis. Annals of the New York Academy of Sciences *1192*, 131-138.

Kawai, M., Green, C.B., Lecka-Czernik, B., Douris, N., Gilbert, M.R., Kojima, S., Ackert-Bicknell, C., Garg, N., Horowitz, M.C., Adamo, M.L.*, et al.* (2010c). A circadian-regulated gene, Nocturnin, promotes adipogenesis by stimulating PPAR-gamma nuclear translocation. Proceedings of the National Academy of Sciences of the United States of America *107*, 10508-10513.

Kawai, M., and Rosen, C.J. (2010). PPARgamma: a circadian transcription factor in adipogenesis and osteogenesis. Nature reviews Endocrinology *6*, 629-636.

Kim, J.H., and Richter, J.D. (2008). Measuring CPEB-mediated cytoplasmic polyadenylation-deadenylation in Xenopus laevis oocytes and egg extracts. Methods in enzymology *448*, 119-138.

Kohler, A., and Hurt, E. (2007). Exporting RNA from the nucleus to the cytoplasm. Nature reviews Molecular cell biology *8*, 761-773.

Kojima, S., Gatfield, D., Esau, C.C., and Green, C.B. (2010). MicroRNA-122 modulates the rhythmic expression profile of the circadian deadenylase Nocturnin in mouse liver. PloS one *5*, e11264.

Kojima, S., Gendreau, K.L., Sher-Chen, E.L., Gao, P., and Green, C.B. (2015). Changes in poly(A) tail length dynamics from the loss of the circadian deadenylase Nocturnin. Scientific reports *5*, 17059.

Kornmann, B., Schaad, O., Bujard, H., Takahashi, J.S., and Schibler, U. (2007). System-driven and oscillator-dependent circadian transcription in mice with a conditionally active liver clock. PLoS Biol *5*, e34.

Kozak, M. (1991a). Effects of long 5' leader sequences on initiation by eukaryotic ribosomes in vitro. Gene Expr *1*, 117-125.

Kozak, M. (1991b). A short leader sequence impairs the fidelity of initiation by eukaryotic ribosomes. Gene Expr *1*, 111-115.

Kubota, K., Nakahara, K., Ohtsuka, T., Yoshida, S., Kawaguchi, J., Fujita, Y., Ozeki, Y., Hara, A., Yoshimura, C., Furukawa, H.*, et al.* (2004). Identification of 2'-phosphodiesterase, which plays a role in the 2-5A system regulated by interferon. The Journal of biological chemistry *279*, 37832-37841.

Kuhn, U., Gundel, M., Knoth, A., Kerwitz, Y., Rudel, S., and Wahle, E. (2009). Poly(A) tail length is controlled by the nuclear poly(A)-binding protein regulating the interaction between poly(A) polymerase and the cleavage and polyadenylation specificity factor. The Journal of biological chemistry *284*, 22803-22814.

Kyburz, A., Friedlein, A., Langen, H., and Keller, W. (2006). Direct interactions between subunits of CPSF and the U2 snRNP contribute to the coupling of pre-mRNA 3' end processing and splicing. Mol Cell *23*, 195-205.

LaCava, J., Houseley, J., Saveanu, C., Petfalski, E., Thompson, E., Jacquier, A., and Tollervey, D. (2005). RNA degradation by the exosome is promoted by a nuclear polyadenylation complex. Cell *121*, 713-724.

Landau, M., Mayrose, I., Rosenberg, Y., Glaser, F., Martz, E., Pupko, T., and Ben-Tal, N. (2005). ConSurf 2005: the projection of evolutionary conservation scores of residues on protein structures. Nucleic acids research *33*, W299-302.

Lee, K.M., and Tarn, W.Y. (2013). Coupling pre-mRNA processing to transcription on the RNA factory assembly line. RNA Biol *10*, 380-390.

Li, R., Yue, J., Zhang, Y., Zhou, L., Hao, W., Yuan, J., Qiang, B., Ding, J.M., Peng, X., and Cao, J.M. (2008). CLOCK/BMAL1 regulates human nocturnin transcription through binding to the E-box of nocturnin promoter. Mol Cell Biochem *317*, 169-177.

Li, X.L., Ezelle, H.J., Hsi, T.Y., and Hassel, B.A. (2011). A central role for RNA in the induction and biological activities of type 1 interferons. Wiley interdisciplinary reviews RNA *2*, 58-78.

Ling, S.H., Qamra, R., and Song, H. (2011). Structural and functional insights into eukaryotic mRNA decapping. Wiley interdisciplinary reviews RNA *2*, 193-208.

Liu, H., and Kiledjian, M. (2005). Scavenger decapping activity facilitates 5' to 3' mRNA decay. Molecular and cellular biology *25*, 9764-9772.

Liu, Q., Greimann, J.C., and Lima, C.D. (2006). Reconstitution, activities, and structure of the eukaryotic RNA exosome. Cell *127*, 1223-1237.

Lykke-Andersen, J., and Wagner, E. (2005). Recruitment and activation of mRNA decay enzymes by two ARE-mediated decay activation domains in the proteins TTP and BRF-1. Genes Dev *19*, 351-361.

Mandon, E.C., Trueman, S.F., and Gilmore, R. (2013). Protein translocation across the rough endoplasmic reticulum. Cold Spring Harb Perspect Biol *5*.

Marintchev, A. (2013). Roles of helicases in translation initiation: a mechanistic view. Biochimica et biophysica acta *1829*, 799-809.

Martineau, Y., Derry, M.C., Wang, X., Yanagiya, A., Berlanga, J.J., Shyu, A.B., Imataka, H., Gehring, K., and Sonenberg, N. (2008). Poly(A)-binding protein-interacting protein 1 binds to eukaryotic translation initiation factor 3 to stimulate translation. Molecular and cellular biology *28*, 6658-6667.

Maryati, M., Airhihen, B., and Winkler, G.S. (2015). The enzyme activities of Caf1 and Ccr4 are both required for deadenylation by the human Ccr4-Not nuclease module. Biochem J *469*, 169-176.

Mayeuf-Louchart, A., Zecchin, M., Staels, B., and Duez, H. (2017). Circadian control of metabolism and pathological consequences of clock perturbations. Biochimie *143*, 42-50.

Michel, Y.M., Poncet, D., Piron, M., Kean, K.M., and Borman, A.M. (2000). Cap-Poly(A) synergy in mammalian cell-free extracts. Investigation of the requirements for poly(A)-mediated stimulation of translation initiation. The Journal of biological chemistry *275*, 32268-32276.

Moore, M.J., and Proudfoot, N.J. (2009). Pre-mRNA processing reaches back to transcription and ahead to translation. Cell *136*, 688-700.

Muckenthaler, M., Gray, N.K., and Hentze, M.W. (1998). IRP-1 binding to ferritin mRNA prevents the recruitment of the small ribosomal subunit by the cap-binding complex eIF4F. Mol Cell *2*, 383-388.

Nagarajan, V.K., Jones, C.I., Newbury, S.F., and Green, P.J. (2013). XRN 5'-->3' exoribonucleases: structure, mechanisms and functions. Biochimica et biophysica acta *1829*, 590-603.

Nandagopal, N., and Roux, P.P. (2015). Regulation of global and specific mRNA translation by the mTOR signaling pathway. Translation (Austin) *3*, e983402.

Neve, J., Patel, R., Wang, Z., Louey, A., and Furger, A.M. (2017). Cleavage and polyadenylation: Ending the message expands gene regulation. RNA Biol *14*, 865-890.

Oishi, K., Miyazaki, K., Kadota, K., Kikuno, R., Nagase, T., Atsumi, G., Ohkura, N., Azama, T., Mesaki, M., Yukimasa, S.*, et al.* (2003). Genome-wide expression analysis of mouse liver reveals CLOCK-regulated circadian output genes. The Journal of biological chemistry *278*, 41519-41527.

Paik, J.H., Kollipara, R., Chu, G., Ji, H., Xiao, Y., Ding, Z., Miao, L., Tothova, Z., Horner, J.W., Carrasco, D.R.*, et al.* (2007). FoxOs are lineage-restricted redundant tumor suppressors and regulate endothelial cell homeostasis. Cell *128*, 309-323.

Pan, X., and Hussain, M.M. (2007). Diurnal regulation of microsomal triglyceride transfer protein and plasma lipid levels. The Journal of biological chemistry *282*, 24707-24719.

Pan, X., and Hussain, M.M. (2009). Clock is important for food and circadian regulation of macronutrient absorption in mice. J Lipid Res *50*, 1800-1813.

Parsyan, A., Svitkin, Y., Shahbazian, D., Gkogkas, C., Lasko, P., Merrick, W.C., and Sonenberg, N. (2011). mRNA helicases: the tacticians of translational control. Nature reviews Molecular cell biology *12*, 235-245.

Pearce, S.F., Rorbach, J., Van Haute, L., D'Souza, A.R., Rebelo-Guiomar, P., Powell, C.A., Brierley, I., Firth, A.E., and Minczuk, M. (2017). Maturation of selected human mitochondrial tRNAs requires deadenylation. Elife *6*.

Perrin, L., Loizides-Mangold, U., Chanon, S., Gobet, C., Hulo, N., Isenegger, L., Weger, B.D., Migliavacca, E., Charpagne, A., Betts, J.A*., et al.* (2018). Transcriptomic analyses reveal rhythmic and CLOCK-driven pathways in human skeletal muscle. Elife *7*.

Piccirillo, C., Khanna, R., and Kiledjian, M. (2003). Functional characterization of the mammalian mRNA decapping enzyme hDcp2. Rna *9*, 1138-1147.

Poulsen, J.B., Andersen, K.R., Kjaer, K.H., Durand, F., Faou, P., Vestergaard, A.L., Talbo, G.H., Hoogenraad, N., Brodersen, D.E., Justesen, J*., et al.* (2011). Human 2'-phosphodiesterase localizes to the mitochondrial matrix with a putative function in mitochondrial RNA turnover. Nucleic acids research *39*, 3754-3770.

Qi, Y., Jiang, C., Cheng, J., Krausz, K.W., Li, T., Ferrell, J.M., Gonzalez, F.J., and Chiang, J.Y. (2015). Bile acid signaling in lipid metabolism: metabolomic and lipidomic analysis of lipid and bile acid markers linked to anti-obesity and anti-diabetes in mice. Biochimica et biophysica acta *1851*, 19-29.

Ramanathan, A., Robb, G.B., and Chan, S.H. (2016). mRNA capping: biological functions and applications. Nucleic acids research *44*, 7511-7526.

Reis, F.P., Pobre, V., Silva, I.J., Malecki, M., and Arraiano, C.M. (2013). The RNase II/RNB family of exoribonucleases: putting the 'Dis' in disease. Wiley interdisciplinary reviews RNA *4*, 607-615.

Ren, Y.G., Martinez, J., Kirsebom, L.A., and Virtanen, A. (2002). Inhibition of Klenow DNA polymerase and poly(A)-specific ribonuclease by aminoglycosides. Rna *8*, 1393-1400.

Rorbach, J., Nicholls, T.J., and Minczuk, M. (2011). PDE12 removes mitochondrial RNA poly(A) tails and controls translation in human mitochondria. Nucleic acids research *39*, 7750-7763.

Roy, G., De Crescenzo, G., Khaleghpour, K., Kahvejian, A., O'Connor-McCourt, M., and Sonenberg, N. (2002). Paip1 interacts with poly(A) binding protein through two independent binding motifs. Molecular and cellular biology *22*, 3769-3782.

Ryder, S.P. (2016). Protein-mRNA interactome capture: cartography of the mRNP landscape. F1000Res *5*, 2627.

She, M., Decker, C.J., Svergun, D.I., Round, A., Chen, N., Muhlrad, D., Parker, R., and Song, H. (2008). Structural basis of dcp2 recognition and activation by dcp1. Mol Cell *29*, 337-349.

Shi, Y. (2017a). Mechanistic insights into precursor messenger RNA splicing by the spliceosome. Nature reviews Molecular cell biology *18*, 655-670.

Shi, Y. (2017b). The Spliceosome: A Protein-Directed Metalloribozyme. J Mol Biol *429*, 2640-2653.

Silverman, R.H. (2007). Viral encounters with 2',5'-oligoadenylate synthetase and RNase L during the interferon antiviral response. J Virol *81*, 12720-12729.

Sinturel, F., Brechemier-Baey, D., Kiledjian, M., Condon, C., and Benard, L. (2012). Activation of 5'-3' exoribonuclease Xrn1 by cofactor Dcs1 is essential for mitochondrial function in yeast. Proceedings of the National Academy of Sciences of the United States of America *109*, 8264-8269.

Siwaszek, A., Ukleja, M., and Dziembowski, A. (2014). Proteins involved in the degradation of cytoplasmic mRNA in the major eukaryotic model systems. RNA Biol *11*, 1122-1136.

Soheilypour, M., and Mofrad, M.R.K. (2018). Quality control of mRNAs at the entry of the nuclear pore: Cooperation in a complex molecular system. Nucleus *9*, 202-211.

Stebbins-Boaz, B., Cao, Q., de Moor, C.H., Mendez, R., and Richter, J.D. (1999). Maskin is a CPEB-associated factor that transiently interacts with eIF-4E. Mol Cell *4*, 1017-1027.

Stewart, M. (2019). Polyadenylation and nuclear export of mRNAs. The Journal of biological chemistry *294*, 2977-2987.

Stubblefield, J.J., Gao, P., Kilaru, G., Mukadam, B., Terrien, J., and Green, C.B. (2018). Temporal Control of Metabolic Amplitude by Nocturnin. Cell reports *22*, 1225-1235.

Temme, C., Zhang, L., Kremmer, E., Ihling, C., Chartier, A., Sinz, A., Simonelig, M., and Wahle, E. (2010). Subunits of the Drosophila CCR4-NOT complex and their roles in mRNA deadenylation. Rna *16*, 1356-1370.

Tharun, S., and Parker, R. (2001). Targeting an mRNA for decapping: displacement of translation factors and association of the Lsm1p-7p complex on deadenylated yeast mRNAs. Mol Cell *8*, 1075-1083.

Thomson, A.M., Rogers, J.T., and Leedman, P.J. (1999). Iron-regulatory proteins, iron-responsive elements and ferritin mRNA translation. Int J Biochem Cell Biol *31*, 1139-1152.

Tian, B., and Manley, J.L. (2013). Alternative cleavage and polyadenylation: the long and short of it. Trends Biochem Sci *38*, 312-320.

Tritschler, F., Huntzinger, E., and Izaurralde, E. (2010). Role of GW182 proteins and PABPC1 in the miRNA pathway: a sense of deja vu. Nature reviews Molecular cell biology *11*, 379-384.

Trotman, J.B., and Schoenberg, D.R. (2019). A recap of RNA recapping. Wiley interdisciplinary reviews RNA *10*, e1504.

Tucker, M., Staples, R.R., Valencia-Sanchez, M.A., Muhlrad, D., and Parker, R. (2002). Ccr4p is the catalytic subunit of a Ccr4p/Pop2p/Notp mRNA deadenylase complex in Saccharomyces cerevisiae. The EMBO journal *21*, 1427-1436.

Tucker, M., Valencia-Sanchez, M.A., Staples, R.R., Chen, J., Denis, C.L., and Parker, R. (2001). The transcription factor associated Ccr4 and Caf1 proteins are components of the major cytoplasmic mRNA deadenylase in Saccharomyces cerevisiae. Cell *104*, 377-386.

Wang, H., Morita, M., Yang, X., Suzuki, T., Yang, W., Wang, J., Ito, K., Wang, Q., Zhao, C., Bartlam, M.*, et al.* (2010). Crystal structure of the human CNOT6L nuclease domain reveals strict poly(A) substrate specificity. The EMBO journal *29*, 2566-2576.

Wang, Y., Osterbur, D.L., Megaw, P.L., Tosini, G., Fukuhara, C., Green, C.B., and Besharse, J.C. (2001). Rhythmic expression of Nocturnin mRNA in multiple tissues of the mouse. BMC developmental biology *1*, 9.

Wang, Z., Jiao, X., Carr-Schmid, A., and Kiledjian, M. (2002). The hDcp2 protein is a mammalian mRNA decapping enzyme. Proceedings of the National Academy of Sciences of the United States of America *99*, 12663-12668.

Wasmuth, E.V., and Lima, C.D. (2017). The Rrp6 C-terminal domain binds RNA and activates the nuclear RNA exosome. Nucleic acids research *45*, 846-860.

Webster, M.W., Chen, Y.H., Stowell, J.A.W., Alhusaini, N., Sweet, T., Graveley, B.R., Coller, J., and Passmore, L.A. (2018). mRNA Deadenylation Is Coupled to Translation Rates by the Differential Activities of Ccr4-Not Nucleases. Mol Cell *70*, 1089-1100 e1088.

Weidmann, C.A., Raynard, N.A., Blewett, N.H., Van Etten, J., and Goldstrohm, A.C. (2014). The RNA binding domain of Pumilio antagonizes poly-adenosine binding protein and accelerates deadenylation. Rna *20*, 1298-1319.

Wood, E.R., Bledsoe, R., Chai, J., Daka, P., Deng, H., Ding, Y., Harris-Gurley, S., Kryn, L.H., Nartey, E., Nichols, J.*, et al.* (2015). The Role of Phosphodiesterase 12 (PDE12) as a Negative Regulator of the Innate Immune Response and the Discovery of Antiviral Inhibitors. The Journal of biological chemistry *290*, 19681-19696.

Yamashita, A., Chang, T.C., Yamashita, Y., Zhu, W., Zhong, Z., Chen, C.Y., and Shyu, A.B. (2005). Concerted action of poly(A) nucleases and decapping enzyme in mammalian mRNA turnover. Nature structural & molecular biology *12*, 1054-1063.

Yi, H., Park, J., Ha, M., Lim, J., Chang, H., and Kim, V.N. (2018). PABP Cooperates with the CCR4-NOT Complex to Promote mRNA Deadenylation and Block Precocious Decay. Mol Cell *70*, 1081-1088 e1085.

Yoon, J.H., Abdelmohsen, K., and Gorospe, M. (2013). Posttranscriptional gene regulation by long noncoding RNA. J Mol Biol *425*, 3723-3730.

You, L., Pan, L., Chen, L., Gu, W., and Chen, J. (2016). MiR-27a is Essential for the Shift from Osteogenic Differentiation to Adipogenic Differentiation of Mesenchymal Stem Cells in Postmenopausal Osteoporosis. Cell Physiol Biochem *39*, 253-265.

Yu, J.H., Yang, W.H., Gulick, T., Bloch, K.D., and Bloch, D.B. (2005). Ge-1 is a central component of the mammalian cytoplasmic mRNA processing body. Rna *11*, 1795-1802.

Yuan, Z., Li, Q., Luo, S., Liu, Z., Luo, D., Zhang, B., Zhang, D., Rao, P., and Xiao, J. (2016). PPARgamma and Wnt Signaling in Adipogenic and Osteogenic Differentiation of Mesenchymal Stem Cells. Curr Stem Cell Res Ther *11*, 216-225.

Zekri, L., Kuzuoglu-Ozturk, D., and Izaurralde, E. (2013). GW182 proteins cause PABP dissociation from silenced miRNA targets in the absence of deadenylation. The EMBO journal *32*, 1052-1065.

Zhang, Q., Yan, D., Guo, E., Ding, B., Yang, W., Liu, R., Yamamoto, T., and Bartlam, M. (2016). Structural basis for inhibition of the deadenylase activity of human CNOT6L. FEBS Lett *590*, 1270-1279.

Zheng, D., Wang, R., Ding, Q., Wang, T., Xie, B., Wei, L., Zhong, Z., and Tian, B. (2018). Cellular stress alters 3'UTR landscape through alternative polyadenylation and isoform-specific degradation. Nature communications *9*, 2268.

Zieker, D., Jenne, I., Koenigsrainer, I., Zdichavsky, M., Nieselt, K., Buck, K., Zieker, J., Beckert, S., Glatzle, J., Spanagel, R.*, et al.* (2010). Circadian expression of clock- and tumor suppressor genes in human oral mucosa. Cell Physiol Biochem *26*, 155-166.

Zinder, J.C., and Lima, C.D. (2017). Targeting RNA for processing or destruction by the eukaryotic RNA exosome and its cofactors. Genes Dev *31*, 88-100.

CHAPTER 2

Purification and Structural Characterization of NOCT[1]

2.1 Abstract

NOCT is a human Endonuclease, Exonuclease, and Phosphatase (EEP) superfamily enzyme that is most closely related to the human EEP deadenylases CNOT6L and PDE12. As structures of the NOCT enzyme had yet to be reported, we undertook structural studies of the human homolog. To ascertain the structure of the NOCT catalytic domain, we first optimized the purification of recombinant NOCT. While purified $NOCT_{120-431}$ was sufficiently pure for crystallization, our observations of nucleic acid co-purification prompted further optimization of purification of a longer $NOCT_{64-431}$ construct for use in biochemistry. We then solved a pair of crystal structures of the NOCT catalytic domain and observed that NOCT forms a highly conserved α/β hydrolase fold that defines the EEP catalytic domain. Conserved active site residues form a core active site that is nearly identical to those of CNOT6L and PDE12. However, unique features of

---

[1] This work was contributed to by Dr. Paul Del Rizzo, who designed, cloned, and did initial expression testing of the human NOCT constructs, and by Jennifer Chasseur, who performed the crystallization of NOCT with 3´ , 5; ADP. This work was published in the following publication: Abshire, E.T., Chasseur, J., Bohn, J.A., Del Rizzo, P.A., Freddolino, P.L., Goldstrohm, A.C., and Trievel, R.C. (2018). The structure of human Nocturnin reveals a conserved ribonuclease domain that represses target transcript translation and abundance in cells. Nucleic acids research *46*, 6257-6270.

the NOCT catalytic domain suggest that NOCT may recognize and bind its substrates through unique interactions, differentiating it from other members of the EEP family.

2.2 Introduction

NOCT was initially described as an EEP deadenylase based on the homology of the *Xenopus* enzyme (xNOCT) with yeast Ccr4, the founding member of the EEP deadenylase subfamily (Baggs and Green, 2003). This study found that the C-terminal domains of Ccr4 and xNOCT were conserved with 31% sequence identity and 41% sequence similarity (Green and Besharse, 1996). These regions of xNOCT and Ccr4 correspond to the EEP catalytic domain, and NOCT was therefore classified as a member of the EEP deadenylase subfamily of Mg^{2+}-ribonucleases. This group of enzymes include CNOT6L, the human homolog of Ccr4 and PDE12, a mitochondrial tRNA deadenylase, both of which have been structurally characterized (Pearce et al., 2017; Wang et al., 2010; Wood et al., 2015). The NOCT EEP domain has 22% and 24% sequence identity to CNOT6L and PDE12 EEP domains, respectively.

Structural characterization of NOCT is an important step forward in the field, as no structures of the enzyme had been determined prior to publication of the data presented in this chapter (Abshire et al., 2018). Structures of the NOCT catalytic domain will confirm that NOCT forms the predicted EEP hydrolase fold and that its active site is formed by the conserved residues predicted to be involved in catalysis. More importantly, NOCT structures will also identify unique features that differentiate NOCT from CNOT6L and PDE12, which may predict differences in function. Identifying variations in predicted substrate binding residues may explain why NOCT recognizes a specific subset of RNAs involved in lipid metabolism.

Here, we describe the steps taken to purify recombinant NOCT enzyme for both structural and biochemical characterization. In addition to the high degree of purity required for protein crystallization, removal of co-purifying nucleic acid was a major impetus for optimizing purification of NOCT used inn biochemical analysis. Residual nucleic acids are known to bring along nucleases that can confound the results of ribonuclease activity assays (Hrit et al., 2014). We then describe a pair of crystal structures of the NOCT EEP catalytic domain and discuss unique features of NOCT relative to CNOT6L and PDE12. NOCT is also compared to the EEP abasic site endonuclease APE1, which has a well-characterized catalytic mechanism and can be used to predict the potential function of NOCT active site residues. The structure of the NOCT catalytic domain can therefore be used as a powerful predictive tool in designing studies to test the importance of active site residues in NOCT activity.

2.3 Results

2.3.I Purification of NOCT$_{120-431}$ for structural studies.

In our initial purifications, we attempted to purify the full-length enzyme NOCT$_{1-431}$ but encountered difficulty with solubility and degradation of recombinant NOCT expressed in *E.coli*. We examined the predicted secondary structure elements of NOCT using the GlobPlot secondary structure prediction tool and comparison to structures of the closely related human enzyme CNOT6L. These analyses were then utilized in construct design to better predict EEP catalytic domain boundaries (Linding et al., 2003; Wang et al., 2010). According to these predictions, the N-terminus of NOCT is highly disordered and was therefore truncated in the constructs expressing recombinant NOCT. Addition of an N-terminal Sumo tag improved solubility and stability of NOCT after cell lysis.

Figure 2.1: Initial purification of NOCT$_{120-431}$ yielded unstable protein containing nucleic acid contaminants. (A) Schematic of NOCT purification protocol prior to optimization. **(B)** Size exclusion chromatogram of NOCT$_{120-431}$ purification on a Superdex 75 size exclusion column. The peak eluting at 45 mL represents protein aggregates eluting in the column void volume. Monomeric NOCT$_{120-431}$ elutes at approximately 70 mL. The fractions were pooled for analysis in three groups by size, corresponding to the fractions corresponding to the peak with an absorbance maximum at 55 mL, 63 mL, and 74 mL (Pool A, B, and C, respectively). **(C)** Fractions from Pool A, B, and C were visualized on an agarose gel with ethidium bromide.

The Strep(II)-Sumo-NOCT$_{120-431}$ fusion protein was selected for further optimization of purity and solubility. NOCT initially was expressed in Rosetta 2 DE3 *E.coli* and purified in a two-step scheme starting with a Strep-Tactin affinity purification followed by removal of the Strep(II)-Sumo tag by Sumo protease and finally, size exclusion chromatography. Purified NOCT eluted as multiple species, including a large peak corresponding to aggregated protein, indicating that the purified protein was fairly unstable (Figure 2.1). Additionally, the purified protein co-purified with nucleic acid (Figure 2.1). As the identity of this nucleic acid was unknown and was assumed to be heterogeneous, we sought to remove it from our protein purifications. Addition of an anion
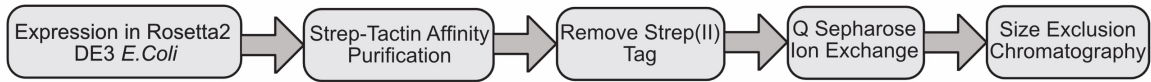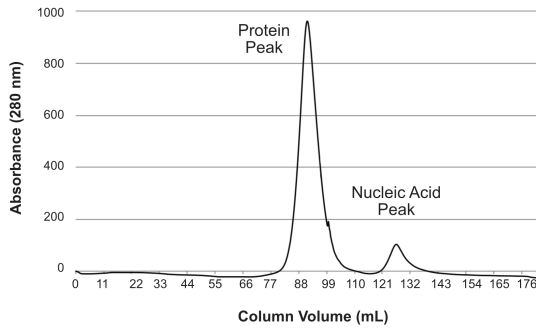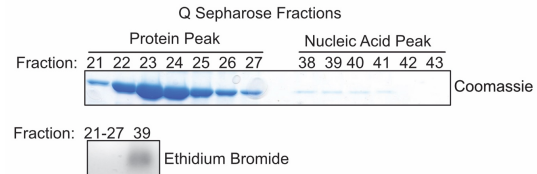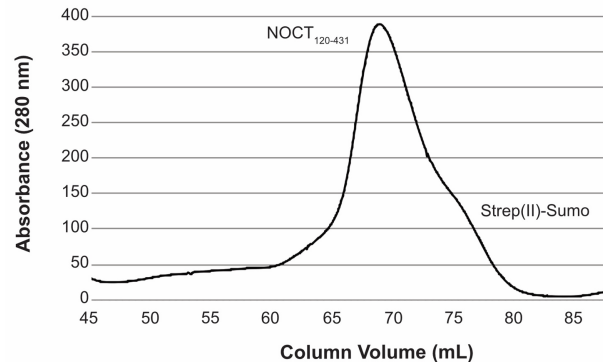
Figure 2.2: NOCT protein purifications are optimized for purity and stability. (A) Schematic of the optimized NOCT$_{120-431}$ purification protocol. **(B)** Chromatogram of NOCT$_{120-431}$ purification on a Q Sepharose anion exchange column. The peak eluting at 90 mL represents Strep-Tactin affinity-purified Strep(II)- Sumo-NOCT. **(C)** The fractions associated with this peak can be visualized on an SDS PAGE gel using Coomassie Blue stain. The second peak at 125 mL represents nucleic acid-containing fractions and samples from this peak are not strongly Coomassie stain but are able to be visualized using ethidium bromide. **(D)** Chromatogram of NOCT$_{120-431}$ purification on a Superdex75 Size Exclusion column. NOCT$_{120-431}$ elutes in a single peak at approximately 70mL, the retention volume corresponding to the NOCT$_{120-431}$ monomer. The shoulder peak at ~75mL corresponds to the Strep(II)-Sumo Tag, which partially copurifies with NOCT$_{120-431}$.

exchange chromatography step to purifications resulted in differential elution of protein

and nucleic acid, effectively removing nucleic acid from recombinant NOCT (Figure 2.2).

Analysis of our final purified protein by measuring the A260/A280 ratios, which were in

the range of 1.8-1.9, indicating that nucleic acid comprised less than 5% of the final purified protein.

NOCT purification buffer was optimized to include sulfate and glycerol additives, which increased protein stability. Size exclusion chromatography showed a single peak corresponding to the $NOCT_{120-431}$ monomer and had very little signal corresponding to protein aggregate (Figure 2.2). The resulting protein from these purifications was sufficiently pure for use in crystallization trials.

2.3.II Purification of NOCT$_{64-431}$ for biochemical studies.

For biochemical assays, we revisited purification of a longer $NOCT_{64-431}$ construct with the optimized purification protocol, as recombinant NOCT was significantly more stable and soluble. The $NOCT_{64-431}$ construct was selected based on the proximity of the N-terminus truncation boundary to a second translation start site in the native *NOCT* mRNA sequence. This start site, which initiates at Met67, excludes predicted disordered regions but approximates a possible native NOCT truncated isoform. To obtain highly purified recombinant enzyme, we isolated $NOCT_{64-431}$ using a stringent, multistep protocol modified from the method used to purify $NOCT_{120-431}$ (see Materials and Methods for purification details) (Figures 2.2 and 2.3). Briefly, NOCT protein was expressed in a modified BL21 STAR *E.coli* cell line with mutant RNase E, which is catalytically inactive against mRNAs but still able to process rRNAs in bacteria (Lopez et al., 1999). Use of this cell line eliminated one potential source of co-purifying nuclease activity. The Strep(II)-tag affinity step was amended to include a 1M NaCl wash to better remove co-purifying RNAs, and the Q Sepharose column used in the anion exchange

Figure 2.3: Purification of human NOCT for biochemical characterization. **(A)** Optimized purification of NOCT$_{64-431}$. **(B)** Initial purification of NOCT using Strep-Tactin affinity purification. **(C)** NOCT is purified using ion exchange chromatography as a second purification step. **(D)** Purification of NOCT using gel filtration chromatography.
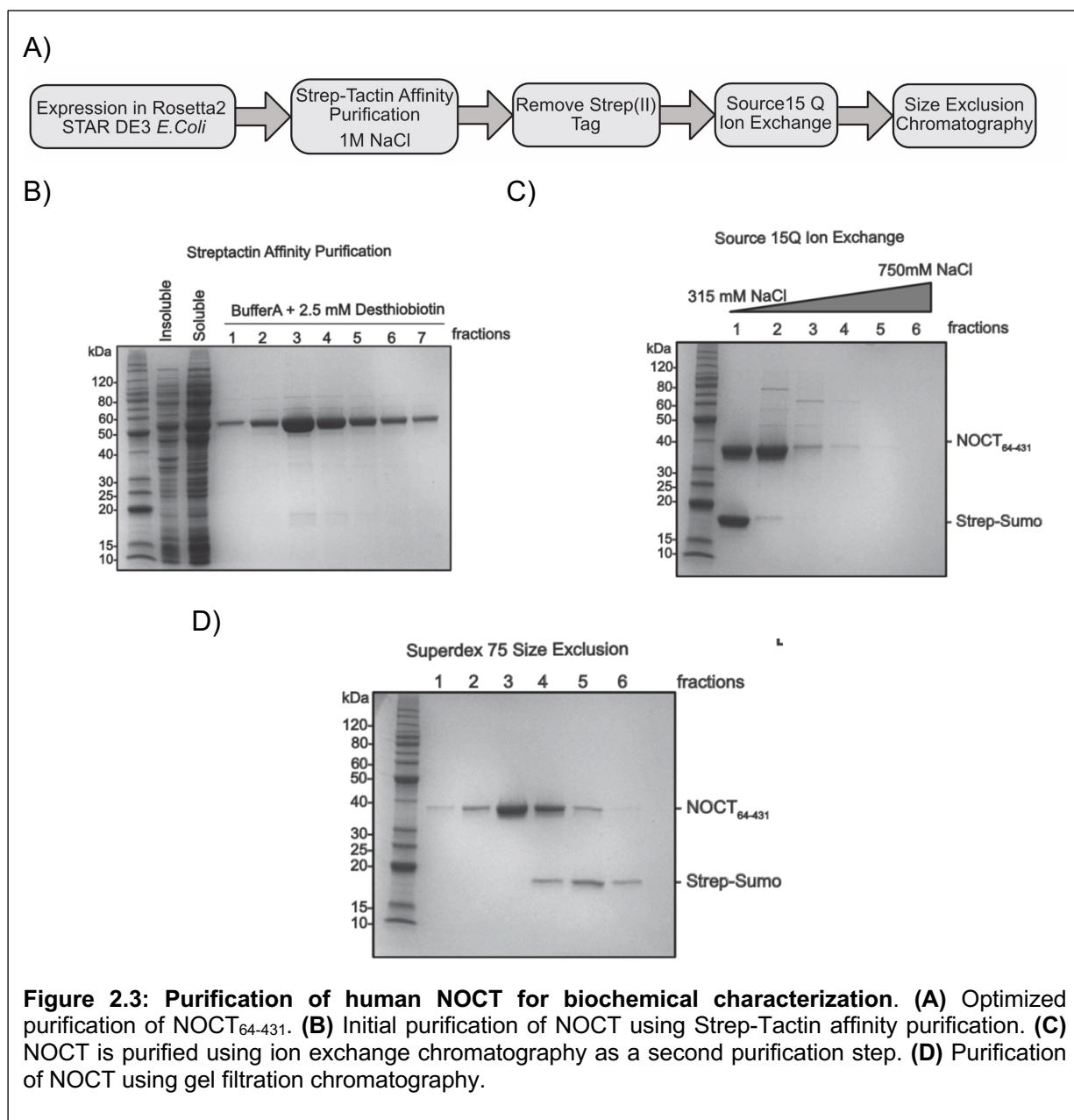
chromatography step was replaced with a higher resolution Source 15Q resin (to improve separation of protein from nucleic acid. This purification scheme resulted in highly purified protein (Figures 2.3 and 2.4).

One commonly described challenge in purifying RNAses is co-purification of contaminating nucleases, which can result from bound nucleic acid remaining associated with the purified protein of interest, carrying other RNA binding proteins (RBPs) and

nucleases (Hrit et al., 2014). In early activity assays used to assess purifications of NOCT,

ribonuclease activity was only observed at relatively high concentrations of NOCT (5 µM),

which is greater than concentrations typically observed in conditions used to assay other

deadenylases (Wang et al., 2010; Wood et al., 2015). This putative was highly variable

between different purifications purported NOCT catalytic mutants purified as controls also

had detectable ribonuclease activity. Thus, the observed RNA degradation was attributed

to a co-purifying nuclease, and NOCT purifications were further optimized to eliminate

this activity (Figures 2.3). As earlier studies reported that crude purifications of mouse

NOCT had activity, we additionally cloned and purified the analogous NOCT construct

from mouse (mNOCT$_{62-429}$) in order to compare activity from the mouse and human

enzymes (Figure 2.4) (Wang et al., 2001).

Figure 2.4: Stringent purification of recombinant NOCT produced highly pure protein for use in biochemical analysis. (A) A representative SDS-PAGE gel of 5 µg of purified NOCT64-431, visualized using silver stain to detect impurities. The developing step was extended to facilitate detection of low-abundance impurities. The expected molecular weight of purified recombinant NOCT64-431 is 41.7 kDa. **(B)** A representative SDS-PAGE gel of 1 µg of purified hNOCT$_{64-431}$, mNOCT$_{62-429}$, single-step Strep-Tactin purified Strep(II)-Sumo- mNOCT$_{62-429}$, CNOT6L$_{158-555}$, and CNOT6L$_{158-555}$ E240A visualized using silver stain.

As positive and negative controls for RNA degradation, we purified the human deadenylase $CNOT6L_{158-555}$ and the catalytically inactive $CNOT6L_{158-555}$ E240A mutant using the same expression and purification procedure used for NOCT. This approach would confirm that the modified purification protocol was compatible with deadenylase activity. The longer $NOCT_{64-431}$ construct used in these studies has analogous N-terminal boundary to the $CNOT6L_{158-555}$ construct used in these assays and previously published studies (Wang et al., 2010). The purity of all enzymes used in biochemical characterization was assessed using silver staining and quantified using densitometry, which determined that the proteins used in biochemical studies have fraction purities ranging from 0.96 to 1.0 (Figure 2.7). Biochemical assays performed with these enzymes will be discussed in Chapter 3.

2.3.III Overall structure of the NOCT catalytic domain.

Highly purified recombinant $NOCT_{120-431}$ was used to solve structures of the NOCT catalytic domain in two different space groups for three separate crystals grown in different conditions (see Materials and Methods for details). An initial NOCT structure was solved at 2.15 Å resolution and was used as a search model for molecular replacement (MR) for the two subsequent and higher-quality structures that were reported (Abshire et al., 2018). These structures were comprised of a 2.41 Å resolution crystal structure (PDP 6BT2) solved in the $P4_12_12$ space group and a 1.48 Å resolution crystal structure (PDB 6BT1) solved in the $P2_12_12_1$ space group (Figure 2.5 and Table 2.1). The NOCT catalytic domain forms an α/β sandwich hydrolase fold in which two anti-parallel β-sheets face each other in the core of the protein. One sheet is formed from the β4- β5- β3- β2- β15- β14 and β1 strands and the other by the β6- β7- β8- β9- β13 and β10 strands.

These β-sheets are surrounded by α and 3_{10} helices, which bury most of the protein core (Figure 2.5).

Figure 2.5: The crystal structure of the NOCT catalytic domain. (A) A cartoon representation of the 1.48 Å resolution structure of the NOCT$_{120-431}$. **(B)** The NOCT 1.48 Å resolution and 2.41 Å resolution structures were aligned (RMSD 0.268 Å over 1619 atoms). No conformational differences are observed with the different Mg^{2+} occupancies. **(C)** Cartoon representation of the NOCT$_{120-431}$ 1.48 Å resolution structure with secondary structural elements of the NOCT catalytic domain colored using PyMOL chainbow settings.

	NOCT (6BT2)	NOCT (6BT1)
Data collection		
Space group	$P2_12_12_1$	$P4_12_12$
Cell dimensions		
a, b, c (Å)	62.0, 69.4, 153.5	61.9, 61.9, 155.5
α, β, γ (°)	90, 90, 90	90, 90, 90
Resolution (Å)	48.20 - 2.41 (2.48-2.41)[a]	42.10 – 1.48 (1.51-1.48)[a]
R_{merge},	0.104 (0.461)	0.085 (0.629)
R_{meas}	0.110 (0.503)	0.112 (0.500)
$I/\sigma(I)$	13.82 (3.34)	15.78 (2.72)
$CC_{1/2}$	(0.895)	(0.680)
CC^*	(0.972)	(0.953)
Completeness (%)	98.9 (90.7)	99.8 (98.0)
Redundancy	8.9 (5.5)	11.6 (4.2)
Refinement		
Resolution (Å)	2.41	1.48
No. reflections	26012 (1301)	51337 (2567)
R_{work} / R_{free}	0.16/0.22	0.15/0.19
No. atoms	4544	2614
Protein	4306	2348
Ligand/ion	14	2
Water	224	264
B factors (Å^2)	42.0	23.8
Protein (Å^2)	41.9	22.4
Ligand/ion (Å^2)	47.5	21.7
Water (Å^2)	44.3	36.8
R.m.s. deviations		
Bond lengths (Å)	0.008	0.006
Bond angles (°)	1.05	1.03
Ramachandran (%)		
Favored (%)	97.0	97.6
Allowed (%)	2.4	2.1
Outliers (%)	0.6	0.3

[a] Values in parentheses are for highest-resolution shell.

Table 2.1: X-ray Data Collection and Refinement Statistics

The NOCT structure also has an unusual π helix, a secondary structural element defined by an i, i+5 hydrogen bonding pattern (in contrast to the *i, i+4* canonical hydrogen bonding pattern for an α-helix or *i, i+3* hydrogen bonding for 3_{10} helices). Functional roles for π helices have been previously described in the literature for other proteins such as fumarase C and lipoxygenase. For the π helices described in these structures, the π helix participates in ligand and/or substrate binding and are located close to these binding sites (Weaver, 2000). The hydrogen bonding patterns of π helices are less energetically favourable than those of α-helices, so it is likely that previously described structures with

Figure 2.6: The NOCT catalytic domain contains a surface-exposed π helix. (A) A cartoon representation of the surface-exposed π helix, located immediately before the α3 helix. **(B)** The alternative hydrogen-bonding pattern found in the π helix can be attributed to Pro205, located in the NOCT α3 helix. **(C)** A depiction of the *i*, i+5 hydrogen bonding pattern for the π helix and α3 helix. Hydrogen bonds are depicted with dashed lines. Hydrogen bond distance cutoffs were assigned an upper limit of 3.3 Å.

π helices stabilize the secondary structure by hydrogen bonding to reinforce the *i*, *i+5* hydrogen bonding pattern in the π helix. This finding is supported by the observation that a subset of these described π helices contain proline residues, which are known "helix breaking" side chains. Ultimately, the functional relevance of the NOCT π helix is unknown, and future studies may identify a role for this structural feature in roles such as protein-protein interactions.

2.3.IV Active site of the NOCT catalytic domain.

Prior to solving the structure of the NOCT catalytic domain, putative active site residues were identified based on homology to other EEP family members (Figure 1.4 and Figure 2.11) (Wang et al., 2010; Wood et al., 2015). These residues cluster together in a solvent-exposed pocket with bound Mg^{2+} ions. The active site is composed of residues Asn149 on β2, Glu195 on β5, His286 on β8, Asp324 on β9, Asp377 on β13, His414, located in the β15 - α6 loop, and Asn326 in the β9 - α4 loop (Figure 2.7).

