An investigation into the structural role of the CCR4-NOT complex in mRNA stability

---

Hannah Brazier

Wolfson College

A thesis submitted for the degree of Doctor of Philosophy in Molecular and Cellular Medicine at the Kennedy Institute of Rheumatology

Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal sciences

University of Oxford

Supervisors: Dr Nicola Burgess-Brown, Dr Wyatt Yue

Trinity term 2017

Declaration

I declare that no parts of this research or thesis have been submitted for another award, degree or diploma at any other university or learning institution. This thesis contains no other person's work, except where stated in text.

Acknowledgements

I would like to thank a number of people for their input in allowing the work described here to be realised. Firstly, I would like to thank my supervisors Dr Nicola Burgess-Brown, Dr Wyatt Yue and Dr Jonathan Dean who have offered invaluable guidance and advice throughout this project. For the time and energy they bestowed upon me, I am truly grateful.

Great thanks must be attributed to the Kennedy Institute of Rheumatology Trust, which provided funding to allow the project to proceed, and to Wolfson College for the support of academic staff and the great atmosphere of the college community.

I would also like to thank my friends and colleagues, past and present, at the SGC but particularly those in the Biotechnology and MOB group who proved to be a wealth of knowledge and acted as a sounding board for my ideas. Specifically, I would like to thank Dr Claire Strain-Damerell for her patience and endless cloning advice and Dr Nicholas Fox and Dr Caroline Sanvitale-Rudolf for their light-hearted jokes which helped me face late nights and failed experiments.

My final thanks must be dedicated to my parents who have always supported me throughout my studies and without whom this would not be possible.

Abstract

The CCR4-NOT complex is a global regulator of gene expression which controls mRNA levels by removing the poly-(A) tail, a step known as deadenylation and one that constitutes the rate-limiting step in mRNA decay. The human complex is comprised of eight stably associated CNOT subunits where CNOT1 forms the scaffold onto which CNOT2-11 bind. Although much has been learnt about the CCR4-NOT complex, questions still remain. Thus, this study focused on a number of sub-complexes of CNOT subunits and associated proteins to determine the mode of interaction with a hope to explore the mechanism of deadenylation by the CCR4-NOT complex. Firstly, the complex of CNOT10:CNOT11, found only in higher eukaryotes, was reconstituted for the first time using recombinant proteins. Crystallisation trials, limited proteolysis and mass spectrometry were used to isolate novel interaction regions between CNOT10 and CNOT11 which may provide direction for future structural and functional studies. Secondly, the interaction between CNOT9 and TTP was characterised. TTP is a RNA-binding protein which targets inflammatory mRNAs for deadenylation by recruiting the CCR4-NOT complex. This study highlights novel interactions between TTP and both CNOT2 and CNOT9. Moreover, BioLayer interferometry (BLI), peptide arrays and site-directed mutagenesis identified that TTP interacts with CNOT9 in a tryptophan-mediated manner. These findings change the known interface between TTP and the CCR4-NOT complex. Lastly, the MultiBac system was used to reconstitute a number of human CNOT sub-complexes, one of which was shown to effectively degrade a poly-(A) substrate, demonstrating it is enzymatically active. This achievement provides a tool for the future study of the CCR4-NOT complex. In summary, this study highlights novel interactions and characterises previously unknown binding mechanisms between CCR4-NOT subunits which expands our current understanding of the complex.

Contents

List of Tables

List of Figures

Abbreviations

3′ UTR	3′ untranslated region
aa	amino acid
AREs	AU-rich elements
BLI	BioLayer interferometry
CAF1	carbon catabolite repressor protein (CCR) 4-CCR4-associated factor 1
CCR4	carbon catabolite repressor protein 4
cDNA	complementary DNA
CDE	Constitutive decay element
CNOT	CCR4-negative on TATA-less
cryo-EM	Cryo-electron microscopy
CV	column volume
DNA	deoxyribonucleic acid
DSF	differential scanning fluorimetry
DTT	Dithiothreitol
DUF	Domain of unknown function
E. coli	*Escherichia coli*
EMSA	Electrophoretic mobility shift assay
EDTA	Ethylenediaminetetraacetic acid
ExoT	Exonuclease T
FBS	fetal bovine serum
Flc	fluorescein
GST	Glutathione S-transferase
HCS	Hampton crystal screen

HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulphonic acid
HIN	Hampton index screen
HPLC	high performance liquid chromatography
HRP	horseradish peroxidase
HTP	high-throughput
ITC	Isothermal titration calorimetry
JCSG	Joint Center for Structural Genomics
LB	Luria broth
LC-QTOF	liquid chromatography-Quadrupole time-of-flight
LIC	ligation independent cloning
LPS	lipopolysaccharide
µl	microlitre
MAPK	Mitogen-activated protein kinase
MBP	maltose binding protein
min	minutes
miRNA	micro ribonucleic acid
MK2	MAPK kinase 2
ml	millilitre
mM	millimolar
mRNA	messenger ribonucleic acid
MS	mass spectrometry
MWCO	molecular weight cut off
NaCl	sodium chloride
NEB	New England Biolabs
nl	nanolitre

NOT	negative on TATA-less
PCR	polymerase chain reaction
poly-(A)	poly-adenosine
PUF	Pumilio family
QTOF	Quadrupole time-of-flight
RA	Rheumatoid arthritis
RISC	RNA-induced silencing complex
RNA	ribonucleic acid
RNBP	RNA-binding protein
rpm	revolutions per minute
S. cerevisiae	*Saccharomyces cerevisiae*
SDS-PAGE	sodium dodecyl sulphate polyacrylamide gel electrophoresis
sec	seconds
Sf9	*Spodoptera frugiperda*
SFM	serum-free medium
TAE	Tris base, acetic acid and EDTA
tandem MS	tryptic digest mass spectrometry
TB	terrific broth
Tb	*Trypanosoma brucei*
TEV	Tobacco Etch virus
TNF-α	Tumour necrosis factor-α
TTP	Tristetraprolin
Trx	Thioredoxin
Y2H	Yeast-2-hybrid

1. Introduction

1.1 Inflammation

Inflammation is the innate immune system's normal physiological response to tissue injury and infection. This process isolates the damaged area and recruits immune cells to promote the healing process. The innate immune cells (e.g. macrophages, fibroblasts) recognise the invasion of pathogens or tissue injury using pattern recognition receptors (PRR) on the cell surface which detect pathogen-associated molecular patterns (PAMPs) or damage-associated molecular patterns (DAMPs) (e.g. nucleic acids, LPS, carbohydrates) (Tang et al., 2012). Once the PRRs detect the PAMP/DAMPs, downstream proteins are signalled to activate the major inflammatory pathways to regulate transcription of inflammatory genes (e.g. chemokine (C-X-C motif) ligand (CXCL)-1, CXCL-2, interleukin (IL)-1α) (Newton and Dixit, 2012).

1.2 Control of inflammatory gene expression

Controlling inflammatory gene expression is critical to allow genes to be rapidly switched on/off. This is imperative for proteins that require a burst of gene expression for their cellular function, such as tumour necrosis factor-α (TNF-α), a pro-inflammatory cytokine. Inflammatory gene expression should peak to protect the affected tissue and then swiftly return to baseline levels to prevent further damage (Garneau et al., 2007). If left uncontrolled, aberrant expression may lead to chronic inflammatory diseases e.g. rheumatoid arthritis. Therefore, a number of elements have evolved to carefully control the half-life of mRNA and thus the corresponding protein expression levels. These elements include the 3′-poly-adenosine (poly-(A)) tail, 5′-methylguanosine cap (5′ cap) and cis-acting elements in the 3′ untranslated region (3′ UTR).

1.2.1 Elements affecting mRNA half-life

During transcription, two stabilising modifications are added to mRNAs; the poly-(A) tail and the 5′cap (**Figure 1.1**) (Garneau et al., 2007). The function of these structures is three-fold. Firstly the poly-(A) tail is bound by poly-(A) binding proteins (PABP) which protect the mRNA from degradation by nuclease (Garneau et al., 2007). Secondly, the 5′ cap and the poly-(A) tail are required for efficient translation; the 5′ cap is an essential recognition site for the binding of eukaryotic initiation factors needed for translational initiation, and the poly-(A) tail is a docking site for poly-(A) binding proteins, known to interact with the ribosome and facilitate translation (Kahvejian et al., 2005). Thirdly, these elements allow mRNA circularisation as the PABPs bound to the poly-(A) tail can associate with eukaryotic initiation factors bound to the 5′ cap (Gallie, 1998, Kahvejian et al., 2005). This circularisation, which has been visualised by atomic force microscopy, is believed to promote translation efficiency (Gallie, 1998, Labno et al., 2016, Wells et al., 1998).

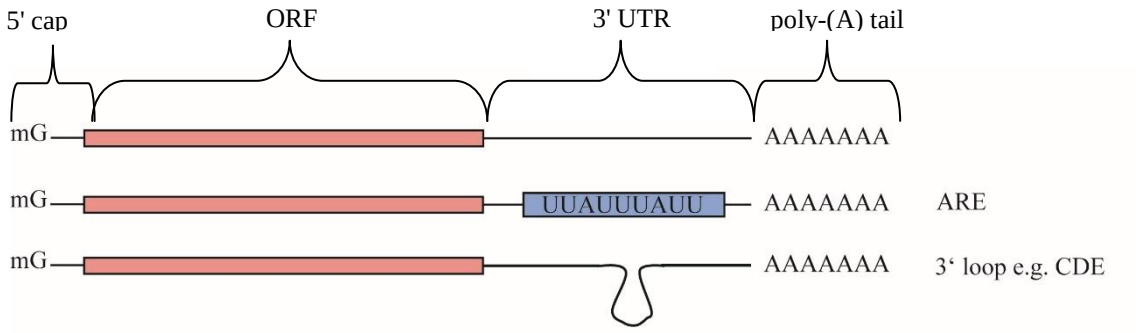


Figure 1.1 mRNA stability and decay structures

A number of cis-acting mRNA decay elements are located in the 3’ UTR of mRNAs to control the half-life of these mRNAs and thus the protein expression levels. Two examples of cis-acting elements are shown: AU-rich elements (AREs) and 3’ loops e.g. constitutive decay elements (CDEs). Both are found in TNF-α and cause the recruitment of RNBPs which tether the mRNA to deadenylases. The deadenylases remove the poly-(A) tail of mRNA which allows subsequent decay of the mRNA body.

Further to the 5′ cap and poly-(A) tail, there are a number of cis-acting stability elements located in the 3′ UTR of mRNA. Certain stem loop structures in the 3′ UTR can act as

stability regulating elements. One such example is the constitutive decay element (CDE) found in TNF-α mRNA. The 23-mer loop is recognised by RNA-binding proteins (RNBPs) including the Roquin family, which then recruit proteins to degrade the mRNA, described in section 1.2.3 (Sakurai et al., 2015).

AU-rich elements (AREs) are another cis-acting stability structure in the 3′ UTR (**Figure 1.1**). AREs frequently contain a nonameric sequence (UUAUUUAUU) and are common in short-lived mRNAs to control stability and expression (Zubiaga et al., 1995). ARE-containing mRNAs include cytokines, growth factors, proto-oncogenes and other transiently expressed proteins (e.g. TNF-α, interleukin-2, interleukin-10, interleukin-6, granulocyte-macrophage colony-stimulating factor, keratinocyte-derived chemokine and cyclooxygenase-2) (Zhao et al., 2011, Stoecklin et al., 2008, Chen and Shyu, 1995). As with CDEs, AREs are recognised by RNBPs, for example the Tis11 family, which then recruit proteins to deadenylate the mRNA.

1.2.2 Deadenylation

Deadenylation, the process of removing the mRNA poly-(A) tail, has many functions throughout the life cycle of mRNA. Newly transcribed mRNAs are polyadenylated extensively (~250 nucleotides in mammals). Hence, deadenylases are first required in the cytoplasm to trim the long poly-(A) tail of nascent mRNAs to their proper length (Wiederhold and Passmore, 2010, Goldstrohm and Wickens, 2008). The steady state length of the poly-(A) tail, typically thought to be 50-100 nucleotides in mammalian cells, is dictated by the cellular fate of the mRNA (Chang et al., 2014). During early development, mRNAs containing longer poly-(A) tails are translated more effectively (Richter, 1999). Deadenylation is later required as the first obligatory step of mRNA degradation. Decay can only occur once the poly-(A)-tail is removed or if the mRNA is

internally cleaved (**Figure 1.2**) (Garneau et al., 2007). However, deadenylation is reversible and some mRNAs contain a signal to be re-polyadenylated and then can continue to be translated (Huarte et al., 1992).

The process of deadenylation is catalysed by specific nucleases known as deadenylases. Deadenylases use Mg^{2+} ions bound in the active site to cleave the phosphodiester bond between nucleotides to remove the poly-(A) tail of mRNA. There are 10 known deadenylases encoded by the human genome, including isoforms CNOT6/CNOT6L (CCR4 in yeast), isoforms CNOT7/CNOT8 (CAF1 in yeast) (which are members of the CCR4-NOT complex that will be later discussed in depth), Angel, PAN2:PAN3, PARN, PNLDC1, Nocturnin and PDE12 (Goldstrohm and Wickens, 2008, Skeparnias et al., 2017). Deadenylases belong to two main groups; the exonuclease-endonuclease-phosphatase (EEP) class and DEDD class, which are characterised by the residues in the active site that chelate magnesium ions (Asp-Glu-Asp-Asp). CNOT7 and PARN are examples of the DEDD family, and CNOT6 and Angel are EEP family members. Further information on the specific deadenylases is provided in **Table 1.1**.

In mammalian cells, a complex of PAN2 and PAN3, where PAN2 contains deadenylase activity and PAN3 confers RNA-binding, mediates the initial shortening of the poly-(A) tail which is trimmed to ~25 nucleotides (Wolf et al., 2014). Beyond this, PAN2:PAN3 (where ':' indicates complex formation) activity is limited due to lack of PABPs which the complex requires for docking. Subsequently, the CCR4-NOT complex, comprising of CNOT6/6L and CNOT7/8 deadenylases, is recruited to remove the remainder of the poly-(A) tail (Garneau et al., 2007). How these transitions are coordinated and whether deadenylases respond to different cellular signalling events remain important unanswered questions. Certain deadenylases can form sub-complexes e.g. CNOT6/6L and CNOT7/8 within the CCR4-NOT complex (Petit et al., 2012). Complex formation may act to

Name	DEDD or EEP	Function	Reference
CNOT6/CNOT6L (CCR4 in yeast)	EEP	Associates with CNOT7/CNOT8 as part of the CCR4-NOT complex	(Petit et al., 2012)
CNOT7/CNOT8 (CAF1 in yeast)	DEDD	Associates with CNOT6/CNOT6L as part of the CCR4-NOT complex	(Petit et al., 2012)
Angel	EEP	Thought to possess deadenylase activity due to similarity to CCR4 but not confirmed in vitro	(Godwin et al., 2013)
Pan2	DEDD	Associates with Pan3 to remove the poly-(A) tail, leaving ~25 nucleotides	(Wolf et al., 2014)
PARN	DEDD	Deadenylates mRNA, small rRNA and telomerase RNA	(Skeparnias et al., 2017)
PNLDC1	DEDD	Putative deadenylase that may be involved in pre-piRNA trimming activity in mammals	(Izumi et al., 2016)
PDE12	EEP	Involved in the interferon response	(Kubota et al., 2004)
Nocturnin	EEP	Circadian deadenylase where expression peaks at night, for reasons as yet unknown	(Godwin et al., 2013)

Table 1.1 Human deadenylases

There are 10 deadenylases encoded by the human genome. Deadenylases are characterised into two classes; the exonuclease-endonuclease-phosphatase (EEP) and DEDD class, which are characterised by the residues in the active site that chelate magnesium ions (Asp-Glu-Asp-Asp). The varied functions of the deadenylases and a reference to provide more information is listed.

additionally modulate the activity of the deadenylases. Although CNOT6/CNOT6L and CNOT7/CNOT8 are active as monomers, it has been shown they have increased activity as part of the deadenylase module within the CCR4-NOT complex (Stowell et al., 2016), but the cellular function for this remains unclear.

