Interactions of Spt5 with RNA and Xrn2 in transcriptional termination

Madeleine Eaton

Lincoln College



Supervised by Professors Jason Schnell and Lidia Vasilieva

Department of Biochemistry, University of Oxford

A thesis submitted in partial fulfilment of the requirements for the degree of Master of Science by Research in Biochemistry at the University of Oxford

April 2025

Acknowledgements

First and foremost, I would like to express my gratitude to my supervisors Jason Schnell and Lidia Vasilieva. Thank you for your guidance and advice; your enthusiasm for science is both inspiring and infectious.

I am grateful to all the members of the Schnell and Vasilieva labs, particularly Laurentien Jungkamp and Kiana Moradian, for all their kindness, support, and endless discussions about research, making it a pleasure to be in the lab.

Special thanks as well to Daniel Molland for his advice, supervision, and generous donation of the S75 column. Your insatiable thirst for knowledge and science is truly motivating and pushes me to become a better scientist.

And to Izzy Rhyu - "*'We'll be Friends Forever, won't we, Pooh?' asked Piglet. 'Even longer,' Pooh answered.*" Thank you for the distractions when needed, encouragement when it seemed impossible to continue, and always humbling me in the daily crossword.

Many thanks to Abby Southers for the great emotional support in and out of the lab and for the endless cups of coffee while cheering each other on. You have made Oxford feel more like home, and I truly appreciate your friendship and kindness throughout this journey.

My gratitude goes to Christina Redfield, for her expertise and support when conducting NMR experiments. I would also like to thank Krzysztof Kus, who has enough knowledge of transcription to fill an ocean. I am deeply grateful for his supervision at the bench and patience with my endless questions.

And finally, to my family for their unwavering love and support. Your belief in me has always been my greatest strength.

Abstract

RNA polymerase II (Pol II) mediates the transcription of protein coding genes, a fundamental process that is essential for cellular function. This process is tightly regulated by accessory factors, including Spt5. Spt5 is a conserved accessory elongation factor involved across all stages of transcription, including initiation, elongation, and termination. However, while the importance of Spt5 during transcriptional termination is well established, the mechanistic basis of its involvement is unclear.

In this study, I use nuclear magnetic resonance spectroscopy (NMR) to investigate the interactions of the *Schizosaccharomyces pombe* Spt5 KOWx45 domains with RNA and Xrn2. Phosphorylation of the C-terminal repeat region (CTR) of the KOWx45 domains greatly enhanced its interactions with RNA. The RNA-interacting residues localise to surface-exposed regions of the KOWx4 domain, suggesting a role in guiding nascent transcript exit. NMR was also used to characterize the interactions of the KOWx45 domains with the 5'-3' exoribonuclease Xrn2, primarily responsible for degrading the cleaved nascent transcript and dislodging Pol II from the DNA template. The KOW5 domain was implicated in mediating binding to Xrn2, with partial overlap between the identified binding interface and previously described sites of Pol II contact with transcription factors (Spt5). These results support the notion of remodelling and destabilisation of the Pol II complex during termination.

The findings presented here provide new insights into the regulatory role of Spt5 in contacting RNA and modulating activity of other accessory factors to regulate

transcriptional termination. In conclusion, this offers new perspectives on the current understanding of transcriptional termination.

Table of Contents

Acknowledgements *3*

Abstract *5*

Table of Contents *7*

List of Figures *9*

List of Tables *11*

Declaration *12*

Abbreviations *13*

Chapter 1: Introduction *16*

Transcription by RNA polymerase II **16**

Transcription Initiation 17

Transcription Elongation 19

Transcription Termination and nascent mRNA processing 20

Xrn2 ..... **23**

Spt5: a key conserved transcription accessory factor **27**

Spt5 domain architecture and structure 27

Role of Spt5 throughout transcription 29

Structural insights into Spt5 function 31

Novel Insights into transcriptional termination **33**

Links between Spt5 and Xrn2 33

Spt5 enhances Xrn2 activity via the KOW5 domain 36

Aims of this Study **37**

Use of NMR as an experimental technique 38

Focus on KOWx45 domains 38

Chapter 2: Purification and Assignment of KOWx45 construct via NMR *40*

Purification of KOWx45 Construct **40**

Mass Spectrometry of Purified KOWx45 Construct **42**

Characterization & Assignment of KOWx45 construct via NMR **43**

Chapter 3: Characterizing KOWx45 Interactions with RNA *48*

Identifying the RNA binding interface of KOWx45 and RNA **48**

***In vitro* phosphorylation of KOWx45 Construct** **54**

CTR Phosphorylation enhances KOWx45 RNA binding **57**

Chapter 4: Characterizing KOWx45 Interactions with Xrn2 ..... *64*

Purification of Xrn2 ..... **64**

Validation of Xrn2 Activity 65
Mapping of the Xrn2 binding site on Spt5 via NMR 67
Characterization of loop region of Xrn2 via NMR 72
Chapter 5: Discussion ***76***
What is the role of Spt5 CTR dephosphorylation in prompting termination? 77
How does Xrn2 bind to the Pol II complex? 80
Limitations & Future Directions 85
Conclusion 89
Chapter 6: Materials and Methods ***91***
Expression plasmids 91
RNA scaffolds 91
KOWx45 & Xrn2 construct cloning 91
Expression & Purification of KOWx45 93
SDS-Page Gel Analysis 95
Preparation of in-gel digest samples for mass spectrometry 96
Preparation of in-solution digest samples for mass spectrometry 97
Labelled KOWx45 & Xrn2 purifications 98
KOWx45 & RNA NMR titration setup 100
***In vitro* KOWx45 phosphorylation** 101
Expression & Purification of Xrn2 101
KOWx45 + Xrn2 NMR setup 102
Xrn2 EMSA 103
Xrn2 RNA activity assay 103
NMR experiments 104
Appendix ***106***
Protein Sequences 106
His-tagged KOWx45 construct 106
His-tagged Xrn2 construct 106
His-tagged Xrn2 cysteine mutant construct 106
Mass Spectrometry Data 107
Report of KOWx45 Chemical Shift Assignments 109
References ***119***

List of Figures

Figure 1.1 Illustration schematic of proposed models of transcription termination.

Figure 1.2 Xrn2 conservation and structure.

Figure 1.3 Spt5 domain architecture and structure.

Figure 1.4 Comparison of Spt5 domain architecture between species.

Figure 1.5 Cryo-EM structure of DSIF (Spt5 & Spt4) bound to Pol II elongation complex.

Figure 1.6 Evidence for a direct interaction between Spt5 and Xrn2.

Figure 1.7 Structure of a pre-termination complex with Xrn2 bound to Pol II.

Figure 1.8 Fluorescence anisotropy assay comparing RNA degradation of WT Xrn2 vs. WT Xrn2 in the presence of the KOW5-sCTR domain showing stimulation of Xrn2. activity.

Figure 1.9 AlphaFold3 prediction of KOWx45 construct focused on in this study.

Figure 2.1 Purification of KOWx45 protein construct.

Figure 2.2 Characterization and assignment of KOWx45 by NMR.

Figure 2.3 Talos-n predicted secondary structure for KOWx45 construct.

Figure 3.1 Chemical shift perturbation (ppm) plots of RNA addition into unphosphorylated KOWx45.

Figure 3.2 Plots showing global loss of signal intensity with RNA addition.

Figure 3.3 Identifying RNA binding interface of KOWx45.

Figure 3.4 Chemical shift perturbation (ppm) plots of 7mer RNA addition into unphosphorylated KOWx45.

Figure 3.5 *In vitro* phosphorylation of KOWx45 sCTR region.

Figure 3.6 Chemical shift perturbation (ppm) plots of RNA addition into phosphorylated KOWx45.

Figure 3.7 Identifying the RNA binding interface of pKOWx45.

Figure 3.8 Comparison of CSP residues with electrostatic surface potential of KOWx45 construct.

Figure 3.9 Structural prediction of phosphorylated KOWx45 construct

Figure 3.10 Identification of CSPs upon phosphorylation of KOWx45 construct

Figure 4.1 Optimization of purification of WT Xrn2.

Figure 4.2 Mobility shift assay characterizing WT Xrn2 RNA binding ability.

Figure 4.3 RNA degradation assay characterizing WT Xrn2 degradation of RNA.

Figure 4.4 Chemical shift perturbation (ppm) plots of Xrn2 addition into unphosphorylated KOWx45 vs. phosphorylated KOWx45.

Figure 4.5 Identifying the Xrn2 binding interface of KOWx45.

Figure 4.6 Chemical shift perturbation (ppm) plots of Xrn2 addition into KOWx45 & RNA complexes.

Figure 4.7 Testing interactions of RNA & KOWx45 with the loop of Xrn2.

Figure 5.1 Multiple sequence alignment of KOWx45 linker.

Figure 5.2 Insights into Spt5’s role in transcriptional termination.

Appendix 4.1 Depiction of magnetization transfer in triple resonance experiments.

Appendix 4.2 CARA strip representation of assignment workflow.

List of Tables

Table 6.1 Acquisition Parameters for NMR experiments

Appendix 1 Table of Hits from Mass Spectrometry Analysis of In-solution digest of KOWx45 Sample.

Appendix 2 Table of Hits from Mass Spectrometry Analysis of In-gel digest of top band of KOWx45 Sample.

Appendix 3 Table of Hits from Mass Spectrometry Analysis of In-gel digest of bottom band of KOWx45 Sample.

Appendix 5 Chemical shifts of backbone resonance assignments of KOWx45 construct

Declaration

All data described in this thesis was generated by the author, except as listed below:

- Introductory figures 1.6 & 1.8 were prepared by Dr. Krzysztof Kuś, Vasilieva lab.
- Mass spectrometry analysis of purified KOWx45 construct was performed by the Advanced Proteomics Facility, Department of Biochemistry, University of Oxford.
- CDK9 stock was purified by Dr. Tea Kecman, Vasilieva lab.

Abbreviations

AF3	AlphaFold3
TBP	TATA-box binding protein
B-TROSY / BEST-TROSY	Band selective excitation short-transient transverse relaxation optimized spectroscopy
BME	Beta-mercaptoethanol
BTFA	Bistrifluoroacetamide
C. thermophilum	*Chaetomium thermophilum*
CDK9	Cyclin dependent kinase 9
ChIP-seq	Chromatin Immunoprecipitation sequencing
CPF	Cleavage and polyadenylation factor
CPMG	Carr-Purcell-Meiboom-Gill relaxation dispersion experiments
CTD	C-terminal domain (RNA polymerase II)
CTR	C-terminal region (Spt5)
DSIF	DRB sensitivity inducing factor
DSS in D2O	Sodium trimethylsilylpropanesulfonate in deuterium oxide
DTT	Dithiothreitol
E. coli	*Escherichia Coli*
EMSA	Electrophoretic mobility shift assay
EtBr	Ethidium Bromide
FAM	5'-fluorecin amidite

H. sapiens	*Homo sapiens*
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
IMAC	Immobilized metal affinity chromatography
kDa	Kilodaltons
KOW	Kyrpides-Ouzounis-Woese
LB	Luria broth
mRNA	messenger RNA
MQ	Milli-Q water
NELF	Negative elongation factor
NET-seq	Native elongation transcript sequencing
PCR	Polymerase chain reaction
PIC	Pre-initiation complex
PMSF	Phenylmethylsulphonyl fluoride
ppm	Parts per million
Pol II	RNA polymerase II
PAS	Polyadenylation/Poly(A) site
PTC	Pre-termination complex
PTEF-b	Positive transcription elongation factor b
S. cerevisiae	*Saccharomyces cerevisiae* (budding yeast)
SEC	Size exclusion chromatography

S. pombe	*Saccharomyces pombe* (fission yeast)
SDS-PAGE	Sodium dodecyl sulphate polyacrylamide gel electrophoresis
Tris	Tris(hydroxymethyl)aminomethane
TSS	Transcriptional start site
TT-seq	Transient transcriptome sequencing
WT	Wild type

Chapter 1: Introduction

Transcription by RNA polymerase II

Transcription of the eukaryotic genome is performed by DNA-dependent RNA polymerases I, II, and III. RNA polymerase I (Pol I) transcribes ribosomal (r)RNA in the nucleolus, RNA polymerase III (Pol III) transcribes many small transcripts including transfer (t)RNA and spliceosomal small nuclear (sn)RNA U6, and RNA polymerase II (Pol II) transcribes messenger (m)RNA and a diverse class of non-coding RNAs (Girbig et al., 2022).

The transcription cycle consists of three main stages: initiation, elongation, and termination. Regulation at each stage affects gene transcription, maturation and processing of pre-mRNA into functional mRNA molecules, and contributes to mRNA levels (Kwak & Lis, 2013). Transcription is fundamental to life, and as such, dysregulation leads to a wide range of diseases, making an understanding of this process therapeutically relevant. For instance, viral infection of the host involves remodelling of normal cellular transcriptional pathways, aberrant gene expression is a hallmark of cancer, and several transcription factors are targets for treatment of neurodegenerative diseases such as Alzheimer's and Huntington's (Bradner et al., 2017; Jin et al., 2019; Ungerleider et al., 2024).

Therefore, a significant amount of work has been dedicated to gaining detailed mechanistic insights into the structural organisation of Pol II complexes associated with transcription and RNA processing factors to understand their function throughout transcription (Bernecky et al., 2016; Ehara et al., 2017; Han et al., 2023; Yanagisawa

et al., 2024). Our understanding of transcriptional termination is comparatively limited relative to that of transcriptional initiation and elongation (Girbig et al., 2022; Kwak & Lis, 2013; Sainsbury et al., 2015). Therefore, my thesis research focuses on gaining mechanistic insights into termination (using *Schizosaccharomyces pombe* as a model organism), with particular focus on the role of the transcriptional accessory factor Spt5 in this process.

This chapter outlines the current understanding of Spt5's role during the transcriptional cycle and highlights existing gaps in the literature regarding its role in termination. This aims to establish the main objectives of this study as outlined further below.

Transcription Initiation

Eukaryotic transcriptional initiation begins with stepwise formation of the pre-initiation complex (PIC) at the transcriptional start site (Sainsbury et al., 2015). Initially, TFIIA and TFIID are deposited on the DNA promoter via the TATA-box binding protein, TBP (Sainsbury et al., 2015). These general transcription factors (GTFs) recruit the 12-subunit Pol II, and remaining GTFs TFIIB, TFIIF, TFIIE, and TFIIH (Girbig et al., 2022). The Mediator complex bridges the DNA promoter, Pol II, and GTFs to enhance PIC assembly (Grünberg & Hahn, 2013).

In higher eukaryotes, Pol II undergoes promotor proximal pausing after transcription of an initial 20-120 nucleotide stretch of RNA (Rasmussen & Lis, 1993). Pause release into productive elongation is mediated by the CDK9 kinase subunit of positive transcription elongation factor b (PTEF-b) that mediates release of paused Pol II into elongation by phosphorylating the C-terminal tail (CTD) of the largest subunit of Pol II

and the C-terminal region of conserved elongation factor Spt5 (Chen et al., 2018). Rpb1 is the largest of the 12 subunits of Pol II and contains the active site as well as the repetitive CTD. The human Pol II CTD has 52 repeats (21 consensus and 31 non-consensus) while *S. pombe* Pol II CTD has 29 repeats with the consensus sequence YSPTSPS (Buratowski, 2009; Jasnovidova & Stefl, 2013; Proudfoot, 2004). Reversible phosphorylation of the Pol II CTD regulates transcription and RNA processing by mediating recruitment of factors specific to each stage of transcription. Ser2 and Ser5 are the major phosphorylation sites marking each stage of transcription: Ser5 phosphorylation occurs during initiation & elongation and recruits capping enzymes, and Ser2 phosphorylation occurs during termination and recruits the 3' end processing machinery (Eick & Geyer, 2013; Komarnitsky et al., 2000).

Spt5 and its cofactor Spt4 together form DRB sensitivity inducing factor (DSIF), which associates with Pol II throughout the entire transcription cycle. Spt5 phosphorylation allows release of promoter proximal paused Pol II during initiation(Yamada et al., 2006). Although not conserved in yeast, in higher eukaryotes, CDK9 also phosphorylates negative elongation factor (NELF), leading to its dissociation from Pol II. NELF is recruited to Pol II by unphosphorylated Spt5 leading to pausing (Vos et al., 2018). Phosphorylation of the Pol II CTD on Ser2 recruits RNA processing factors and phosphorylation of the DSIF CTR allows RTF1 binding, a subunit of the PAF1 complex which stimulates Pol II activity (Vos et al., 2020). PAF, along with transcription factor Spt6, allosterically stimulate Pol II transcription by binding to the Pol II core to aid in the transition towards productive elongation (Girbig et al., 2022). Spt6 also aids transcription through nucleosomes (Aoi et al., 2022; Vos et al., 2018).

Structures of paused Pol II with DSIF and NELF bound showed that the DNA-RNA hybrid scaffold becomes tilted, preventing binding of nucleotide triphosphates and thus further elongation (Vos et al., 2018). Binding of NELF is also mutually exclusive with TFIIS binding, an elongation factor which activates paused Pol II after backtracking (Vos et al., 2018). This promotor-proximal pausing is proposed to allow a 'window of opportunity' for recruitment of additional RNA capping, processing, and elongation factors.

Transcription Elongation

Synthesis of pre-mRNA nascent transcripts during elongation is coupled to forward translocation of Pol II via a Brownian ratchet mechanism. In this mechanism, the pre-translocation state that contains the base-paired DNA-RNA hybrid in the Pol II active site 'ratchets' forward into the post-translocation state, thus allowing nucleotide triphosphate incorporation and bond formation (Dangkulwanich et al., 2013).

Pol II pausing or stalling during elongation can be caused by a multitude of factors, including RNA secondary structure, polymerase backtracking, the DNA template sequence, or nucleosomes (Noe Gonzalez et al., 2021). For example, weaker A-T repetitive DNA sequences lead to weaker DNA-RNA hybrids and more frequent pausing of Pol II (Zamft et al., 2012). Native elongation transcript sequencing (NET-seq) studies in yeast showed that greater than 70% of pause sites in gene bodies were enriched in adenosine at the RNA 3' end (Churchman & Weissman, 2011). From this paused state, Pol II can either backtrack, terminate, or exit from the pause to continue elongation (Noe Gonzalez et al., 2021). Backtracked Pol II complexes are arrested and unable to proceed unless released by endonucleolytic cleavage of the transcript. This

process is catalysed by the Pol II active site and activated by transcription factor TFIIS (in eukaryotes) or GreB (in bacteria) (Schweikhard et al., 2014).

To achieve productive transcriptional elongation, the presence of regulatory factors is essential to control Pol II progression and prevent premature termination. Accessory elongation factor Spt5 (discussed further below) enhances processivity and prevents pausing by maintaining a closed Pol II active centre cleft. In other cases, Spt5/NusG can also induce pausing through interaction with the non-template strand of DNA or via facilitating recruitment of NELF in higher eukaryotes (Hu et al., 2021a).

Transcription Termination and nascent mRNA processing

In eukaryotes, transcriptional termination requires endonucleolytic cleavage of the nascent transcript downstream of the consensus element polyadenylation signal (PAS) by the pre-mRNA 3' end processing machinery cleavage polyadenylation factor (CPF) and destabilisation of the elongation complex by the 5' to 3' exoribonuclease Xrn2 (Kim et al., 2004; Kuś et al., 2025; West et al., 2004; Yanagisawa et al., 2024; Zeng et al., 2024). Xrn2 is recruited to the Pol II associated 5'-monophosphorylated downstream RNA cleavage product and dislodges Pol II from DNA. Transitioning into termination is facilitated via dephosphorylation of Spt5 and the Pol II CTD (on the Ser5 position in *S. pombe*). This is accomplished by protein phosphatase 1 (PP1) in mammals and PP1 homolog Dis2 in fission yeast (Kecman et al., 2018; Parua et al., 2018). Dis2 is a component of the phosphatase module of the larger cleavage and polyadenylation factor (CPF) (Ho, Vasilieva et al., 2025, unpublished manuscript). After transcription through the poly(A) site, the CPF binds the Ser2 phosphorylated Pol II CTD before cleavage and polyadenylation of the nascent mRNA (Girbig et al., 2022).

Transcriptional termination is also tightly linked to RNA 3' end processing. Preventing cleavage of the nascent pre-mRNA downstream of the PAS by the CPF leads to failure to terminate inactivating nonnucleolytic activity (Birse et al., 1998). This is because cleavage of pre-mRNA by the CPF generates 5' monophosphorylated ends able to be degraded by Xrn2 (Kecman et al., 2018). Transcription through the poly(A) site is also associated with deceleration of Pol II, with transcription rate dropping from greater than 2kb/min to less than 1kb/min (Cortazar et al., 2019). Deceleration correlates with poly(A) site recognition and dephosphorylation of the Spt5 CTR by Dis2/PP1. Blocking poly(A) site recognition through mutation of the consensus sequence prevented Pol II deceleration and led to accumulation at the 3' end of genes. Furthermore, Spt5 CTR phosphorylation correlates with specific stages of transcription. Phosphorylation by P-TEFb during initiation converts Spt5 into a positive regulator of elongation and allows release of promoter proximal paused Pol II, while poly(A) site recognition during termination correlates with loss of Spt5 phosphorylation. Mutation of the PP1 complex both reduces Spt5 dephosphorylation during termination and increases readthrough transcription through gene bodies and 3' regions, preventing efficient termination (Cortazar et al., 2019).

Two models for removal of Pol II after nascent mRNA cleavage have been proposed: the allosteric and torpedo models (Fig. 1.1). Both are supported by experimental evidence and likely are not mutually exclusive (Eaton et al., 2020). The allosteric model proposes that transcription past the poly(A) site and binding of termination factors causes conformational changes in the elongation complex leading to dissociation. Pcf11, a component of the cleavage and polyadenylation complex is sufficient to dissociate the elongation complex alone via binding to Pol II and RNA, however these

studies were conducted *in vitro* (Zhang & Gilmour, 2006). In contrast, the torpedo model proposes that exoribonucleolytic degradation of the nascent mRNA dissociates Pol II from DNA. Depletion of the exoribonuclease Xrn2 leads to read-through transcription, while CPF depletion also causes defective termination due to the requirement for cleavage to generate appropriate substrates for Xrn2 (Eaton et al., 2018; Nojima et al., 2015; Birse et al., 1998). Despite this, Xrn2 mediated degradation of nascent RNA alone is not enough to lead to termination, implying additional factors or steps are necessary for this process (Luo et al., 2006).

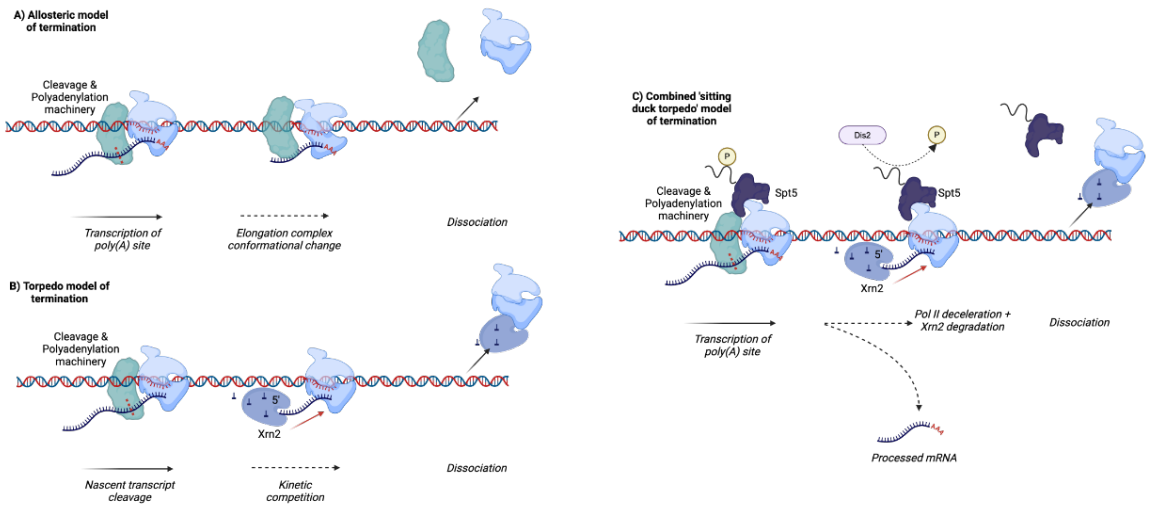


Figure 1.1. Illustration schematic of proposed models of transcription termination. *A) Allosteric model. B) Torpedo model. C) Combined 'sitting duck torpedo' model.*

It is likely that transcriptional termination involves aspects of both models, culminating in a proposed 'sitting duck torpedo' model of termination (Cortazar et al., 2019). Here, Spt5 dephosphorylation, triggered by poly(A) site transcription, leads to a conformational change and complex deceleration. Deceleration of Pol II and the elongation complex post-cleavage of the nascent mRNA is important as it allows the ultimate dislodgement of Pol II from the DNA template by Xrn2. CPF cleavage of the

nascent transcript provides a free 5' phosphate for Xrn2 to degrade and 'chase down' the still transcribing Pol II. Kinetic competition between the two moving components of Pol II deceleration and Xrn2 degradation determines the site of termination. Supporting this, Pol II mutants with increased or decreased transcription rates compared to their wild-type counterpart led to faster or delayed termination respectively (Fong et al., 2015). Additionally, there is evidence supporting the idea that recruitment of Xrn2 and RNA processing factors is interdependent and that these processes are heavily linked (Luo et al., 2006). Depletion of Xrn2 leads to reduced efficiency of transcript processing at poly(A) sites, and recent structures indicate Xrn2 binding contributes to remodelling of the elongation complex to allow termination and processing factors to bind to the Pol II surface (Dikunova et al., 2025; Kuś et al., 2025; Yanagisawa et al., 2024).

Continual degradation of the cleaved nascent transcript increases RNA tension between Xrn2 and Pol II. When Xrn2 contacts the Pol II complex, it destabilizes the transcription bubble and eventually triggering a collapse and RNA release and dislodgement of Pol II from the DNA template.

Despite this, the mechanism of how Spt5 dephosphorylation leads to deceleration of the elongation complex initially is unknown.

Xrn2

Xrn2 is a 5'-3' exoribonuclease that is important for termination of gene transcription, removal of aberrant mRNAs, processing small nucleolar RNA, and degradation of miRNAs and tRNAs (Chang et al., 2011). As a result, it is an essential protein in both *Saccharomyces cerevisiae* and *Schizosaccharomyces pombe (*Amberg et al., 1992;

Chernyakov et al., 2008; Kuś et al., 2025). Xrn2 is selective for RNA substrates with 5' monophosphorylated ends and is highly processive. In yeast, Xrn2 has an interacting partner Rai1 which enhances Xrn2 activity and stability (Xiang et al., 2009). Rai1 has putative pyrophosphohydrolase activity allowing it to produce new monophosphorylated RNA substrates suitable for Xrn2 binding (Xiang et al., 2009).

Crystal structures of *S. pombe* Xrn2 in complex with its interacting partner Rai1 show that Xrn2 contains two conserved regions CR1 & CR2, which together form the catalytic centre (Xiang et al., 2009). CR1 and CR2 are connected by a long unstructured region (Fig. 1.2b). Seven conserved acidic residues in the active site coordinate two divalent cations required for hydrolysis of nucleotide phosphodiester bonds (Fig. 1.2a). Substrate selectivity is achieved through a basic binding pocket, which sterically accommodates only 5' monophosphorylated RNA substrates without clashes (Fig. 1.2c). Two conserved histidine (His) and tryptophan (Trp) residues also base stack with substrate nucleotides and ensure enzyme processivity.

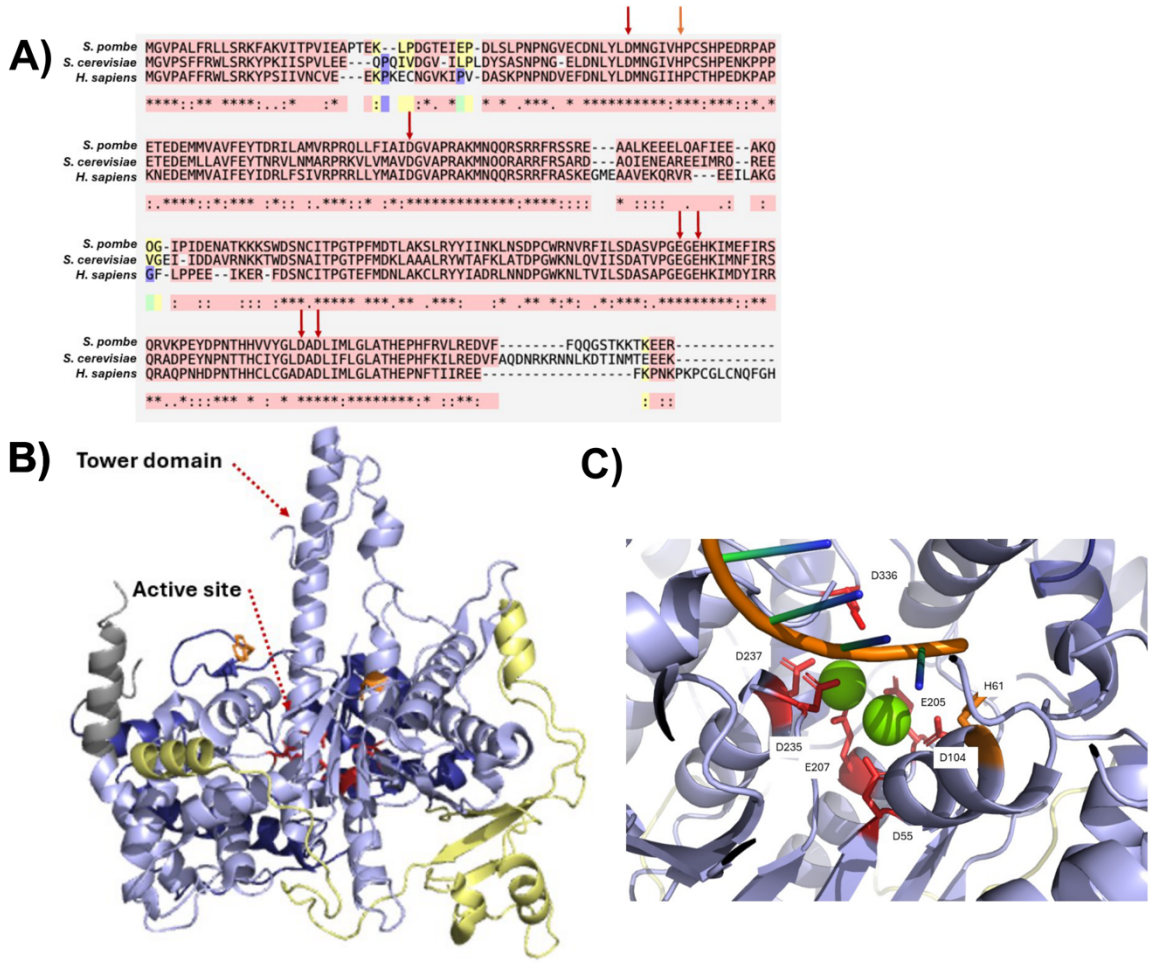


Figure 1.2. Xrn2 conservation and structure. *A) Sequence alignment of S. pombe, S. cerevisiae, and H. sapiens Xrn2 with highly conserved residues highlighted in orange. Alignment produced using TCoffee software. Conserved base stacking His/Trp residues indicated in orange and conserved active site Asp/Glu residues indicated in red. B) Crystal structure of S. pombe Xrn2 (PDB 3FQD) residues 1-885. Conserved region 1 (CR1) shown in light blue, conserved region 2 (CR2) shown in dark blue, linker shown in gray (density not fully modelled), and C-terminal region shown in yellow. Active site and tower domain also indicated. Base stacking residues His/Trp highlighted in orange and active site residues highlighted in red with side chains shown. C) AF3 predicted structure of full-length S. pombe Xrn2 in complex with 7mer RNA scaffold and 2 divalent cations. Conserved region 1 (CR1) shown in light blue, conserved region 2 (CR2) shown in dark blue, linker shown in gray (density not fully modelled), and C-terminal region shown in yellow. Conserved His residue (orange) stacks with 5' phosphate end of RNA. Active site residues (red) coordinate divalent cations in active site to aid hydrolysis.*

A combination of static structural studies as well as NMR dynamics experiments have elucidated the catalytic cycle of Xrn2. The active site of the enzyme is found in front of the tower domain helix α4, which contributes several conserved active site residues.

Interestingly, homologs of Xrn1 (a related exoribonuclease) lack the C-terminal portion of the tower domain helix. This indicates the domain could be specific to Xrn2 and important for its unique function during transcriptional termination (Xiang et al., 2009). In the pre-translocation state, three bases of RNA substrate are sandwiched between the conserved Trp/His residues. Translocation of the 5' phosphate into the basic binding pocket leads to a conformational change into the active state (Overbeck et al., 2022). This is accompanied by rearrangement of the N-terminal α1 helix in CR1. This helix is proposed to act as a Brownian ratchet supporting substrate translocation into the active state. Translocation to the active state is followed by hydrolysis of the nucleotide phosphodiester bond and product release into the post-translocation state, allowing the RNA substrate to move one base along and readopt the pre-translocation state (Overbeck et al., 2022).