Two variations of bound ligands were found in the crystal structures of NOCT, with one bound Mg^{2+} in the active site of the 1.48 Å resolution structure and two Mg^{2+} ions in the 2.41 Å resolution structure. Both NOCT structures have a Mg^{2+} coordinated directly to Glu195, which is the primary Mg^{2+}-binding residue in EEP superfamily members. Mutation of this residue leads to catalytic inactivation of many EEP family enzymes, which

Figure 2.7: The arrangement of residue side chains and Mg^{2+} ions in the active site of human NOCT. (A) A detailed view of the active site of the 2.41 Å resolution structure. Two Mg2+ ions are bound in the active site (one coordinated to Glu195 and the other to Asp324) and a sulfate anion is bound between Lys219 and Lys288 adjacent to the active site. Hydrogen bond distance cutoffs were assigned an upper limit of 3.3 Å. **(B)** A detailed view of the active site of the 2.41 Å resolution structure with the waters coordinated to the Mg^{2+} ions illustrated in teal dashed lines. **C)** A detailed view of the active site of the 1.48 Å resolution structure. A single Mg2+ ion is bound in the active site, coordinated to Glu195. Hydrogen bonds are depicted with dashed lines. **D)** A detailed view of the active site of the 1.48 Å resolution structure with the waters coordinated to the Mg^{2+} ion illustrated in teal dashed lines.

would predict that this residue is also critical for NOCT catalytic activity (Freudenthal et al., 2015; Schellenberg et al., 2012; Wang et al., 2010). The second Mg^{2+} observed in the 2.41 Å resolution structure was coordinated to Glu324, which is the analogous to the second Mg^{2+}-binding residue in the CNOT6L apoenzyme structure (Glu410, PDB 3NGQ). This Mg^{2+} binding residue has not been tested in mutagenesis studies and it is unclear whether this Mg^{2+} ion plays a role in catalysis.

2.3.V Electrostatic and conservation analysis of the NOCT catalytic domain.

The electrostatic surface of NOCT has a pronounced basic patch approximately 13 Å long and 7.5 Å wide along the surface adjacent to the putative active site, which may be indicative of an RNA binding site. Similar, albeit less prominent, basic clefts are observed on CNOT6L and PDE12 structures, although the functions of these features have yet to be determined (Figure 2.8). We observed a sulfate anion from the buffer

NOCT (6BT1)   CNOT6L (3NGQ)   PDE12 (4Z0V)

-3 kT/e   +3 kT/e

Figure 2.8: Comparison of surface charges of the human deadenylases NOCT, CNOT6L, and PDE12. Comparison of the electrostatic surface potential of the NOCT, CNOT6L, and PDE12 catalytic domains. Renderings of surface charges are contoured to +/-3 kT/e for NOCT, CNOT6L, and PDE12. Active site Mg2+ ions are depicted in light grey.
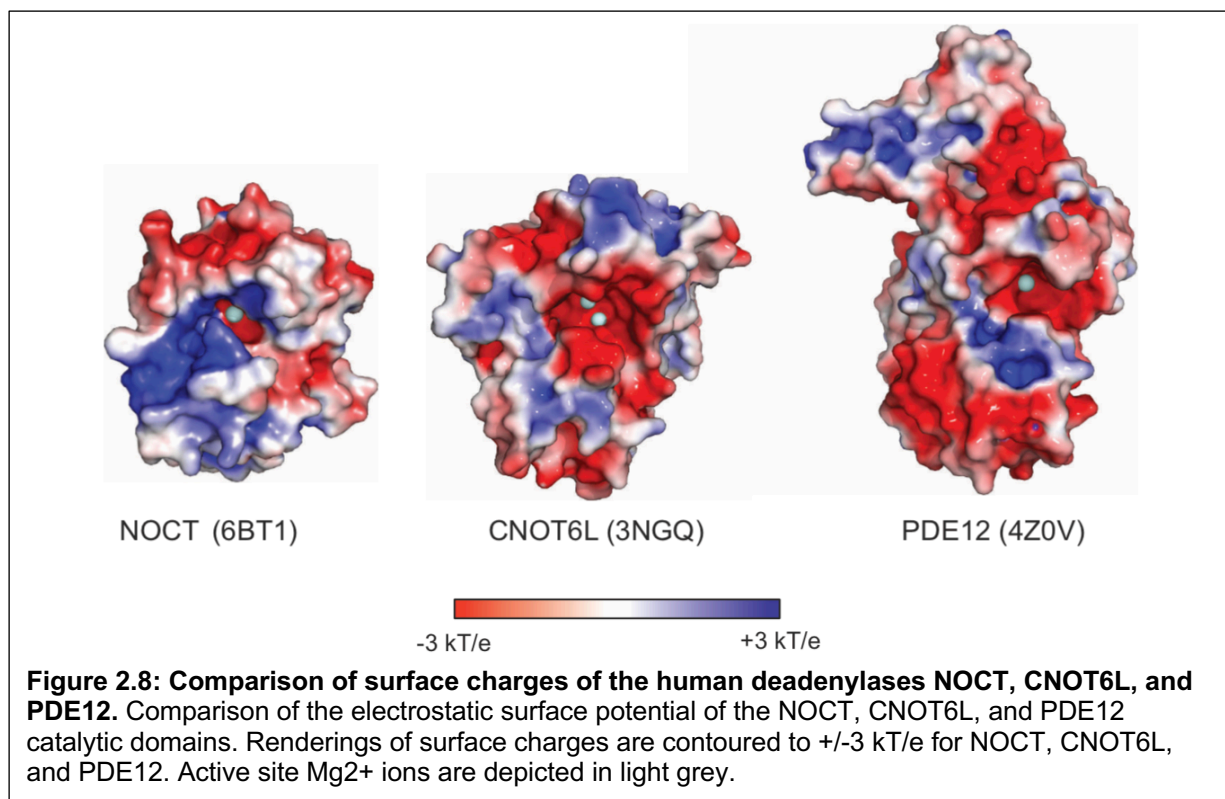
bound in the basic cleft between two lysine residues (Lys219 and Lys288), potentially

mimicking binding of the phosphodiester backbone of RNA (Figure 2.9). Although this

observation has not been directly tested, the basic patch may at least partially explain the

apparent affinity of recombinant NOCT for bacterial nucleic acids, which tend to

persistently co-purify with the NOCT catalytic domain (Figure 2.2).
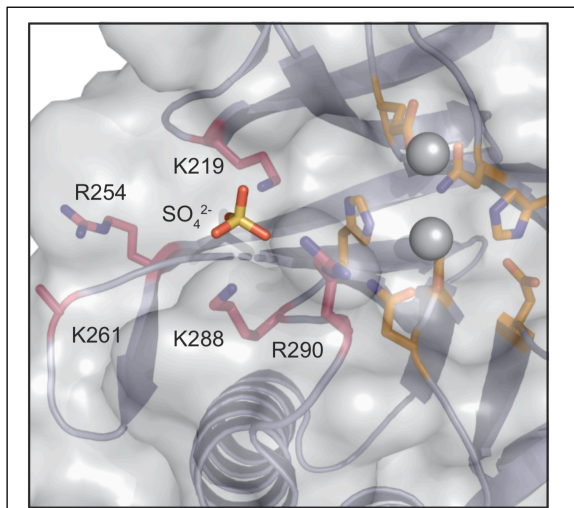
Figure 2.9: Basic patch residues adjacent to the NOCT active site coordinate a sulfate anion. The active site of the 2.41 Å resolution structure and the adjacent basic residues are shown with the enzyme surface corresponding to the active site and basic patch superimposed.

We also examined the sequence conservation of the NOCT active site and basic patch residues among NOCT homologs. ConSurf analysis identifies the most variable and conserved residues of a protein and maps them to a surface representation of the structure. This conservation map indicates that the most conserved residues cluster in the predicted active site and extend into the basic cleft, indicating that the residues located in these regions likely serve important functional

roles (Figure 2.10) (Ashkenazy et al., 2016; Ashkenazy et al., 2010; Celniker G., 2013;

Glaser et al., 2003; Landau et al., 2005). In summary, the NOCT structure illustrates a

conserved hydrolase domain with a conserved active site and divalent metal ion cofactors

located adjacent to a basic cleft, consistent with its proposed function as a ribonuclease.

Figure 2.10: Conservation of residues in the human NOCT catalytic domain and basic patch. (A) The ConSurf server was used to generate a color-coded conservation alignment mapped to the NOCT structure surface representation. The most highly conserved residues are clustered in the active site and the adjacent basic cleft. **(B)** A surface representation of the NOCT 1.48 Å resolution structure. Residues are coloured by conservation based on a multiple sequence alignment generated by ConSurf. The residues that are conserved are coloured in violet, and the most variable residues are coloured in teal. Neutral residues that are neither highly variable nor conserved are white. The most conserved residues are found in the active site and adjacent basic cleft, predicting that these side chains are important for NOCT function.

2.4 Discussion

Here we describe the process for obtaining highly-purified recombinant NOCT protein, which was critical for structural and biochemical studies. The NOCT catalytic forms the highly conserved α/β hydrolase fold that defines members of the EEP deadenylase subfamily. Conserved active site residues form the NOCT active site and

include a Mg^{2+}-binding Glu residue found in many of the EEP enzymes. (Freudenthal et al., 2015; Schellenberg et al., 2012; Wang et al., 2010; Wood et al., 2015). We also identify unique features of NOCT, including a surface exposed π helix and a charged basic patch adjacent to the active site. In addition to our structures, another NOCT catalytic domain structure at 2.7 Å resolution was later published (Estrella et al., 2018). This structure was very similar to the structures discussed here (RMSD= 0.255 over 1861 atoms). Very few differences were observed between these structures, but the structure by Estrella and colleagues was incompletely built in areas, limiting comparison. The authors noted that Arg290 may serve as a lid to obscure the active site; however, the position of this side chain is different in each of the published structures, suggesting that this side chain adopts a variety of conformations. Therefore, there is very little evidence of this residue serving as a gatekeeper to the NOCT active site, as biochemical studies of Arg290 mutants have not been performed to date (Estrella et al., 2018). Overall, the published NOCT structures serve to characterize the general features of the active site in the absence of ligand-bound structures.

The EEP deadenylases are highly structurally homologous enzymes with nearly identical active site conformations, as observed in the structures of human NOCT, CNOT6L, and PDE12 (Figure 2.11). CNOT6L and PDE12 have robust activity against poly(A) substrates (Rorbach et al., 2011; Wang et al., 2010). PDE12 also hydrolyzes the 2´,5´-A oligoadenylate nucleotide (Wood et al., 2015). CNOT6L and PDE12 have been co-crystalized with non-native ligands with chemical structures distinct from the structures of their ligands, indicating that there is some inherent flexibility in the catalytic domain for ligand binding (discussed in Chapter 1, Table 1.1).

NOCT is most similar to the poly(A) specific enzyme CNOT6L, a well-documented CCR4-type deadenylase, but the precise determinants of deadenylase poly(A) specificity remain incompletely understood. Structural and functional studies of CNOT6L have provided some clues regarding this specificity and 3´-exoribonuclease mechanism (Wang et al., 2010). Alignments of NOCT with the EEP enzymes CNOT6L, PDE12, and APE1 illustrate that the aspartate and histidine form a conserved pair that interact through side chain-mediated hydrogen bonding (Figure 2.11) (Freudenthal et al., 2015; Wang et al., 2010; Wood et al., 2015). The imidazolium cation of this histidine may promote active site binding of the reactive phosphate of the nucleic acid substrate through hydrogen bonding, as proposed for the catalytic mechanism of APE1 (Freudenthal et al., 2015). In CNOT6L and NOCT structures, the carboxylate groups of conserved glutamates coordinate Mg^{2+} ions (Figure 2.11). In CNOT6L, coordination of these Mg^{2+} ions mediate the binding of poly(A) DNA. The loss of catalytic activity observed with the CNOT6L E240A mutant likely then results from a loss of Mg^{2+} binding (Wang et al., 2010). Alignment of the NOCT and

Figure 2.11: Comparison of active site structures of NOCT with human deadenylases and APE1. The NOCT 1.48 Å resolution structure (light blue, Mg2+ ion in grey) was aligned with APE1 (5CFG in pink, Mg^{2+} ion in violet, RMSD 8.692Å), CNOT6L (3NGQ in yellow, Mg^{2+} ion in sand, RMSD 1.154Å), and PDE12 (4Z0V in green, Mg^{2+} ion in green, RMSD 1.170 Å). A subset of active site residues is conserved between NOCT and APE1, while the active sites of NOCT, CNOT6L, and PDE12 are highly conserved.

67

CNOT6L illustrates putative differences in the mechanisms of ligand binding for these enzymes (discussed in Chapter 1). The Phe484 residue in CNOT6L reported to be important in adenine binding is not conserved in NOCT which has a cystine (Cys372) at this position. Strikingly, Pro365 in CNOT6L is at the approximate position of Arg290 in NOCT, which is one of the residues in the NOCT basic patch (Figure 2.12). NOCT and CNOT6L may therefore recognize their ligands in distinct manners.

Figure 2.12: Comparison of active site and adjacent residues in the structures of NOCT and CNOT6L. The NOCT 1.48 Å resolution structure (6BT1, light blue, Mg2+ ion in grey) was aligned with CNOT6L (3NGQ in yellow, Mg^{2+} ion in sand, RMSD 1.154Å). The residues described to be involved in base stacking interactions with CNOT6L ligands are not conserved in NOCT.
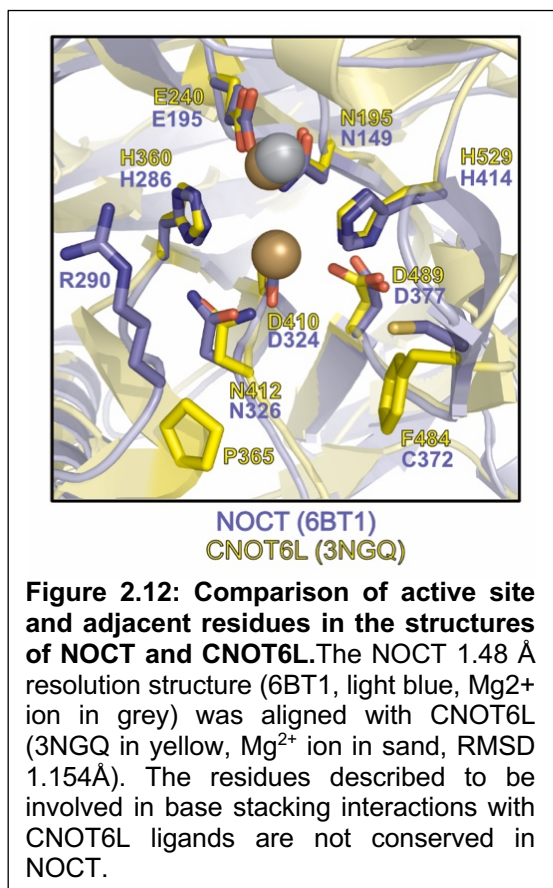
Further supporting these putative differences in predicted RNA binding modes are the differences in surface charges. NOCT has a prominent basic patch adjacent to the active site relative to CNOT6L and PDE12 (Figure 2.8). The contribution of these basic patch residues to substrate residues is unknown and requires identification of an RNA substrate that can be bound by NOCT, as the ability of recombinant NOCT to interact with poly(A) RNA is currently unclear (discussed further in Chapter 3). While the function of some of the features identified in the NOCT structure have yet to be determined, the studies here represent an important step forward in comparative studies of EEP deadenylases by reporting the first crystal structures of NOCT.

2.5 Materials and Methods

Plasmids.

To generate the vectors for protein expression, Nocturnin$_{120\text{-}431}$ or Nocturnin$_{64\text{-}431}$ was cloned into the pSumo vector using BamHI and XhoI restriction sites. The pSumo vector was digested using BamHI and XhoI restriction sites, and the NOCT insert was amplified by PCR, adding a 5′ BamHI overhang and 3′ XhoI overhang. The PCR insert and digested pSumo vector were joined using T4 DNA ligase.

Protein purification for crystallography.

Strep(II)-Sumo-NOCT$_{120\text{-}431}$ was expressed overnight at 18 °C in BL21 (DE3) *E.coli* cells (Millipore Sigma) grown in 2xYT media (Fisher Scientific). Cells were collected by centrifugation and resuspended in 50 mM phosphate buffer (pH 7.5), 150 mM NaCl, 1 mM MgCl$_2$, 5mM β-mercaptoethanol (βME), and 10% glycerol. For 35 mL of resuspended cells, 0.5 mg lysozyme, Thermo Scientific Halt Protease Inhibitor Cocktail (0.5x), NEB Micrococcal Nuclease (100U), and 5.0 mM CaCl$_2$ were added to lyse the cells. Lysis reactions were incubated for 90 minutes at 4 °C, followed by sonication. Lysis was clarified using ultracentrifugation and then filtered using a 0.45 um syringe filter unit. The clarified lysate was purified on a BioRad NGC chromatography system, and lysate was passed over a Strep-Tactin Superflow Plus resin (Qiagen) in 50 mM phosphate buffer (pH 7.5), 300 mM NaCl, 1 mM MgCl$_2$, 5mM βME, and 10% glycerol and eluted with 2.5 mM desthiobiotin. The resulting fractions were analyzed for purity using SDS-PAGE. 15 µL of each fraction was mixed with 5 µL of 4x Laemmli Buffer (BioRad) and samples were heated at 95 °C. Samples were resolved on a 12% SDS PAGE gel and were stained using Coomassie Blue Stain and (50% (v/v) methanol, 10% (v/v) acetic acid, and 0.1%

(w/v) Coomassie Brilliant Blue) destained in deionized (dI) water. All samples were then pooled and the Strep (II) Sumo tag was cleaved overnight at 4 °C in Snake Skin 10,000 MW cutoff dialysis tubing (Thermo) with Ulp1 Sumo protease in 10 mM phosphate buffer (pH 7.5), 20 mM NaCl, 1 mM $MgCl_2$, 10% glycerol, and 5 mM βME. $NOCT_{120-431}$ protein was purified further by ion exchange chromatography using Q Sepharose Fast Flow resin (GE Healthcare) in 10 mM phosphate buffer (pH 7.5), 20 mM NaCl, 1 mM $MgCl_2$, 10% glycerol, and 5 mM βME and was eluted using a 20mM to 1M NaCl gradient. The fractions corresponding to the expected $NOCT_{120-431}$ elution profile were collected and analyzed by SDS PAGE as previously described. Fractions were pooled and concentrated using a 30,000 MW cutoff spin concentrator (Amicon) to a volume of 2 mL. The protein used to produce the 2.41 Å crystal structure was then purified in 50 mM Tris (pH 8.0), 50 mM Na_2SO_4, 1 mM $MgCl_2$, 5% glycerol and 1 mM TCEP over a Superdex 75 gel filtration column (GE Healthcare). The protein used to produce the 1.48 Å crystal structure was purified using the same conditions with minor modifications to the buffer conditions, where 50 mM Tris (pH 8.0), 75 mM Na_2SO_4, 4 mM $MgCl_2$, 10% glycerol and 1 mM TCEP was used. The fractions corresponding to the expected $NOCT_{120-431}$ elution profile were collected and analyzed by SDS PAGE as previously described. These fractions were pooled and concentrated using a 15 mL 30,000 MW cutoff spin concentrator (Amicon) to a volume corresponding to a final protein concentration of approximately 3.5 mg/mL $\pm$ 0.5 mg/mL, as analyzed by absorbance of the protein solution at 280 nm. Protein was aliquoted, flash frozen in liquid nitrogen, and stored at -80 °C.

Protein purification for biochemical assays.

For biochemical assays, Strep(II)-Sumo-NOCT$_{64-431}$, Strep(II)-Sumo-mNOCT$_{62-429}$, Strep(II)-Sumo-CNOT6L$_{158-555}$, and Strep(II)-Sumo-CNOT6L$_{158-555}$ E240A were expressed using a Strep(II) pSumo vector in BL21 STAR (DE3) E. coli cells (ThermoFisher) transformed with the pRare2 plasmid (MilliporeSigma) and purified as described above, with the addition of a 1.0 M NaCl wash (50 mM phosphate buffer (pH 7.5), 1.0 M NaCl, 1 mM MgCl$_2$, 5mM βME and 10% glycerol) after the addition of lysate onto the Strep-Tactin Superflow Plus resin. The single step-purified mNOCT was concentrated and stored after the Strep-Tactin purification step. For all other proteins, the purifications were completed as described above. The ion exchange step was performed using Source 15Q resin (GE Healthcare) and the CNOT6L WT and E240A mutant proteins were purified in the phosphate and Tris buffers at pH 8.0, as described in the previous section. Human NOCT constructs were cloned by P. Del Rizzo.

X-ray crystallography.

For the crystals used to determine the 2.15 Å resolution X-ray structure, aliquots of frozen NOCT$_{120-431}$ were thawed and supplemented with 1.5 mM AMP and 8 mM MgCl$_2$. Solutions were allowed to equilibrate on ice for 30 min before being concentrated in a 500 μL 10,000 MW cutoff spin concentrator (Amicon) to a volume corresponding to a final protein concentration of approximately 9 mg/mL $\pm$ 0.5 mg/mL. Protein solutions were clarified by centrifugation at 18,000 x g and 4 °C for 1 minute. Crystals were grown using hanging drop vapor diffusion at room temperature (20 – 25 °C) in 2 μL drop volumes with a 1:1 ratio of protein solution and well solution. Crystals formed in 1-3 days in the conditions optimized for crystallography (100 mM HEPES buffer (pH 7.5), 100mM

ammonium sulfate, 14% PEG 10,000, and 4% ethylene glycol). Crystals were harvested on LithoLoops (Molecular Dimensions) and were flash frozen in liquid nitrogen in cryoprotectant (100 mM HEPES buffer (pH 7.5), 100mM ammonium sulfate, 28% PEG 10,000, and 4% ethylene glycol).

For the crystals used to solve the 2.41 Å resolution crystal structure, aliquots of frozen $NOCT_{120-431}$ were thawed and supplemented with 25 mM AMP and 10 mM $MgCl_2$. Solutions were allowed to equilibrate on ice for 30 min before solutions were concentrated in a 500uL 10,000 MW cutoff spin concentrator (Amicon) to a volume corresponding to a final protein concentration of approximately 9 mg/mL $\pm$ 0.5 mg/mL. Protein solutions were clarified by centrifugation at 18,000 x g and 4 °C for 1 minute. Crystals were grown using hanging drop vapor diffusion at room temperature (20 – 25 °C) in 2 µL drop volumes with a 1:1 ratio of protein solution and well solution. Crystals formed in 1-3 days in the conditions optimized for crystallography (100 mM HEPES buffer (pH 7.5), 5% isopropanol, 200 mM sodium acetate, and 10% PEG 4000). Crystals were harvested on LithoLoops (Molecular Dimensions) and were flash frozen in liquid nitrogen in cryoprotectant (100 mM HEPES buffer (pH 7.5), 5% isopropanol, 200 mM sodium acetate, 10% PEG 4000, 20% PEG 400, 10mM AMP, and 10mM $MgCl_2$).

For the crystals used to determine the 1.48 Å resolution crystal structure, samples were treated as described above for the 2.41 resolution crystal structure with the modifications indicated here. Briefly, $NOCT_{120-431}$ protein solutions were supplemented with 10 mM 3´ , 5´ adenosine diphosphate and 6.0 mM $MgCl_2$. Crystals were grown using hanging drop vapor diffusion with 3 µL drop volumes with a 1:1:1 ratio of protein solution, distilled water, and well solution. Crystals formed over 1-3 days and were grown in 100

mM Tris (pH 8.5), 100 mM $MgCl_2$, 100 mM sodium citrate, and 14% PEG 4000. Crystals were harvested on LithoLoops (Molecular Dimensions) and were flash frozen in liquid nitrogen in cryoprotectant (100 mM Tris (pH 8.5), 100 mM $MgCl_2$, 100 mM sodium citrate, 14% PEG 4000, 20% PEG 400, 10 mM 3´, 5´ adenosine diphosphate and 6.0 mM $MgCl_2$). Crystallization of the crystals used to determine the 1.48 Å resolution crystal structure was determined by J. Chasseur.

Data collection and structure determination.

X-ray diffraction data were collected at the Argonne National Laboratories on the LS-CAT beamline 21 ID-F for the 2.15 Å initial crystal structure and on beamline 21 ID-G for the 2.41 Å and 1.48 Å crystal structures. Data were then indexed, integrated, and scaled using HKL-2000 (Otwinowski and Minor, 1997). The 2.15 Å dataset belonged to the space group $P2_12_12_1$. This structure was solved by molecular replacement (MR) using a homology model of $NOCT_{120-431}$ generated using I-TASSER (Roy et al., 2010). The generated homology model was based off of previously published structures of the closely related EEP deadenylase CNOT6L as the homology modeling template. The resulting MR solution was then used as a starting model for AutoBuild in Phenix, followed by iterative rounds of refinement in Phenix refine and model building in Coot (Emsley et al., 2010). Metal ions present in the active sites of these structures were determined by generating an omit Fo-Fc omit map in models lacking Mg^{2+} ions and the coordinated water molecules, and the map was contoured to 5σ. Strong positive density peaks were observed, indicating that these peaks in the active site correspond to bound metal ions. These peaks were assumed to be Mg^{2+} ions based on the composition of our protein and well solutions used in generating crystals, and the Mg^{2+} and coordinating waters were

modeled into the active site. No density corresponding to the AMP nucleotide used in crystallization conditions was observed. The 2.41 Å and 1.48 Å structures were solved in the P4$_1$2$_1$2 and P2$_1$2$_1$2$_1$, respectively. Structures were determined using MR, with the initial 2.15 Å structure as the search model for MR of the 2.41 Å dataset and the 2.14 Å structure was used for MR of the 2.41 Å dataset. Refinement and model building performed as described above.

Surface electrostatic potentials were calculated for the 1.48 Å resolution structure in MacPyMOLx11Hybrid using the APBS Tools 2.1 Plugin. Input files were created using PDB2PQR using AMBER charges (ff99) and radii, which excluded the solvation shell and Mg^{2+} ions (Baker et al., 2001; Dolinsky et al., 2007; Dolinsky et al., 2004). Surface electrostatics were compared between NOCT, PDE12 (PDB 4Z0V) and CNOT6L (PDB 3NGQ) by contouring each to +/- 3 kT/e. Surface renderings of the NOCT 1.48 Å resolution structure colored by conservation were generated using the ConSurf Server (http://consurf.tau.ac.il/2016/). The homolog search was done using the HMMER algorithm and the UniProt protein database. The minimal ID threshold was set to 15% to reflect the low sequence homology of NOCT and the most closely related human protein homologs CNOT6L (18% sequence identity) and PDE12 (16% sequence identity). The multiple sequence alignment was generated with MAFFT-L-INS-I (Ashkenazy et al., 2016; Ashkenazy et al., 2010; Celniker G., 2013; Glaser et al., 2003; Landau et al., 2005).

2.6 References

Abshire, E.T., Chasseur, J., Bohn, J.A., Del Rizzo, P.A., Freddolino, P.L., Goldstrohm, A.C., and Trievel, R.C. (2018). The structure of human Nocturnin reveals a conserved ribonuclease domain that represses target transcript translation and abundance in cells. Nucleic acids research *46*, 6257-6270.

Ashkenazy, H., Abadi, S., Martz, E., Chay, O., Mayrose, I., Pupko, T., and Ben-Tal, N. (2016). ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules. Nucleic acids research *44*, W344-350.

Ashkenazy, H., Erez, E., Martz, E., Pupko, T., and Ben-Tal, N. (2010). ConSurf 2010: calculating evolutionary conservation in sequence and structure of proteins and nucleic acids. Nucleic acids research *38*, W529-533.

Baggs, J.E., and Green, C.B. (2003). Nocturnin, a deadenylase in Xenopus laevis retina: a mechanism for posttranscriptional control of circadian-related mRNA. Current biology : CB *13*, 189-198.

Baker, N.A., Sept, D., Joseph, S., Holst, M.J., and McCammon, J.A. (2001). Electrostatics of nanosystems: application to microtubules and the ribosome. Proceedings of the National Academy of Sciences of the United States of America *98*, 10037-10041.

Celniker G., N.G., Ashkenazy H., Glaser F., Martz E., Mayrose I., Pupko T., and Ben-Tal N. (2013). ConSurf: Using Evolutionary Data to Raise Testable Hypotheses about Protein Function. Israel Journal of Chemistry.

Dolinsky, T.J., Czodrowski, P., Li, H., Nielsen, J.E., Jensen, J.H., Klebe, G., and Baker, N.A. (2007). PDB2PQR: expanding and upgrading automated preparation of biomolecular structures for molecular simulations. Nucleic acids research *35*, W522-525.

Dolinsky, T.J., Nielsen, J.E., McCammon, J.A., and Baker, N.A. (2004). PDB2PQR: an automated pipeline for the setup of Poisson-Boltzmann electrostatics calculations. Nucleic acids research *32*, W665-667.

Emsley, P., Lohkamp, B., Scott, W.G., and Cowtan, K. (2010). Features and development of Coot. Acta Crystallogr D Biol Crystallogr *66*, 486-501.

Estrella, M.A., Du, J., and Korennykh, A. (2018). Crystal Structure of Human Nocturnin Catalytic Domain. Scientific reports *8*, 16294.

Freudenthal, B.D., Beard, W.A., Cuneo, M.J., Dyrkheeva, N.S., and Wilson, S.H. (2015). Capturing snapshots of APE1 processing DNA damage. Nature structural & molecular biology *22*, 924-931.

Glaser, F., Pupko, T., Paz, I., Bell, R.E., Bechor-Shental, D., Martz, E., and Ben-Tal, N. (2003). ConSurf: identification of functional regions in proteins by surface-mapping of phylogenetic information. Bioinformatics *19*, 163-164.

Green, C.B., and Besharse, J.C. (1996). Identification of a novel vertebrate circadian clock-regulated gene encoding the protein nocturnin. Proceedings of the National Academy of Sciences of the United States of America *93*, 14884-14888.

Hrit, J., Raynard, N., Van Etten, J., Sankar, K., Petterson, A., and Goldstrohm, A.C. (2014). In vitro analysis of RNA degradation catalyzed by deadenylase enzymes. Methods in molecular biology *1125*, 325-339.

Landau, M., Mayrose, I., Rosenberg, Y., Glaser, F., Martz, E., Pupko, T., and Ben-Tal, N. (2005). ConSurf 2005: the projection of evolutionary conservation scores of residues on protein structures. Nucleic acids research *33*, W299-302.

Linding, R., Russell, R.B., Neduva, V., and Gibson, T.J. (2003). GlobPlot: Exploring protein sequences for globularity and disorder. Nucleic acids research *31*, 3701-3708.

Lopez, P.J., Marchand, I., Joyce, S.A., and Dreyfus, M. (1999). The C-terminal half of RNase E, which organizes the Escherichia coli degradosome, participates in mRNA degradation but not rRNA processing in vivo. Mol Microbiol *33*, 188-199.

Otwinowski, Z., and Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. Methods in enzymology *276*, 307-326.

Pearce, S.F., Rorbach, J., Van Haute, L., D'Souza, A.R., Rebelo-Guiomar, P., Powell, C.A., Brierley, I., Firth, A.E., and Minczuk, M. (2017). Maturation of selected human mitochondrial tRNAs requires deadenylation. Elife *6*.

Rorbach, J., Nicholls, T.J., and Minczuk, M. (2011). PDE12 removes mitochondrial RNA poly(A) tails and controls translation in human mitochondria. Nucleic acids research *39*, 7750-7763.

Roy, A., Kucukural, A., and Zhang, Y. (2010). I-TASSER: a unified platform for automated protein structure and function prediction. Nat Protoc *5*, 725-738.

Schellenberg, M.J., Appel, C.D., Adhikari, S., Robertson, P.D., Ramsden, D.A., and Williams, R.S. (2012). Mechanism of repair of 5´ -topoisomerase II-DNA adducts by mammalian tyrosyl-DNA phosphodiesterase 2. Nature structural & molecular biology *19*, 1363-1371.

Wang, H., Morita, M., Yang, X., Suzuki, T., Yang, W., Wang, J., Ito, K., Wang, Q., Zhao, C., Bartlam, M.*, et al.* (2010). Crystal structure of the human CNOT6L nuclease domain reveals strict poly(A) substrate specificity. The EMBO journal *29*, 2566-2576.

Wang, Y., Osterbur, D.L., Megaw, P.L., Tosini, G., Fukuhara, C., Green, C.B., and Besharse, J.C. (2001). Rhythmic expression of Nocturnin mRNA in multiple tissues of the mouse. BMC developmental biology *1*, 9.

Weaver, T.M. (2000). The pi-helix translates structure into function. Protein Sci *9*, 201-206.

Wood, E.R., Bledsoe, R., Chai, J., Daka, P., Deng, H., Ding, Y., Harris-Gurley, S., Kryn, L.H., Nartey, E., Nichols, J*., et al.* (2015). The Role of Phosphodiesterase 12 (PDE12) as a Negative Regulator of the Innate Immune Response and the Discovery of Antiviral Inhibitors. The Journal of biological chemistry *290*, 19681-19696.

CHAPTER 3

Biochemical Characterization of NOCT[1]

3.1 Abstract

The circadian protein Nocturnin (NOCT) belongs to the exonuclease, endonuclease, and phosphatase (EEP) superfamily and is most similar to the CCR4-class of deadenylases that degrade the 3´ poly-adenosine (poly(A)) tails of mRNAs. The active site of NOCT is highly conserved with other exoribonucleases such as human CNOT6L and PDE12. In contrast to the related deadenylases, purified recombinant NOCT lacks *in vitro* ribonuclease activity, suggesting that unidentified factors are necessary for enzymatic activity. In addition to lacking ribonuclease activity, NOCT is not active against substrates of other EEP family members such as DNA and a select panel of phosphorylated signaling molecules. These studies demonstrate that NOCT behaves in a manner distinct from other EEP deadenylases and suggests that NOCT activity is regulated in the cellular context through mechanisms such as post-translational modification or by protein partners.

---

[1] This work is published in Abshire, E.T., Chasseur, J., Bohn, J.A., Del Rizzo, P.A., Freddolino, P.L., Goldstrohm, A.C., and Trievel, R.C. (2018). The structure of human Nocturnin reveals a conserved ribonuclease domain that represses target transcript translation and abundance in cells. Nucleic acids research *46*, 6257-6270.

3.2 Introduction

Much of what is known about the activity of human EEP deadenylases enzymes has resulted from studies of CNOT6L, a Ccr4 homolog, and PDE12. CNOT6L catalyzes removal of mRNA poly(A) tails as a subunit of the human CCR4-NOT complex, which regulates mRNA decay and translational repression globally (Doidge et al., 2012; Maryati et al., 2015; Mittal et al., 2011). PDE12 is a more specialized deadenylase localized to the mitochondria and plays a role in tRNA quality control. Mitochondrial tRNAs that are erroneously 3´ polyadenylated by the mitochondrial poly(A) polymerase cannot be properly charged with their cognate amino acids, which impairs translation. Therefore, PDE12 is needed to remove these poly(A) tails in order to maintain tRNA quality in the mitochondria (Pearce et al., 2017). The biochemical activities of the human EEP enzymes therefore maintain the quality and abundance of multiple types of RNAs in the cell. Human CNOT6L and PDE12 are both active against poly(A) RNA in *in vitro* assays and the structural homology between NOCT and these enzymes suggest that NOCT has ribonuclease activity against poly(A) RNA as well (discussed in Chapter 2) (Rorbach et al., 2011; Wang et al., 2010). Mutagenesis studies of residues in the CNOT6L active site include those that are conserved in NOCT, suggesting that these same residues may be important for NOCT function as well (Wang et al., 2010).

NOCT was first discovered in the retina of *Xenopus laevis* (xNOCT) in a screen for genes exhibiting circadian behavior (Green and Besharse, 1996a, b, 1997). Based on sequence homology, xNOCT and its homologs were described as Ccr4-type deadenylases and predicted to act through enzymatic removal of 3´ poly(A) tails of its target transcripts in order to promote mRNA decay (Baggs and Green, 2003, 2006).

79

Putative RNA decay activity was reported for xNOCT against short RNA with mixed sequence and a 3′ 100-mer poly(A) tail. The observed ribonuclease activity was ablated upon addition of EDTA, indicating that RNA degradation was Mg^{2+}-dependent. Mutagenesis of the predicted Mg^{2+}-binding residue of xNOCT (E152A) had the same effect. This study also indicated that NOCT prefers poly(A) RNA over poly(A) DNA or mixed sequence RNA substrates (Baggs and Green, 2003). Ribonuclease activity has also been reported for the murine homolog of NOCT (mNOCT). The E193A Mg^{2+} binding residue was reported to also eliminate activity; however, reaction products were still observed using mutant mNOCT in reactions against polyadenylated RNA substrates (Garbarino-Pico et al., 2007). A comprehensive biochemical analysis of NOCT activity is important to understanding its biological function, and here we discuss characterization of the human enzyme, which had not been previously studied.

3.3 Results
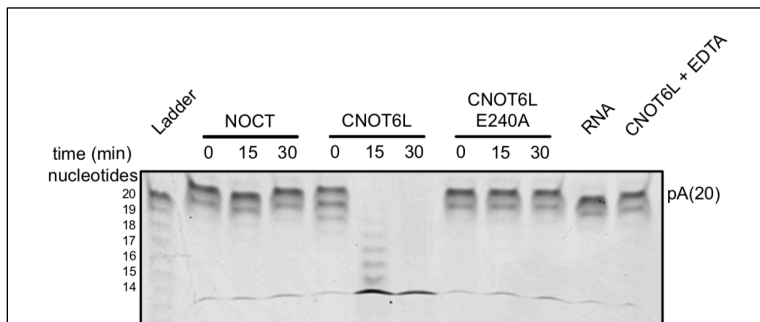
Figure 3.1: NOCT is inactive against a poly(A)$_{20}$ RNA substrate *in vitro*. Recombinant NOCT$_{64-431}$, CNOT6L$_{158-555}$, and CNOT6L$_{158-555}$ E240A were incubated with 5′ Cy3 labelled poly(A)$_{20}$ RNA substrate. A representative gel is shown here. Where indicated, the chelator EDTA was added to stop the ribonuclease reaction.

We measured the activity of recombinant wild-type (WT) NOCT$_{64-431}$ against a poly(A)$_{20}$ RNA substrate in the reaction buffer previously reported for CNOT6L, as these conditions were known to be compatible with deadenylase activity. In

Figure 3.2: NOCT is inactive in a range of ionic strengths and Mg^{2+} ion concentrations. (A) The Mg^{2+} ion concentration was titrated in NOCT ribonuclease reactions and RNA product accumulation was measured. A representative gel is shown. **(B)** The ionic strength of the assay buffer was titrated in NOCT ribonuclease reactions.
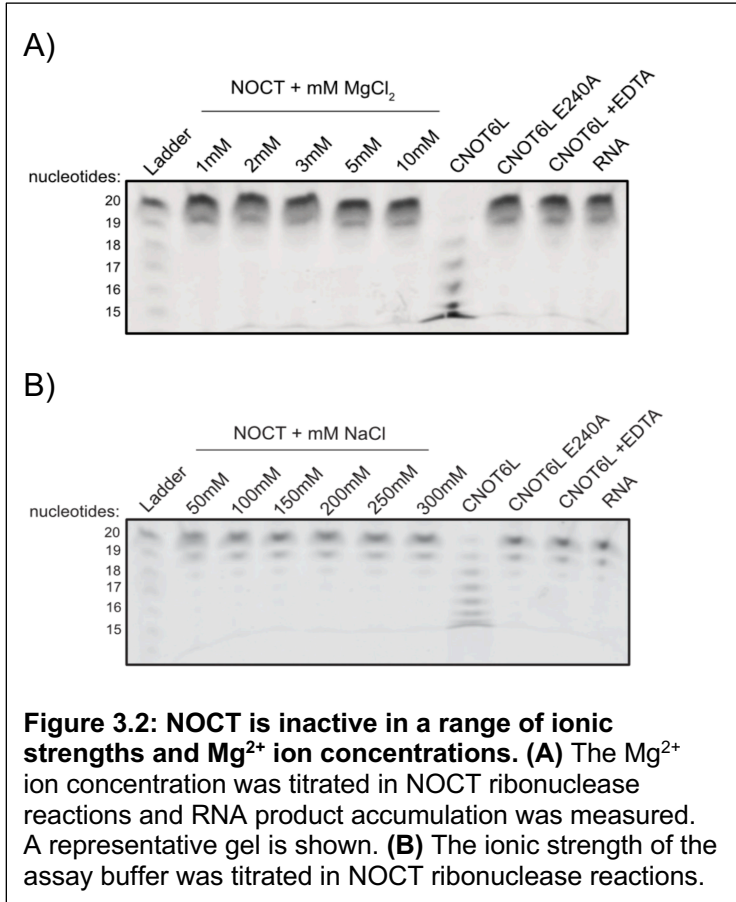
these conditions, WT NOCT$_{64-431}$ did not degrade the poly(A)$_{20}$ RNA substrate, whereas WT CNOT6L had robust ribonuclease activity (Figure 3.1). Consistent with prior studies, the CNOT6L E240A catalytic mutant was inactive and chelation of Mg^{2+} with EDTA inhibited WT CNOT6L activity (Wang et al., 2010).

We considered that the reaction conditions may be suboptimal for human NOCT, and thus explored the effects of divalent and monovalent ions. First, we titrated Mg^{2+}and NaCl concentrations, spanning concentrations necessary for activities of other deadenylases (Figure 3.2) (Lowell et al., 1992; Wang et al., 2010). NOCT activity was not observed in any of the conditions that we tested. Following these assays, we titrated the amount of poly(A) RNA substrate from concentrations below that of NOCT to concentrations well in excess of the enzyme, again observing that NOCT was inactive in all of the tested conditions (Figure 3.3).
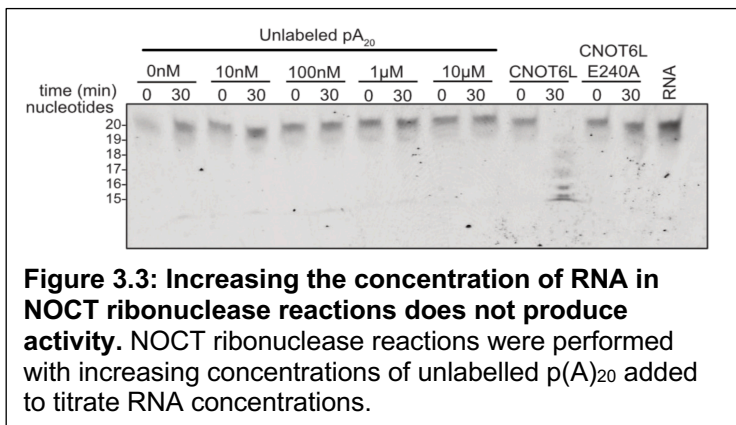
Figure 3.3: Increasing the concentration of RNA in NOCT ribonuclease reactions does not produce activity. NOCT ribonuclease reactions were performed with increasing concentrations of unlabelled p(A)$_{20}$ added to titrate RNA concentrations.