1.2.3 mRNA degradation

After the poly-(A) tail has been removed by deadenylases (e.g. PAN2-PAN3 or the CCR4-NOT complex), the mRNA body can be degraded. Degradation is an important process to prevent aberrant translation of mRNAs which would lead to uncontrolled expression of proteins affecting the normal cell function. Degradation occurs by two pathways (**Figure 1.2**). The pathway on the right of **Figure 1.2a** depicts degradation of the mRNA by the exosome, a complex of 3′-5′ exonucleases (Mitchell et al., 1997, Chekanova et al., 2000, Brouwer et al., 2001, Estevez et al., 2001). The exonucleases from the exosome degrade the mRNA body in the 3′-5′ direction to which they now have

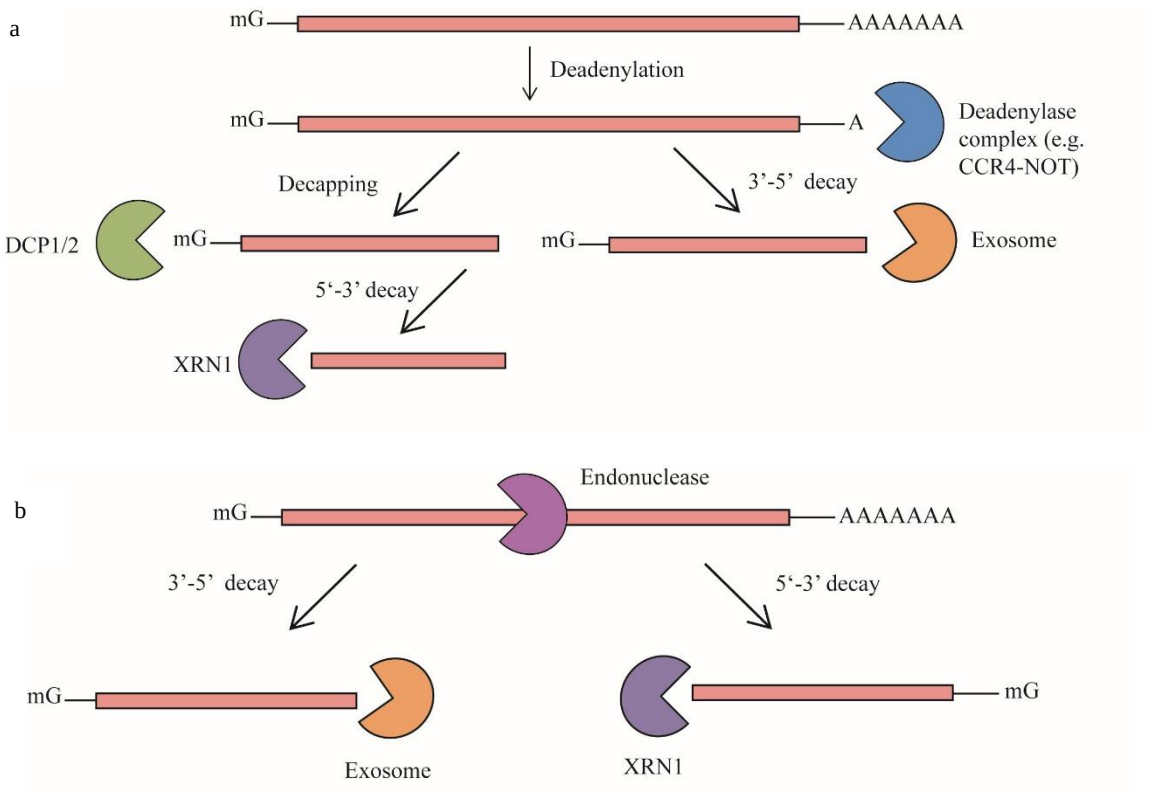


Figure 1.2 mRNA decay pathways

Schematic depicting the pathways in mRNA decay. (a) Firstly, the poly-(A) tail is removed by a deadenylase e.g. the CCR4-NOT complex. The next step is either 5′-3′ decay (left) or 3′-5′ decay (right). (b) The pathway of degradation when the mRNA is internally cleaved by an endonuclease. The cleaved 5′ is attacked by nucleases in the exosome and degraded 3′-5′. The cut 5' end is degraded by XRN1 in a 5′-3′ direction.

access to the 3′ end of the mRNA body because the poly-(A) tail has been removed. The left of **Figure 1.2a** shows the mRNA degradation pathway that is initiated by the removal of the mRNA 5′ cap by a DCP1/2 complex, where DCP2 provides the catalytic domain and DCP1 contains the regulatory domain (Mugridge et al., 2016). The removal of the 5′ cap prevents translational initiation. After decapping, the mRNA body is degraded in the 5′-3′ direction by the XRN1 exoribonuclease.

An alternate mRNA degradation pathway is shown in **Figure 1.2b**. This process is initiated by internal cleavage of the mRNA body by an endonuclease (e.g. Ago2, PMR1, IRE1 (Garneau et al., 2007, Chernokalskaya et al., 1998)). Internal cleavage produces two unprotected ends which are subsequently attacked by both the exosome and XRN1 which effectively degrades the mRNA from the incision site (**Figure 1.2**) (Labno et al., 2016). One mechanism which drives internal cleavage of mRNAs is the targeting by microRNAs (miRNA). miRNAs are small RNA molecules of approximately ~25 nucleotides which are bound by Agonaute proteins (Ago), forming the RNA-induced silencing complex (RISC). The RISC complex is targeted to mRNAs that are complementary to the loaded miRNA, which then binds to the target, and the RISC complex internally cleaves the mRNA within the region of the bound miRNA (Tolia and Joshua-Tor, 2007). Humans have four Ago proteins but only Ago2 possesses the nuclease activity required to cleave mRNA. The internal cleavage triggers mRNA degradation via the exosome and XRN1, as in **Figure 1.3a** (Iwakawa and Tomari, 2015). Alternatively, the RISC complex can recruit the CCR4-NOT complex, a multi-subunit deadenylase, via interactions with GW182 (**Figure 1.3b)**. Thus, the deadenylase complex is targetted in an miRNA-mediated manner to remove the poly-(A) tail from the mRNA, which is the first step of mRNA degradation; after which, the mRNA body can be degraded.

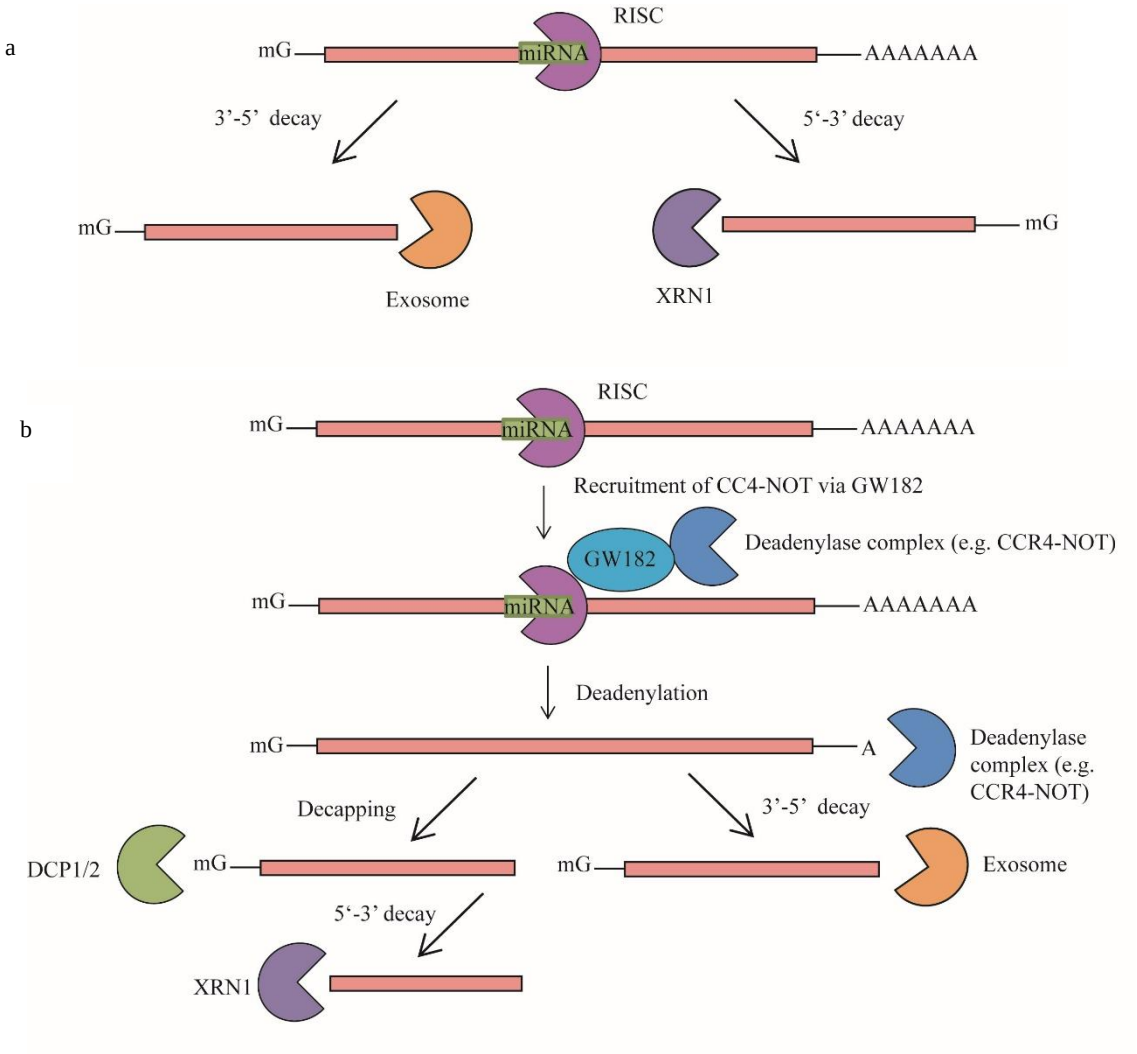


Figure 1.3 miRNA-mediated mRNA decay pathways

Schematic depicting the pathways of miRNA-mediated mRNA decay. (a) The miRNA is bound by Agonaute proteins which together form the RISC complex. The miRNA targets the RISC complex to complementary mRNA and the RISC can internally cleave the mRNA at the region where the miRNA has bound. After cleavage, the unprotected 3′ and 5′ ends are attacked by XRN1 and the nucleases from the exosome. (b) Additionally, once the RISC complex has been targeted to mRNA, it can recruit the CCR4-NOT deadenylase complex via an interaction with GW182. The CCR4-NOT complex removes the poly-(A) tail and this allows for subsequent decapping by DCP1/2 and degradation by XRN1 and nucleases in the exosome.

1.3 The CCR4-NOT complex

The CCR4-NOT complex is an evolutionarily conserved multi-protein complex discovered in yeast thirty years ago by means of a genetic screen (Reed, 1980). The CCR4-NOT complex was named after the catalytic subunits, carbon catabolite repressor protein (CCR) 4-CCR4-associated factor (CAF) 1 complex (CCR4 and CAF1), and the structural subunits, CCR4-negative on TATA-less (NOT). The CCR4-NOT complex is a large (>1 MDa) assembly comprising eight stably associated subunits in human; CNOT1-11. The human CCR4-NOT complex contains two mutually exclusive isoforms of the two deadenylase subunits (CNOT6L or CNOT6 and CNOT7 or CNOT8) (**Figure 1.4, Table 1.2**) (Lui, 1998, Mauxion et al., 2013). Beyond the stable subunits, a plethora of additional factors transiently interact with the human CCR-NOT complex, including CNOT4. The CNOT subunits in the human CCR4-NOT are bound to the core CNOT1 subunit in sub-complexes or modules; CNOT2 and CNOT3, CNOT10 and CNOT11 and CNOT6/6L and CNOT7/8 interact with another subunit as well as docking onto CNOT1.

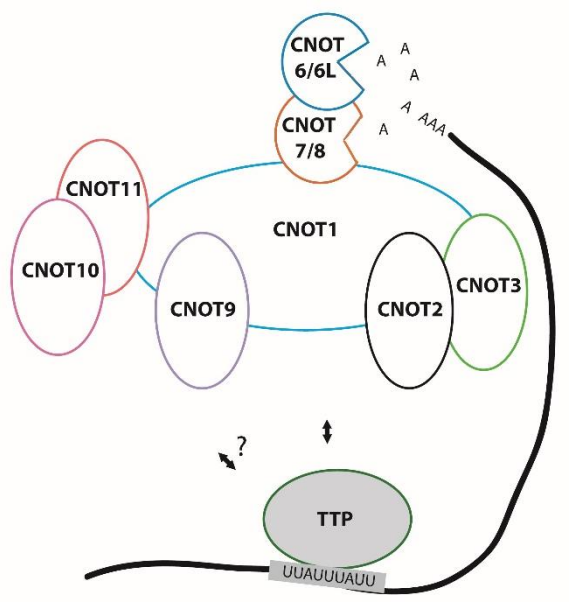


Figure 1.4 The CCR4-NOT complex

Schematic of the arrangement of the subunits within the human CCR4-NOT complex. CNOT1 is the main scaffold, onto which docks CNOT2-11. TTP, bound to mRNA via an AU-rich element (ARE) (UUAUUUAUU), is also associated with the complex via the middle region of CNOT1, indicated by double-headed arrow. However, it remains to be determined if further interactions between TTP and the CCR4-NOT complex are required for recruitment, indicated by the question mark.

Human nomenclature	Function	Yeast nomenclature	Reference
CNOT1	Scaffold	NOT1	(Boland et al., 2013, Chen et al., 2014, Mauxion et al., 2013)
CNOT2	Complex stability	NOT2	(Boland et al., 2013)
CNOT3	Complex stability, stem cell differentiation	NOT3/NOT5	(Boland et al., 2013)
CNOT4	E3 ligase	NOT4	(Albert et al., 2002, Lau et al., 2009)
CNOT6/6L	EEP deadenylase	CCR4	(Lau et al., 2009)
CNOT7/8	DEDD deadenylase	CAF1	(Lau et al., 2009)
CNOT9	Interactions with GW182 to recruit CCR4-NOT to RISC complex for miRNA mediated decay	CAF40	(Chen et al., 2014, Mathys et al., 2014, Garces et al., 2007)
CNOT10	Function unknown		(Mauxion et al., 2013)
CNOT11	Function unknown		(Mauxion et al., 2013)

Table 1.2 The CCR4-NOT complex subunits

A table detailing the CCR4-NOT subunits with the names of human and yeast homologs. The currently understood function of each subunit is listed with a reference for further information. CNOT1 forms the main scaffold of the complex with CNOT2 and CNOT3 aiding the stability of the complex. CNOT6/6L and CNOT7/8 are deadenylases which can remove the poly-(A) tail of mRNAs. CNOT9 has been shown to interact with GW182 to facilitate miRNA-mediated mRNA decay. CNOT10 is thought to be recruited to the complex via an interaction between CNOT11 and CNOT1, but beyond this their role is unclear.

The CCR4-NOT complex was originally annotated as a transcriptional inhibitor evidenced by sensitivity of CCR4-NOT mutant strains to transcription elongation inhibitors. Since its discovery, CCR4-NOT has been found to both activate and repress transcription (Reese, 2013). Beyond a role in transcription, characterisation of the two

deadenylase subunits of the complex, CCR4 and CAF1, suggested a role in regulating glucose repressible alcohol dehydrogenase 2 in yeast (Malvar et al., 1992). Further functional analysis of CCR4/CAF1 subunits revealed a role in the deadenylation of mRNA, described above. The CCR4-NOT complex has a varied and extensive role within cells that involves transcription, decapping, miRNA-mediated degradation, protein-mediated degradation, among many new functions slowly emerging (i.e. stem cell differentiation (Zheng et al., 2012)). Therefore, the function of the CCR4-NOT is both broad and convoluted, deeming it a master regulator of gene expression.