Interestingly, cryo-EM structures comparing the Xrn2-Rai1 complex from both *S. cerevisiae* and *C. thermophilum* identified that while the core structures and fold remained the same, the interaction interfaces of Xrn2 to Rai1 and Pol II varied between species (Dikunova et al., 2025). The binding mechanism of *S. cerevisiae* Xrn2 to Pol II allows for increased flexibility in binding to the Pol II CTD, whereas the *C. thermophilum* complex has restricted conformational space for CTD binding (Dikunova et al., 2025).

This shows the potential for differences in identical Xrn2 termination factor complexes between organisms and opens interesting questions about whether this is relevant for the NMR studies of the Xrn2 catalytic cycle and dynamics, which were performed with Xrn2 from *C. thermophilum.* This observation informed the decision to build on these

studies here using Xrn2 from *S. pombe* to gain specific insight into its yeast form and regulation.

Spt5: a key conserved transcription accessory factor

Spt5 domain architecture and structure

RNA polymerase II (Pol II) requires a multitude of accessory factors for productive transcriptional elongation. *In vitro* studies demonstrated that Pol II alone is unable to efficiently transcribe chromatin at the rates seen *in vivo*, suggesting an essential role for additional proteins within the elongation complex throughout transcription (Izban & Luse, 1992).

Spt5 is one such accessory elongation factor that in eukaryotes acts as a heterodimer in complex with Spt4, collectively known as DSIF (DRB sensitivity inducing factor). Eukaryotic Spt5 is a large multidomain protein, consisting of a series of individual structured domains all connected by long and flexible linkers (Fig. 1.3b). The N-terminal acidic region is followed by a NusG N-terminal (NGN) domain responsible for heterodimer formation as well as five Kyrpides-Ouzounis-Woese (KOW) domains and a final C-terminal repetitive region (Decker, 2021) (Fig. 1.3a).

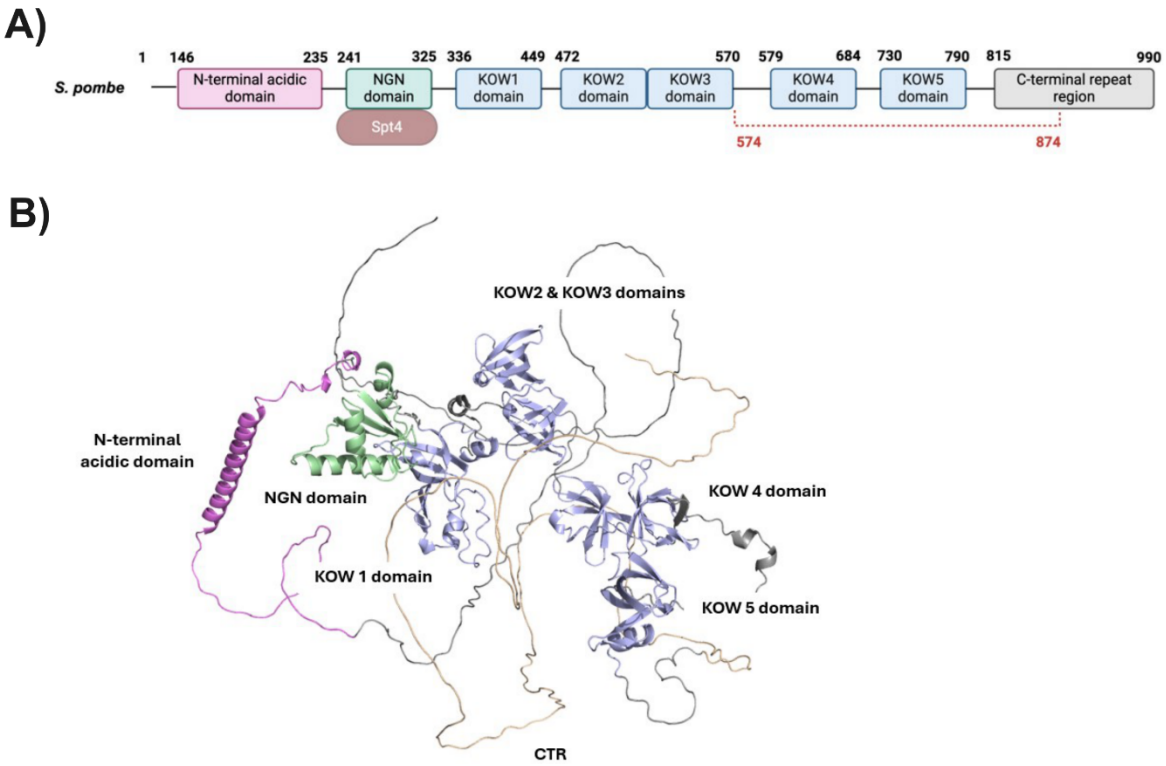


Figure 1.3. Spt5 domain architecture and structure. *A) Domain architecture schematic of S. pombe Spt5. Domain boundaries of the construct used in this study are highlighted in red. B) AF3 predicted structure of full-length S. Pombe Spt5 with domains labelled showing large amount of disorder present due to linkers.*

KOW domains adopt the same fold as the larger Tudor domain family, small proteins which mediate protein-RNA and protein-protein interactions (Lasko, 2010; Ponting, 1997). The conserved architecture consists of a five stranded beta-barrel structure (Zuber et al., 2022). Metazoan and plant Spt5 contain putative additional KOW domains 6 & 7 after the CTR region (Zuber et al., 2018). The bacterial Spt5 homolog NusG is comprised only of the NGN domain followed by one KOW domain, however, bacteria have no homolog of Spt4 (Meyer et al., 2015).

Interestingly, previous NMR solution structures of the KOW4 domain revealed the presence of an N-terminal structural element, collectively referred to here as KOWx4. The domain consists of the canonical KOW4 domain β-barrel fold covered by a ‘lid’

composed of a β-hairpin. The N-terminal element is connected to the domain via a cationic linker and consists of an antiparallel β-sheet formed by four β-strands but lacks the C-terminal β-strand present in typical KOW domains (Zuber et al., 2018). The KOWx element is required for stabilization of the greater KOWx4 domain overall.

Role of Spt5 throughout transcription

The Spt5/NusG protein family is the only transcription factor conserved between eukaryotes, bacteria, and archaea (Harris et al., 2003; Peterlin & Price, 2006) (Fig. 1.4). The Spt5/NusG gene is also essential for viability in yeast, zebrafish, Drosophila, and mammalian & bacterial cells, further reinforcing its ancient and central role in regulating the transcriptional cycle (Hartzog & Fu, 2013).

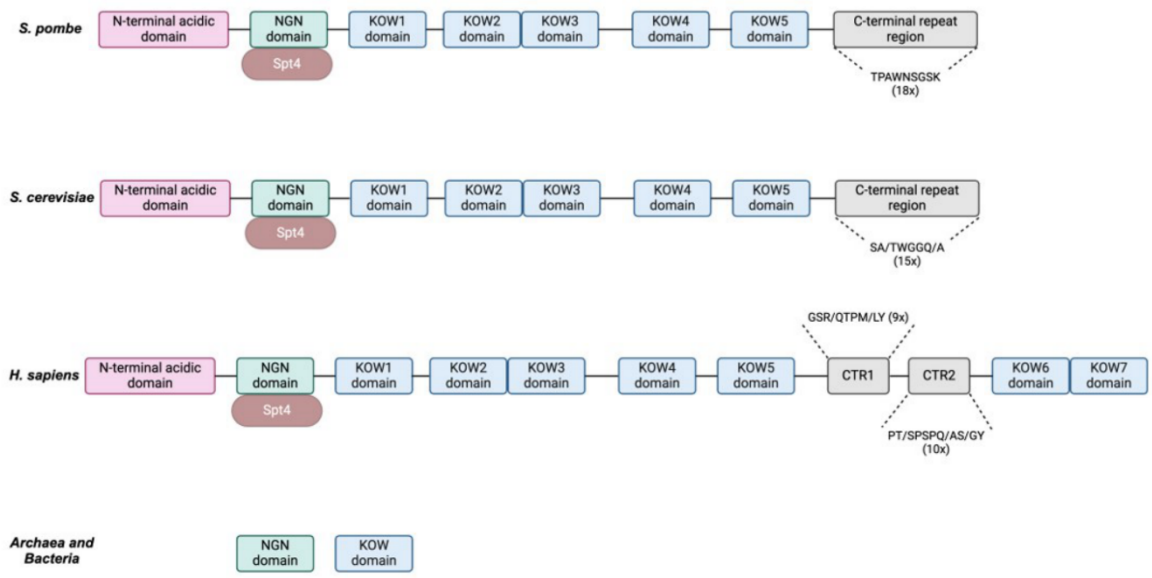


Figure 1.4. Comparison of Spt5 domain architecture between species. *CTR consensus sequence and repeat numbers are shown. Domain organisation is based on (Shetty et al., 2017; Zhou et al. 2009).*

Chromatin immunoprecipitation sequencing (ChIP-seq) studies demonstrated that DSIF associates with the elongation complex throughout transcription from the

transcriptional start site (TSS) to the termination site, with individual mutation of the *spt4* or *spt5* gene negatively impacting Pol II elongation rate and processivity *in vivo* (Glover-Cutter et al., 2008; Mason & Struhl, 2005).

Transient transcript sequencing (TT-seq) experiments using rapid auxin degradation of Spt5 resulted in overall reduced RNA synthesis, delayed termination, and defects in splicing and 3' end processing (Kuś et al., 2025). Furthermore, elongation complexes containing Spt5 mutants lacking the KOW5 domain and CTR accumulated 3' end extended RNAs (Baejen et al., 2017). This supports the idea that one important function of Spt5 is supporting transition towards termination and 3' end processing.

Spt5 also participates in licencing functional Pol II complexes at gene promoters, allowing them to progress into elongation (Kuś et al., 2025). Spt5 loss correlates with accumulation of transcriptionally engaged Pol II at the promoter-proximal pause site and reduced global RNA transcript levels. The small percentage of Pol II complexes which are able to progress towards elongation despite this loss have defective splicing and RNA processing, functions which rely at least partially on Spt5 for factor recruitment to Pol II.

Additionally, Spt5 is implicated in multiple other aspects of transcription (Aoi et al., 2021; Decker, 2021). Rapid degradation of Spt5 in cells led to significant decreases of the major Pol II subunit Rpb1, indicating a role in maintaining stability of the polymerase (Hu et al., 2021b).

Structural insights into Spt5 function

CryoEM structures of transcribing mammalian Pol II showed that Spt5's flexible nature imparts a modular structure, in which individual domains contact different regions of the Pol II surface to accomplish varied functions. The Spt5 KOW1 and NGN domains, with Spt4 strongly associated, form a DNA exit tunnel, bridging the Rpb1 clamp and Rpb2 core modules (Bernecky et al., 2016) (Fig. 1.5). The KOW1 domain also contacts the minor groove of template DNA via basic and polar residues. This DNA exit tunnel clamp locks DNA into the appropriate orientation in the active site, separates the upstream and downstream DNA strands, and helps to maintain a closed active centre cleft for Pol II processivity (Ehara et al., 2017a). This is a key function of Spt5, and mutation of KOW1 in yeast is lethal (Dollinger et al., 2023).

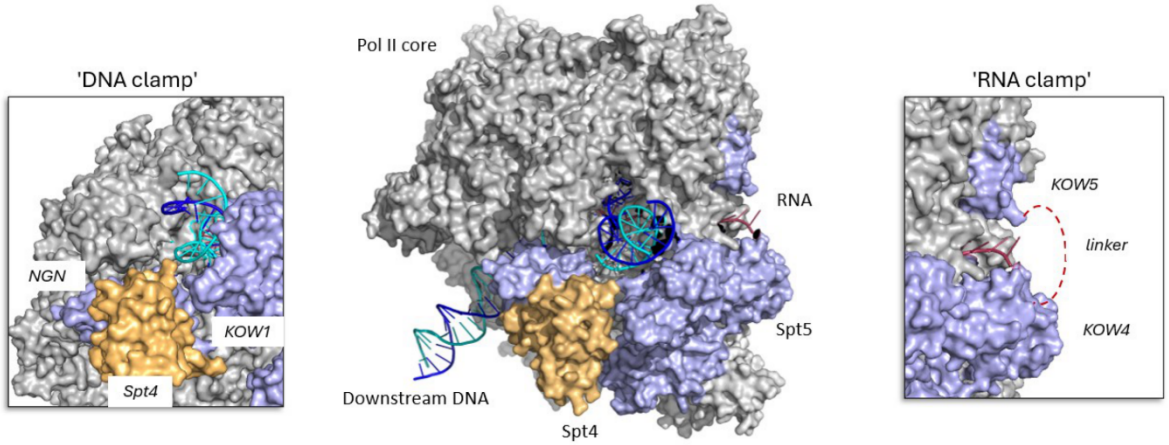


Figure 1.5. Cryo-EM structure of DSIF (Spt5 & Spt4) bound to Pol II elongation complex *(PDB 5OIK).*

Accordingly, in Spt5 depleted cells, Pol II preferentially accumulates and prematurely terminates on T-rich DNA regions, indicating a dependence on Spt5 for maintaining stability and processivity during transcript elongation (Han et al., 2023).

Similarly to the KOW1 and NGN domains, the KOW4 and KOW5 domains are predicted to form an RNA clamp around exiting nascent RNA, preventing Pol II dissociation and increasing stability (Bernecky et al., 2016) (Fig. 1.5). Crosslinking data showed KOW5 contacts exiting RNA, and KOW5 binding to Pol II blocks one of two RNA exit paths, thus it plays a role in directing exiting RNA appropriately (Kuś et al., 2025; Qiu & Gilmour, 2017). KOW4 contains a basic surface patch thought to contribute to binding to RNA along with the Rpb7 stalk, and UV crosslinking results showed potential transient contacts between KOW5 and exiting RNA (Ehara et al., 2017; Qiu & Gilmour, 2017). Mutation of the KOWx45 domains abolished RNA binding and caused transcriptional defects, leading to the conclusion that these domains are particularly important in maintaining Spt5 binding to RNA and the Pol II surface (Qiu & Gilmour, 2017).

KOW5 is also thought to restrict the core modules of Pol II to further enhance processivity of the elongation complex. While not able to be visualized structurally, the CTR is predicted to protrude from the complex, serving as a recruitment platform for RNA processing and termination factors as well as modulating the state of transcription via dynamic phosphorylation/dephosphorylation (Ehara et al., 2017b).

However, structural methods such as cryoEM and crystallography have been unable to visualize the linker regions and CTR of Spt5 due to their high disorder. This has left a gap in understanding about their roles in transcription and in particular termination, where these regions are known to play key roles, but mechanistic insight into how these functions are achieved is lacking.

Novel Insights into transcriptional termination

As previously discussed, the current model of transcriptional termination is known as the ‘sitting duck torpedo’ model, where Xrn2 chases down the transcribing Pol II to dislodge it from the DNA template.

However, novel data published by the Vasilieva lab has expanded on this conventional view of termination, showing that Xrn2 interacts with Pol II and Spt5 directly and possibly much earlier in the transcriptional cycle than previously proposed (Kuś et al., 2025; Yanagisawa et al., 2024; Zeng et al., 2024). In fact, Xrn2 exonucleolytic activity is enhanced by Spt5, however the mechanism for this is not understood (Kuś et al., 2025). These results open many questions into the nature of Spt5’s role in this interaction and lay the foundation for much of what this thesis aims to address.

Links between Spt5 and Xrn2

Proteomics work analysing interacting components from a pull-down of FLAG-tagged Spt5 identified high enrichment of termination machinery components (Fig. 1.6a). Alongside known Spt5 interactors including Pol II subunits and Spt4, Xrn2 and its partner Rai1 were also enriched. Reconstitution of the Pol II-DSIF elongation complex found that Xrn2 and Rai1 were able to stably bind, forming a proposed ‘pre-termination complex’ (Kuś et al., 2025). This suggests a novel identified interaction between Spt5 and Xrn2 that could occur earlier than previously thought.

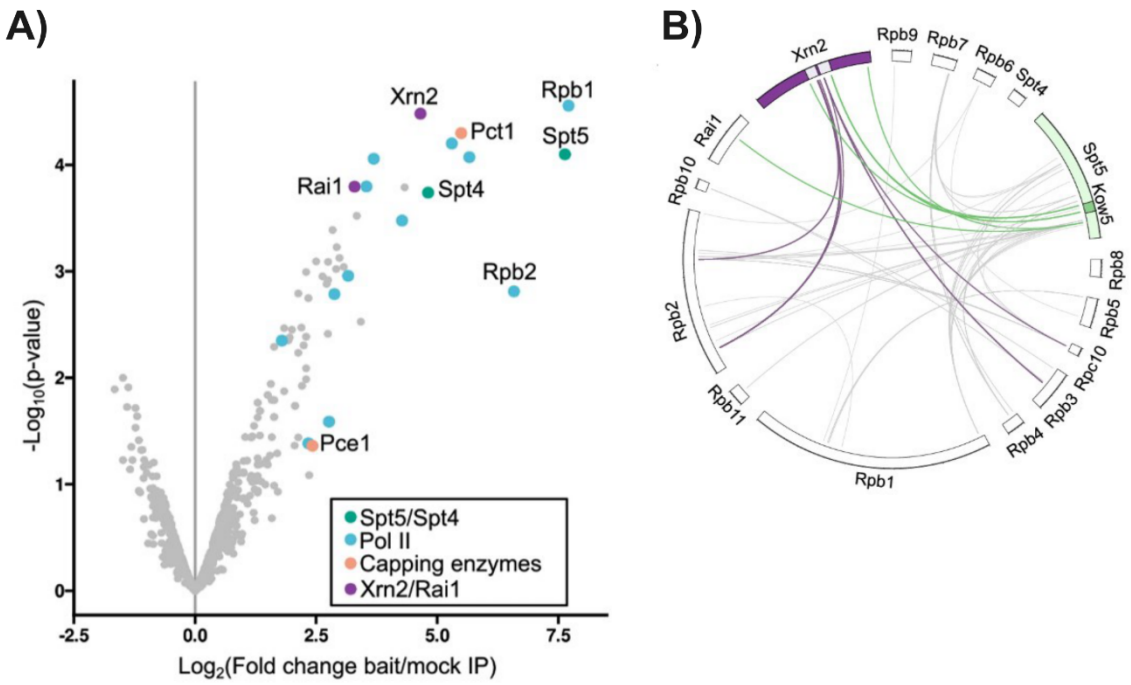


Figure 1.6. Evidence for an association between Spt5 and Xrn2. *A) Volcano plot showing enrichment of termination complex proteins such as Xrn2 and Rai1 after Spt5-Flag purification. B) BS3 crosslinking of inactivated Xrn2, Rai1, DSIF, Pol II and DNA-RNA hybrid scaffold following by mass spectrometry analysis, showing preferential crosslinking of Xrn2 to the KOW5-sCTR region. Crosslinks between Xrn2/Rai1 and Spt5 are highlighted in green.*

Crosslinking of this reconstituted pre-termination complex (PTC) followed by mass spectrometry showed specific interactions between Xrn2 and components of the Pol II surface as well as within the KOW5 domain of Spt5 (Kuś et al., 2025) (Fig. 1.6b).

CryoEM structures of the PTC lacked significant Xrn2 density but indicated that Xrn2 binds at least partially to Pol II using an α-helical region within its large unstructured loop. Ionic interactions between arginine residues within the helix and a negatively charged Pol II pocket are proposed to anchor Xrn2 to the surface. Mutation or removal of these charged residues reduced interactions between Xrn2 and Pol II both *in vitro* and *in vivo* (Kuś et al., 2025).

However, the binding interface of the U1 splicing machinery complex also utilizes this negatively charged Pol II pocket, indicating that there may be competition between splicing and termination components (Kuś et al., 2025). Xrn2 may be prevented from forming a pre-termination complex with Pol II during elongation (even when present in the gene body) due to binding of the U1 splicing complex. Xrn2 binding was also demonstrated to be more extensive in other studies, encompassing other regions of Pol II (Dikunova et al., 2025; Yanagisawa et al., 2024; Zeng et al., 2024).

A separate cryoEM structure of Xrn2/Rai1 associated with the Pol II complex found overlaps between the proposed Xrn2-Pol II binding interface and the Spt5-Pol II binding interface, specifically the NGN and KOW1-4 domains, indicating binding of Xrn2 and Spt5/Spt6/PAF1 may be mutually exclusive (Yanagisawa et al., 2024) (Fig. 1.7). Spt5 may remain bound to Pol II via the KOW5 domain, yet removal of the DNA exit tunnel and weakening of the RNA clamp would in theory contribute to destabilization of Pol II and ultimate termination (Zeng et al., 2024). Transcription factor swapping is well documented during the initiation-elongation transition, with genome-wide occupancy profiles of Pol II transcription factors in *S. cerevisiae* showing a distinct transition between initiation and elongation factors, including DSIF, 150 base pairs downstream of the TSS (Mayer et al., 2010). Analogously, a two-step transition was observed at the poly(A) site involving exit of elongation factors and recruitment of termination factors, such as Pcf11. The above structural studies align and expand on this idea, with Xrn2 binding potentially contributing to an elongation-termination transition via partial removal of elongation factor Spt5. The termination factor Seb1 binding interface also overlaps with that of Spt5, and Spt5 dephosphorylation is required to trigger its recruitment (Kecman et al., 2018). Additionally, interactions

between KOW5 and Xrn2 connect the nascent RNA exit tunnel to the RNA entry channel of Xrn2, further facilitating Pol II displacement (Zeng et al., 2024).

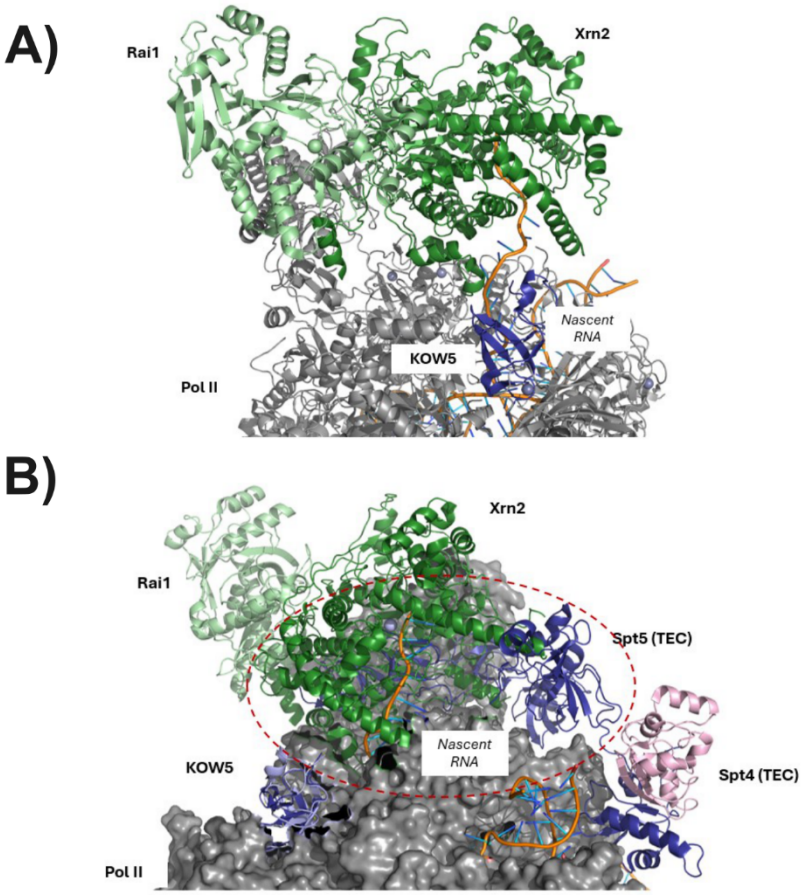


Figure 1.7. Structure of a pre-termination complex with Xrn2 bound to Pol II. *A) Cryo-EM structure of S. cerevisiae Xrn2 (green) & Rai (light green) bound to the Pol II surface (ray) to form a pre-termination complex (PTTC) (PDB 8JCH). Only density for the KOW5 domain (dark blue) is observed. B) Overlay of transcription elongating complex (TEC) (PDB 7XN7) onto PTTC structure showing severe steric clash between KOW domains 1-3 (dark blue). Spt4 is shown in pink. The KOW5 domain from both structures overlay identically without overlap with Xrn2 binding interface.*

Spt5 enhances Xrn2 activity via the KOW5 domain

Finally, and perhaps most interestingly, RNA degradation assays indicated that not only is there an association between Spt5 and Xrn2, but that the presence of Spt5 enhances Xrn2 exoribonucleolytic activity (Kuś et al., 2025). This effect was narrowed down to

the KOW5-CTR domain specifically, with RNA substrate half-life decreasing 3-fold when KOW5-CTR was added to Xrn2 (Fig. 1.8). In conjunction with the above results in which Xrn2 binding to the PTC results in partial removal of Spt5, it implicates KOW5 as having a significant role in facilitating both Xrn2 binding and activity.

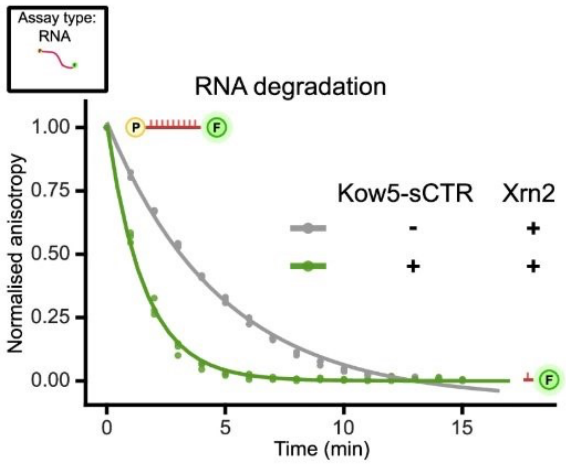


Figure 1.8. Fluorescence anisotropy assay *comparing RNA degradation rate of WT Xrn2 vs. WT Xrn2 in the presence of the KOW5-sCTR domain showing stimulation of Xrn2 activity.*

Aims of this Study

Evidently, Spt5 is a central accessory transcription factor involved throughout the entire transcriptional cycle. Despite the progress made in understanding Spt5's structure and function as well as its interactions with Pol II, its role during termination remains unclear. Specifically, mechanistic insights into the interactions of the KOWx45 domains during termination are lacking. These domains were proposed to be central to both RNA and Xrn2 interactions and so are focused on here.

I aim to address this through three main objectives in this work:

a) Do the KOWx45 domains of Spt5 interact with RNA and does this contribute to activation of Xrn2?

b) How does CTR phosphorylation impact this interaction?

c) How do the KOWx45 domains of Spt5 interact with Xrn2 and how does CTR phosphorylation impact this interaction?

Use of NMR as an experimental technique

Due to the flexible and disordered nature of Spt5, nuclear magnetic resonance (NMR) was deemed the appropriate method to gain mechanistic insight into its function. Previous crystal and cryoEM structures of Spt5 lack density for the linker regions and the CTR and are missing significant regions of some structured domains. This has complicated the understanding of Spt5 function. In contrast, NMR is suited for visualization of these disordered regions, allowing residue specific insights into the dynamics and binding interfaces of Spt5.

Focus on KOWx45 domains

Due to the proposed roles of the KOW4 and KOW5 domains, and difficulties in both producing sufficient quantities and analysing by NMR a full-length Spt5 construct, the construct used was comprised of the KOW4 and KOW5 domains along with the interdomain linker and a portion of the CTR (sCTR) region with 8 conserved C-terminal repeats (residues 574 -874) (Fig. 1.9).

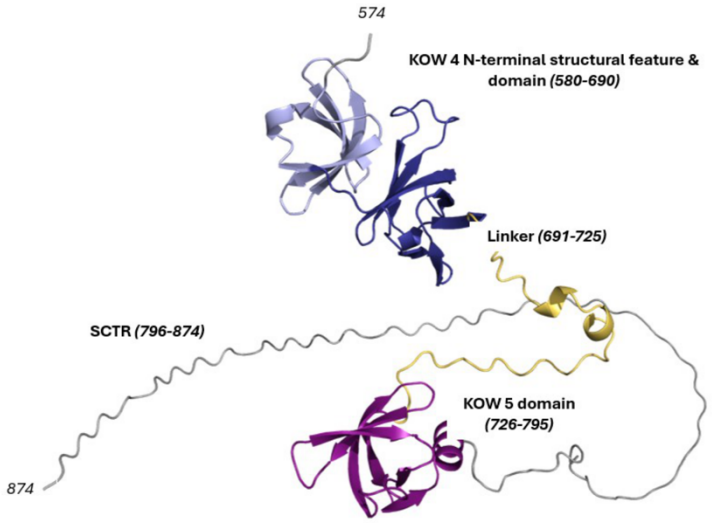


Figure 1.9. AF3 prediction of KOWx45 construct focused on here. *KOW4 domain is highlighted in blue, with the 4x N-terminal structural feature important for stability highlighted in light blue. KOW5 is highlighted in purple with the linker between them shown in yellow.*

Chapter 2: Purification and Assignment of KOWx45 construct via NMR

Purification of KOWx45 Construct

A key objective of this study was to determine how the KOWx45 domains of Spt5 contribute to transcriptional termination, particularly focusing on their interactions with RNA and Xrn2. Static structural studies suggested that Spt5 regulates transcription partially through the formation of an 'RNA clamp' by the KOWx45 domains near the RNA exit channel. To investigate these interactions, it was first essential to optimize a protocol to purify isotopically labelled KOWx45 protein for full backbone resonance assignment. This chapter outlines the purification and structural characterization of the KOWx5 construct using NMR to facilitate identification of residues perturbed upon addition of binding partners in downstream analysis.

Initial purification of KOWx45 was performed via immobilized metal affinity chromatography (IMAC) using an adapted protocol from the Vasilieva lab. Protein yield from 2 litres of starting culture produced 1ml of protein at 350uM (11.9mg/ml), and purity of the final product was checked via SDS-page gel electrophoresis and mass spectrometry analysis, with KOWx45 (32kDa) identified as the higher molecular weight band in the gel (Fig. 2.1). However, there exists a smear directly below this band of interest as well as a low molecular weight species (~17kDa) on the gel. This likely arises due to cleavage of the long flexible and disordered regions within the construct. The smearing is likely due to cleavage of the short CTR regions at varied points within the repetitive sequence with the clear lower molecular weight band caused by cleavage of the linker between the KOW4 and KOW5 domains. As IMAC was

performed using an N-terminal His-tag preceding the KOW4 domain, separation of these products was only possible following size exclusion chromatography (SEC).

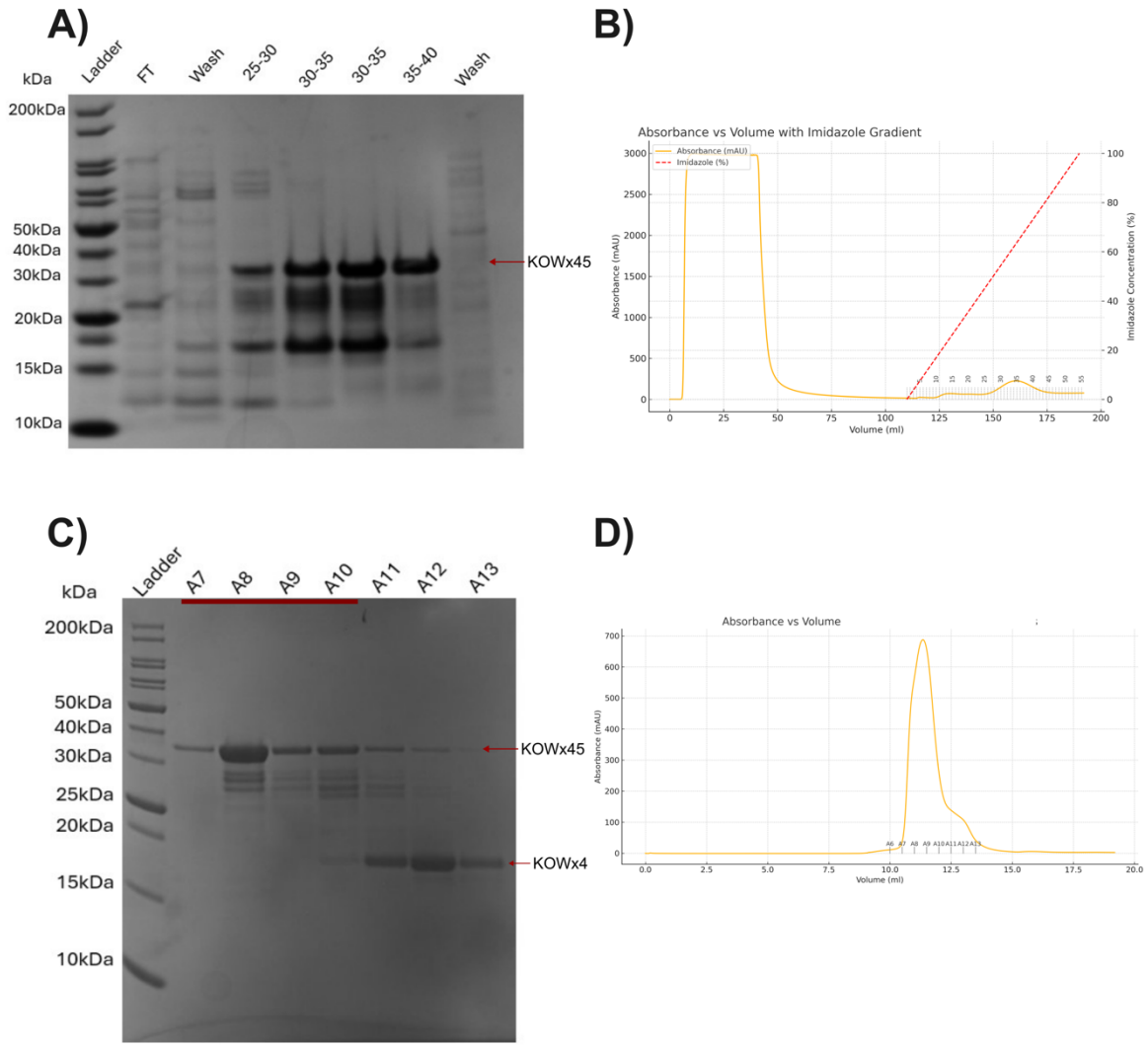


Figure 2.1. Purification of KOWx45 protein construct. *A) SDS-page gel of HisTrap purified KOWx45 showing presence of both KOWx45 construct (32kDa) and co-purified lower molecular weight species KOWx4 B) HisTrap purified KOWx45 Akta UV trace with lanes run on gel and concentrated for SEC highlighted in red. C) SDS-Page gel of SEC purified KOWx45 showing efficient separation of full-length KOWx45 (32kDa) D) SEC purified KOWx45 Akta UV trace with lanes concentrated for downstream analysis highlighted in red.*

In order to minimize cleavage, the purification protocol was adapted to include 10mM of phenylmethanesulfonylfluoride (PMSF) to resuspended cell pellets before purification, for further irreversible inhibition of cellular serine proteases. Additionally, cell lysates and protein elution fractions were kept on ice at all times, and the protocol was conducted within a single day to minimize elution fractions sitting overnight. While these mitigations improved the yield of full-length KOWx45, evidence of some proteolytic digestion remained, likely occurring *in vivo* before cell collection and lysis. However, the yield and purity of the full-length construct was deemed high enough for this not to be a significant problem moving forward.