As many EEP enzymes and all EEP deadenylases are Mg^{2+}-dependent activity, we had expected NOCT to be similarly dependent on Mg^{2+} for ribonuclease activity (Freudenthal et al., 2015; Schellenberg et al., 2012; Wang et al., 2010; Wood et al., 2015). Additionally, in the NOCT crystal
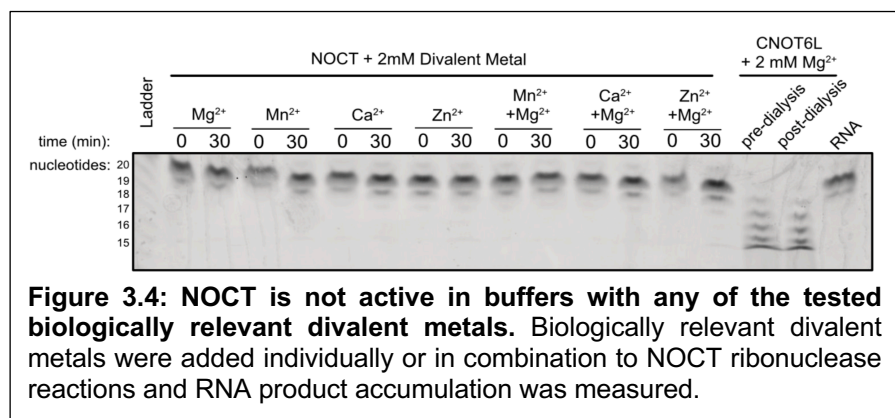
Figure 3.4: NOCT is not active in buffers with any of the tested biologically relevant divalent metals. Biologically relevant divalent metals were added individually or in combination to NOCT ribonuclease reactions and RNA product accumulation was measured.

structures, the metals bound in the active site show octahedral coordination to Glu195 and active site waters, consistent with Mg^{2+} ion binding. To ascertain whether NOCT instead uses an alternative metal, we screened biologically relevant divalent metal cations including Mn^{2+}, Ca^{2+}, and Zn^{2+} individually and in combination with Mg^{2+}. NOCT activity was not observed in any of the conditions tested, which indicates that the inactivity observed in our exoribonuclease assays is not due to the inavailability of one the metal cofactors tested in these assays.

Though the human and murine homologs of NOCT are very highly conserved (91% sequence identity) and are therefore unlikely to have distinct activities, we also considered the potential for species-specific differences in NOCT activity. We compared the activities of hNOCT and mouse NOCT (mNOCT$_{62-429}$) purified under our multi-step, stringent conditions. Strep(II)-Sumo-mNOCT$_{62-429}$ was additionally purified using Strep-Tactin resin and a high salt wash (single step mNOCT) (detailed in the Chapter 2 Materials and Methods) (Figure 2.3). Using the conditions reported for CNOT6L, neither

human NOCT, mNOCT, nor single step purified mNOCT had detectible ribonuclease activity (Figure 3.5, top panel) (Wang et al., 2010).

We then assayed NOCT and mNOCT activity in two deadenylase reaction buffers that were previously reported to support activity of mNOCT or xNOCT to determine if the CNOT6L assay buffer was incompatible with NOCT ribonuclease activity (Figure 3.5) (Baggs and Green, 2003; Garbarino-Pico et al., 2007). It is noteworthy that the reaction
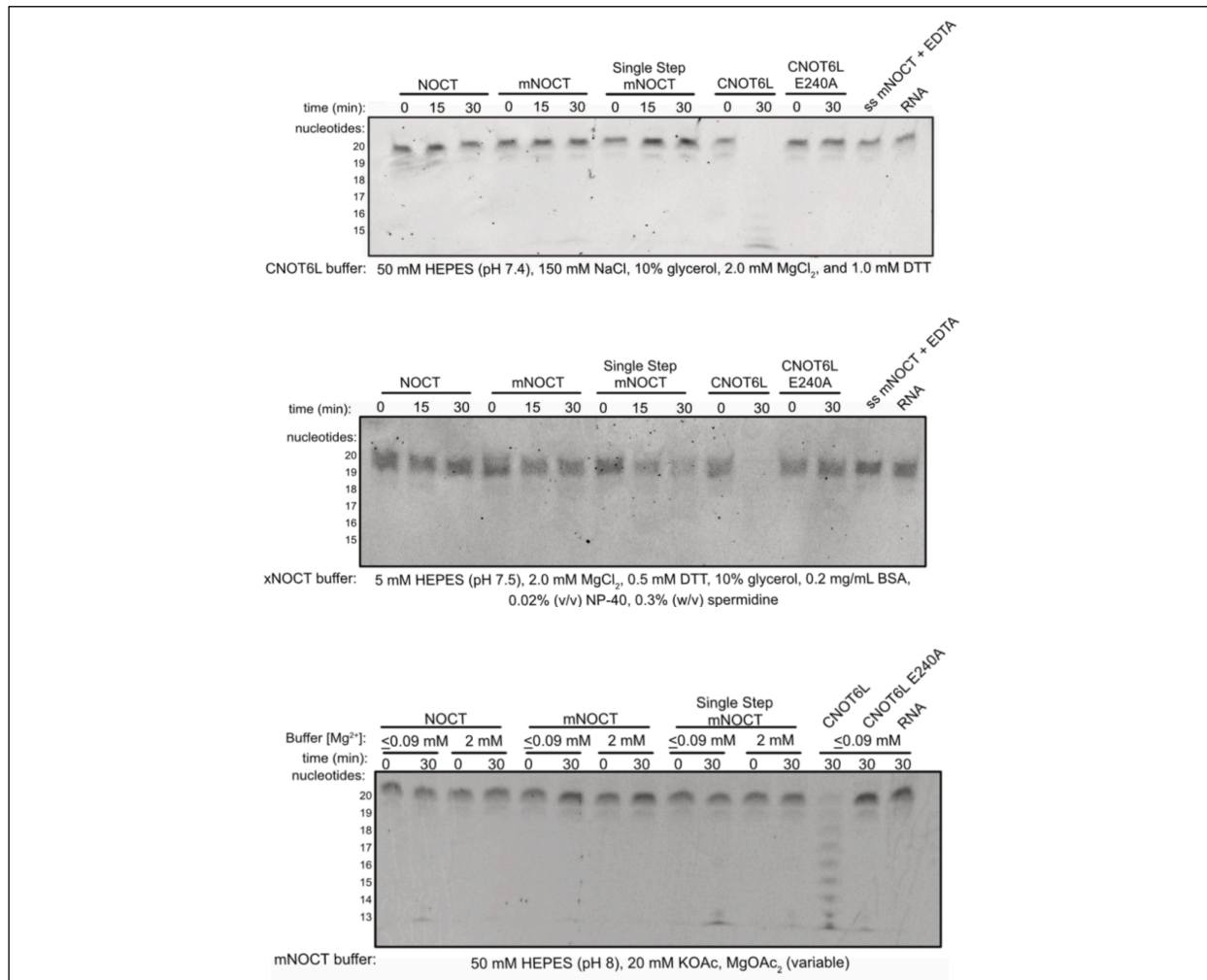
Figure 3.5: Human and mouse NOCT are not active in previously reported conditions used in biochemical studies of mNOCT and xNOCT. NOCT ribonuclease reactions were performed with NOCT, mNOCT, or single step purified mNOCT. NOCT activity in CNOT6L, xNOCT, and mNOCT assay buffers were compared. In the mNOCT buffer conditions with 0.01 mM MgOAc$_2$, the final concentration of Mg^{2+} carried over from the protein purification is accounted for. Final Mg^{2+} concentrations are sub-physiological and are as follows: 0.04 mM (NOCT), 0.05 mM (mNOCT), 0.06 mM (single step (ss) purified mNOCT), 0.05 mM (CNOT6L), and 0.09 mM (CNOT6L E240A).

conditions previously employed for measuring mNOCT or xNOCT activity contained either very low ionic strength or sub-physiological Mg^{2+} concentrations. In contrast, we used physiological ionic strength and Mg^{2+} concentrations in our initial conditions (Figures 3.1-3.4), which were similar to those used to characterize the catalytic activity of human CNOT6L and PDE12. We did not observe deadenylase activity for NOCT in the xNOCT or mNOCT reaction buffers, whereas CNOT6L, but not the E240A mutant, could readily degrade poly(A) RNA (Figure 3.5). We conclude that purified recombinant human or mouse NOCT do not have detectable *in vitro* deadenylase activity under these conditions. Corroborating this result, a subsequent paper demonstrated no detectable human NOCT ribonuclease activity against poly(A) RNA (Estrella et al., 2018).

We then turned to examining other potential substrates for NOCT to see if another substrate would elicit activity. To examine whether NOCT has RNA endonuclease activity, we tested a long ~1100 nucleotide (nt) RNA substrate containing the *Renilla* Luciferase ORF with a short 3′ UTR and a 60 nt poly(A) tail. (Figure 3.6). It should be noted that the electrophoretic conditions for this assay are insufficient to resolve the size difference between the full-length substrate and a deadenylated substrate lacking the 60 nt poly(A) tail. Therefore, this assay is not likely to resolve exoribonuclease activity unless the RNA were to be degraded past the poly(A)$_{60}$ tail. For
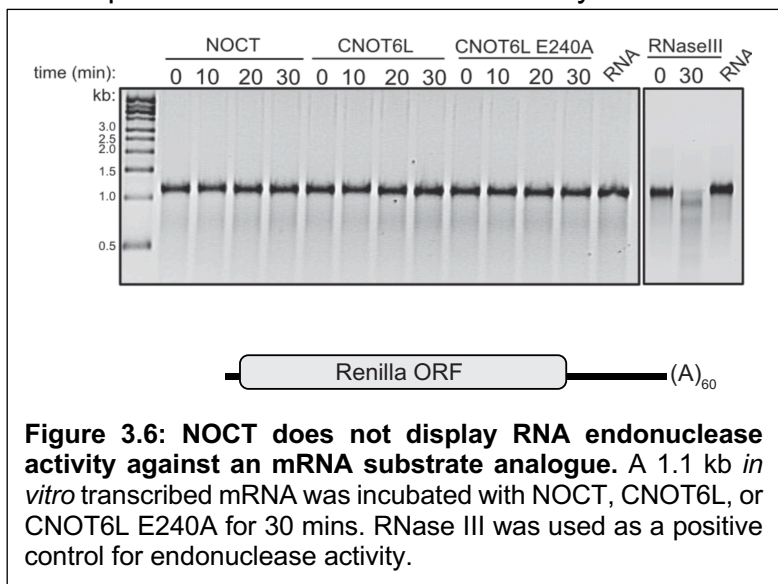
Figure 3.6: NOCT does not display RNA endonuclease activity against an mRNA substrate analogue. A 1.1 kb *in vitro* transcribed mRNA was incubated with NOCT, CNOT6L, or CNOT6L E240A for 30 mins. RNase III was used as a positive control for endonuclease activity.

example, CNOT6L should readily deadenylate this mRNA substrate in this assay but this activity is not apparent. As a positive control for endonuclease activity, RNase III activity was assessed and shown to be distinguishable using these experimental conditions.

Figure 3.7: NOCT is not active against DNA substrates. NOCT was incubated with ssDNA sand dsDNA 24-mers and reaction products were resolved on a gel. Micrococcal Nuclease (MNase) was used as a positive control for ssDNA degradation and DNase I was used as a positive control for dsDNA degradation.
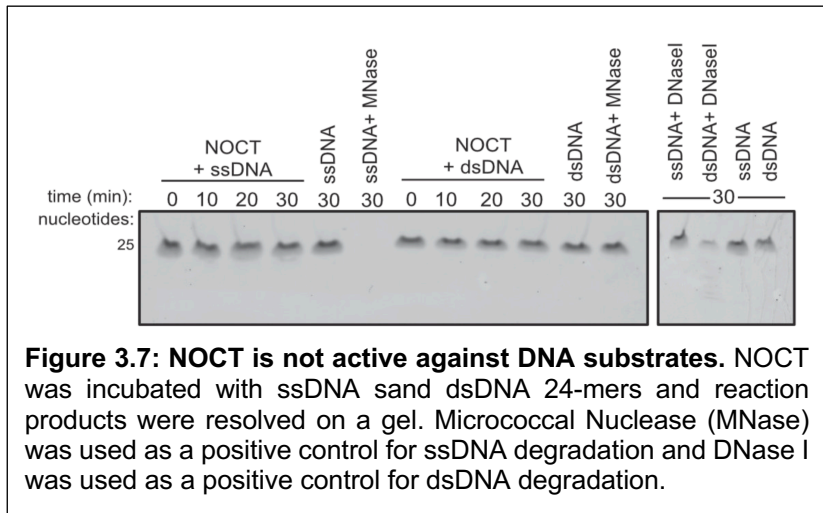
Other members of the EEP superfamily, such as DNase I and APE1, exhibit endonuclease activity against DNA substrates (Freudenthal et al., 2015; Shiokawa and Tanuma, 2001). To test for activity against DNA, we tested whether NOCT can hydrolyze single stranded DNA (ssDNA) and double stranded DNA (dsDNA) 24-mers of mixed sequence. While this does not comprehensively sample the possible DNA sequences that could be NOCT substrates, a more thorough screen of substrates was not undertaken. Currently, there are no indications that NOCT is active against DNA substrates, as closely related enzymes CNOT6L and PDE12 do not display significant DNA exonuclease activity (Rorbach et al., 2011; Wang et al., 2010).

NOCT is a member of the larger EEP superfamily which includes enzymes that hydrolyze a wide range of phosphorylated substrates. A subset of EEP family members have phosphatase activity, including the inositol phosphatases. Notably, these enzymes have hydrolase activity against a wide range of substrates in *in vitro* assays, potentially allowing for identification of NOCT phosphatase activity against suboptimal or non-

physiological substrates (Dyson et al., 2012). We therefore screened a collection of

phosphorylated signalling molecules, nucleotides, lipids and carbohydrates to determine

if NOCT has phosphatase activity against any of these potential substrates (Table 3.1).

Additionally, we included a panel of phosphorylated compounds representing known

metabolic intermediates from glycolysis and the pentose phosphate pathway. Regulation

of such intermediates would be potentially congruent with the metabolic phenotypes

observed in NOCT knockout mice (Green et al., 2007; Stubblefield et al., 2018). However,

NOCT did not have detectable phosphatase activity against any of these compounds

(Figure 3.7). Together, our results indicate that the recombinant NOCT catalytic domain

does not have ribonuclease activity on its own, and that NOCT is not active against DNA

or any of the alternative nucleotides or phosphorylated signalling molecules tested here.

Figure 3.8: NOCT does not display phosphatase activity towards various phosphorylated molecules. NOCT activity was measured against a panel of phosphate molecules (Table 3.1) including nucleotides, inositol phosphates, sphingomyelins, and sugar phosphates. Phosphatase activity was measured using a Malachite Green Phosphate Assay. Calf intestinal phosphatase was used as a positive control for a subset of substrates.

3.4 Discussion

Here, we report that recombinant NOCT does not have observable activity in ribonuclease, deoxyribonuclease, or phosphatase assays. The lack of exoribonuclease activity described in our *in vitro* assays seemingly contradicts previous reports that mNOCT and xNOCT could degrade poly(A) RNA substrates *in vitro* (Baggs and Green, 2003; Garbarino-Pico et al., 2007). These assays utilized recombinant GST fusions of either full length mNOCT or xNOCT that were isolated by a single step affinity purification scheme using glutathione resin and wash steps of moderate stringency (150 mM NaCl) (Baggs and Green, 2003; Garbarino-Pico et al., 2007). The stringency of this purification scheme is not likely enough to prevent co-purification of contaminating nuclease bacterial nucleases (Hrit et al., 2014). In addition, the mNOCT and xNOCT used in these analyses were not eluted from the resin used in purification, which is problematic in precise quantification of enzyme concentrations that enable comparison between samples (Baggs and Green, 2003; Garbarino-Pico et al., 2007). Furthermore, the lower molecular weight product described as the result of NOCT ribonuclease activity was only a small fraction of RNA relative to the full-length substrate, though the activity reported for xNOCT appears stronger than that for mNOCT. Mutant NOCT protein with a glutamate to alanine substitution in the active site at the Mg^{2+} binding residue was assayed as a negative control. This magnesium-binding residue is well described as a critical side chain for EEP deadenylase function, as analogous mutations in other family members abolish activity (Freudenthal et al., 2015; Wang et al., 2010). The structures of the NOCT catalytic domain confirm that this residue (E195 in the human enzyme) directly coordinates Mg^{2+} in the active site (Chapter 2). Reactions with mutant NOCT did not exhibit observable

accumulation of the lower molecular weight RNA, although TLC analysis of the mNOCT

reaction products indicate that AMP that would be released from exoribonuclease activity

is detectable in mNOCT E193A reactions (Garbarino-Pico et al., 2007). Therefore, the

methods by which NOCT was purified and assessed for deadenylase activity suggest that

variations in purity and/or protein concentration could explain the apparent nuclease

activity of NOCT. Due to these considerations, previous reports of NOCT deadenylase

activity are ambiguous but may be reporting on activity due to contamination rather than

bona fide NOCT activity (Baggs and Green, 2003; Garbarino-Pico et al., 2007).

In order to obtain soluble NOCT protein, we had to truncate the N-terminal 63

amino acids, which could feasibly have an effect on activity. It is noteworthy that NOCT

protein isoforms are annotated in the GenBank database that start at the equivalent

position to Met67. Moreover, $NOCT_{64-431}$, which contains the entire EEP domain, is

functional in cells (see Chapter 4), and similar truncations in other CCR4-family members,

such as CNOT6L, yeast Ccr4, and PDE12 remain catalytically active (Figure 3.5)

(Viswanathan et al., 2003; Wang et al., 2010; Wood et al., 2015). Therefore, it is unlikely

that the lack of observed NOCT ribonuclease activity is due to deletion of N-terminal

residues that are critical for activity. The lack of *in vitro* activity may be explained by the

lack of a protein partner, which may act through one or more mechanisms of action. One

possibility is that a protein partner may regulate NOCT activity by facilitating substrate

binding. Precedent for this idea is exemplified by the Pan2 deadenylase, which is inactive

without its RNA-binding protein partner, Pan3 (Brown et al., 1996; Wolf et al., 2014).

Second, NOCT itself may have an intrinsic ability to bind to and regulate substrate RNAs

but may be in an inactive conformation without the binding of a protein partner or cofactor

that allosterically activates the protein. One example of this type of regulation is the activation of RNase L, which multimerizes into an active state upon binding of 2,5-A to an allosteric site (Chakrabarti et al., 2011). In another example, the deadenylase PARN is activated by interaction with the 7mG mRNA cap (Virtanen et al., 2013).

NOCT may bind to mRNA substrates by recognizing specific sequence motifs or three-dimensional RNA structures. In this case, poly(A) RNA would not be the consensus recognition sequence for NOCT, differentiating NOCT further from other Ccr4-type deadenylases. CNOT6L and PDE12 ribonuclease activity against poly(A) RNA *in vitro* suggests that poly(A) sequence is sufficient for substrate recognition by other human deadenylases (Rorbach et al., 2011; Wang et al., 2010). Finally, it is conceivable that NOCT activity may require post-translational modifications, which would be absent in the recombinant protein expressed in bacteria. Determination of the requirements for NOCT activity may be better informed by identification of endogenous substrates, protein partners, and potential post-translational modifications that occur when NOCT is expressed in human cells.

3.5 Materials and Methods

Biochemical Assays.

Biochemical mRNA degradation reactions were performed in CNOT6L buffer (50 mM HEPES (4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid) pH 7.4, 150 mM NaCl, 10% glycerol, 2.0 mM $MgCl_2$, and 1.0 mM dithiothreitol (DTT)). $NOCT_{64-431}$, $CNOT6L_{158-555}$, or $CNOT6L_{158-555}$ E240A (1 µM) were mixed with 10 nM 5′ Cy3 labelled $p(A)_{20}$ (Dharmacon), and RNasin RNase A inhibitor (Promega) and incubated over a 30-minute time course at 37 °C before stopping the reactions in 95% formamide and 25mM EDTA

(Ethylenediaminetetraacetic acid). To titrate the ionic strength or Mg^{2+} concentrations, individual buffers were made with the indicated NaCl or $MgCl_2$ concentrations in the buffer described above. To test various divalent metals cations, $NOCT_{64-431}$ or $CNOT6L_{158-555}$ were treated with 15 mM EDTA to chelate Mg^{2+} from the enzyme and protein buffer before dialysis in assay buffer lacking $MgCl_2$. The divalent metal ions tested (Mg^{2+}, Mn^{2+}, Ca^{2+}, and Zn^{2+}) were added to assay buffer at a final concentration of 2.0 mM and combined with $NOCT_{64-431}$ or $CNOT6L_{158-555}$. To control for the effect of dialysis on activity, $CNOT6L_{158-555}$ activity was compared pre- and post-dialysis in buffer containing Mg^{2+}. To measure NOCT activity in a range of substrate RNA concentrations, 10 nM 5′ Cy3 labelled $p(A)_{20}$ (Dharmacon) was held constant and unlabelled $p(A)_{19}$ (Dharmacon) was titrated into reactions at concentrations ranging from 0 nM to 10 µM unlabelled RNA. Reaction products were resolved on a 15% polyacrylamide sequencing gel and detected using a Typhoon Trio + (GE Healthcare).

Additional biochemical assays were performed in the assay buffers described in Baggs and Green and in Garbarino-Pico *et al.* (Baggs and Green, 2003; Garbarino-Pico et al., 2007; Viswanathan et al., 2003). Briefly, mRNA degradation assays were performed in 1) xNOCT buffer (5 mM HEPES (pH 7.5), 2.0 mM $MgCl_2$, 0.5 mM DTT, 10% glycerol, 0.2 mg/mL bovine serum albumin (BSA), 0.02% (v/v) NP-40, and 0.3% (w/v) spermidine); and 2) mNOCT buffer (50 mM HEPES (pH 8.0), 0.01 mM $Mg(OAc)_2$, and 20 mM KOAc). $NOCT_{64-431}$, $mNOCT_{62-429}$, $CNOT6L_{158-555}$, or $CNOT6L_{158-555}$ E240A (1 µM) were mixed with 10 nM 5′ Cy3 labelled $p(A)_{20}$ (Dharmacon), and RNasin RNase A inhibitor (Promega) and incubated over a 30-minute time course at 37 °C. Reactions were stopped, resolved on a polyacrylamide gel, and imaged as described above.

To measure NOCT activity against a long mRNA substrate analog, the mRNA substrate template was generated by PCR using primers listed in Supplementary Table 2 and a plasmid containing the *Renilla* Luciferase ORF. The mRNA substrate was produced using the HiScribe T7 High Yield RNA Synthesis Kit (NEB) and purified using an RNA Clean and Concentrator -25 Kit (Zymo Research). Reactions were incubated with the long RNA substrate and RNasin at 37 °C for a 30-minute time course. As a positive control, RNase III (NEB) was incubated with the long substrate RNA in 1x NEB RNase III buffer at 37 °C for a 30-minute time course. Products were resolved on a 2.5% agarose/ 4% (v/v) formaldehyde/ 1x MOPS (200 mM 3-N-Morpholino propanesulfonic acid, 50 mM NaOAc, 10 mM EDTA) denaturing gel and detected by ethidium bromide (EtBr) staining and UV detection on a Syngene Gene Genius Bio Imaging System.

NOCT$_{64-431}$ activity against ssDNA and dsDNA was tested using 5´ FAM- labeled DNA 25-mers. The dsDNA substrate contained the ssDNA substrate forward strand sequence (5´-FAM- CGATAGCATCCTACCTTCTCTCCAT-3´) annealed to the reverse complement sequence (5´- ATGGAGAGAAGGTAGGATGCTATCG-3´). NOCT$_{64-431}$, Micrococcal Nuclease (NEB), or DNase I (Sigma) were incubated with ssDNA or dsDNA in assay buffer for 30 minutes and resolved on a polyacrylamide gel as described above.

Phosphatase activity assays against a panel of compounds (Table 2.3) were conducted using 500 nM NOCT$_{64-431}$ and 100 µM of each phosphorylated compound incubated in 50 mM HEPES (pH 7.4), 150 mM NaCl, 10% glycerol, 2.0 mM MgCl$_2$, and 1.0 mM DTT for 30 minutes, and the resulting concentrations of phosphate were measured using a Malachite Green Assay (Cayman Chemical) with reactions performed in duplicate. To account for differences in the rate of phosphate production via

spontaneous hydrolysis of each compound, each compound was tested in assay buffer lacking protein. Background phosphate production was then subtracted individually for each compound to account for non-enzymatic phosphate formation. Calf intestinal alkaline phosphatase (NEB) was used as a positive control.

Number:	Name:	Source:
1	ATP	NEB
2	CTP	NEB
3	GTP	NEB
4	TTP	Promega
5	UTP	NEB
6	AMP	Sigma
7	3', 5' ADP	Sigma
8	3' UMP	Cayman Chemical
9	ADP	Cayman Chemical
10	UDP	Cayman Chemical
11	GMP	Cayman Chemical
12	CMP	Sigma
13	IMP	Sigma
14	TMP	Sigma
15	GDP	Sigma
16	Cyclic AMP	Cayman Chemical
17	Cyclic GMP	Cayman Chemical
18	PtdIns-(5)-P1 (1,2-dihexanoyl)	Cayman Chemical
19	PtdIns-(4,5)-P2 (1,2-dipalmitoyl)	Cayman Chemical
20	D-myo-Inositol-1,4,5-triphosphate	Cayman Chemical
21	PtdIns-(3,4,5)-P3 (1,2-dihexanoyl)	Cayman Chemical
22	PtdIns-(3)-P1 (1,2-dioctanoyl)	Cayman Chemical
23	D-myo-Inositol-4,5-diphosphate	Cayman Chemical
24	PtdIns-(3,5)-P2 (1,2-dihexanoyl)	Cayman Chemical
25	D-myo-Inositol-2,4-diphosphate	Cayman Chemical
26	D-myo-Inositol-1,5,6-triphosphate	Cayman Chemical
27	D-myo-Inositol-2,3,5-triphosphate	Cayman Chemical
28	D-myo-Inositol-3,4,5-triphosphate	Cayman Chemical
29	D-myo-Inositol-1,4,6-triphosphate	Cayman Chemical
30	D-myo-Inositol-1,4-diphosphate	Cayman Chemical
31	D-myo-Inositol-1,3,4,6-tetraphosphate	Cayman Chemical
32	D-myo-Inositol-1,5-diphosphate	Cayman Chemical
33	D-myo-Inositol-1,2,4,5,6-pentaphosphate	Cayman Chemical
34	D-myo-Inositol-1,2,3,5-tetraphosphate	Cayman Chemical
35	D-myo-Inositol-1,2,4,5-tetraphosphate	Cayman Chemical
36	Palmitoyl Sphingomyelin	Cayman Chemical
37	Sphingosine-1-Phosphate	Cayman Chemical
38	Glucose 6 Phosphate	Cayman Chemical
39	Ribulose 5 Phosphate	Cayman Chemical
40	Mannose 6 Phosphate	Cayman Chemical
41	Fructose 6 Phosphate	Cayman Chemical
42	Fructose 1,6 bisphosphate	Cayman Chemical
43	Glyceraldehyde 3 Phosphate	Cayman Chemical
44	3 Phosphoglycerate	Cayman Chemical
45	Phosphoenol pyruvate	Cayman Chemical
46	Ribose 5 phosphate	Sigma
47	Polyphosphate	Calgon
48	B glycerophosphate	Sigma

Table 3.1: **Phosphorylated compounds used in screening for NOCT phosphatase activity.**

3.6 References

Abshire, E.T., Chasseur, J., Bohn, J.A., Del Rizzo, P.A., Freddolino, P.L., Goldstrohm, A.C., and Trievel, R.C. (2018). The structure of human Nocturnin reveals a conserved ribonuclease domain that represses target transcript translation and abundance in cells. Nucleic acids research *46*, 6257-6270.

Baggs, J.E., and Green, C.B. (2003). Nocturnin, a deadenylase in Xenopus laevis retina: a mechanism for posttranscriptional control of circadian-related mRNA. Current biology : CB *13*, 189-198.

Baggs, J.E., and Green, C.B. (2006). Functional analysis of nocturnin: a circadian clock-regulated gene identified by differential display. Methods in molecular biology *317*, 243-254.

Brown, C.E., Tarun, S.Z., Jr., Boeck, R., and Sachs, A.B. (1996). PAN3 encodes a subunit of the Pab1p-dependent poly(A) nuclease in Saccharomyces cerevisiae. Molecular and cellular biology *16*, 5744-5753.

Chakrabarti, A., Jha, B.K., and Silverman, R.H. (2011). New insights into the role of RNase L in innate immunity. J Interferon Cytokine Res *31*, 49-57.

Doidge, R., Mittal, S., Aslam, A., and Winkler, G.S. (2012). Deadenylation of cytoplasmic mRNA by the mammalian Ccr4-Not complex. Biochem Soc Trans *40*, 896-901.

Dyson, J.M., Fedele, C.G., Davies, E.M., Becanovic, J., and Mitchell, C.A. (2012). Phosphoinositide phosphatases: just as important as the kinases. Subcell Biochem *58*, 215-279.

Estrella, M.A., Du, J., and Korennykh, A. (2018). Crystal Structure of Human Nocturnin Catalytic Domain. Scientific reports *8*, 16294.

Freudenthal, B.D., Beard, W.A., Cuneo, M.J., Dyrkheeva, N.S., and Wilson, S.H. (2015). Capturing snapshots of APE1 processing DNA damage. Nature structural & molecular biology *22*, 924-931.

Garbarino-Pico, E., Niu, S., Rollag, M.D., Strayer, C.A., Besharse, J.C., and Green, C.B. (2007). Immediate early response of the circadian polyA ribonuclease nocturnin to two extracellular stimuli. Rna *13*, 745-755.

Green, C.B., and Besharse, J.C. (1996a). Identification of a novel vertebrate circadian clock-regulated gene encoding the protein nocturnin. Proceedings of the National Academy of Sciences of the United States of America *93*, 14884-14888.

Green, C.B., and Besharse, J.C. (1996b). Use of a high stringency differential display screen for identification of retinal mRNAs that are regulated by a circadian clock. Brain Res Mol Brain Res *37*, 157-165.

Green, C.B., and Besharse, J.C. (1997). Identification of vertebrate circadian clock-regulated genes by differential display. Methods in molecular biology *85*, 219-230.

Green, C.B., Douris, N., Kojima, S., Strayer, C.A., Fogerty, J., Lourim, D., Keller, S.R., and Besharse, J.C. (2007). Loss of Nocturnin, a circadian deadenylase, confers resistance to hepatic steatosis and diet-induced obesity. Proceedings of the National Academy of Sciences of the United States of America *104*, 9888-9893.

Hrit, J., Raynard, N., Van Etten, J., Sankar, K., Petterson, A., and Goldstrohm, A.C. (2014). In vitro analysis of RNA degradation catalyzed by deadenylase enzymes. Methods in molecular biology *1125*, 325-339.

Lowell, J.E., Rudner, D.Z., and Sachs, A.B. (1992). 3´ -UTR-dependent deadenylation by the yeast poly(A) nuclease. Genes Dev *6*, 2088-2099.

Maryati, M., Airhihen, B., and Winkler, G.S. (2015). The enzyme activities of Caf1 and Ccr4 are both required for deadenylation by the human Ccr4-Not nuclease module. Biochem J *469*, 169-176.

Mittal, S., Aslam, A., Doidge, R., Medica, R., and Winkler, G.S. (2011). The Ccr4a (CNOT6) and Ccr4b (CNOT6L) deadenylase subunits of the human Ccr4-Not complex contribute to the prevention of cell death and senescence. Mol Biol Cell *22*, 748-758.

Pearce, S.F., Rorbach, J., Van Haute, L., D'Souza, A.R., Rebelo-Guiomar, P., Powell, C.A., Brierley, I., Firth, A.E., and Minczuk, M. (2017). Maturation of selected human mitochondrial tRNAs requires deadenylation. Elife *6*.

Rorbach, J., Nicholls, T.J., and Minczuk, M. (2011). PDE12 removes mitochondrial RNA poly(A) tails and controls translation in human mitochondria. Nucleic acids research *39*, 7750-7763.

Schellenberg, M.J., Appel, C.D., Adhikari, S., Robertson, P.D., Ramsden, D.A., and Williams, R.S. (2012). Mechanism of repair of 5´-topoisomerase II-DNA adducts by mammalian tyrosyl-DNA phosphodiesterase 2. Nature structural & molecular biology *19*, 1363-1371.

Shiokawa, D., and Tanuma, S. (2001). Characterization of human DNase I family endonucleases and activation of DNase gamma during apoptosis. Biochemistry *40*, 143-152.

Stubblefield, J.J., Gao, P., Kilaru, G., Mukadam, B., Terrien, J., and Green, C.B. (2018). Temporal Control of Metabolic Amplitude by Nocturnin. Cell reports *22*, 1225-1235.

Virtanen, A., Henriksson, N., Nilsson, P., and Nissbeck, M. (2013). Poly(A)-specific ribonuclease (PARN): an allosterically regulated, processive and mRNA cap-interacting deadenylase. Crit Rev Biochem Mol Biol *48*, 192-209.

Viswanathan, P., Chen, J., Chiang, Y.C., and Denis, C.L. (2003). Identification of multiple RNA features that influence CCR4 deadenylation activity. The Journal of biological chemistry *278*, 14949-14955.

Wang, H., Morita, M., Yang, X., Suzuki, T., Yang, W., Wang, J., Ito, K., Wang, Q., Zhao, C., Bartlam, M*., et al.* (2010). Crystal structure of the human CNOT6L nuclease domain reveals strict poly(A) substrate specificity. The EMBO journal *29*, 2566-2576.

Wolf, J., Valkov, E., Allen, M.D., Meineke, B., Gordiyenko, Y., McLaughlin, S.H., Olsen, T.M., Robinson, C.V., Bycroft, M., Stewart, M*., et al.* (2014). Structural basis for Pan3 binding to Pan2 and its function in mRNA recruitment and deadenylation. The EMBO journal *33*, 1514-1526.

Wood, E.R., Bledsoe, R., Chai, J., Daka, P., Deng, H., Ding, Y., Harris-Gurley, S., Kryn, L.H., Nartey, E., Nichols, J*., et al.* (2015). The Role of Phosphodiesterase 12 (PDE12) as a Negative Regulator of the Innate Immune Response and the Discovery of Antiviral Inhibitors. The Journal of biological chemistry *290*, 19681-19696.

CHAPTER 4

NOCT Activity in Cell-based Reporter Assays [1]

4.1 Abstract

NOCT was first classified as an EEP deadenylase based on its homology to the founding member of the protein family, yeast Ccr4. We hypothesized that NOCT has an additional requirement for activity present in the cellular context but not in *in vitro* experiments. Based on this prediction, we turned to cell-based assays to determine if NOCT could repress mRNA expression when directed to an mRNA reporter in cells. Here, we describe the first evidence that human NOCT represses the expression of a reporter mRNA *in vivo* and in a manner dependent on conserved residues in the active site. This effect was stronger at the level of protein expression than at the level of mRNA abundance, suggesting that translational repression may play a role in the mechanism of NOCT-mediated repression. Furthermore, we determined that NOCT-mediated repression is dependent on the 3´ end of the mRNA and that NOCT does not recruit 5´-acting factors that contribute to mRNA decay and translational repression. Together, these data show that NOCT is able to cause mRNA decay and represses translation of an mRNA reporter.

---

[1] This work was contributed to by Dr. Peter Freddolino, who performed the statistical analysis of data from the tethered function assays and Northern blotting. This work was published in Abshire, E.T., Chasseur, J., Bohn, J.A., Del Rizzo, P.A., Freddolino, P.L., Goldstrohm, A.C., and Trievel, R.C. (2018). The structure of human Nocturnin reveals a conserved ribonuclease domain that represses target transcript translation and abundance in cells. Nucleic acids research *46*, 6257-6270.

4.2 Introduction

The majority of studies on NOCT function have been performed in knockout mice, which report that NOCT plays a role in regulating lipid metabolism through regulation of lipid trafficking, adipogenesis, and lipid synthesis (Douris et al., 2011; Green et al., 2007; Kawai et al., 2010; Stubblefield et al., 2018). These studies also uncovered a potential role for NOCT in negative regulation of osteogenesis, as the knockout mice have increased bone density and overexpression of NOCT in pre-osteoblasts inhibits differentiation (Green et al., 2007; Kawai et al., 2010).

The increased bone mass in NOCT KO mice suggests that NOCT may negatively regulate mRNAs that promote osteoblastogenic differentiation. Conversely, overexpression of NOCT in MC3T3-E1 osteoblasts impairs differentiation, and the mRNA levels of osteoblastogenic markers *Osteocalcin, Alkaline phosphatase,* and *Runx2* were decreased relative to GFP overexpression controls (Kawai et al., 2010). These mRNAs have the potential to be direct targets of NOCT but have not yet been tested as direct targets. To date, only one study has described the direct association of NOCT with an endogenous mRNA and has utilized reporter assays to examine regulation of a presumed target mRNA, the pro-osteoblastogenic factor *Igf1*. Igf1 is a circulating factor synthesized in the liver that acts as an endocrine factor to regulate bone accumulation in both mice and in humans (Elis et al., 2010; Langlois et al., 1998; Liu et al., 1993; Yakar et al., 2002). *Igf1* mRNA expression appears to fluctuate in a circadian manner antiphase to *Noct* mRNA expression, raising the possibility that NOCT's presumed deadenylase activity could explain expression pattern of *Igf1* transcripts (Green and Besharse, 1996; Green et al., 2007; Yakar et al., 2002).

Figure 4.1: Schematic of tethered function assays measuring the effect of NOCT directed to a reporter mRNA. A plasmid expressing MS2 coat protein fused to NOCT or HaloTag control is co-transfected with ffLuc and NanoLuc 4xMS2BS plasmids. The interaction between MS2 and the RNA stem loops targets NOCT to the reporter RNA and the resulting effect on reporter RNA expression is measured using a dual-glo luciferase assay.

As the mechanism by which NOCT recognizes its targets has not been identified, we utilized a tethered function assay to examine whether NOCT could repress expression from a reporter mRNA in cells. This approach is widely used in studying mRNA regulation and enables quantitation of NOCT's ability to regulate the expression of protein from a NanoLuciferase (NLuc) reporter mRNA in cells and the steady-state levels of the reporter mRNA (Figure 4.1) (Bos et al., 2016; Clement and Lykke-Andersen, 2008; Coller and Wickens, 2002; Van Etten et al., 2013). In these studies, we utilized a reporter construct that expresses an RNA with an ORF encoding the NLuc protein, followed by a 3′UTR bearing four binding sites for the MS2 coat protein followed by a cleavage and polyadenylation site (NLuc 4xMS2BS p(A)). The use of the MS2 RNA stem loops as the recognition element in these assays is advantageous as the mechanism by which NOCT recognizes its substrates is unknown and provides the binding interaction between NOCT and a reporter substrate. We attempted to control for the potential effects of endogenous NOCT by utilizing HEK293 cells, which express low levels of endogenous NOCT protein in Western blotting experiments (data not shown). Using these assays, we examined the effect of NOCT on mRNAs in the cellular context as well as the role of individual active site residues in

NOCT-mediated repression of mRNA expression, demonstrating that NOCT represses mRNA expression in cells. Additionally, we showed this repression is dependent on a subset of NOCT active site residues as well as the sequence and/or structure of the 3´ end of the mRNA.

4.3 Results

To characterize the effect of NOCT on a reporter mRNA in the cellular context, we measured luciferase expression from the NLuc 4xMS2BS p(A) reporter when co-transfected with MS2-NOCT in a tethered function assay. Tethered MS2-NOCT reproducibly repressed reporter RNA expression by 2.7 to 3.0-fold (Figure 4.2) (Abshire et al., 2018). A fusion of MS2-HT to the DEDD domain deadenylase CNOT7 (MS2-HT-CNOT7) also repressed the mRNA, consistent with its well-documented ability to inhibit translation and deadenylate mRNA (Aslam et al., 2009; Doidge et al., 2012; Viswanathan et al., 2004). This is the first evidence that human NOCT can repress gene expression when directed to an mRNA *in vivo*.
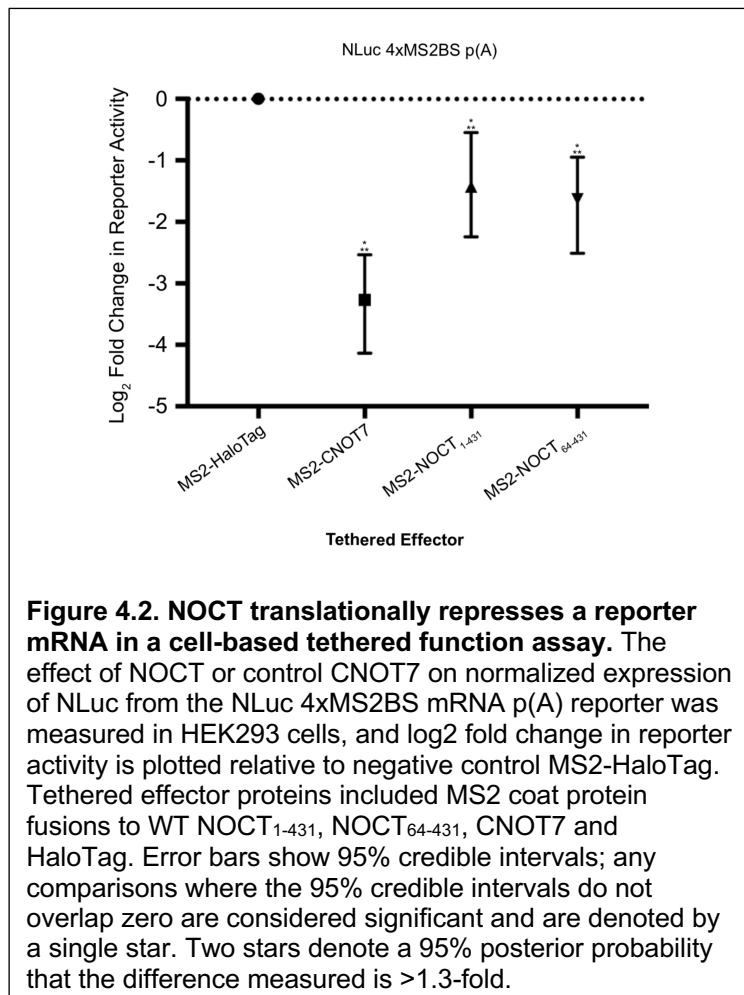
Figure 4.2. NOCT translationally represses a reporter mRNA in a cell-based tethered function assay. The effect of NOCT or control CNOT7 on normalized expression of NLuc from the NLuc 4xMS2BS mRNA p(A) reporter was measured in HEK293 cells, and log2 fold change in reporter activity is plotted relative to negative control MS2-HaloTag. Tethered effector proteins included MS2 coat protein fusions to WT $NOCT_{1-431}$, $NOCT_{64-431}$, CNOT7 and HaloTag. Error bars show 95% credible intervals; any comparisons where the 95% credible intervals do not overlap zero are considered significant and are denoted by a single star. Two stars denote a 95% posterior probability that the difference measured is >1.3-fold.