1.3.1 Composition of the CCR4-NOT complex

The largest subunit of the complex, NOT1, is the only essential protein in yeast but deletion of the other subunits leads to a slow growth phenotype (Maillet et al., 2000). Pull-down studies identified CNOT1 as the main scaffold onto which the other subunits dock (Maillet et al., 2000, Ukleja et al., 2016b, Boland et al., 2013, Chen et al., 2014). The interactions between CNOT subunits and transient binding partners will be described in terms of the three regions of CNOT1: an N-terminal domain, a middle region and a C-terminal domain (**Figure 1.5**).

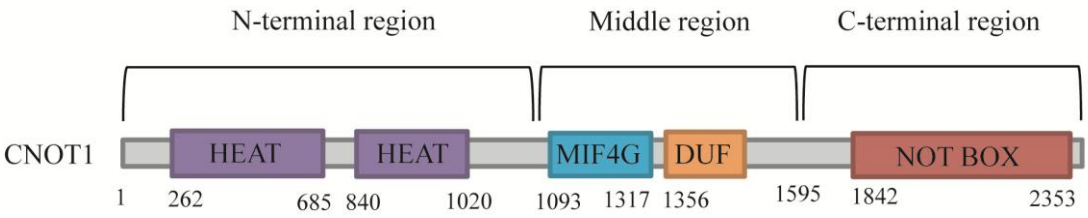


Figure 1.5 Schematic of the domains within CNOT1

Schematic of the domain arrangement of CNOT1. CNOT1 is described in three regions; the N-terminal region, the middle region and the C-terminal region. The N-terminal region contains HEAT repeat domains which associate with a dimer of CNOT10 and CNOT11. These HEAT repeats are also used by TTP (described below) to recruit the CCR4-NOT complex for targeted deadenylation of mRNAs. The middle region contains a MIF4G domain which binds the deadenylase subunits, CNOT6/CNOT6L and CNOT7/CNOT8. The middle region also contains a domain of unknown function (DUF) which has been shown to interact with CNOT9. The final region is the C-terminus which contains a NOT box that binds CNOT2 and CNOT3 via their NOT box domains.

1.3.2 Interactions mediated by the N-terminal region of CNOT1

The N-terminal region (aa 1-1092, **Figure 1.5**) of CNOT1 contains HEAT repeats which typically consist of ~50 residues folded into two anti-parallel helices. These motifs are named after just four of the diverse number of proteins which contain these repeats (Huntingtin, elongation factor 3 (EF3), protein phosphatase 2A (PP2A), and the yeast kinase TOR1) and often provide a platform for protein-protein interactions (Andrade et al., 2001). The first HEAT repeat region (aa 1-551) has been shown, via a pull-down assay, to directly interact with Dead End 1 (DND1) (Yamaji et al., 2017). DND1 is an RNBP which affects the mRNA levels of genes associated with regulating pluripotent stem cells and cancer. It is hypothesised that DND1 binds target mRNA and recruits the CCR4-NOT complex via CNOT1 to remove the poly-(A) tail of the mRNA and target it for subsequent decay. Further work is required to understand the nature and purpose of this interaction.

The extreme N-terminus of CNOT1 (aa 1-240) has also been shown to bind a module of CNOT10 and CNOT11 where, although CNOT10 is essential in *Trypanosoma brucei (T. brucei),* very little additional information is known about these proteins (Farber et al., 2013). This module is present is the human complex but a similar module has yet to be detected in yeast, suggesting the function may be mammalian specific. Results obtained from a yeast-2-hydrid (Y2H) screen suggest that CNOT10 interacts with CNOT1 via CNOT11 (Mauxion et al., 2013). Beyond this, the motifs required for binding and the function of this module remain to be determined. Structural predictions using psiPRED indicate CNOT10 contains regions which are intrinsically disordered but may have a function in protein binding. CNOT11 is also predicted to contain intrinsically disordered regions but a domain of unknown function is also predicted towards the C-terminus. Further research into this module may provide insight into whether CNOT10 and

CNOT11 confer any additional functions or regulatory mechanisms that are not present in yeast.

The HEAT repeats of CNOT1 (aa 800-1015) also boast an interaction with a C-terminal peptide of Tristetraprolin (TTP), an RNBP, as depicted in a co-crystal structure (PDB ID:4J8S) (**Figure 1.6**) (Fabian et al., 2013). The TTP peptide appears to lie in a small cavity on the CNOT1 surface. TTP binds inflammatory mRNAs via a central zinc-finger domain and recruits them to the CCR4-NOT complex via this interaction with CNOT1. This targets the inflammatory mRNA for deadenylation, by the CNOT6/CNOT6L and CNOT7/CNOT8 subunits, and subsequent decay, as in **Figure 1.2**. However, the interaction between CNOT1 and TTP, mediated by hydrogen bonds and salt bridges, is weak (2 μM as measured by isothermal titration calorimetry (ITC) (Fabian et al., 2013)) and one may hypothesise that further subunits of the CCR4-NOT complex may be involved in recruiting TTP. The individually weak interactions may collectively increase the affinity by avidity in order to draw together the TTP, mRNA and the catalytic subunits of the CCR4-NOT.

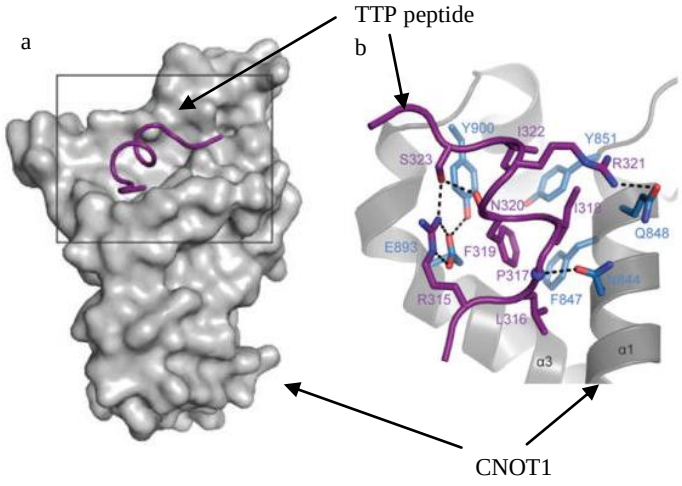


Figure 1.6 Structure of CNOT1 in complex with a TTP peptide adapted from Fabian et al (PDB ID:4J8S) (Fabian et al., 2013))

(a) The HEAT repeat of CNOT1 (aa 800-1015) contains a hydrophobic groove which binds a C-terminal 10 residue peptide of TTP. (b) TTP binds CNOT1 using a number of electrostatic interactions and most prominently a salt bridge between Glu893 of CNOT1 and Arg315 of TTP. The TTP peptide was found due to its high sequence conservation across different organisms and the CNOT1 domain boundaries were established by pull-downs using different truncations of CNOT1.

1.3.3 Interactions mediated by the middle region of CNOT1

The middle domain of CNOT1 (aa 1093-1594) contains a MIF4G-like domain and DUF3819 domain (**Figure 1.5**). The DUF3819 domain is known to bind CNOT9 and will now be referred to as the CNOT9-binding domain (Chen et al., 2014, Mathys et al., 2014). CNOT9 is a 33 kDa protein which is present as a stable component of the CCR4-NOT complex but also exists in cells as a free homodimer (Garces et al., 2007). A structure of the complex of CNOT1:CNOT9 shows CNOT1 binding to the convex side of CNOT9 (PDB ID:4CRU) (Mathys et al., 2014, Chen et al., 2014). CNOT9 contains a central armadillo domain, composed of 17 helices which fold to form a positively charged cleft along the concave side of CNOT9 (**Figure 1.7**, grey). The cleft has been shown to bind nucleic acids, though a preference for poly-(A) nucleic acids was not detected (Garces et al., 2007). Thus, the CNOT9 cleft may function to secure the mRNA

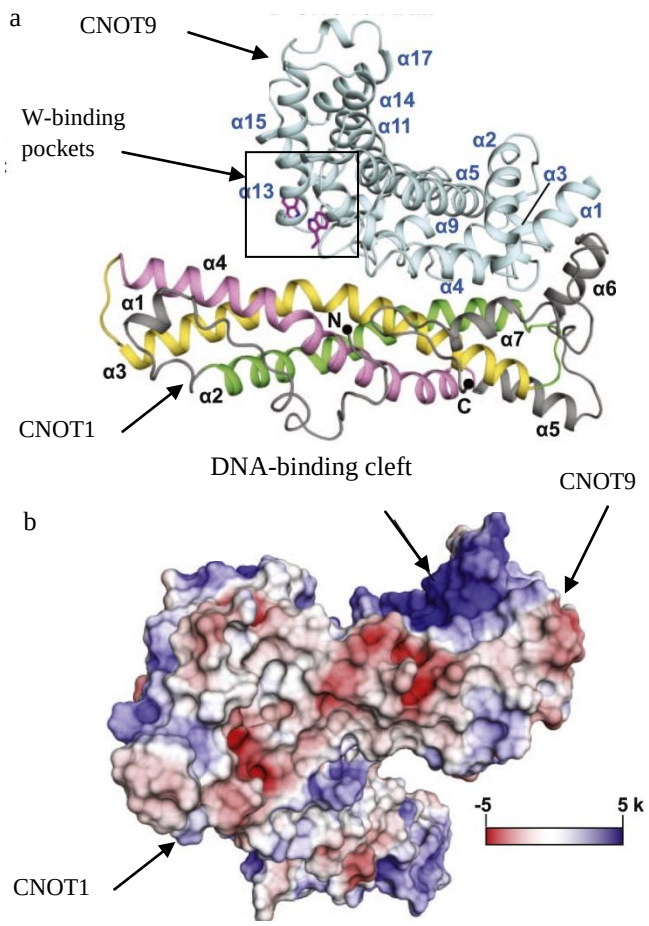


Figure 1.7 Structure of the heterodimer of CNOT1 and CNOT9 adapted from Chen et al (PDB ID:4CRU) (Chen et al., 2014)

Structure depicts the CNOT9-binding domain of CNOT1 located in the middle region. (a) This domain of CNOT1 (shown by pink, yellow and green helices) binds to the convex side of CNOT9 armadillo domain (shown in grey). The armadillo domain of CNOT9 is composed of 17 helices which fold into a crescent shape. (b) The concave side of CNOT9 is positively charged (blue) and has been shown to bind nucleic acids, though specificity for poly-(A) has not been proven.

body in close proximity to the catalytic subunits. The CNOT9-binding domain of CNOT1 (**Figure 1.7a**, pink, yellow and green helices) is a helical region which binds the convex surface of CNOT9. The interface between the CNOT9 dimer and the heterodimer of CNOT1:CNOT9 encompasses the same region but is larger between CNOT1:CNOT9, thus the heterodimer is hypothesised to be the favourable conformation. The crystal structure of the heterodimer also reveals two pockets in CNOT9 that bind tryptophan residues (**Figure 1.8,** boxed) (PDB ID:4CT7) (Chen et al., 2014, Mathys et al., 2014). CNOT9 utilises these motifs to interact with a number of binding partners, e.g. the RISC complex component GW182 (Pfaff et al., 2013, Chen et al., 2014). Thus, CNOT9 bridges the gap between the mRNA targeted by the RISC complex and the deadenylase subunits of the CCR4-NOT complex which remove the poly-(A) tail of the targeted mRNA and initiate mRNA decay. CNOT9 is also known to interact with Grb10-interacting GYF protein-2 (GIGYF2) using tryptophan residues and a role for CNOT9:GIGYF2 interaction has been suggested during epidermal growth factor receptor (EGFR)

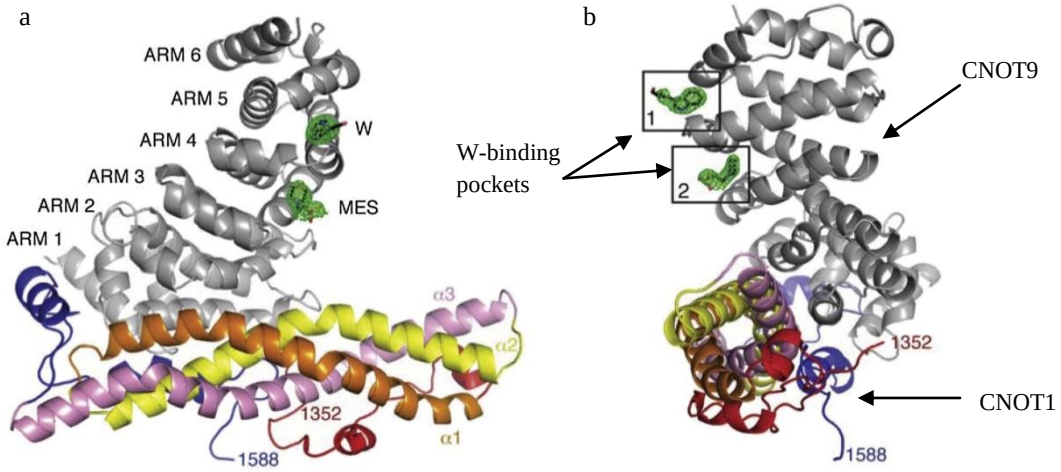


Figure 1.8 Crystal structure of the heterodimer of CNOT1 bound to CNOT9 adapted from Chen et al (PDB ID:4CT7) (Chen et al., 2014)

(a) The crystal structure models the helices which associate to form the CNOT9-binding domain of CNOT1 (coloured). This CNOT1 domain is bound by the armadillo domain of CNOT9 which is composed of 6 arm repeats folded into a crescent shape. (b) The tryptophan/MES which occupy tryptophan binding pockets within CNOT9 have been boxed to highlight their location. Tryptophan-binding pockets are situated on the convex surface of CNOT9 and thus do not conflict with the binding of CNOT1 or the nucleic acid binding groove, which is located on the concave surface of CNOT9. Thus, CNOT9 may simultaneously bind CNOT1, tryptophan-mediated binding partners and nucleic acids.

signalling in cervical and breast cancer (Ajiro et al., 2009). Therefore, tryptophan-mediated interactions may be a conserved mechanism for binding CNOT9.

The MIF4G domain of CNOT1 (aa 1093-1317), located in the middle region, binds the deadenylase module of the complex, where the interaction between yeast CNOT7 and CNOT1 is shown in **Figure 1.9a** (PDB ID:4GMJ) (Petit et al., 2012). The crystal structure shows the MIF4G domain of CNOT1 (lilac) which is formed by 10 helices. CNOT1 binds CNOT7 at an interaction surface opposite to the active site of CNOT7, where magnesium ions (green) are bound. To complete the nuclease module, CNOT6 binds CNOT7 via a domain containing five leucine rich repeats. This orientates the nuclease domain of CNOT6 opposite CNOT1, leave both active sites accessible to potential substrates (**Figure 1.9b**) (PDB ID:4B8C) (Basquin et al., 2012).