Mass Spectrometry of Purified KOWx45 Construct

In order to confirm the identity of the proposed band of interest, in-solution and in-gel digests of the sample were conducted for downstream mass spectrometry analysis. A sample of purified and concentrated KOWx45 solution was digested with LysC, while both the band of interest and the predicted lower cleavage product band were isolated from the gel and digested with LysC and chymotrypsin.

Peptides resulting from digestion were identified using peptide mass fingerprinting, where results were searched against both the KOWx45 protein sequence as well as the *E. coli* TrEMBL database containing translated coding sequences from the BL21 proteome. The in-solution digest gave 62% coverage of full-length KOWx45 with 29 unique peptides at high confidence. No major contaminants were identified. In contrast, the in-gel digests gave hits corresponding to KOWx45 but with a low peptide count. This is likely due to the 20-fold dilution of protein sample before loading into a gel to prevent overloading and allow clear interpretation of results. The upper molecular

weight band gave 2 high confidence unique peptides with a 9% coverage of the full-length construct that included regions from both KOW4 and KOW5 domains. The lower molecular weight band also gave 2 high confidence peptides with a 5% coverage of the construct. These two peptides both originated only from the KOW4 domain. No other major high confidence contaminant peptides were identified. Overall, despite the low peptide abundance from the in-gel digests, mass spectrometry data confirmed the higher molecular weight band to be the KOWx45 protein of interest and that it was highly pure, thus making it suitable for downstream studies.

Confirmation of the ability to purify the KOWx45 construct at high yield and with high purity ultimately allowed for growth and purification of isotopically labelled KOWx45 samples for characterization via NMR.

Characterization & Assignment of KOWx45 construct via NMR

Following confirmation of KOWx45 purification, the protein was isotopically labelled with ^{15}N and ^{13}C, allowing for a suite of triple resonance NMR experiments and sequential backbone assignment of the protein. 2D B-TROSY experiments gave an initial idea of protein characteristics. Peaks were narrow with high signal to noise, and a substantial proportion of identified peaks showed poor dispersion yet very sharp peaks. This is due to rapid tumbling originating from the highly disordered and flexible linker and sCTR regions. The 3D NMR experiments recorded were: HNCO, HN(CA)CO, HNCA, HN(CO)CA, and HNCACB. The backbone amides of the 311 residues present in the construct were assigned to 97% completion (not including prolines and 6xHis-tag sequence) (Fig. 2.2a). The remaining unassigned regions comprised only short

stretches between 1-4 residues long and are located in the KOWx4 element of the structured KOW4 domain region (Fig. 2.2b), partially clustered around the interface between this connection domain and the main KOW4 beta barrel.

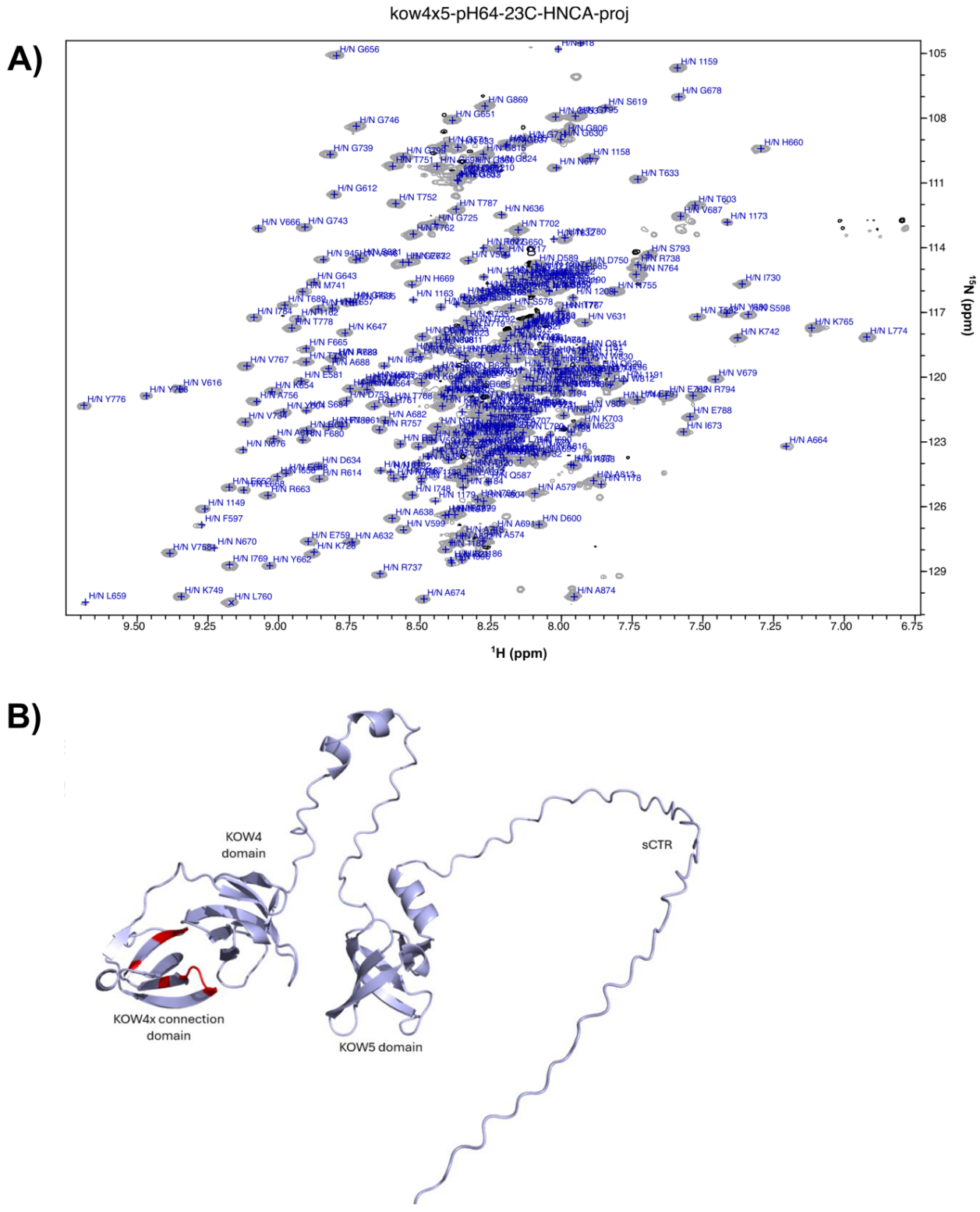


Figure 2.2. Characterization and assignment of KOWx45 by NMR. *A) HNCA spectrum of ^{15}N, ^{13}C isotopically labelled KOWx45 with residue backbone resonance assignments shown. B) AlphaFold3 prediction of purified KOWx45 construct with domains labelled and unassigned residues centred in KOWx4 connection domain and highlighted in red.*

Talos-n was used to calculate secondary structure propensity of the assigned systems, which also provided a check on whether any chemical shifts were outside the expected ranges for each amino acid type (Fig. 2.3) (Shen, Bax, 2013). A further measurement of the confidence in the chemical shift assignments came from comparison of the secondary structure prediction from chemical shifts with the secondary structures observed in an AlphaFold3 (AF3) model prediction and past experimental structures derived by crystallography and cryo-EM. Significant secondary structure elements are found only in the KOW4 and KOW5 domains and are predominantly composed of β-sheets. The conserved architecture of KOW domains consists of a five stranded beta-barrel. The linker and sCTR regions both lack any confidently predicted structural elements. Despite some potential regions with helical propensity, particularly within the linker, the confidence is low. As these regions are highly disordered and flexible, it is unlikely they contain very structured elements, although they may adopt some transient secondary structure dependent on post-translational modifications, presence of specific binding partners, or transcriptional stage.

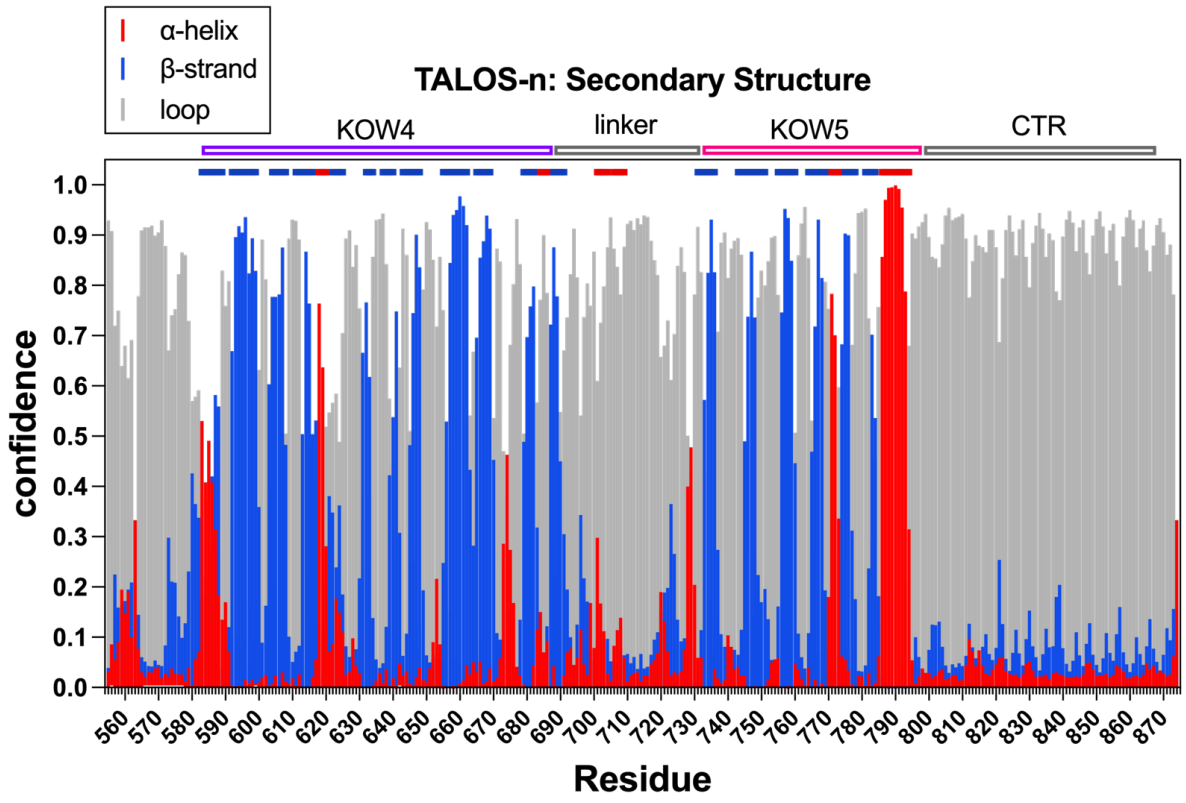


Figure 2.3. Talos-n predicted secondary structure for KOWx45 construct *(predicted α-helix shown in red and predicted β-strand shown in blue). Residue number indicated on x axis and confidence in predicted structure indicated on y axis. Secondary structure elements predicted by Alphafold3 shown below domain markers (predicted α-helix shown in red and predicted β-strand shown in blue), showing good correlation between AF3 vs. Talos-n predicted structure elements.*

These results from purification and characterization of the KOWx45 construct using solution state NMR established a foundational framework for subsequent investigations, fulfilling the first objective of the thesis. Validation of the construct for visualization via NMR enabled comparative analysis of spectra between the apo and predicted ligand-bound KOWx45 models to identify binding interfaces and gave confidence in the ability to identify residues affected by ligand binding.

Chapter 3: Characterizing KOWx45 Interactions with RNA

Identifying the RNA binding interface of KOWx45 and RNA

Previous cryo-EM structures proposed that the KOWx45 domains form an 'RNA clamp' around exiting nascent RNA, with mutation of this region abolishing RNA binding and causing general transcriptional defects (Bernecky et al., 2016; Qiu & Gilmour, 2017). Given these observations, I sought to map the binding interface of RNA on these domains.

2D $^{1}H,^{15}N$ HSQC experiments were conducted on ^{15}N labelled KOWx45 protein. RNA scaffolds of varying length (5, 10, & 26mers) were added to identify any chemical shift perturbations (CSPs) in the protein (Fig. 3.1), given in parts per million (ppm). Only minor CSPs were observed for all three RNA lengths. The maximum CSP values for 5mer RNA at both 1:1 and 1:4 molar ratios were 0.03ppm and 0.05ppm, respectively. The 10mer RNA at 1:1 molar ratio of RNA to protein had a maximum CSP value of 0.06ppm and the 26mer RNA at a 1:0.3 molar ratio of RNA to protein had a maximum CSP value of 0.065ppm. The lack of large CSPs, even at RNA concentrations well above that of the protein in certain cases, indicating these interactions are weak.

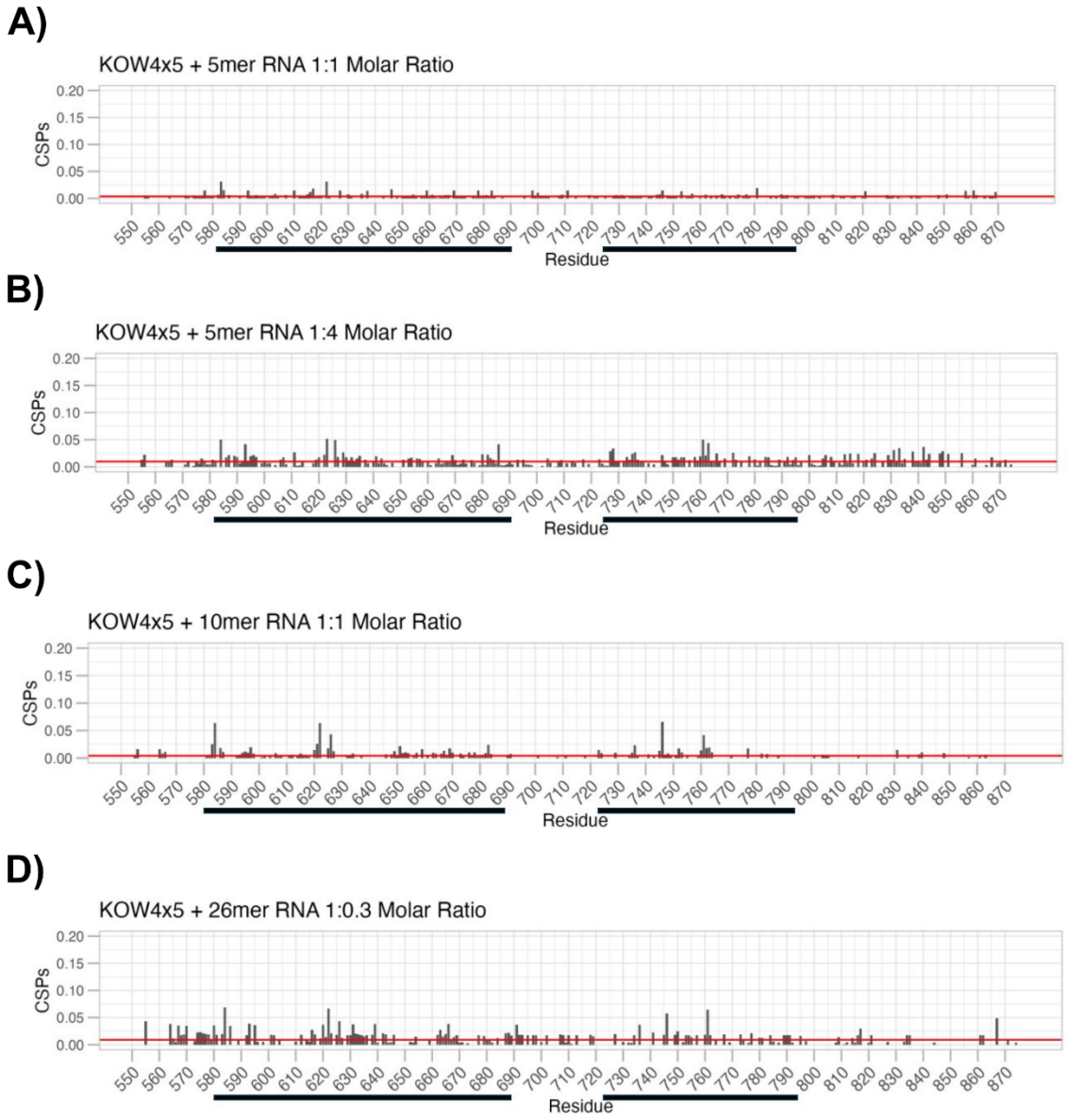


Figure 3.1. Chemical shift perturbation (ppm) plots of RNA addition into unphosphorylated KOWx45 with mean CSP indicated in red. *Horizontal bars indicate domain boundaries from KOWx4 and KOW5. A) CSP plot of 5mer RNA scaffold added to KOWx45 at a molar ratio of 1:1. B) CSP plot of 10mer RNA scaffold added to KOWx45 at a molar ratio of 4:1. C) CSP plot of 10mer RNA scaffold added to KOWx45 at a molar ratio of 1:1. D) CSP plot of 26mer RNA scaffold added to KOWx45 at a molar ratio of 0.3:1.*

These minor CSPs were, however, accompanied by global signal loss as RNA concentration increased, with more pronounced effects as the length of the RNA increased. As seen in the signal intensity plots below, the signal intensity drops to below 50% at a 1:0.2 molar ratio of protein to RNA when using the 26mer (Fig. 3.2b).

In contrast, at the same RNA ratio with the 10mer, the signal intensity is still around 80% (Fig. 3.2a). This effect compounded analysis difficulties, with the 26mer data only able to be confidently analysed up to a 1:0.3 molar ratio of KOWx45:RNA, despite being added up to a 1:4 molar ratio. Loss of signal intensity could be attributed to formation of a larger complex upon RNA binding, potentially due to binding of multiple protein molecules to one RNA scaffold. A larger complex would be unable to be detected due to slow tumbling, leading to a significant drop in signal intensity. Alternatively, signal loss could result from transient aggregation or non-specific interactions at high RNA concentrations reducing the detectable signal.

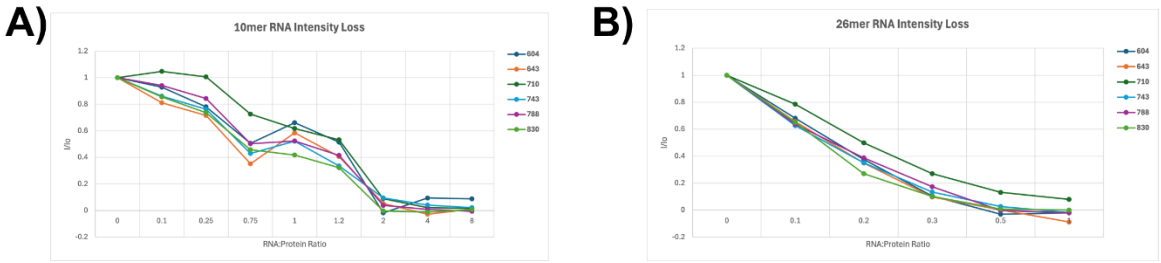


Figure 3.2. Global signal intensity loss with RNA addition. *A) Plot showing global loss of signal intensity with increasing concentration of 26mer RNA. Residues chosen to be plotted represent all regions of the Spt5 construct: 604 and 643 from the KOW4 domain, 710 from the linker, 743 and 788 from the KOW5 domain, and 830 from the CTR. B) Plot showing global loss of signal intensity with increasing concentration of 10mer RNA.*

Despite the small magnitude of the CSPs, the same residues were affected by the different RNA lengths, suggesting a correlation. The CSPs, though small, were evaluated given the fact that signal intensity losses provided an indication that interaction may be occurring. The majority of the residues with at least minor CSPs were found in the KOWx4 domain, with the strongest CSPs for residues in the interface between the KOWx element and KOW4 (residues 584 & 593) as well as within the connecting linker (Fig. 3.3). Two residues in the KOW5 domain were also consistently

affected (residues 746 and 761), found in the loop between β-strands three and four of the β-barrel KOW fold.

For further analysis, affected residues were mapped onto the cryoEM structure of DSIF bound to a transcribing human Pol II complex at a nominal resolution of 3.4Å (Fig. 3.3d). While some affected residues in the KOWx4 domain seem to overlap with regions used to interact with the Pol II surface, some form an exposed patch allowing the β-barrel to contact the nascent RNA exit channel.

The residues affected in the KOW5 domain clearly overlap with the binding interface to Pol II. This could either indicate these are artefactual, as they would not be exposed if the KOW5 domain is constitutively bound to Pol II, or that RNA binding is mutually exclusive with Pol II binding for this domain. Due to a lack of structures of full-length *S. pombe* DSIF bound to Pol II, residues affected by RNA addition were translated to their human Spt5 equivalents for visualization here, but this may not be an accurate comparison if there are unconserved structural features. It is important to note that that interactions of *S. pombe* DSIF may not bind to Pol II or nascent RNA in an identical manner to that of the human ortholog.

One structure of *S. pombe* DSIF bound to Pol II was recently published by Kuś et al., but only 72 residues of the KOW5 domain (between residues 726-797) were visible. Mapping those found to be affected onto this structure agreed with the previous findings, with the affected KOW5 residues 746, 761, & 763 overlapping with contacts made to the Pol II surface (Fig. 3.3e).

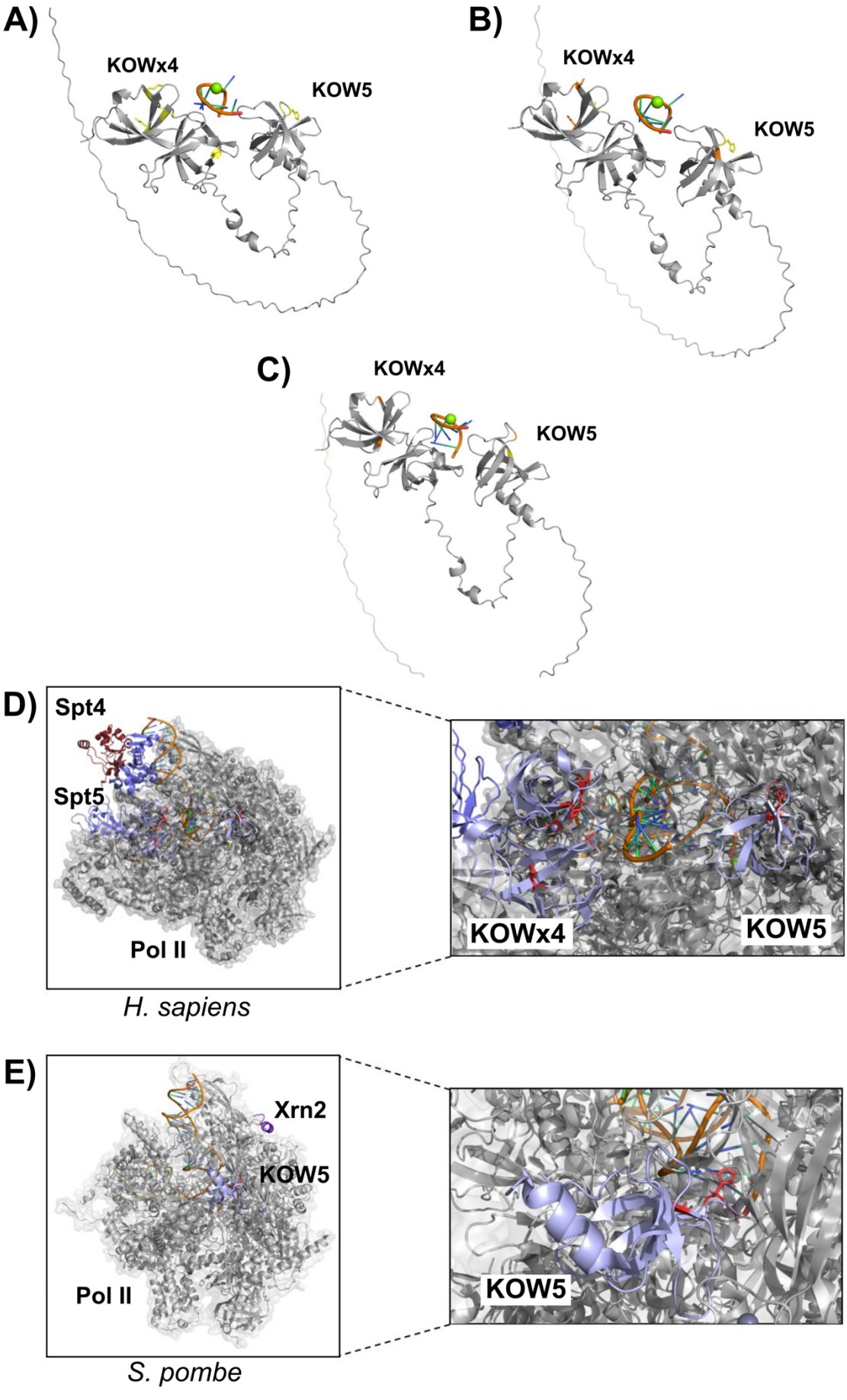

A)
KOWx4
KOW5
B)
KOWx4
KOW5
C)
KOWx4
KOW5
D)
Spt4
Spt5
Pol II
KOWx4
KOW5
H. sapiens
E)
Xrn2
KOW5
Pol II
KOW5
S. pombe

Figure 3.3. Identifying RNA binding interface of KOWx45. *CSPs are coloured based on magnitude (yellow = 0.04-0.06ppm, orange = 0.06-0.08ppm, red > 0.08ppm). A) CSPs from 5mer RNA mapped onto AF3 prediction of KOWx45 construct. B) CSPs from 10mer RNA mapped onto AF3 prediction of KOWx45 construct. C) CSPs from 26mer RNA mapped onto AF3 prediction of KOWx45 construct. D) CSPs (red) from all RNA titrations mapped onto human Pol II elongation complex (PDB 5OIK). Pol II shown in grey, Spt5 domains shown in light blue, Spt4 shown in magenta. E) CSPs (red) from all RNA titrations mapped onto S. pombe Pol II elongation complex (PDB 8QSZ). Pol II shown in grey, Spt5 domains shown in light blue, Xrn2 helix shown in purple.*

A 7mer RNA scaffold length was also titrated, but the observed results differed from the previous observations made with other RNA lengths (Fig. 3.4). Affected residues were found predominantly in the CTR region, and at high molar ratios of 1:3 protein to RNA, within the linker region. However, these CSPs were potentially non-specific effects in these disordered regions being caused by changes in free ion concentration in the NMR buffer as RNA was titrated in. Therefore, the experiment was repeated, but the 7mer RNA was instead solubilized in NMR buffer containing 2mM $MgCl_2$ to remove weaker or non-specific interactions and account for changes in free Mg^{2+} concentration upon RNA addition to the protein solution. This abolished the previous CSPs observed, which could indicate those CSPs were non-specific. The maximum CSP value seen also decreased from 0.072ppm to 0.043ppm. Remaining minor CSPs are now seen in the structured regions, such as between residues 620-630 in the KOWx4 domain (Fig. 3.4b). These are regions previously seen to be affected by RNA titration using other scaffold lengths. However, even at 1:4 molar ratios of protein to RNA, these perturbations were relatively small, even when compared to previous titrations with the 5mer RNA scaffold.

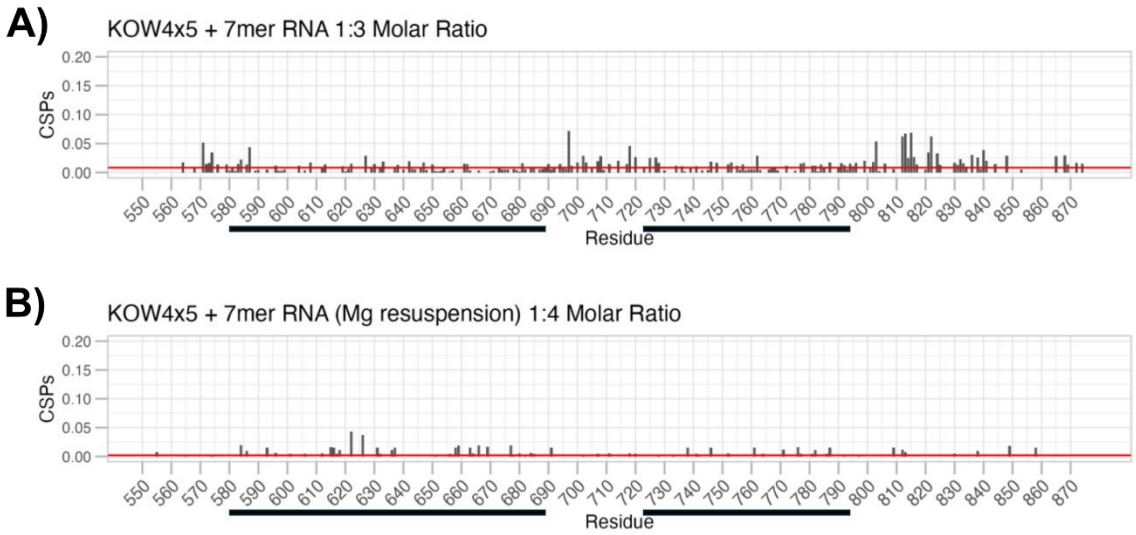


Figure 3.4. Chemical shift perturbation (ppm) plots of 7mer RNA addition into unphosphorylated KOWx45 *with mean CSP indicated in red. Horizontal bars indicate domain boundaries from KOWx4 and KOW5. A) CSP plot of 7mer RNA scaffold added to KOWx45 at a molar ratio of 3:1. B) CSP plot of 7mer RNA scaffold resuspended into NMR buffer added to KOWx45 at a molar ratio of 4:1.*

Overall, these observations indicated that KOWx45 lacked strong interactions with the RNAs that were tested. These results led me to the question of whether the interaction may be affected on post-translational modifications, such as phosphorylation, that are extensive during transcription.

In vitro phosphorylation of KOWx45 Construct

Initial NMR titrations revealed only small CSPs between KOWx45 and RNA, a finding that was initially unexpected given prior literature suggesting these domains play a significant role in interacting with nascent RNA. As the literature identifies phosphorylation of the Spt5 C-terminal repeat region to be an important modulator of its function throughout transcription, particularly during transcriptional termination, I decided to investigate whether phosphorylation of the sCTR region of my KOWx45 construct would impact RNA binding ability. The phosphorylation state of the CTR is

functionally significant, with dephosphorylation at the 3' end of genes correlating with Pol II deceleration and termination. Consequently, it is plausible that the phosphorylation state of this region influences the RNA-binding capacity of KOWx45.

In order to generate a phosphorylated KOWx45 construct for these studies, four batches of CDK9 stock from the Vasilieva lab were tested in small scale *in vitro* phosphorylation reactions. Phosphorylation was confirmed via SDS-page gel analysis using self-casted Phos-tag™ gels. Phos-tag™ is a small molecule that binds to phosphorylated residues, slowing their mobility in the gel compared to their unphosphorylated counterparts. Batches 1-3 clearly show a mobility shift in the gel, thus making them appropriate to use in larger scale NMR phosphorylation experiments (Fig. 3.5a). Notably, there is a small proportion of KOWx45 in the phosphorylated samples that remains unphosphorylated just below the full-length KOWx45 sample. This is likely due to CTR region cleavage and removal of Threonine (Thr) residues, which was also noted in original purifications of the construct. Phosphorylation of KOWx45 by batch 4 was less efficient, potentially due to protein degradation and was not used further.