We then compared the activities of the truncated MS2-NOCT$_{64-431}$ to MS2-NOCT$_{1-431}$ in the tethered function assay. This comparison was used to ensure that the lack of activity observed in the *in vitro* assays previously described in Chapter 2 was not due to truncation of the NOCT N-terminus. Repression of reporter expression by NOCT$_{64-431}$ and NOCT$_{1-431}$ was comparable, indicating that the truncation of the N-terminus does not negatively affect NOCT activity in tethered function assays (Figure 4.2).

The use of the tethered function assay allows us to interrogate the role of the NOCT active site residues in a cellular context, where NOCT activity is observable. We therefore individually mutated Glu195, Asn149, His286, Asp324, Asn326, Asp377, and His414 to alanine in the MS2-NOCT construct. Additionally, we examined the effect of mutating His414 to asparagine, which is a more conservative mutation more likely to preserve hydrogen-bonding interactions of His414 within the active site (Figure 2.8). Mutation of His414 to either alanine or asparagine (H414A and H414N) reduced reporter repression, which indicated that the His414 side chain itself serves an important function. N149A and D377A also reduced NOCT-mediated repression of the reporter (Figure 4.3). In contrast, the H286A, E324A, and N326A substitutions did not appreciably affect NOCT-mediated repression. Strikingly, the E195A mutation slightly enhanced repression activity, which is the opposite of what would be expected from mutation of the residue that coordinates a catalytic Mg^{2+} ion (Figure 2.8). This mutation is broadly described as an inactivating mutant in EEP superfamily enzymes in general, including Ccr4, PDE12, APE1, and TDP2, though this effect has largely been shown in *in vitro* experiments (Chen et al., 2002; Freudenthal et al., 2015; Rorbach et al., 2011; Schellenberg et al., 2012).
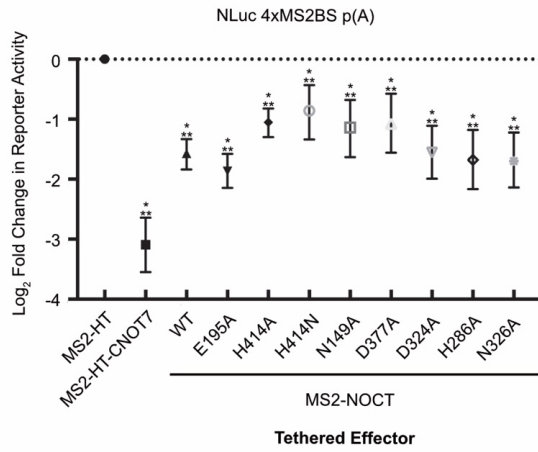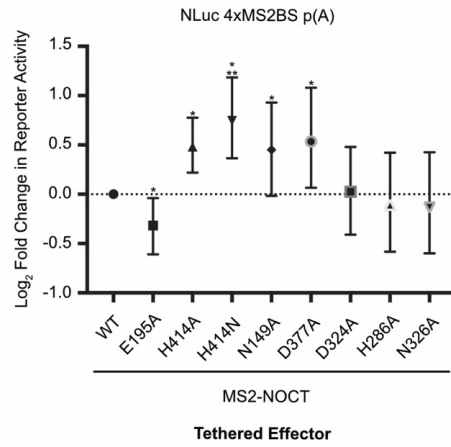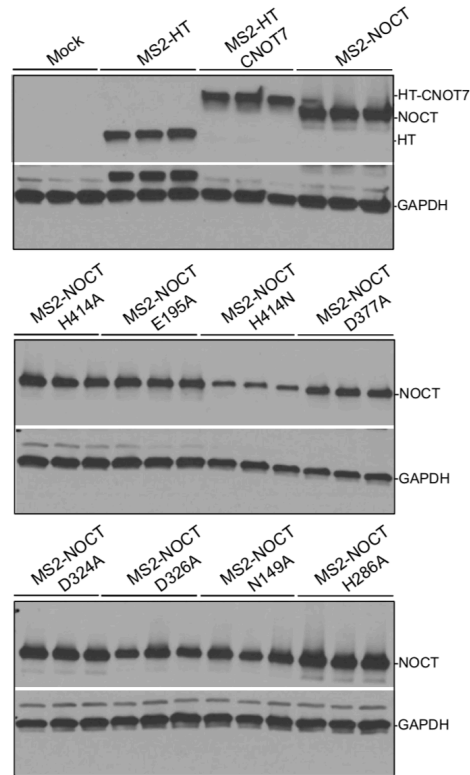
Figure 4.3: NOCT repression of protein expression depends on conserved active site residues. (A) Mutation of NOCT the active site reduced but did not eliminate translational repression of the NLuc 4xMS2BS p(A) reporter. Tethered effectors fused to MS2 coat protein are listed on the bottom of the graph. Log$_2$ fold change in reporter protein expression calculated relative to negative control MS2-HaloTag (MS2-HT) is plotted for each tethered effector. Error bars and significance stars are as in Figure 4.2. **(B)** As in (A), but with activity shown relative to WT NOCT. **(C)** Western blot detection of tethered effectors for three replicate samples using anti-V5 monoclonal antibody (WT and mutant MS2-NOCT effectors) or anti-HaloTag monoclonal antibody (MS2-HT and MS2-HT-CNOT7).

To ensure that the measured differences in reporter repression were not due to differential expression, we examined the steady state protein levels of the overexpressed NOCT catalytic mutants as compared to WT NOCT by Western blotting. This showed that differences in the ability of NOCT to repress reporter mRNAs following mutagenesis the active site cannot be attributed to differences in expression of the NOCT fusion proteins. All constructs were well-expressed, and slight differences among expression levels of the mutant proteins do not correlate with differences in reporter expression (Figure 4.3). This also serves as confirmation that the mutations in the NOCT active site are not causing significant differences in protein stability to the extent that the active site mutants are significantly affected relative to wild-type NOCT. Together, our observations indicate that NOCT possesses the ability to repress protein expression when directed to a specific mRNA, and that this repressive activity is dependent at least in part on specific conserved active site residues.

These tethered function assays were performed in HEK293 cells, as we have detected low levels of NOCT protein expression in this cell line, we elected to use it in order to minimize the potential effects of endogenous NOCT, especially in experiments using catalytic mutants of NOCT. HEK293 cells have a neuroadrenal phenotype, which is not necessarily a biologically relevant cell line with respect to NOCT expression (Lin et al., 2014; Shaw et al., 2002). Analysis of RNA expression in various human tissues indicated that NOCT is most highly expressed in adipose tissue and liver, as well as in lung and mammary tissues and transformed fibroblast and lymphocyte cell lines. Expression in brain tissue and adrenal glands were not among the tissues with the highest expression of NOCT mRNA (Figure 4.4) (Carithers and Moore, 2015; Consortium, 2013,

Figure 4.4: Expression of NOCT mRNA across different tissue types. The plot of mRNA expression data was obtained using the GTex database. Expression of *NOCT* mRNA is plotted as log10 transcripts per million (TPM) for each tissue type. *NOCT* expression is highest in adipose, liver, and mammary tissues. Plot is generated by GTEx release v6p (http://gtexportal.org).
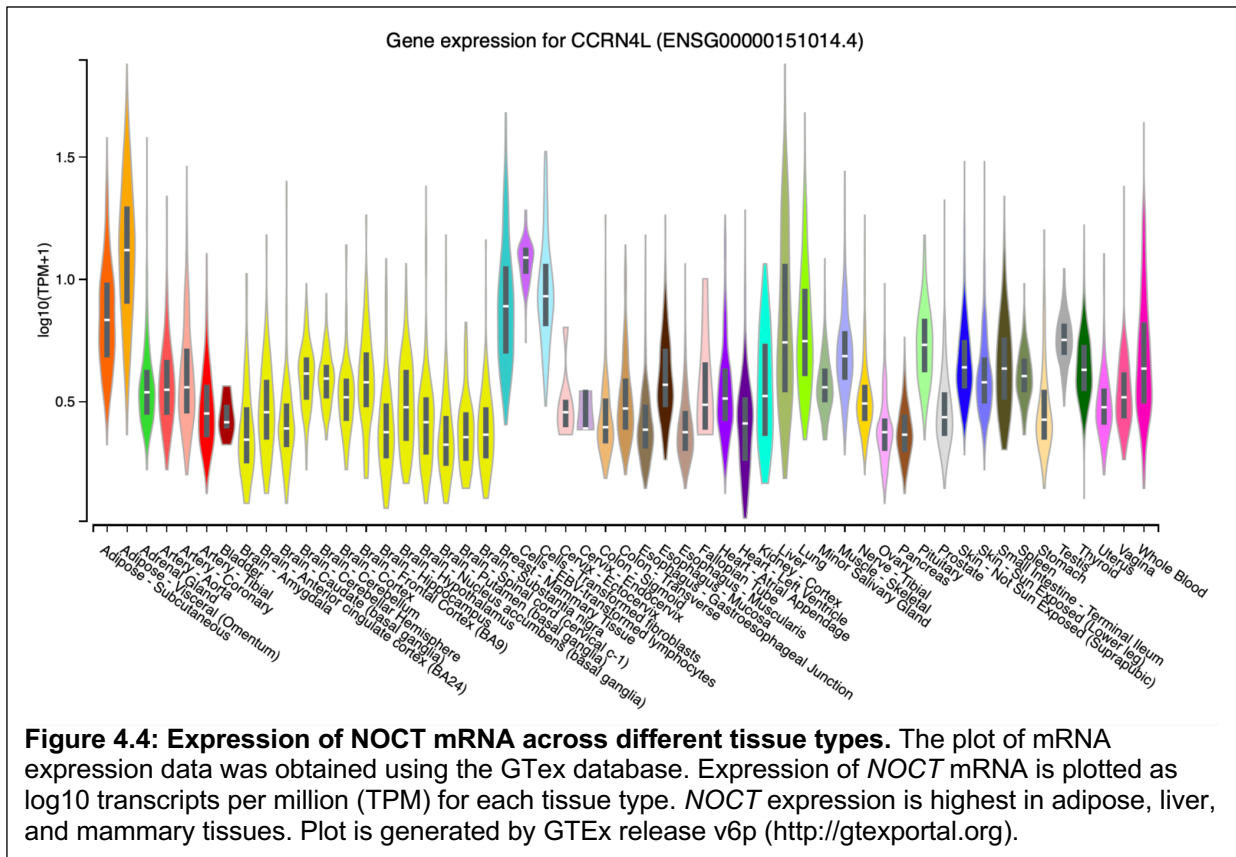
2015; Keen and Moore, 2015). While NOCT has observed activity in HEK293 cells, we could not rule out that the behaviour of NOCT constructs would differ in cell lines that correspond to tissues that express higher levels of endogenous NOCT and may also express unidentified factors that affect NOCT activity. For this experiment we turned to HepG2 human hepatocarcinoma cells, which express higher levels of endogenous NOCT and are more representative of liver tissue, which has significant expression of endogenous *Noct* mRNA in mice (Green et al., 2007; Wang et al., 2001). Additionally, we have detected significantly more NOCT in HepG2 cells than in HEK293 cells (discussed in Chapter 5, Figure 5.2). Therefore, we repeated the tethered function assay in HepG2 cells with WT MS2-NOCT and the H414A mutant and observed that NOCT has the same effect on mRNA reporters as it does in the assay performed in HEK293 cells (Figure 4.5).

Together, these data show that NOCT reproducibly represses mRNA reporter expression

in another cell line.

Figure 4.5: NOCT repression of reporter RNAs displays the same trend of repressive activity in HepG2 cells as in HEK293 cells. (A) The effect of NOCT or control CNOT7 on normalized expression of NLuc from the NLuc 4xMS2BS mRNA p(A) reporter was measured in HepG2 cells, and log2 fold change in reporter activity is plotted relative to negative control MS2-HaloTag. Tethered effector proteins included MS2 coat protein fusions to WT NOCT1-431, NOCT64-431, CNOT7 and HaloTag. Error bars and significance stars are as in Figure 4.2. **(B)** Western blot detection of tethered effectors using anti-V5 monoclonal antibody (WT and mutant MS2-NOCT effectors) or anti-HaloTag monoclonal antibody (MS2-HT and MS2-HT-CNOT7). Endogenous levels of NOCT in HepG2 cells were detected using an anti-NOCT polyclonal antibody to assess if overexpression of NOCT tethered effectors alters levels of endogenous NOCT protein. Detection of GAPDH with an anti-GAPDH monoclonal antibody is used to assess differences in loading.

Given the relationship of NOCT to CCR4-type 3′ exoribonucleases, we next

explored the effect of the 3′ end of the reporter mRNA on repression. The CCR4-NOT

complex is known to act on the 3′ end of mRNAs through its deadenylase subunits, but

also recruits factors that repress translation and promote 5′ directed decay (Alhusaini and

Coller, 2016; Behm-Ansmant et al., 2006; Cooke et al., 2010). If this were true for NOCT

we would expect tethered NOCT to recruit these factors, which would contribute to

repression. To facilitate this analysis, we compared NOCT-mediated repression of a NLuc

4xMS2BS reporter terminating with a 3′ poly(A) tail to a similar reporter terminating with

Figure 4.6: NOCT repression of protein expression depends on the identity of the mRNA 3′ end. (A) The reporters used in 3′ end analysis was derived from the original pNLP 4xMS2BS reporter with a 3′ cleavage and polyadenylation signal, allowing for 3′ end processing. The MALAT1 reporter contained the MALAT1 stable triple helical reporter in place of the cleavage and polyadenylation signal. **(B)** Tethered function assays (as in Figure 4.2) comparing repression of NLuc targets with the indicated 3′ ends by MS2-tethered effectors relative to that of MS2-HT. Error bars and significance stars are as in Figure 4.2.

a 3′ MALAT1 triple-helical structure (Figure 4.6). The MALAT1 3′ end is processed by RNase P and does not undergo canonical cleavage and polyadenylation (Wilusz, 2016). Instead, a short, template-encoded poly(A) sequence is incorporated into a triple helix structure (Brown et al., 2014). As expected, WT NOCT repressed the poly(A) reporter, as did the positive control CNOT7. The NOCT H414A mutant exhibited reduced activity towards the poly(A) reporter compared to the WT enzyme. Strikingly, NOCT did not have any substantial activity towards the MALAT1 reporter, suggesting that

the mechanism of repression is dependent on either poly(A) and/or the accessibility of the 3′ end (Figure 4.6). Tethered CNOT7 retained partial inhibitory function on the MALAT1 reporter, consistent with the documented ability of the CNOT complex to cause translational repression and promote 5′ decapping and decay (Alhusaini and Coller, 2016; Behm-Ansmant et al., 2006; Rouya et al., 2014). Taken together, these observations

indicate that NOCT acts at the 3´ end of the mRNA, consistent with CCR4-type exoribonucleases, and unlike CNOT7, does not appear to recruit 5´-directed decay factors or endonucleases to target mRNAs.

Figure 4.7: WT NOCT degrades poly(A) RNA and mutating active site residues stabilizes NLuc RNA. (A) Fusion proteins of MS2 coat proteins with WT NOCT or NOCT active site mutants were transfected into HEK293 cells along with the NLuc 4xMS2 poly(A) reporter. Resulting levels of NLuc RNA relative to the MS2-HT control were measured by Northern blot analysis. WT NOCT reduced the steady state levels of the NLuc reporter RNA. Mutating the NOCT active site residues to alanine appeared to reduce degradation; however, the width of the credible intervals precludes a firm conclusion as to whether or not there is residual activity relative to HT. **(B)** As in panel (A), but with data shown relative to WT NOCT. The data suggest that all tested mutations reduce activity relative to WT, although only the H414A and N326A mutations showed strong evidence of loss of activity. **(C)** Representative Northern blot of NLuc and Firefly Luciferase mRNAs from tethered function assays. RNA levels in the presence of WT NOCT and NOCT mutants with polyadenylated substrates.

We subsequently analysed the effect of NOCT on NLuc 4xMS2BS p(A) reporter mRNA levels using Northern blot analysis to determine whether NOCT-mediated repression occurs through inhibition of translation, enhanced RNA decay, or a

combination of both mechanisms. Relative to the MS2-HT control, tethered MS2-NOCT reduced the reporter mRNA levels by 2.2-fold (Figure 4.7).

In contrast, based on our statistical criteria, the NOCT active site mutants did not discernibly reduce reporter mRNA levels relative to the MS2-HT negative control, suggesting that the mutations compromised their ability to degrade the mRNA. When compared to WT NOCT, loss of activity could only be firmly established for the H414A mutant (Figure 4.7). The other active site mutants showed some evidence of impaired activity relative to WT NOCT, but measurement noise precludes a firm conclusion on their relative activity compared with both HT and WT NOCT. Nevertheless, these data (particularly regarding the H414A mutant) indicate that the catalytic site of NOCT contributes to RNA decay *in vivo*.
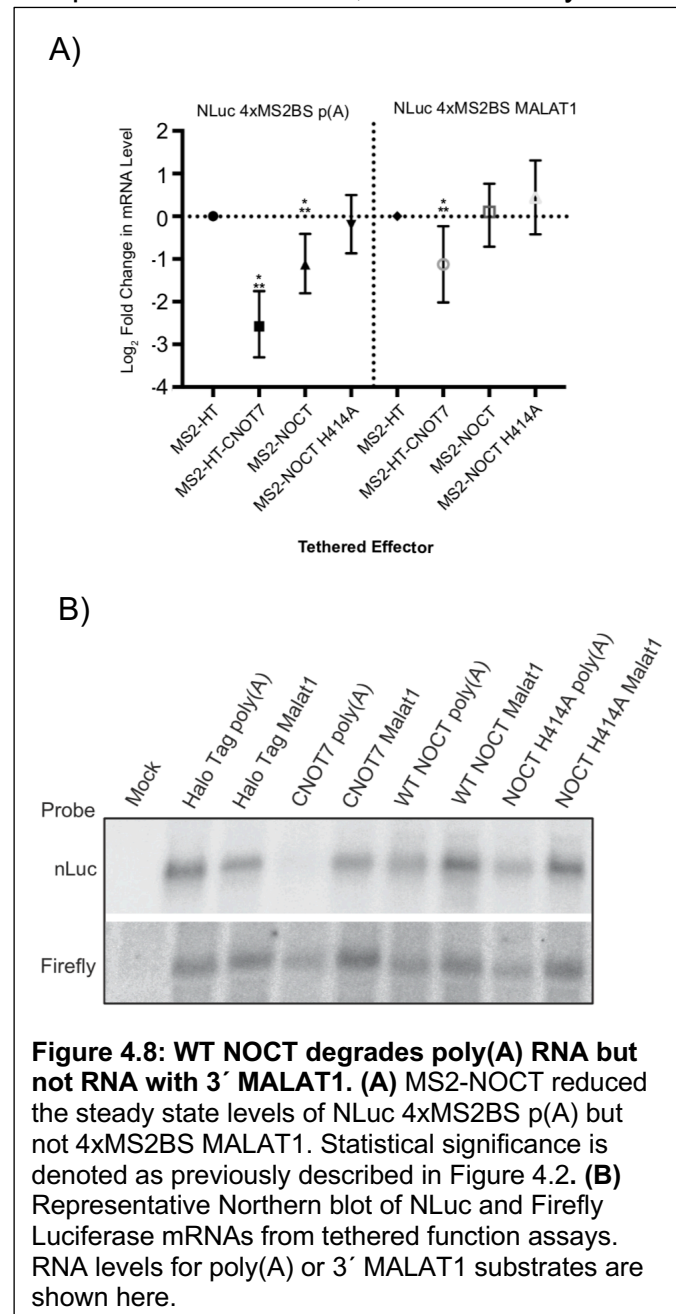
We also tested the ability of tethered NOCT to reduce levels of the NLuc 4xMS2BS 3′ MALAT1 mRNA and observed that NOCT was incapable of degrading the RNA, consistent with its inability to translationally repress the MALAT1 reporter (Figure 4.8). The

Figure 4.8: WT NOCT degrades poly(A) RNA but not RNA with 3′ MALAT1. (A) MS2-NOCT reduced the steady state levels of NLuc 4xMS2BS p(A) but not 4xMS2BS MALAT1. Statistical significance is denoted as previously described in Figure 4.2. **(B)** Representative Northern blot of NLuc and Firefly Luciferase mRNAs from tethered function assays. RNA levels for poly(A) or 3′ MALAT1 substrates are shown here.

positive control MS2-HT-CNOT7 was capable of degrading the MALAT1 mRNA, albeit with reduced effectiveness compared to the poly(A) mRNA, consistent with CNOT-mediated RNA degradation and 5′-directed decay. These observations indicate that NOCT-mediated repression involves mRNA degradation and is dependent on either the sequence and/or structure of the 3′ end.

4.4 Data Tables

Log2 Difference in Reporter Expression

Tethered Effector	Log 2 diff.	CI Low	CI High	P (diff)	P (within +/- 50% HT)	N=
Halo Tag	0	0	0	0	1	3
CNOT7	-3.268007	-4.134348	-2.536429	1	0	3
NOCT 1-431	-1.428992	-2.243524	-0.547318	0.989167	0.027500	3
NOCT 64-431	-1.632109	-2.509334	-0.946863	0.998333	0.008333	3

Log 2 diff. = Log2 difference in mean activity of the tethered effector, relative to the effector used in normalization
CI Low = the lower limit of the 95% credible interval
CI High = the upper limit of the 95% credible interval
P (diff) = the posterior probability that the activity is significantly different from the activity of the effector used for normalization
P (within +/- 50% HT) = the posterior probability that the activity is within +/- 50% of Halo Tag activity

Table 4.1: Data analysis and statistics for Figure 4.2

A)

Log2 Difference in Reporter Expression

Tethered Effector	Log 2 diff.	CI Low	CI High	P (>1.3 fold)	P (diff)	N=
Halo Tag	0	0	0	0	0	13
CNOT7	-3.092478	-3.549057	-2.641468	1	1	13
WT NOCT	-1.567695	-1.838397	-1.330357	1	1	13
E195A	-1.867749	-2.146525	-1.577590	1	1	3
H414A	-1.052285	-1.297275	-0.820126	1	1	13
H414N	-0.857976	-1.337804	-0.432482	0.973889	0.998889	3
N149A	-1.138216	-1.632248	-0.679595	0.996111	0.999444	3
D377A	-1.074130	-1.558051	-0.575742	0.995556	1	3
D324A	-1.554471	-1.990244	-1.108768	1	1	3
H286A	-1.678415	-2.166031	-1.177155	1	1	3
N326A	-1.696757	-2.139532	-1.220465	0.999444	1	3

B)

Log2 Difference in Reporter Expression

Tethered Effector	Log 2 diff.	CI Low	CI High	P (>1.3 fold)	P (diff)	N=
Halo Tag	1.551282	1.312687	1.783630	1	1	13
CNOT7	-1.526692	-1.996743	-1.034947	1	1	13
WT NOCT	0	0	0	0	0	13
E195A	-0.318382	-0.610368	-0.038701	0.332778	0.982222	3
H414A	0.482661	0.218778	0.776423	0.781111	0.998889	13
H414N	0.744425	0.365347	1.184073	0.966111	0.997222	3
N149A	0.450435	-0.014844	0.930759	0.609444	0.970556	3
D377A	0.535362	0.065910	1.079517	0.751111	0.983333	3
D324A	0.025472	-0.409087	0.479551	0.052778	0.558333	3
H286A	-0.107071	-0.581834	0.421675	0.134444	0.666111	3
N326A	-0.1344603	-0.5997928	0.4264901	0.140000	0.728889	3

Log 2 diff. = Log2 difference in mean activity of the tethered effector, relative to the effector used in normalization
CI Low = the lower limit of the 95% credible interval
CI High = the upper limit of the 95% credible interval
P (>1.3 fold) = the posterior probability that the difference observed is >1.3 fold vs the effector used for normalization
P (diff) = the posterior probability that the difference observed is signifcant vs the effector used for normalization

Table 4.2: Data analysis and statistics for Figure 4.3

Log2 Difference in Reporter Expression

Tethered Effector	Log 2 diff.	CI Low	CI High	P (>1.3 fold)	P (diff)	N=
Halo Tag, pA	0	0	0	0	0	13
CNOT7, pA	-3.092478	-3.549057	-2.641468	1	1	13
WT NOCT, pA	-1.567695	-1.838397	-1.330357	1	1	13
H414A, pA	-1.052285	-1.297275	-0.820126	1	1	13
Halo Tag, Malat1	0	0	0	0	0	10
CNOT7, Malat1	-1.356700	-1.851665	-0.841266	0.999444	1	10
WT NOCT, Malat1	-0.212492	-0.582603	0.199218	0.166667	0.880556	10
H414A, Malat1	-0.291699	-0.668252	0.084193	0.084193	0.939444	10

Log 2 diff. = Log2 difference in mean activity of the tethered effector, relative to the effector used in normalization
CI Low = the lower limit of the 95% credible interval
CI High = the upper limit of the 95% credible interval
P (>1.3 fold) = the posterior probability that the difference observed is >1.3 fold vs the effector used for normalization
P (diff) = the posterior probability that the difference observed is signifcant vs the effector used for normalization

Table 4.3: Data analysis and statistics for Figure 4.6

A)

Log2 Difference in mRNA Level

Tethered Effector	Log 2 diff.	CI Low	CI High	P (>1.3 fold)	P (insig)
Halo Tag	0	0	0	0	1
CNOT7	-2.578428	-3.303737	-1.754760	1	0
WT NOCT	-1.127761	-1.800538	-0.414883	0.982333	0.057667
E195A	-0.745265	-1.739478	0.420948	0.750667	0.375000
H414A	-0.204617	-0.866746	0.499182	0.300333	0.898158
H414N	-0.458411	-1.549117	0.508112	0.559000	0.592333
N149A	-0.403574	-1.483374	0.547570	0.512667	0.623000
D377A	-0.477115	-1.503212	0.545772	0.576333	0.565000
D324A	-0.705992	-1.786779	0.252236	0.743333	0.396667
H286A	-0.539628	-1.573578	0.500926	0.615333	0.526667
N326A	-0.122037	-1.181752	0.829487	0.294333	0.748000

B)

Log2 Difference in mRNA Level

Tethered Effector	Log 2 diff.	CI Low	CI High	P (>1.3 fold)	P (insig)	N
Halo Tag	1.127761	0.414883	1.800538	0.982333	0.039079	8
CNOT7	-1.450667	-2.249169	-0.662008	0.995333	0.018000	8
WT NOCT	0	0	0	0	1	8
E195A	0.382496	-0.698603	1.478946	0.497333	0.611333	3
H414A	0.923144	0.254750	1.617683	0.946000	0.156000	8
H414N	0.669350	-0.325970	1.684291	0.722000	0.410000	3
N149A	0.724187	-0.330081	1.741319	0.755667	0.313158	3
D377A	0.650647	-0.340460	1.674736	0.709667	0.438667	3
D324A	0.421769	-0.649930	1.485722	0.539667	0.588947	3
H286A	0.616744	-0.367767	1.559078	0.694605	0.593000	3
N326A	1.005724	-0.042730	2.017179	0.890333	0.194333	3

Log 2 diff. = Log2 difference in mean activity of the tethered effector, relative to the effector used in normalization
CI Low = the lower limit of the 95% credible interval
CI High = the upper limit of the 95% credible interval
P (>1.3 fold) = the posterior probability that the difference observed is >1.3 fold vs the effector used for normalizatio
P (insig) = the posterior probability that the difference observed is insignificant vs the effector used for normalization

Table 4.4: Data analysis and statistics for Figure 4.7

Log2 Difference in mRNA Level

Tethered Effector	Log 2 diff.	CI Low	CI High	P (>1.3 fold)	P (insig)	N=
Halo Tag, pA	0	0	0	0	1	8
CNOT7, pA	-2.578428	-3.303737	-1.754760	1	0	8
WT NOCT, pA	-1.127761	-1.800538	-0.414883	0.982333	0.057667	8
H414A, pA	-0.204617	-0.866746	0.499182	0.300333	0.898158	8
Halo Tag, Malat1	0	0	0	0	1	6
CNOT7, Malat1	-1.128485	-2.019830	-0.230167	0.955000	0.109211	6
WT NOCT, Malat1	0.100366	-0.714934	0.764189	0.222237	0.878947	6
H414A, Malat1	0.442343	-0.422298	1.309356	0.569473	0.625526	6

Log 2 diff. = Log2 difference in mean activity of the tethered effector, relative to the effector used in normalization
CI Low = the lower limit of the 95% credible interval
CI High = the upper limit of the 95% credible interval
P (>1.3 fold) = the posterior probability that the difference observed is >1.3 fold vs the effector used for normalization
P (insig) = the posterior probability that the difference observed is insignificant vs the effector used for normalization

Table 4.5: Data analysis and statistics for Figure 4.8

4.5 Discussion

The results presented here represent a major advance in the field of NOCT biology; we demonstrate that NOCT can translationally repress and degrade an mRNA when directed to that transcript in vivo, providing the first evidence of NOCT-mediated repression of mRNAs in human cells. We additionally tested the $NOCT_{64-431}$ construct utilized in biochemical studies and compared its activity to $NOCT_{1-431}$, which was not able to be assayed in *in vitro* assays (discussed in Chapters 2 and 3). While this result indicates that the first 63 amino acids of NOCT are dispensable for activity in the tethered function assay, it is still possible that these constructs may differ with respect to recognition and binding of endogenous target mRNAs.

The structural homology between NOCT and CNOT6L (discussed in Chapter 2) suggests a role for NOCT in RNA degradation, which is supported by the *in vivo* data

presented here, demonstrating that NOCT reduces the levels of reporter mRNAs in a manner dependent on key conserved active site residues (Figures 4.3 and 4.7). A subset of these mutants behaves as expected based on homology with the other EEP deadenylases—mutational analysis of conserved actives site residues in CNOT6L demonstrated that E240A, D489A, and H529A mutations abolished deadenylase activity *in vitro*. Similarly, the D377A and H414A (and H414N) mutants of NOCT had reduced activity in the tethered function assays, though the CNOT6L and NOCT mutant activities were measured in very different contexts (Figures 4.3). Additionally, the NOCT H414A mutant reduced mRNA degradation in vivo relative to the WT enzyme, illustrating the catalytic defects associated with this mutant (Figure 4.7).

Contrary to expectations, mutagenesis of NOCT Glu195 with alanine does not significantly alter mRNA repression or degradation in vivo (Figures 4.3 and 4.7), in contrast to the corresponding E240A mutant in CNOT6L. It is unknown why the E195A mutation in NOCT does not have an effect in tethered function assays and may reflect binding of this mutant to the reporter in a manner that inhibits translation of the transcript. This is one mechanism by which a catalytically inactive form of NOCT could potentially still repress reporter expression. In Northern blotting experiments the effect of E195A mutant on mRNA steady states was not able to be determined. Further studies are needed in order to determine if mRNA decay is abolished with the E195A mutant, and techniques such as transcription shutoff with Actinomycin D and half-life measurements may be more informative than the steady-state measurements used here.

NOCT repression activity is also dependent on the sequence and/or structure of the 3´ end of the mRNA, as introduction of the MALAT1 structure blocks NOCT-mediated

decay and translational repression (Figures 4.6 and 4.8). If NOCT facilitated decay through recruitment of endonucleases, or if NOCT recruiting 5´ acting factors such as those that promote decapping, then a 3´ MALAT1 structure would not be expected to completely prevent RNA degradation and repression. While the structure and/or accessibility of the 3´ end is crucial, the potential for 3´ end sequence specificity of NOCT remains to be determined, as *in vitro* studies have not been able to demonstrate that NOCT recognizes poly(A) RNA. Both the 3´ poly(A) and 3´ MALAT1 reporter mRNAs terminate in adenosine residues, but the 3´ end of MALAT1 is sequestered within a stable triple helix, suggesting that the MALAT1 structure prevents NOCT from accessing the 3´ end of the reporter (Brown et al., 2014).

In addition to observing repression of reports at the protein levels, we also turned to Northern blotting, which can be used to both measure relative RNA abundances and may resolve decay intermediates. No decay intermediates were observed, such as a shorter reporter mRNA that would correspond to a deadenylated mRNA. We cannot rule out that NOCT works in concert with other ribonucleases to promote mRNA decay. For instance, NOCT may deadenylate mRNA targets before another enzyme, such as the exosome, degrades the remainder of the RNA message (Zinder and Lima, 2017). In this case, NOCT may have a functional interaction with the canonical decay pathways in which deadenylated mRNAs are further degraded by either the 3´ to 5´ or 5´ to 3´ decay pathways (Goldstrohm and Wickens, 2008). In these cases, decay would likely be too rapid for intermediates to be resolved in our assays. This hypothesis remains untested but could explain why mutagenesis of the NOCT active site did not completely ablate

repression of the reporter mRNAs, in the case that NOCT can recruit downstream decay factors, even when the active site is mutated.

Based on these observations discussed here, which shows that NOCT represses expression and reduces steady-state levels of mRNA reporters in cell-based assays, and those in Chapter 2, in which NOCT had no detectable activity *in vitro*, we propose that NOCT requires one or more unidentified protein partners to be active. Such a partner may facilitate substrate binding, such as an RNA-binding protein that would direct NOCT to target mRNAs, thereby regulating its substrate specificity. Precedent for this idea is exemplified by the Pan2 deadenylase, which is inactive without its RNA-binding protein partner, Pan3 (Brown et al., 1996; Wolf et al., 2014). The tethered function assay used in this study may bypass the necessity of such a partner, but such a partner may also be necessary for activation of NOCT through mechanisms such as allosteric activation. Additionally, it cannot be ruled out that NOCT binds to mRNAs by recognizing specific sequence motifs or three-dimensional RNA structures, which were missing from the substrates tested *in vitro*. Finally, it is conceivable that NOCT activity may require post-translational modifications, which would be absent in the recombinant protein expressed in bacteria.

Our data suggest that NOCT may function in both translational repression and mRNA decay, functioning similarly to other described deadenylases. First, the H414A mutation diminishes NOCT-mediated decay of poly(A) reporter mRNA, indicating that its exoribonuclease function is important, yet this same mutant retains some translational repression activity (Figures 4.3 and 4.7). This suggests that a combination of exonuclease activity and an additional mechanism involving translational repression is responsible for

NOCT-mediated repression of polyadenylated target mRNAs. At this time, we cannot definitively discriminate between the contributions of these mechanisms. Second, NOCT-mediated mRNA decay and translational repression activities are abrogated on the MALAT1 reporter, suggesting that both activities of NOCT are dependent on its ability to antagonize 3′ poly(A) and poly(A) binding protein function.

A crucial challenge is to identify the natural mRNAs that are bound and regulated by NOCT. Our findings indicate that NOCT-regulated mRNAs are unlikely to accumulate as deadenylated intermediates, which could explain difficulties in identifying direct NOCT targets by surveying increased poly(A) tail length in NOCT-deficient mice (Kojima et al., 2015; Stubblefield et al., 2018). NOCT repression is additionally manifested in reduced protein expression from its target mRNA; therefore, quantitative proteomics approaches may also facilitate identification of its regulatory impact on natural target mRNAs. Such approaches, conducted in physiologically relevant tissues and cell types, will be necessary to discover the post-transcriptional regulatory networks that underlie NOCT's important physiological functions.

4.6 Materials and Methods

Cell-based Assays.

HEK293 cells (ATCC) were cultured in DMEM with glucose, 10% (v/v) FBS, 100 U/mL penicillin, 100U/mL streptomycin, and 292 µg/mL glutamine (penicillin/streptomycin/glutamine from Gibco), at 37 °C and 5% CO_2. The full-length NOCT coding region was cloned into the pF5K vector (Promega) to express MS2-NOCT in HEK293 cells. All protein constructs contained a V5 epitope tag for western blot detection. As a negative control, HaloTag (HT) fused to MS2 coat protein in the pFN21A

vector was used (Promega). Point mutations were generated using the oligonucleotide mediated site-directed mutagenesis method (Quikchange, Agilent). The reporter plasmid, pNLP 4xMS2BS p(A), was derived from the plasmid psiCheck1 (Promega) and contained the SV40 promoter driving NanoLuc Luciferase (NLuc) with a 3´ UTR containing 4 stem-loop binding sites for phage MS2 coat protein. The NLuc 4xMS2BS p(A) and NLuc 4xMS2BS MALAT1 reporters contained either a minimal synthetic 3´ cleavage and polyadenylation signal or a 3´ MALAT1 triple helix, respectively. The Firefly Luciferase internal control plasmid was pGL4.13 (Promega).

For transfections, cells were plated at 600,000 cells per well in 6-well culturing plates. 24-hours post plating, DNA was transfected into cells using a 3µL:1ng ratio of ViaFect (Promega) to DNA including 1.25 µg of effector plasmid, 500 ng of Firefly plasmid, and 1.25 µg of the NLuc reporter plasmid. 48 hours post-transfection, cells were harvested in 1x PBS (pH 7.5). Cell suspensions were then plated into white-walled 96-well plates and assayed for luminescence using the NanoGlo Dual Luciferase Assay System and a GloMax Multi+ luminometer according to the manufacturer's instructions (Promega).

Tethered function assay data were analysed using a hierarchical Bayesian model in order to account for the differences in variation between technical and biological replicates. The input data were the log2 ratios of NLuc activity to that of the Firefly Luciferase internal control (ffLuc). We modelled the average value for each biological replicate as arising from a t-distribution centred on the overall average value μ_{prot} for the particular protein variant of interest; a separately inferred average value for the ffLuc intensity on each particular day ($\mu_{ffluc,day}$, shared across all proteins considered on that

day) was applied as an additive offset to the day-wise values. The values for the technical replicates were further modelled with t-distributions centred on the appropriate daily average value $\mu_{prot,day}$ - $\mu_{ffluc,day}$. The protein-wise scale parameters σ_{prot} were modelled as arising from a common gamma distribution with uninformative hyperpriors, whereas the other scale parameters, and the degree of freedom parameters for the t distributions were simply assigned uninformative priors. We fitted the models using JAGS via the rjags interface, running four independent Monte Carlo chains for 250,000 iterations with 25,000 steps of burn-in; we ensured convergence by verifying that the Gelman-Rubin shrinkage statistic for all parameters of interest was less than 1.1 (Gelman, 1992; Plummer, 2003a). Reported credible intervals were calculated using the highest posterior density approach with the R boa package; reported probabilities were calculated directly from the posterior distributions (Smith, 2007). We performed separate fits for each reporter construct. The main topic of interest in all cases is the central log ratio value μ_{prot} for each protein, which reflects the relative Luciferase level observed when that protein is present in the assay. We most frequently report the differences observed for the μ_{prot} value of a protein of interest relative to either that of a HT negative control or of WT NOCT. All raw data, number of replicates, and statistics from the fitted models are also shown in Section 4.5. The data analysis described here was designed and performed by P. Freddolino.

Western Blotting.

Cells harvested from the transfected cells used in the tethered function assays were harvested in 1x PBS, pelleted, and resuspended in radioactive immunoprecipitation assay buffer (RIPA buffer, 50 mM Tris (pH 8.0), 150 mM NaCl, 1% (v/v) IGEPAL CA-630, 0.5% (w/v) sodium deoxycholate, 0.1% (w/v) sodium dodecyl sulphate) for lysis. Lysates

were homogenized using QIAShredder Columns (Qiagen) and sample concentrations were determined using the BioRad DC Lowry assay. Each sample (8 µg total protein) was resolved on 4-20% TGX gradient gels (BioRad) and then transferred to Millipore EMD Immobilon membrane. Transfected NOCT proteins were detected using anti-V5 antibody (Invitrogen, R960-25) and HT was detected using anti-HT antibody (Promega, G9211). Blots were probed using HRP-conjugated Secondary Antibody (Sigma, A1047) and Pierce ECL Western Blotting substrate before exposure to autoradiography film. As a loading control, Glyceraldehyde-3-phosphate dehydrogenase (GAPDH) was detected on the same blots using anti-GAPDH antibody (Ambion, AM4300).

Northern Blotting.

Cells, transfected as described in the cell-based assay method above, were harvested in 1x PBS and total cellular RNA was extracted using the ReliaPrep RNA Cell Miniprep System (Promega). For each experiment, triplicate RNA samples were pooled. RNA from each sample (3-5 µg total RNA) was prepared for electrophoresis in 1x MOPS,

Northern Blotting:

Target:	Forward:	Reverse:
nLuc	5' CACTCGAAGATTTCGTTGGGGAC 3'	5' GGATCCTAATACGACTCACTATAGGGGATGCGAGCTGAAGCACAAGC 3' **
Firefly	5' CGAGATGAGCGTTCGGCTGGCAGAA 3'	5' GGATCCTAATACGACTCACTATAGGGCCGAAGCCGTGGTGAAATGGCA 3' **

** Primer contains T7 Promoter Sequence

Table 4.6: Sequences of primers used to generate the probes for Northern blotting to detect NLuc and Firefly mRNAs in cell-based assays.

3.7% (v/v) formaldehyde, 25% (v/v) formamide, 5% (v/v) glycerol, 1 mM EDTA, 0.025% (w/v) bromophenol blue and 0.025% (w/v) xylene cyanol FF. Samples were resolved on a 1% MOPS agarose gel containing 4% formaldehyde before transfer to Millipore EMD Ny+ membrane using 20x SSC. The membrane was UV crosslinked and prehybridized with Ultrahyb buffer (Invitrogen).