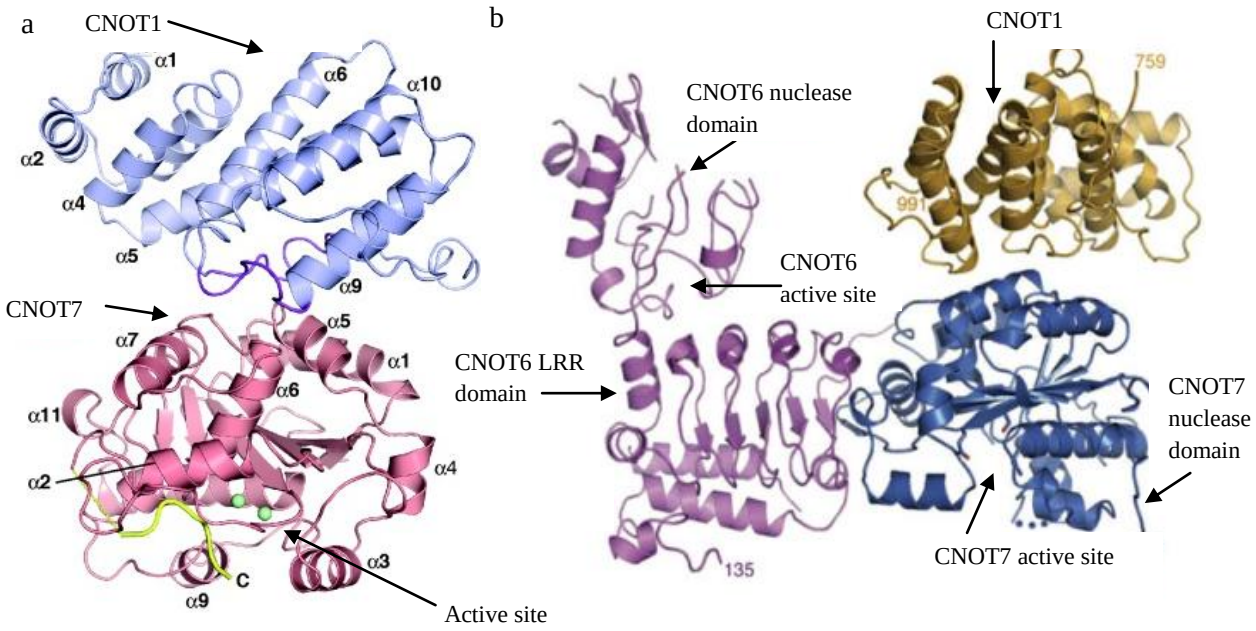


Figure 1.9 Crystal structure of the heterodimer of CNOT1 bound to CNOT7 (PDB ID:4GMJ) adapted from Petit et al and the nuclease module of the CCR4-NOT complex adapted from Basquin et al (PDB ID:4B8C) (Petit et al., 2012 and Basquin et al., 2012)

(a) The crystal structure depicts the CNOT1 MIF4G domain (lilac) bound to the deadenylase subunit CNOT7 (pink). The magnesium ions chelated in the active site (yellow) are shown in green. CAF1 (CNOT7) additionally interacts with CNOT6/CNOT6L (not shown in model). (b) The model shows the nuclease module of the yeast CCR4-NOT complex where CNOT7 is bound by both CNOT1 and CNOT6.

Lau et al used tagged CNOT subunits at near endogenous levels to isolate the complex and found that CNOT7 and CNOT8 were part of distinct CCR4-NOT complexes, as were CNOT6 and CNOT6L, indicating the two isoforms of the deadenylase subunits were mutually exclusive (Lau et al., 2009). Thus, the deadenylase module is comprised of either CNOT6/6L (or CCR4 in yeast) and CNOT7/8 (or CAF1 in yeast). CNOT6/CNOT6L bind CNOT7 which then associates with CNOT1, shown in **Figure 1.9**. Mutations to disrupt the residues which bind magnesium ions in the catalytic site of each of the individual subunits had little effect on the activity of the complex, suggesting overlapping functions (Stowell et al., 2016). It will be interesting to structurally characterise the interaction between the deadenylase subunits to determine how the arrangement allows access to both catalytic sites. Any differences between solved structures containing varied isoforms may also help to reveal contrasting features that may explain the requirement of the two isoforms. There are 10 human deadenylases so it is proposed that each may degrade a specific set of mRNAs e.g. CNOT8 binds PUF proteins more tightly than CNOT7 (Van Etten et al., 2012). However, this proposal raises several further questions concerning the affinity of deadenylase subunits for different subsets of mRNAs. Moreover, it remains to be determined if the cell type or position in the cell cycle effects the expression levels of the deadenylases.

1.3.4 Interactions mediated by the C-terminal region of CNOT1

The C-terminal region (aa 1595-2373, **Figure 1.5**) of CNOT1 contains a NOT box domain which provides a platform for the interaction with CNOT2 and CNOT3 NOT boxes. It is hypothesised that this module provides structural stability to the complex. Recent data also suggests a role in the differentiation of human embryonic stem cells (ESCs) as CNOT1, CNOT2 and CNOT3 were found to be highly expressed in human H1 ESCs (Zheng et al., 2012). Moreover, when sequentially silenced, a differentiation-like

morphology was observed, suggesting CNOT1, CNOT2 and CNOT3 may inhibit extra-embryonic lineage differentiation. Evidence concerning the participation of CNOT3 in self-renewal led scientists to explore the role of CNOT3 in cancer. CNOT3 is mutated in human cancers including prostate and urinary tract cancers, thus establishing a function during tumourigenesis (Forbes et al., 2011).

CNOT4 (aa 227-433) has been shown by Y2H to interact with the C-terminus of CNOT1 (aa 1106-2376) (Albert et al., 2000). CNOT4 is anomalous to the other CNOT subunits in that it only interacts transiently with the human CCR4-NOT complex (Lau et al., 2009). It was established by gel filtration and mass spectrometry (MS) to purify in a ~200 kDa complex without the other CNOT subunits. Thus, further studies are required to determine what effects the recruitment of CNOT4 by the CCR4-NOT complex. Y2H screens and ubiquitin assays identified that the N-terminus of CNOT4 has E3 ligase activity, mediated by a novel C_4C_4 RING finger fold (**Figure 1.10**) (Albert et al., 2002). CNOT4 also contains a C_3H-type zinc-finger and an RNA-recognition motif (RRM). The precise function of this subunit within the complex remains to be determined and further analysis is required to explore in which situations CNOT4 is bound to the CCR4-NOT complex.

Figure 1.10 Domain schematic of CNOT4

Schematic depicts the domain architecture of CNOT4. CNOT4 contains 3 domains: a RING domain which is essential for its E3 ligase function, an RNA-recognition motif (RRM) domain and a C_3H-type zinc-finger.

TAB182, a tankyrase interacting protein that has a role in mitosis, has been found to co-purify with the CCR4-NOT complex. However, TAB182 is not detected in every preparation of the complex, suggesting it may associate under certain currently unknown

circumstances (Lau et al., 2009, Mauxion et al., 2013). It is unclear with which subunits TAB182 interacts and what the biological function is of this interaction.

The interactions described above have been summarised in a 2-dimensional schematic in **Figure 1.11**. The domain arrangement of CNOT1 is shown at the top in colour. The CNOT subunits which interact with CNOT1 domains are indicated by a downwards facing arrow e.g. CNOT2 interacts with the NOT box domain at the C-terminus of CNOT1.

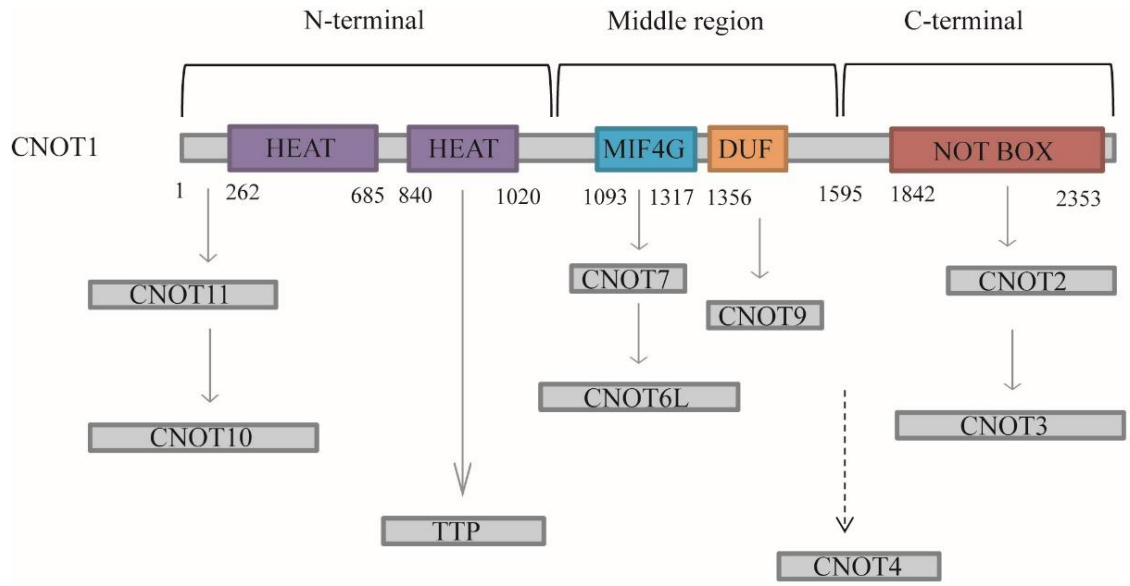


Figure 1.11 Domain schematic of CNOT1 and its interaction with the CNOT subunits

Schematic depicts the domain architecture of CNOT1, which is divided into three regions; an N-terminal domain, a middle region and a C-terminal domain. The numbers indicate the amino acid boundaries for the different domains. The extreme N-terminus of CNOT1 has been shown to interact with CNOT11, which then also interacts with CNOT10. CNOT1 contains two HEAT repeats at the N-terminus where the second has been shown to interact with TTP. A MIF4G domain is located within the middle region of CNOT1 and it has been shown to bind the deadenylase module of CNOT6/6L and CNOT7/8. A domain of unknown function (DUF) is also present in the middle region and this has been shown to recruit CNOT9. The C-terminal region of CNOT1 contains a NOT box domain shown to bind the subunits CNOT2 and CNOT3. The final CNOT subunit, CNOT4, has been shown to co-purify in some preparations of the complex thus it is anticipated that CNOT4 only transiently interacts with the other CNOT subunits.

1.3.5 Current structural information of the whole CCR4-NOT complex

There are a number of crystal structures of individual domains or sub-complexes from the yeast and human CCR4-NOT complex as described above which are detailed in **Appendix Table 7.1**. Although informative in expanding our understanding of the individual proteins and sub-complexes, the wider interactions between the complex as a whole should be the focus for future research to ascertain how these sub-complexes or different modules are co-ordinated by CNOT1. Moreover, the yeast complex is found in two distinct sizes; 1 MDa and 1.9 MDa. Studying the complex as a whole may reveal the purpose of the two different compositions (Lui, 1998, Bai et al., 1999).

In an effort to study the whole complex, two cryo-electron microscopy (EM) models of the yeast complex have been generated. The first model determined used the yeast strain *S. cerevisiae* where NOT1 was tandem affinity purification (TAP) tagged. The complex was captured using IgG resin and then further purified using a MoBiTech column. The complex was purified and cross-linked simultaneously using a glycerol gradient. This resulted in a sample containing NOT1-5, CCR4, CAF1, CAF40 (equivalent to human CNOT9) and CAF130. Analysis by cryo-EM produced a ~30 Å model of the CCR4-NOT complex. The model showed an L-shaped particle with two arms of ~600 kDa and ~300 kDa (Nasertorabi et al., 2011). The authors hypothesise, based on size considerations, that CNOT1 constitutes the longer arm and the other CNOT proteins as the shorter, with the deadenylase subunits in the hinge region. However, resolution was limited by heterogeneity of the complex.

Five years later, another cryo-EM model was generated but with improved resolution of ~20 Å (**Figure 1.12**) (Ukleja et al., 2016a). Again, the complex was purified from yeast but an alternative species was utilised, *S. pombe* where NOT2 was TAP-tagged.

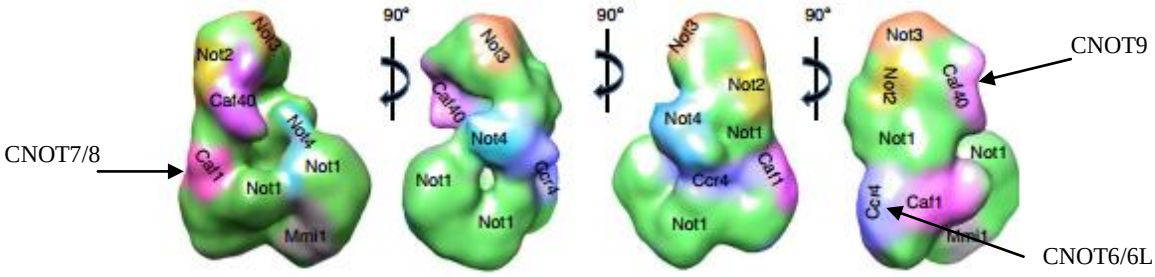


Figure 1.12 Cryo-EM reconstruction of the yeast CCR4-NOT complex adapted from Ukleja et al (Ukleja et al., 2016a)

Model derived by cryo-EM of the CCR4-NOT complex (Ukleja et al., 2016a). Inconsistent naming between the yeast and human complexes has been clarified by labels.

Purification again used IgG resin to capture the protein and glycerol gradient centrifugation to purify and cross-link the sample. NOT1-5, CCR4, CAF1 and CAF40 were detected. Similarly, the cryo-EM model depicted an L-shaped particle. To determine exact positions of the CNOT subunits, a green fluorescent protein (GFP) was fused to different CNOT proteins and the complex purified. An anti-GFP antibody was added to the purified complex and the sample was analysed using cryo-EM. The difference between the antibody-bound and unbound models highlighted the location of the antibody which was used to identify the position of the GFP-fused subunit. As a result, CCR4 and CAF1 were found in the hinge region of the L-shape particle, as previously hypothesised, with NOT1 forming two different length arms. NOT2 and NOT3 were localised to one arm of NOT1. Computational modelling using PyRy3D software was then used to predict the positions of the other subunits. Known crystal structures of yeast sub-complexes were modelled into the L-shaped particle determined by cryo-EM. This revealed CAF40 within the arm containing NOT2/NOT3 and NOT4. In the other arm, Meiotic mRNA interceptor 1 (Mmi1) was modelled. Mmi1 is an RNBP that is an essential gene in *S. pombe*. Mmi1 is known to target cis-acting elements in meiotic mRNA known as determinants of selective removal (DSRs) and directs them for subsequent decay, by first initiating deadenylation by recruiting the CCR4-NOT complex

(Stowell et al., 2016). Although speculative and computational, this annotated model of the CCR4-NOT complex provides the best identification of complex composition currently known. It provides some structural information on the overall complex and allows hypotheses to be made on how additional subunits may interact with the complex.

While isolating the endogenous complex has proved successful and enabled structural determination at low resolution, a more advanced method of complex generation is now required to achieve the higher resolution and greater flexibility essential for advancing our understanding of the CCR4-NOT complex. Recombinant expression of the entire complex may offer several benefits including greater control over subunit composition, less heterogeneity, the ability to introduce disease mutants and potentially greater protein yield. These benefits may be of importance in future structural/biochemical studies e.g. recombinant expression would allow the introduction of mutations within CNOT3 which are thought to have a link to certain cancers (Forbes et al., 2011).

Our understanding is advancing on the composition of the yeast CCR4-NOT complex, but it still remains to be determined how the human complex is arranged. Given the sequence homology of individual subunits and the additional CNOT10:CNOT11 module, it may be that the human and yeast CCR4-NOT complexes show a diverse architecture. Therefore, the overall formation of the human complex, how and where the CNOT10:CNOT11 module functions remain important questions and efforts in structural biology should be focussed to answer these questions. Reconstitution of the whole complex by recombinant expression may also allow interaction between the CCR4-NOT complex and specific RNBPs to be studied more closely.

1.4 RNA-binding proteins that regulate inflammation

RNBPs (e.g. HuR, TTP, Mmi1) are targeted to mRNA by sequences commonly found in the 3′UTRs (e.g. AREs). RNBPs recruit the CCR4-NOT complex, forming a physical link between the bound mRNA and the CCR4-NOT complex which facilitates the degradation of the mRNA. The mechanism of one such RNBP, TTP, will be discussed below.

1.4.1 Discovery of TTP

TTP was discovered in the early 1990s from a cDNA library derived from insulin-stimulated cells and was named after its three Pro-Pro-Pro-Pro repeats (Lai et al., 1990). These tetra-proline repeats are located in intrinsically disordered N- and C-termini which flank two C_3H-type zinc-fingers. A nuclear magnetic resonance (NMR) structure of the tandem zinc-finger has been solved of a TTP family member (PDB ID:1RGO) (**Figure 1.13**) (Hudson et al., 2004). This domain is largely conserved across the TTP family of proteins. The structure depicts the two zinc-fingers bound by RNA where the zinc ions are shown in grey and residues that chelate zinc in green. TTP belongs to the Tis11 family of RNBPs which bind pro-inflammatory mRNAs, thus making it an important target for understanding inflammatory gene expression (Carballo et al., 1998).