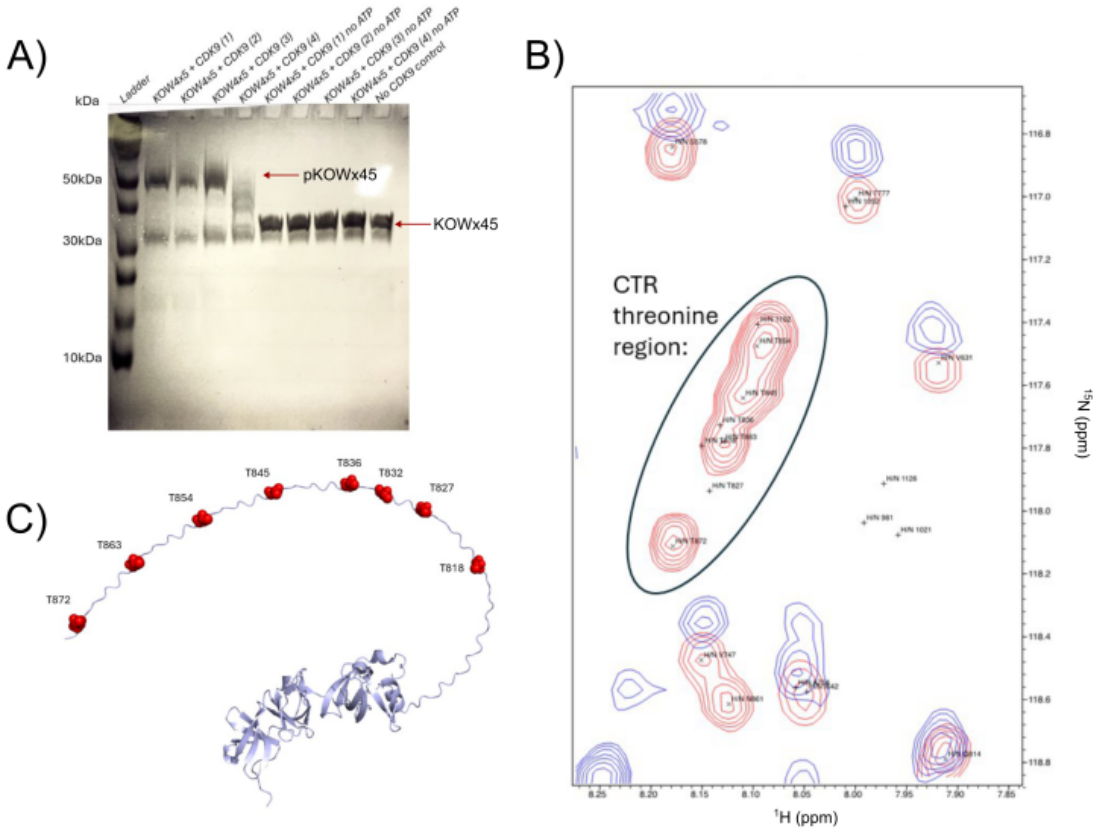


Figure 3.5. In vitro phosphorylation of KOWx45 sCTR region. *A) Phos-tag gel with lanes 1-4 showing KOWx45 slowed gel mobility indicative of phosphorylation and lanes 4-9 as positive controls B) ^{1}H,^{15}N HQSC spectrum of ^{15}N isotopically labelled KOWx45 (unphosphorylated in blue with phosphorylated overlay in red) showing disappearance of threonine residues in sCTR region, which was attributed to phosphorylation. C) AlphaFold 3 prediction of KOWx45 with threonine residues present in the sCTR and targeted for phosphorylation by CDK9 highlighted as red spheres.*

However, while Phos-tag™ gel analysis confirmed the presence of phosphorylated residues, it lacked confirmation of site-specific phosphorylation of the threonine residues present within the sCTR region of KOWx45. NMR served as an alternative technique to do so. A larger scale phosphorylation reaction was assembled, at a concentration of ^{15}N isotopically labelled KOWx45 previously identified as sufficient for NMR studies. Initial ^{1}H,^{15}N HSQC spectra collection of the phosphorylated construct failed, with the reason identified as the presence of Mn^{2+} ions. Manganese ions are the

superior divalent cation cofactor for supporting kinase activity as compared to magnesium, zinc, or copper ions (Pei & Shuman, 2003). However, manganese is also paramagnetic, leading to short relaxation times and broadened linewidths that significantly affect spectra quality (Lehr et al., 2020). Mn^{2+} was removed via dialysis and spectra of the dialyzed sample showed clear spectral overlap with the unphosphorylated HSQC spectra and specific disappearance of all threonine residues present in the sCTR region of the spectra (Fig. 3.5b & c). This confirmed threonine specific phosphorylation by CDK9, which aligns with findings from the literature, and allowed expanded NMR titration studies examining effects of sCTR phosphorylation on RNA binding ability (Pei & Shuman, 2003).

CTR Phosphorylation enhances KOWx45 RNA binding

After confirming *in vitro* site-specific phosphorylation of the KOWx45 CTR, effects on nucleic acid binding were tested by addition of either 7mer or 26mer RNA scaffolds. Interestingly, significant chemical shift perturbations were observed compared to the lack of CSPs in the unphosphorylated protein (Fig. 3.6).

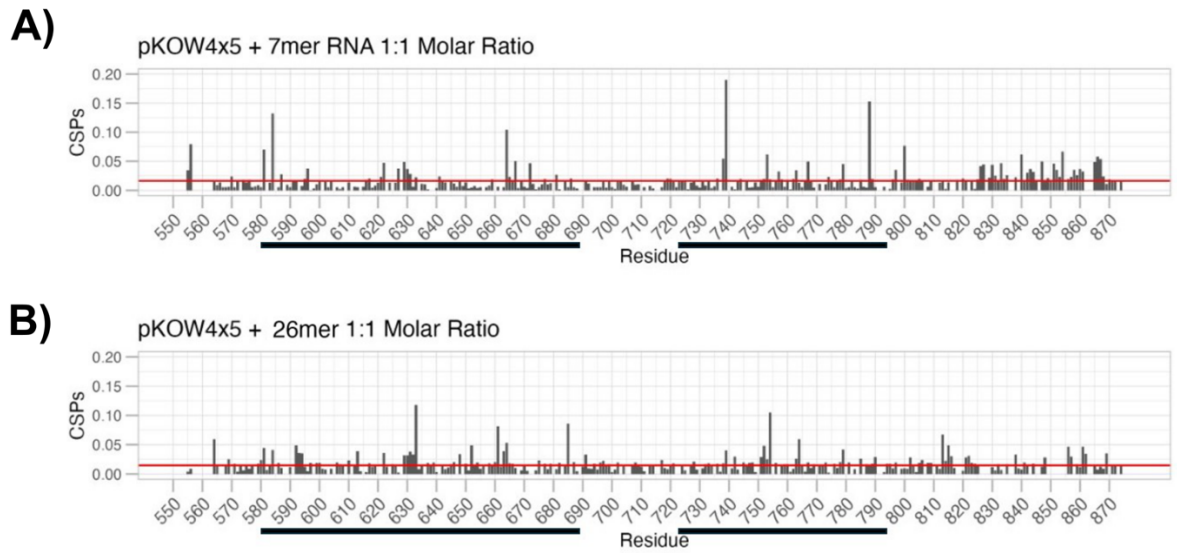


Figure 3.6. Chemical shift perturbation (ppm) plots of RNA addition into phosphorylated KOWx45 *with mean CSP indicated in red. Horizontal bars indicate domain boundaries from KOWx4 and KOW5. A) CSP plot of 7mer RNA scaffold added to pKOWx45 at a molar ratio of 1:1. B) CSP plot of 26mer RNA scaffold added to KOWx45 at a molar ratio of 1:1.*

When mapped onto the predicted structure of the protein, affected residues centred around the predicted RNA binding site (Fig. 3.7a & b). This included several positively charged residues (R629, R685, & R738) as well as some polar, uncharged amino acids such as threonine and asparagine. Several residues at the interface between the KOWx structured element and the KOW4 domain were also perturbed, similarly to RNA titrations done on the unphosphorylated KOWx45 protein. RNA binding potentially leads to conformational rearrangement of these regions relative to each other, which could expose surface patches or residues available for contacting exiting nucleic acids.

Affected residues were again mapped onto the cryoEM structure of DSIF bound to a transcribing human Pol II complex and the *S. pombe* KOW5 domain bound to Pol II (Fig. 3.7c & d). Similar to RNA titrations done with unphosphorylated KOWx45, affected residues from the KOW5 domain overlapped with the Pol II binding surface.

However, several affected residues in the KOWx4 domain form an exposed patch that could allow the KOW β-barrel to contact the exit channel and guide exiting nascent RNA. These results suggest that the KOWx4 domain is most important for binding nascent RNA, which is supported by data from studies of human Spt5 domains.

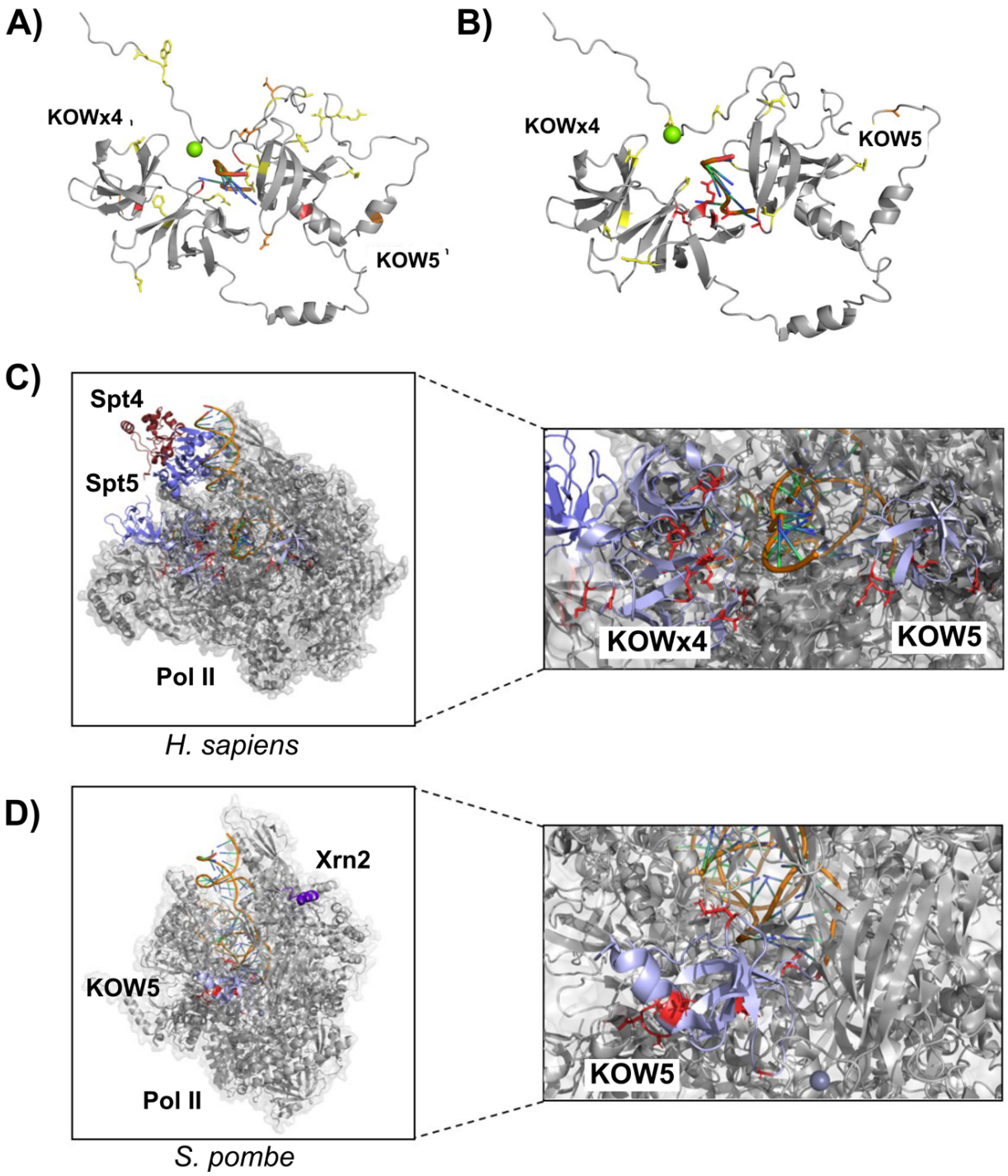


Figure 3.7. Identifying RNA binding interface of pKOWx45. *CSPs are coloured based on magnitude (yellow = 0.04-0.06ppm, orange = 0.06-0.08ppm, red > 0.08ppm). A) CSPs from 7mer RNA mapped onto AF3 prediction of pKOWx45 construct. Note that although the AF3 model indicates potential contact between the KOWx45 domains, this is not supported by the experimental NMR data. B) CSPs from 26mer RNA mapped onto AF3 prediction of pKOWx45 construct. Note that although the AF3 model indicates potential contact between the KOWx45 domains, this is not supported by the experimental NMR data. C) CSPs (red) from all RNA titrations mapped onto human Pol II elongation complex (PDB 5OIK). Pol II shown in grey, Spt5 domains shown in light blue, Spt4 shown in magenta. D) CSPs (red) from all RNA titrations mapped onto S. pombe Pol II elongation complex (PDB 8QSZ). Pol II shown in grey, Spt5 domains shown in light blue, Xrn2 helix shown in purple.*

When compared to the electrostatic surface potential of the construct, several affected residues correlate with positively charged patches on the protein surface that would enhance nucleic acid binding (Fig. 3.8).

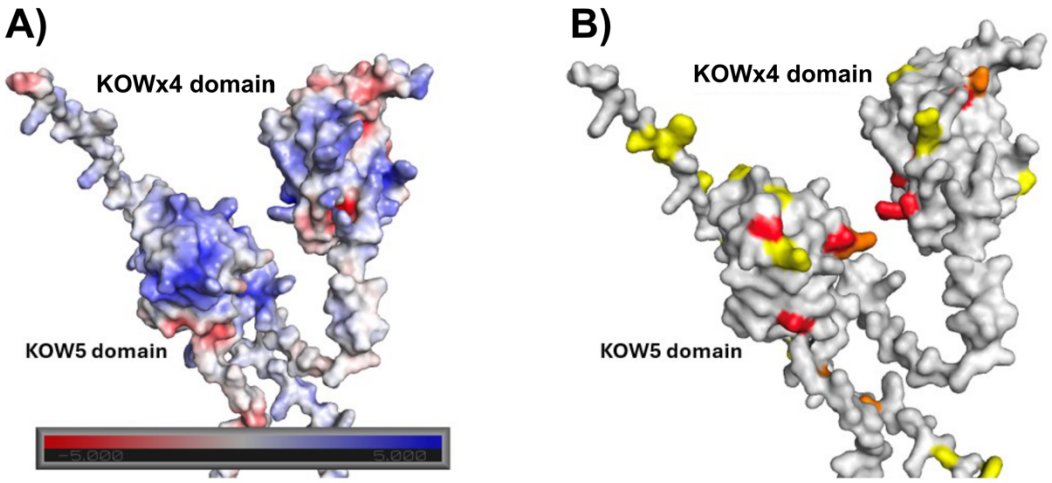


Figure 3.8. Comparison of CSP residues with electrostatic surface potential of KOWx45 construct. *A) Electrostatic surface potential of KOWx45 calculated with APBS program (coloured from -5 kT/e in red to +5 kT/e in blue). B) Residues showing chemical shift perturbations upon RNA addition to pKOWx45 are coloured in yellow (weak), orange (medium) and red (strong).*

My results indicated that phosphorylation of the CTR region of KOWx45 domains greatly impacted its ability to bind to RNA, yet how this occurred remained unknown.

Analysis of chemical shift perturbations indicated multiple residues were involved in binding to RNA, which aligns with the idea of weaker transient and multivalent nucleic acid interactions within the KOWx4 domain. As nascent RNA is continuously produced and exiting the RNA clamp during transcription, having strong binding based on only a few key residues would hinder this process.

I hypothesized that phosphorylation of the CTR region could induce transient structural changes in this region or allosterically within the KOWx45 domains, although no significant changes were seen in AlphaFold3 predictions (Fig. 3.9).

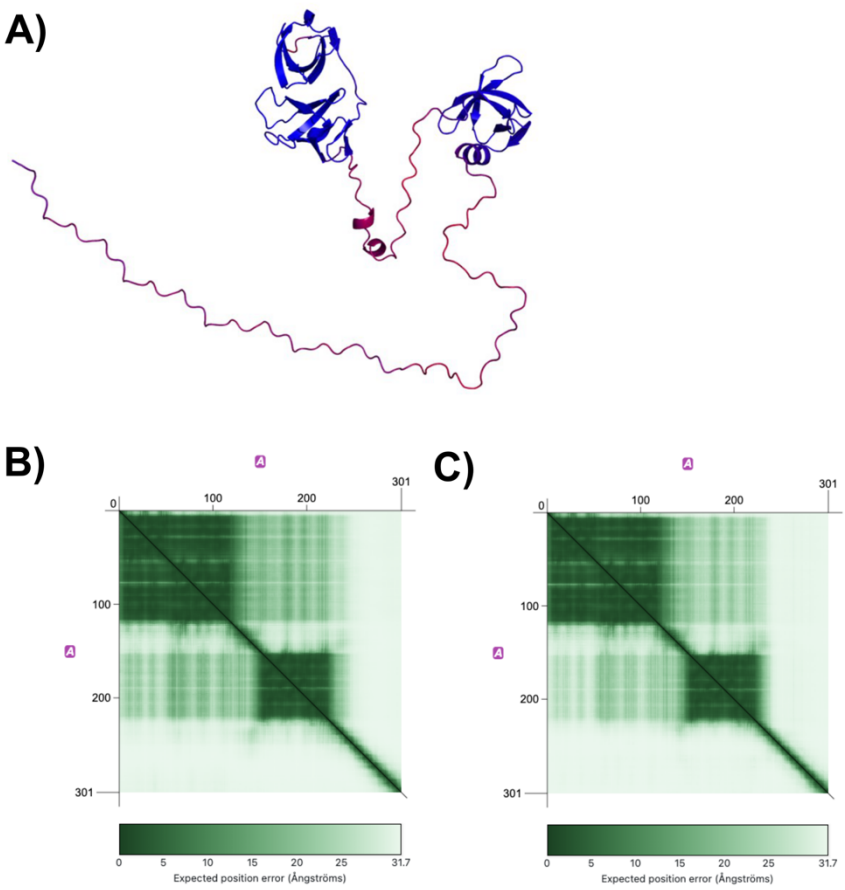


Figure 3.9. Structural prediction of phosphorylated KOWx45 construct. *A) AlphaFold3 prediction of KOWx45 construct with phosphorylated sCTR region (coloured blue to red based on pLDDT values). B) Predicted aligned error (PAE) plot of phosphorylated KOWx45 construct indicating increased confidence in potential residue interactions within the sCTR region. C) Predicted aligned error (PAE) plot of unphosphorylated KOWx45 construct.*

To gain experimental insight into whether phosphorylation led to structural changes in the construct that could explain the increased RNA binding, I conducted a minimal CSP analysis comparing peak shifts between the backbone resonance assignments of the unphosphorylated and phosphorylated KOWx45 spectra (Fig. 3.10). The results indicated larger shifts within the CTR than previously seen in RNA binding analyses, supporting the idea of phosphorylation-induced local rearrangements of this region. The strongest perturbations were observed at the interface between the KOWx structured element and the KOW4 domain, also suggesting the potential for conformational changes beyond the immediate phosphorylation site. Phosphorylation may lead to allosteric changes or prevent this region from folding back onto the structured domains, both of which could expose surfaces suitable for interaction with RNA or even Xrn2.

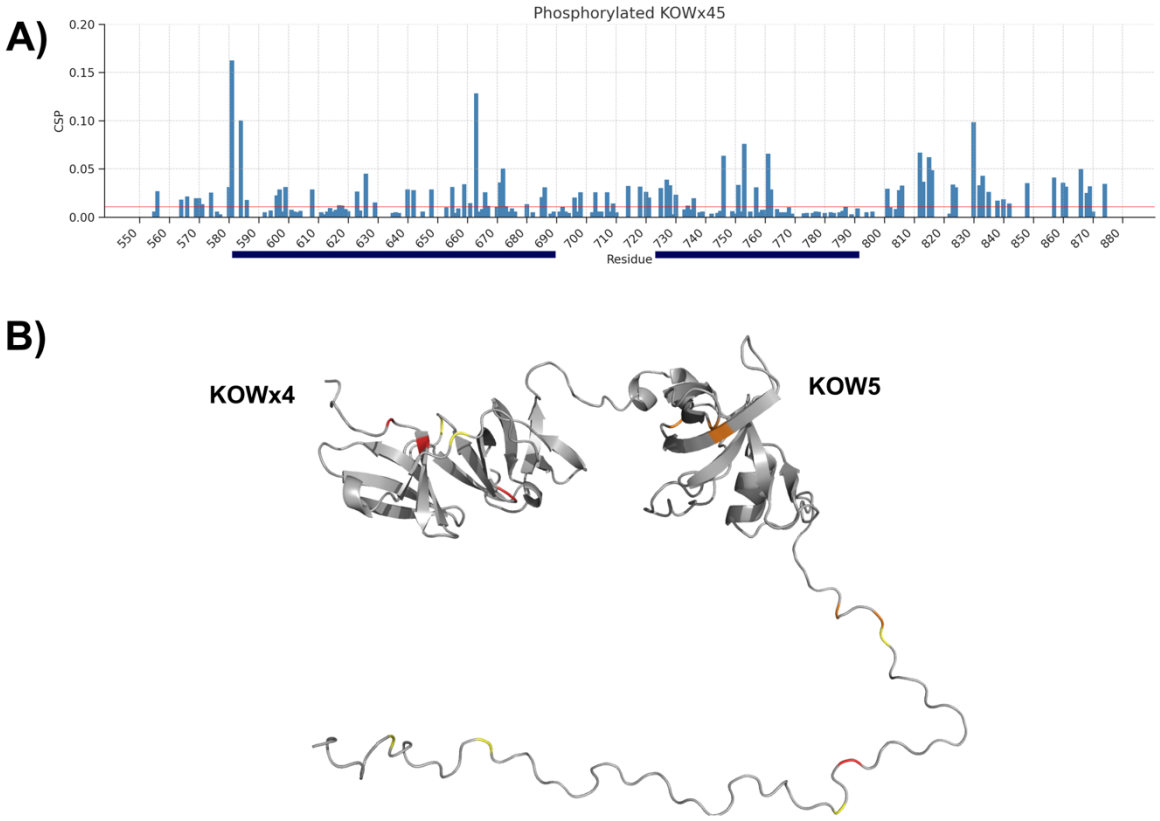


Figure 3.10. Identification of CSPs upon phosphorylation of KOWx45 construct. *A) Chemical shift perturbation (ppm) plots of RNA titration into phosphorylated KOWx45 with mean CSP indicated in red. Horizontal bars indicate domain boundaries from KOWx4 and KOW5. B) CSPs mapped onto AF3 prediction of pKOWx45 construct. CSPs are coloured based on magnitude (yellow = 0.04-0.06ppm, orange = 0.06-0.08ppm, red > 0.08ppm).*

Overall, these results show that phosphorylation of the sCTR region of KOWx45 enhances its RNA binding ability, particularly through multiple interactions with the KOWx4 domain. Observed CSPs align with positively charged surface patches and surface regions found in previous structural studies to be exposed to the exiting RNA tunnel of Pol II. Further biophysical studies are needed to confirm how this post-translational modification leads to such an effect, but these findings reinforce the idea that phosphorylation can serve as a regulatory switch modulating interactions between Pol II, regulatory factors such as Spt5, and RNA.

Chapter 4: Characterizing KOWx45 Interactions with Xrn2

Purification of Xrn2

Having established that phosphorylation alters RNA binding, we next investigated whether this modification impacts the interaction of Spt5 with the exoribonuclease Xrn2, another key component of the termination machinery. Xrn2 dislodges Pol II from the DNA template through progressive degradation of the cleaved nascent RNA. Crosslinking studies also identified interactions between Xrn2 and Spt5 (Kuś et al., 2025). Given this, I optimized purification of Xrn2 for use in further NMR studies to identify the binding interface between these two components.

Wild-type Xrn2 (residues 1-885) was purified from a pRSFDuet plasmid loaned by the Vasilieva lab. Initial attempts to purify the protein with an identical protocol as that used for KOWx45 purification resulted in protein precipitation and loss. Further test purifications resulted in an optimized final purification protocol that yielded pure Xrn2 (Fig. 4.1). Increasing salt concentration as well as avoiding over-concentration of Xrn2 reduced precipitation and improved protein stability greatly.

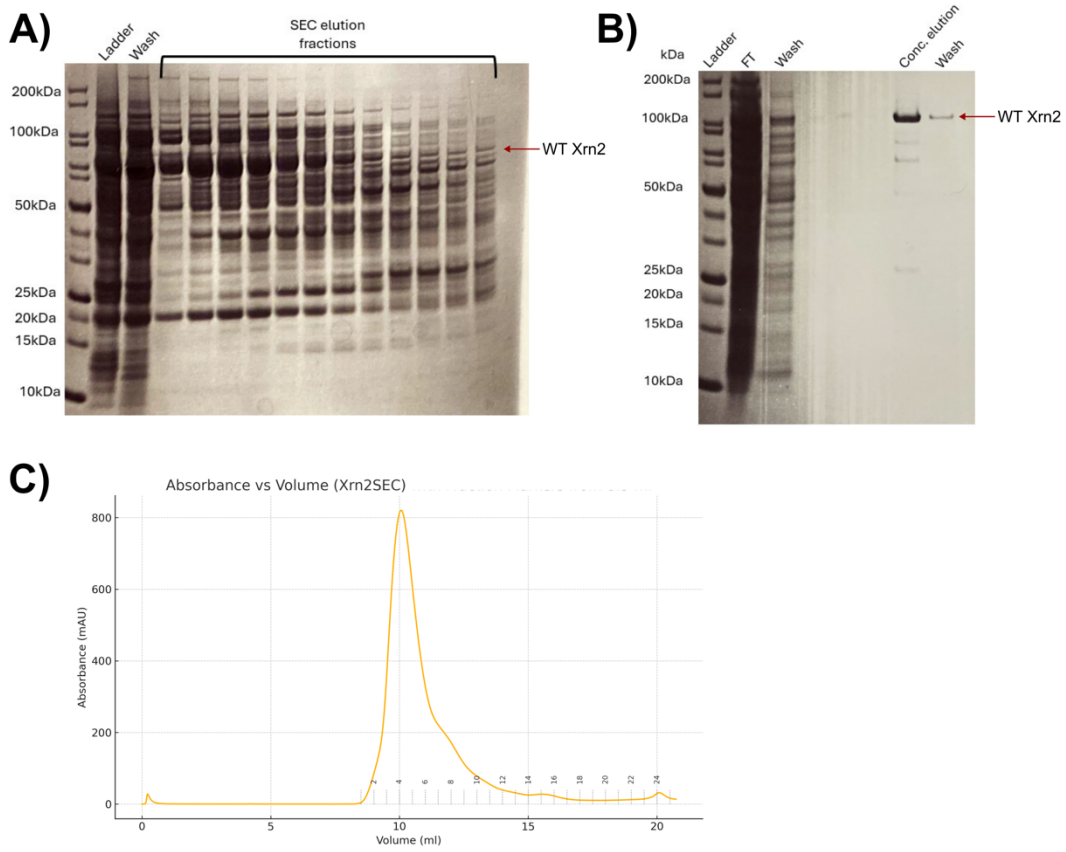


Figure 4.1. Optimization of purification of WT Xrn2. *A) SDS-Page gel of initial SEC purified WT Xrn2 (90kDa) showing low yield and purity of protein. B) SDS-Page gel of concentrated SEC purified WT Xrn2 (90kDa) using optimized protocol showing high sample purity C) Optimized SEC purified WT Xrn2 Akta UV trace. Lanes 1-8 were concentrated for downstream use.*

Validation of Xrn2 Activity

In order to confirm the RNA binding activity of WT Xrn2, an electrophoretic mobility shift assay (EMSA) was conducted using a non-cleavable RNA scaffold (ncRNA). Visualization by ethidium bromide (EtBr) clearly shows a shift in ncRNA mobility upon addition of Xrn2, even at low stoichiometric ratios of 1:0.5 RNA:Xrn2 (Fig. 4.2).

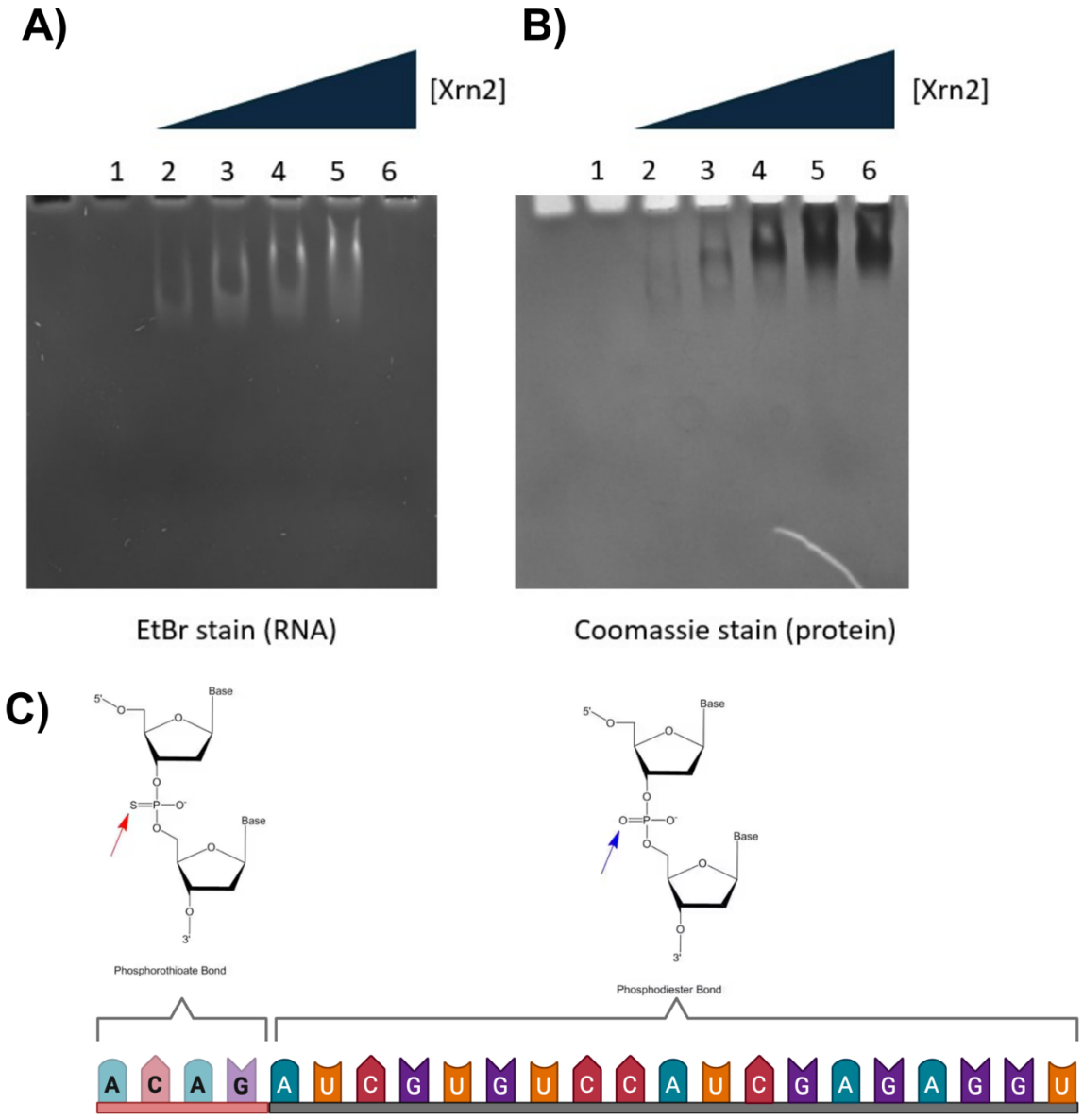


Figure 4.2. Mobility shift assay characterizing WT Xrn2 RNA binding ability. *Lane 1 contains RNA only and lane 6 contains Xrn2 only. A) EtBr-stained DNA retardation gel showing clear mobility shift of noncleavable RNA scaffold with increased concentration of Xrn2. B) Coomassie-stained DNA retardation gel showing presence and increased concentration of Xrn2. C) Schematic of noncleavable RNA scaffold used with phosphorothioate bond substitutions (created in Biorender and adapted from Sigma Aldrich).*

Furthermore, an RNA cleavage assay was conducted to assess WT Xrn2 degradation activity (Fig. 4.3). Visualization by EtBr showed disappearance of a cleavable RNA scaffold after less than one minute of incubation with WT Xrn2. In contrast, the

noncleavable RNA scaffold remained intact even after one hour of incubation, with slight smearing below the primary RNA band potentially indicative of a slight decrease in stability over longer periods of time.