Antisense Northern probe templates for NLuc and Firefly were PCR amplified with primers containing the T7 promoter sequence (Table 4.6). Probes were transcribed using the MAXIscript T7 Transcription Kit (Ambion) with αP^{32} UTP (20 µCi at 3 µCi/mmol) and cold UTP (8 µM). Radiolabeled probes were purified using Sephadex G-25 resin (GE Healthcare). The final activity of the probes ranged from 1.3 x10^6 – 2.5 x10^6 counts per million/µg probe. Blots were probed overnight at 68 °C before washing with 2x SSC + 0.1% (w/v) SDS and then 0.1x SSC + 0.1% (w/v) SDS. Blots were exposed to phosphoimager screens for 3 days and then imaged using a Typhoon FLA 9500. Image Quant (GE Healthcare) was used to quantify the intensity of the bands for NLuc and Firefly for each sample. Background correction was performed using the rolling ball method with a radius set to 200 to determine the baseline before peak integration.

The quantitation data were analysed using a Bayesian hierarchical model similar to that described above for the tethered function assays, using the log$_2$ ratio of NLuc to ffLuc peak areas as a key statistic for each sample. As before, we modelled the log$_2$ ratio values as arising from a t-distribution, centred on the difference between day-wise values for NLuc (independently inferred for each target) and ffLuc (shared for all targets on that day). All such t-distributions used shared values for the scale and degree of freedom parameters, using an uninformative prior for the scale and the Gamma(2,0.1) prior suggested for the degrees of freedom (Juárez and Steel, 2010). Day-wise means were, in turn, modelled as arising from a normal distribution centered on $\mu_{prot,day}$ - $\mu_{ffluc,day}$ with a shared variance parameter representing technical noise in the measurements themselves. The prior for the square root of the variance was set to be a Gamma(1,2) distribution, amounting to a weak assumption that the technical noise would make no

more than 1-2 $\log_2$ units of difference in the abundance measurements. The model was

fitted in JAGS (Plummer, 2003b) using 8 independent Monte Carlo chains, with 500,000

iterations per chain and a 25,000 step burn-in period, and convergence assessed as

described above. All data, number of replicates, and statistics are shown in section 4.6.

The data analysis described here was designed and performed by P. Freddolino.

4.7 References

Abshire, E.T., Chasseur, J., Bohn, J.A., Del Rizzo, P.A., Freddolino, P.L., Goldstrohm, A.C., and Trievel, R.C. (2018). The structure of human Nocturnin reveals a conserved ribonuclease domain that represses target transcript translation and abundance in cells. Nucleic acids research *46*, 6257-6270.

Alhusaini, N., and Coller, J. (2016). The deadenylase components Not2p, Not3p, and Not5p promote mRNA decapping. Rna *22*, 709-721.

Aslam, A., Mittal, S., Koch, F., Andrau, J.C., and Winkler, G.S. (2009). The Ccr4-NOT deadenylase subunits CNOT7 and CNOT8 have overlapping roles and modulate cell proliferation. Mol Biol Cell *20*, 3840-3850.

Behm-Ansmant, I., Rehwinkel, J., Doerks, T., Stark, A., Bork, P., and Izaurralde, E. (2006). mRNA degradation by miRNAs and GW182 requires both CCR4:NOT deadenylase and DCP1:DCP2 decapping complexes. Genes Dev *20*, 1885-1898.

Bos, T.J., Nussbacher, J.K., Aigner, S., and Yeo, G.W. (2016). Tethered Function Assays as Tools to Elucidate the Molecular Roles of RNA-Binding Proteins. Adv Exp Med Biol *907*, 61-88.

Brown, C.E., Tarun, S.Z., Jr., Boeck, R., and Sachs, A.B. (1996). PAN3 encodes a subunit of the Pab1p-dependent poly(A) nuclease in Saccharomyces cerevisiae. Molecular and cellular biology *16*, 5744-5753.

Brown, J.A., Bulkley, D., Wang, J., Valenstein, M.L., Yario, T.A., Steitz, T.A., and Steitz, J.A. (2014). Structural insights into the stabilization of MALAT1 noncoding RNA by a bipartite triple helix. Nature structural & molecular biology *21*, 633-640.

Carithers, L.J., and Moore, H.M. (2015). The Genotype-Tissue Expression (GTEx) Project. Biopreserv Biobank *13*, 307-308.

Chen, J., Chiang, Y.C., and Denis, C.L. (2002). CCR4, a 3′-5′ poly(A) RNA and ssDNA exonuclease, is the catalytic component of the cytoplasmic deadenylase. The EMBO journal *21*, 1414-1426.

Clement, S.L., and Lykke-Andersen, J. (2008). A tethering approach to study proteins that activate mRNA turnover in human cells. Methods in molecular biology *419*, 121-133.

Coller, J., and Wickens, M. (2002). Tethered function assays using 3′ untranslated regions. Methods *26*, 142-150.

Consortium, G.T. (2013). The Genotype-Tissue Expression (GTEx) project. Nat Genet *45*, 580-585.

Consortium, G.T. (2015). Human genomics. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans. Science *348*, 648-660.

Cooke, A., Prigge, A., and Wickens, M. (2010). Translational repression by deadenylases. The Journal of biological chemistry *285*, 28506-28513.

Doidge, R., Mittal, S., Aslam, A., and Winkler, G.S. (2012). The anti-proliferative activity of BTG/TOB proteins is mediated via the Caf1a (CNOT7) and Caf1b (CNOT8) deadenylase subunits of the Ccr4-not complex. PloS one *7*, e51331.

Douris, N., Kojima, S., Pan, X., Lerch-Gaggl, A.F., Duong, S.Q., Hussain, M.M., and Green, C.B. (2011). Nocturnin regulates circadian trafficking of dietary lipid in intestinal enterocytes. Current biology : CB *21*, 1347-1355.

Elis, S., Courtland, H.W., Wu, Y., Rosen, C.J., Sun, H., Jepsen, K.J., Majeska, R.J., and Yakar, S. (2010). Elevated serum levels of IGF-1 are sufficient to establish normal body size and skeletal properties even in the absence of tissue IGF-1. J Bone Miner Res *25*, 1257-1266.

Freudenthal, B.D., Beard, W.A., Cuneo, M.J., Dyrkheeva, N.S., and Wilson, S.H. (2015). Capturing snapshots of APE1 processing DNA damage. Nature structural & molecular biology *22*, 924-931.

Gelman, A., Rubin D.B. (1992). Inference from Iterative Simulation Using Multiple Sequences. Statistical Science *7*, 457-472.

Goldstrohm, A.C., and Wickens, M. (2008). Multifunctional deadenylase complexes diversify mRNA control. Nature reviews Molecular cell biology *9*, 337-344.

Green, C.B., and Besharse, J.C. (1996). Identification of a novel vertebrate circadian clock-regulated gene encoding the protein nocturnin. Proceedings of the National Academy of Sciences of the United States of America *93*, 14884-14888.

Green, C.B., Douris, N., Kojima, S., Strayer, C.A., Fogerty, J., Lourim, D., Keller, S.R., and Besharse, J.C. (2007). Loss of Nocturnin, a circadian deadenylase, confers resistance to hepatic steatosis and diet-induced obesity. Proceedings of the National Academy of Sciences of the United States of America *104*, 9888-9893.

Juárez, M.A., and Steel, M.F.J. (2010). Model-Based Clustering of Non-Gaussian Panel Data Based on Skew-t Distributions. Journal of Business & Economic Statistics *28*, 52-66.

Kawai, M., Green, C.B., Lecka-Czernik, B., Douris, N., Gilbert, M.R., Kojima, S., Ackert-Bicknell, C., Garg, N., Horowitz, M.C., Adamo, M.L.*, et al.* (2010). A circadian-regulated gene, Nocturnin, promotes adipogenesis by stimulating PPAR-gamma nuclear translocation. Proceedings of the National Academy of Sciences of the United States of America *107*, 10508-10513.

Keen, J.C., and Moore, H.M. (2015). The Genotype-Tissue Expression (GTEx) Project: Linking Clinical Data with Molecular Analysis to Advance Personalized Medicine. J Pers Med *5*, 22-29.

Kojima, S., Gendreau, K.L., Sher-Chen, E.L., Gao, P., and Green, C.B. (2015). Changes in poly(A) tail length dynamics from the loss of the circadian deadenylase Nocturnin. Scientific reports *5*, 17059.

Langlois, J.A., Rosen, C.J., Visser, M., Hannan, M.T., Harris, T., Wilson, P.W., and Kiel, D.P. (1998). Association between insulin-like growth factor I and bone mineral density in older women and men: the Framingham Heart Study. The Journal of clinical endocrinology and metabolism *83*, 4257-4262.

Lin, Y.C., Boone, M., Meuris, L., Lemmens, I., Van Roy, N., Soete, A., Reumers, J., Moisse, M., Plaisance, S., Drmanac, R.*, et al.* (2014). Genome dynamics of the human embryonic kidney 293 lineage in response to cell biology manipulations. Nature communications *5*, 4767.

Liu, J.P., Baker, J., Perkins, A.S., Robertson, E.J., and Efstratiadis, A. (1993). Mice carrying null mutations of the genes encoding insulin-like growth factor I (Igf-1) and type 1 IGF receptor (Igf1r). Cell *75*, 59-72.

Plummer, M. (2003a). JAGS: A Program for Analysis of Bayesian Graphical Models Using Gibbs Sampling. Proceedings of the 3rd International Workshop on Distributed Statistical Computing, 1-10.

Plummer, M. (2003b). JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling.

Rorbach, J., Nicholls, T.J., and Minczuk, M. (2011). PDE12 removes mitochondrial RNA poly(A) tails and controls translation in human mitochondria. Nucleic acids research *39*, 7750-7763.

Rouya, C., Siddiqui, N., Morita, M., Duchaine, T.F., Fabian, M.R., and Sonenberg, N. (2014). Human DDX6 effects miRNA-mediated gene silencing via direct binding to CNOT1. Rna *20*, 1398-1409.

Schellenberg, M.J., Appel, C.D., Adhikari, S., Robertson, P.D., Ramsden, D.A., and Williams, R.S. (2012). Mechanism of repair of 5´-topoisomerase II-DNA adducts by mammalian tyrosyl-DNA phosphodiesterase 2. Nature structural & molecular biology *19*, 1363-1371.

Shaw, G., Morse, S., Ararat, M., and Graham, F.L. (2002). Preferential transformation of human neuronal cells by human adenoviruses and the origin of HEK 293 cells. FASEB J *16*, 869-871.

Smith, B.J. (2007). boa: An R Package for MCMC Output Convergence Assessment and Posterior Inference. Journal of Statistical Software *21*.

Stubblefield, J.J., Gao, P., Kilaru, G., Mukadam, B., Terrien, J., and Green, C.B. (2018). Temporal Control of Metabolic Amplitude by Nocturnin. Cell reports *22*, 1225-1235.

Van Etten, J., Schagat, T.L., and Goldstrohm, A.C. (2013). A guide to design and optimization of reporter assays for 3´ untranslated region mediated regulation of mammalian messenger RNAs. Methods *63*, 110-118.

Viswanathan, P., Ohn, T., Chiang, Y.C., Chen, J., and Denis, C.L. (2004). Mouse CAF1 can function as a processive deadenylase/3´-5´-exonuclease in vitro but in yeast the deadenylase function of CAF1 is not required for mRNA poly(A) removal. The Journal of biological chemistry *279*, 23988-23995.

Wang, Y., Osterbur, D.L., Megaw, P.L., Tosini, G., Fukuhara, C., Green, C.B., and Besharse, J.C. (2001). Rhythmic expression of Nocturnin mRNA in multiple tissues of the mouse. BMC developmental biology *1*, 9.

Wilusz, J.E. (2016). Long noncoding RNAs: Re-writing dogmas of RNA processing and stability. Biochimica et biophysica acta *1859*, 128-138.

Wolf, J., Valkov, E., Allen, M.D., Meineke, B., Gordiyenko, Y., McLaughlin, S.H., Olsen, T.M., Robinson, C.V., Bycroft, M., Stewart, M.*, et al.* (2014). Structural basis for Pan3 binding to Pan2 and its function in mRNA recruitment and deadenylation. The EMBO journal *33*, 1514-1526.

Yakar, S., Rosen, C.J., Beamer, W.G., Ackert-Bicknell, C.L., Wu, Y., Liu, J.L., Ooi, G.T., Setser, J., Frystyk, J., Boisclair, Y.R.*, et al.* (2002). Circulating levels of IGF-1 directly regulate bone growth and density. J Clin Invest *110*, 771-781.

Zinder, J.C., and Lima, C.D. (2017). Targeting RNA for processing or destruction by the eukaryotic RNA exosome and its cofactors. Genes Dev *31*, 88-100.

CHAPTER 5

Analysis of NOCT-mediated Regulation of Gene Expression and Metabolism[1]

5.1 Abstract

NOCT-mediated regulation is thought to occur through repression of mRNAs involved in metabolic processes such as lipid trafficking and metabolism and regulation of body mass. However, identification of these mRNAs has proved difficult, and definitive evidence of direct regulation remains elusive. As the intracellular localization of NOCT protein would considerably change the pool of potential targets, we examined the NOCT protein sequence for motifs indicating subcellular localization. Sequence features of the NOCT protein suggests that full-length NOCT is a preprotein that may be localized to the mitochondria, wherein it can be proteolytically processed. Consistent with this hypothesis, mutagenesis of predicted mitochondrial targeting sequence features prevents processing of NOCT. Interestingly, processing of NOCT appears to occur in a cell and tissue type-specific manner, suggesting the potential for differential localization of NOCT to the cytoplasm or mitochondria. Identification of endogenous NOCT substrates within these

---

[1]The work described in this chapter was a collaboration with Dr. Kelsey Hughes, who created the cell lines and RNA samples for the RNA sequencing experiments. Dr./ Hughes also performed the analysis of NOCT protein processing. Dr. Peter Freddolino and Rucheng Diao performed bioinformatic analysis of the RNA-Seq datasets from these cell lines.

subcellular compartments is crucial for understanding and characterizing the recognition and specificity of NOCT substrates. To identify NOCT-mediated changes in gene expression, we expressed mutant version of NOCT that is restricted to the cytoplasm and measured resulting changes in mRNA levels by RNA-Seq. These results suggest that in a neuro-adrenal cell line, NOCT overexpression in the cytoplasm downregulates transcripts in a transcript isoform-dependent manner. We additionally confirmed a report that NOCT can dephosphorylate nicotinamide adenine dinucleotide substrates, providing a potential direct role for NOCT in regulation of metabolism and/or redox biology. Together, our results suggest that NOCT is capable of affecting metabolism as a dual-function enzyme involved in post-transcriptional gene regulation and modulation of metabolite levels.

5.2 Introduction

Loss of NOCT expression has a profound effect on body mass, adipogenesis, and lipid metabolism and trafficking (Douris et al., 2011; Green et al., 2007; Kawai et al., 2010b). This effect has been predicted to be mediated by through mechanisms of mRNA regulation, as NOCT is most similar to Ccr4-type deadenylases and represses expression of reporter mRNAs *in vivo* (Abshire et al., 2018; Baggs and Green, 2003; Douris et al., 2011; Green et al., 2007; Kawai et al., 2010b). This hypothesis is supported by global studies that have identified changes in mRNA transcript levels upon perturbation of NOCT expression (Hee et al., 2012; Kojima et al., 2015; Stubblefield et al., 2018). However, some of these previous studies are limited in that it is assumed that mRNA targets of NOCT will display circadian behavior and decreased poly(A) tail lengths (discussed in Chapter 1) (Kojima et al., 2015; Stubblefield et al., 2018). Thus, while these studies

provide evidence that NOCT affects mRNA levels *in vivo*, the criteria for target identification may be overly restrictive.

Identification of putative NOCT targets have not taken into account the potential for localization of NOCT into different cellular compartments. NOCT has previously been described to localize to the nucleus and cytoplasm in murine oocytes and the perinucleus in HEK293 cells (Kawai et al., 2010b; Nishikawa et al., 2013). A pair of recent articles have additionally provided evidence that NOCT localizes to the mitochondria (Estrella et al., 2019; Le et al., 2019). The pools of mRNAs in cellular compartments can differ significantly, especially in the mitochondria, which has its own genome and transcribes and translates 13 mRNAs distinct from nuclear encoded mRNAs (D'Souza and Minczuk, 2018). Therefore, determining NOCT localization can provide a greater understanding of the potential mRNAs it could affect.

While identification of endogenous NOCT targets has largely been limited to mRNAs, the potential for NOCT regulation of non-mRNA targets must be considered. This has been shown to be the case for the human EEP deadenylase PDE12, which has activity against both mitochondrial tRNAs and 2′, 5′ oligoadenylate (2′,5′-A) (Pearce et al., 2017; Wood et al., 2015). During the writing of this thesis, a preprint publication reported NOCT activity against non-RNA substrates $NADP^+$ and NADPH, an enzymatic activity that has only recently been suggested to exist in mammals (Ding et al., 2018; Estrella et al., 2019). This result suggests that NOCT may have non-RNA substrates and that this phosphatase activity may have significant effects on cellular metabolism.

Here we describe the features of the NOCT mRNA and protein that suggest that NOCT could be localized to the cytoplasm or to the mitochondria. These features are

interrogated through mutagenesis and we show that in HEK293 and HCT116 cells, the N-terminus of NOCT facilitates processing of the NOCT protein in a manner consistent with mitochondrial localization. In a survey of different mouse and human tissue types, we demonstrate that NOCT processing occurs in a tissue-dependent manner. We additionally discuss ongoing efforts to characterize the endogenous substrates of NOCT using RNA-Seq, demonstrating that overexpression of cytoplasmic NOCT alters steady state levels of mRNA transcripts. Additionally, we corroborate the reports of $NADP^+$ and NADPH phosphatase activity. NOCT targets will be discussed in the context of NOCT phenotypes as well as the additional biological processes that may be influenced by NOCT activity.

5.3 Results

5.3.I NOCT is processed in a manner consistent with mitochondrial localization

To accurately assess NOCT-mediated differential gene expression, the localization of NOCT first needed to be determined. We examined the NOCT amino acid sequence using MitoFates to identify sequence elements predicted to influence

Figure 5.1: Features of full-length NOCT predicted to influence subcellular localization. The NOCT N-terminus has a predicted mitochondrial localization sequence (aa1-15) followed by a consensus sequence for processing by mitochondrial peptidases. Additionally, there are two Met residues with Kozak sequences of varying optimality for initiating translation, one of which would exclude the mitochondrial localization sequence.

mitochondrial localization. MitoFates scores the probability of localization by identifying N-terminal protein sequences for enrichment of arginine residues within a sequence an amphipathic α-helix required for recognition by TOM20. These sequences facilitate import of polypeptides through the TOM and TIM translocase complexes into the mitochondrial matrix (Schmidt et al., 2010). This analysis identified a predicted mitochondrial targeting sequence (MTS) in the first 15 amino acids of the NOCT N-terminus (Probability (P)=0.94) (Fukasawa et al., 2015). Deletion of residues 2-15 from the NOCT protein sequence reduced the MitoFates prediction of localization to P=0.35. Downstream of this putative localization sequence are predicted cleavage sites for mitochondrial matrix processing peptidase (MPP) and intermediate cleaving peptidase 55 (Icp55), which would remove mitochondrial pre-sequences after Leu74 and then remove Tyr75, respectively (Figure 5.1) (Fukasawa et al., 2015). Based on these observations, we designed a set of constructs with mutations in the NOCT N-terminus to address whether NOCT is processed in a manner consistent with mitochondrial localization and subsequent proteolytic processing.

We examined the molecular weights of endogenous NOCT expressed in HEK293 and HepG2 cell lines for evidence of proteolytic processing. For these experiments, we generated a polyclonal α-NOCT antibody against full-length NOCT and verified its reactivity against recombinant human NOCT$_{64-431}$ (Figure 5.2). Full-length NOCT has an apparent molecular weight of 48.2 kDa based on amino acid sequence (Figure 5.2 A). In HepG2 cells, the major band is ~55 kDa, which is larger than predicted but may be indicative of the charge of the NOCT N-terminus, which is enriched in basic residues (Figure 5.2 B). In addition to the ~55 kDa band, a smaller, ~41 kDa band was observed

both in HepG2 and HEK293 cells, which may suggest that multiple isoforms of NOCT are generated in cells. While these sizes are consistent with the change in molecular weight predicted for processing by MPP and/or Icp55, we considered that these different NOCT isoforms may be produced by either differential translation or processing of *NOCT* mRNA.

Previous studies have observed the human, mouse and *Xenopus Noct* mRNA sequences have two different potential start codons in the context of Kozak sequences, suggesting that multiple isoforms of NOCT may be translated (Wang et al., 2001). We investigated the potential usage of two translation start sites at Met1 and Met67 of the human NOCT mRNA sequence. Both Kozak sequences (Met1 5´-CCCGC<u>AUG</u>U-3´ and Met67 5´-UGUUCC<u>AUG</u>G-3´) are divergent from the canonical sequence (5´-gccRcc<u>AUG</u>G-3´), suggesting that leaky scanning through the Met1 Kozak sequence may allow for translation of a long and a short isoform of NOCT. While we had also considered that differential processing of *NOCT* mRNA could produce transcripts expressing different NOCT isoforms, we did not find evidence of multiple isoforms of *NOCT* mRNA in humans.

We examined the utilization the two observed NOCT translation initiation sites by cloning the NOCT ORF and 5´ leader sequences into expression vectors. These constructs were mutated either at the Met1 start codon (M1X) or at the Met67 start codon (M67G) and were then transfected into HEK293 cells. NOCT protein expression was assessed using Western blotting with α-NOCT. We observed an increase in the amount of the ~41 kDa band with overexpression of WT NOCT constructs, which is close to the size either produced by Met67 translation initiation (41.3 kDa) or by Met1 initiation followed by proteolytic processing to the ~41 kDa form (40.4 kDa). Mutation of Met1 produced a ~41 kDa protein, indicating that leaky scanning can result in translation from

130

Met67. When Met67 is mutated, the observed protein is also ~41 kDa, indicating that

Met1 initiation is followed by processing of NOCT to the shorter isoform (Figure 5.2 B).

We also generated a construct where the putative MTS is deleted (NOCT Δ2-15), which
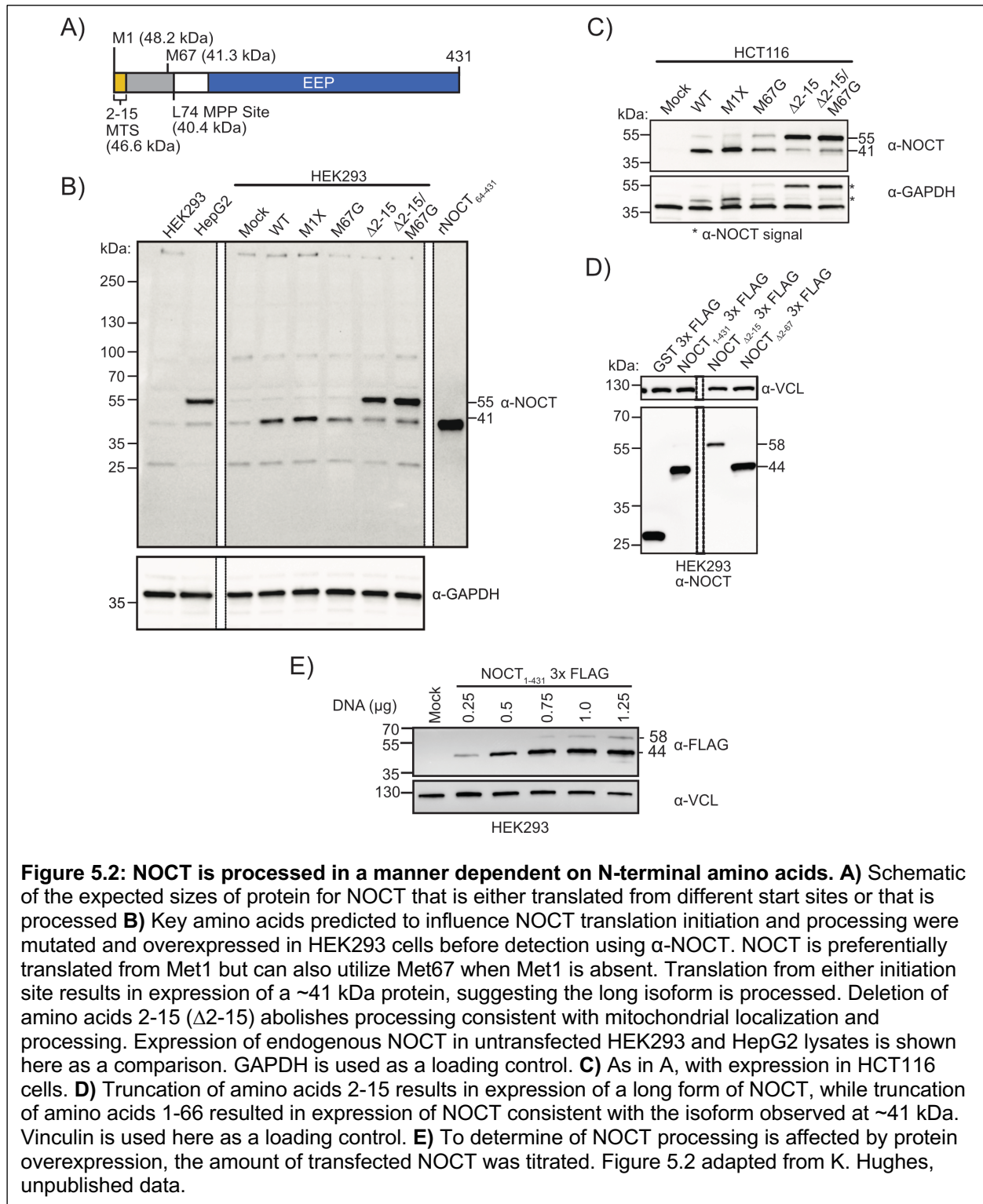
Figure 5.2: NOCT is processed in a manner dependent on N-terminal amino acids. A) Schematic of the expected sizes of protein for NOCT that is either translated from different start sites or that is processed **B)** Key amino acids predicted to influence NOCT translation initiation and processing were mutated and overexpressed in HEK293 cells before detection using α-NOCT. NOCT is preferentially translated from Met1 but can also utilize Met67 when Met1 is absent. Translation from either initiation site results in expression of a ~41 kDa protein, suggesting the long isoform is processed. Deletion of amino acids 2-15 (Δ2-15) abolishes processing consistent with mitochondrial localization and processing. Expression of endogenous NOCT in untransfected HEK293 and HepG2 lysates is shown here as a comparison. GAPDH is used as a loading control. **C)** As in A, with expression in HCT116 cells. **D)** Truncation of amino acids 2-15 results in expression of a long form of NOCT, while truncation of amino acids 1-66 resulted in expression of NOCT consistent with the isoform observed at ~41 kDa. Vinculin is used here as a loading control. **E)** To determine of NOCT processing is affected by protein overexpression, the amount of transfected NOCT was titrated. Figure 5.2 adapted from K. Hughes, unpublished data.

resulted in the expression of a ~55 kDa protein as the predominant isoform. We also mutated Met67 in the NOCT Δ2-15 construct (Δ2-15/M67G), which has a very similar pattern of protein expression compared to the NOCT Δ2-15 construct (Figure 5.2 B). This indicates that Met1 is the preferred translation initiation site, and that translation initiation from Met67 is observed when Met1 is mutated. Together, these data support a hypothesis wherein *NOCT* translation is initiated at the Met1 start codon, followed by proteolytic processing of the protein to a ~41 kDa form.

We further corroborated this hypothesis by transfecting the WT and mutant NOCT constructs into HCT116 cells. We observed that a ~41 kDa protein is the predominant form of NOCT that is produced unless the MTS is deleted, demonstrating that processing is not specific to the HEK293 cell line. Additionally, we examined processing of NOCT constructs with a 3x C-terminal FLAG tag as an alternative method of detecting molecular weight changes (Figure 5.2 C). Full length NOCT 3x FLAG was detected at ~44 kDa (which corresponds to the size expected for processed NOCT plus the ~3 kDa of mass added by the tag). NOCT Δ2-15 3x FLAG has an apparent molecular weight of ~58 kDa, which is consistent with the deletion of the MTS preventing processing (Figure 5.2 D). As a point of comparison, we transfected a NOCT Δ2-67 3x FLAG construct, which is observed at its expected size of ~44 kDa (Figure 5.2 D). As we have determined that expression of full-length (~55 kDa) NOCT does not necessarily result in processing (as is observed in HepG2 cells), we verified that the processing observed in HEK293 cells was not due to protein overexpression by titrated the amount of transfected WT NOCT. We observed that the ~41 kDa species is seen even with smaller amounts of DNA transfected. The ~55 kDa band is observed with transfection of 0.75 µg of DNA or more,

indicating that WT NOCT is saturating the proteolytic machinery responsible for processing (Figure 5.2 E). These results suggest that the ~55 kDa species observed in HepG2 cells is a pre-protein that is processed to generate the ~41 kDa species. Based on our MitoFates analysis, we predict that processing occurs in the mitochondria, as deletion of the putative MTS prevents expression of the ~41 kDa protein (Figure 5.2 B).

Figure 5.3: The N-terminus of NOCT is sufficient for processing. Fusion of the 86 N-terminal amino acids of NOCT to EGFP results in processing of EGFP consistent with mitochondrial localization and processing. Figure 5.3 adapted from K. Hughes, unpublished data.
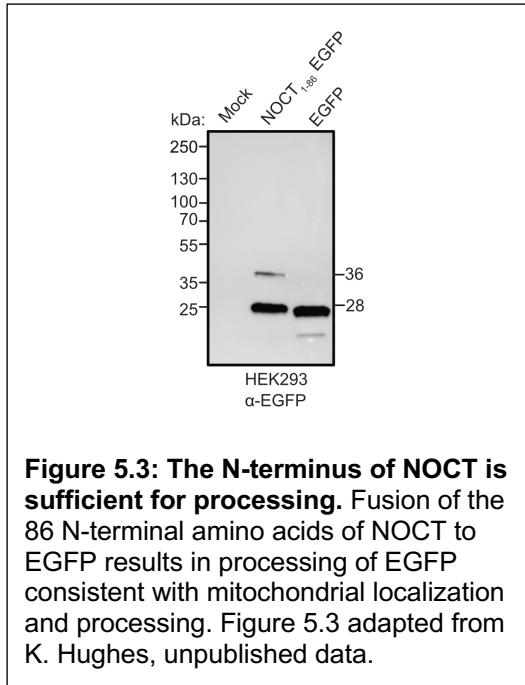
To assess if the N-terminus of NOCT is sufficient for processing, we created an EGFP fusion protein with the first 86 amino acids of NOCT at the N-terminus. We observed a minor band at ~36 kDa, which is the expected molecular weight for the full-length $NOCT_{1-86}$- EGFP fusion protein. The major species has an apparent molecular weight of ~28 kDa, which is consistent with mitochondrial peptidases processing the $NOCT_{1-86}$- EGFP fusion protein, leaving 12 remaining amino acids of the NOCT N-terminus (amino acids 74-86) fused to EGFP (Figure 5.3). Together, our data indicate that in HEK293 and HCT116 cells, NOCT is processed in a manner consistent with predicted localization to the mitochondria, and that the N-terminus of NOCT is sufficient for processing of the presumed NOCT pre-protein.

As we analyzed the apparent sizes of endogenous NOCT in various cell types, we observed that in HepG2 cells, the ~55 kDa form of NOCT is the major form (Figure 5.2 B). The expression of the long isoform suggests that there are cases where NOCT is translated from the Met1 initiation site but is not processed. Following this observation,

Figure 5.4: NOCT appears to be differentially processed in different tissue types. Western blots of tissue extracts from either **A)** human or **B)** mouse tissues were probed using α-NOCT. Bands consistent with the ~41 kDa and ~55 kDa forms of NOCT were observed, indicating that processing and possibly subcellular localization of endogenous NOCT may occur in a tissue-specific manner. Figure 5.4 adapted from K. Hughes, unpublished data.

we next asked whether NOCT processing was specific to tissue type. We obtained blots with both mouse and human tissue extracts and probed for NOCT with our polyclonal α-NOCT antibody (Figure 5.4). This revealed that the ~55 kDa band is present in multiple tissues such as kidney, brain, and human liver, as well as in stomach and brain in mouse tissue. We also observed bands around ~41 kDa, which would correspond to the smaller processed form in brain and skeletal muscle in human and in brain, heart, and skeletal muscle in mouse tissues. Other bands outside of the 40-55 kDa range cannot be explained by processing or full-length NOCT expression and may be proteins cross-reacting with our polyclonal antibody. In tissues such as human spleen or mouse ovary, NOCT protein expression was below the limit of detection. The presence of bands at ~55 kDa and ~41 kDa in various human and mouse tissues supports our hypothesis that NOCT processing occurs in a tissue-dependent manner.

In summary, our results suggest that multiple NOCT isoforms are expressed, and that the expression of the shorter, ~41 kDa isoform of NOCT is dependent on the predicted MTS. The other major species detected is a ~55 kDa form consistent with expression of full-length, unprocessed NOCT, which suggests that NOCT is not processed in all cases. Furthermore, the patterns of expression and sizes of protein vary with cell and tissue type, suggesting that NOCT expression and processing is tissue specific (Figures 5.3 and 5.4). Using this information, current and future studies are focused on the cytoplasmic mRNA targets of NOCT as well as protein partners that are themselves localized to the cytoplasm.

5.3.II Cytoplasmic NOCT overexpression results in differential regulation of mRNA transcripts.

Figure 5.5: Verification of stable expression of FLAG-tagged fusion proteins in HEK293 cell lines. NOCT Δ2-15 3x FLAG and GST 3x FLAG expression was verified by Western blotting against the 3x FLAG epitope tag. Vinculin was used as a loading control. Figure 5.5 adapted from K. Hughes, unpublished data.
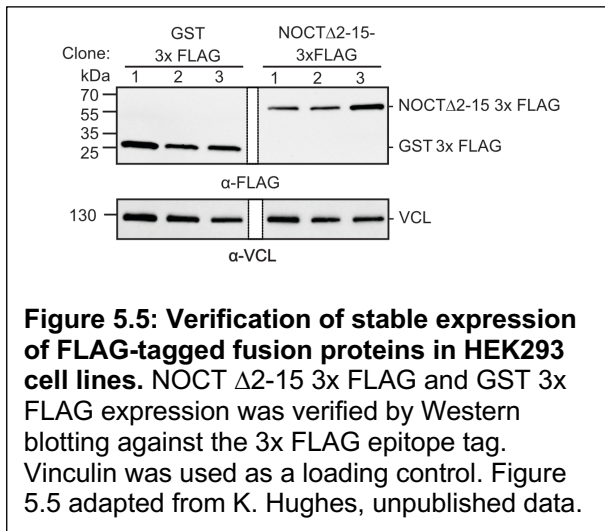
We next asked whether overexpression of a cytoplasmic form of NOCT (NOCT Δ2-15 3x FLAG) would result in changes in expression levels of nuclear-encoded mRNAs. We performed RNA-Seq analysis of RNA from cell lines that constitutively overexpress NOCT Δ2-15 3x FLAG in HEK293 cells. Use of these cells is advantageous, as HEK293 express low levels of endogenous NOCT and are easily genetically manipulated. Clonal lines of HEK293 cell lines that stably express NOCT Δ2-15 3x FLAG; or, as a negative control, C-terminal FLAG-tagged Glutathione S-transferase (GST 3x FLAG), were created in order to assess the effect of NOCT overexpression (Figure 5.5). RNA samples from three clonal lines for

each condition were submitted for RNA sequencing, resulting in a high-confidence dataset for changes in transcript abundance occurring with NOCT overexpression. Our analysis used rRNA depletion rather than oligo d(T) purification, so as to remove abundant rRNAs and identify regulated transcripts independently of poly(A) tail length. Furthermore, we did not restrict our analysis to circadian transcripts, as NOCT may regulate transcripts that do not show rhythmic expression.

Using statistical significance criteria (adjusted p-value $\geq$ -log10(0.05)) for expression changes, we identified 255 upregulated transcripts, with 32 transcripts regulated between 2 and 4-fold and 2 transcripts with greater than 4-fold changes. We also identified 235 downregulated transcripts, with 31 transcripts negatively regulated between 2 and 4-fold and 7 transcripts downregulated with greater than 4-fold changes (Figure 5.6 and Table 5.1). Differential regulation of mRNA transcripts upon NOCT overexpression indicates that NOCT is able to modulate the abundances of endogenous nuclear-encoded mRNA transcripts (Table 5.1).
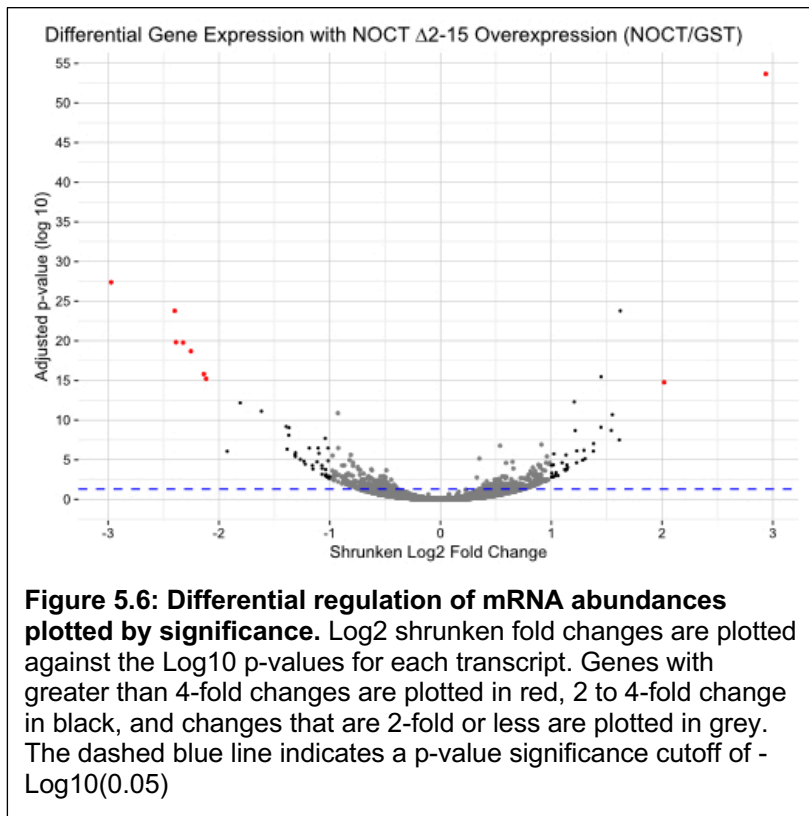
Figure 5.6: Differential regulation of mRNA abundances plotted by significance. Log2 shrunken fold changes are plotted against the Log10 p-values for each transcript. Genes with greater than 4-fold changes are plotted in red, 2 to 4-fold change in black, and changes that are 2-fold or less are plotted in grey. The dashed blue line indicates a p-value significance cutoff of -Log10(0.05)

Target	Avg TPM (GST)	Avg TPM (NOCT)	RNAseq Log2 Fold Change	qPCR Log2 Fold Change	Function
DLGAP1	0.5305	0.0112	-2.44	-2.94	Neuron synapse, NMDA signaling
RORB	2.5833	0.3179	-2.02	-2.40	Circadian regualtion, neurodevelopment
TMEFF2	0.8877	0.1155	-2.64	-2.32	Neuron survival factor
PTPRZ1	0.1660	0.0337	-2.23	-2.13	Glutamate signaling, AMPA receptor
FRRS1L	1.0470	0.4481	-4.12	-2.12	Oligodenrocyte differentiation, receptor tyrosine phosphatase

Table 5.1: Log$_2$ Fold Changes observed for 5 selected transcripts in the top ten downregulated targets observed with NOCT overexpression. For each target the average transcripts per million (TPM) count for each HEK293 cell line is reported in addition to the Log2 fold changes measured in RNA-Seq and qPCR.
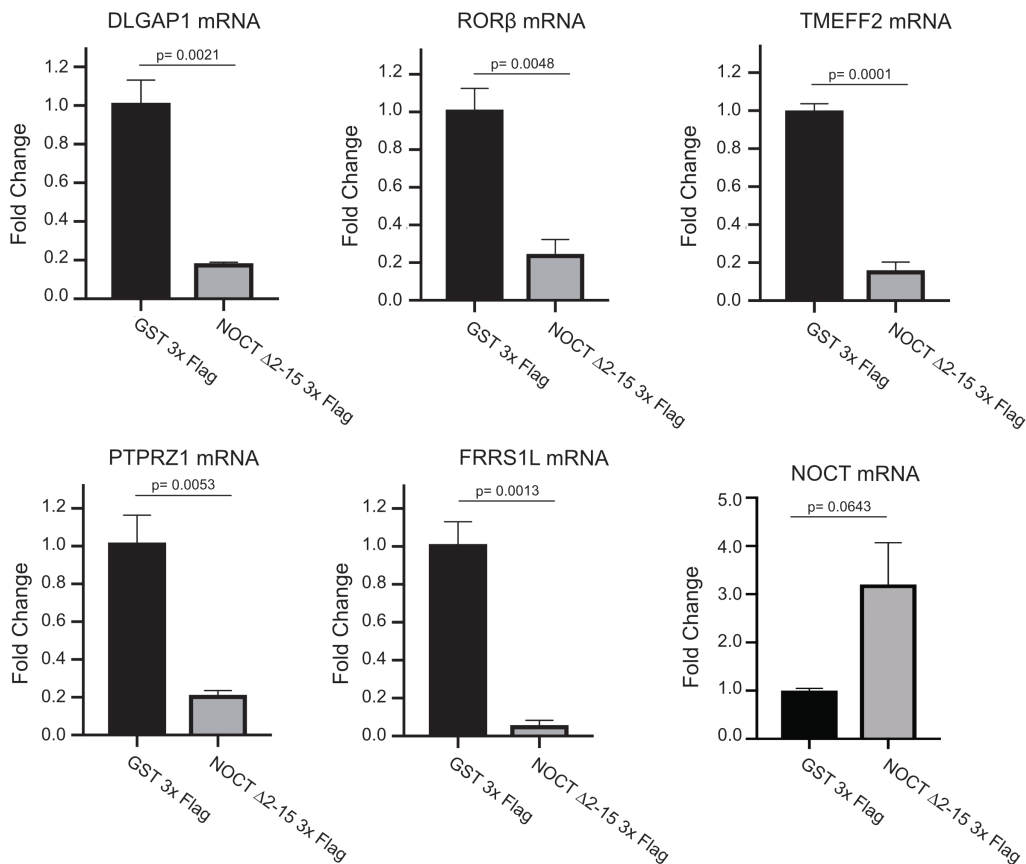
Figure 5.7: Fold change expression of top downregulated mRNA targets in RNA-Seq experiments. Fold changes in expression for DLGAP1, RORβ, TMEFF2, PTPRZ1, FRRS1L, and NOCT were measured in the clonal lines overexpressing either GST 3x FLAG or NOCT Δ2-15 3x FLAG constructs. Differences in mRNA abundances for each cell line were normalized to TATA binding protein (TBP). In each case, NOCT overexpression is associated with reduced levels of mRNA abundances for these transcripts.