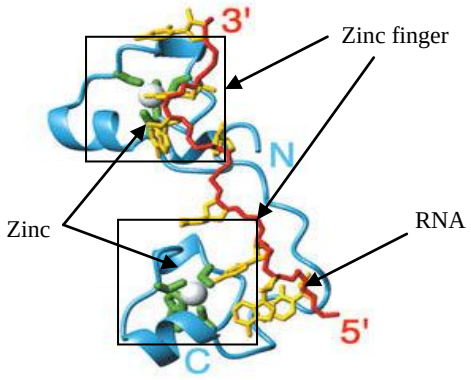


Figure 1.13 NMR model of zinc-finger domain of Tis11d adapted from Hudson et al (Hudson et al., 2004)

The NMR structure of a TTP family member has been solved. The zinc-finger domain is bound to a stretch of RNA. Zinc ions are shown as grey balls and residues which chelate them are highlighted in green. The zinc-finger is conserved across the TTP family and is the main structured element, flanked by largely disordered N- and C-termini.

Interest in TTP has led to the production of novel mouse strains which provided insight into the protein's biology. Taylor et al utilised gene targeting to create mice deficient in TTP (Taylor et al., 1996a). The mice appeared normal at birth but after 1-8 weeks, they developed alopecia, dermatitis, erosive arthritis, cachexia and conjunctivitis. Development of these symptoms demonstrates $TTP^{-/-}$ mice to be a useful model for human autoimmune diseases. The phenotype of TTP-deficient mice was noted to resemble chronic TNF-α stimulation. TNF-α is a pro-inflammatory cytokine which functions in signalling pathways regulating angiogenesis, cell survival and invasion, and TNF-α has been shown to promote tumour formation in mice. The similarity between phenotypes of TTP and TNF-α knockout mice prompted an attempt to 'rescue' the mice using an anti-TNF-α antibody (Taylor et al., 1996a). The successful rescue experiment revealed that TTP regulates TNF-α.

Transgenic mice have also been influential to study the stability of TNF-α mRNA, where one mouse strain includes the targeted deletion of the TNF-α ARE (ΔARE) (Kontoyiannis et al., 1999). These mice weigh less and have increased mortality rates compared to wild-type littermates. Swollen and distorted joints impaired movement, and indicated the development of inflammatory arthritis (Kontoyiannis et al., 1999). It can be reasoned that if the TNF-α mRNA does not contain an ARE, it will not be targeted for degradation by TTP. This results in high levels of mRNA/protein and an inflamed phenotype (Qiu et al., 2012). Thus, it was established that TTP facilitates the degradation TNF-α mRNA via the recognition of an ARE and recruitment of the CCR4-NOT complex (Williams, 2012).

1.4.2 Mechanism of action of TTP

TTP is regulated by an upstream kinase cascade which is initiated by p38 mitogen-activated kinase (MAPK). The p38 pathway (**Figure 1.14**) is stimulated by pro-inflammatory signals which initiate the phosphorylation of MAPK kinase (MK2). MK2 phosphorylates TTP at two major sites, Ser-52 and Ser-178 (murine numbering) (Mahtani et al., 2001, Chrestensen et al., 2004). It has been shown that MK2-mediated phosphorylation of TTP directly prevents the recruitment by TTP of the CCR4-NOT complex to mRNA (Marchese et al., 2010). Ser-323 is also a minor MK2 site (Cao et al., 2007), phosphorylation of which blocks CNOT1 binding (Fabian et al., 2013). Thus, phosphorylation of TTP acts a molecular activation switch. This has been studied using a

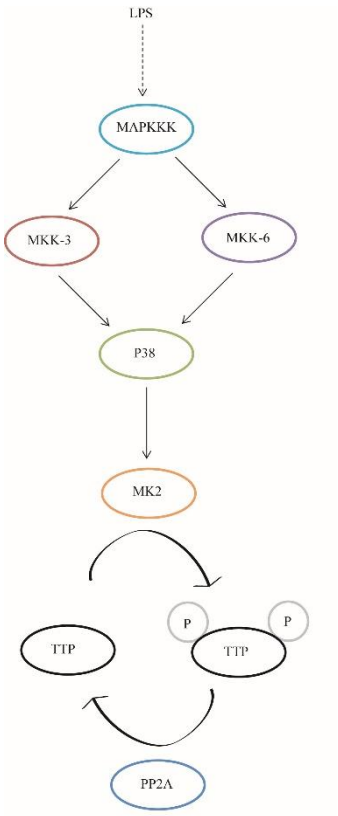


Figure 1.14 p38 MAPK signally pathway

The p38 MAPK signalling pathway is initiated by pro-inflammatory stimuli e.g. lipopolysaccharide (LPS). What follows is a kinase cascade leading to the phosphorylation of TTP by MK2. When phosphorylated, TTP is prevented from recruiting the CCR4-NOT complex to direct mRNA degradation. But when TTP is dephosphorylated by PP2A, TTP can recruit the CCR4-NOT complex and participate in mRNA degradation.

mouse model where the endogenous TTP was replaced by a non-MK2 phosphorylatable equivalent which produces constitutively active TTP (Ross et al., 2015). When the mice were exposed to lipopolysaccharide (LPS) to initiate an inflammatory response, these

animals exhibited a less severe response, indicating that the constitutively active TTP was responsible for elevated ARE-mediated mRNA decay. This model demonstrated that it is the unphosphorylated form of TTP that actively degrades mRNAs. Cellular studies and mouse models suggest the model shown in **Figure 1.15**. TTP binds the ARE of inflammatory mRNAs and targets them for either degradation or stabilisation. When TTP is unphosphorylated, it can recruit the CCR4-NOT complex, bringing together the inflammatory mRNAs and the catalytic subunits which then degrade the poly-(A) tail. Therefore, the amount of inflammatory mRNA is reduced and subsequently less cytokine protein is translated which returns inflammation to basal levels.

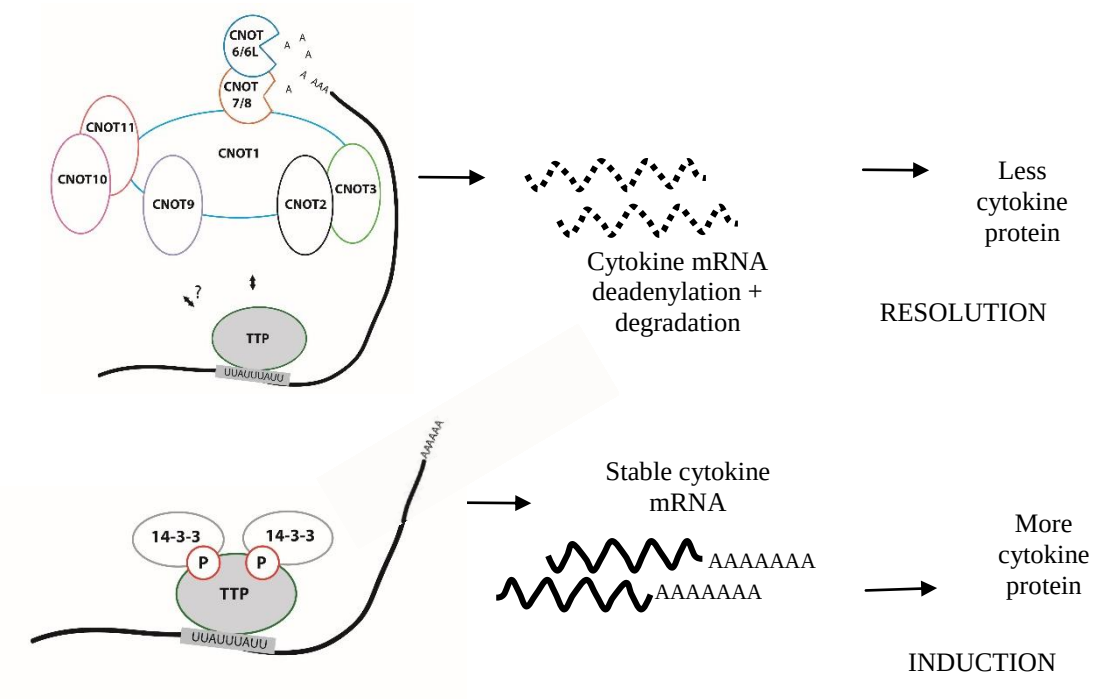


Figure 1.15 The effect of phosphorylation on recruitment of CCR4-NOT by TTP

Unphosphorylated TTP recruits the CCR4-NOT complex which allows the catalytic subunits of the complex to degrade the poly-(A) tail, the first step of mRNA degradation. Therefore, there is less cytokine mRNA and less cytokine protein so inflammation is resolved. However, when TTP is phosphorylated by MK2, it can no longer recruit the complex and is instead bound by 14-3-3 which stabilises inflammatory mRNAs and so there is increased expression, more cytokine protein and inflammation is induced.

Conversely, phosphorylated TTP cannot recruit the complex and is instead bound by 14-3-3, a protein which binds Ser/Thr phosphorylated proteins (Martens et al., 1992). 14-3-3-bound TTP stabilises mRNA which increases cytokine production, inducing inflammation (Stoecklin et al., 2004). Binding of 14-3-3 has been suggested to promote nuclear localisation of TTP which means it would be no longer be available for interaction with the complex (Taylor et al., 1996b). Phosphorylation of TTP also prevents it from being degraded by the proteasome (Ngoc et al., 2014). Therefore, phosphorylation of TTP causes stabilisation and inactivation and this allows TTP to be maintained in the inactive state as a reserve pool. However, when p38 MAPK stimulation dissipates, PP2A, a phosphorylase, can dephosphorylate TTP which activates it (**Figure 1.14**). Thus, the phosphorylation state of TTP acts as molecular switch between the active, unphosphorylated form and the stable, inactive, phosphorylated form which

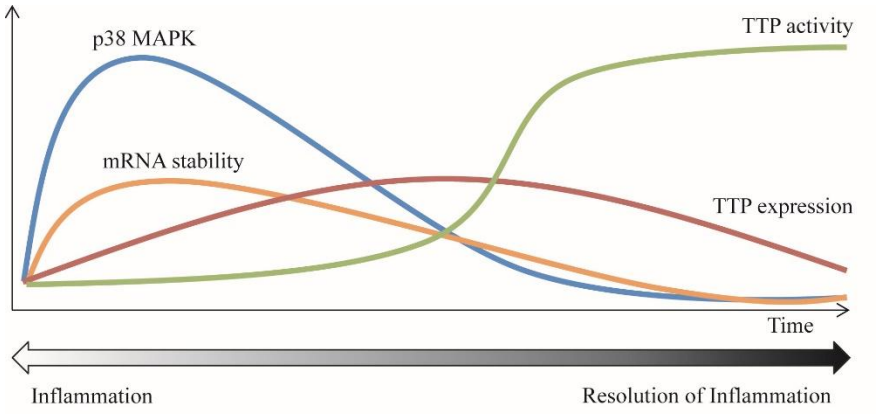


Figure 1.16 Interconnecting expression levels of TTP and p38 MAPK

As p38 MAPK is stimulated by inflammatory signals, mRNA stability is increased which allows a range of inflammatory cytokines to be translated to combat the inflamed tissues. As p38 MAPK stimulation dissipates, TTP can be dephosphorylated by PP2A which decreases the stability of mRNAs and reduces cytokine translation, bringing a resolution to inflammation.

allows careful regulation over mRNA stability and consequently cytokine protein levels (**Figure 1.16**) (Mahtani et al., 2001, Chrestensen et al., 2004) (Tiedje et al., 2016).

Although the general mechanism of the regulation of TTP by phosphorylation appears clear, smaller details require further clarification. It remains unclear whether phosphorylated and unphosphorylated forms of TTP bind AREs with similar affinity (Clement, 2011, Hitti et al., 2006, de Boer et al., 2014, Marchese et al., 2010). Individual nucleotide resolution crosslinking and immunoprecipitation (iCLIP) was used to demonstrate that a phosphorylation mutant of TTP more readily binds the 3′UTR of mRNA compared to the wildtype TTP. However, it still remains to be determined if there is a phosphorylation-dependent effect on how the TTP zinc-finger binds mRNA.

In addition to a role in inflammation, TTP has been shown to regulate the half-life of proto-oncogenic mRNAs and all ten of the molecular hallmarks of cancer include genes which contain AREs (Stoecklin et al., 2003). TTP loss in cervical cancer leads to the stabilisation of E6-AP ubiquitin ligase which degrades p53 (Sanduja et al., 2009) and low expression levels of TTP in pancreatic cancer, among others, are linked to a poor outcome of patient survival (Stoecklin et al., 2003, Brennan et al., 2009, Wei et al., 2016). Given these observations, TTP has been termed a tumour suppressor gene. Therefore, understanding TTP biology may be vital for a number of conditions including cancer and chronic inflammatory conditions such as rheumatoid arthritis.

1.5 Rheumatoid Arthritis

Rheumatoid arthritis (RA) is a chronic autoimmune disease which affects 400,000 people in the United Kingdom, and presents in the fourth or fifth decade of life (Koopman, 2003). Currently, the etiological cause remains unknown. However, certain genetic associations have been established, the strongest of which being within genes located in the major histocompatibility complex, a region on chromosome 6 which spans nearly 4 megabases (Mb). Within this region are human leukocyte antigen (HLA) genes, DR1 and

DR4, which are an indicator of disease susceptibility (Kapitany et al., 2005, Fishman. P., 2010).

Once the disease is triggered, macrophages, among other cells, invade the synovium. Macrophages release inflammatory cytokines, such as TNF-α, which perpetuates the inflammatory response in rheumatoid joints (Remmers et al., 2007). The overproduction of cytokines induces a secondary inflammatory response involving the recruitment of further inflammatory cells and joint destruction factors (Fishman. P., 2010). Synovial inflammation constitutes the main symptom of the disease which leads to pain, swelling, cartilage destruction, bone erosion and joint deformation. Joint deformation occurs in about 90% of patients in the first two years after diagnosis and patients typically develop a moderate disability (Fishman. P., 2010). These implications impose a significant financial and emotional burden on patients and their families. Thus, investigation into the mechanism of this disease and potential therapeutic strategies remains an important goal.

1.5.1 Treatment for RA

Newly diagnosed RA patients are prescribed disease modifying rheumatic drugs (DMARDs) within 3 months of diagnosis, methotrexate being the most notable example (Romao et al., 2013). Further to this, a number of biological agents have been developed e.g. anti-TNF-α (Gaffo et al., 2006) (Choy et al., 2013). Enternacept, the first anti-TNF-α inhibitor developed in 1998, acts as a decoy receptor to bind soluble TNF-α and prevent propagation of the signal (Mohler et al., 1993). However, this biologic was given subcutaneously once a week and was superseded by infliximab, where treatment is intravenous (Seymour et al., 2001). Although successful in reducing the symptoms of RA, there are problems with biologics including high cost, patient-wide efficacy and risk

of lymphoma (Romao et al., 2013). Hence, research has been invested to pursue the discovery of small molecule inhibitors.

P38 MAPK was a natural focus of drug discovery as it is a well characterised protein and there is a precedent for targeting kinases to inhibit phosphorylation (Melnikova and Golden, 2004). The rationale is that since activation of p38 MAPK causes the production of pro-inflammatory cytokines, inhibiting this protein will reduce the inflammation. Further fuelling the hunt was the discovery that an inhibitor of p38 MAPK blocked TNF-α and IL1 production by LPS treated monocytes, and the identification of its post-transcriptional mechanism of action (Lee et al., 1994, Dean et al., 1999). This consolidated p38 MAPK as a potential target and attracted many drug companies (Dominguez et al., 2005). As with most kinases, small molecules were designed to be competitive inhibitors of the ATP-binding site e.g. imatinib for Bcl-Abl translocations in chronic myeloid leukaemia or gefitinib for EGFR which is an approved therapy for non-small cell lung carcinoma (Melnikova and Golden, 2004). Although many compounds inhibited p38 MAPK, a number of issues plagued drug development including limited long-term efficacy problems (Hammaker and Firestein, 2010). Progress still needs to be made in the field of anti-inflammatory treatments and future drug discovery may focus on other proteins within the pathway such as MK2 or MAP kinase kinase (MKK) 3 or 6 (Hammaker and Firestein, 2010) (**Figure 1.14**).