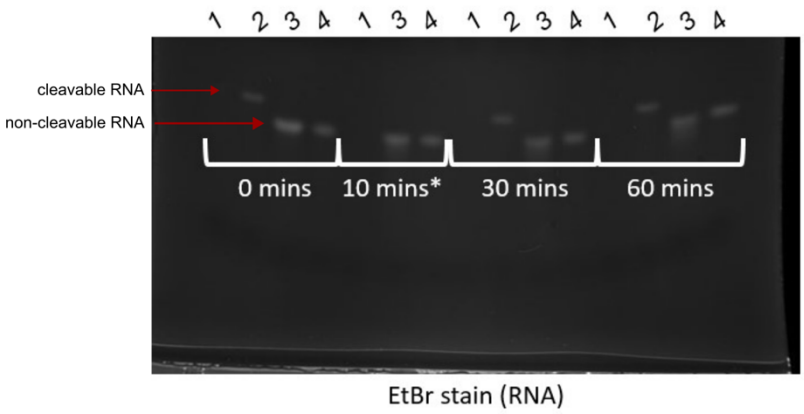


Figure 4.3. RNA degradation assay characterizing WT Xrn2 degradation of RNA and inability to degrade noncleavable RNA scaffold. *Reaction 1 contained both Xrn2 and RNA, reaction 2 contained only RNA, reaction 3 contained Xrn2 and noncleavable RNA scaffold, and reaction 4 contained only noncleavable RNA scaffold. Timepoints of each reaction was quenched at 4 timepoints and ran on gel above. Note that due to lane space limitations, reaction 2 after 10 minutes was omitted.*

Mapping of the Xrn2 binding site on Spt5 via NMR

Although the preliminary data from the Vasilieva lab indicated the presence of a direct interaction between Spt5 and Xrn2, the exact binding site was unknown. To map the binding interface between Spt5 and Xrn2, we titrated unlabelled Xrn2 into ^{15}N isotopically labelled KOWx45 samples and observed CSPs using NMR.

Similar to the RNA titration results, only minor CSPs were observed when Xrn2 was added to unphosphorylated KOWx45. In contrast, significant CSPs, primarily clustered

within the KOW5 domain, were observed when Xrn2 was added to phosphorylated KOWx45, indicating this is the primary interaction interface between the two proteins (Fig. 4.4). This correlated with previous crosslinking data between Xrn2 and full-length Spt5 from the Vasilieva lab. Crosslinks were primarily found between the KOW5 domain, along with a small number to the CTR, although this was in the context of Pol II.

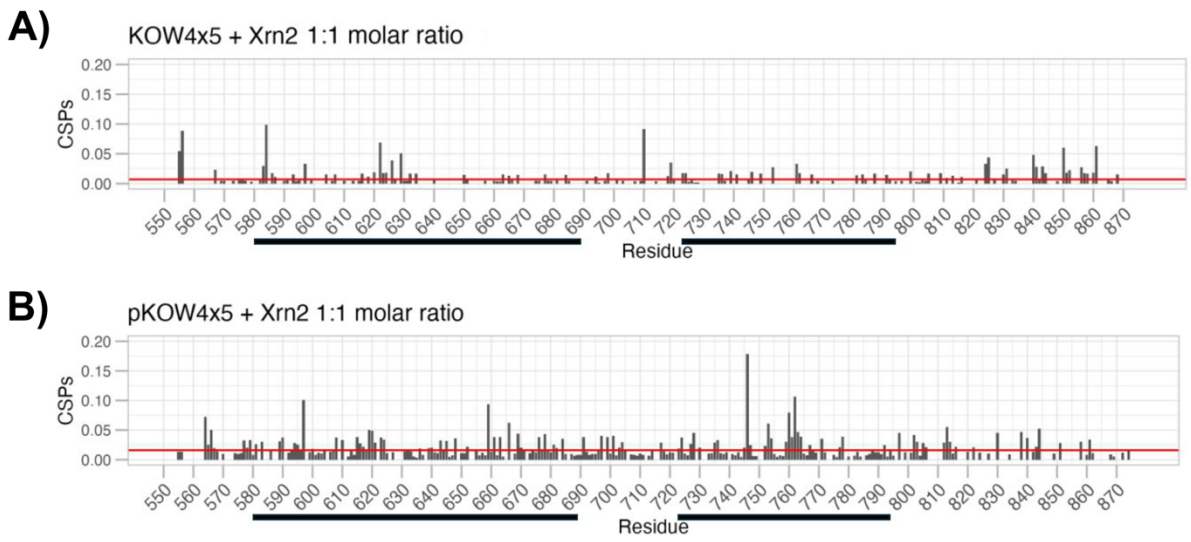


Figure 4.4. Chemical shift perturbation (ppm) plots of Xrn2 addition into unphosphorylated KOWx45 vs. Phosphorylated KOWx45 *with mean CSP indicated in red. Horizontal bars indicate domain boundaries from KOWx4 and KOW5. A) CSP plot of Xrn2 added to KOWx45 at a molar ratio of 1:1. B) CSP plot of Xrn2 added to pKOWx45 at a molar ratio of 1:1.*

When mapped onto the KOWx45 structure, affected residues are found within the second to fourth β-strands within the KOW β-barrel fold, with a particular cluster comprising the loop between the third and fourth strands (residues L760, H761, T762, G763). These CSP clusters align with the AlphaFold3 prediction of the interaction between the two proteins, where the loop and strand region from the KOW5 domain specifically makes contacts with the Xrn2 structured region. Some CSPs are also

found within the KOWx4 domain, and the AlphaFold3 structure predicts this domain to contribute to the overall KOWx45 binding. However, affected residues are predominantly clustered at the interface between the KOW4 domain and KOWx structured element. This, combined with the lack of crosslinking data supporting the idea of interactions between the KOWx4 domain and Xrn2 and lower confidence in the predicted interacting complex structure, led me to the conclusion that this is artefactual and does not form part of the interaction interface. It is more likely that the observed CSPs originate from conformational changes in the orientation of the KOW4 and KOWx elements relative to each other upon Xrn2 binding.

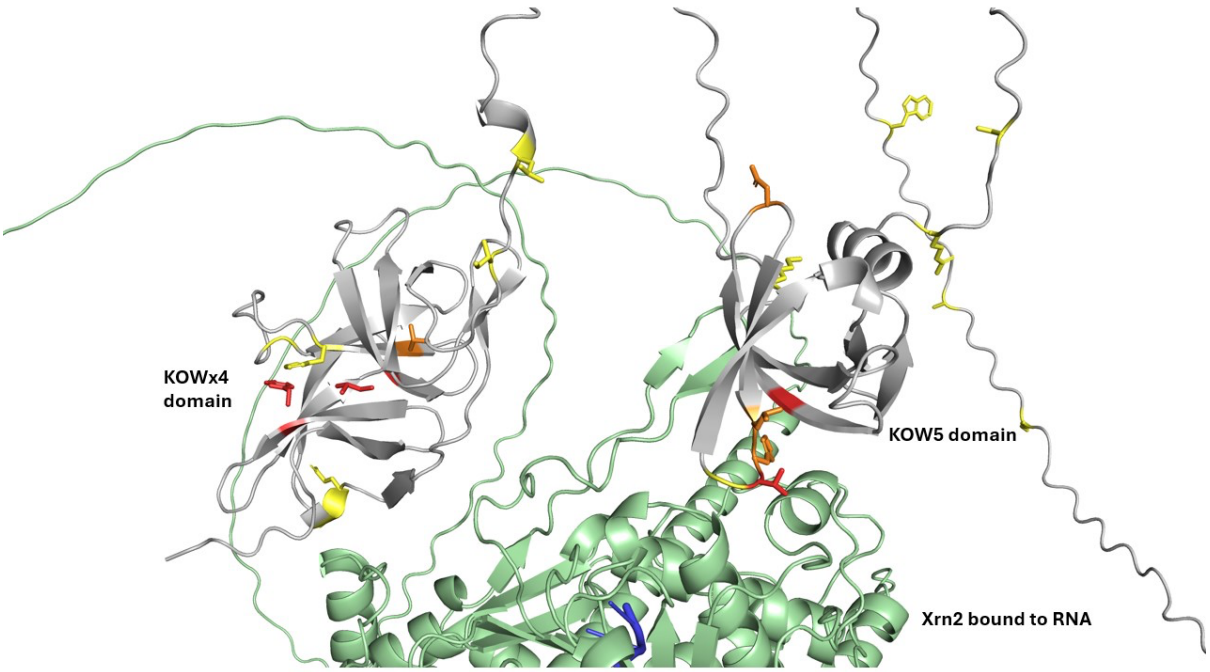


Figure 4.5. Mapping the Xrn2 binding interface of KOWx45. *CSPs are coloured based on magnitude (yellow = 0.04-0.06ppm, orange = 0.06-0.08ppm, red > 0.08ppm). CSPs from Xrn2 addition to pKOWx45 mapped onto AF3 prediction of KOWx45 construct bound to WT Xrn2 (with non-cleavable RNA scaffold added).*

Moving forward, we wanted to identify how RNA presence affects this interaction. A non-cleavable RNA scaffold (ncRNA) was additionally added either before or after

addition of Xrn2. When Xrn2 was added to a pre-formed ^{15}N KOWx45 & RNA complex, the perturbed residues are the same as when Xrn2 is added to KOWx45 alone, including the loop between the third and fourth β-strands of the KOW5 domain (Fig. 4.5a & b). This indicates the binding interface is unaffected by the presence of RNA. However, additional residues are affected that are not seen in the Xrn2 & KOWx45 complex alone, in particular residues in the KOWx4 domain. The overlay of these residues correlates with those affected when ncRNA scaffold is added to KOWx45 alone, potentially indicating that Xrn2 has higher affinity for RNA and is thus able to 'steal' it from KOWx45 (Fig. 4.5c). The proposed interaction would lead to perturbations of the same residues as those used to bind RNA initially.

Experiments adding ncRNA to the pre-formed ^{15}N KOWx45 & Xrn2 complex showed no strong CSPs (Fig. 4.5d). This likely indicates that the RNA scaffold is preferentially binding to Xrn2 due to its higher affinity.

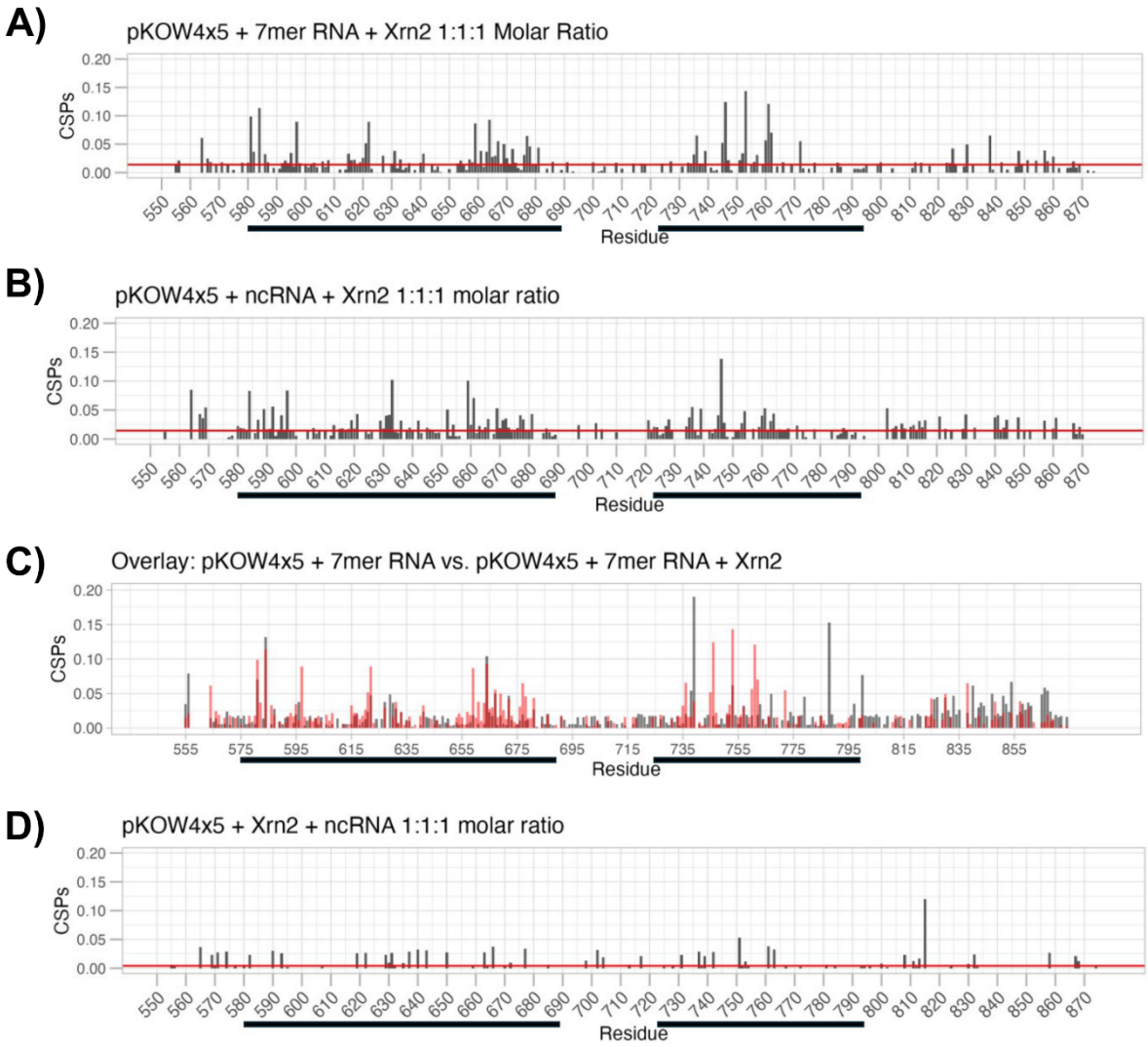


Figure 4.6. Chemical shift perturbation (ppm) plots of Xrn2 addition into KOWx45 & RNA complexes *with mean CSP indicated in red. Horizontal bars indicate domain boundaries from KOWx4 and KOW5. A) CSP plot of Xrn2 added to preformed pKOWx45 & 7mer RNA complex at a molar ratio of 1:1:1. B) CSP plot of Xrn2 added to preformed pKOWx45 & non-cleavable 26mer RNA complex at a molar ratio of 1:1:1. C) Overlay CSP plot of 7mer RNA added to pKOWx45 at a molar ratio of 1:1 (red) vs. Xrn2 added to preformed pKOWx45 & 7mer RNA complex at a molar ratio of 1:1:1. (black) D) CSP plot of noncleavable 26mer RNA addition to preformed pKOWx45 & Xrn2 complex at a molar ratio of 1:1:1.*

Characterization of loop region of Xrn2 via NMR

I also attempted to map the binding interface of the KOWx45 construct on the Xrn2 protein. Full-length Xrn2 is a 112 kDa protein comprised of 991 amino acids, which is at the upper limit of protein NMR capabilities. Increased molecular weight leads to slow tumbling and broadening of the signal. However, Xrn2 is comprised of two conserved domains (CR1 & CR2) linked by a poorly conserved and disordered loop region comprised of residues 390-589. The sequence-based predictor of disordered protein regions, PrDos, also predicts a 176-residue region of disorder between residues 436 and 591. Previous high-resolution structures of Xrn2 in complex with the binding partner Rai1 were unable to visualize this loop, yet crosslinking data from the Vasilieva lab implicated it in forming tethering interactions on the Pol II surface (Kuś et al., 2025; Xiang et al., 2009). Clusters of crosslinks were identified between the loop region of Xrn2 and the Rpb2, Rpb3, and Rpc10 subunits of Pol II. Based on cryoEM structures of a pre-termination complex composed of Spt5, Xrn2, and Rai1, the authors proposed that a positively charged 16-residue helical region of Xrn2 (residues 481-496) within this unstructured loop binds to a negatively charged pocket on the Pol II surface. While the larger unstructured loop has low sequence conservation, this charged helical structure is predicted in human and mouse Xrn2 orthologs. Furthermore, mutation of these arginine residues or loop deletion reduced Xrn2 recruitment to Pol II *in vivo*.

Although the larger structured Xrn2 domains and catalytic centre cannot be visualized using NMR, the high disorder within the loop region reduces line broadening and increase signal to noise, making visualization of this region possible (Fig. 4.6a). While residues within the KOWx45 domains involved in Xrn2 binding were previously

identified, I wanted to investigate whether these contacts are facilitated via the disordered loop region of Xrn2.

Using an identical optimized purification protocol discussed above, WT Xrn2 was isotopically enriched in ^{15}N and ^{1}H, ^{15}N HSQC experiments were conducted on the apo-enzyme as well as following addition of non-cleavable RNA scaffold and the unlabeled KOWx45 domain. 122 peaks were identified in the spectrum compared to the 176 residues predicted in the disordered loop (which includes 9 prolines), all with narrow chemical shift dispersion and sharp peaks, as expected from a disordered protein (Fig. 4.6a).

No significant residue shifts were observed upon addition of the KOW5 domain to the sample (Fig. 4.6b). Three residue movements were observed upon addition of the noncleavable RNA scaffold, but lacking backbone resonance assignments, their identity was unknown (Fig. 4.6c). Due to the significant sample yield and cost required, it was not considered a productive avenue to attempt to collect triple resonance data from this sample. TOCSY and NOESY spectra were collected to gain further information about residue identity, but the signal to noise ratio was too low for effective residue assignment.

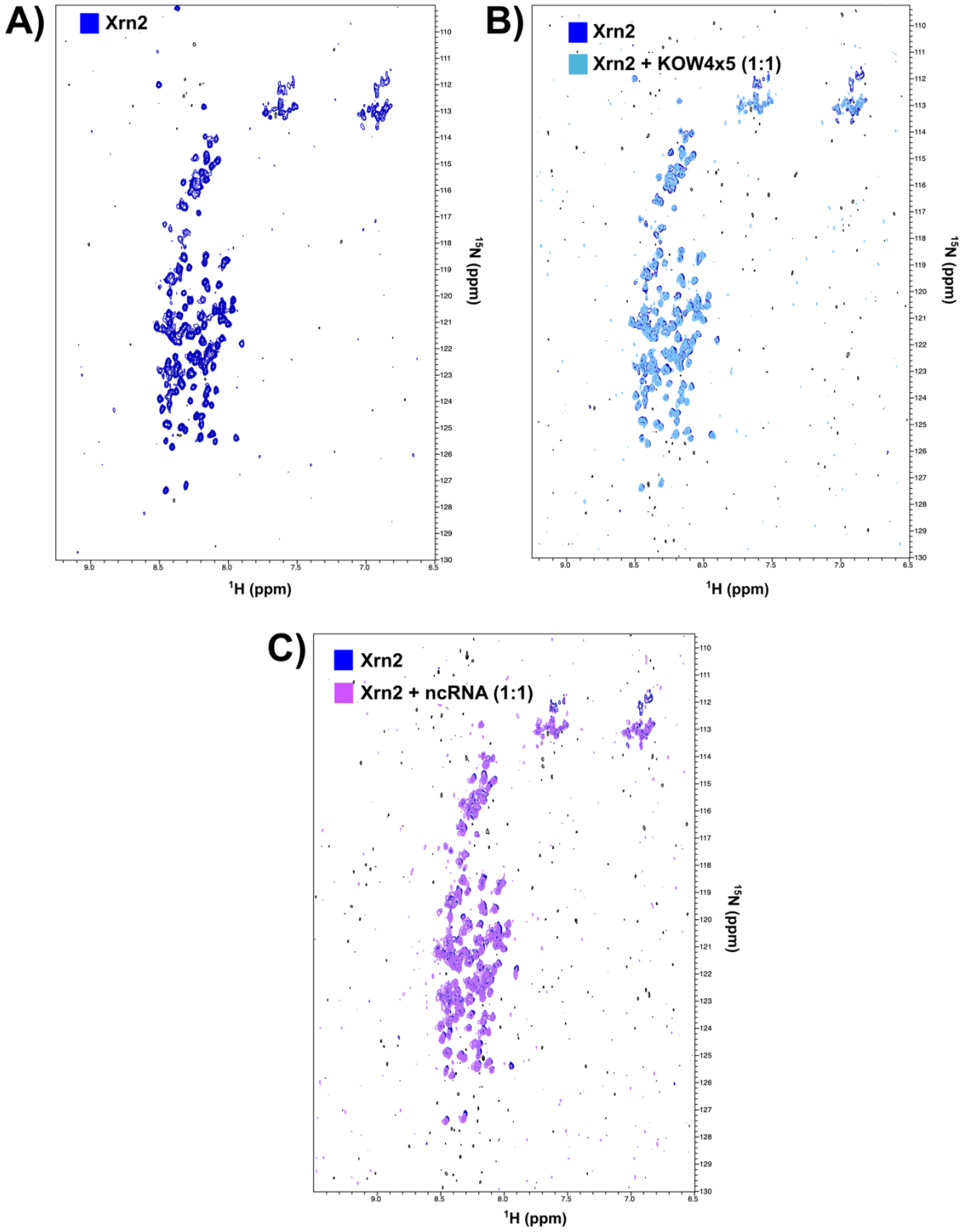


Figure 4.7. Testing interactions of RNA & KOWx45 with the loop of Xrn2. *A) B-TROSY HSQC spectrum of ^{15}N isotopically labelled WT Xrn2 (peaks shown in blue). B) B-TROSY HSQC spectrum of ^{15}N isotopically labelled WT Xrn2 (peaks shown in blue) with overlay of HSQC with unlabelled KOWx45 domain added (peaks shown in cyan). C) B-TROSY HSQC spectrum of ^{15}N isotopically labelled WT Xrn2 (peaks shown in blue) with overlay of HSQC with non-cleavable RNA scaffold added to 1:1 molar ratio (peaks shown in pink).*

From this, we concluded that Xrn2 does not use this unstructured loop region to bind to the KOWx45 domains or to RNA. This aligns with the proposed RNA cleavage mechanism in the conserved and structured active site, with CSPs observed upon addition of the noncleavable scaffold likely occurring due to conformational changes within the protein overall.

Altogether, the findings presented in this chapter support a hypothesis of functional coupling between Spt5 post-translational modification state and engagement with Xrn2. Similar to RNA interaction, binding to Xrn2 is strengthened by KOWx45 CTR phosphorylation. Finally, this interaction interface is likely within the Xrn2 catalytic core as opposed to the disordered loop region of the enzyme. These results provide a starting point for further investigation into the regions of Xrn2 used to contact the KOW5 domain and studies into Xrn2 catalysis and dynamics.

Chapter 5: Discussion

The primary aim of this thesis was to contribute to the wider understanding of the role of Spt5 in transcriptional termination, specifically through investigating interactions of the KOWx45 domains with RNA and Xrn2. To this end, I used NMR to structurally characterize the KOWx45 domains with near full assignment of all backbone resonance assignments. I further mapped the binding interfaces of RNA and the 5-3' exoribonuclease Xrn2 onto these domains. To do so, I optimized a method for *S. pombe* Xrn2 purification at higher yield and stability then previously obtainable for NMR studies, which also has uses beyond those described here.

The data presented here is also supported by *in vivo* research documenting the importance of both Spt5 and Xrn2 presence and activity for efficient and regulated termination. TT-seq experiments show that rapid degradation of Spt5 reduces RNA levels and causes 3’ end processing defects (Kuś et al., 2025). Rapid depletion of Spt5 and Xrn2 leads to defects in termination (Kuś et al., 2025). Additionally, structural data corroborates the findings presented here of the KOW5 domain being implicated in contacting Xrn2.

This work can be used as a platform to further understand and analyse interactions of these domains with additional termination factors in future experiments. Spt5 is involved in other stages of transcription, and its depletion is associated with failure to process pre-mRNA co-transcriptionally by the spliceosome and recruit of 5’ to 3’ exoribonuclease Xrn2 mediating dislodgment of Pol II from DNA (Kuś et al., 2025).

What is the role of Spt5 CTR dephosphorylation in prompting termination?

Spt5 is an essential transcription elongation factor highly conserved in both structure and function, yet an understanding its regulation of transcriptional termination remains incomplete. To address how Spt5 may regulate termination through interactions with RNA or the exoribonuclease Xrn2, I focused on the KOWx45 domains of Spt5, a region implicated in these interactions.

Achieving the near complete chemical shift assignment of the KOWx45 backbone resonances provided a robust foundation for mapping interaction sites and comparing the unbound structure to those complexed with RNA or Xrn2. These results allow identification of binding interfaces and begin to shed light on the structural basis by which Spt5 may influence termination dynamics.

Addition of RNA into ^{15}N isotopically labelled KOWx45 and phosphorylated KOWx45 samples suggested that phosphorylation of Spt5 is a modulator of its function. Dephosphorylation by Dis2 after Pol II passes the poly(A) site leads to Pol II deceleration, Spt5 dislodgement and eventual termination. As the KOWx45 domains are proposed to be important in contacting exiting RNA and maintaining the 'RNA clamp', the phosphorylation of Spt5 may be linked to RNA binding and/or removal (Bernecky et al., 2016; Zuber et al., 2018). Analysis of CSPs showed strongly perturbed residues primarily in the KOWx4 domain, correlating with positively charged regions of the protein that would allow contacts with the phosphate backbone of RNA. The identified region of RNA binding supports cryo-EM structural data which positions the KOWx4 β-barrel near the exiting RNA tunnel, forming the RNA clamp.

NMR results from CSP analysis support the hypothesis that CTR phosphorylation may lead to local rearrangement of the CTR region and potentially conformation shifts in the structured domains that would explain the enhanced RNA binding. Small-angle scattering (SAXS) or hydrogen-deuterium exchange mass spectrometry (HDX-MS) would be useful secondary techniques to identify general changes in secondary structure or protein compaction upon phosphorylation. These techniques would give complementary information on large-scale or global structural rearrangements complementary to residue-specific changes obtained from CSP analysis and NMR.

Interestingly, no CSPs were observed in the linker between the KOWx4 and KOW5 domains, despite this region being previously implicated in RNA binding and Spt5 regulation during transcription (Zuber et al., 2018). However, these observations were made with human Spt5, and when analysed, sequence conservation of the linker between species is poor (Fig. 5.1). Additionally, when examining the electrostatic surface potential of the *S. pombe* KOWx45 domains, no strong positively charged patches of residues that would enhance nucleic acid binding are seen. Although binding to single stranded RNA can also be mediated by pi stacking of aromatic residues such as phenylalanine and tyrosine with the exposed apolar bases (Jones et al., 2001), no aromatic residues are identified in the *S. pombe* linker region.

Thus, the structure and function of Spt5 could have diverged between species, allowing diverse levels of regulation. This is further seen in human Spt5, where phosphorylation at residue S666 within the linker is a key step in releasing promoter proximal paused Pol II. This serine is not conserved between human and *S. pombe*

Spt5 orthologs, indicating that this is potentially a gained trait required for transcription and regulation of the increased complexity of the human genome.

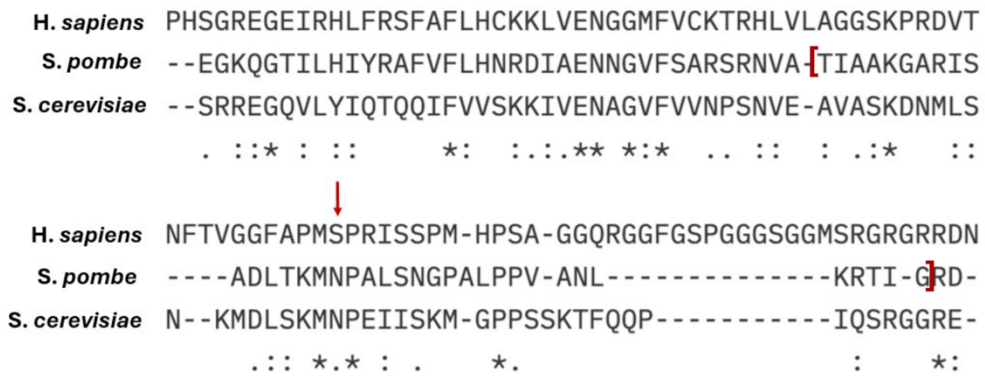


Figure 5.1. Multiple sequence alignment (produced using Clustal Omega) of Spt5 linker sequence from H. sapiens, S. pombe, & S. cerevisiae *with S666 residue highlighted showing lack of conservation of regulatory phosphorylated residue. Boundaries of S. pombe KOWx45 linker are highlighted in red.*

As mentioned, a notable effect observed in the unphosphorylated sample was a loss of signal intensity upon RNA addition. This effect was also dependent on RNA length, with signal loss becoming more pronounced as RNA length increased and is likely due to formation of a larger complex upon RNA addition. If more than one protein molecule binds to the same piece of RNA, a large species would be formed, experiencing slower tumbling and therefore longer relaxation times that would affect the detectable signal. However, if this was the case, it remains unclear why this effect was not observed during addition of RNA into phosphorylated KOWx45. Further work is required to determine the nature of this effect, potentially testing binding to a wider range of RNA lengths or using individual domains to prevent a ‘daisy-chaining’ effect.

In the context of nucleic acid interactions by the KOW domains of Spt5, my findings complement those reported by Zuber et al. (2018). In comparison to Zuber et al., here

we studied the combined KOWx45 construct, capturing a more physiologically relevant idea of interactions with RNA. Importantly, this study diverges in methodological detail concerning the ionic environment of the construct. Zuber et al. did not explicitly control for divalent cation concentrations in their binding assays. This nuance in approach may explain the apparent discrepancies in binding strength between studies.

How does Xrn2 bind to the Pol II complex?

In addition, this study aimed to identify how Spt5 interacts with the exoribonuclease Xrn2. Combined crosslinking and cryoEM studies identified novel interactions between Xrn2 and the KOW5 domain of Spt5. To determine how Spt5 interacts with Xrn2, we conducted NMR experiments adding Xrn2 to isotopically labelled KOWx45 domains. This confirmed the proposed contribution of the KOW5 domain to this interaction and mapped the binding interface. Chemical shift perturbations were observed primarily in the second, third, and fourth β-strands within the KOW β-barrel fold. The loop between the third and fourth strands (residues 760-763) showed the strongest response. This region lies within the KOW5 barrel structure previously implicated in Pol II interaction, potentially suggesting that Spt5 serves as a dynamic interface interacting with different termination complex components.

These results could provide insights into how Xrn2 is recruited to Pol II and how Spt5 is dislodged from the polymerase surface. In the ‘sitting duck torpedo’ model of termination, Spt5 does not directly interact with Xrn2, and nascent RNA degradation precedes contact and dislodgement of Pol II from the DNA template. However, my NMR studies indicate a direct interaction between the two components, supported by

crosslinking data from the Vasilieva lab (Kuś et al., 2025). Furthermore, ChIP-seq studies show that Spt5 depletion also leads to near complete loss of Xrn2 association with chromatin. Taken together, these insights shift our current understanding of the model of transcriptional termination.

Interestingly, the residues in the KOW5 domain proposed to form part of Xrn2 interaction interface overlap with those identified in previous cryoEM structures to bind to the Pol II surface. This raises interesting questions as to the role of Xrn2 in potentially also binding and displacing Spt5 from the polymerase surface, which could facilitate termination, as Spt5 presence is required to maintain both the DNA and RNA 'clamps' for Pol II processivity. Exchange and displacement of different transcription factors is a known phenomenon, and some of these dynamic interactions may be mutually exclusive. Previous structural studies from the Vasilieva lab indicated that Xrn2 tethers to the Pol II surface using a small helical region within its large unstructured loop. The binding interface of the Xrn2 short helix on the Pol II body overlaps with that of the U1 splicing complex, which is associated with the elongation complex. The interaction could anchor the enzyme, allowing more stable binding to Pol II and contact with Spt5 and nascent RNA. The ability to form a more stable interaction with Spt5 post-tethering could contribute to remodelling of the complex into one more predisposed for termination.

Additional NMR studies and Xrn2 activity assays indicated that Xrn2 has a high affinity for RNA, shown by the near complete lack of CSPs when ncRNA was added to a preformed ^{15}N KOWx45 & Xrn2 complex. This 'stealing' of nascent exiting RNA by Xrn2 would destabilize both binding of Spt5 domains to Pol II and the maintenance of

processive Pol II while also increasing RNA tension and favouring release of the polymerase entirely (Rengachari et al., 2025).

In the context of transcriptional termination, my data supports a model in which Spt5 dephosphorylation past the poly(A) site could weaken maintenance of the 'RNA clamp'. Dephosphorylation of the CTR region seems to weaken contacts between Spt5 and nascent exiting RNA, which are important for maintaining Pol II stability and processivity. Thus, dephosphorylation may contribute to Pol II deceleration during termination. Phosphorylation of the Spt5 CTR may also have indirect impacts on transcription, like the Pol II CTD, by tuning which accessory factors can bind to the Pol II complex, which in turn influence transcriptional progression and promote termination. Simultaneously, Xrn2 binding to the KOW5 domain to displace it from the polymerase surface, and Xrn2 preferentially binding to the nascent RNA transcript could combine to facilitate termination at the ends of genes (Fig. 5.2).