We validated out results by measuring the changes in mRNA expression levels using 5 of the top 10 downregulated transcripts by qPCR (Figure 5.7). In all cases, the fold-change measurements in the qPCR data were very similar to those observed in the RNA-Seq dataset, which shows that we can measure NOCT-mediated changes in mRNA abundance using two different methods of quantitation (Table 5.1). To determine if the decrease in mRNA levels corresponded to a decrease in the steady state levels of protein encoded by these transcripts, we additionally examined the changes in protein levels in the NOCT Δ2-15 3x FLAG and GST 3x FLAG overexpression cell lines. Contrary to what was observed for the mRNA levels of *FRRS1L*, *PTPRZ1*, and *RORβ*, protein levels did not decrease with overexpression of NOCT Δ2-15 3x FLAG (Figure 5.8). These results are preliminary but suggest that other factors could regulate protein abundance for these genes, such that NOCT-mediated downregulation mRNA abundance is not sufficient to decrease steady state levels of protein. Additionally, steady state measurements of these proteins may not capture decreased translation from a pool of mRNAs with reduced expression levels.
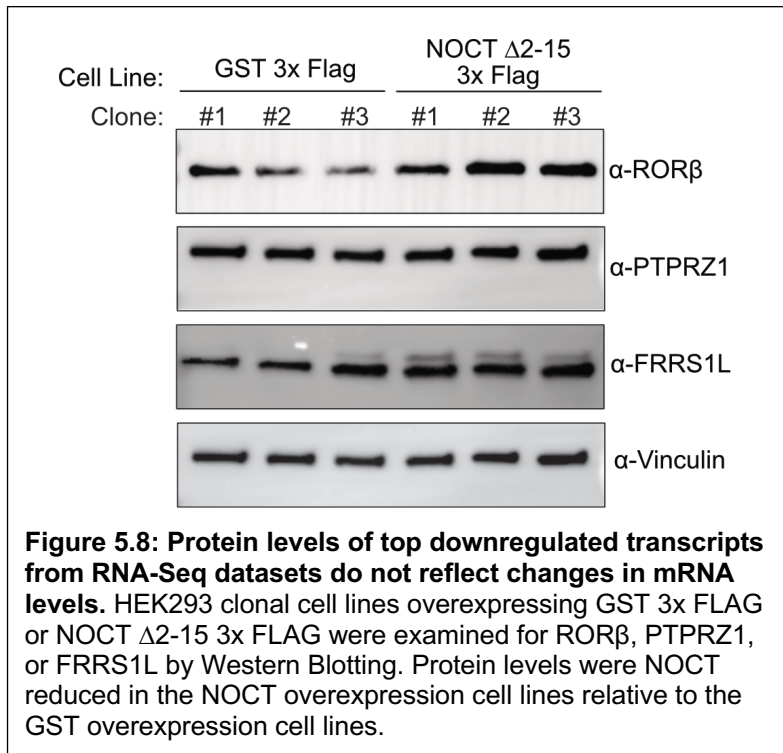
The RNA-Seq data was then examined at the transcript level to determine if differential regulation mRNA transcript

Figure 5.8: Protein levels of top downregulated transcripts from RNA-Seq datasets do not reflect changes in mRNA levels. HEK293 clonal cell lines overexpressing GST 3x FLAG or NOCT Δ2-15 3x FLAG were examined for RORβ, PTPRZ1, or FRRS1L by Western Blotting. Protein levels were NOCT reduced in the NOCT overexpression cell lines relative to the GST overexpression cell lines.

isoforms are evident in the sequencing data. To do this, the RNA-Seq data were reanalyzed for log2 fold changes in expression at the level of individual transcript isoforms of *DLGAP1, TMEFF2,* and *PTPRZ1* (Table 5.2). Within the different transcript isoforms of *TMEFF2*, the log2 fold changes for isoforms such as *TMEFF2-202* (log2 FC = -4.8) and *TMEFF2-203* (log2 FC= -2.18), were variable, suggesting that not all transcript isoforms are regulated to the same extent. Furthermore, not all isoforms for a given gene are significantly regulated. For example, *DLGAP1* has five different isoforms detected in our RNA-Seq data, all of which have decreased mRNA abundance with NOCT Δ2-15 versus GST overexpression; however, only a subset of these isoforms have changes that reach statistical significance (false discovery rate adjusted p-values (q-values) below the threshold cutoff of q=0.05) (Table 5.2). This demonstrates that while changes in mRNA abundance are observed at the gene level of analysis, not all isoforms of a given gene

Ensembl Accession:	Transcript Name:	Biotype:	Log2 fold change with NOCT OE:	q-value:
ENST00000315677	DLGAP1-201	mRNA, protein coding	-3.86	9.72E-05
ENST00000400150	DLGAP1-205	mRNA, protein coding	-3.25	0.9
ENST00000581527	DLGAP1-218	mRNA, protein coding	-4.36	4.67E-05
ENST00000485480	DLGAP1-209	lncRNA, processed transcript	-0.16	1
ENST00000581550	DLGAP1-219	lncRNA, processed transcript	-3.04	0.13
ENST00000272771	TMEFF2-201	mRNA, protein coding	-2.75	0.01
ENST00000392314	TMEFF2-202	mRNA, protein coding	-4.8	3.06E-07
ENST00000409056	TMEFF2-203	mRNA, protein coding	-2.18	2.68E-04
ENST00000393386	PTPRZ1-201	mRNA, protein coding	-1.58	0.89
ENST00000449182	PTPRZ1-202	mRNA, protein coding	-3.14	0.17
ENST00000471837	PTPRZ1-205	lncRNA, processed transcript	-0.67	0.98

Table 5.2: Differential regulation observed at the transcript level in RNA-Seq datasets. For these transcript isoforms, the decrease in mRNA abundance between NOCT Δ2-15 overexpression versus GST overexpression is reported. The significance is reported as a q-value, where significant changes in expression have q values $\leq$ 0.05.

may be regulated by overexpression of NOCT Δ2-15. Further validation of these results

is needed to determine if these observed differences in mRNA isoform abundance results

in biologically significant regulation of one isoform relative to another for a given gene and

if regulation of certain isoforms relative to others results in functional changes at the

protein level.

Figure 5.9: iPAGE analysis of the RNA-Seq datasets reveals GO terms over and underrepresented in differentially regulated mRNA transcripts. 51 individual GO terms were identified that have changes in the abundances of their associated mRNAs with respect to NOCT expression/ GST expression. For each category, the mRNAs included within the category are equally distributed into bins, with the most downregulated transcripts on the left and most upregulated transcripts on the right of the heat map. Figure 5.9 was adapted from R.Diao, unpublished data.

To assess what biological processes and pathways are enriched in the genes

differentially regulated by NOCT overexpression, we performed a gene ontology (GO)

analysis using iPAGE (Goodarzi et al., 2009). This analysis identified a total of 51

categories in which mRNA abundances changed significantly in NOCT Δ2-15 versus GST overexpression cell lines (Figure 5.9). Within each category, the individual genes were equally partitioned into discrete "bins" so that overrepresentation and underrepresentation of the individual GO term category could be assessed with respect to mRNAs that have increased abundance or decreased abundance. The downregulated transcripts are enriched for terms such as "voltage gated potassium channels" and "negative regulation of osteoblast differentiation", while terms such as "kinetochore" and "DNA replication" are underrepresented (Figure 5.9 and Table 5.3). It should be noted that the "kinetochore" and "DNA replication" terms are similarly underrepresented among the upregulated regulated genes (Table 5.4). For the genes that are upregulated with NOCT Δ2-15 overexpression, terms such as "retinol binding" and "positive regulation of cytosolic calcium ion concentration" are overrepresented.

GO Term	GO Term Description	Regulation of genes in GO Term (NOCT/GST):	Representation of GO Terms in Dataset:	p-Value:
GO:0007156	homophilic cell adhesion via plasma membrane adhesion molecules	Downregulated	Overrepresented	2.89E-08
GO:0005249	voltage-gated potassium channel activity	Downregulated	Overrepresented	8.02E-07
GO:0045668	negative regulation of osteoblast differentiation	Downregulated	Overrepresented	1.02E-06
GO:0004714	transmembrane receptor protein tyrosine kinase activity	Downregulated	Overrepresented	2.61E-05
GO:0004970	ionotropic glutamate receptor activity	Downregulated	Overrepresented	7.14E-05
GO:0000776	kinetochore	Downregulated	Underrepresented	4.78E-05
GO:0003735	structural constituent of ribosome	Downregulated	Underrepresented	9.44E-05
GO:0006260	DNA replication	Downregulated	Underrepresented	4.51E-04
GO:0051082	unfolded protein binding	Downregulated	Underrepresented	1.01E-03
GO:0005681	spliceosomal complex	Downregulated	Underrepresented	1.48E-03
GO:0005604	basement membrane	Upregulated	Overrepresented	2.31E-11
GO:0019841	retinol binding	Upregulated	Overrepresented	1.87E-07
GO:0007204	positive regulation of cytosolic calcium ion concentration	Upregulated	Overrepresented	1.21E-06
GO:0008083	growth factor activity	Upregulated	Overrepresented	1.55E-06
GO:0010757	negative regulation of plasminogen activation	Upregulated	Overrepresented	1.69E-05
GO:0003735	structural constituent of ribosome	Upregulated	Underrepresented	1.34E-08
GO:0008168	methyltransferase activity	Upregulated	Underrepresented	3.00E-06
GO:0006260	DNA replication	Upregulated	Underrepresented	2.27E-05
GO:0043161	proteasome-mediated ubiquitin-dependent protein catabolic process	Upregulated	Underrepresented	3.73E-05
GO:0000776	kinetochore	Upregulated	Underrepresented	4.73E-05

Table 5.3: GO terms over and underrepresented for the most significantly downregulated and upregulated mRNA transcripts. Terms represented in multiple categories are likely not biologically significant results.
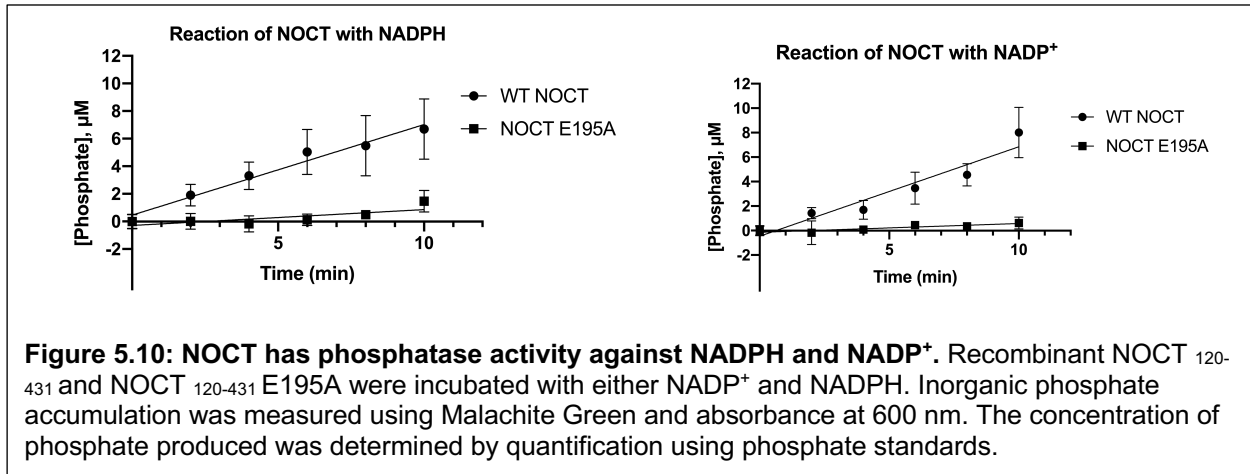
While GO terms in any of these four categories can yield insight into the regulatory role of NOCT, we considered the terms that are overrepresented in the set of genes that are downregulated with NOCT overexpression, as these transcripts are regulated in a

manner that would be consistent with the putative repressive activity of NOCT. Strikingly, the top five GO terms for the most significantly downregulated transcripts include the terms "voltage gated potassium channels," and "ionotropic glutamate receptor activity," which are associated with the downregulation of transcripts that are associated with function in neurons. Such transcripts include KCNQ5 and KCNQ2, which are voltage-gated potassium channels that are expressed in the brain and are implicated in neurological disorders (Lehman et al., 2017; Niday et al., 2017). These results provide new evidence that NOCT is capable of causing differential regulation of transcripts related to neurological function and demonstrate that the RNA-Seq analysis reported here may uncover new biological roles associated with NOCT regulation of mRNA abundance.

5.3.III NOCT has phosphatase activity against NADPH and NADP$^+$ in vitro.

During the preparation of this dissertation, a preprint manuscript reported that NOCT has activity against non-mRNA nucleotide metabolites, specifically NADP(H) (Estrella et al., 2019). This finding has significant implications, especially within the context of NOCT mitochondrial localization and its impact on metabolism. Additionally, the changes in mRNA steady state levels observed in our RNA-Seq data set could also potentially reflect downstream responses in gene expression as a consequence of NOCT phosphatase activity. Using our highly purified, recombinant NOCT, we have now corroborated the phosphatase activity against NADP$^+$ and NADPH (Figure 5.10). Using a Malachite Green phosphate assay we observed that free phosphate was produced in reactions with WT NOCT$_{120-431}$ but not in reactions with NOCT$_{120-431}$ E195A. Therefore, the effect of this mutation is consistent with *bona fide* NOCT enzymatic activity against nicotinamide dinucleotide substrates. These data, in conjunction with the ability of NOCT

to repress mRNAs in cell-based assays (discussed in Chapter 4), suggests that NOCT

may act as a multifunctional enzyme.

Figure 5.10: NOCT has phosphatase activity against NADPH and NADP$^+$. Recombinant NOCT $_{120-431}$ and NOCT $_{120-431}$ E195A were incubated with either NADP$^+$ and NADPH. Inorganic phosphate accumulation was measured using Malachite Green and absorbance at 600 nm. The concentration of phosphate produced was determined by quantification using phosphate standards.

5.4 Discussion

Early studies of NOCT classified the enzyme as a circadian deadenylase and suggested that the observed phenotypes in mice, namely resistance to diet-induced obesity, increased bone mass, and defects in lipid trafficking, were a result of NOCT mediated mRNA decay of associated metabolic transcripts (Douris et al., 2011; Green et al., 2007; Kawai et al., 2010a). The data presented here suggest a more complex model of NOCT mediated regulation, by which NOCT is predicted to be localized and processed in a tissue dependent manner and potentially affects metabolic function through regulation of both mRNA targets and metabolites such as NADP$^+$ and NADPH.

We observed that the NOCT N-terminus contains sequence motifs consistent with mitochondrial localization and processing by mitochondrial peptidases. In addition to these features in the NOCT protein sequence, the *NOCT* mRNA sequence has two translation initiation sites at Met1 and Met67, suggesting that leaky scanning may lead to expression of a short and a long form of NOCT (Figure 5.1). Using constructs with the

native *NOCT* 5´ leader sequence and ORF, we determined that in HCT116 and HEK293 cells, NOCT that is expressed from the Met1 initiation site is subsequently processed to form a ~41 kDa product (Figure 5.2). This processing is consistent with the predicted localization of NOCT to the mitochondria, which is supported by microscopy studies that observe NOCT localization in the mitochondria (Figure 5.2) (Estrella et al., 2019; Le et al., 2019). Fractionation experiments by Estrella, et al detected NOCT in the mitochondria and nucleus but not the cytoplasm, though nuclear localization was not evident by microscopy. This study assessed mitochondrial localization by using NOCT overexpression constructs in A549 cells, and it is possible in this case that overexpression of NOCT altered protein localization in this cell type or that in this specific cell type NOCT localizes strictly to the mitochondria (Estrella et al., 2019). While we also used NOCT expression constructs to analyze processing, there are several advantages to our analysis. First, we used cell lines that have low levels of endogenous NOCT and overexpressed levels of NOCT protein comparable to endogenous levels in cell lines such as HepG2. Second, we titrated transfected WT NOCT and demonstrated that processing still occurred at lower levels of expression (Figure 5.2 E). Third, we observed expression of endogenous NOCT isoforms consistent with expression of full-length and processed forms in multiple human and murine tissues, indicating that processing is tissue specific. (Figure 5.4). Further validation of mitochondrial localization using detection of endogenous NOCT protein will be needed to confirm localization of NOCT.

NOCT may play distinct regulatory roles in different tissues based on the observed differential processing of endogenous NOCT. The presence of the ~55 kDa isoform suggests that NOCT is cytoplasmic in tissues expressing this form and that NOCT would

regulate nuclear-encoded mRNA expression consistent with the changes in mRNA abundance observed when NOCT expression is perturbed (Figure 5.6) (Hee et al., 2012; Kojima et al., 2015; Stubblefield et al., 2018). In tissues where NOCT would be solely localized to the mitochondria, this would greatly restrict the pool of potential target mRNAs to 13 transcripts. However, the presence of the short form of NOCT is not necessarily an indication of localization to only the mitochondria, as Met67 translation initiation may occur more frequently in other cell types. It has not yet been determined if NOCT regulates mitochondrial mRNAs, and NOCT may also regulate non-mRNAs, as is the case for PDE12, which processes polyadenylated mt-tRNAs to ensure efficient translation (Pearce et al., 2017). Mitochondrial transcripts encode a subset of the subunits of the electron transport chain, and regulation of mt-RNAs could have profound effects on metabolism and cellular energy states (D'Souza and Minczuk, 2018). Additionally, NOCT localization to the cytoplasm and/or the mitochondria would target two largely separate NADP(H) pools, as these metabolites are not shuttled across the mitochondrial membrane (Nikiforov et al., 2011). Differential localization of NOCT indicates that the regulatory effects on mRNAs and NADP(H) pools are therefore likely to be tissue specific.

We addressed the potential regulatory effects of cytoplasmic NOCT (NOCT Δ2-15 3x FLAG) by measuring expression levels of nuclear-encoded mRNA transcripts when NOCT is overexpressed. Hundreds of mRNA transcripts were regulated (490 transcripts in total), which is the first evidence that NOCT overexpression in the cytoplasm changes mRNA abundances on a larger scale. 31 specific GO Terms were enriched in our dataset, suggesting that NOCT regulates specific biological processes. 235 of the regulated transcripts had decreased expression levels with NOCT Δ2-15 overexpression, which is
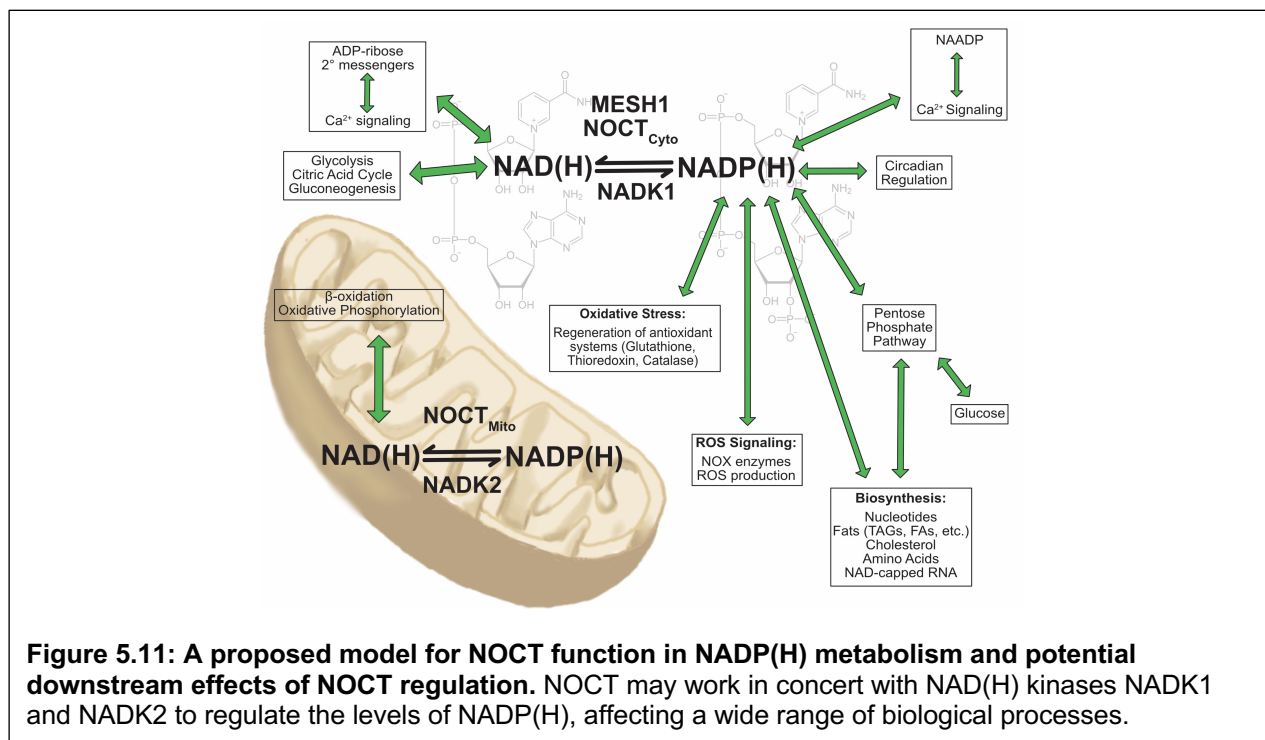
consistent with NOCT-mediated repression of these mRNAs. A subset of these transcripts may represent direct targets that are bound and repressed by NOCT, which must be validated through expressions such as RNA-Immunoprecipitation (RIP) to measure direct association of the mRNAs with NOCT. Determining enriched motifs in the mRNAs regulated by NOCT would potentially identify a sequence element that facilitates NOCT association with its direct targets and would aid in identification of other potential target mRNAs harboring this putative binding site. The activated mRNAs in our RNA-Seq data likely represent indirect targets that are affected by NOCT regulation of another target, such as a repressor. A fraction of the downregulated transcripts may also be indirect, if NOCT represses positive regulators of these mRNAs.

During validation of our RNA-Seq results, we examined the steady state protein levels for three of our top regulated transcripts, *FRRS1L, PTPRZ1,* and *RORβ.* Contrary to the downregulation of these mRNAs, the protein levels did not decrease with NOCT overexpression (Figure 5.8) This observation may be indicate that these proteins are very stable, or that the reduction in mRNA levels was not sufficient to significantly impact translation. Additionally, these three transcripts represent a very small sample of the 235 downregulated transcripts and decreased steady state levels of protein may be more evident for other genes in our dataset. Analysis of NOCT-mediated effects on mRNA levels may underestimate its impact on the proteome, given our observation of translational repression in tethered function assays (Figures 4.3 and 4.7). Therefore, it is possible that a subset of the transcripts in our dataset could be regulated by NOCT-mediated translational repression, such that changes in abundance at the mRNA level could be more modest than the changes observed at the protein level.

We found that multiple GO terms overrepresented in the genes downregulated with overexpression of NOCT Δ2-15 represented neurological functions (Table 5.4). While this is likely in part due to the neuro-adrenal phenotype of the HEK293 cells used in this study, it is possible that NOCT regulates transcripts in the brain, as endogenous NOCT appears to be expressed in both human and mouse brain tissue lysates (Figure 5.4). Neurological phenotypes have not been reported in NOCK KO models to determine its effects on gene expression and metabolism in brain tissue. There is precedence for mechanisms of post-transcriptional control in neurons effecting energy homeostasis; for example, in POMC neurons, Dicer activity and miRNA biogenesis is crucial for cell survival and POMC-mediated regulation of melanocortin signaling, which regulates metabolism and insulin signaling (Cansell and Luquet, 2012; Derghal et al., 2017; Schneeberger et al., 2012). The role of NOCT regulation in the brain should be further investigated using tools such as brain-specific knockout models in order to determine the neurological phenotypes of NOCT regulation.

New insights into non-mRNA targets of NOCT provide another layer of complexity in interpreting the changes in gene expression when NOCT is overexpressed. Consistent with a recent report of NOCT activity against nicotinamide adenine dinucleotides, we observed that NOCT has phosphatase activity against $NADP^+$ and NADPH (Figure 5.10) (Estrella et al., 2019). It is therefore important to evaluate the evidence for and against NOCT regulation of mRNAs to assess whether NOCT may have multiple functions. NOCT has phosphatase activity active against NADP(H) *in vitro* and represses reporter mRNA repression *in vivo* (Figure 4.2 and Figure 5.10). NOCT may be able to accommodate both mRNA and nicotinamide adenine dinucleotides in its active site, as the plasticity of EEP

deadenylase active sites has been shown to allow for binding of chemically diverse molecules (discussed in Chapter 2) (Wang et al., 2010; Wood et al., 2015; Zhang et al., 2016). NOCT activity in tethered function assays is perturbed by mutation of the NOCT active site and is sensitive to the sequence and/or structure of the mRNA 3´end, suggesting that NOCT can function as a ribonuclease and has RNA substrate preferences in the cellular context (Figures 4.2 and 4.6)

Figure 5.11: A proposed model for NOCT function in NADP(H) metabolism and potential downstream effects of NOCT regulation. NOCT may work in concert with NAD(H) kinases NADK1 and NADK2 to regulate the levels of NADP(H), affecting a wide range of biological processes.

Understanding if and/or how NOCT affects cellular metabolism through the regulation of NADP(H) is a complex question, as these metabolites are used in a large range of biological functions ranging from reactive oxygen species (ROS) signaling to fatty acid biosynthesis (Figure 5.11). In this model, mitochondrial NOCT and the NAD(H) kinase NADK2 regulate the relative levels of NADP(H) and NAD(H), while NOCT and MESH1 would oppose the activity of NADK1 in the cytoplasm (Ding et al., 2018; Kawai and Murata, 2008; Ohashi et al., 2012). Metabolic changes that result from changing

nicotinamide dinucleotide levels may cause downstream affects that alter gene expression at multiple levels. This is an important consideration in interpreting the phenotypes observed when NOCT expression levels are altered in overexpression or loss-of function models. The changes observed in RNA-Seq experiments may indicate NOCT-mediated repression of mRNA abundance and/or translation, or through processes responding to NOCT phosphatase activity against NADP(H), such as differential transcription, mRNA processing, or NOCT-independent mRNA decay.

The obesity-resistant phenotype of the NOCT KO mice can be at least partly explained by the putative NOCT phosphatase activity in tissues. Loss of NADPH oxidase NOX4 sensitizes mice to diet-induced obesity and fatty liver disease, which are phenotypes observed in NOCT KO mice. Loss of NOCT would potentially provide more NADPH as a substrate for NOX4, increasing ROS signaling pathways that mediate glucose uptake into adipocytes and repress adipogenesis (Green et al., 2007; Li et al., 2012). Therefore, the resistance to obesity in NOCT KO mice could be partially explained by increased NOX4 signaling. Not all reported effects observed in the NOCT KO mice can be attributed to the increase in NADP(H) levels predicted with the loss of NOCT phosphatase activity. For example, NOCT KO would be predicted to increase the pool of cytoplasmic NADPH, potentially increasing efficiency of fatty acid synthesis. Despite predicted increases in NADP(H) levels, the mass of fat pads and hepatic fat deposits are decreased. Furthermore, on an HFD, the conditions where loss of NOCT has the most significant effects on deceasing fat deposition in adipose and liver tissues, fatty acid synthesis is downregulated (Currie et al., 2013; Green et al., 2007; Parks and Parks, 2002). While a loss of NOCT function is expected to decrease NAD(H) levels, *increased*

NAD$^+$ and NADH is correlated with the same phenotypes seen in NOCT KO models, i.e., resistance to diet induced obesity. In one report, treatment of mice with an NAD$^+$ precursor showed that increased levels of NAD$^+$ are negatively correlated with HFD-induced obesity and mitochondrial dysfunction (Canto et al., 2012). NOCT KO mice would be expected to have decreased NAD$^+$, which could potentially sensitize mice to the effects of HFD and could decrease fatty acid oxidation (Currie et al., 2013). NOCT phenotypes related to bone mass also conflict with the expected effect of reduced NADPH levels. Loss of NADPH oxidase activity from NOX2 has been reported to protect against loss of bone mass in mice fed an HFD (Rahman et al., 2018). Therefore, loss of NOCT expression would be expected to increase the amount of NADPH cofactor for NOX2, resulting in a potential increase in activity that would promote loss of bone mass. In contrast, NOCT KO mice have increased bone mass relative to WT mice (Kawai et al., 2010b). The NOCT phenotypes that do not appear to be attributable to altered NADP(H) metabolism may instead be more indicative of NOCT regulation of mRNA transcripts related to processes such as osteoblastogenesis and lipid metabolism. Ultimately, we cannot yet extrapolate NADP(H) phosphatase activity to the NOCT knockout phenotypes, as the effect of NOCT on NADP(H) has not been determined in cells or tissues. NOCT regulation of both mRNAs and NADP(H) levels remain possible mechanisms by which NOCT modulates metabolism, and it is possible that of both play a role in NOCT-mediated regulation. Understanding the relative contributions of NOCT ribonuclease and phosphatase activities to NOCT-mediated regulation will be important for fully understanding NOCT regulation in the cellular context.

5.5 Materials and Methods

Cloning of plasmids

The pFC3F plasmid backbone was generated from the pF5A vector (Promega) to generate C-terminal 3x FLAG tag fusions. GST and NOCT$_{1-431}$ were restriction cloned into pFC3F. The NOCT Δ2-15 3x FLAG pFC3F construct was generated by inverse PCR. For the cloning of the NOCT constructs with N-terminal point mutants, pCR4 TOPO NOCT was purchased from Open Biosystems and point mutants were generated using quikchange PCR to generate pCR4 TOPO constructs of NOCT$_{1-431}$ M1X and NOCT$_{1-431}$ M67G and NOCT$_{1-431}$ R15M. Inverse PCR was used to generate constructs NOCT Δ2-15 and NOCT Δ2-15 M67G. PCR fragments and the pF5A backbone were digested and ligated using the Flexi Cloning System (Promega, C8640) according to the manufacturer's instructions. The NOCT$_{1-86}$ EGFP plasmid was generated by PCR amplification of the EGFP ORF from the pEGFPN1 vector (Clontech) with appended 5´ AfeI and 3´ EcoRI restriction sites. The PCR products and pF5A NOCT were restriction digested and the PCR amplicon was used to replace the AfeI – EcoRI fragment of NOCT$_{1-431}$ pF5A. Cloning was performed by K. Hughes.

Generation of α-NOCT Antibody

The α-NOCT polyclonal antibody was produced in rabbits inoculated with MBP-NOCT$_{1-431}$. Briefly, MBP-NOCT$_{1-431}$ was recombinantly expressed in BL21 DE3 cells and purified over amylose resin. Crude serum was verified for reactivity against HEK293 cell extracts overexpressing MS2-NOCT$_{1-431}$. Briefly, HEK293 cells were transfected with 3 µg of MS2-NOCT$_{1-431}$ pFN21a vector or HaloTag-MS2 in pFN21a using using a 3 µL:1 ng ratio of FuGene HD (Promega). Cells were harvested, pelleted, and lysed. 10 µg of each

sample were resolved using SDS PAGE and then transferred to Millipore EMD Immobilon membrane. Membranes were probed with a 1:1000 dilution of either pre-immune serum or terminal bleed serum and HRP conjugated goat α-mouse IgG secondary antibody (Sigma). Blots were visualized using Pierce ECL Western blotting substrate before exposure to autoradiography film. Verified serum was purified using ammonium sulfate fractionation and antigen affinity purification using recombinantly expressed HaloTag-NOCT$_{1-431}$ immobilized on HaloLink resin (Promega). The serum was purified over the immobilized NOCT column before elution with 5 M NaI and 1 mM ammonium thiocyanate. Antibody was subsequently dialyzed into 1x PBS (137 mM NaCl, 27 mM KCl, 10 mM Na$_2$HPO$_4$, and 18 mM KH$_2$PO$_4$) for 1 hour at 4 °C. Dialysis was repeated two additional times with the final dialysis step performed overnight. Insoluble material was removed by centrifugation at 18,000 xg and the antibody was concentrated to 1 mg/mL using a Centricon spin concentrator. Purified polyclonal antibody was verified for reactivity against recombinant purified NOCT$_{64-431}$. MBP-NOCT expression and purification was performed by N. Raynard. Serum testing was performed by E. Abshire. HaloTag NOCT expression and purification, NOCT antibody purification, and antibody purification performed by K. Hughes.

Analysis of Human and Mouse Tissue blots

Pre-transferred blots of tissue extract panels from human and mouse were purchased from Novus Biologicals (NBP2-31378 and NBP2-20111). Blots were hydrated and washed before probing with α-NOCT antibody. Blots were probed using HRP conjugated goat α-mouse IgG secondary antibody (Sigma) and Pierce ECL Western

blotting substrate before imaging on a Chemidoc Touch (BioRad). Analysis was performed by K. Hughes.

Analysis of Differential NOCT processing by Western blotting

HEK293 cells were cultured in DMEM (Fisher, 11965118), 1% Pen/ Strep/ Glutamine (Fisher, 10378016) and 10% FBS (Thermo, 26140079) at 37 °C and 5% CO_2. Cells were transfected with NOCT N-terminal mutant constructs using FugeneHD (Promega, E2312) in a ratio of 3 µL transfection reagent per 1 µg DNA. HCT116 were cultured as HEK 293 cells but in McCoy 5A (Fisher, 16600082), 1% Pen/ Strep/ Glutamine and 10% FBS. Cells were transfected using FugeneHD in a ratio of 4 µL transfection reagent per 1 µg DNA. Both HEK293 and HCT116 cells were harvested, pelleted, and resuspended in modified radioactive immunoprecipitation assay buffer (mRIPA buffer, 10 mM Tris (pH 8.0), 140 mM NaCl, 1 mM Diaminoethane-tetraacetic acid (EDTA), 0.5 mM Triethylene glycol diamine tetraacetic acid (EGTA),1% (v/v) Triton X-100, 0.1% (w/v) sodium deoxycholate, and 0.1% (w/v) sodium dodecyl sulphate) for lysis. Lysates were homogenized using a handheld cell disruptor and sample concentrations were determined using the BioRad DC Lowry assay. Each sample was resolved using SDS PAGE and then transferred to Millipore EMD Immobilon membrane. Blots were probed using the following primary antibodies: α-NOCT (detecting NOCT MTS mutants expressed from pF5A), α-EGFP (detecting EGFP and $NOCT_{1-86}$ EGFP, Clontech clone JL-8, 632380), and α-FLAG (detecting C-terminal 3x FLAG-tagged constructs). Blots were probed using HRP-conjugated Secondary Antibody (Sigma A0545 for α-NOCT blots, Sigma A1047 for α-EGFP and α-FLAG blots) and Millipore Immobilon Western Blotting substrate before imaging on a Chemidoc Touch (BioRad). As a loading control,

Vinculin was detected on the same blots using anti-Vinculin antibody (Fisher, 700062).

Analysis was performed by K. Hughes.

Generation of Cell lines

To produce stably transfected HEK293 cell lines (ATCC), cells were transfected with a 1:10 molar ratio of pEF6 HisA (Invitrogen) to either GST 3x FLAG pFC3F or NOCT Δ2-15 3x FLAG pFC3F using FugeneHD (Promega, E2312). 5 μg/mL Blasticidin S (Thermo Fisher, A1113903) was used to select Blasticidin resistant populations that were then subcloned using the dilution method. Blasticidin resistant clones expressing 3x FLAG-tagged fusion proteins for each construct were identified by screening clones using anti-FLAG Western blots. Cells were harvested in 1x PBS, pelleted, and frozen. Prior to use, cells were resuspended in mRIPA supplemented with protease inhibitor (Millipore Sigma, 5892791001) for lysis. Lysates were homogenized using a handheld cell disruptor and sample concentrations were determined using the BioRad DC Lowry assay. Each sample (10 μg total protein) was resolved on 4-20% TGX gradient gels (BioRad) and then transferred to Millipore EMD Immobilon membrane. Blots were probed using α-FLAG (Millipore Sigma F3165). Blots were probed using HRP-conjugated Secondary Antibody (Sigma, A1047) and Pierce ECL Western Blotting substrate before imaging on a Chemidoc Touch (BioRad). Cell lines were generated and validated by K. Hughes.

RNA-Seq Sample Preparation

For each clonal cell line, two 10 cm dishes were plated with $3x10^6$ cells in DMEM + 10% FBS and incubated for three days at 37 °C and 5% CO_2. Cells were harvested by dissociation with TrypLE (ThermoFisher Scientific; 12605028) and were washed twice

with cold 1x PBS, pH 7.4 (Gibco, 10010023) by resuspension and centrifugation at 3,000 xg for 5 min at room temp. Cells were resuspended in cold PBS and pelleted by centrifugation at 4°C for 5 minutes at 21,100 xg before storage at -80 °C. FLAG expression was verified by anti-FLAG western blot as described above. Total RNA was isolated from frozen cell pellets using a Maxwell RSCSimplyRNA Cells kit (Promega; AS1390) using a 2x concentration of DNAse I. RNA was eluted in 50 µL nuclease-free water. Total RNA (5 µg) was combined with ERCC Spike-In Standards (5 µl of 1:50 diluted stock solution; Invitrogen, Carlsbad, CA; 4456700) and submitted to the University of Minnesota Genomics Center for library generation and sequencing. TruSeq strand-specific 50 bp dual indexed RNA libraries were prepared using RiboZero rRNA depletion and sequenced on a HiSeq 2500 in High-output mode. Sample preparation was performed by K. Hughes. Library generation and sequencing was performed by the University of Minnesota Genomics Center.

RNA-Seq methods

RNA-Seq data was analyzed for data quality using FastQC and MultiQC (Andrews, 2010; Ewels et al., 2016). Due to overrepresentation of adaptor sequences in the sequencing reads, the first 10 bases were removed from each read using Cutadapt (version 1.8.1) followed by Trimmomatics (version 0.33) to remove low quality scores at the ends of each read and to exclude reads that are less than 20 base pairs after trimming (Bolger et al., 2014; Martin, 2011). Paired survival rates for all six samples were determined to be >99% after reads were filtered and trimmed.   STAR (version 2.5.2b) was used to build an annotated genome index starting with the GRCh38 human genome and incorporating the NOCT Δ2-15 3x FLAG pFC3F and GST 3X FLAG pFC3F plasmid

sequences and the ERCC spike-in controls. Reads were subsequently aligned to the genome with transcriptome guidance (Dobin et al., 2013; Dobin and Gingeras, 2015, 2016). Gene level quantification was performed using HTSeq to obtain a raw count of reads per gene (Anders et al., 2015). We next normalized raw counts and performed differential gene expression analysis using DESeq2 (Love et al., 2014). Once we determined a dataset of genes differentially regulated in NOCT Δ2-15 3x FLAG versus GST 3x FLAG samples, we used iPAGE to determine which GO Terms were overrepresented and underrepresented in the differentially expressed genes (Goodarzi et al., 2009). For transcript-level analysis, we quantified the transcripts using kallisto (version 0.43.0) and a reference transcriptome built from cDNAs and ncRNAs from GRCh38 and the transcript sequences for NOCT Δ2-15 3x FLAG and GST 3x FLAG (Bray et al., 2016a, b). Quantification was then analyzed using sleuth (version 0.29.0), including normalization and a Wald test to approximate Log2 fold changes (Pimentel et al., 2017). Data visualization was done in R using ggplot2 (Wickham, 2016). RNA-Seq data analysis was performed by R. Diao and P. Freddolino.

Reverse Transcription and qPCR.

Three replicate samples of each of the clonal cell lines produced for RNA-Seq were grown in T75 flasks at 37C and 5% CO_2. Confluent flasks were trypsinized with TrypLE (Thermo, 12605028) and cells were resuspended in 10 mL DMEM + 10% FBS (Thermo, 16000044) 2 aliquots of cell suspensions, 3 mL each, were collected by centrifugation at 1000 x g. Supernatants were removed and cell pellets were stored at -80 °C until further use. RNA was harvested from one cell pellet from each cell line using the Maxwell RSC SimplyRNA Cells Kit (Promega). Briefly, cell pellets were resuspended in 200 μL of

Maxwell Lysis buffer and then mixed with 200 µL of Maxwell Homogenization buffer with 1-Thioglycerol. Lysates were added to the Maxwell RSC SimplyRNA Cells Kit cartridges along with a 2x concentration of Maxwell DNaseI solution per cartridge. The Maxwell RSC SimplyRNA Cells protocol was run without modification on a Maxwell RSC Instrument (Promega) and RNA was eluted in 40 µL of RNase free water and was stored at -80 °C.

RNA was quantified using absorbance at 260 nm on a Nanodrop One UV spectrophotometer (Thermo Fisher). Reverse transcription reactions were carried out using 500 ng of total RNA from each sample and the GoScript Reverse Transcription Mix, Random Primers Kit (Promega, A2800) according to the manufacturer's protocol. For each sample, control reactions without reverse transcriptase (RT) were generated to assess background from other sources of DNA. RT+ and RT- reverse transcription reactions were diluted 67-fold in RNase/DNase free water and RT-qPCR was performed on the CFX96 Real Time System (BioRad) using GoTaq qPCR Master Mix (Promega, A6002). Cycling parameters were as follows: 1) 95 °C for 2 min, 2) 95 °C for 15 s, 3) variable annealing temperatures for 1 min, 4) 72 °C for 1 min. Steps 2-4 were repeated for a total of 40 cycles. Control reactions without template and without RT were performed for all primer sets. For each transcript, PrimeTime Assays (IDT) were used in qPCR experiments. The primers and annealing temperatures used are listed in Table 5.5.