1.6 Aims

As described above, the CCR4-NOT complex is an important master regulator of gene expression and although some efforts have been made to characterise the entire yeast complex, further work is required, especially focusing on the human complex. Increasing our understanding of the human CCR4-NOT complex and the proteins it recruits will

allow us to better understand how the complex orchestrates deadenylation and consequently regulates gene expression.

This project had two main aims, the first of which was to identify interacting regions between subunits of the human CCR4-NOT complex. This aim was further subdivided to focus firstly on the relatively unknown module of CNOT10:CNOT11 and investigate the motifs required to form the module and to recruit it to CNOT1. Crystallisation trials, limited proteolysis and MS were used to confirm the Y2H screen finding that CNOT10 and CNOT11 form a complex. These techniques were also employed to identify the interacting domains of CNOT10 and CNOT11, as described in chapter 3. The second focus of this aim was to study a larger assembly of CNOT subunits. As such, assemblies of CNOT subunits were reconstituted using the MultiBac system for baculovirus expression of protein complexes. Reconstituted complexes were assayed for deadenylase function to confirm they are biologically active, as described in chapter 5. Producing biologically active assemblies of the CCR4-NOT complex provides a useful tool for the field to facilitate experiments to improve the current understanding of the CCR4-NOT complex.

The second aim was to characterise the interface between the CCR4-NOT complex and TTP. A 10-residue C-terminal peptide of TTP has been identified to bind to the middle region of CNOT1, the main scaffold protein of the complex, thus recruiting TTP to the complex (Fabian et al., 2013). However, given the relatively weak binding affinity of the TTP peptide to CNOT1 (2 μM) and the literature highlighting the importance of the N-terminus of TTP for targeted deadenylation (Lykke-Andersen and Wagner, 2005), it remained to be determined if further CNOT subunits were required to recruit TTP. Researchers from the Kennedy Institute used a siRNA knockdown screen to determine if any additional CNOT subunits were important for the recruitment of TTP by the CCR4-

NOT complex. Each subunit of the complex was sequentially knocked down and the ability of TTP to recruit the complex was monitored by GST pull-down. Results indicated that when CNOT9 or CNOT2 is depleted, TTP is less able to recruit the complex (Bulbrook, 2015), (**Appendix Figure 7.11**). Biophysical techniques and mutational studies were utilised to first validate the interaction between TTP and CNOT9/CNOT2. Subsequently, these techniques were used to explore the motifs required in both proteins to recruit TTP to the CCR4-NOT complex for targeted deadenylation, a process described in chapter 4. Determining the structure of the complex and the proteins it recruits will increase the current understanding of gene expression as the structural interface between the complex and its RNBPs may provide an explanation of the regulatory processes in place to control deadenylation, e.g. how phosphorylation of TTP by MK2 prevents recruitment by the CCR4-NOT complex (Marchese et al., 2010).

2. Materials and Methods

Materials

2.1 Bacterial strains

Name	Description of strain
MACH1	T1 phage-resistant strain for the preparation of DNA.
BL21(DE3)-R3-pRARE2	Adapted from BL21(DE3) and Rosetta2 (Merck) to produce a phage-resistant derivative of BL21(DE3) termed BL21(DE3)-R3. This bacterial strain was then transformed with plasmid pRARE2 (isolated from Rosetta2 cells), which carries seven rare-codon tRNA genes. The resulting chloramphenicol-resistant strain BL21(DE3)-R3-pRARE2.
BL21(DE3)-R3-pRARE2 BirA	Adapted from the BL21(DE3)-R3-pRARE2 strain by transformation with the BirA plasmid which encodes the gene for *Escherichia coli (E. coli)* biotin-protein ligase.
DH10Bac	Contains a baculovirus shuttle vector (bMON14272, kanamycin) and a helper plasmid (pMON7142, tetracycline). This allows site-specific recombination between baculovirus cloning vectors and bMON14272 to produce recombinant bacmids for transfection and subsequent expression in insect cells.
DH10MBac	Adapted from the DH10Bac strain (Thermo Fisher) but containing disruption of the *viral cathepsin-type cysteine protease (V-CATH)* gene and the *chiA* gene, encoding chitinase. Disruption of genes both genes served to a) eliminate V-CATH activity and b) enable chitin-affinity chromatography for purification without interference from the *chiA* gene product.

Table 2.1 Bacterial strains

Five different bacterial strains were used during the course of this project. MACH1 *E. coli* were used for the propagation of DNA and cloning of CCR4-NOT subunits and associated proteins. BL21(DE3)-R3-pRARE2 *E. coli* were used for the expression of proteins. BL21(DE3)-R3-pRARE2 BirA *E. coli* were used for the expression of proteins which required in vivo biotinylation for the end use of the protein e.g. biotinylation of proteins allows the attachment to a streptavidin biosensor in BLI. DH10Bac and DH10MBac were used for the transposition of baculovirus expression vectors into a recombinant bacmid within the *E. coli*.

2.2 Vectors

All vectors have been adapted at SGC to contain compatible ends that facilitates sub-cloning using the ligation independent cloning (LIC) approach, and the *SacB* gene, which is used for screening of correct clones. When the vector is fully digested, the *SacB* gene is removed and the gene of interest is cloned in its place. The presence of *SacB* causes toxicity in gram-negative bacteria when plated on sucrose-containing plates. It is thought that the levansucrase, encoded by the *SacB* gene, catalyses the accumulation of levans in the periplasm of gram-negative bacteria which leads to cell death. Therefore, when the ligated products from LIC cloning are transformed, the bacteria are spread onto 5% sucrose plates (plus the antibiotic selection), allowing negative selection. Bacteria which have been transformed with vectors which have been digested correctly and thus contain the gene of interest will grow successfully on the plate, whereas colonies which contain vectors with an intact *SacB* gene will not, ensuring the correct clones are chosen for future experiments.

2.2.1 *E. coli* expression vectors

Vector name	Description
pNIC28-Bsa4	pET-based expression vector with His_6 tag in N-terminal fusion peptide, with tabacco etch virus (TEV) protease cleavage site.
pNH-TrxT	pET-based expression vector with His_6–Trx (*E. coli* thioredoxin) in N-terminal fusion peptide, with TEV protease cleavage site.
pGTVL2	pET-based expression vector with His_6-GST tag at the N-terminal fusion, followed by TEV protease cleavage site.
pNIC-Bio3	pET-based expression vector with His_6 tag in N-terminal fusion peptide, with TEV protease cleavage site, and a biotinylation sequence in C-terminal fusion.
pNIC-MBP	pET-based expression vector with MBP-His_6 tag in N-terminal fusion peptide, with TEV protease cleavage site.
pSUMO-LIC	pET-based expression vector with His_6-SUMO (96-aa sequence derived from SUMO1 protein) tags followed by a SUMO protease cleavage site. If no stop codon is included in reverse primer, protein of interest will be fused to biotinylation tag followed by additional His_6 tag.

Table 2.2 *E. coli* expression vectors

Six different bacterial expression vectors were used during the course of this project. All vectors contained a His_6 fusion tag to allow nickel affinity capture in expression tests. But vectors also contained a number of other possible fusions to improve solubility (e.g. SUMO, MBP, GST) and to change functionality (e.g. Avi tag which facilitates the in vivo biotinylation of the protein to allow the attachment to a streptavidin biosensor in BLI).

2.2.2 Baculovirus-mediated insect cell expression vectors

Vector name	Vector description
pFB-LIC-Bse	Baculovirus transfer vector with His_6 tag in 22-aa N-terminal fusion peptide, with TEV protease cleavage site.
pFB-HGT-LIC	Baculovirus transfer vector with His_6-GST tag in 243-aa N-terminal fusion peptide, with TEV protease cleavage site.
pFB-CT10HF-LIC	Baculovirus transfer vector with C-terminal His_{10}-FLAG tag, preceded by a TEV protease cleavage site.
pFL-NT6H-LIC	Baculovirus transfer vector with N-terminal His_6 tag, succeeded by a TEV protease cleavage site. A Cre-LoxP site is present for recombination with other MultiBac vectors.
pSPL-CTF-LIC	Baculovirus transfer vector with C-terminal FLAG tag, succeeded by a TEV protease cleavage site. A Cre-LoxP site is present for recombination with other MultiBac vectors.
pUCDM-CTF-LIC	Baculovirus transfer vector with C-terminal FLAG tag, succeeded by a TEV protease cleavage site. A Cre-LoxP site is present for recombination with other MultiBac vectors.
pFL-CT10HStII-LIC	Baculovirus transfer vector with C-terminal His_{10} twin Streptactin tag, succeeded by a TEV protease cleavage site. A Cre-LoxP site is present for recombination with other MultiBac vectors.

Table 2.3 Baculovirus-mediated insect cell expression vectors

Seven different Baculovirus-mediated expression vectors were used during the course of this project. All vectors contained a His_6 fusion tag to allow nickel affinity capture in expression tests. But vectors also contained a number of other possible fusions to improve solution (e.g. GST). Additionally, MultiBac vectors, which are designed with a Cre-LoxP site to allow the recombination of multi-gene plasmids, were used during this project for the expression of CCR4-NOT sub-complexes in insect cells.

2.3 Primers

A full list of primers used for the sub-cloning of CCR4-NOT subunits and associated proteins are provided in the appendix (**Appendix Table 7.2 and 7.3**).

2.4 Media

Luria Broth (LB)

25 g of premixed LB powder (MERCK, comprising of yeast extract, peptone from casein and sodium chloride in a 1:2:2 ratio) was solubilised in 1 L of double distilled Milli Q water in 1 L glass bottle and autoclaved.

Terrific Broth (TB)

12 g of tryptone (or peptone from casein) was mixed with 24 g of yeast extract and 4 ml of glycerol and solubilised to a final volume of 900 ml double distilled Milli Q water in a 1 L glass bottle and autoclaved.

TB salt

2.31 g of KH_2PO_4 was mixed with 12.54 g of K_2HPO_4 and solubilised in a final volume of 100 ml of double distilled Milli Q water and autoclaved. Once cool, the 100 ml of TB salt is added to 900 ml of TB before inoculation with bacteria.

LB-agar

25 g of premixed LB powder (MERCK) was mixed with 13.5 g of agar and solubilised with 1 L of double distilled Milli Q water in 1 L glass bottle and autoclaved.

2.5 Construct design

To increase the likelihood of obtaining soluble protein for certain regions of the target protein, a multi-construct approach was adopted (Gileadi et al., 2008), surveying combinations of N- and C-terminal boundaries that were designed to correlate with either previously solved structures, or secondary structure prediction tools (e.g. PFAM search, psiPRED, DOMpred, disoPRED) (Buchan et al., 2013, Finn et al., 2016, Jones and Cozzetto, 2015) (**Figure 2.1**).

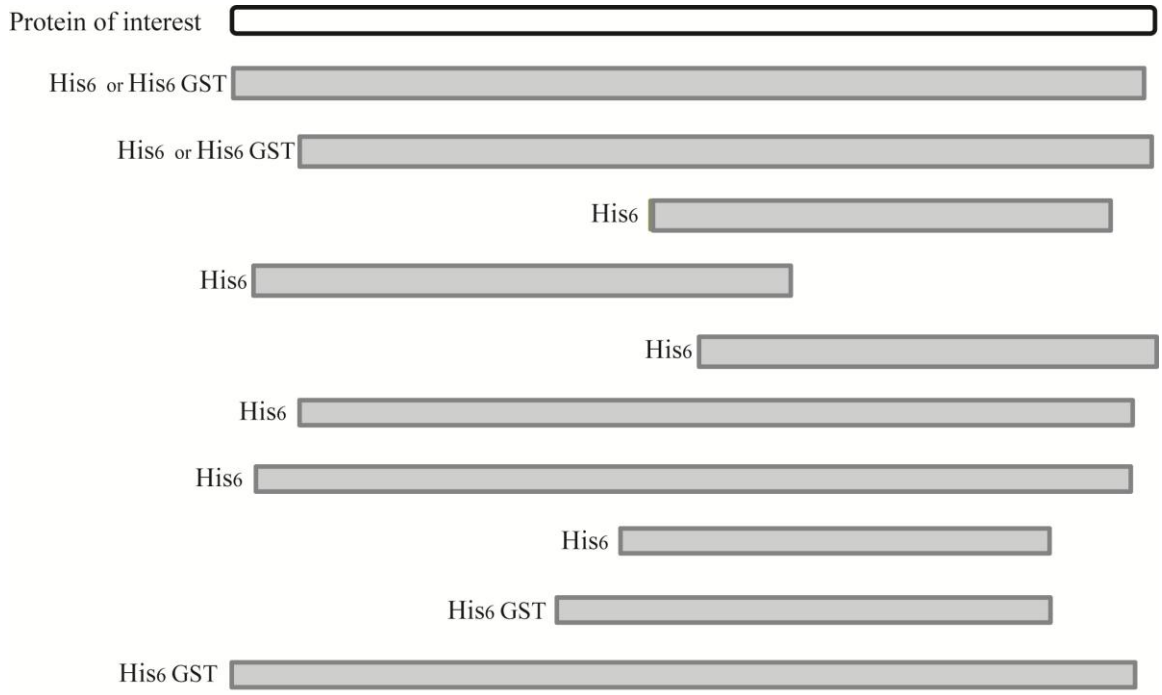


Figure 2.1 Example schematic of nested construct design

A multi-construct approach was used where constructs with both N- and C-terminal truncations were designed to explore the soluble regions of the protein of interest. Different fusion tags were also trialled to improve solubility or add functionality (e.g. an Avi tag allows in vivo biotinylation of constructs which can then be bound to a streptavidin biosensor for protein interaction experiments using BLI).

2.6 Preparation of DNA plasmids for sub-cloning and transformation of *E. coli*

Human cDNA clones for the CCR4-NOT subunits and associated proteins were obtained from the Mammalian Gene Collection (MGC) at the SGC and are listed in the **Appendix Table 7.4**. Plasmid DNA was prepared using the QIAspin miniprep kit (QIAGEN), following the manufacturers′ instructions.

2.7 Ligation Independent Cloning (LIC)

To achieve high-throughput (HTP) cloning of 96 constructs at a time, the LIC strategy was employed (Savitsky et al., 2010). LIC is a simple and cost-effective tool to produce constructs from multiple targets in parallel. Firstly, genes of interest are amplified by PCR using gene-specific primers which also contain a LIC compatible sequence in the 5′

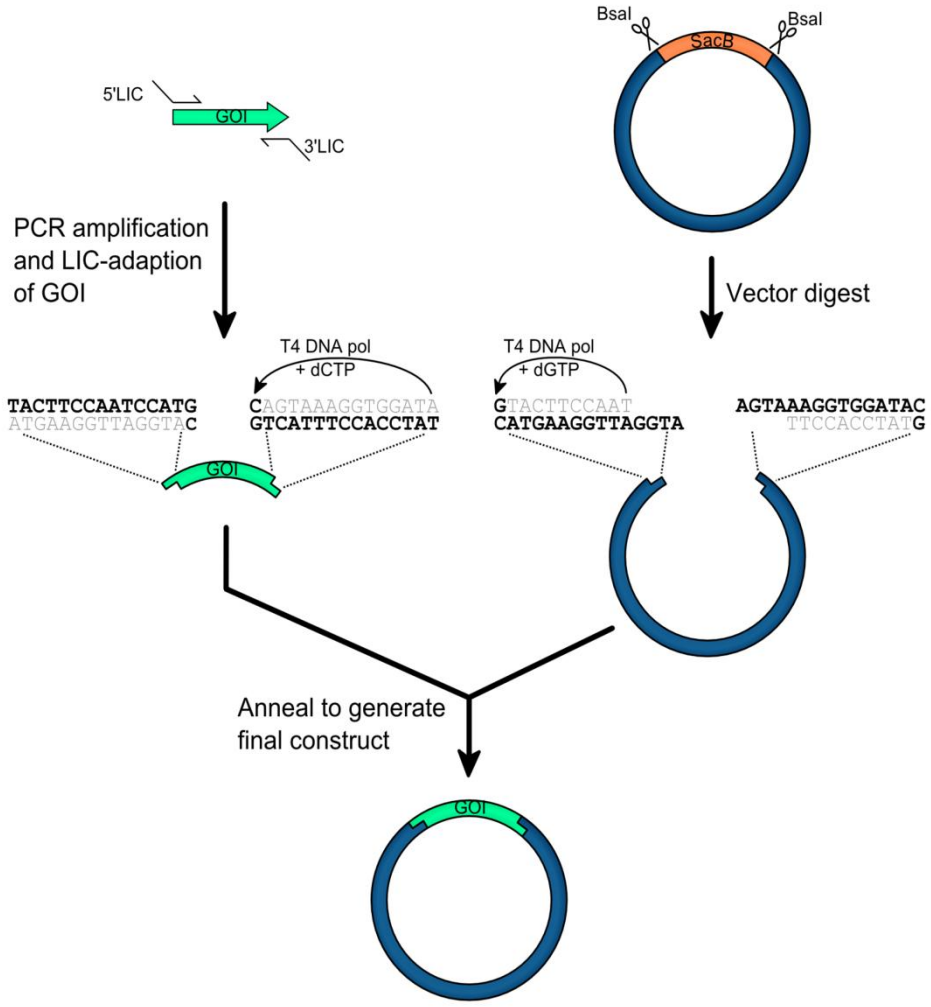


Figure 2.2 LIC schematic

Firstly 5′ (TACTTCCAATCCATG) and 3′ (TATCCACCTTTACTG) primers with LIC specific termini are used to generate a DNA fragment of the gene of interest (GOI) by polymerase chain reaction (PCR). The vector is digested using restriction endonucleases (e.g. BsaI) which remove the toxic *SacB* gene used for negative selection of colonies. Following this, the PCR product and the digested vector are treated with T4 DNA polymerase in the presence of an excess of dCTP for the insert and dGTP for the vector where the 3′-5 exonuclease activity creates compatible ends which anneal without the need for DNA ligase.