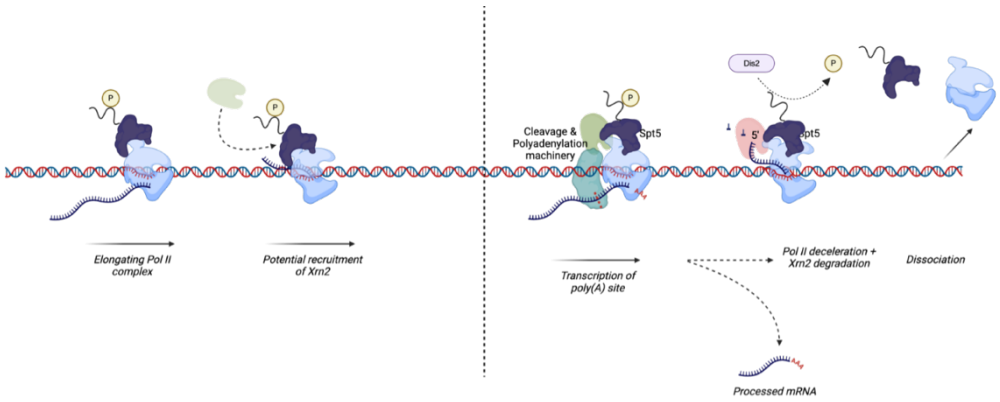


Figure 5.2. Insights into Spt5's role in transcriptional termination. *Xrn2 can interact with phosphorylated Spt5, particularly via the KOW5 domain. After transcript cleavage past the poly(A) site, Spt5 is dephosphorylated, which plays a role in remodelling of the Pol II complex. This includes removal of several Spt5 domains, weakening of the RNA clamp of Pol II and recruitment of further termination factors. Xrn2 degrades the cleaved nascent transcript, enhanced by the KOWx45 domains, leading to destabilisation of the complex and eventual Pol II dissociation.*

However, it remains unknown what triggers the initial dephosphorylation of the Spt5 CTR by Dis2 phosphatase. Transcription past the poly(A) site could lead to conformational or post-translational modification changes in the Pol II complex that allow recruitment of specific termination factors. Dis2 is known to be important for dephosphorylation of both Spt5 and the Pol II CTD at position Thr4 in *S. pombe*. In *S. cerevisiae* and *H. Sapiens*, Dis2 homologs remove Tyr1 phosphorylation and Ser5 phosphorylation of the Pol II CTD respectively (Schreieck et al., 2014; Washington et al., 2002).

Dis2 phosphatase activity mediates recruitment of termination factors such as Seb1 to the Pol II CTD. Seb1 is a pause-promoting factor as well as a regulator of co-transcriptional RNA processing (Duval et al., 2023). Crosslinking studies of Seb1 to Pol II identified links to regions of the Pol II surface in proximity to regions identified as Spt5 binding surfaces (Kecman et al., 2018). This reinforces the notion that removal of

Spt5 both destabilizes Pol II elongation and permits binding of termination and transcript processing factors, and that this is partially mediated by dephosphorylation at the 3' ends of genes.

Dis2 is a component of the phosphatase module of the cleavage and polyadenylation factor (CPF), which recognizes the polyadenylation signal, cleaves, and polyadenylates the mRNA. Theoretically, association of the CPF after selection of the poly(A) site would bring Dis2 into proximity allowing dephosphorylation of its substrates, including Spt5. However, PP1 (mammalian Dis2 homolog) has also been identified to associate with the complex earlier than the termination site and has roles in promoter-proximal pausing at the 5' end of genes (Kecman et al., 2018; Parua et al., 2018). This leaves uncertainties regarding what triggers the dephosphorylation of the Spt5 CTR (and Pol II) during termination.

Additionally, it is still uncertain why the phosphorylated KOWx45 construct would show significantly larger CSPs upon binding to Xrn2. As discussed above, dephosphorylation of Spt5 occurring at the transcriptional end site is thought to be a significant factor triggering termination. Thus, it is puzzling why phosphorylated Spt5, which is found in the elongating Pol II complex would interact much stronger with Xrn2. It could be that Xrn2 is recruited earlier than during termination and activated only after CPF cleavage of the nascent transcript providing an entry point for the endoribonuclease. Preferential engagement with the phosphorylated form could serve as a regulatory checkpoint ensuring all appropriate termination and processing factors are present when required.

Xrn2 itself has been identified as a phosphorylation target of CDK9, and mutations preventing phosphorylation impair its recruitment to chromatin (Sansó et al., 2016). This supports the idea of dynamic phosphorylation/dephosphorylation during transcription by CDK9 being an important regulatory mechanism ensuring that termination factors are correctly positioned. ChIP-seq studies after depletion of Spt5 showed severely compromised recruitment of Xrn2 and loss of Xrn2 association with chromatin, indicating Spt5 is central for ensuring termination proceeds smoothly (Kuś et al., 2025). Phosphorylation of the sCTR may provide a more favorable environment strengthening the interactions with Xrn2 after it 'tethers' to the Pol II surface using the helix found within the loop region. Alternately, the observed effect could be a result of secondary effects of Spt5 rearrangements upon phosphorylation.

However, further studies, such as cross-linking mass spectrometry, *in vitro* transcriptional assays, or *in vivo* ChIP-seq studies investigating whether Xrn2 recruitment to the Pol II complex overall is affected by Spt5 CTR (or even Pol II CTR) phosphorylation state are needed to confirm this.

Limitations & Future Directions

The most immediate future direction for this project is mutational analysis based on the NMR results presented here. Based on my RNA binding data, I designed and cloned three KOWx45 mutant constructs to test their effects on RNA interactions. Residues R685 & I661 were chosen to be mutated to alanine as they were highly perturbed by RNA binding. Additionally, a double H761A/T762A mutant was generated to test the effect on Xrn2 interaction. These residues are found within the KOW5 loop implicated in Xrn2 binding. These mutants now need to be purified and investigated to

evaluate whether these residues are involved in Xrn2 binding. If the double mutant impairs Xrn2 binding then single mutants should be tested, as my NMR results suggested that the interaction is multivalent, possibly involving multiple weaker contacts.

An additional avenue for further experimentation centres around characterizing the dynamics of the Xrn2 catalytic centre. Xrn2 is a highly processive enzyme, which allows it to degrade the cleaved nascent transcript. Static structural evidence indicated that the N-terminal α1 helix of the enzyme is key for the catalytic cycle of the enzyme, acting as a Brownian ratchet to translocate the substrate into the catalytic centre. Substrate entry is characterized by conformational change into an active state followed by hydrolysis and product release.

Previous work from the Vasilieva lab indicated the KOW5 domain was able to stimulate enzymatic activity, but did not give mechanistic insight into how this occurred. To investigate this, I began a series of ^{19}F relaxation dispersion experiments to characterize the conformational dynamics of *S. pombe* Xrn2. This technique monitors microsecond to millisecond timescale motions by measuring changes in transverse relaxation rates under varying Carr-Purcell-Meiboom-Gill (CPMG) pulse conditions, providing insight into conformational exchange processes. Bromotrifluoroacetone (BTFA) was used as a cysteine-labelling ^{19}F probe to label catalytically active wild-type Xrn2, wild-type Xrn2 with a cysteine introduced at residue 12 at the end of the α1-helix (C12 mutant), and a C12 mutant Xrn2 with all other background cysteines removed.

It is predicted that the presence of non-cleavable RNA would shift the conformational equilibrium of the enzyme into a catalytically active pre-hydrolysis state and that a

similar trend would be observed in the presence of the KOWx45 construct. Although constructs were successfully generated and labelled, completion of the full experimental workflow was not possible. It would be interesting to not only investigate the impacts of KOWx45 & non cleavable RNA on the exchange rate of Xrn2 but also examine the effects of the phosphorylated construct and the KOW5 loop mutant predicted to destabilize the interaction between the two binding partners. Gaining a mechanistic understanding of the contribution of the KOWx45 construct to Xrn2 activity would complement the binding results presented here and contribute to the overall model of Xrn2 recruitment and activation during termination.

Furthermore, transcriptional termination is tightly coupled to RNA capping, splicing, and 3' end processing. While not addressed here, DSIF (Spt4 & Spt5), and specifically Spt5's CTR, has been implicated as a scaffold for assembly of processing, capping, and polyadenylation & cleavage factors. In yeast, mutation of DSIF genes is synthetically lethal with capping enzyme mutations and leads to altered splicing and 3' end formation (Lindstrom et al., 2003). Human Spt5 was also shown to stimulate mRNA capping enzymes (Wen & Shatkin, 1999). It would be interesting to use the NMR protocol described here to identify further binding partners during termination, with a focus on the impact of the phosphorylation state of the sCTR.

Additionally, the mechanism of regulation of 3' end processing likely differs between species, analogous to the additional negative regulatory function of the linker seen in humans but not in *S. pombe*. Metazoan Spt5 has a more complex CTR region with two additional KOW domains 6 & 7 (Zuber et al., 2018). While the specific function of these domains is unclear, mutation leads to significant development defects, and it is

proposed they have roles in recruitment and cooperative assembly of processing factors. This would make it interesting to use NMR to identify species specific differences in Spt5 function and identify interactions of different domains with RNA processing or splicing factors. Due to the disordered nature of many of the regions involved in these interactions, NMR represents an optimal method for this investigation.

The major limitation of this study is the simplified model used to investigate Spt5 domain function. Transcription is a complex process involving an array of different transcription factors involved at various stages. Due to the size limitations of NMR, this study was reductionist, investigating interactions of Spt5 with RNA and Xrn2 alone without the presence of additional binding partners or more pertinently, Pol II. It would have been insightful to conduct titrations of RNA or Xrn2 into a pre-formed KOWx45 & Pol II complex. However, the size of the complex would have caused very slow tumbling and greatly decreased signal to noise as a result.

One experimental approach in the future that could circumvent this issue would be a competition-based titration with Xrn2, as my results proposed overlapping binding interfaces between the Pol II surface and the KOW5 domain. If Xrn2 were added into a pre-formed complex with isotopically labelled KOWx45 and was able to displace KOW5, in theory the signal would become detectable as KOWx45 alone gives sharp peaks and high signal to noise for a protein of its size. This could help to confirm the proposed competition between Xrn2 binding and Spt5 displacement.

Finally, the study lacks *in vivo* data in *S. pombe* supporting the proposed hypothesis. Transient transcriptome sequencing (TT-seq) or chromatin immunoprecipitation

sequencing (ChIP-seq) of Xrn2 or Spt5 could be methods used to confirm the role and presence of individual Spt5 domains in interactions with RNA and Xrn2. However, this is complicated by the multifaceted nature of Spt5 during transcription. As it is an essential accessory factor, mutation or depletion *in vivo* may lead to widespread deleterious effects throughout transcription, making it difficult to identify termination specific defects.

Conclusion

Here, I present a near-complete NMR assignment of the solution-state KOWx45 domain construct from Spt5 and provide evidence that phosphorylation of the sCTR region enhances its interaction with both RNA and the termination factor Xrn2. Using this, I mapped the RNA-binding interface to the KOWx4 domain and identified the Xrn2 interaction surface on the KOW5 domain, offering insight into Spt5's role in transcriptional termination.

My data support a model in which KOWx45 engagement with nascent RNA and termination machinery is phosphorylation-dependent, suggesting a regulatory mechanism tied to transcriptional progression and Pol II deceleration. These findings establish a framework for deeper mechanistic studies and raise the possibility that modulation of Spt5's post-translational state could influence transcription termination efficiency. Further work is needed to determine the dynamic nature of these interactions in the full-length protein and in the context of the larger elongation/termination complex.

Overall, an increased understanding of the foundational biological process of transcription mechanisms and transcriptional termination will allow insights into gene

regulation and stability, informing future strategies for targeting dysregulated gene transcription in disease.

Chapter 6: Materials and Methods

Expression plasmids

A pRSFDuet plasmid containing WT Spt5 (N-terminally His-tagged) and Spt4 as well as a pRSFDuet plasmid containing WT C-terminally truncated Xrn2 (residues 1-885) (C-terminally His-tagged) were generously provided by the Vasilieva lab, Department of Biochemistry, Oxford.

RNA scaffolds

RNA scaffolds were obtained from Eurofins with the following sequences: 5mer RNA (AUACA), 7mer RNA (AUACACU), 10mer RNA (AUACUCUGAC), 26mer RNA (UGAAUCUAUUUCUUUUAUCGAGAGGU), Non-cleavable RNA (ncRNA) (5P-A*C*A*G*AUCGUGUCCAUCGAGAGGU), FAM-labelled RNA (5P-FAM-UGAAUCUAUUUCUUUUAUCGAGAGGU)

KOWx45 & Xrn2 construct cloning

KOWx45 domain construct with 8 C-terminal domain repeats was cloned from pRSFDuet vector containing full-length Spt5 & Spt4. KOWx45 domain insert was amplified using PCR with the following reaction mixture on ice: 25µl Phire master mix (ThermoFisher Scientific), 2.5µl forward & reverse primers (10mM stock), 1µl template plasmid (100ng/µl stock diluted 1:10), 19µl Milli-Q (MQ) water. pRSFDuet vector was amplified using PCR with the following reaction mixture on ice: 10µl Q5 5x buffer, 1µl dNTPs (10mM stock), 0.5µl Q5 enzyme, 2.5µl forward & reverse primers (10mM stock), 1µl template plasmid (100ng/µl stock diluted 1:10), 32.5µl MQ water. Samples were transferred to a thermocycler for PCR amplification.

1µl Dpn1 (New England Biolabs) was added to both PCR products and incubated for 1 hour at 37°C to remove cell-derived plasmid template. PCR products were purified using QIAquick PCR Purification Kit for PCR Cleanup. DNA concentration was quantified using Nanodrop spectrophotometer (ThermoFisher Scientific). PCR products were assembled via Gibon assembly with the follow reaction mixture on ice: 15ng of KOWx45 insert (82ng/µl stock), 60ng of pRSFDuet vector (48.9ng/µl stock), 5µl NEBuilder HiFi DNA Assembly master mix (New England Biolabs), 3µl MQ water. Assembly reaction was incubated 1 hour at 50°C and stored on ice.

Xrn12 R12C mutant was cloned from pRSFDuet vector containing WT Xrn2 (residues 1-885). Mutant primers were designed in SnapGene and purchased from Invitrogen. Xrn2 was amplified using PCR with the following reaction mixture on ice: 25µl Phire master mix, 2.5µl forward & reverse mutant primers (10mM stock), 1µl template plasmid (100ng/µl stock diluted 1:10), 19µl MQ water. Samples were transferred to a thermocycler for PCR amplification.

1µl Dpn1 (New England Biolabs) was added to both PCR products and incubated for 1 hour at 37°C to remove cell-derived plasmid template. PCR products were purified using QIAquick PCR Purification Kit for PCR Cleanup. DNA concentration was quantified using Nanodrop spectrophotometer (ThermoFisher Scientific).

The seven native cysteine residues were removed from the Xrn12 R12C mutant in a similar protocol. Two gene blocks were ordered covering the N and C terminal regions of the protein where all cysteines mutated to alanine. Primers were ordered to amplify the remaining middle region of the protein. Middle insert and plasmid vector were amplified using PCR with the following reaction mixture on ice: 25µl Phire master mix,

2.5µl forward & reverse mutant primers (10mM stock), 1µl template plasmid (100ng/µl stock diluted 1:10), 19µl MQ water. Samples were transferred to PCR thermocycler for PCR. PCR products had DNA loading dye added and were run on a 1% agarose gel. Products were purified using Monarch DNA gel extraction kit. PCR products were assembled via Gibon assembly with the follow reaction mixture on ice: 40ng of Xrn2 R12C mutant vector, 15ng of middle insert, 10ng of each gene block, 2.5µl NEBuilder HiFi DNA Assembly master mix (New England Biolabs), 1.7µl MQ water. Assembly reaction was incubated 1 hour at 50°C and stored on ice.

2µl of each respective Gibson product for KOWx45, Xrn2 R12C mutant, and Xrn2 R12C full mutant was incubated with 200µl of DH5α competent cells on ice for 30 minutes before being heat shocked for 45 seconds at 42°C. Cells were recovered on ice for 2 minutes before addition of 400µl of SOC recovery media and incubation for 1 hour at 37°C with shaking (180 rpm). 50µl of cells were then plated onto kanamycin LB (high salt) plates with working concentration 100mg/L, and plates were incubated overnight at 37°C. Single colonies were picked from the plate and added to starter cultures of 5ml LB (high salt) media with kanamycin (100ug/ml). Starter cultures were incubated overnight at 37°C with shaking (180 rpm). Plasmids were then purified using QIAprep Spin Miniprep kit, and 5µl at 100ng/µl of plasmid product was sent for sequencing (Source Biosciences).

Expression & Purification of KOWx45

2µl of KOWx45 construct plasmid was incubated with 200µl of BL21(DE3) competent cells on ice for 30 minutes before being heat shocked for 45 seconds at 42°C. Cells were recovered on ice for 2 minutes before addition of 400µl of SOC recovery media

and incubation for 1 hour at 37°C with shaking (180 rpm). 50µl of cells were then plated onto kanamycin LB (high salt) plates with working concentration 100mg/L, and plates were incubated overnight at 37°C.

A single colony was picked from the plate and added to a starter culture of 100ml of LB (high salt) media with kanamycin (100ug/ml). The culture was incubated overnight at 37°C with shaking (180 rpm). 50ml of the overnight starter culture and 1ml of kanamycin (100ug/ml) were added per 1 litre of LB (high salt) and grown to OD 0.8 at 37°C with shaking (180 rpm). The cultures were induced with 1mM IPTG overnight at 18°C with shaking (180 rpm).

Cells were harvested by centrifugation for 30 minutes in a Beckman Avanti centrifuge with JLA 8.1 rotor at 5,000x *g* and 4°C and cell pellets were frozen at -80°C. Pellets were thawed and resuspended in 40ml of lysis buffer (25mM Tris, 500mM NaCl, 1mM BME, 10mM Imidazole, 5% glycerol, pH 7.5) with 2 protease inhibitor tablets added (Roche cOmplete mini EDTA-free Protease Inhibitor Cocktail). Cells were lysed using QSonica Q700 sonicator at 45 J with active sonication time of 2:30 minutes (5 seconds of sonication with 25 seconds of rest). Cell lysates were cleared by centrifugation for 30 minutes in a Beckman Avanti centrifuge with JA 25.50 rotor at 30,000x *g* and 4°C.

5ml HisTrap High Performance IMAC column (Cytiva) was connected to ÄKTA pure protein purification system and equilibrated with 10 CV of MQ water followed by 10 V lysis buffer (flow rate 5ml/min). Cell lysates were loaded onto column using Äkta sample pump S9H (flow rate 2ml/min) and washed with 10 CV of lysis buffer to remove non-specific binders. Protein was eluted with increasing imidazole gradient of elution buffer (25mM Tris, 500mM NaCl, 1mM BME, 200mM Imidazole, 5% glycerol, pH 7.5).

1.5 ml elution fractions were collected via fraction collector and those containing protein of interest were identified using UV-readings of UNICORN software and confirmed via SDS-Page gel. Column was cleaned with 5 CV of lysis buffer followed by 10 CV MQ water and 5 CV 20% ethanol.

Superdex 75 Increase 10/300 GL SEC column was connected to ÄKTA pure protein purification system and equilibrated with 2 CV of water followed by 2 CV of SEC buffer (10mM HEPES, 100mM NaCl, 1mM $MgCl_2$, 1mM BME, pH 7.5) at a flow rate 0.8ml/min. Protein elution fractions were concentrated to 1ml using Amicon Ultra-15 centrifugal filter concentrator (10K MWCO) before injection into 2ml Superloop (Cytiva). Proteins were eluted over 1 CV of SEC buffer and collected in 0.5 fractions. Fractions containing protein of interest were identified using UV-readings of UNICORN software and confirmed via SDS-page gel. Column was cleaned with 2 CV of MQ water and 1 CV 20% ethanol.

Protein elution fractions were concentrated to 1ml using Amicon Ultra-15 centrifugal filter concentrator (10K MWCO) and flash-frozen and stored at –80°C.

SDS-Page Gel Analysis

30µl of each sample was added to 10µl of NuPage sample buffer (5x) and loaded into a Bolt 12% Bis-Tris Plus gel (Invitrogen). Gels were run at 180V for 35 minutes with Bolt MES SDS running buffer (1x) and stained with InstantBlue Coomassie Blue (Expedeon).

Preparation of in-gel digest samples for mass spectrometry

Gel was first rinsed with MQ water for 15 minutes at room temperature. Bands containing KOWx45 protein of interest as well as lower band with cleavage product of interest were excised from SDS-page gel and transferred into Lobind 1.5ml Eppendorf tubes. Using a pipette tip, bands were cut into small cubes. Gel pieces were incubated with 100µl of destain solution (250µl of 0.1M ammonium bicarbonate, 250µl of 100% acetonitrile) for 30 minutes at room temperature with vortexing every 10 minutes before addition of 500µl of 100% acetonitrile for 5 minutes. Acetonitrile was then removed.

Add 50µl of 10mM TCEP in 0.1M ammonium bicarbonate and incubate at room temperature for 30 minutes to reduce protein before addition of 500µl of 100% acetonitrile for 5 minutes and subsequent removal. Add 50µl of 50mM chloroacetamide in 0.1 ammonium bicarbonate and incubate at room temperature for 30 minutes in the dark to alkylate protein before addition of 500µl of 100% acetonitrile for 5 minutes and subsequent removal.

Gel pieces were digested with 50µl trypsin (0.1ug/µl stock diluted with 10mM ammonium bicarbonate in 10% acetonitrile for 0.5ug enzyme per gel piece) for 20 hours at 37°C with shaking (650 rpm). Trypsin solution was collected from gel pieces, transferred into a Lobind 1.5ml Eppendorf tube, and labelled as fraction 1. 100µl of extraction buffer (250µl 5% formic acid, 500µl 100% acetonitrile) was added to gel pieces and incubated for 15 minutes at 37°C with shaking (650 rpm). Fraction 2 solution was collected from gel pieces, transferred into a LoBind 1.5ml Eppendorf tube, and pooled with fraction 1. 100µl of extraction buffer was added again to gel pieces

and incubated for 15 minutes at 37°C with shaking (650 rpm). Fraction 3 solution was collected from gel pieces, transferred into a LoBind 1.5ml Eppendorf tube, and pooled with fractions 1 & 2. Using speedvac pump, dry peptides, and store at –20°C.

Preparation of in-solution digest samples for mass spectrometry

50µl of fresh 8M urea solution in 100mM ammonium bicarbonate buffer was added to 50µl of KOWx45 protein at 1mg/ml (50ug of protein) and incubated for 10 minutes at room temperature with shaking (650 rpm). Cysteines were reduced via addition of 2µl of 0.5M TCEP and incubation for 30 minutes at room temperature. Reduced cysteines were alkylated via addition of 10.2µl of 0.5M chloroacetamide (prepared in 8M urea & 100mM ammonium bicarbonate solution) and incubation for 30 minutes at room temperature in the dark. KOWx45 sample was predigested using 0.5µl of LysC (1ug LysC per 100ug of sample) for 2 hours at 37°C with shaking (800 rpm). 112.7µl of 0.1M ammonium bicarbonate solution was added to dilute urea concentration to 2M followed by addition of 0.25µl of 1M calcium carbonate solution. KOWx45 sample was digested using 10µl of trypsin (1ug trypsin per 40ug of sample) for 20 hours at 37°C with shaking (800 rpm).

Trypsin digest was stopped via addition of 11.38µl of 100% formic acid, and sample was stored at –20°C until C18 desalting and purification. To prepare C18 tip, use a 2mm diameter syringe to cut a piece of C18 resin and insert into 10µl pipette tip placed into a LoBind 1.5ml Eppendorf tube to create a column. C18 resin was activated by centrifuging 60µl of 100% acetonitrile (HPLC grade) at 1500xg at room temperature through the column twice. Resin was washed by centrifuging 60µl of 0.1%

trifluoroacetic acid (HPLC grade) at 1500xg at room temperature through the column. After a final centrifugation for 2 minutes at 6000xg at room temperature, the solution was discarded. Digested protein samples were centrifuged at 18500xg and 4°C for removal of aggregates and supernatant was collected. Sample was loaded onto C18 resin 60µl at a time and centrifuged until liquid goes through at 1500xg at room temperature. After all sample is loaded onto the resin, a final centrifugation for 2 minutes at 6000xg at room temperature was conducted. Resin was washed by centrifuging 60µl of 0.1% trifluoroacetic acid (HPLC grade) at 1500xg at room temperature through the column. After a final centrifugation for 2 minutes at 6000xg at room temperature and flowthrough solution was collected.

C18 resin tip was placed into a clean LoBind 1.5ml Eppendorf tube and peptides were eluted by centrifuging 60µl of 50% acetonitrile in 0.1% trifluoroacetic acid (HPLC grade) through the column twice. After a final centrifugation for 2 minutes at 6000xg at room temperature, eluted peptides were dried using speedvac pump and stored at –20°C.

Labelled KOWx45 & Xrn2 purifications

For preparation and purification of isotopically labelled KOWx45 and Xrn2 protein samples, a single colony was picked from either KOWx45 or Xrn2 transformation plate (as described above) and added to a starter culture of 100ml of LB (high salt) media with kanamycin (100ug/ml). The culture was incubated overnight at 37°C with shaking (180 rpm). 50ml of the overnight starter culture and 1ml of kanamycin (100ug/ml) were added per 1 litre of LB (high salt) media (with a minimum of 2 litres of culture grown) and grown to OD 0.8 at 37°C with shaking (180 rpm). Under sterile conditions, 1x M9

salt media (100ml 10x M9 salts for NMR, 900ml MQ water) and 2x M9 salt media (100ml 10x M9 salts for NMR, 800ml MQ water, 2ml 1M magnesium sulphate, 500µl trace metals, 1g ^{15}N ammonium sulphate, 20ml 20% glucose or 2g ^{13}C glucose if double-labelling protein, 1ml each of biotin, thiamine, and kanamycin all at 100mg/ml stock concentration) were prepared and mixed.

After LB (high salt) cultures reached OD 0.8, cell pellets were harvested by centrifugation for 30 minutes in a Beckman Avanti centrifuge with JLA 8.1 rotor at 5,000x *g* and 4°C. Pellets from 2 litres of culture were condensed and resuspended in 1 litre of 1x M9 media. Cell pellets were again harvested by centrifugation for 30 minutes in a Beckman Avanti centrifuge with JLA 8.1 rotor at 5,000x *g* and 4°C. Pellets were resuspended in 1 litre of 2x M9 media and left to rest for 30 minutes at room temperature. The cultures were induced with 2mM IPTG overnight at 18°C with shaking (180 rpm) before purification according to above protocol for unlabelled sample.

Backbone resonance assignment of KOWx45 construct

Triple resonance experiments allowed sequence specific backbone resonance assignment of the KOWx45 protein construct. Each peak in the ^{1}H-^{15}N HSQC spectrum represents one amino acid, and the triple resonance workflow allowed peak identification by correlating the Cα, Cβ, and Co shifts of neighbouring residues. Five different triple resonance experiments were used here to complete the assignment: HNCO, HNCA, HN(CO)CA, HN(CA)CO, and HNCACB (Appendix 4.1).

After picking all peaks from the HSQC spectrum, pairs of spectra such as HNCO and HN(CA)CO, were used to establish connectivity. Each HNCO strip provides only one

peak corresponding to that of the carbonyl carbon of the preceding residue (C'_{i-1}), while the HN(CA)CO strip provides two peaks corresponding to the intra-residue carbonyl carbon and the preceding carbonyl carbon (C'_i and C'_{i-1}). Matching C'_i with C'_{i-1} of the next residue allowed 'walking' along the protein backbone sequence (Appendix 4.2). A similar procedure was used with the HNCA and HN(CO)CA strips to confirm the proposed connectivity, with the HNCA providing both intra-residue and inter-residue (preceding) Cα signals and the HN(CO)CA providing only the inter-residue Cα.

Once several residues were confidently linked together using the above workflow, the CARA software fragment alignment feature was used to propose possible positions for the linked resonances within the KOWx45 sequence. The final sequence alignment was confirmed using the HNCACB strips, which provide Cα and side chain Cβ shifts that are distinct for each amino acid type.

KOWx45 & RNA NMR titration setup

KOWx45 protein fractions were thawed and buffer exchanged into NMR buffer using Amicon Ultra-15 centrifugal filter concentrator (10K MWCO) before dilution to desired concentration. NMR samples were mixed with RNA stock using RNase-free pipette tips in LoBind 1.5ml Eppendorf tubes. Samples were pH'd to pH 6.4 before transfer into 3ml NMR tubes using Pasteur pipettes for NMR and loaded into SampleJet sample rack in 950 MHz NMR spectrometer.

In vitro KOWx45 phosphorylation

In vitro phosphorylation of KOWx45 CTR was performed with 30µl of 1mg/ml KOWx45, 4µl of 0.4mg/ml CDK9 (purified by Dr. Tea Kecman, Vasilieva lab), 4µl 10mM $MnCl_2$, 4µl 10mM ATP. Sample was incubated for 1 hour at 25°C and 7µl aliquot was added to 3µl gel loading dye. Phosphorylation was confirmed via analysis on Phos-tag™ gel (0.85ml MQ water, 1.9ml 1M Tris pH 8.8, 2.1ml 30% acrylamide, 25µl 5mM Phos-tag™, 50µl 10% SDS, 25ul, 10mM $MnCl_2$, 20µl 3% APS, 15µl TEMED). Gel was stained with Coomassie blue and visualized.

Reaction volumes were scaled up appropriately for larger-scale phosphorylation of ^{15}N labelled KOWx45 samples used for NMR. Before collection of HSQC spectra, paramagnetic Mn^{2+} ions were removed via dialysis into NMR buffer using Slide-A-Lyzer MINI Dialysis Devices 3.5K MWCO, 0.5ML.

Expression & Purification of Xrn2

2µl of WT C-terminally truncated Xrn2 plasmid was incubated with 200µl of BL21(DE3) competent cells on ice for 30 minutes before being heat shocked for 45 seconds at 42°C. Cells were recovered on ice for 2 minutes before addition of 400µl of SOC recovery media and incubation for 1 hour at 37°C with shaking (180 rpm). 50µl of cells were then plated onto kanamycin LB (high salt) plates with working concentration 100mg/L, and plates were incubated overnight at 37°C.

A single colony was picked from the plate and added to a starter culture of 100ml of LB (high salt) media with kanamycin (100ug/ml). The culture was incubated overnight at 37°C with shaking (180 rpm). 50ml of the overnight starter culture and 1ml of

kanamycin (100ug/ml) were added per 1 litre of LB (high salt) and grown to OD 0.8 at 37°C with shaking (180 rpm). The cultures were induced with 1mM IPTG overnight at 18°C with shaking (180 rpm).

Cells were harvested by centrifugation for 30 minutes in a Beckman Avanti centrifuge with JLA 8.1 rotor at 5,000x *g* and 4°C and cell pellets were frozen at -80°C. Pellets were thawed and resuspended in 40ml of lysis buffer 2 (25mM Tris, 500mM NaCl, 40mM Imidazole 1mM ME, 5% glycerol) with 1 protease inhibitor tablet added (Roche cOmplete mini EDTA-free Protease Inhibitor Cocktail). Cells were lysed via sonication at amplitude 45 J with active sonication time of 2:30 minutes (5 seconds of sonication with 25 seconds of rest). Cell lysates were cleared by centrifugation for 30 minutes in a Beckman Avanti centrifuge with JA 25.50 rotor at 30,000x *g* and 4°C.

2ml of 50% Ni-NTA agarose affinity resin slurry (ThermoFisher Scientific) was applied to gravity flow column and equilibrated with 12CV of MQ water followed by 12CV of lysis buffer 2. Beads were incubated with cleared cell lysate for 1 hour, applied to column, and washed with 12CV of lysis buffer 2. 12CV elution buffer 2 (25mM Tris, 500mM NaCl, 200mM Imidazole, 1mM bME, 5% glycerol) was incubated with beads for 60 minutes before collection of elution fraction 1. Subsequent elution fractions 2 and 3 were collected in an identical manner. Elution fractions were concentrated to 1ml using Amicon Ultra-15 centrifugal filter concentrator (10K MWCO) and flash-frozen and stored at –80°C.

KOWx45 + Xrn2 NMR setup

KOWx45 and Xrn2 protein fractions were thawed and buffer exchanged into NMR buffer using Amicon Ultra-15 centrifugal filter concentrator (10K MWCO) before

dilution to desired concentration. NMR samples were mixed in 1.5ml Eppendorf tubes and pH'd before transfer into 3ml NMR tubes using Pasteur pipettes for NMR and loaded into Bruker SampleJet sample rack in 950 MHz NMR spectrometer.

Xrn2 EMSA

Non-cleavable RNA (Eurofins) was used for EMSA assays. A set of samples containing 5uM ncRNA with 1:0.5, 1:1, 1:2, 1:4 molar ratios of WT C-terminally truncated Xrn2 added to a final sample volume of 10ul. Samples were incubated for 20 minutes at room temperature and 3µl of loading dye was added. Samples were analysed on a pre-cast 6% DNA retardation gel (Invitrogen) in 0.5x TBE buffer at 150 V. The gel was stained with ethidium bromide followed by Coomassie blue stain and visualized with FLA-600 Fujifilm image analyser.

Xrn2 RNA activity assay

RNA degradation assay reactions were set up to a final volume of 20ul, mixing either non-cleavable RNA or hydrolysable RNA with WT C-terminally truncated Xrn2 at 20uM each (1:1 molar ratio). 4µl aliquots of each reaction were taken at 0, 10, 30, and 60 minutes and quenched in STOP buffer (10mM EDTA, 3.5M urea, 50ug/ml heparin, 0.01% bromophenol blue, 0.015% xylencyanol in formamide). Samples were resolved using Urea SDS-Page gel (3.3ml 3% acrylamide, 100mg Tris, 55mg boric acid, 20µl 500mM EDTA, 4.8g urea, 6.7ml (MQ) water, 14µl 10% APS, 7µl TEMED) in 1X TBE buffer for 30 minutes at 150V and visualized via ethidium bromide stain.