Gene	PrimeTime Assay	Annealing Temperature	Primer Efficiency
DLGAP1	Hs.PT.58.22742606	60 °C	91%
RORB	Hs.PT.58.39397316	65 °C	89%
TMEFF2	Hs.PT.58.15573952	65 °C	91%
PTPRZ1	Hs.PT.58.27038488	62 °C	93%
FRRS1L	Hs.PT.58.15557564	65 °C	92%
NOCT	Hs.PT.58.21201364	63.5 °C	92%
TBP	Hs.PT.58v.39858774	61 °C	97%

Table 5.5: Primers used in RT-qPCR experiments

Annealing temperatures for each primer set were optimized to 90-110 % efficiency at 200 nM primer concentration. Cycle threshold (Ct) values were measured using the CFX Manager software and analyzed using the Pfaffl method (Pfaffl, 2001) NOCT overexpression was additionally verified using RT-qPCR. For each cell line, the Ct values for the three experimental replicates were averaged to generate the values $Ct_{NOCT, clonal}$ $_{average}$ and $Ct_{GST, clonal\ average}$. For analysis, the Ct values for each set of clonal cell lines were averaged to generate the values $Ct_{NOCT, total\ average}$ and $Ct_{GST, total\ average}$. Fold change was calculated for each individual clonal cell line using the following:

$$\Delta Ct = Ct_{GST, total\ average} - Ct_{clonal\ average}$$

So that ΔCt values for the target and TATA binding protein (TBP) control are calculated for each clonal cell line, generating the values ΔCt_{target} and ΔCt_{TBP}. E values were calculated using the following equation, where the slope is the linear fit produced when graphing Ct values for a 3-fold dilution series of cDNA in the qPCR reaction:

$$E = 10^{-1/slope}$$

Fold changes for each clonal cell line were calculated using the following equation:

$$FC = \frac{E_{target}^{\Delta Ct, target}}{E_{TBP}^{\Delta Ct, TBP}}$$

Western Blotting for RNA-Seq Targets.

Cells harvested from the three clonal cell lines overexpressing GST 3x FLAG or NOCT Δ2-15 3x FLAG were harvested in 1x PBS, pelleted, and resuspended in radioactive immunoprecipitation assay buffer (RIPA buffer, 50 mM Tris (pH 8.0), 150 mM NaCl, 1% (v/v) IGEPAL CA-630, 0.5% (w/v) sodium deoxycholate, 0.1% (w/v) sodium dodecyl sulphate) for lysis. Lysates were homogenized using a handheld cell disruptor and sample

158

concentrations were determined using the BioRad DC Lowry assay. Each sample (10 µg total protein for analysis of FRRS1L and PTPRZ1, 25 µg total protein for analysis of RORB) was resolved on 4-20% TGX gradient gels (BioRad) and then transferred to Millipore EMD Immobilon membrane. Blots were probed using the following primary antibodies: α-PTPRZ1 (R&D, MAB2688), α-FRRS1L (Abcam, ab98878), and α-RORB (Abcam, ab187657). Blots were probed using HRP-conjugated Secondary Antibody (Sigma A0545 for FRRS1L and RORB blots, and CST 7077 for PTPRZ1 blots) and Millipore Immobilon Western Blotting substrate before imaging on a Chemidoc Touch (BioRad). As a loading control, Vinculin was detected on the same blots using anti-Vinculin antibody (Fisher, 700062).

Purification of Recombinant NOCT for Malachite Green Assays

The Strep(II)- $NOCT_{120-431}$ pSumo expression construct was cloned previously (Chapter 2) (Abshire et al., 2018). Briefly, the $Nocturnin_{120-431}$ was PCR amplified with BamHI and XhoI overhangs and subcloned into BamHI and XhoI digested pSumo. The $NOCT_{120-431}$ E195A mutant was generated using quikchange PCR to generate the point mutation. Strep(II)-Sumo-$NOCT_{120-431}$ and Strep(II)-Sumo-$NOCT_{120-431}$ E195A were expressed using a Strep(II) pSumo vector in BL21 STAR (DE3) E. coli cells (ThermoFisher) expressing the pRare2 plasmid (MilliporeSigma) and purified in 50 mM Tris, pH 8.0, 75 mM Na_2SO_4, 10% glycerol, 5 mM βME, and 4 mM $MgCl_2$, with the addition of a 1.0 M NaCl wash (in 50 mM Tris, pH 8.0, 75 mM Na_2SO_4, 1 M NaCl, 10% glycerol, 5 mM βME, and 4 mM $MgCl_2$) after the addition of lysate onto the Strep-Tactin Superflow Plus resin. All fractions were then pooled and the Strep (II) Sumo tag was cleaved overnight at 4 °C in Snake Skin 10,000 MW cutoff dialysis tubing (Thermo) with

Ulp1 Sumo protease in 50 mM Tris, pH 8.0, 75 mM Na_2SO_4, 10% glycerol, 5 mM βME, and 4 mM $MgCl_2$. $NOCT_{120-431}$ and $NOCT_{120-431}$ E195A were then concentrated to ~6 mg/mL and stored after the Strep-Tactin purification step.

Malachite green NADP and NADPH phosphatase assays

NOCT phosphatase activity assays against NAD and NADPH were conducted in triplicate using a Malachite Green Phosphate Assay Kit (Cayman Chemical 10009325) and a modified protocol from the manufacturer's instructions. Reactions were performed using 50 nM $NOCT_{120-431}$ or $NOCT_{120-431}$ E195A and 1 mM of either NADPH (Roche, 10107824001) or NADP (Roche, 10128031001) incubated in 20 mM Tris (pH 8.0), 70 mM NaCl, and 2.0 mM $MgCl_2$ for 10 minutes. Reactions were stopped with the MG Acid Solution (Cayman Chemical) and precipitates were cleared by centrifugation of reactions at 18,000xg for 5 minutes. The resulting concentrations of phosphate were measured using a Malachite Green Assay (Cayman Chemical) and absorbance at 600 nm was measured using a GloMax Discover plate reader (Promega). Phosphate concentrations were quantified using a phosphate standard curve.

5.6 References

Abshire, E.T., Chasseur, J., Bohn, J.A., Del Rizzo, P.A., Freddolino, P.L., Goldstrohm, A.C., and Trievel, R.C. (2018). The structure of human Nocturnin reveals a conserved ribonuclease domain that represses target transcript translation and abundance in cells. Nucleic acids research *46*, 6257-6270.

Anders, S., Pyl, P.T., and Huber, W. (2015). HTSeq--a Python framework to work with high-throughput sequencing data. Bioinformatics *31*, 166-169.

Andrews, S. (2010). FastQC: a quality control tool for high throughput sequence data.

Baggs, J.E., and Green, C.B. (2003). Nocturnin, a deadenylase in Xenopus laevis retina: a mechanism for posttranscriptional control of circadian-related mRNA. Current biology : CB *13*, 189-198.

Bolger, A.M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics *30*, 2114-2120.

Bray, N.L., Pimentel, H., Melsted, P., and Pachter, L. (2016a). Erratum: Near-optimal probabilistic RNA-seq quantification. Nat Biotechnol *34*, 888.

Bray, N.L., Pimentel, H., Melsted, P., and Pachter, L. (2016b). Near-optimal probabilistic RNA-seq quantification. Nat Biotechnol *34*, 525-527.

Cansell, C., and Luquet, S. (2012). Hypothalamic regulation of energy balance: a key role for DICER miRNA processing in arcuate POMC neurons. Mol Metab *2*, 55-57.

Canto, C., Houtkooper, R.H., Pirinen, E., Youn, D.Y., Oosterveer, M.H., Cen, Y., Fernandez-Marcos, P.J., Yamamoto, H., Andreux, P.A., Cettour-Rose, P.*, et al.* (2012). The NAD(+) precursor nicotinamide riboside enhances oxidative metabolism and protects against high-fat diet-induced obesity. Cell Metab *15*, 838-847.

Currie, E., Schulze, A., Zechner, R., Walther, T.C., and Farese, R.V., Jr. (2013). Cellular fatty acid metabolism and cancer. Cell Metab *18*, 153-161.

D'Souza, A.R., and Minczuk, M. (2018). Mitochondrial transcription and translation: overview. Essays Biochem *62*, 309-320.

Derghal, A., Djelloul, M., Trouslard, J., and Mounien, L. (2017). The Role of MicroRNA in the Modulation of the Melanocortinergic System. Front Neurosci *11*, 181.

Ding, C.C., Rose, J., Wu, J., Sun, T., Chen, K., Chen, P., Xu, E., Tian, S., Akinwuntan, J., Guan, Z.*, et al.* (2018). Mammalian stringent-like response mediated by the cytosolic NADPH phosphatase MESH1. Biorxiv [Preprint].

Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics *29*, 15-21.

Dobin, A., and Gingeras, T.R. (2015). Mapping RNA-seq Reads with STAR. Curr Protoc Bioinformatics *51*, 11 14 11-19.

Dobin, A., and Gingeras, T.R. (2016). Optimizing RNA-Seq Mapping with STAR. Methods in molecular biology *1415*, 245-262.

Douris, N., Kojima, S., Pan, X., Lerch-Gaggl, A.F., Duong, S.Q., Hussain, M.M., and Green, C.B. (2011). Nocturnin regulates circadian trafficking of dietary lipid in intestinal enterocytes. Current biology : CB *21*, 1347-1355.

Estrella, M.A., Du, J., Chen, L., Rath, S., Prangley, E., Chitrakar, A., Aoki, T., Schedl, P., Rabinowitz, J., and Korennykh, A. (2019). The Metabolites NADP+ and NADPH are the Targets of the Circadian Protein Nocturnin (Curled). Biorxiv [Preprint].

Ewels, P., Magnusson, M., Lundin, S., and Kaller, M. (2016). MultiQC: summarize analysis results for multiple tools and samples in a single report. Bioinformatics *32*, 3047-3048.

Fukasawa, Y., Tsuji, J., Fu, S.C., Tomii, K., Horton, P., and Imai, K. (2015). MitoFates: improved prediction of mitochondrial targeting sequences and their cleavage sites. Mol Cell Proteomics *14*, 1113-1126.

Goodarzi, H., Elemento, O., and Tavazoie, S. (2009). Revealing global regulatory perturbations across human cancers. Mol Cell *36*, 900-911.

Green, C.B., Douris, N., Kojima, S., Strayer, C.A., Fogerty, J., Lourim, D., Keller, S.R., and Besharse, J.C. (2007). Loss of Nocturnin, a circadian deadenylase, confers resistance to hepatic steatosis and diet-induced obesity. Proceedings of the National Academy of Sciences of the United States of America *104*, 9888-9893.

Hee, S.W., Tsai, S.H., Chang, Y.C., Chang, C.J., Yu, I.S., Lee, P.C., Lee, W.J., Yun-Chia Chang, E., and Chuang, L.M. (2012). The role of nocturnin in early adipogenesis and modulation of systemic insulin resistance in human. Obesity *20*, 1558-1565.

Kawai, M., Green, C.B., Horowitz, M., Ackert-Bicknell, C., Lecka-Czernik, B., and Rosen, C.J. (2010a). Nocturnin: a circadian target of Pparg-induced adipogenesis. Annals of the New York Academy of Sciences *1192*, 131-138.

Kawai, M., Green, C.B., Lecka-Czernik, B., Douris, N., Gilbert, M.R., Kojima, S., Ackert-Bicknell, C., Garg, N., Horowitz, M.C., Adamo, M.L.*, et al.* (2010b). A circadian-regulated gene, Nocturnin, promotes adipogenesis by stimulating PPAR-gamma nuclear translocation. Proceedings of the National Academy of Sciences of the United States of America *107*, 10508-10513.

Kawai, S., and Murata, K. (2008). Structure and function of NAD kinase and NADP phosphatase: key enzymes that regulate the intracellular balance of NAD(H) and NADP(H). Biosci Biotechnol Biochem *72*, 919-930.

Kojima, S., Gendreau, K.L., Sher-Chen, E.L., Gao, P., and Green, C.B. (2015). Changes in poly(A) tail length dynamics from the loss of the circadian deadenylase Nocturnin. Scientific reports *5*, 17059.

Le, P.T., Bornstein, S.A., Motyl, K.J., Tian, L., Stubblefield, J.J., Hong, H.K., Takahashi, J.S., Green, C.B., Rosen, C.J., and Guntur, A.R. (2019). A novel mouse model overexpressing Nocturnin results in decreased fat mass in male mice. J Cell Physiol.

Lehman, A., Thouta, S., Mancini, G.M.S., Naidu, S., van Slegtenhorst, M., McWalter, K., Person, R., Mwenifumbo, J., Salvarinova, R., Study, C.*, et al.* (2017). Loss-of-Function and Gain-of-Function Mutations in KCNQ5 Cause Intellectual Disability or Epileptic Encephalopathy. Am J Hum Genet *101*, 65-74.

162

Li, Y., Mouche, S., Sajic, T., Veyrat-Durebex, C., Supale, R., Pierroz, D., Ferrari, S., Negro, F., Hasler, U., Feraille, E.*, et al.* (2012). Deficiency in the NADPH oxidase 4 predisposes towards diet-induced obesity. Int J Obes (Lond) *36*, 1503-1513.

Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome biology *15*, 550.

Martin, M. (2011). Cutadapt Removes Adapter Sequences From High-Throughput Sequencing Reads. EMBnetjournal *17*.

Niday, Z., Hawkins, V.E., Soh, H., Mulkey, D.K., and Tzingounis, A.V. (2017). Epilepsy-Associated KCNQ2 Channels Regulate Multiple Intrinsic Properties of Layer 2/3 Pyramidal Neurons. J Neurosci *37*, 576-586.

Nikiforov, A., Dolle, C., Niere, M., and Ziegler, M. (2011). Pathways and subcellular compartmentation of NAD biosynthesis in human cells: from entry of extracellular precursors to mitochondrial NAD generation. The Journal of biological chemistry *286*, 21767-21778.

Nishikawa, S., Hatanaka, Y., Tokoro, M., Shin, S.W., Shimizu, N., Nishihara, T., Kato, R., Takemoto, A., Amano, T., Anzai, M.*, et al.* (2013). Functional analysis of nocturnin, a circadian deadenylase, at maternal-to-zygotic transition in mice. J Reprod Dev *59*, 258-265.

Ohashi, K., Kawai, S., and Murata, K. (2012). Identification and characterization of a human mitochondrial NAD kinase. Nature communications *3*, 1248.

Parks, E.J., and Parks, E.J. (2002). Changes in fat synthesis influenced by dietary macronutrient content. Proc Nutr Soc *61*, 281-286.

Pearce, S.F., Rorbach, J., Van Haute, L., D'Souza, A.R., Rebelo-Guiomar, P., Powell, C.A., Brierley, I., Firth, A.E., and Minczuk, M. (2017). Maturation of selected human mitochondrial tRNAs requires deadenylation. Elife *6*.

Pfaffl, M.W. (2001). A new mathematical model for relative quantification in real-time RT-PCR. Nucleic acids research *29*, e45.

Pimentel, H., Bray, N.L., Puente, S., Melsted, P., and Pachter, L. (2017). Differential analysis of RNA-seq incorporating quantification uncertainty. Nat Methods *14*, 687-690.

Rahman, M.M., El Jamali, A., Halade, G.V., Ouhtit, A., Abou-Saleh, H., and Pintus, G. (2018). Nox2 Activity Is Required in Obesity-Mediated Alteration of Bone Remodeling. Oxid Med Cell Longev *2018*, 6054361.

Schmidt, O., Pfanner, N., and Meisinger, C. (2010). Mitochondrial protein import: from proteomics to functional mechanisms. Nature reviews Molecular cell biology *11*, 655-667.

Schneeberger, M., Altirriba, J., Garcia, A., Esteban, Y., Castano, C., Garcia-Lavandeira, M., Alvarez, C.V., Gomis, R., and Claret, M. (2012). Deletion of miRNA processing enzyme Dicer in POMC-expressing cells leads to pituitary dysfunction, neurodegeneration and development of obesity. Mol Metab *2*, 74-85.

Stubblefield, J.J., Gao, P., Kilaru, G., Mukadam, B., Terrien, J., and Green, C.B. (2018). Temporal Control of Metabolic Amplitude by Nocturnin. Cell reports *22*, 1225-1235.

Wang, H., Morita, M., Yang, X., Suzuki, T., Yang, W., Wang, J., Ito, K., Wang, Q., Zhao, C., Bartlam, M.*, et al.* (2010). Crystal structure of the human CNOT6L nuclease domain reveals strict poly(A) substrate specificity. The EMBO journal *29*, 2566-2576.

Wang, Y., Osterbur, D.L., Megaw, P.L., Tosini, G., Fukuhara, C., Green, C.B., and Besharse, J.C. (2001). Rhythmic expression of Nocturnin mRNA in multiple tissues of the mouse. BMC developmental biology *1*, 9.

Wickham, H. (2016). ggplot2: Elegant Graphics for Data Analysis. Spriger- Verlag New York.

Wood, E.R., Bledsoe, R., Chai, J., Daka, P., Deng, H., Ding, Y., Harris-Gurley, S., Kryn, L.H., Nartey, E., Nichols, J.*, et al.* (2015). The Role of Phosphodiesterase 12 (PDE12) as a Negative Regulator of the Innate Immune Response and the Discovery of Antiviral Inhibitors. The Journal of biological chemistry *290*, 19681-19696.

Zhang, Q., Yan, D., Guo, E., Ding, B., Yang, W., Liu, R., Yamamoto, T., and Bartlam, M. (2016). Structural basis for inhibition of the deadenylase activity of human CNOT6L. FEBS Lett *590*, 1270-1279.

CHAPTER 6

Concluding Remarks and Future Directions

6.1 Summary of Thesis

6.1.I Structure and activity of NOCT

The structural studies described here in this dissertation represent a major advance in understanding the functions of NOCT. The structures are the first that have been reported for NOCT and contribute to a growing body of structures of human EEP deadenylases (Abshire et al., 2018; Wang et al., 2010; Wood et al., 2015; Zhang et al., 2016). As the functions of these enzymes are uncovered it has become increasingly apparent that while the arrangement of core conserved active site residues are nearly identical, there must be fundamental differences between NOCT, CNOT6L, and PDE12 that distinguish between their different substrates and ultimately impact function (Abshire et al., 2018; Wang et al., 2010; Wood et al., 2015; Zhang et al., 2016). For example, CNOT6L has much less activity against 2´,5´ -A, which is readily hydrolyzed by PDE12, which is indicative of difference in the EEP deadenylase active sites that confer different specificities, despite the conservation of the core catalytic residues (Figures 2.1 and 2.14) (Wood et al., 2015). Structures of NOCT suggest that its mechanism of substrate recognition may be distinct from CNOT6L and PDE12, as it lacks the conserved Pro and Phe sidechains that appear to contribute to ligand binding and has a unique basic patch adjacent to the active site (Chapter 2)

(Abshire et al., 2018; Wang et al., 2010; Wood et al., 2015). The mechanisms of specificity conferred at the level of individual residues and local structural elements have yet to be determined, but the growing body of literature on EEP catalytic domain structures will assist in designing studies that aim to understand how EEP deadenylases identify their substrates.

Our ability to determine the structure of the NOCT catalytic domain and to accurately characterize the *in vitro* activity of NOCT relied heavily on obtaining highly pure enzyme. We performed the first characterization of the 3′ exoribonuclease activity of human NOCT against poly(A) RNA, the presumed substrate for human NOCT based on previous reports of mNOCT and xNOCT activity against poly(A) RNA and the poly(A) exoribonuclease activities of human PDE12 and CNOT6L (Baggs and Green, 2003; Garbarino-Pico et al., 2007; Rorbach et al., 2011; Wang et al., 2010). Previous reports of NOCT deadenylase activity utilized enzyme obtained from low stringency purifications of NOCT (Baggs and Green, 2003; Garbarino-Pico et al., 2007). Thus, it is possible that the observed activity in these assays may be exoribonuclease activity from co-purifying bacterial nucleases. Using an optimized purification protocol, we demonstrated that the human and mouse NOCT catalytic domains do not deadenylate poly(A) RNA *in vitro*, which was later corroborated in a subsequent publication from another group (Abshire et al., 2018; Estrella et al., 2018). This observation provides another example where EEP deadenylases with highly conserved active sites differ in activity against the same substrate, as PDE12 and CNOT6L both deadenylate short poly(A) RNAs (Rorbach et al., 2011; Wang et al., 2010). The reason behind the lack of *in vitro* ribonuclease activity observed for NOCT has yet to be determined.

We then considered potential substrates from other groups of EEP enzymes, which hydrolyze substrates ranging from nucleic acids to sphingomyelins and inositol polyphosphates (Freudenthal et al., 2015; Gorelik et al., 2016; Tresaugues et al., 2014; Wang et al., 2010). While NOCT shares the greatest degree of homology with the EEP deadenylases, this homology did not rule out that NOCT possesses endonuclease, DNase, or phosphatase activity against phosphosubstrates that are more characteristic substrates of other EEP subfamilies. A candidate-based screen of other compounds including nucleotides, sphingomyelins, phosphorylated glycolytic intermediates, and inositol phosphates did not identify a substrate for NOCT (Figures 3.7 and 3.8) (Abshire et al., 2018). While this was not an exhaustive screen, we determined that NOCT was not likely able to hydrolyze these alternative substrates.

At this point, our ability to perform *in vitro* assays characterizing the roles of individual NOCT active site residues was limited. Ideally, NOCT active site residues would have been mutated individually to observe the effect on enzymatic activity; however, the lack of *in vitro* activity means that we can only predict the functions of NOCT active site residues based on comparisons to other EEP enzymes. This type of comparative analysis has been performed for CNOT6L and APE1, an enzyme that removes abasic sites from DNA during repair pathways (Freudenthal et al., 2015). The endonuclease mechanism of APE1 has been studied in greater detail than the 3′ exoribonuclease mechanism of CNOT6L and the presumed catalytic mechanism of NOCT (Freudenthal et al., 2015; Wang et al., 2010). As the active sites of CNOT6L and NOCT are highly conserved, we performed this same comparison between APE1 and NOCT (Figure 6.1). EEP enzymes share a conserved divalent metal-dependent hydrolase mechanism, which suggests that

Figure 6.1 Aligning the NOCT and APE1 active sites suggests potential functions for the conserved residues in NOCT. The active site of APE1 (shown in pink, with the Mg²⁺ ion shown in gold), a well-described EEP superfamily endonuclease that is partially conserved with NOCT and other EEP family members. This conservation predicts that specific NOCT active site residues (shown in blue, with the Mg²⁺ ion shown in grey) are involved in nucleophilic activation, hydrolysis, and product stabilization.

there may be functional overlap between active site residues that are conserved between NOCT and APE1 (Figures 2.14 and 6.1). Though these types of comparisons can be highly informative in predicting the function of active site residues, there are several caveats to this approach. APE1 has endonuclease activity against the backbone of DNA abasic sites, while NOCT is a NADP(H) phosphatase and putative exoribonuclease, meaning that there are important differences in the active sites of these enzymes that confer specificity to different substrates, even if the mechanism of hydrolysis is conserved (Abshire et al., 2018; Estrella et al., 2019; Freudenthal et al., 2015). Furthermore, descriptions of the APE1 catalytic mechanism have not reached consensus. Similar to NOCT, APE1 structures have been reported with either one or two metal ions bound in the active site, and it is unclear whether one or both metals are used in catalysis (Borjigin et al., 2012; Freudenthal et al., 2015; Oezguen et al., 2007). As such, the hypothetical functions of NOCT active site residues can only be considered in more general terms with the understanding that the suggested roles for NOCT active site residues will require further biochemical characterization.

In many of the EEP superfamily enzymes, a conserved glutamate (Glu195 in NOCT and Glu96 in APE1) residue coordinates a Mg^{2+} ion in the active site, which functions in stabilization of the pentavalent transition state and product stabilization in APE1 (Freudenthal et al., 2015; Mol et al., 1995). A second active site glutamate (Glu324 and Glu210 in NOCT and APE1, respectively) activates an active site water nucleophile for hydrolysis of the phosphate backbone in the APE1 endonuclease mechanism. Other active site residues are generally described in substrate orientation and product stabilization as well as forming a hydrogen bond network assisting in shifting the pK_a of the nucleophile-activating sidechain (Freudenthal et al., 2015). These predictions can guide both interpretations of structural studies of NOCT in complex with its substrates and efforts to characterize the NOCT active site in *in vitro* experiments.

The recent discovery of NOCT *in vitro* activity against NADP(H) has allowed for the contribution of NOCT active site residues to its catalytic activity to be revisited. A recent preprint article examined the effect that mutagenizing residues in and around the active site of NOCT would have on NADP(H) phosphatase activity (Estrella et al., 2019). We replicated this activity in our own assays, also observing that the $NOCT_{120-431}$ E195A point mutation abolished activity (Figure 5.10) (Estrella et al., 2019). Mutation of this glutamate residue similarly abolishes the activities of other EEP enzymes (Freudenthal et al., 2015; Schellenberg et al., 2012; Wang et al., 2010; Wood et al., 2015). The only other active site mutant that was tested by Estrella and colleagues (H286N) also appears to render NOCT inactive (Estrella et al., 2019). Other mutations around the active site had varying effects on NADP(H) phosphatase activity, however, it is still difficult to assess the exact function of these mutants as their proposed role in substrate binding is unclear.

The preprint article by Estrella and colleagues reports a structure of NOCT in complex with NADPH, however, the poor resolution of NADPH in the active site (NADPH average B-factors = 73.6 $Å^2$ versus 48.7 $Å^2$ for the protein) makes it difficult to interpret the proposed contacts between NOCT and NADP(H) (Estrella et al., 2019). Now that NOCT NADP(H) phosphatase activity has been characterized in preliminary studies, additional mutagenesis studies should also be performed in order to assess the necessity of the other active site residues in NOCT NADP(H) substrate binding and phosphatase activity (Estrella et al., 2019). By using amino acid substitutions that are more or less conservative, the function of individual active site residues may become clearer. Future work on identifying direct mRNA targets and substrates of NOCT will aid structural and biochemical studies that characterize how NOCT binds its various substrates and how this compares to the binding mechanisms of other EEP deadenylases.

Though the biochemical results presented in Chapter 3 of this thesis did not find *in vitro* conditions in which NOCT was active, these results present an important finding in the field at the time of publication (Abshire et al., 2018). Previous work had assumed that NOCT is an active deadenylase in designing and interpreting the results of experiments utilizing NOCT active site mutants (Baggs and Green, 2003; Kawai et al., 2010a; Kawai et al., 2010b; Kojima et al., 2015; Stubblefield et al., 2018). While NOCT deadenylase activity is not ruled out by a lack of *in vitro* ribonuclease activity, these findings indicate that NOCT function is more complex than the enzyme targeting mRNAs on its own *in vivo*. NOCT may require a protein partner to be active, which has been observed for other deadenylases (Goldstrohm and Wickens, 2008). Such a protein partner may facilitate NOCT binding and recognition of target mRNAs or may allosterically activate NOCT.

Recombinant NOCT purified from bacteria may be missing a post-translational modification that is added to NOCT when it is expressed in mammalian cells. Additionally, the RNA substrates tested in our biochemical assays may not contain the sequence and/or structure context necessary for NOCT binding and activity.

6.1.II Activity of NOCT in cell-based assays

Though we did not identify conditions under which NOCT has *in vitro* ribonuclease activity, the results of cell-based tethered function assays suggest that NOCT has repressive activity against mRNAs in cells (Figure 4.2). Furthermore, a subset of the NOCT active site residues reduced repression by NOCT, suggesting that the catalytic site was at least in part responsible for repression of the reporter mRNA (Figure 4.3). None of the active site mutations completely ablated repression, indicating that NOCT repression may also include mechanisms of translational repression.

We determined that the identity of the 3´ end of the RNA is crucial in determining whether NOCT can repress reporter mRNAs. While NOCT can readily repress an mRNA reporter with a 3´ cleavage and polyadenylation signal, a reporter with a MALAT1 triple helix at the 3´ end completely relieved repression (Figure 4.6). This indicates that NOCT-mediated repression is dependent on the sequence and/or structure of the 3´ end of the mRNA, but also suggests that NOCT does not recruit factors that act on the 5´ end of the transcript. This provides another point of comparison between NOCT and CNOT6L; as a member of the CCR4-NOT deadenylase complex, CNOT6L-mediated repression involves recruiting the 5´-acting decapping factors to target mRNAs by CCR4-NOT (Alhusaini and Coller, 2016). This observation partially explains the residual repression of the MALAT1 reporter by CNOT7 in these assays, as this deadenylase is also a subunit

of CCR4-NOT (Doidge et al., 2012). Repression of reporter mRNAs by NOCT is perturbed by altering either the active site residues or the features of its mRNA target, suggesting that NOCT directly mediates repression of mRNAs with specific sequences or structures.

The mechanism of NOCT-mediated repression is a complex question, as it likely involves the identification of other protein factors that facilitate repression of NOCT target mRNAs and may include mechanisms of translational repression. This is reflected in the results of Northern blotting experiments, in which stabilization of the reporter mRNA levels were detected with the NOCT H414A mutant such that the mRNA levels were similar to those with the negative control (Figure 4.7). However, effect upon mRNA abundance was unable to be determined in the case of the other active site residues so that it is unclear whether these other active site residues play a role in NOCT-mediated mRNA decay (Figure 4.7). Overall, the effect of NOCT appears to be more significant when protein expression from the reporter is measured versus the observed effects steady state levels of reporter mRNA. This may indicate that the contribution of RNA decay in NOCT-mediated mRNA repression is modest and that the combined effect of translational repression and mRNA decay is what ultimately represses expression from the mRNA reporter. The effects of NOCT on mRNA abundance may also be transcript dependent with respect to mRNA decay and translational repression, as overexpression of NOCT significantly affected the abundance of endogenous mRNAs in HEK293 cells (Figure 5.6). This observed regulation of endogenous transcripts underscores the importance of identification of endogenous NOCT targets to assess mechanisms of NOCT-mediated repression of mRNA transcripts. Thus, our cell-based assays suggest a model by which NOCT represses mRNAs through a combinatorial mechanism of RNA decay and

translational repression in a manner dependent on the sequence and/or structure of the

3′ end of the mRNA (Figure 6.2).

Figure 6.2 A putative model of NOCT repression of mRNAs and regulation of NOCT function.
NOCT may directly or indirectly associate with mRNA transcripts through recognition of a sequence
and/or structural element, here referred to as a recognition element (RE). In the case of indirect
association, this element may be a binding site for an RBP that recruits NOCT, which then mediates
association between NOCT and its targets. Non-RBP binding partners may control NOCT activity
through allosteric mechanisms. As NOCT is not active against mRNAs *in vitro,* such a partner (either
RBP or non-RBP) may be necessary for NOCT activity against mRNAs. Additionally, NOCT activity
could be dependent on post-translational modifications. These putative regulatory mechanisms have
yet to be identified or ruled out and a combination of these mechanisms may modulate NOCT-
mediated mRNA repression. This figure has been adapted from a review by K. Hughes et al (Hughes

6.1.III Endogenous targets of NOCT

To accurately determine the effect of NOCT on endogenous mRNAs, we first

considered the localization of NOCT, as the protein sequence has elements that suggest

localization of NOCT to the mitochondria. The observation of multiple molecular weight

species in Western blots of cell lysates indicates that NOCT is processed (Figure 5.2).

Upon examination of the N-terminus using the MitoFates prediction software, we

identified an arginine-enriched sequence predicted to form an amphipathic helix

consistent with a mitochondrial targeting sequence (residues 2-15), as well as

downstream consensus sequences for mitochondrial peptidases cleaving at L74/Y75.

These sequence features contribute to a significant probability score for putative

localization of NOCT to the mitochondria (P= 0.94) (Fukasawa et al., 2015). As the pools

of available substrate mRNAs are highly different in the cytoplasm versus the

mitochondria, localization would greatly impact the target identification (D'Souza and

Minczuk, 2018; Solnestam et al., 2012). Furthermore, the NADP(H) pools in these cellular

compartments are largely distinct, suggesting that localization would also greatly impact

the downstream effects of NOCT phosphatase activity (Nikiforov et al., 2011).

Consistent with this predicted localization, transfected full-length NOCT has an

apparent molecular weight of ~41 kDa, which is the size expected if NOCT is processed

(Figure 5.2). However, this observation is not necessarily an indication of processing, as

there is a downstream Met67 residue and Kozak sequence that would allow for translation

of a short isoform of NOCT that also would produce a ~41 kDa protein. We then examined

the effect of mutagenizing the sequences predicted to be associated with either NOCT

localization and processing or alternative translation initiation. Truncation of the

amphipathic helix (residues 2-15) resulted in detection of a longer, ~55 kDa form of

NOCT, suggesting that mitochondrial localization leads to processing. Furthermore, the

presence of a ~41 kDa band attributable to Met67 initiation was only observed in when

Met1 was mutated, suggesting that Met1 is the preferred translation initiation site in

HEK293 and HCT116 cell lines (Figure 5.2). Fusion of the NOCT N-terminus (residues 1-

86) to EGFP was also sufficient to cause processing (Figure 5.3). Together, these data

suggest that NOCT can processed in a manner consistent with mitochondrial localization.

We then examined the molecular weights of endogenous NOCT in different mouse

and human tissue extracts and observed multiple bands consistent with the molecular

weights expected for the different predicted isoforms that would result from full-length

NOCT or the form produced by mitochondrial localization and processing and/or Met67 translation initiation (Figure 5.4). These expression patterns varied by tissue, suggesting that processing is dependent on tissue type. The ~55 kDa form of endogenous NOCT corresponds to Met1 initiated NOCT that is not localized or processed, which is presumed to be cytoplasmic. It is unknown why a full-length form of NOCT that contains the sequence elements necessary for localization and processing would not be imported into the mitochondria and cleaved into its short form. This result suggests that there may be additional mechanisms of regulation that determine localization patterns in various cell types, which will be important in understanding how NOCT activity is regulated.

We next addressed the effect that overexpression of cytoplasmic NOCT has on differential expression of endogenous mRNA transcripts. RNA-Seq was performed in HEK293 cells overexpressing a NOCT Δ2-15 3x FLAG. We analyzed this data without constraining our dataset to circadian transcripts with differences in poly(A) tail lengths as had been done in previous work (Kojima et al., 2015; Stubblefield et al., 2018). This identifying a total of 490 regulated genes, including 235 downregulated transcripts (Figures 5.6 and 5.7). To validate our results, we confirmed that the mRNA abundance was reduced in a subset of the top downregulated genes identified by RNA-Seq (*DLAGP1, TMEFF2, PTPRZ1, RORβ,* and *FRRS1L*) using qPCR (Figure 5.6). When we attempted to demonstrate that overexpression of NOCT also resulted in a decrease in steady-state protein levels for a subset of the top downregulated targets, we did not observe a decrease in protein levels for PTPRZ1, RORβ, or FRRS1L (Figure 5.8). This may be due to effects of protein stability such that steady state measurements are unable to detect differential regulation, or efficient translation of the remaining transcripts may be

175

enough to mask the effect of NOCT. Additionally, further validation of the antibodies used in these experiments is needed. As these targets represent only a small number of regulated mRNAs, we will need to experiment further to determine the functional outcome of NOCT-mediated changes in mRNA abundance. We examined the RNA-Seq data at the transcript level for evidence that NOCT regulates mRNAs in an isoform-specific manner. Isoform-level differences in fold change expression were observed for *DLAGP1, TMEFF2,* and *PTPRZ1* mRNAs, suggesting that NOCT regulates some transcript isoform abundances more than others. Though we observed differential changes in overall abundance for these mRNAs, further validation is needed to determine biological significance.

In addition to differences in data analysis, we turned to the neuro-adrenal HEK293 cell line, which has low levels of endogenous NOCT protein expression and is easily genetically manipulated (Figure 5.2). We also anticipate that these cells have the necessary cofactors needed for NOCT function, based on the results of the tethered function assays (Figure 4.2). While this piece of evidence does not mean that the endogenous NOCT targets are also expressed in this cell type, *NOCT* mRNA is expressed to some extent in brain, kidney, and adrenal gland and could feasibly regulate mRNAs in these tissues (Figure 4.4) (Wang et al., 2001).

In this thesis, we confirmed reports of NOCT NADP(H) phosphatase activity in *in vitro* assays, demonstrating that NOCT could potentially regulate metabolite levels *in vivo*. While it first needs to be validated that NOCT phosphatase activity has a significant effect on NADP(H) in cells, this observation must be considered in the interpretation of our RNA-Seq results, which could also be measuring changes in the transcriptome that occur in

response to modulation of NADP(H) levels. Parsing the different effects of NOCT repression of mRNA targets and phosphatase activity against nicotinamide adenine dinucleotides may prove challenging in light of the many effects that may occur downstream of post-transcriptional regulation of a large number of mRNAs (Figure 5.6) and fluctuations in NADP(H) levels (Figure 5.11). Both mechanisms may contribute to NOCT-mediated regulation. Some phenotypes of the NOCT KO mice can be explained by increased NADP(H) levels, such as the inhibition of ROS-signaling mediated adipogenesis that would be predicted by increased NADP(H) levels (discussed in Chapter 5) (Li et al., 2012). Conversely, other phenotypes do not seem to correlate with the putative effects of NOCT KO. These observations include the increased NADPH levels in KO mice that would be predicted to increase fatty acid synthesis and decrease bone mass, contrary to what is observed, and that loss of NOCT may increase NAD^+ levels, which may sensitize mice to HFD-induced obesity (Canto et al., 2012; Currie et al., 2013; Rahman et al., 2018) (discussed in Chapter 5). These phenotypes may therefore be more indicative of NOCT regulation of transcripts rather than regulation through modulation of NADP(H) levels. Ultimately, these are predicted effects of NOCT-mediated regulation and the effect of NOCT on cellular NADP(H) levels in different cellular compartments must first be determined.

6.2 Future Directions: Identifying protein partners of NOCT

NOCT represses mRNA expression *in vivo* but has no measurable ribonuclease activity *in vitro* (discussed in Chapter 3). This observation predicts that protein partners may regulate NOCT activity *in* vivo, though none have been identified to date. Therefore, it is still possible that the cellular functions of NOCT are not regulated by binding partners.

This may be the case for NOCT phosphatase activity, as the catalytic domain displays this activity *in vitro*; however, for the ribonuclease activity of NOCT this would indicate that while the active site of NOCT is nearly identical to those of PDE12 and CNOT6L (discussed in Chapter 2), it does not have the enzymatic activity against poly(A) RNA substrates observed for PDE12 and CNOT6L (Rorbach et al., 2011; Wang et al., 2010).
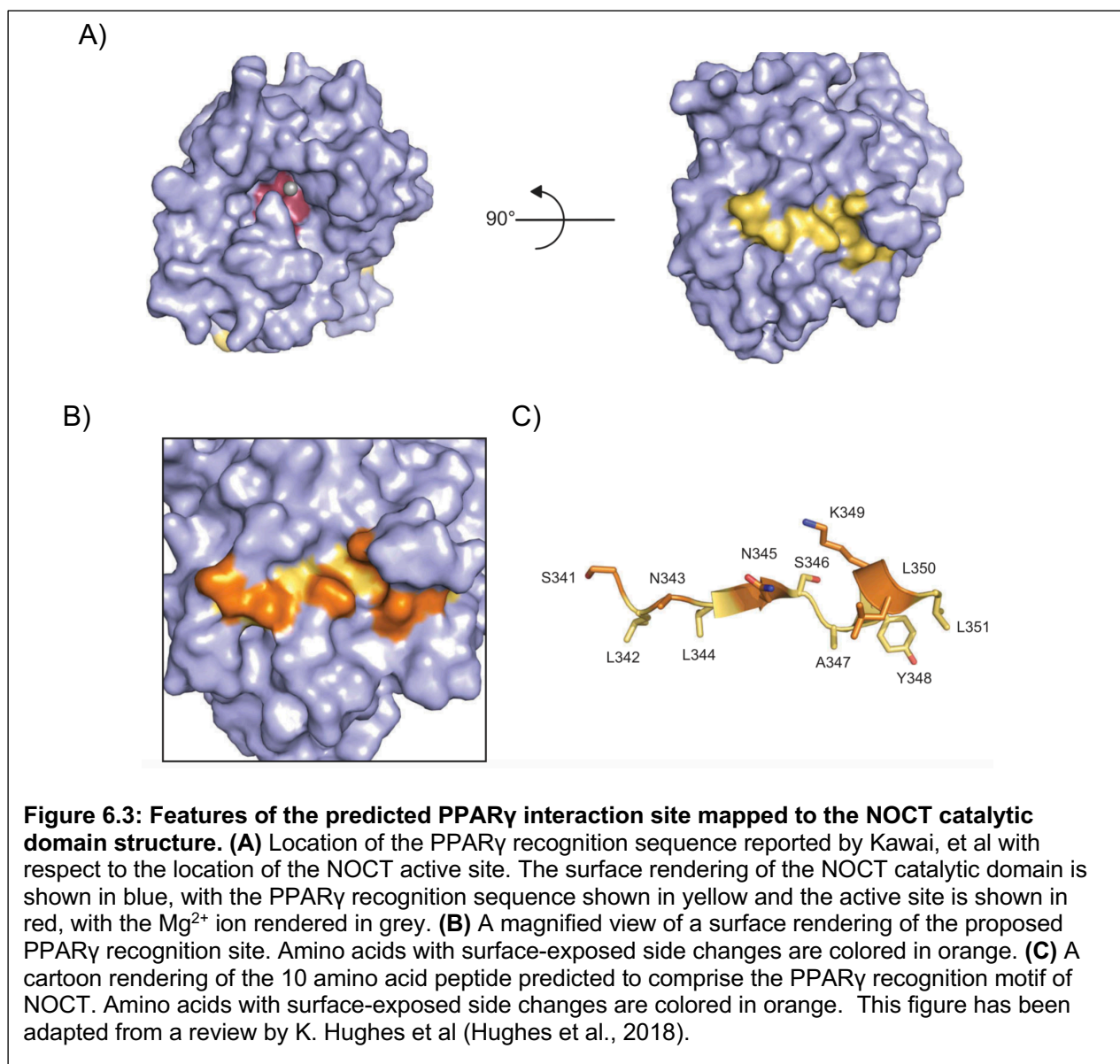
Our tethered function assays provide the primary evidence for NOCT ribonuclease activity in cells (Figure 4.3). While repression of reporter mRNAs may occur through mechanisms other than NOCT activity, differences were observed in NOCT-mediated repression that were dependent on the active site of NOCT. Mutagenesis of the active site residues His414, Asp377, and Asn149 to alanine moderately relieved repression of NLuc mRNA reporters, indicating that perturbation of conserved active site residues reduces repression of the reporter mRNA. Northern Blotting for the NLuc reporter co-transfected with the NOCT H414A mutant suggests that active site mutagenesis stabilizes the levels of reporter mRNA consistent with NOCT promoting the decay of the reporter mRNA in a manner dependent on its active site. These results suggest that NOCT has ribonuclease activity in cells and that a protein partner or post-translational modification not found in our *in vitro* conditions may be necessary for NOCT activity.