(TACTTCCAATCCATG) and 3′ (TATCCACCTTTACTG) terminus of the primer, indicated in **Figure 2.2**. The LIC-adapted vector of choice is digested with a restriction enzyme (e.g. BsaI) which removes the *SacB* gene used for sucrose selection, as described in section 2.2. After which, both the insert and the vector are incubated with T4 DNA polymerase and an excess of a single nucleotide. The 3′-5′ exonuclease activity of T4 DNA polymerase creates compatible ends between vector and insert that anneal without DNA ligase. This methodology allows the same procedure to be applied to many targets, removing the need for tailoring of specific restriction enzyme to each target, thus increasing throughput of cloning.

2.7.1 Polymerase chain reaction (PCR)

Forward and reverse primer mixes (final concentration 5 μM) (synthesised to HPLC-grade by Eurofins) were aliquoted into a 96-well PCR plate (STARLAB). A full list of the primers used for sub-cloning is shown in the appendix (**Appendix Table 7.2 and 7.3**). Template DNA was diluted to 2.5 ng/μl with nuclease-free water. A PCR master mix (for 100 reactions) composed of 500 μl of 5 x HercII buffer (Agilent), 75 μl of 10 mM dNTP mix (Invitrogen), 25 μl of HercII Polymerase (Agilent), 1.5 ml of nuclease-free water (Ambion) was prepared. 21 μl of the master mix was added to each well of a 96-well PCR plate, followed by primers (1.5 μl) and templates (2.5 μl) separately. The PCR plate was covered with an adhesive film (ABgene, Thermo Fischer Scientific) and placed in a Bio-Rad C100 Touch thermocycler (Bio-Rad) to execute the programme shown below.

95 °C, 10 min
(95 °C, 30 sec; 68 °C, 30 sec; 68 °C, 1-3 min*) x 5 cycles
(95 °C, 30 sec; 60 °C, 30 sec; 68 °C, 1-3 min*) x 5 cycles
(95 °C, 30 sec; 55 °C, 30 sec; 68 °C, 1-3 min*) x 5 cycles

(95 °C, 30 sec; 50 °C, 30 sec; 68 °C, 1-3 min*) x 10 cycles

68 °C, 10 min

15 °C hold

(*Extension time dependent on length of PCR product – e.g. 1 min per 1 kb fragment)

2.7.2 T4 DNA polymerase treatment of PCR products and vectors

The PCR products were treated with 1 µl DpnI enzyme (1 in 20 in Buffer 2 (NEB)) and incubated at 37 ^{o}C for 1 h to remove the template DNA. The inserts were diluted to 100 µl with nuclease-free water, transferred to a PCR purification plate (Millipore) and filtered using a vacuum manifold. The PCR products were washed with 100 µl of water and eluted with 50 µl of elution buffer. Purified inserts were T4 treated using 5 µl of DNA, 2.15 µl of nuclease-free water, 1 µl of 10 x T4 polymerase buffer (Novagen), 1 µl of 25 mM dCTP (Invitrogen), 0.5 µl of 100 mM DTT (Apollo Scientific Ltd), 0.1 µl of bovine serum albumin (BSA) (NEB) and 0.25 µl of T4 DNA polymerase (Novagen). The T4 treatment reaction was incubated for 30 min at 22 ^{o}C followed by 20 min at 75 ^{o}C in a thermocycler.

2.7.3 Vector digest

To remove the *SacB* gene and prepare vectors for LIC, 5 µg of the chosen vector (volume adjustable according to DNA concentration) from **Table 2.2** or **Table 2.3** was digested using 10 µl of Buffer 3 (NEB), 1 µl of BSA and 3 µl of BsaI (NEB). The digestion mixture was made up to 100 µl with nuclease-free water. The reaction was incubated for 2 h at 50 ^{o}C in a thermocycler and the digestion was confirmed on an agarose gel (Section 2.11). The digested vector was purified using a PureLink® PCR purification kit (Thermo Fischer Scientific). The vector was treated with T4 DNA polymerase using 21.5 µl of nuclease-free water, 50 µl of BsaI-digested plasmid, 10 µl of NEB Buffer 2, 10 µl

of 25 mM dGTP, 1 µl of BSA, 5 µl of 100 mM DTT and 2.5 µl of T4 DNA polymerase (NEB). The reaction mixture was incubated for 30 min at 22 ^{o}C followed by 20 min at 75 ^{o}C in a thermocycler.

2.7.4 Annealing and transformation

To generate the final construct, 1 µl of T4 polymerase treated vector and 2 µl of T4 polymerase treated PCR product were combined. The ligation was incubated for 30 min at room temperature then transferred to ice where 40 µl of MACH1 competent cells (produced in-house) were added. The cells were incubated on ice for 30 min, heat-shocked at 42 ^{o}C for 45 sec then chilled briefly on ice. 100 µl of SOC medium (2% tryptone, 0.5% yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM $MgCl_2$, 10 mM $MgSO_4$, and 20 mM glucose) was added. The cells were incubated for at least 60 min at 37 ^{o}C. 100 µl of transformation was plated onto LB-agar with the correct antibiotic selection, and incubated at 37 ^{o}C overnight. Colonies appeared the next day and DNA was extracted as described in section 2.6. This DNA could then be used to transform expression strains for expression tests or transposition.

2.8 Restriction cloning

Alongside LIC cloning, restriction cloning was also used to insert the CCR4-NOT genes into the MultiBac vectors. Each MultiBac vector contains two promoters under which genes may be inserted. The multiple cloning site under one promoter has been adapted to facilitate LIC cloning, whereas the second promoter still contains the multiple cloning site, thus restriction cloning is required to insert the gene fragments. Restriction digests were completed using 1 µg of the chosen vector, digested with 1 µl of each restriction enzyme (NEB) and 5 µl of CutSmart buffer (NEB) in 50 µl final reaction volume.

Reactions were typically incubated for 2 h at 37 ^{o}C unless otherwise recommended by NEB.

2.9 DNA site-directed mutagenesis

Substitutions in genes were introduced using the mega-primer extension method (Ling and Robinson, 1997). Two LIC primers, the 5′ and 3′ primers which bind the gene and contain either TACTTCCAATCCATG as part of the 5′ primers or TATCCACCTTTACTG as part of the 3′ primer, and a mutational primer which contains the desired change were used, indicated by a cross in **Figure 2.3**. The 5′ mutational primer and the 3′ LIC primer sequence were used to generate the mega-primer (megaR where one strand is black and one is red). The mega primer is then used in a second PCR reaction along with the 5′ LIC primer to amplify the whole gene fragment. The thermocycler conditions are listed below.

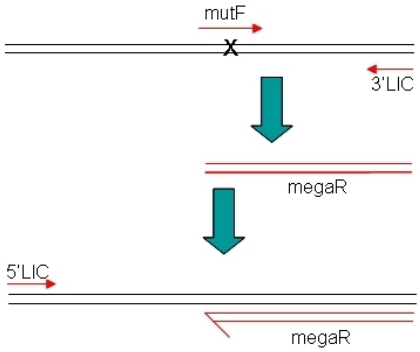


Figure 2.3 Schematic of mega-primer extension PCR

The schematic shows an example of the process of generating mutants by mega-primer extension PCR. Point of mutation is indicated by a cross. The 3′ LIC primer and 5′ mutational primer are used to amplify a section of the gene of interest. The amplified product (megaR) becomes a mega-primer in a second round of PCR. The mega-primer and the 5′ LIC primer are used to then amplify the entire gene.

For example:

First round of PCR

3′ primer & Mutagenic forward

10 x Pfx buffer (2 x)	10 µl
50 mM $MgSO_4$	1 µl

10 mM dNTP mix	1.5 µl
Platinum Pfx	0.4 µl
Water	29.1 µl
Primers (5 µM)	3.0 µl
Template (approx. 2.5 ng/µl)	5.0 µl
Final:	50 µl

Thermocycler conditions:

94 °C, 5 min

(94 °C, 30 sec; 52 °C, 30 sec; 68 °C, 1-3 min*) x 25 cycles

68 °C, 10 min

15 °C hold

(*Extension time dependent on length of PCR product – e.g. 1 min per 1 kb fragment)

The product from the first PCR was visualised on a 1.5% agarose gel (section 2.11) and PCR purified. The final product was eluted in 30 µl of TE (10 mM Tris-HCl, 1 mM disodium EDTA, pH 8.0)

Second round of PCR

5′ primer & MegaR primer

10 x Pfx buffer (2 x)	5.0 µl
50 mM $MgSO_4$	0.5 µl
10 mM dNTP mix	0.75 µl
Platinum Pfx	0.2 µl
Water	10.3 µl
Megaprimer	5 µl
5′ or 3′ LIC Primer (5 µM)	0.75 µl
Template DNA (approx. 50 ng)	2.5 µl
Final:	25 µl

Thermocycler conditions:

94 °C, 5 min

(94 °C, 30 sec; 52 °C, 1 min; 68 °C, 1-3 min*) x 15-20 cycles

68 °C, 10 min

15 °C hold

(*Extension time dependent on length of PCR product – e.g. 1 min per 1 kb fragment)

The final PCR product (25 µl) was incubated with 1 µl of DpnI (1 in 20 dilution in NEB2) to remove template DNA and subsequently purified, as instructed by the manufacturer, using a Qiagen PCR purification kit (Invitrogen). The final product was eluted in 30 µl of EB (10 mM Tris-HCl, pH 8.5) and annealed to the chosen vector.

2.10 DNA sequencing

After selection of correct clones, all expression plasmids were sequenced by Source BioScience using Sanger sequencing. For each sequencing read, 10 µl of plasmid DNA and 5 µl of primer, both in EB buffer, were required. The sequences were analysed using Contig Express Project provided by Vector NTI.

2.11 Agarose gel electrophoresis

PCR products were run on a 1.5% agarose gel to ascertain if the correct size of fragment has been generated. A 1.5% gel was chosen as this percentage readily resolves fragments of 200–3,000 bp, which encompasses the majority of targets. Small gel casts containing 50 ml agarose were made using 0.75 g of agarose (Web Scientific) resuspended in 50 ml of 1 X TAE (40 mM Tris acetate, 1 mM EDTA, pH 8.3) and microwaved until the agarose has dissolved. Large casts required 200 ml, thus 3 g of agarose was resuspended in 200 ml of 1 X TAE and microwaved until dissolved. A 1 kb plus ladder (Invitrogen)

was used for size comparison of PCR fragments. Small gels were run for <30 min at 120 V whereas larger gels were run for 1 h at 150 V.

Successfully cloned and sequence verified constructs were either transformed into BL21 (DE3)-R3-pRARE2 *E. coli* for expression testing, as detailed in section 2.16.1 or used to produce recombinant baculoviruses for the expression of protein in insect cells, as described in section 2.12.

2.12 Production of recombinant baculoviruses

Successfully cloned constructs destined for expression in insect cells (**Figure 2.4**, step 1) were transposed into a viral DNA backbone and propagated in DH10Bac *E. coli*. The resulting DNA product, known as a bacmid, and the cloned expression plasmid contain transposition sites which allow the gene of interest to transpose into the bacmid DNA. This transposition disrupts a *lacZ* gene within the transposition sites of the bacmid. *LacZ* encodes for the enzyme β-galatosidase that hydrolyses lactose to glucose and galactose (Juers et al., 2012). Disrupting of this gene prevents the enzyme from converting 5-bromo-4-chloro-3-indoyl-β-d-galactopyranoside, commercially known as X-gal (Glycosynth), to a blue product, thus colonies where the expression plasmid has successfully transposed appear white (**Figure 2.4,** step 2).

After transformation into DH10Bac *E. coli*, correctly transposed clones were selected using agar plates containing kanamycin (to select for the bacmid), tetracycline (to select for the transposase enzyme plasmid), isopropyl β-D-1-thiogalactopyranoside (IPTG) (to induce expression of *lacZ* gene), gentamycin (to select for the acceptor in the recombined plasmid) and X-gal.

The recombinant bacmid is purified from the white colonies (step 3) and used to transfect Spodoptera frugiperda (*Sf9*) insect cells (**Figure 2.4,** step 4). The bacmid DNA enters the *Sf9* cells and initiates expression of viral proteins that produce the baculovirus. The baculovirus is harvested from the medium and used to infect *Sf9* cells where the virus allows expression of the protein of interest. Each generation of viral amplification is identified with a number, where P0 indicates the virus produced from transfection and P1, P2 etc. Finally, expression levels are determination at small scale (**Figure 2.4,** step 5), detailed in section 2.16.2. The following sections (2.12.1-2.12.4) describe the methodology details of each step used to implement this procedure.