NMR experiments

Triple resonance spectra for unphosphorylated and phosphorylated KOWx45 samples (500µl at 100uM with 5% D2O) were collected on a 950MHz Bruker NMR spectrometer. The band selective excitation short-transient transverse relaxation optimized spectroscopy (BEST-TROSY) versions of HNCA, HNCO, HN(CA)CO, HN(CO)CA, & HN(CA)CB spectra (Solyom et al., 2013) collected for each sample.

RNA and Xrn2 titration spectra were collected on a 750MHz Bruker NMR spectrometer (200µl at 50uM with 5% D2O). The 2D BEST-TROSY pulse sequence (Solyom et al., 2013) was used for all interaction experiments.

Table 6.1 Acquisition Parameters for NMR experiments

Experiment Name	Pulse Programme Name	^{1}H Freq.	Carrier			Spectral Widths			Time Domain			Reference Chemical Shift
			^{1}H	^{15}N	^{13}C	^{1}H	^{15}N	^{13}C	^{1}H	^{15}N	^{13}C	^{1}H
B-TROSY	BT_TROSY	950 MHz	4.791	118.596	-	12.34	27.00	-	1276	256	-	DSS at 0ppm
HNCO	BT_HNCO	750 MHz	4.791	118.596	175.725	12.02	27.00	15.00	1276	72	110	DSS at 0ppm
HNCACO	BT_HNCACO	750 MHz	4.791	118.596	175.725	12.02	27.00	13.00	1276	64	110	DSS at 0ppm
HNCA	BT_HNCA	750 MHz	4.791	118.596	56.753	12.02	27.00	30.00	1276	72	120	DSS at 0ppm
HNCACB	BT_HNCACB	750 MHz	4.791	118.596	41.754	12.02	27.00	70.00	1276	66	120	DSS at 0ppm

Chemical shift perturbations (CSPs) were calculated with the formula below via a minimal shift approach based on overlay of each titration spectrum on the assigned ^{1}H-^{15}N HSQC spectrum peaks (Williamson, 2013)

$$CSP = \sqrt{0.5 \times (\delta_H^2 + (0.2 \times \delta_N^2)}$$

where δ_H^2 and δ_N^2 are the chemical shift difference in the ^1H and ^{15}N dimensions, respectively. NMR spectra were processed using NMRPipe (Delaglio et al., 1995). Visualization and analysis of the processed spectra was done using CARA (Keller, 1996).

Appendix

Protein Sequences

His-tagged KOWx45 construct

QAVNSAYELHDLVQLDVNTVACIFSVDRDTYKVIDQNGGVRTVLASQITMRHSNRRGVATD
RNGAEIRIGDKVKEVGGEGKQGTILHIYRAFVFLHNRDIAENNGVFSARSRNVATIAAKGARI
SADLTKMNPALSNGPALPPVANLKRTIGRDKAIGATVRIRRGPMKGLLGVIKDTTDANARVE
LHTGNKMVTIPKENLLYTTKTGELISYTEFIERSRGIRPGSISTADGPNVPNWAQGARTPAVA
NGSRTPAWNTGSRTPAWNSGSKTPAWNSGSRTPAWNSGNKTPAWNAGSRTPAW

His-tagged Xrn2 construct

MGVPALFRLLSRKFAKVITPVIEAPTEKLPDGTEIEPDLSLPNPNGVECDNLYLDMNGIV
HPCSHPEDRPAPETEDEMMVAVFEYTDRILAMVRPRQLLFIAIDGVAPRAKMNQQRSR
RFRSSREAALKEEELQAFIEEAKQQGIPIDENATKKKSWDSNCITPGTPFMDTLAKSLRY
YIINKLNSDPCWRNVRFILSDASVPGEGEHKIMEFIRSQRVKPEYDPNTHHVVYGLDADL
IMLGLATHEPHFRVLREDVFFQQGSTKKTKEERLGIKRLDDVSETNKVPVKKPFIWLNVS
ILREYLEVELYVPNLPFPFDLERAIDDWVFFIFFVGNDFLPHLPSLDIRDGAVERLTEIWRA
SLPHMGGYLTLDGSVNLARAEVILSAVGNQEDDIFKRLKQQEDRRNENYRRRQQRESN
QESESYVDNVVIQRSVETQSTEVVTSSKSTSVDTKPPKKTQKIDAPAPVDLVNLSEKTS
NRSLGATNRELINNRAANRLGLSREAAAVSSVNKLAASALKAQLVSNETLQNVPLEDSI
ASSSAYEDTDSIESSTPVVHPIDTKVSNVGQKRKAPDSTEENENTDTVRLYEPGYRERY
YEQKFHISPDEPEKIREAVKHYVHGLCWVLLYYYQGCPSWTWYYPYHYAPFAADFKDL
ASIDVKFELNQPFKPYEQLLGVLPAASKNNLPEKLQTLMTDENSEIIDFYPENFTIDLNGK
KFEWQGVALLPFIDENRLLNAVSKIYPQLTEEESKRNEDGSTLLFISEHHPMFSELVKQL
YSKKRQGKPLKLSGKMAHGLFGKVNTNDSVIPNVSVQCPIDVTSADALQKYGSIDDNQ
SISLVFEVPKSHFVHKSMLLRGVKMPNRVLTPEDINQVRAERSFSSRRNNGNSGSGHH
HHHHHH

His-tagged Xrn2 cysteine mutant construct

MGVPALFRLLSCKFAKVITPVIEAPTEKLPDGTEIEPDLSLPNPNGVEADNLYLDMNGIV
HPASHPEDRPAPETEDEMMVAVFEYTDRILAMVRPRQLLFIAIDGVAPRAKMNQQRSR
RFRSSREAALKEEELQAFIEEAKQQGIPIDENATKKKSWDSNAITPGTPFMDTLAKSLRY
YIINKLNSDPAWRNVRFILSDASVPGEGEHKIMEFIRSQRVKPEYDPNTHHVVYGLDADL
IMLGLATHEPHFRVLREDVFFQQGSTKKTKEERLGIKRLDDVSETNKVPVKKPFIWLNVS
ILREYLEVELYVPNLPFPFDLERAIDDWVFFIFFVGNDFLPHLPSLDIRDGAVERLTEIWRA
SLPHMGGYLTLDGSVNLARAEVILSAVGNQEDDIFKRLKQQEDRRNENYRRRQQRESN
QESESYVDNVVIQRSVETQSTEVVTSSKSTSVDTKPPKKTQKIDAPAPVDLVNLSEKTS
NRSLGATNRELINNRAANRLGLSREAAAVSSVNKLAASALKAQLVSNETLQNVPLEDSI
ASSSAYEDTDSIESSTPVVHPIDTKVSNVGQKRKAPDSTEENENTDTVRLYEPGYRERY
YEQKFHISPDEPEKIREAVKHYVHGLAWVLLYYYQGAPSWTWYYPYHYAPFAADFKDL

ASIDVKFELNQPFKPYEQLLGVLPAASKNNLPEKLQTLMTDENSEIIDFYPENFTIDLNGK
KFEWQGVALLPFIDENRLLNAVSKIYPQLTEEESKRNEDGSTLLFISEHHPMFSELVKQL
YSKKRQGKPLKLSGKMAHGLFGKVNTNDSVIPNVSVQAPIDVTSADALQKYGSIDDNQ
SISLVFEVPKSHFVHKSMLLRGVKMPNRVLTPEDINQVRAERSFSSRRNNGNSGSGHH
HHHHHH

Mass Spectrometry Data

Table 1: Table of Hits from Mass Spectrometry Analysis of In-solution digest of KOWx45 Sample.

A0A140NCN8	Heat shock protein 70 OS=Escherichia coli (strain B / BL21-DE3) OX=469008 GN=ECBD_3003 PE=3 SV=1	6	1	1	1	556	61.9	5.17	4.46	1
A0A140NFV8	Rhamnulokinase OS=Escherichia coli (strain B / BL21-DE3) OX=469008 GN=rhaB PE=3 SV=1	8	1	1	1	489	54.1	5.2	4.13	1
A0A140N9W8	Pertactin central region domain-containing protein OS=Escherichia coli (strain B / BL21-DE3) OX=469008 GN=ECBD_1979 PE=4 SV=1	4	1	1	1	418	42.9	4.53	3.85	1
A0A140N770	4-hydroxy-tetrahydrodipicolinate synthase OS=Escherichia coli (strain B / BL21-DE3) OX=469008 GN=dapA PE=3 SV=1	12	1	1	1	292	31.3	6.44	3.13	1
A0A140N7C2	Transcriptional regulator, AraC family OS=Escherichia coli (strain B / BL21-DE3) OX=469008 GN=ECBD_1446 PE=4 SV=1	5	1	1	1	354	39.3	8.79	3.02	1
A0A140NH57	Phosphoenolpyruvate-protein phosphotransferase OS=Escherichia coli (strain B / BL21-DE3) OX=469008 GN=ECBD_4076 PE=3 SV=1	2	1	1	1	833	91.8	5.44	2.97	1
A0A140N4M3	Nickel-responsive regulator OS=Escherichia coli (strain B / BL21-DE3) OX=469008 GN=nikR PE=3 SV=1	12	1	1	1	133	15.1	6.24	2.8	1
A0A140N5I9	Lysozyme OS=Escherichia coli (strain B / BL21-DE3) OX=469008 GN=ECBD_0412 PE=4 SV=1	3	1	1	1	686	74.8	6.13	2.69	1
A0A140N855	Tyrosine recombinase XerD OS=Escherichia coli (strain B / BL21-DE3) OX=469008 GN=xerD PE=3 SV=1	4	1	1	1	298	34.2	8.81	2.64	1
A0A140NCZ2	DUF3592 domain-containing protein OS=Escherichia coli (strain B / BL21-DE3) OX=469008 GN=ECBD_2477 PE=4 SV=1	13	1	1	1	135	15.2	5.68	2.48	1
A0A140N923	Iron-sulfur cluster assembly scaffold protein IscU OS=Escherichia coli (strain B / BL21-DE3) OX=469008 GN=ECBD_1155 PE=3 SV=1	9	1	1	1	128	13.8	4.93	2.36	1
A0A140N8N4	Iron-sulfur cluster insertion protein ErpA OS=Escherichia coli (strain B / BL21-DE3) OX=469008 GN=erpA PE=3 SV=1	11	1	1	1	114	12.1	4.31	2.36	1
A0A140N4H3	Small ribosomal subunit protein uS12 OS=Escherichia coli (strain B / BL21-DE3) OX=469008 GN=rpsL PE=1 SV=1	10	1	1	1	124	13.7	10.87	2.24	1
A0A140N2T1	Large ribosomal subunit protein uL6 OS=Escherichia coli (strain B / BL21-DE3) OX=469008 GN=rplF PE=1 SV=1	7	1	1	1	177	18.9	9.7	2.05	1
A0A140N7P2	Multidrug resistance protein MdtA OS=Escherichia coli (strain B / BL21-DE3) OX=469008 GN=mdtA PE=2 SV=1	9	1	1	1	415	44.5	8.66	0	1
A0A140N8Z5	Baseplate J family protein OS=Escherichia coli (strain B / BL21-DE3) OX=469008 GN=ECBD_2756 PE=4 SV=1	12	1	1	1	302	32.5	4.79	0	1

A0A140NCN8	Heat shock protein 70 OS=Escherichia coli (strain B / BL21-DE3) OX=469008 GN=ECBD_3003 PE=3 SV=1	6	1	1	1	556	61.9	5.17	4.46	1
A0A140NFV8	Rhamnulokinase OS=Escherichia coli (strain B / BL21-DE3) OX=469008 GN=rhaB PE=3 SV=1	8	1	1	1	489	54.1	5.2	4.13	1
A0A140N9W8	Pertactin central region domain-containing protein OS=Escherichia coli (strain B / BL21-DE3) OX=469008 GN=ECBD_1979 PE=4 SV=1	4	1	1	1	418	42.9	4.53	3.85	1
A0A140N770	4-hydroxy-tetrahydrodipicolinate synthase OS=Escherichia coli (strain B / BL21-DE3) OX=469008 GN=dapA PE=3 SV=1	12	1	1	1	292	31.3	6.44	3.13	1
A0A140N7C2	Transcriptional regulator, AraC family OS=Escherichia coli (strain B / BL21-DE3) OX=469008 GN=ECBD_1446 PE=4 SV=1	5	1	1	1	354	39.3	8.79	3.02	1
A0A140NH57	Phosphoenolpyruvate-protein phosphotransferase OS=Escherichia coli (strain B / BL21-DE3) OX=469008 GN=ECBD_4076 PE=3 SV=1	2	1	1	1	833	91.8	5.44	2.97	1
A0A140N4M3	Nickel-responsive regulator OS=Escherichia coli (strain B / BL21-DE3) OX=469008 GN=nikR PE=3 SV=1	12	1	1	1	133	15.1	6.24	2.8	1
A0A140N5I9	Lysozyme OS=Escherichia coli (strain B / BL21-DE3) OX=469008 GN=ECBD_0412 PE=4 SV=1	3	1	1	1	686	74.8	6.13	2.69	1
A0A140N855	Tyrosine recombinase XerD OS=Escherichia coli (strain B / BL21-DE3) OX=469008 GN=xerD PE=3 SV=1	4	1	1	1	298	34.2	8.81	2.64	1
A0A140NCZ2	DUF3592 domain-containing protein OS=Escherichia coli (strain B / BL21-DE3) OX=469008 GN=ECBD_2477 PE=4 SV=1	13	1	1	1	135	15.2	5.68	2.48	1
A0A140N923	Iron-sulfur cluster assembly scaffold protein IscU OS=Escherichia coli (strain B / BL21-DE3) OX=469008 GN=ECBD_1155 PE=3 SV=1	9	1	1	1	128	13.8	4.93	2.36	1
A0A140N8N4	Iron-sulfur cluster insertion protein ErpA OS=Escherichia coli (strain B / BL21-DE3) OX=469008 GN=erpA PE=3 SV=1	11	1	1	1	114	12.1	4.31	2.36	1
A0A140N4H3	Small ribosomal subunit protein uS12 OS=Escherichia coli (strain B / BL21-DE3) OX=469008 GN=rpsL PE=1 SV=1	10	1	1	1	124	13.7	10.87	2.24	1
A0A140N2T1	Large ribosomal subunit protein uL6 OS=Escherichia coli (strain B / BL21-DE3) OX=469008 GN=rplF PE=1 SV=1	7	1	1	1	177	18.9	9.7	2.05	1
A0A140N7P2	Multidrug resistance protein MdtA OS=Escherichia coli (strain B / BL21-DE3) OX=469008 GN=mdtA PE=2 SV=1	9	1	1	1	415	44.5	8.66	0	1
A0A140N8Z5	Baseplate J family protein OS=Escherichia coli (strain B / BL21-DE3) OX=469008 GN=ECBD_2756 PE=4 SV=1	12	1	1	1	302	32.5	4.79	0	1

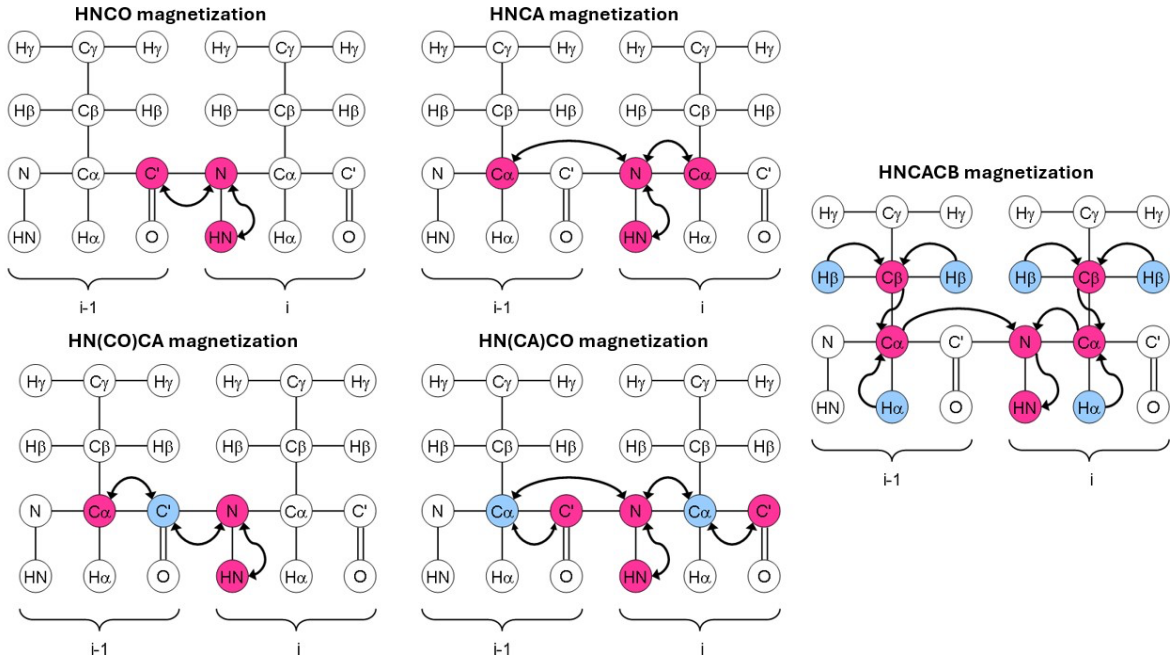
Table 2: Table of Hits from Mass Spectrometry Analysis of In-gel digest of top band of KOWx45 Sample.

Accession	Description	Coverage [%]	# Peptides	# PSMs	# Unique Peptides	# AAs	MW [kDa]	calc. pI	Score Sequest HT: Sequest HT	# Peptides (by Search Engine): Sequest HT
KOW4x5	KOW4x5	9	2	2	2	320	34.5	10.39	4.95	2
Confidence	Sequence	Modifications	# Protein Groups	# Proteins	# PSMs	Master Protein Accessions	# Missed Cleavages	Theo. MH+ [Da]	Confidence (by Search Engine): Sequest HT	XCorr (by Search Engine): Sequest HT
High	NSGNKTPAWNAGSRTPA		1	1	1	KOW4x5	1	1728.84122	High	2.64
High	LHNRDIAENNGVF		1	1	1	KOW4x5	1	1498.73972	High	2.31
Accession	**Description**	**Coverage [%]**	**# Peptides**	**# PSMs**	**# Unique Peptides**	**# AAs**	**MW [kDa]**	**calc. pI**	**Score Sequest HT: Sequest HT**	**# Peptides (by Search Engine): Sequest HT**
A0A140N8I9	Glycosyl transferase group 1 OS=Escherichia coli (strain B / BL21-DE3) OX=469008 GN=ECBD_1605 PE=4 SV=1	4	1	1	1	407	44.9	8.76	1.99	1
Confidence	Sequence	Modifications	# Protein Groups	# Proteins	# PSMs	Master Protein Accessions	# Missed Cleavages	Theo. MH+ [Da]	Confidence (by Search Engine): Sequest HT	XCorr (by Search Engine): Sequest HT
High	ENVIEAADRLRDEPL		1	1	1	A0A140N8I9	1	1739.89225	High	1.99

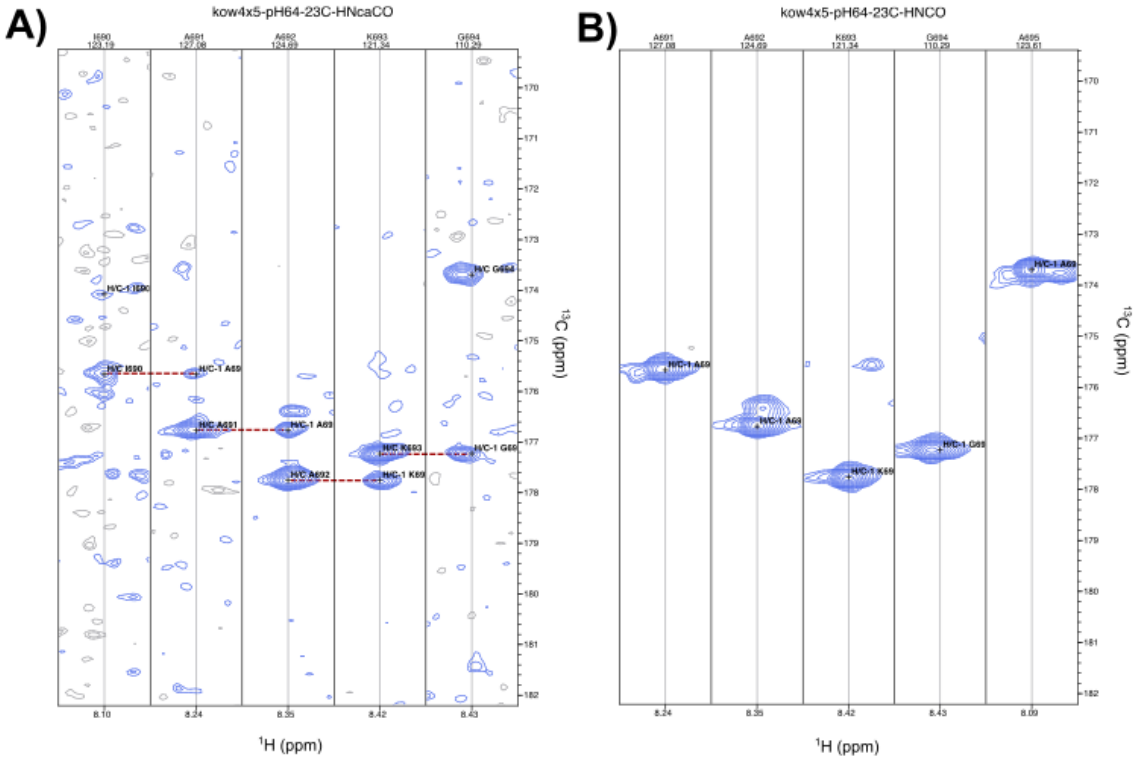
Table 3: Table of Hits from Mass Spectrometry Analysis of In-gel digest of top band of KOWx45 Sample.

Accession	Description	Coverage [%]	# Peptides	# PSMs	# Unique Peptides	# AAs	MW [kDa]	calc. pI	Score Sequest HT: Sequest HT	# Peptides (by Search Engine): Sequest HT
KOW4x5	KOW4x5	5	2	3	2	320	34.5	10.39	7.94	2
Confidence	Sequence	Modifications	# Protein Groups	# Proteins	# PSMs	Master Protein Accessions	# Missed Cleavages	Theo. MH+ [Da]	Confidence (by Search Engine): Sequest HT	XCorr (by Search Engine): Sequest HT
High	VFLHNRDIAENNGVF		1	1	2	KOW4x5	2	1744.87654	High	2.92
High	LHNRDIAENNGVF		1	1	1	KOW4x5	1	1498.73972	High	2.65

KOWx45 Chemical Shift Assignment



Appendix Figure 4.1. Depictions of magnetization transfer between nuclei between different types of triple resonance experiments (diagrams taken from Protein NMR, A Practical Guide).



Appendix Figure 4.2. CARA strip representation of assignment workflow A) Strips taken from the HN(CA)CO spectrum between residues A691-R695 showing links between C'i and C'i-1. B) Strips taken from the HNCO spectrum between residues A691-R695 showing C'i-1 peaks.

Report of KOWx45 Chemical Shift Assignments

Table 4: Chemical shifts of backbone resonance assignments of KOWx45 construct

Residue number	Amino Acid Type	$^{1}H_N$	$^{15}N_H$	$^{13}C_O$	$^{13}C\alpha$	$^{13}C\beta$
563	PRO	-	-	177.023	63.388	32.312
564	MET	8.687	120.579	176.394	55.571	32.312
565	SER	8.272	116.349	174.223	58.351	63.730
566	ASP	8.373	122.547	176.315	54.371	41.034
567	LYS	8.263	121.398	176.708	56.243	32.566
568	SER	8.325	116.594	174.333	58.670	63.730
569	ASP	8.341	122.417	176.461	54.289	41.006
570	LYS	8.293	121.657	177.268	56.635	32.578
571	GLY	8.411	109.286	174.221	45.364	-

572	SER	8.030	115.401	174.445	58.265	63.726
573	VAL	8.108	121.548	175.796	62.152	32.634
574	ALA	8.277	127.591	177.259	52.336	19.197
575	ALA	8.204	123.736	177.636	52.336	19.192
576	VAL	8.060	119.080	175.796	62.115	32.805
577	ASN	8.438	122.302	174.836	53.063	39.133
578	SER	8.178	116.794	173.357	58.262	63.925
579	ALA	8.095	125.383	176.205	52.481	19.616
580	TYR	7.416	117.041	174.128	57.098	39.716
581	GLU	8.920	120.182	175.340	54.045	33.196
582	LEU	8.374	120.177	177.117	57.017	41.831
586	VAL	8.241	121.196	176.223	56.262	32.820
587	GLN	8.258	124.839	174.411	52.481	-
589	ASP	8.090	114.769	176.614	53.360	40.829
590	VAL	8.329	114.603	175.614	65.242	31.577
591	ASN	8.185	115.290	175.500	53.354	41.093
592	THR	7.520	117.195	172.807	63.348	69.730
593	VAL	8.505	123.229	173.515	58.983	34.132
594	ALA	8.747	120.535	173.163	50.542	23.823
595	CYS	8.604	120.248	174.898	56.161	29.815
596	ILE	8.388	128.599	176.224	63.122	-
597	PHE	9.272	126.841	173.965	56.411	41.808
598	SER	7.339	117.095	172.321	58.823	64.393
599	VAL	8.559	127.073	174.400	61.876	33.767
600	ASP	8.079	126.828	175.138	53.146	41.500
601	ARG	8.901	122.485	174.825	59.040	27.560
602	ASP	8.340	119.003	174.228	54.008	40.794
603	THR	7.527	112.041	172.192	61.113	71.425
604	TYR	8.982	121.613	174.681	56.069	40.553
606	VAL	8.500	119.056	174.738	53.245	38.578
607	ILE	7.992	121.776	174.285	59.680	32.566
612	GLY	8.803	111.535	172.640	45.236	-
613	VAL	8.399	123.839	177.060	61.525	33.219
614	ARG	8.856	124.719	174.335	54.263	33.980

615	THR	8.525	118.847	174.738	62.089	69.317
616	VAL	9.349	120.546	173.928	58.648	35.719
617	LEU	8.307	121.098	179.445	53.427	43.829
618	ALA	9.018	122.866	179.255	55.863	18.392
619	SER	7.845	107.528	175.654	59.517	62.401
620	GLN	7.867	119.663	174.887	56.005	29.868
621	ILE	8.391	128.492	177.944	62.072	38.918
622	THR	8.276	114.015	174.782	62.806	69.510
623	MET	7.966	122.544	176.223	56.163	32.871
624	ARG	8.164	120.818	175.591	55.353	31.211
625	HIS	8.416	120.901	176.698	56.989	32.882
626	SER	8.426	116.761	174.303	58.553	63.730
627	ASN	8.266	122.488	176.425	54.286	41.034
628	ARG	8.364	120.094	176.694	56.894	30.687
629	ARG	8.334	119.758	176.299	55.555	29.986
630	GLY	8.003	108.980	171.926	45.935	-
631	VAL	7.917	117.474	174.175	60.310	35.317
632	ALA	8.740	127.635	175.340	50.504	21.822
633	THR	7.730	110.832	174.018	59.825	71.453
634	ASP	8.858	124.139	178.942	51.955	41.354
635	ARG	8.735	116.541	177.060	58.867	30.718
636	ASN	8.212	112.473	175.624	52.481	39.308
637	GLY	8.197	109.298	173.543	45.356	-
638	ALA	8.598	126.544	177.152	51.463	19.153
639	GLU	8.424	120.809	176.708	56.953	30.618
640	ILE	8.626	119.477	173.200	59.534	41.034
641	ARG	8.679	120.251	175.749	53.681	34.513
642	ILE	8.047	118.574	177.684	63.679	36.715
643	GLY	8.875	115.599	174.251	44.930	-
644	ASP	8.132	121.477	174.993	55.047	41.291
645	LYS	8.495	120.247	176.225	55.651	33.710
646	VAL	8.726	114.556	173.059	58.153	36.555
647	LYS	8.765	117.953	176.362	53.826	36.715
648	GLU	8.974	124.464	176.151	57.644	31.465

649	VAL	8.262	122.147	176.312	62.813	31.210
650	GLY	8.217	114.031	173.516	43.979	-
651	GLY	8.385	108.100	176.034	46.645	-
652	GLU	9.175	125.106	177.745	56.626	30.256
653	GLY	8.021	107.947	174.664	47.339	-
654	LYS	9.023	120.665	175.795	57.454	32.831
655	GLN	8.410	118.037	175.642	53.258	32.516
656	GLY	8.796	105.101	-	46.322	-
657	THR	8.810	116.800	175.508	61.788	70.133
658	ILE	9.004	124.607	176.802	63.901	37.688
659	LEU	9.683	130.426	175.733	55.826	45.692
660	HIS	7.295	109.412	173.672	55.319	38.626
661	ILE	8.777	122.282	174.773	61.082	41.034
662	TYR	9.032	128.734	172.728	57.619	41.337
663	ARG	9.038	125.479	175.431	58.303	27.524
664	ALA	7.206	123.205	176.452	52.787	19.084
665	PHE	8.903	118.671	174.866	58.142	41.591
666	VAL	9.072	113.105	173.745	58.827	34.850
667	PHE	8.361	118.962	174.487	57.493	41.204
668	LEU	9.123	125.232	174.390	53.355	44.962
669	HIS	8.529	115.716	175.145	54.369	35.393
670	ASN	9.227	127.909	175.145	51.986	40.286
671	ARG	8.570	123.101	176.500	57.926	29.910
672	ASP	8.492	118.117	176.028	55.124	40.794
673	ILE	7.568	122.533	175.050	61.012	38.922
674	ALA	8.486	130.276	-	54.271	18.963
675	GLU	8.014	119.396	178.064	58.048	29.253
676	ASN	9.126	123.370	173.266	54.217	35.884
677	ASN	8.018	110.295	175.550	53.999	38.220
678	GLY	7.585	107.009	174.187	47.424	-
679	VAL	7.456	120.094	175.900	61.274	32.458
680	PHE	8.915	122.909	171.662	55.972	39.465
681	SER	8.712	114.490	172.835	55.605	66.835
682	ALA	8.623	122.009	176.128	51.179	22.168

683	ARG	8.800	119.127	178.717	55.774	29.799
684	SER	8.902	121.551	175.255	61.643	62.807
685	ARG	7.987	115.144	175.263	56.878	29.323
686	ASN	8.053	115.578	174.575	52.021	39.256
687	VAL	7.579	112.547	172.786	58.344	35.608
688	ALA	8.823	119.600	177.300	50.380	22.237
689	THR	8.978	116.669	174.077	63.246	69.317
690	ILE	8.099	123.186	175.662	61.132	38.438
691	ALA	8.238	127.084	176.771	52.253	19.519
692	ALA	8.349	124.694	177.755	52.142	19.168
693	LYS	8.421	121.339	177.227	56.670	32.789
694	GLY	8.440	110.231	173.695	45.160	-
695	ALA	8.094	123.608	177.627	52.404	19.356
696	ARG	8.330	120.610	176.221	56.002	30.655
697	ILE	8.226	122.506	176.225	60.979	38.629
698	SER	8.354	120.024	174.345	58.248	63.788
699	ALA	8.377	126.368	177.294	52.773	19.173
700	ASP	-	-	176.738	54.289	40.738
701	LEU	8.294	123.301	178.162	55.785	41.676
702	THR	8.153	113.176	174.917	62.933	69.414
703	LYS	7.931	122.215	176.331	56.238	32.690
704	MET	8.101	120.411	175.557	55.364	32.942
705	ASN	8.386	120.854	173.653	51.135	38.764
706	PRO	-	-	176.775	63.619	32.023
707	ALA	8.184	122.308	177.971	52.746	18.757
708	LEU	7.945	119.910	177.488	55.042	42.243
709	SER	8.101	115.666	174.242	58.303	63.738
710	ASN	8.403	120.617	175.200	53.189	39.199
711	GLY	8.127	109.045	171.462	44.692	-
712	PRO	-	-	176.523	62.894	32.131
713	ALA	8.333	124.237	177.377	52.197	18.975
714	LEU	8.193	123.168	174.800	52.746	41.622
716	PRO	-	-	177.022	62.770	31.824
717	VAL	8.124	119.996	176.061	62.333	32.566

718	ALA	8.350	127.351	177.306	52.590	19.173
719	ASN	8.349	117.809	175.151	53.208	38.578
720	LEU	8.135	122.569	177.319	55.426	42.220
721	LYS	8.251	121.783	176.428	56.408	32.566
722	ARG	8.253	122.145	176.310	56.093	31.048
723	THR	8.251	116.243	174.382	61.868	69.768
724	ILE	8.217	123.592	176.512	61.165	38.872
725	GLY	8.447	112.911	173.618	45.074	-
726	ARG	8.146	120.826	175.954	55.856	31.235
727	ASP	8.426	123.420	178.138	54.681	41.616
728	LYS	8.875	128.109	176.663	57.890	32.330
729	ALA	8.798	119.130	177.631	53.429	18.651
730	ILE	7.361	115.684	177.224	64.835	36.689
731	GLY	8.747	116.496	174.033	44.681	-
732	ALA	8.145	123.850	177.287	52.731	18.653
733	THR	8.903	119.296	175.141	63.590	68.947
734	VAL	9.118	122.082	173.517	57.650	36.138
735	ARG	8.336	117.367	176.005	53.279	35.056
736	ILE	8.296	125.664	177.831	63.350	36.184
737	ARG	8.642	129.112	175.041	59.058	32.327
738	ARG	7.730	114.793	175.329	54.800	34.197
739	GLY	8.818	109.677	173.451	44.934	-
740	PRO	-	-	178.238	64.843	32.300
741	MET	8.915	116.040	173.314	54.760	32.163
742	LYS	7.377	118.180	176.961	58.122	32.324
743	GLY	8.907	113.049	174.071	44.845	-
744	LEU	7.794	121.118	175.770	55.778	41.611
745	LEU	7.877	119.907	177.883	53.436	45.432
746	GLY	8.725	108.374	170.195	46.168	-
747	VAL	8.154	118.417	176.144	59.679	35.873
748	ILE	8.527	125.461	176.686	63.749	36.967
749	LYS	9.344	130.168	175.589	55.229	32.859
750	ASP	7.913	114.877	173.944	53.202	42.992
751	THR	8.597	110.222	-	60.079	72.236