Conservation between NOCT, PDE12, and CNOT6L indicate that NOCT has an intact active site that is competent for deadenylase activity. Many cases can be found where enzymes are evolutionarily repurposed for another function and lack ancestral activity, however, these cases usually involve changes in the active site so that key residues once involved in enzymatic activity are mutated in the active site of the pseudoenzyme (Murphy et al., 2017). This finding is not the case for NOCT, suggesting

that the enzyme has deadenylase activity, but not in the *in vitro* conditions tested here in this work. Since many conditions compatible with deadenylase activity were surveyed, (Figures 3.1-8), it seems probable that an unknown factor found *in vivo* would be necessary for NOCT activity. While NOCT diverges from CNOT6L and PDE12 with respect to the Pro and Phe residues that contribute to ligand binding, the presence of a conserved basic patch compatible with RNA binding suggests that NOCT may have a different mechanism of putative mRNA substrate binding, not that it is incapable (discussed in Chapter 2) (Wang et al., 2010; Wood et al., 2015; Zhang et al., 2016).

Regulation of ribonuclease activity by protein partners has been well-described for other enzymes. For example, the Ccr4-NOT complex has two nuclease modules, Ccr4 (CNOT6 or CNOT6L in humans) Caf1 (CNOT7 or CNOT8 in humans), which are Ccr4 regulated by a variety of protein partners such as PABPC1, which stimulates the activity of Ccr4 against mRNA poly(A) tails while inhibiting Caf1-mediated decay (Webster et al., 2018; Yi et al., 2018). Caf1 interacts directly with the anti-proliferative protein BTG2, which stimulates mRNA decay activity of Caf1, and to a lesser extent, Ccr4 (Mauxion et al., 2008; Prevot et al., 2001; Rouault et al., 1998).

To date, only one binding partner of NOCT has been described. PPARγ was shown to physically interact with NOCT in immunoprecipitation (IP) experiments. This observation was further tested by introducing a NOCT peptide with predicted sequence homology to a Src homology 2 (SH2) domain binding motif. This peptide appears to compete with the NOCT-PPARγ interaction observed in IP experiments (Kawai et al., 2010b). The authors of this study defined the SH2 sequence motif as the interaction surface of NOCT with PPARγ; however, this motif was not tested for interaction with

Figure 6.3: Features of the predicted PPARγ interaction site mapped to the NOCT catalytic domain structure. (A) Location of the PPARγ recognition sequence reported by Kawai, et al with respect to the location of the NOCT active site. The surface rendering of the NOCT catalytic domain is shown in blue, with the PPARγ recognition sequence shown in yellow and the active site is shown in red, with the Mg^{2+} ion rendered in grey. **(B)** A magnified view of a surface rendering of the proposed PPARγ recognition site. Amino acids with surface-exposed side changes are colored in orange. **(C)** A cartoon rendering of the 10 amino acid peptide predicted to comprise the PPARγ recognition motif of NOCT. Amino acids with surface-exposed side changes are colored in orange. This figure has been adapted from a review by K. Hughes et al (Hughes et al., 2018).

PPARγ in the context of full-length NOCT via mutational analysis. When mapped to the surface of the NOCT crystal structure, the side chains of the proposed interaction motif (341- 351) are not largely surface exposed and accessible for interaction (Figure 6.3) . Furthermore, SH2 domains are known to be downstream of a phosphotyrosine essential for the binding interaction. While there is a tyrosine residue at position 334 in the NOCT protein sequence, this residue has not been observed to be phosphorylated and was not included in the peptide used in competition studies (Kawai et al., 2010b; Liu et al., 2010).

The interaction between NOCT and PPARγ has been implicated in the pro-adipogenic functions predicted for NOCT (Green et al., 2007; Hee et al., 2012; Kawai et al., 2010b). Interaction with NOCT was reported to increase nuclear translocation of PPARγ to the nucleus, which is presumed to increase the transcriptional of pro-adipogenic PPARγ targets (Lefterova et al., 2014). NOCT was observed to be perinuclear in this context, localizing to the nuclear membrane but not crossing into the nucleus itself (Kawai et al., 2010b). These conclusions must be cautiously interpreted, as the resolution of these experiments was not high enough to be definitive with respect to nuclear localization. While fractionation experiments have detected NOCT in nuclear fractions, this has not been observed in microscopy studies (Estrella et al., 2019). The functional effect of the interaction between NOCT and PPARγ on the transcriptional targets of PPARγ or the endogenous mRNA targets of NOCT is therefore still unknown.

Other candidate-based approaches based on the known function of NOCT may yield identification of NOCT protein partners. Currently, the best candidates for these approaches are proteins involved in RNA decay and translational control pathways. Protein partners that may regulate the phosphatase activity of NOCT are more difficult to assess, as this is not a well-described activity in mammals. As NOCT is most closely related to CCR4-type deadenylases, it is feasible that NOCT could interact with the human CCR4-NOT complex as an alternative deadenylase subunit analogous to CNOT6 and CNOT6L (Doidge et al., 2012). NOCT lacks the leucine-rich motif that facilitates interaction between CNOT6 or CNOT6L and CCR4-NOT (Ng et al., 2011). The lack of this motif does not necessarily rule out alternative mechanisms of interaction between NOCT and CCR4-NOT or that CNOT6L, CNOT6, and NOCT are mutually exclusive

deadenylase subunits of CCR4- NOT. Interaction between NOCT and the CCR4-NOT complex has been shown in *Drosophila* but this may not hold true for human NOCT, as there is very little conservation between the N-termini of *Drosophila* and human NOCT (the 120 N-terminal amino acids in NOCT is 8% identical to the aligned sequence in the *Drosophila* homolog, Curled) (Temme et al., 2010).

6.3 Future Directions: The function of the NOCT N-terminus

NOC has a unique N-terminal domain that is predicted to be largely disordered. Analysis of the NOCT protein sequence using GlobPlot illustrates the high degree of disorder in the N-terminus (Linding et al., 2003; Marchler-Bauer et al., 2017; Marchler-Bauer and Bryant, 2004; Marchler-Bauer et al., 2015; Marchler-Bauer et al., 2011). The data presented in Chapter 5 indicate that the NOCT N-terminus may mediate localization of NOCT to the mitochondria; however, we observed that endogenous NOCT cann be expressed as a ~55 kDa form consistent with full-length unprocessed protein (Figures 5.2 and 5.4). Therefore, additional functions of the NOCT N-terminus may be possible.

The N-terminus of CNOT6L has an identifiable leucine-rich repeat domain (LRR) that is reported to be important for the interaction between CNOT6L (Ccr4) and CCR4-NOT via Caf1 (CNOT7 and CNOT8) (Clark et al., 2004; Ng et al., 2011). One study suggests that the LRR might also be important for subcellular localization of CNOT6L to the nucleus (Mittal et al., 2011). *Xenopus* NOCT is reported to contain a similar LRR in the N-terminus, but this repeat motif appears to be poorly conserved relative to other LRR motifs (Baggs and Green, 2003; Ng et al., 2011). Furthermore, interactions between *Xenopus* Noct and Caf1 have not been demonstrated. NOCT may therefore be

differentiated from canonical CCR4 deadenylases by its lack of an LRR, which makes it unlikely that it interacts with CCR4-NOT.

Our cell-based assays demonstrated that full-length MS2-NOCT$_{1-431}$ and a truncated MS2-NOCT$_{64-431}$ construct equivalently repressed a NLuc reporter mRNA. In these assays, the MS2 domain of the fusion proteins facilitated the interaction between NOCT and the reporter mRNA MS2 stem loops (Figures 4.1 and 4.2). It is possible that the NOCT N-terminus plays a yet undiscovered role in specific recognition of endogenous mRNA targets that is bypassed by MS2-facilitated interaction with the mRNA reporters, or that the NOCT N-terminus could mediates protein-protein interactions in certain contexts. Identification of direct targets will help in determining what elements of NOCT and its targets are required for target recognition.

6.4 Future Directions: Functional significance of NOCT-mediated changes in mRNA abundance

RNA-Seq analysis of the NOCT Δ2-15 overexpression cell lines identified 235 downregulated transcripts. We expect that a subset of these transcripts are repressed as a result of direct association with NOCT. To determine the requirements for NOCT target recognition, we attempted to identify motifs present in the downregulated transcripts at the gene level but were unable to determine a high-confidence motif located in the 3′ UTR. Based on our evidence that regulation is isoform-specific (Table 5.3), we plan to further analyze NOCT targets using RNA-Seq data at the transcript level. Once the most highly regulated isoforms are identified, a high-confidence motif is more likely to be successfully identified, which would either represent a novel binding site directly recognized by NOCT or would represent a known motif of another RBP that could mediate

NOCT association with its targets. We also cannot rule out that the motif is outside the 3´UTR and will therefore include the 5´UTR and ORF in motif analysis.

For further validation of NOCT targets, we plan to clone the putative motifs associated with NOCT regulation into reporter mRNAs and measure repression their repression when NOCT is overexpressed. By showing that these motifs can confer NOCT-mediated regulation on a reporter, we would provide evidence that mRNAs containing these motifs are regulated by NOCT. Additionally, we plan to conduct RNA immunoprecipitation experiments using the targets validated in qPCR experiments (Figure 5.7). By showing association between the overexpressed NOCT Δ2-15 and the differentially regulated transcript, we can identify the detected transcripts as direct targets of NOCT regulation and not changed as a result of indirect effects. These data would provide the first concrete evidence of a direct endogenous mRNA target of human NOCT.

It is also important to illustrate the biological significance of NOCT regulation in brain, based on the enrichment of targets associated with neuronal function. NOCT function has been characterized in liver, adipose, bone, and small intestine, but NOCT function in other tissues cannot be ruled out as NOCT phenotypes have not yet been examined in the brain (Douris et al., 2011; Green et al., 2007; Kawai et al., 2010b; Stubblefield et al., 2018). Enrichment of GO terms associated with neurological function may be related to the neuro-adrenal phenotype of HEK293 cells, but it is certainly feasible that NOCT regulates targets related to neurological function in the brain, as endogenous NOCT is expressed in both human and mouse brain tissue (Figure 5.4). Therefore, we anticipate that NOCT may have the ability to regulate neurological function.

The biological significance of NOCT regulation in brain tissue can be addressed using SH-SY5Y neuroblastoma cells, a well-described model for studying neuronal function that can be differentiated into different neuronal cell types (Kovalevich and Langford, 2013). Using lentiviral transduction, NOCT can be stably expressed or knocked down and functions associated with the most significant GO terms can be assessed. For example, voltage-gated potassium channel activity can be assayed using dyes sensitive to stimulation of potassium channels, which would provide a quantitative measure of the effects of NOCT on expression of potassium channels (Beacham et al., 2010). Measurement of transmembrane tyrosine kinases may be possible through assays that rely on measurement of the amounts of phosphorylated tyrosine, if the appropriate phosphotyrosine antibodies and/or agonists are available (Minor, 2005). These assays would report not just on receptor expression but also activation, which could be affected by NOCT regulation of NADP(H) levels. Mitochondrial function is known to be impacted by Ca^{2+} signaling in SH-SY5Y cells in response to carbachol, a compound that stimulates nicotinic acid adenine dinucleotide phosphate (NAADPH) signaling (Aley et al., 2013; Sherer et al., 2001). This indicates that NOCT overexpression could alter Ca^{2+} signaling in SH-SY5Y cells through depletion of NADP, the molecular precursor of NAADP, in addition to regulation of mRNAs associated with this function (Figure 5.11) (Galione, 2015). Measuring the affect NOCT-mediated regulation in neuronal models would provide an indication of the combined effects of NOCT mRNA repression and phosphatase activities in brain tissue.

6.5 Future Directions: Cellular effects of NOCT activity against non-mRNA substrates

NOCT phosphatase activity against $NADP^+$ and NADPH was reported during the writing of this thesis and was subsequently confirmed in our own studies (Figure 5.10) (Estrella et al., 2019). We had not predicted this activity and included it in our phosphatase screen due to few examples of mammalian $NADP^+$ and NADPH phosphatases (Figure 3.8). While this activity was described in rat liver lysates in the 1980s, very little follow up was performed to identify the responsible enzymes (Navas et al., 1986; Richter, 1987). A second mammalian enzyme, MESH1, was recently reported to have NADP(H) phosphatase activity, which indicates that NADP(H) phosphatase activity in the cytoplasm is not unique to NOCT (Ding et al., 2018). Numerous reports on $NADP^+$ and NADPH phosphatase activity have been published for bacteria, archaea, and plants (Gallais et al., 2000; Kawai et al., 2005; Kawai et al., 2004). The biological significance of NOCT NADP(H) phosphatase activity has not yet been determined in tissues or cells, which is crucial to fully understanding the mechanisms of NOCT-mediated regulation.

The observation that NOCT lacks *in vitro* ribonuclease activity but has NADP(H) phosphatase activity was suggested to exclude the possibility of NOCT activity against mRNAs (Estrella et al., 2019). However, several key pieces of evidence presented in this dissertation suggest otherwise. NOCT represses reporter mRNAs in cell-based tethered function assays, and a subset of mutations in the NOCT active partially alleviate repression, indicating that ribonuclease activity contributes to NOCT repression of mRNAs (Figure 4.3). Additionally, NOCT repression is highly sensitive to the sequence and/or structure of the 3´ end of the mRNA, indicating that NOCT repression is specific

for certain mRNA features (Figure 4.6). For NOCT have both NADP(H) phosphatase and ribonuclease activity, its active site would have to be able to accommodate and hydrolyze both 3´ and 2´ position phosphate bonds in nucleotide substrates. Flexibility within the active sites of other EEP deadenylases have been reported, such as for PDE12, which degrades 2´,5´-A oligoadenylate in addition to poly(A) RNA. CNOT6L has also been shown to have weak activity against 2´,5´-A (Wood et al., 2015). Therefore, we cannot exclude a model wherein NOCT requires a protein partner for activity against mRNA in cells, which may be an important mechanism of regulating the dual activities of NOCT.

We propose here that NOCT acts in two capacities, regulating mRNA transcripts that express gene products related to metabolism as well as modulating the levels of NADP(H). Several examples of RNA-binding metabolic enzymes, or "moonlighting RBPs" have been described in the literature (Castello et al., 2015). In one example, inosine 5´ monophosphate dehydrogenase 1 (IMPDH1), a purine biosynthesis enzyme, has reported disease-associated RNA binding mutants that cause retinal degeneration (Hedstrom, 2008; McLean et al., 2004; Mortimer and Hedstrom, 2005; Mortimer et al., 2008). Other classic examples include the citric acid cycle enzyme aconitase, which binds to RNA stem-loops of transcripts involved in iron homeostasis, and glyceraldehyde-3-phosphate dehydrogenase (GAPDH), which binds to AU-rich elements of many different mRNA transcripts (Constable et al., 1992; Dollenmaier and Weitz, 2003; Hentze et al., 1987; Mullner and Kuhn, 1988). Strikingly, GAPDH binds NAD^+ and RNA in the same binding site (Nagy et al., 2000; Nagy and Rigby, 1995). These examples illustrate how the RNA binding activity of metabolic enzymes can serve highly important functions. Surveys of RNA binding proteins have identified other examples of moonlighting RBPs,

suggesting that many of these dual activity enzymes provide a level of translational control in response to metabolic states (Castello et al., 2015).

NOCT phosphatase activity against NADP(H) would have different effects based on localization to the mitochondria or the cytoplasm. NADP(H) is not in flux between these subcellular compartments, which suggests that differential localization of NOCT may affect distinct pools of NADPH (Nikiforov et al., 2011). Cellular mechanisms that exchange precursors of NADPH between these compartments likely enables a degree of compensation for disruption of the different pools of nicotinamide dinucleotides (Adler et al., 2014). The exchange of these precursors is highly regulated and occurs through specialized shuttle systems (Xiao et al., 2018). These processes may enable adaptation to NOCT-mediated perturbations in NADP(H) levels, especially in the case of using models that constitutively alter NOCT expression. Therefore, the use of inducible knockdown or overexpression models may be especially valuable in these studies. In the cytoplasm, MESH1 phosphatase also regulates NADP(H) levels in the cytoplasm, which may complicate studying the effects of NOCT in the cytoplasm through its redundant activity (Ding et al., 2018). This redundant activity may necessitate knockdown of MESH1 to more easily observe the role of NOCT phosphatase activity in cellular function.
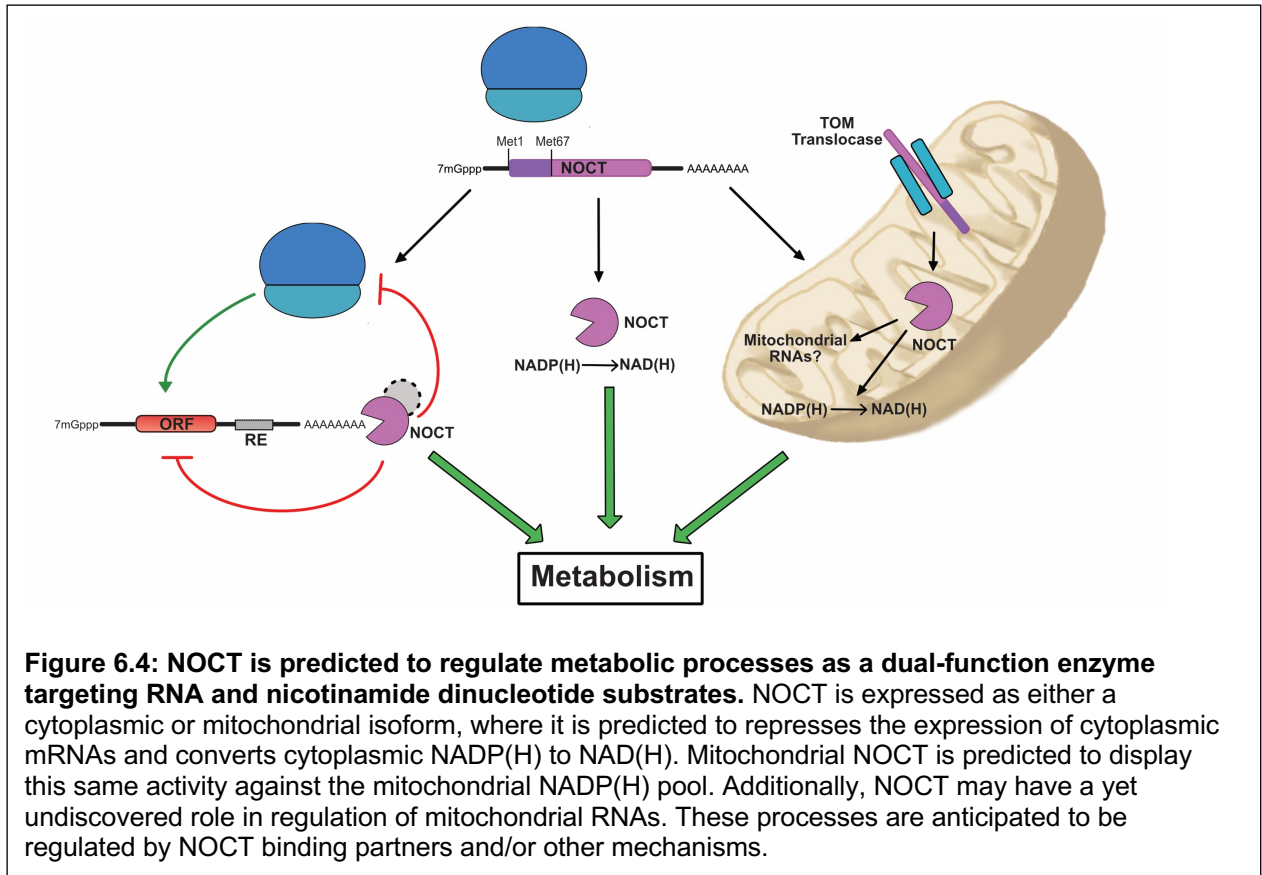
As NADP(H) is involved in a large range of biological processes, there are many directions that could be taken to determine the cellular effects of NOCT phosphatase activity (Figure 5.11). First, the changes in cellular NAD(H) and NADP(H) levels in response to perturbation of NOCT expression will be important to determine the functional significance of NOCT NADP(H) phosphatase activity. Our RNA-Seq dataset may also indicate which pathways are more significantly affected by NOCT. As the changes in gene

expression may indicate cellular effects that are both direct and indirect, some of the observed changes in gene expression may be downstream of NOCT activity against nicotinamide dinucleotides. Comparing the iPAGE analysis (Figure 5.9) to known functions of cellular NADP(H) may help target the pathways most sensitive to NOCT overexpression, such as Ca^{2+} signaling, antioxidant activity, and mitochondrial function (Xiao et al., 2018; Ying, 2008). The role of NOCT phosphatase activity must therefore be determined in order to understand the full effects of NOCT-mediated regulation.

6.6 Concluding Remarks

This work represents several advancements in understanding NOCT biology. We solved the first structure of the NOCT catalytic domain, illustrating the structural homology of NOCT with other EEP deadenylases. We determined that NOCT lacks *in vitro* ribonuclease activity but corroborate that has a unique *in vitro* phosphatase activity against nicotinamide dinucleotides. We also provide the first evidence that human NOCT can repress expression of a reporter mRNA *in vivo* and changes the abundance of hundreds of endogenous transcripts when overexpressed. Furthermore, we identify that the NOCT N-terminus is processed in a tissue-specific manner consistent with mitochondrial and cytoplasmic localization. Broadly, these data suggest that NOCT is dual-function enzyme that is localized to multiple compartments and has activity against mRNA and nicotinamide dinucleotides, resulting in regulation of metabolic processes (Figure 6.4).

This work also opens up many new areas of inquiry. Among the most important include understanding the protein partners, co-factors and/or post translational modifications that enable NOCT to repress mRNA expression *in vivo* that are missing *in*

Figure 6.4: NOCT is predicted to regulate metabolic processes as a dual-function enzyme targeting RNA and nicotinamide dinucleotide substrates. NOCT is expressed as either a cytoplasmic or mitochondrial isoform, where it is predicted to represses the expression of cytoplasmic mRNAs and converts cytoplasmic NADP(H) to NAD(H). Mitochondrial NOCT is predicted to display this same activity against the mitochondrial NADP(H) pool. Additionally, NOCT may have a yet undiscovered role in regulation of mitochondrial RNAs. These processes are anticipated to be regulated by NOCT binding partners and/or other mechanisms.

vitro. It is still not understood how the combination of NOCT activities leads to phenotypes of obesity resistance, impaired dietary fat transport, and increased bone density. We still do not know the mechanism by which NOCT recognizes and targets its mRNA substrates. Despite the many questions that remain, the work here lays a foundation for addressing these questions and understanding the biological functions of NOCT.

6.7 References

Abshire, E.T., Chasseur, J., Bohn, J.A., Del Rizzo, P.A., Freddolino, P.L., Goldstrohm, A.C., and Trievel, R.C. (2018). The structure of human Nocturnin reveals a conserved ribonuclease domain that represses target transcript translation and abundance in cells. Nucleic acids research *46*, 6257-6270.

Adler, L.t., Chen, C., and Koutalos, Y. (2014). Mitochondria contribute to NADPH generation in mouse rod photoreceptors. The Journal of biological chemistry *289*, 1519-1528.

Aley, P.K., Singh, N., Brailoiu, G.C., Brailoiu, E., and Churchill, G.C. (2013). Nicotinic acid adenine dinucleotide phosphate (NAADP) is a second messenger in muscarinic receptor-induced contraction of guinea pig trachea. The Journal of biological chemistry *288*, 10986-10993.

Alhusaini, N., and Coller, J. (2016). The deadenylase components Not2p, Not3p, and Not5p promote mRNA decapping. Rna *22*, 709-721.

Baggs, J.E., and Green, C.B. (2003). Nocturnin, a deadenylase in Xenopus laevis retina: a mechanism for posttranscriptional control of circadian-related mRNA. Current biology : CB *13*, 189-198.

Beacham, D., Blackmer, T., O'Grady, M., and Hanson, G. (2010). Cell-based potassium ion channel screening using the FluxOR assay. J Biomol Screen *15*, 441-446.

Borjigin, M., Arenaz, P., and Stec, B. (2012). Chinese hamster AP endonuclease operates by a two-metal ion assisted catalytic mechanism. FEBS Lett *586*, 242-247.

Canto, C., Houtkooper, R.H., Pirinen, E., Youn, D.Y., Oosterveer, M.H., Cen, Y., Fernandez-Marcos, P.J., Yamamoto, H., Andreux, P.A., Cettour-Rose, P.*, et al.* (2012). The NAD(+) precursor nicotinamide riboside enhances oxidative metabolism and protects against high-fat diet-induced obesity. Cell Metab *15*, 838-847.

Castello, A., Hentze, M.W., and Preiss, T. (2015). Metabolic Enzymes Enjoying New Partnerships as RNA-Binding Proteins. Trends in endocrinology and metabolism: TEM *26*, 746-757.

Clark, L.B., Viswanathan, P., Quigley, G., Chiang, Y.C., McMahon, J.S., Yao, G., Chen, J., Nelsbach, A., and Denis, C.L. (2004). Systematic mutagenesis of the leucine-rich repeat (LRR) domain of CCR4 reveals specific sites for binding to CAF1 and a separate critical role for the LRR in CCR4 deadenylase activity. The Journal of biological chemistry *279*, 13616-13623.

Constable, A., Quick, S., Gray, N.K., and Hentze, M.W. (1992). Modulation of the RNA-binding activity of a regulatory protein by iron in vitro: switching between enzymatic and genetic function? Proceedings of the National Academy of Sciences of the United States of America *89*, 4554-4558.

Currie, E., Schulze, A., Zechner, R., Walther, T.C., and Farese, R.V., Jr. (2013). Cellular fatty acid metabolism and cancer. Cell Metab *18*, 153-161.

D'Souza, A.R., and Minczuk, M. (2018). Mitochondrial transcription and translation: overview. Essays Biochem *62*, 309-320.

Ding, C.C., Rose, J., Wu, J., Sun, T., Chen, K., Chen, P., Xu, E., Tian, S., Akinwuntan, J., Guan, Z.*, et al.* (2018). Mammalian stringent-like response mediated by the cytosolic NADPH phosphatase MESH1. Biorxiv [Preprint].

Doidge, R., Mittal, S., Aslam, A., and Winkler, G.S. (2012). Deadenylation of cytoplasmic mRNA by the mammalian Ccr4-Not complex. Biochem Soc Trans *40*, 896-901.

Dollenmaier, G., and Weitz, M. (2003). Interaction of glyceraldehyde-3-phosphate dehydrogenase with secondary and tertiary RNA structural elements of the hepatitis A virus 3' translated and non-translated regions. J Gen Virol *84*, 403-414.

Douris, N., Kojima, S., Pan, X., Lerch-Gaggl, A.F., Duong, S.Q., Hussain, M.M., and Green, C.B. (2011). Nocturnin regulates circadian trafficking of dietary lipid in intestinal enterocytes. Current biology : CB *21*, 1347-1355.

Estrella, M.A., Du, J., Chen, L., Rath, S., Prangley, E., Chitrakar, A., Aoki, T., Schedl, P., Rabinowitz, J., and Korennykh, A. (2019). The Metabolites NADP+ and NADPH are the Targets of the Circadian Protein Nocturnin (Curled). Biorxiv [Preprint].

Estrella, M.A., Du, J., and Korennykh, A. (2018). Crystal Structure of Human Nocturnin Catalytic Domain. Scientific reports *8*, 16294.

Freudenthal, B.D., Beard, W.A., Cuneo, M.J., Dyrkheeva, N.S., and Wilson, S.H. (2015). Capturing snapshots of APE1 processing DNA damage. Nature structural & molecular biology *22*, 924-931.

Fukasawa, Y., Tsuji, J., Fu, S.C., Tomii, K., Horton, P., and Imai, K. (2015). MitoFates: improved prediction of mitochondrial targeting sequences and their cleavage sites. Mol Cell Proteomics *14*, 1113-1126.

Galione, A. (2015). A primer of NAADP-mediated Ca(2+) signalling: From sea urchin eggs to mammalian cells. Cell Calcium *58*, 27-47.

Gallais, S., de Crescenzo, M.A., and Laval-Martin, D.L. (2000). Evidence of active NADP(+) phosphatase in dormant seeds of Avena sativa L. J Exp Bot *51*, 1389-1394.

Garbarino-Pico, E., Niu, S., Rollag, M.D., Strayer, C.A., Besharse, J.C., and Green, C.B. (2007). Immediate early response of the circadian polyA ribonuclease nocturnin to two extracellular stimuli. Rna *13*, 745-755.

Goldstrohm, A.C., and Wickens, M. (2008). Multifunctional deadenylase complexes diversify mRNA control. Nature reviews Molecular cell biology *9*, 337-344.

Gorelik, A., Heinz, L.X., Illes, K., Superti-Furga, G., and Nagar, B. (2016). Crystal Structure of the Acid Sphingomyelinase-like Phosphodiesterase SMPDL3B Provides Insights into Determinants of Substrate Specificity. The Journal of biological chemistry *291*, 24054-24064.

Green, C.B., Douris, N., Kojima, S., Strayer, C.A., Fogerty, J., Lourim, D., Keller, S.R., and Besharse, J.C. (2007). Loss of Nocturnin, a circadian deadenylase, confers

resistance to hepatic steatosis and diet-induced obesity. Proceedings of the National Academy of Sciences of the United States of America *104*, 9888-9893.

Hedstrom, L. (2008). IMP dehydrogenase-linked retinitis pigmentosa. Nucleosides Nucleotides Nucleic Acids *27*, 839-849.

Hee, S.W., Tsai, S.H., Chang, Y.C., Chang, C.J., Yu, I.S., Lee, P.C., Lee, W.J., Yun-Chia Chang, E., and Chuang, L.M. (2012). The role of nocturnin in early adipogenesis and modulation of systemic insulin resistance in human. Obesity *20*, 1558-1565.

Hentze, M.W., Caughman, S.W., Rouault, T.A., Barriocanal, J.G., Dancis, A., Harford, J.B., and Klausner, R.D. (1987). Identification of the iron-responsive element for the translational regulation of human ferritin mRNA. Science *238*, 1570-1573.

Hughes, K.L., Abshire, E.T., and Goldstrohm, A.C. (2018). Regulatory roles of vertebrate Nocturnin: insights and remaining mysteries. RNA Biol *15*, 1255-1267.

Kawai, M., Delany, A.M., Green, C.B., Adamo, M.L., and Rosen, C.J. (2010a). Nocturnin suppresses igf1 expression in bone by targeting the 3' untranslated region of igf1 mRNA. Endocrinology *151*, 4861-4870.

Kawai, M., Green, C.B., Lecka-Czernik, B., Douris, N., Gilbert, M.R., Kojima, S., Ackert-Bicknell, C., Garg, N., Horowitz, M.C., Adamo, M.L.*, et al.* (2010b). A circadian-regulated gene, Nocturnin, promotes adipogenesis by stimulating PPAR-gamma nuclear translocation. Proceedings of the National Academy of Sciences of the United States of America *107*, 10508-10513.

Kawai, S., Fukuda, C., Mukai, T., and Murata, K. (2005). MJ0917 in archaeon Methanococcus jannaschii is a novel NADP phosphatase/NAD kinase. The Journal of biological chemistry *280*, 39200-39207.

Kawai, S., Mori, S., Mukai, T., and Murata, K. (2004). Cytosolic NADP phosphatases I and II from Arthrobacter sp. strain KM: implication in regulation of NAD+/NADP+ balance. J Basic Microbiol *44*, 185-196.

Kojima, S., Gendreau, K.L., Sher-Chen, E.L., Gao, P., and Green, C.B. (2015). Changes in poly(A) tail length dynamics from the loss of the circadian deadenylase Nocturnin. Scientific reports *5*, 17059.

Kovalevich, J., and Langford, D. (2013). Considerations for the use of SH-SY5Y neuroblastoma cells in neurobiology. Methods in molecular biology *1078*, 9-21.

Lefterova, M.I., Haakonsson, A.K., Lazar, M.A., and Mandrup, S. (2014). PPARgamma and the global map of adipogenesis and beyond. Trends in endocrinology and metabolism: TEM *25*, 293-302.

Li, Y., Mouche, S., Sajic, T., Veyrat-Durebex, C., Supale, R., Pierroz, D., Ferrari, S., Negro, F., Hasler, U., Feraille, E.*, et al.* (2012). Deficiency in the NADPH oxidase 4 predisposes towards diet-induced obesity. Int J Obes (Lond) *36*, 1503-1513.

Linding, R., Russell, R.B., Neduva, V., and Gibson, T.J. (2003). GlobPlot: Exploring protein sequences for globularity and disorder. Nucleic acids research *31*, 3701-3708.

Liu, B.A., Jablonowski, K., Shah, E.E., Engelmann, B.W., Jones, R.B., and Nash, P.D. (2010). SH2 domains recognize contextual peptide sequence information to determine selectivity. Mol Cell Proteomics *9*, 2391-2404.

Marchler-Bauer, A., Bo, Y., Han, L., He, J., Lanczycki, C.J., Lu, S., Chitsaz, F., Derbyshire, M.K., Geer, R.C., Gonzales, N.R.*, et al.* (2017). CDD/SPARCLE: functional classification of proteins via subfamily domain architectures. Nucleic acids research *45*, D200-D203.

Marchler-Bauer, A., and Bryant, S.H. (2004). CD-Search: protein domain annotations on the fly. Nucleic acids research *32*, W327-331.

Marchler-Bauer, A., Derbyshire, M.K., Gonzales, N.R., Lu, S., Chitsaz, F., Geer, L.Y., Geer, R.C., He, J., Gwadz, M., Hurwitz, D.I.*, et al.* (2015). CDD: NCBI's conserved domain database. Nucleic acids research *43*, D222-226.

Marchler-Bauer, A., Lu, S., Anderson, J.B., Chitsaz, F., Derbyshire, M.K., DeWeese-Scott, C., Fong, J.H., Geer, L.Y., Geer, R.C., Gonzales, N.R.*, et al.* (2011). CDD: a Conserved Domain Database for the functional annotation of proteins. Nucleic acids research *39*, D225-229.

Mauxion, F., Faux, C., and Seraphin, B. (2008). The BTG2 protein is a general activator of mRNA deadenylation. The EMBO journal *27*, 1039-1048.

McLean, J.E., Hamaguchi, N., Belenky, P., Mortimer, S.E., Stanton, M., and Hedstrom, L. (2004). Inosine 5'-monophosphate dehydrogenase binds nucleic acids in vitro and in vivo. Biochem J *379*, 243-251.

Minor, L.K. (2005). Assays for membrane tyrosine kinase receptors: methods for high-throughput screening and utility for diagnostics. Expert Rev Mol Diagn *5*, 561-571.

Mittal, S., Aslam, A., Doidge, R., Medica, R., and Winkler, G.S. (2011). The Ccr4a (CNOT6) and Ccr4b (CNOT6L) deadenylase subunits of the human Ccr4-Not complex contribute to the prevention of cell death and senescence. Mol Biol Cell *22*, 748-758.

Mol, C.D., Kuo, C.F., Thayer, M.M., Cunningham, R.P., and Tainer, J.A. (1995). Structure and function of the multifunctional DNA-repair enzyme exonuclease III. Nature *374*, 381-386.

Mortimer, S.E., and Hedstrom, L. (2005). Autosomal dominant retinitis pigmentosa mutations in inosine 5'-monophosphate dehydrogenase type I disrupt nucleic acid binding. Biochem J *390*, 41-47.

Mortimer, S.E., Xu, D., McGrew, D., Hamaguchi, N., Lim, H.C., Bowne, S.J., Daiger, S.P., and Hedstrom, L. (2008). IMP dehydrogenase type 1 associates with polyribosomes translating rhodopsin mRNA. The Journal of biological chemistry *283*, 36354-36360.

Mullner, E.W., and Kuhn, L.C. (1988). A stem-loop in the 3' untranslated region mediates iron-dependent regulation of transferrin receptor mRNA stability in the cytoplasm. Cell *53*, 815-825.

Murphy, J.M., Farhan, H., and Eyers, P.A. (2017). Bio-Zombie: the rise of pseudoenzymes in biology. Biochem Soc Trans *45*, 537-544.

Nagy, E., Henics, T., Eckert, M., Miseta, A., Lightowlers, R.N., and Kellermayer, M. (2000). Identification of the NAD(+)-binding fold of glyceraldehyde-3-phosphate dehydrogenase as a novel RNA-binding domain. Biochem Biophys Res Commun *275*, 253-260.

Nagy, E., and Rigby, W.F. (1995). Glyceraldehyde-3-phosphate dehydrogenase selectively binds AU-rich RNA in the NAD(+)-binding region (Rossmann fold). The Journal of biological chemistry *270*, 2755-2763.

Navas, P., Minnifield, N., Sun, I., and Morre, D.J. (1986). NADP phosphatase as a marker in free-flow electrophoretic separations for cisternae of the Golgi apparatus midregion. Biochimica et biophysica acta *881*, 1-9.

Ng, A.C., Eisenberg, J.M., Heath, R.J., Huett, A., Robinson, C.M., Nau, G.J., and Xavier, R.J. (2011). Human leucine-rich repeat proteins: a genome-wide bioinformatic categorization and functional analysis in innate immunity. Proceedings of the National Academy of Sciences of the United States of America *108 Suppl 1*, 4631-4638.

Nikiforov, A., Dolle, C., Niere, M., and Ziegler, M. (2011). Pathways and subcellular compartmentation of NAD biosynthesis in human cells: from entry of extracellular precursors to mitochondrial NAD generation. The Journal of biological chemistry *286*, 21767-21778.

Oezguen, N., Schein, C.H., Peddi, S.R., Power, T.D., Izumi, T., and Braun, W. (2007). A "moving metal mechanism" for substrate cleavage by the DNA repair endonuclease APE-1. Proteins *68*, 313-323.

Prevot, D., Morel, A.P., Voeltzel, T., Rostan, M.C., Rimokh, R., Magaud, J.P., and Corbo, L. (2001). Relationships of the antiproliferative proteins BTG1 and BTG2 with CAF1, the human homolog of a component of the yeast CCR4 transcriptional complex: involvement in estrogen receptor alpha signaling pathway. The Journal of biological chemistry *276*, 9640-9648.

Rahman, M.M., El Jamali, A., Halade, G.V., Ouhtit, A., Abou-Saleh, H., and Pintus, G. (2018). Nox2 Activity Is Required in Obesity-Mediated Alteration of Bone Remodeling. Oxid Med Cell Longev *2018*, 6054361.

Richter, C. (1987). NADP+ phosphatase: a novel mitochondrial enzyme. Biochem Biophys Res Commun *146*, 253-257.

Rorbach, J., Nicholls, T.J., and Minczuk, M. (2011). PDE12 removes mitochondrial RNA poly(A) tails and controls translation in human mitochondria. Nucleic acids research *39*, 7750-7763.

Rouault, J.P., Prevot, D., Berthet, C., Birot, A.M., Billaud, M., Magaud, J.P., and Corbo, L. (1998). Interaction of BTG1 and p53-regulated BTG2 gene products with mCaf1, the murine homolog of a component of the yeast CCR4 transcriptional regulatory complex. The Journal of biological chemistry *273*, 22563-22569.

Schellenberg, M.J., Appel, C.D., Adhikari, S., Robertson, P.D., Ramsden, D.A., and Williams, R.S. (2012). Mechanism of repair of 5'-topoisomerase II-DNA adducts by mammalian tyrosyl-DNA phosphodiesterase 2. Nature structural & molecular biology *19*, 1363-1371.

Sherer, T.B., Trimmer, P.A., Borland, K., Parks, J.K., Bennett, J.P., Jr., and Tuttle, J.B. (2001). Chronic reduction in complex I function alters calcium signaling in SH-SY5Y neuroblastoma cells. Brain Res *891*, 94-105.

Solnestam, B.W., Stranneheim, H., Hallman, J., Kaller, M., Lundberg, E., Lundeberg, J., and Akan, P. (2012). Comparison of total and cytoplasmic mRNA reveals global regulation by nuclear retention and miRNAs. BMC genomics *13*, 574.

Stubblefield, J.J., Gao, P., Kilaru, G., Mukadam, B., Terrien, J., and Green, C.B. (2018). Temporal Control of Metabolic Amplitude by Nocturnin. Cell reports *22*, 1225-1235.

Temme, C., Zhang, L., Kremmer, E., Ihling, C., Chartier, A., Sinz, A., Simonelig, M., and Wahle, E. (2010). Subunits of the Drosophila CCR4-NOT complex and their roles in mRNA deadenylation. Rna *16*, 1356-1370.

Tresaugues, L., Silvander, C., Flodin, S., Welin, M., Nyman, T., Graslund, S., Hammarstrom, M., Berglund, H., and Nordlund, P. (2014). Structural basis for phosphoinositide substrate recognition, catalysis, and membrane interactions in human inositol polyphosphate 5-phosphatases. Structure *22*, 744-755.

Wang, H., Morita, M., Yang, X., Suzuki, T., Yang, W., Wang, J., Ito, K., Wang, Q., Zhao, C., Bartlam, M.*, et al.* (2010). Crystal structure of the human CNOT6L nuclease domain reveals strict poly(A) substrate specificity. The EMBO journal *29*, 2566-2576.

Wang, Y., Osterbur, D.L., Megaw, P.L., Tosini, G., Fukuhara, C., Green, C.B., and Besharse, J.C. (2001). Rhythmic expression of Nocturnin mRNA in multiple tissues of the mouse. BMC developmental biology *1*, 9.

Webster, M.W., Chen, Y.H., Stowell, J.A.W., Alhusaini, N., Sweet, T., Graveley, B.R., Coller, J., and Passmore, L.A. (2018). mRNA Deadenylation Is Coupled to Translation Rates by the Differential Activities of Ccr4-Not Nucleases. Mol Cell *70*, 1089-1100 e1088.

Wood, E.R., Bledsoe, R., Chai, J., Daka, P., Deng, H., Ding, Y., Harris-Gurley, S., Kryn, L.H., Nartey, E., Nichols, J.*, et al.* (2015). The Role of Phosphodiesterase 12 (PDE12) as a Negative Regulator of the Innate Immune Response and the Discovery of Antiviral Inhibitors. The Journal of biological chemistry *290*, 19681-19696.

Xiao, W., Wang, R.S., Handy, D.E., and Loscalzo, J. (2018). NAD(H) and NADP(H) Redox Couples and Cellular Energy Metabolism. Antioxid Redox Signal *28*, 251-272.

Yi, H., Park, J., Ha, M., Lim, J., Chang, H., and Kim, V.N. (2018). PABP Cooperates with the CCR4-NOT Complex to Promote mRNA Deadenylation and Block Precocious Decay. Mol Cell *70*, 1081-1088 e1085.

Ying, W. (2008). NAD+/NADH and NADP+/NADPH in cellular functions and cell death: regulation and biological consequences. Antioxid Redox Signal *10*, 179-206.

Zhang, Q., Yan, D., Guo, E., Ding, B., Yang, W., Liu, R., Yamamoto, T., and Bartlam, M. (2016). Structural basis for inhibition of the deadenylase activity of human CNOT6L. FEBS Lett *590*, 1270-1279.