752	THR	8.586	111.956	174.580	60.384	71.145
753	ASP	8.760	121.100	176.868	57.806	39.960
754	ALA	8.059	118.562	177.729	52.724	20.863
755	ASN	7.812	116.037	172.663	52.338	42.435
756	ALA	9.078	121.132	174.673	50.152	22.220
757	ARG	8.643	122.432	174.580	55.141	29.868
758	VAL	9.385	128.163	173.757	60.936	34.247
759	GLU	8.895	127.611	176.777	54.755	31.521
760	LEU	9.189	130.484	177.416	55.232	41.627
761	HIS	8.661	121.357	177.222	58.121	30.415
762	THR	8.524	113.373	174.934	62.193	68.844
763	GLY	8.561	114.675	173.824	45.694	-
764	ASN	7.735	115.237	174.029	54.581	37.582
765	LYS	7.115	117.723	175.036	56.161	34.588
766	MET	8.459	122.915	176.682	53.431	32.319
767	VAL	9.112	119.479	174.124	59.526	35.616
768	THR	8.602	121.168	173.762	62.035	69.059
769	ILE	9.174	128.695	172.909	56.081	42.451
770	PRO	-	-	178.334	63.435	32.291
771	LYS	8.593	124.681	179.336	60.154	31.960
772	GLU	8.540	114.678	176.959	57.651	29.249
773	ASN	8.156	119.126	172.667	51.948	37.785
774	LEU	6.921	118.144	175.043	53.235	44.184
775	LEU	8.663	120.177	175.544	53.074	44.997
776	TYR	9.688	121.309	175.680	57.879	40.059
777	THR	7.998	116.944	176.053	62.958	69.456
778	THR	8.953	117.722	177.374	60.473	71.638
780	THR	7.988	113.545	175.214	62.273	69.378
781	GLY	8.368	110.858	174.215	45.382	-
782	GLU	7.630	120.853	174.834	56.403	30.431
783	LEU	8.390	123.313	177.221	53.981	43.524
784	ILE	9.087	117.234	175.140	59.296	41.362
785	SER	-	-	174.112	57.803	65.435
786	TYR	9.467	120.864	176.870	60.784	38.354

787	THR	8.372	112.223	176.612	66.548	68.699
788	GLU	7.545	121.830	177.635	58.983	29.870
789	PHE	8.822	122.292	178.388	61.952	38.870
790	ILE	8.292	120.113	178.863	62.656	36.417
791	GLU	7.731	121.059	179.383	59.368	29.070
792	ARG	8.314	117.602	178.593	58.418	30.143
793	SER	7.693	114.329	174.695	59.844	63.298
794	ARG	7.535	120.858	176.866	57.030	30.662
795	GLY	7.951	107.909	173.719	45.388	-
796	ILE	7.833	119.829	176.002	60.854	38.355
797	ARG	8.398	126.361	174.127	53.753	30.164
798	PRO	-	-	177.426	63.510	32.071
799	GLY	8.560	109.789	174.277	45.310	-
800	SER	-	-	174.391	58.328	63.763
801	ILE	8.184	121.779	176.205	61.092	38.882
802	SER	8.434	119.730	174.836	58.124	63.816
803	THR	8.209	115.919	174.427	61.709	69.584
804	ALA	8.275	125.739	177.259	52.699	19.192
805	ASP	8.194	119.138	176.193	54.276	41.373
806	GLY	7.989	108.756	171.867	44.769	-
807	PRO	-	-	176.720	63.170	32.063
808	ASN	8.458	118.565	174.684	53.208	38.637
809	VAL	7.920	121.529	174.238	59.789	32.334
810	PRO	-	-	176.617	63.052	31.833
811	ASN	8.426	118.563	175.343	53.519	38.396
812	TRP	7.827	120.358	176.209	57.335	29.309
813	ALA	7.887	124.799	177.424	52.572	18.674
814	GLN	7.920	118.730	176.560	56.247	29.057
815	GLY	8.276	109.670	173.766	45.247	-
816	ALA	8.055	123.488	177.630	52.415	19.319
817	ARG	8.285	119.997	176.209	55.862	30.702
818	THR	8.150	117.792	172.707	59.825	69.495
819	PRO	-	-	176.477	63.133	32.194
820	ALA	8.318	124.272	177.746	52.481	19.065

821	VAL	-	-	175.938	62.013	32.904
822	ALA	8.387	127.659	177.474	52.518	19.241
823	ASN	8.405	118.247	175.843	53.354	38.578
824	GLY	8.238	110.171	174.223	45.392	-
825	SER	8.106	115.382	174.371	58.352	63.924
826	ARG	8.331	122.583	176.158	55.899	30.702
827	THR	8.143	117.936	172.712	59.789	69.573
828	PRO	-	-	176.567	63.206	31.803
829	ALA	8.251	123.630	177.509	52.445	18.893
830	TRP	7.903	119.337	175.953	57.171	29.348
831	ASN	8.099	120.173	175.214	52.954	38.578
832	THR	8.027	113.602	176.078	62.152	69.489
833	GLY	8.366	110.914	174.207	45.392	-
834	SER	8.121	115.622	174.584	58.371	63.815
835	ARG	8.313	122.648	-	55.935	30.702
836	THR	8.132	117.727	172.760	59.898	69.489
837	PRO	-	-	175.752	62.806	34.266
838	ALA	8.505	123.976	177.672	52.736	19.222
839	TRP	8.111	120.024	176.032	57.265	29.263
840	ASN	8.070	120.601	174.905	52.845	38.802
841	SER	8.018	115.835	174.993	58.895	63.487
842	GLY	8.360	110.656	174.229	45.395	-
843	SER	8.168	115.456	174.537	58.432	63.830
844	LYS	8.332	122.854	176.409	56.154	32.862
845	THR	8.114	117.597	172.744	59.841	69.635
846	PRO	-	-	176.544	63.202	31.921
847	ALA	8.264	123.639	177.684	52.484	18.944
848	TRP	8.014	120.116	175.953	57.498	29.365
849	ASN	8.015	120.423	174.931	52.809	38.629
850	SER	8.043	115.978	175.072	58.880	63.489
851	GLY	8.354	110.591	174.223	45.387	-
852	SER	8.136	115.454	174.465	58.348	63.769
853	ARG	8.313	122.582	176.132	55.971	30.699
854	THR	8.095	117.491	-	59.825	69.503

855	PRO	-	-	176.495	63.242	32.904
856	ALA	8.242	123.540	177.509	52.494	19.209
857	TRP	8.005	119.972	175.953	57.535	29.348
858	ASN	8.011	120.601	175.198	52.772	38.494
859	SER	8.043	116.032	-	58.843	63.476
860	GLY	8.317	110.234	173.923	45.392	-
861	ASN	8.127	118.459	175.140	53.195	38.865
862	LYS	8.264	121.376	176.404	56.262	32.566
863	THR	8.119	117.671	172.760	59.898	69.489
864	PRO	-	-	176.559	63.198	31.824
865	ALA	8.270	123.736	177.690	52.496	18.943
866	TRP	8.000	119.884	175.953	57.505	29.263
867	ASN	7.962	120.609	174.586	52.699	38.614
868	ALA	7.957	124.086	178.140	53.038	18.934
869	GLY	8.270	107.436	174.279	45.312	-
870	SER	8.047	115.471	174.531	58.349	63.848
871	ARG	8.324	122.641	176.106	55.950	30.719
872	THR	8.180	118.070	172.700	59.846	69.489
873	PRO	-	-	175.591	63.448	31.800
874	ALA	7.955	130.173	169.533	53.764	20.211

References

Amberg, D. C., Goldstein, A. L., & Cole, C. N. (1992). Isolation and characterization of RAT1: an essential gene of Saccharomyces cerevisiae required for the efficient nucleocytoplasmic trafficking of mRNA. *Genes & Development, 6*(7), 1173–1189. 10.1101/gad.6.7.1173

Aoi, Y., Shah, A. P., Ganesan, S., Soliman, S. H. A., Cho, B., Goo, Y. A., Kelleher, N. L., & Shilatifard, A. (2022). SPT6 functions in transcriptional pause/release via PAF1C recruitment. *Molecular Cell, 82*(18), 3412–3423.e5. 10.1016/j.molcel.2022.06.037

Aoi, Y., Takahashi, Y., Shah, A. P., Iwanaszko, M., Rendleman, E. J., Khan, N. H., Cho, B., Goo, Y. A., Ganesan, S., Kelleher, N. L., & Shilatifard, A. (2021). SPT5 stabilization of promoter-proximal RNA polymerase II. *Molecular Cell, 81*(21), 4413–4424.e5. 10.1016/j.molcel.2021.08.006

Baejen, C., Andreani, J., Torkler, P., Battaglia, S., Schwalb, B., Lidschreiber, M., Maier, K. C., Boltendahl, A., Rus, P., Esslinger, S., Söding, J., & Cramer, P. (2017). Genome-wide Analysis of RNA Polymerase II Termination at Protein-Coding Genes. *Molecular Cell, 66*(1), 38–49.e6. 10.1016/j.molcel.2017.02.009

Bernecky, C., Herzog, F., Baumeister, W., Plitzko, J. M., & Cramer, P. (2016). Structure of transcribing mammalian RNA polymerase II. *Nature, 529*(7587), 551–554. 10.1038/nature16482

Birse, C. E., Minvielle-Sebastia, L., Lee, B. A., Keller, W., & Proudfoot, N. J. (1998). Coupling Termination of Transcription to Messenger RNA Maturation in Yeast. *Science, 280*(5361), 298–301. 10.1126/science.280.5361.298

Bradner, J. E., Hnisz, D., & Young, R. A. (2017). Transcriptional Addiction in Cancer. *Cell, 168*(4), 629–643. 10.1016/j.cell.2016.12.013

Buratowski, S. (2009). Progression through the RNA Polymerase II CTD Cycle. *Molecular Cell, 36*(4), 541–546. 10.1016/j.molcel.2009.10.019

Chang, J. H., Xiang, S., Xiang, K., Manley, J. L., & Tong, L. (2011). Structural and biochemical studies of the 5′→3′ exoribonuclease Xrn1. *Nature Structural & Molecular Biology, 18*(3), 270–276. 10.1038/nsmb.1984

Chen, F. X., Smith, E. R., & Shilatifard, A. (2018). Born to run: control of transcription elongation by RNA polymerase II. *Nature Reviews Molecular Cell Biology, 19*(7), 464–478. 10.1038/s41580-018-0010-5

Chernyakov, I., Whipple, J. M., Kotelawala, L., Grayhack, E. J., & Phizicky, E. M. (2008). Degradation of several hypomodified mature tRNA species in Saccharomyces cerevisiae is mediated by Met22 and the 5′–3′ exonucleases Rat1 and Xrn1. *Genes & Development, 22*(10), 1369–1380. 10.1101/gad.1654308

Churchman, L. S., & Weissman, J. S. (2011). Nascent transcript sequencing visualizes transcription at nucleotide resolution. *Nature (London), 469*(7330), 368–373. 10.1038/nature09652

Cortazar, M. A., Sheridan, R. M., Erickson, B., Fong, N., Glover-Cutter, K., Brannan, K., & Bentley, D. L. (2019). Control of RNA Pol II Speed by PNUTS-PP1 and Spt5 Dephosphorylation Facilitates Termination by a "Sitting Duck Torpedo" Mechanism. *Molecular Cell, 76*(6), 896–908.e4. 10.1016/j.molcel.2019.09.031

Dangkulwanich, M., Ishibashi, T., Liu, S., Kireeva, M. L., Lubkowska, L., Kashlev, M., & Bustamante, C. J. (2013). Complete dissection of transcription elongation reveals slow translocation of RNA polymerase II in a linear ratchet mechanism. *eLife, 2*, e00971. 10.7554/eLife.00971

Decker, T. (2021). Mechanisms of Transcription Elongation Factor DSIF (Spt4–Spt5). *Journal of Molecular Biology, 433*(14), 166657. 10.1016/j.jmb.2020.09.016

Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., & Bax, A. (1995). NMRPipe: A multidimensional spectral processing system based on UNIX pipes. *Journal of Biomolecular NMR, 6*(3), 277–293. 10.1007/BF00197809

Dikunova, A., Noskova, N., Overbeck, J. H., Polak, M., Stelzig, D., Zapletal, D., Kubicek, K., Novacek, J., Sprangers, R., & Stefl, R. (2025). Assembly of the Xrn2/Rat1–Rai1–Rtt103 termination complexes in mesophilic and thermophilic organisms. *Structure (London), 33*(2), 300–310.e6. 10.1016/j.str.2024.11.010

Dollinger, R., Deng, E. B., Schultz, J., Wu, S., Deorio, H. R., & Gilmour, D. S. (2023). Assessment of the roles of Spt5-nucleic acid contacts in promoter proximal pausing of RNA polymerase II. *The Journal of Biological Chemistry, 299*(9), 105106. 10.1016/j.jbc.2023.105106

Duval, M., Yague-Sanz, C., Turowski, T. W., Petfalski, E., Tollervey, D., & Bachand, F. (2023). The conserved RNA-binding protein Seb1 promotes cotranscriptional ribosomal RNA processing by controlling RNA polymerase I progression. *Nature Communications, 14*(1), 3013. 10.1038/s41467-023-38826-6

Eaton, J. D., Davidson, L., Bauer, D. L. V., Natsume, T., Kanemaki, M. T., & West, S. (2018). Xrn2 accelerates termination by RNA polymerase II, which is underpinned by CPSF73 activity. *Genes & Development, 32*(2), 127–139. 10.1101/gad.308528.117

Eaton, J. D., Francis, L., Davidson, L., & West, S. (2020). A unified allosteric/torpedo mechanism for transcriptional termination on human protein-coding genes. *Genes & Development, 34*(1), 132–145. 10.1101/gad.332833.119

Ehara, H., Yokoyama, T., Shigematsu, H., Yokoyama, S., Shirouzu, M., & Sekine, S. (2017a). Structure of the complete elongation complex of RNA polymerase II with basal factors. *Science (American Association for the Advancement of Science), 357*(6354), 921–924. 10.1126/science.aan8552

Ehara, H., Yokoyama, T., Shigematsu, H., Yokoyama, S., Shirouzu, M., & Sekine, S. (2017b). Structure of the complete elongation complex of RNA polymerase II with basal factors. *Science (American Association for the Advancement of Science), 357*(6354), 921–924. 10.1126/science.aan8552

Eick, D., & Geyer, M. (2013). The RNA Polymerase II Carboxy-Terminal Domain (CTD) Code. *Chemical Reviews, 113*(11), 8456–8490. 10.1021/cr400071f

Fong, N., Brannan, K., Erickson, B., Kim, H., Cortazar, M., Sheridan, R., Nguyen, T., Karp, S., & Bentley, D. (2015). Effects of Transcription Elongation Rate and Xrn2 Exonuclease Activity on RNA Polymerase II Termination Suggest Widespread Kinetic Competition. *Molecular Cell, 60*(2), 256–267. 10.1016/j.molcel.2015.09.026

Girbig, M., Misiaszek, A. D., & Müller, C. W. (2022). Structural insights into nuclear transcription by eukaryotic DNA-dependent RNA polymerases. *Nature Reviews. Molecular Cell Biology, 23*(9), 603–622. 10.1038/s41580-022-00476-9

Glover-Cutter, K., Kim, S., Espinosa, J., & Bentley, D. L. (2008). RNA polymerase II pauses and associates with pre-mRNA processing factors at both ends of genes. *Nature Structural & Molecular Biology, 15*(1), 71–78. 10.1038/nsmb1352

Grünberg, S., & Hahn, S. (2013). Structural insights into transcription initiation by RNA polymerase II. *Trends in Biochemical Sciences (Amsterdam. Regular Ed.), 38*(12), 603–611. 10.1016/j.tibs.2013.09.002

Han, Z., Moore, G. A., Mitter, R., Lopez Martinez, D., Wan, L., Dirac Svejstrup, A. B., Rueda, D. S., & Svejstrup, J. Q. (2023). DNA-directed termination of RNA polymerase II transcription. *Molecular Cell, 83*(18), 3253–3267.e7. 10.1016/j.molcel.2023.08.007

Harris, J. K., Kelley, S. T., Spiegelman, G. B., & Pace, N. R. (2003). The Genetic Core of the Universal Ancestor. *Genome Research, 13*(3), 407–412. 10.1101/gr.652803

Hartzog, G. A., & Fu, J. (2013). The Spt4–Spt5 complex: A multi-faceted regulator of transcription elongation. *Biochimica Et Biophysica Acta (BBA) - Gene Regulatory Mechanisms, 1829*(1), 105–115. 10.1016/j.bbagrm.2012.08.007

Ho, Vasilieva. (2025) unpublished manuscript

Hu, S., Peng, L., Xu, C., Wang, Z., Song, A., & Chen, F. X. (2021a). SPT5 stabilizes RNA polymerase II, orchestrates transcription cycles, and maintains the enhancer landscape. *Molecular Cell, 81*(21), 4425–4439.e6. 10.1016/j.molcel.2021.08.029

Hu, S., Peng, L., Xu, C., Wang, Z., Song, A., & Chen, F. X. (2021b). SPT5 stabilizes RNA polymerase II, orchestrates transcription cycles, and maintains the enhancer landscape. *Molecular Cell, 81*(21), 4425–4439.e6. 10.1016/j.molcel.2021.08.029

Izban, M. G., & Luse, D. S. (1992). Factor-stimulated RNA polymerase II transcribes at physiological elongation rates on naked DNA but very poorly on chromatin templates. *Journal of Biological Chemistry, 267*(19), 13647–13655. 10.1016/s0021-9258(18)42262-4

Jasnovidova, O., & Stefl, R. (2013). The CTD code of RNA polymerase II: a structural view. *WIREs RNA, 4*(1), 1–16. 10.1002/wrna.1138

Jin, W., Qazi, T. J., Quan, Z., Li, N., & Qing, H. (2019). *Dysregulation of Transcription Factors: A Key Culprit Behind Neurodegenerative Disorders*. SAGE Publications. 10.1177/1073858418811787

Jones, S., Daley, D. T., Luscombe, N. M., Berman, H. M., & Thornton, J. M. (2001). Protein-RNA interactions: a structural analysis. *Nucleic Acids Research, 29*(4), 943–954. 10.1093/nar/29.4.943

Kecman, T., Kuś, K., Heo, D., Duckett, K., Birot, A., Liberatori, S., Mohammed, S., Geis-Asteggiante, L., Robinson, C. V., & Vasiljeva, L. (2018). Elongation/Termination Factor Exchange Mediated by PP1 Phosphatase Orchestrates Transcription Termination. *Cell Reports, 25*(1), 259–269.e5. 10.1016/j.celrep.2018.09.007

Keller, R. (1996). *The Computer Aided Resonance Assignment Tutorial*. Institute for Molecular Biology and Biophysics, Swiss Federal Institute of Technology.

Kelley, J. R., Dimitrova, E., Maciuszek, M., Nguyen, H. T., Szczurek, A. T., Hughes, A. L., Blackledge, N. P., Kettenbach, A. N., & Klose, R. J. (2024). The PNUTS phosphatase complex controls transcription pause release. *Molecular Cell, 84*(24), 4843–4861.e8. 10.1016/j.molcel.2024.10.045

Kim, M., Krogan, N. J., Vasiljeva, L., Rando, O. J., Nedea, E., Greenblatt, J. F., & Buratowski, S. (2004). yeast Rat1 exonuclease promotes transcription termination by RNA polymerase II. [Erratum: 2005 Feb. 10, v. 433, no. 7026, p. 661.]. *Nature, 432*(7016), 517–522. 10.1038/nature03041

Komarnitsky, P., Cho, E. J., & Buratowski, S. (2000). Different phosphorylated forms of RNA polymerase II and associated mRNA processing factors during transcription. *Genes & Development, 14*(19), 2452–2460. 10.1101/gad.824700

Kuś, K., Carrique, L., Kecman, T., Fournier, M., Hassanein, S. S., Aydin, E., Kilchert, C., Grimes, J. M., & Vasiljeva, L. (2025). DSIF factor Spt5 coordinates transcription, maturation and exoribonucleolysis of RNA polymerase II transcripts. *Nature Communications, 16*(1), 10–19. 10.1038/s41467-024-55063-7

Kwak, H., & Lis, J. T. (2013). Control of Transcriptional Elongation. *Annual Review of Genetics, 47*(1), 483–508. 10.1146/annurev-genet-110711-155440

Lasko, P. (2010). Tudor Domain. *Current Biology, 20*(16), R666–R667. https://www.cell.com/current-biology/fulltext/S0960-9822(10)00696-2

Lehr, M., Paschelke, T., Trumpf, E., Vogt, A., Näther, C., Sönnichsen, F. D., & McConnell, A. J. (2020). A Paramagnetic NMR Spectroscopy Toolbox for the Characterisation of Paramagnetic/Spin-Crossover Coordination Complexes and Metal–Organic Cages. *Angewandte Chemie International Edition, 59*(43), 19344–19351. 10.1002/anie.202008439

Lindstrom, D. L., Squazzo, S. L., Muster, N., Burckin, T. A., Wachter, K. C., Emigh, C. A., McCleery, J. A., Yates, J. R., & Hartzog, G. A. (2003). Dual Roles for Spt5 in Pre-mRNA Processing and Transcription Elongation Revealed by Identification of Spt5-Associated Proteins. *Molecular and Cellular Biology, 23*(4), 1368–1378. 10.1128/MCB.23.4.1368-1378.2003

Mason, P. B., & Struhl, K. (2005). Distinction and Relationship between Elongation Rate and Processivity of RNA Polymerase II In Vivo. *Molecular Cell, 17*(6), 831–840. 10.1016/j.molcel.2005.02.017

Mayer, A., Lidschreiber, M., Siebert, M., Leike, K., Söding, J., & Cramer, P. (2010). Uniform transitions of the general RNA polymerase II transcription complex. *Nature Structural & Molecular Biology, 17*(10), 1272–1278. 10.1038/nsmb.1903

Meyer, P. A., Li, S., Zhang, M., Yamada, K., Takagi, Y., Hartzog, G. A., & Fu, J. (2015). Structures and Functions of the Multiple KOW Domains of Transcription Elongation Factor Spt5. *Molecular and Cellular Biology, 35*(19), 3354–3369. 10.1128/MCB.00520-15

Noe Gonzalez, M., Blears, D., & Svejstrup, J. Q. (2021). Causes and consequences of RNA polymerase II stalling during transcript elongation. *Nature Reviews. Molecular Cell Biology, 22*(1), 3–21. 10.1038/s41580-020-00308-8

Nojima, T., Gomes, T., Grosso, A. R., Kimura, H., Dye, M., Dhir, S., Carmo-Fonseca, M., & Proudfoot, N. (2015). Mammalian NET-Seq Reveals Genome-wide Nascent Transcription Coupled to RNA Processing. *Cell, 161*(3), 526–540. 10.1016/j.cell.2015.03.027

Overbeck, J. H., Stelzig, D., Fuchs, A., Wurm, J. P., & Sprangers, R. (2022). Observation of conformational changes that underlie the catalytic cycle of Xrn2. *Nature Chemical Biology, 18*(10), 1152–1160. 10.1038/s41589-022-01111-6

Pei, Y., & Shuman, S. (2003a). Characterization of the Schizosaccharomyces pombe Cdk9/Pch1 protein kinase: Spt5 phosphorylation, autophosphorylation, and mutational analysis. *The Journal of Biological Chemistry, 278*(44), 43346–43356. 10.1074/jbc.M307319200

Pei, Y., & Shuman, S. (2003b). Characterization of the Schizosaccharomyces pombe Cdk9/Pch1 protein kinase: Spt5 phosphorylation, autophosphorylation, and mutational analysis. *The Journal of Biological Chemistry, 278*(44), 43346–43356. 10.1074/jbc.M307319200

Peterlin, B. M., & Price, D. H. (2006). Controlling the Elongation Phase of Transcription with P-TEFb. *Molecular Cell, 23*(3), 297–305. 10.1016/j.molcel.2006.06.014

Ponting, C. P. (1997). Tudor domains in proteins that interact with RNA. *Trends in Biochemical Sciences (Amsterdam. Regular Ed.), 22*(2), 51–52. 10.1016/S0968-0004(96)30049-2

Higman, V. 2019, April-last update*, Protein NMR, a practical guide*. Available: https://protein-nmr.org.uk/.

Proudfoot, N. (2004). New perspectives on connecting messenger RNA 3′ end formation to transcription. *Current Opinion in Cell Biology, 16*(3), 272–278. 10.1016/j.ceb.2004.03.007

Qiu, Y., & Gilmour, D. S. (2017). Identification of Regions in the Spt5 Subunit of DRB Sensitivity-inducing Factor (DSIF) That Are Involved in Promoter-proximal Pausing. *The Journal of Biological Chemistry, 292*(13), 5555–5570. 10.1074/jbc.M116.760751

Rasmussen, E. B., & Lis, J. T. (1993). In vivo transcriptional pausing and cap formation on three Drosophila heat shock genes. *Proceedings of the National Academy of Sciences, 90*(17), 7923–7927. 10.1073/pnas.90.17.7923

Rengachari, S., Hainthaler, T., Oberthuer, C., Lidschreiber, M., & Cramer, P. (2025). Mechanism of polyadenylation-independent RNA polymerase II termination. *Nature Structural & Molecular Biology, 32*(2), 339–345. 10.1038/s41594-024-01409-0

Sainsbury, S., Bernecky, C., & Cramer, P. (2015). Structural basis of transcription initiation by RNA polymerase II. *Nature Reviews Molecular Cell Biology, 16*(3), 129–143. 10.1038/nrm3952

Sansó, M., Levin, R. S., Lipp, J. J., Wang, V. Y., Greifenberg, A. K., Quezada, E. M., Ali, A., Ghosh, A., Larochelle, S., Rana, T. M., Geyer, M., Tong, L., Shokat, K. M., & Fisher, R. P. (2016). P-TEFb regulation of transcription termination factor Xrn2 revealed by a chemical genetic screen for Cdk9 substrates. *Genes & Development, 30*(1), 117–131. 10.1101/gad.269589.115

Schreieck, A., Easter, A. D., Etzold, S., Wiederhold, K., Lidschreiber, M., Cramer, P., & Passmore, L. A. (2014). RNA polymerase II termination involves C-terminal-domain tyrosine dephosphorylation by CPF subunit Glc7. *Nature Structural & Molecular Biology, 21*(2), 175–179. 10.1038/nsmb.2753

Schweikhard, V., Meng, C., Murakami, K., Kaplan, C. D., Kornberg, R. D., & Block, S. M. (2014). Transcription factors TFIIF and TFIIS promote transcript elongation by RNA polymerase II by synergistic and independent mechanisms. *Proceedings of the National Academy of Sciences, 111*(18), 6642–6647. 10.1073/pnas.1405181111

Shen, Y. & Bax, A. 2013, "Protein backbone and sidechain torsion angles predicted from NMR chemical shifts using artificial neural networks", *Journal of Biomolecular NMR,* vol. 56, no. 3, pp. 227−241.

Shetty, A., Kallgren, S. P., Demel, C., Maier, K. C., Spatt, D., Alver, B. H., Cramer, P., Park, P. J., & Winston, F. (2017). Spt5 Plays Vital Roles in the Control of Sense and Antisense Transcription Elongation. *Molecular Cell, 66*(1), 77–88.e5. 10.1016/j.molcel.2017.02.023

Solyom, Z., Schwarten, M., Geist, L., Konrat, R., Willbold, D., & Brutscher, B. (2013). BEST-TROSY experiments for time-efficient sequential resonance assignment of large

disordered proteins. *Journal of Biomolecular NMR, 55*(4), 311–321. 10.1007/s10858-013-9715-0

Ungerleider, N. A., Roberts, C., O'Grady, T. M., Nguyen, T. T., Baddoo, M., Wang, J., Ishaq, E., Concha, M., Lam, M., Bass, J., Nguyen, T. D., Van Otterloo, N., Wickramarachchige-Dona, N., Wyczechowska, D., Morales, M., Ma, T., Dong, Y., & Flemington, E. K. (2024). Viral reprogramming of host transcription initiation. *Nucleic Acids Research, 52*(9), 5016–5032. 10.1093/nar/gkae175

Vos, S. M., Farnung, L., Boehning, M., Wigge, C., Linden, A., Urlaub, H., & Cramer, P. (2018a). Structure of activated transcription complex Pol II–DSIF–PAF–SPT6. *Nature, 560*(7720), 607–612. 10.1038/s41586-018-0440-4

Vos, S. M., Farnung, L., Urlaub, H., & Cramer, P. (2018b). Structure of paused transcription complex Pol II–DSIF–NELF. *Nature (London), 560*(7720), 601–606. 10.1038/s41586-018-0442-2

Washington, K., Ammosova, T., Beullens, M., Jerebtsova, M., Kumar, A., Bollen, M., & Nekhai, S. (2002). Protein Phosphatase-1 Dephosphorylates the C-terminal Domain of RNA Polymerase-II. *Journal of Biological Chemistry, 277*(43), 40442–40448. 10.1074/jbc.M205687200

Wen, Y., & Shatkin, A. J. (1999). Transcription elongation factor hSPT5 stimulates mRNA capping. *Genes & Development, 13*(14), 1774–1779. 10.1101/gad.13.14.1774

West, S., Gromak, N., & Proudfoot, N. J. (2004). Human 5′ → 3′ exonuclease Xrn2 promotes transcription termination at co-transcriptional cleavage sites. *Nature, 432*(7016), 522–525. 10.1038/nature03035

Williamson, M. P. (2013). Using chemical shift perturbation to characterise ligand binding. *Progress in Nuclear Magnetic Resonance Spectroscopy, 73*, 1–16. 10.1016/j.pnmrs.2013.02.001

Xiang, S., Cooper-Morgan, A., Jiao, X., Kiledjian, M., Manley, J. L., & Tong, L. (2009). Structure and function of the 5′→3′ exoribonuclease Rat1 and its activating partner Rai1. *Nature, 458*(7239), 784–788. 10.1038/nature07731

Yamada, T., Yamaguchi, Y., Inukai, N., Okamoto, S., Mura, T., & Handa, H. (2006). P-TEFb-Mediated Phosphorylation of hSpt5 C-Terminal Repeats Is Critical for Processive Transcription Elongation. *Molecular Cell, 21*(2), 227–237. 10.1016/j.molcel.2005.11.024

Yanagisawa, T., Murayama, Y., Ehara, H., Goto, M., Aoki, M., & Sekine, S. (2024). Structural basis of eukaryotic transcription termination by the Rat1 exonuclease complex. *Nature Communications, 15*(1), 7854–12. 10.1038/s41467-024-52157-0

Zamft, B., Bintu, L., Ishibashi, T., & Bustamante, C. (2012). Nascent RNA structure modulates the transcriptional dynamics of RNA polymerases. *Proceedings of the National Academy of Sciences - PNAS, 109*(23), 8948–8953. 10.1073/pnas.1205063109

Zeng, Y., Zhang, H., Wu, X., & Zhang, Y. (2024). Structural basis of exoribonuclease-mediated mRNA transcription termination. *Nature (London), 628*(8009), 887–893. 10.1038/s41586-024-07240-3

Zhang, Z., & Gilmour, D. S. (2006). Pcf11 Is a Termination Factor in Drosophila that Dismantles the Elongation Complex by Bridging the CTD of RNA Polymerase II to the Nascent Transcript. *Molecular Cell, 21*(1), 65–74. 10.1016/j.molcel.2005.11.002

Zhou, K., Kuo, W. H. W., Fillingham, J., & Greenblatt, J. F. (2009). Control of transcriptional elongation and cotranscriptional histone modification by the yeast BUR kinase substrate Spt5. *Proceedings of the National Academy of Sciences, 106*(17), 6956–6961. 10.1073/pnas.0806302106

Zuber, P. K., Hahn, L., Reinl, A., Schweimer, K., Knauer, S. H., Gottesman, M. E., Rösch, P., & Wöhrl, B. M. (2018). Structure and nucleic acid binding properties of KOW domains 4 and 6–7 of human transcription elongation factor DSIF. *Scientific Reports, 8*(1), 11660–12. 10.1038/s41598-018-30042-3

Zuber, P. K., Daviter, T., Heißmann, R., Persau, U., Schweimer, K., & Knauer, S. H. (2022). Structural and thermodynamic analyses of the β-to-α transformation in RfaH reveal principles of fold-switching proteins. *eLife, 11*10.7554/eLife.76630