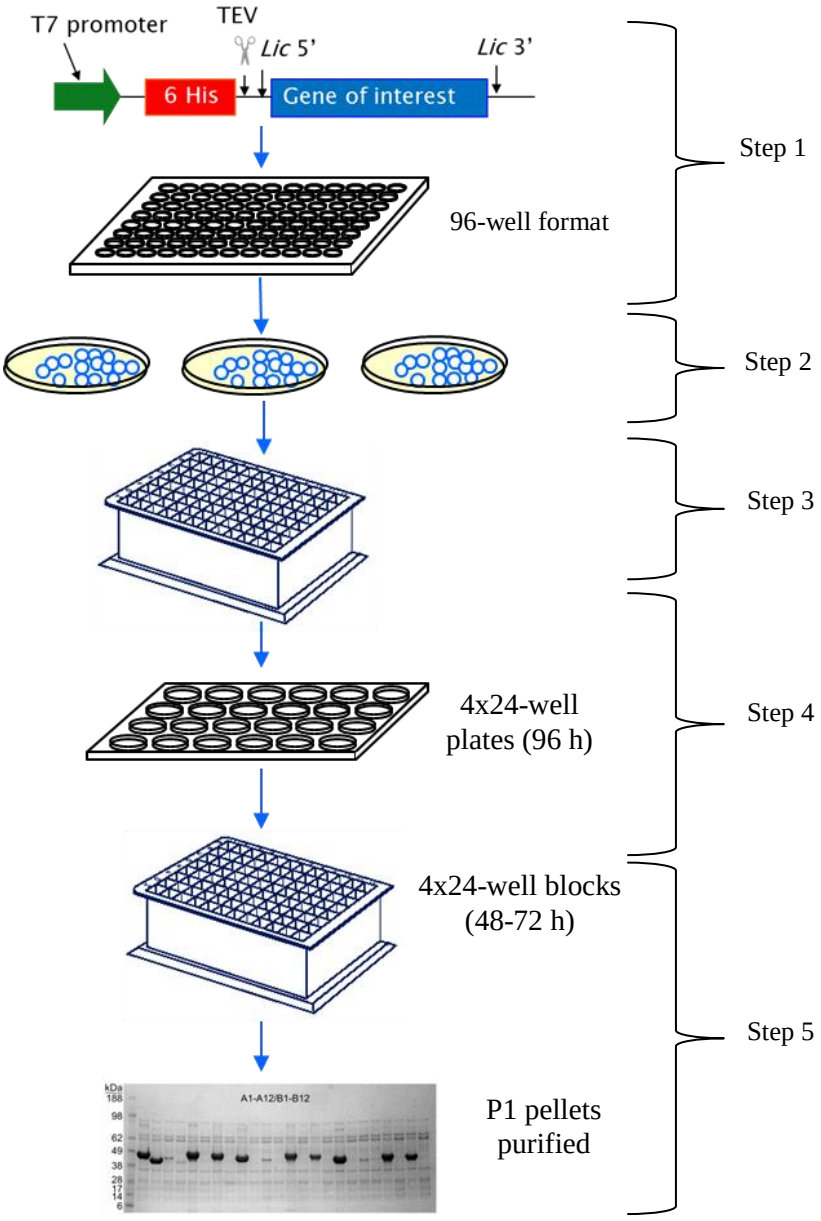


Figure 2.4 Process of the generation of baculviruses

The schematic shows the 5 steps required to generate baculoviruses in HTP. Firstly, successfully cloned constructs (step 1) were transposed into a viral DNA backbone and propagated in DH10Bac *E. coli*. The resulting DNA product, known as a bacmid, is selected for based on blue/white screening. Colonies where the expression plasmid has successfully transposed appear white (step 2). The recombinant bacmid is purified from the white colonies (step 3) and used to transfect Spodoptera frugiperda (*Sf9*) insect cells (step 4). The bacmid DNA enters the *Sf9* cells and initiates expression of viral proteins that produce the baculovirus. The baculovirus is harvested from the medium and used to infect *Sf9* cells where the virus allows expression of the protein of interest. Finally, expression levels are determination at small scale (step 5), detailed in section 2.16.2.

2.12.1 Transposition of cloned vector into bacmid DNA

3 µl of baculoviral expression plasmid was transposed into 30 µl of DH10Bac or DH10MBac to produce recombinant bacmids for viral production. Once the DNA was

added, the DH10Bac were incubated on ice for 30 min. The cells were heat shocked for 45 sec in a 42 ^{o}C water bath and returned briefly to ice. Following this, 900 µl of pre-warmed SOC medium containing 50 µg/ml kanamycin, 10 µg/ml tetracycline and 1 µg/ml gentamycin was added. The cells were placed in a Glas-col shaker at 37 ^{o}C with shaking at 700 rpm for 5 h. 50 µl of the mixture was spread onto the DH10Bac selective plates (50 µg/ml kanamycin (to select the bacmid) (Apollo Scientific Ltd); 7 µg/ml gentamycin (to select pFB-LIC-Bse plasmid), and 10 µg/ml tetracycline (to select the plasmid encoding transposase), 40 µg/ml IPTG (Apollo Scientific Ltd) and 100 µg/ml 5-Bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-gal) (Glycosynth)). The plates were incubated at 37 ^{o}C for 48 h covered with foil. X-gal was added to allow blue-white screening. The baculoviral expression plasmid was transposed into the Tn7 sites present on the bacmid contained within the DH10Bac cells. White colonies contain the recombinant DNA and the blue ones do not.

2.12.2 Bacmid purification

To purify the recombinant bacmid, the white colony was inoculated into 1 ml of 2 x LB medium containing 50 µg/ml kanamycin, 7 µg/ml gentamycin and 10 µg/ml tetracycline and incubated at 37 ^{o}C overnight at 700 rpm in a shaker (Glas-col). Cells were pelleted at 2,600 x g for 30 mins and the supernatant discarded. The bacmid was purified from cells using a Millipore miniprep plasmid kit, following manufacturer′s instructions.

2.12.3 Bacmid PCR Screen

To confirm that the gene of interest had been successfully transposed into the purified bacmid, the DNA was screened using PCR. The bacmids were diluted 50-fold in water and screened by PCR using FBAC-1 and FBAC-2 or gene-specific primers (**Appendix Table 7.2 and Table 7.3**). Each individual reaction contained; 4 µl of Bioline 5x Buffer,

14.9 µl of nuclease-free water, 1.0 µl of primer mix (10 µM), 0.1 µl of Bioline MyTaq DNA polymerase and 1 µl of diluted bacmid.

Thermocycler conditions:

5 min 94 ^{o}C

94 ^{o}C for 45 sec

50 ^{o}C for 45 sec

72 ^{o}C for 2.5 min for 30 cycles

5 min at 72 ^{o}C.

The PCR fragment was visualised by agarose gel and a fragment of the size of the fragment cloned should be detected.

2.12.4 Transfection of insect cells with recombinant bacmid

Once the bacmid has been confirmed to contain the gene of interest, it is then transfected into insect cells to generate baculoviruses which cause the expression of the genes of interest. A volume of 1 ml ($2x10^5$ cells) of *Sf9* cells were required for each transfection. The cells were incubated at 27 ^{o}C for 1 h to allow cell attachment. 200 µl of jetPRIME (Polyplus) was added into 5 ml of jetPRIME buffer in a 15 ml tube and gently vortexed for 10 sec to mix. 50 µl of the mixture was dispensed into a sterile 96-well microtitre plate. 5 µl of recombinant bacmid DNA (at >0.5 µg/µl) was transferred into each well and mixed by shaking the plate gently. The DNA:jetPRIME mixture was incubated inside the laminar flow hood for 30 min. 50 µl of the DNA:jetPRIME mixture was added to the cells and incubated for 4 h at 27 ^{o}C. Controls included cells with jetPRIME only and one well for cells only. After 4 h, an additional 0.8 ml of Sf-900 II serum free medium (SFM) (Life Technologies) containing 2% (v/v) fetal bovine serum (FBS) (Life Technologies) and antibiotics Pen/Strep 0.1% (v/v) (Pen Strep=50 U penicillin and 50 µg streptomycin per ml medium) was added to each well. Cells were incubated at 27 ^{o}C for

6 days. Afterwards, the liquid contents were transferred to a sterile tube and centrifuged at 1,500 x g for 20 min at room temperature. The clear supernatant (P0 virus) (<700 µl) was collected.

2.13 MultiBac expression system

The MultiBac system was developed for the generation of protein:protein complexes in insect cells (Bieniossek et al., 2008, Fitzgerald et al., 2006), based on the original baculoviral technology described in section 2.12. The system allows genes of interest to be cloned into specific vectors, termed acceptors and donors (**Figure 2.5**).

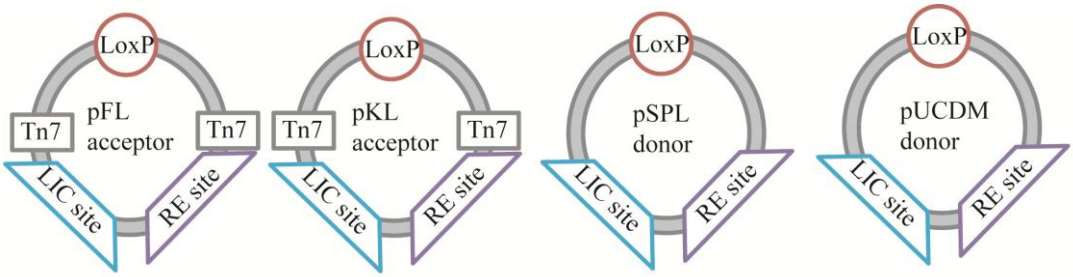


Figure 2.5 MultiBac vectors

Schematic shows the four MultiBac vectors into which the genes of interest can be cloned. Each vector has both a site which facilitates LIC cloning (blue) and restriction cloning (purple) and a LoxP site (red) which allows site-specific recombination of multiple vectors. Each acceptor also contains Tn7 transposable sites (grey) for transposition into the recombinant bacmid.

The vectors, supplied by Prof Imre Berger (University of Bristol), contain two promoters, polH and p10, thus allowing two genes to be expressed by each plasmid. The site under polH promotion was adapted to facilitate LIC cloning (section 2.7) and restriction cloning was available in the second site, under p10 promotion. In order to generate MultiBac baculoviruses which express the subunits of the CCR4-NOT complex, full length alleles of CNOT1, CNOT3, CNOT7, CNOT9, CNOT11 and TAB182 were cloned using LIC into MultiBac vectors under the polH promoter as described above. Full length CNOT9, CNOT2, CNOT6L and CNOT10 were restriction cloned under the p10 promoter. A summary table of the cloning strategy utilised is shown in **Table 5.1**.

Once the discrete vectors were cloned, a Cre-LoxP driven recombination allows individual vectors to combine into one multigene vector as described (Fitzgerald et al., 2006). Acceptor (1 µg) and donor(s) (1.5 µg) plasmids were recombined using Cre recombinase (1.5 µl, NEB) in the presence of Cre recombinase buffer (2 µl, NEB) made up to a total volume of 20 µl. Each complex included a single acceptor vector, containing the Tn7 attachment sites for recombinant bacmid generation, and at least one donor vector, containing a distinct antibiotic resistance cassette on a vector with a PI structural gene (pir)-dependent origin of replication (R6Kγ). Isolation of the multigene vector was done by transformation of the entire recombination reaction into 100 µl of a pir^- strain (MACH1) and with antibiotic selection for each component vector, only those that have recombined were able to express each resistance cassette in the absence of the *pir* gene. DNA was prepared of the recombined vector, which was then transposed and transfected as described in section 2.15 and 2.18.

2.14 SDS-PAGE

15 µl of sample protein was mixed with 5 µl of SDS-PAGE loading buffer (3 parts Novex sample buffer, 1 part DTT (Thermo Fischer Scientific). Samples were boiled at 80 ^oC for 10 min before loading onto 4–12% Bis-Tris Novex® NuPAGE® SDS-PAGE gels (Thermo Fischer Scientific). 5 µl of SeeBlue Plus 2 molecular weight marker (Thermo Fischer Scientific) was also loaded. Bands were resolved at 150 V for 1 h, stained with InstantBlue protein stain (Expedeon), and visualised with a Bio-Rad gel doc imager.

2.15 Buffers used for purification

2.15.1 Buffers for purification of *E. coli*-expressed recombinant proteins

Binding buffer: 50 mM HEPES pH 7.5, 300 mM NaCl, 5% glycerol, 20 mM Imidazole, protease inhibitors (1:1000 v/v), 1 mM TCEP

Wash buffer: 50 mM HEPES pH 7.5, 300 mM NaCl, 5% glycerol, 40 mM Imidazole, 1 mM TCEP

Elution buffer: 50 mM HEPES pH 7.5, 300 mM NaCl, 5% glycerol, 300 mM Imidazole, 1 mM TCEP

Gel filtration buffer: 50 mM HEPES pH 7.5, 300 mM NaCl, 5% glycerol, 1 mM TCEP

2.15.2 Buffers for purification of Baculovirus-expressed recombinant proteins

Binding buffer: 50 mM HEPES pH 7.5, 500 mM NaCl, 5% glycerol, 20 mM Imidazole, protease inhibitors (1:1000 v/v), 1 mM TCEP

Wash buffer: 50 mM HEPES pH 7.5, 500 mM NaCl, 5% glycerol, 40 mM Imidazole, 1 mM TCEP

Elution buffer: 50 mM HEPES pH 7.5, 500 mM NaCl, 5% glycerol, 300 mM Imidazole, 1 mM TCEP

Gel filtration buffer: 50 mM HEPES pH 7.5, 500 mM NaCl, 5% glycerol, 1 mM TCEP

2.16 Small scale expression/purification tests

2.16.1 *E. coli*

Once constructed, the *E.coli* expression plasmids were tested for expression to determine which were worth scaling to litre-scale growth. Expression plasmids were used to transform *E. coli* BL21 (DE3)-R3-pRARE2 competent cells by a heat-shock procedure. Three colonies were inoculated into 1 ml of 2 x Luria Broth (LB) medium containing 50 µg/ml kanamycin and 34 µg/ml chloramphenicol (AppliChem) in a 96-deep well block. The starter cultures were incubated in a Glas-Col shaker (overnight at 37 °C, with shaking at 700 rpm). The following morning, 20 µl of the culture was used to inoculate 1 ml of complete Terrific Broth (TB) medium containing 50 µg/ml kanamycin in two 96-deep-well blocks; one for monitoring the optical density (OD) and one for purification tests. Cultures were incubated (37 °C and 700 rpm) until the absorbance at 600 nm was 2-3, after which they were removed from the incubator and placed stationary at room

temperature for 30 min while the incubator cooled to 18 °C. Expression was induced by adding 0.1 mM IPTG to each well. The flasks were then returned to the incubator and bacteria cultured overnight at 18 °C, with shaking at 700 rpm. Cells were harvested at 3000 x *g* for 20 min and re-suspended in 200 µl of binding buffer containing 0.1% DDM and 0.5 mg/ml lysozyme. The cells were freeze-thawed at -80 °C for 20 min to aid lysis. Cell debris was removed by centrifugation at 3000 x *g* for 10 min. The supernatant was incubated with 50 µl of 50% (v/v) Nickel Sepharose 6 FF (GE Healthcare) at 18 °C for 60 min with shaking at 400 rpm. A vacuum manifold was used to remove the flow through and the resin was washed three times with 200 µl of wash buffer. The plate was incubated with 40 µl of elution buffer for 20 min at 18 °C, shaking at 400 rpm. Bound protein was eluted by centrifugation at 300 x *g* for 3 min. Eluates were analysed by SDS-PAGE to ascertain solubility.

2.16.2 Baculovirus/insect cells

A similar procedure of expression testing was performed for baculovirus-mediated expression. *Sf9* cells cultured at a density of 2 x 10^6 cells/ml in 3 ml were seeded into four 24-deep well blocks. The cells were infected with 120 µl of P0 baculovirus stock in Sf 900-II SFM medium containing 2% (v/v) FBS and incubated at 27 °C, with shaking at 450 rpm in a Glas-Col shaker for 66-72 h. Cells were pelleted by centrifugation at 1500 x *g* for 20 min and the supernatant was harvested. Pellets were washed with 1 ml of PBS (Sigma Aldrich) and resuspended in 1 ml of binding buffer. The pellets were sonicated, transferred into a 96-deep well block and centrifuged at 2600 x *g* for 30 min at 4 °C. To the clarified supernatant in each well, 100 µl of 50% (v/v) Nickel Sepharose 6 FF (GE Healthcare) pre-equilibrated in binding buffer were added. The block was incubated at 18 °C for 1 h, with shaking at 90 rpm. The beads were washed once with 800 µl of binding buffer and three times with 800 µl of wash buffer. The protein was eluted with 50 µl of

elution buffer. Eluates were analysed by SDS-PAGE to ascertain solubility.

2.17 Large scale expression of recombinant proteins

2.17.1 *E. coli*

Based on small-scale tests, constructs that showed a soluble and well-expressed band were followed up to explore the interactions between the CCR4-NOT subunits and associated proteins. Expression plasmids were used to transform *E. coli* BL21 (DE3)-R3-pRARE2 competent cells by a heat-shock procedure. To prepare a starter culture, approximately three colonies were inoculated into 10 ml of LB medium per 1 L of culture to be inoculated containing 50 µg/ml kanamycin and grown overnight at 37 °C, shaking at 200 rpm. Each 10 ml overnight culture was used to inoculate 1 L of complete TB medium supplemented with 50 µg/ml kanamycin in 3 L flasks, and this was grown until the absorbance at 600 nm was 1-2. Typically, proteins were expressed using 6 L of *E. coli*. Once the cultures reached the desired absorbance (~ 6 h), they were removed from the incubator and placed stationary at room temperature while the incubator cooled to 18 °C. Expression was induced by adding 0.1 mM IPTG to each 1 L flask. The flasks were returned to the incubator and grown overnight at 18 °C, with shaking at 200 rpm.

Constructs which had been sub-cloned into a pNIC-Bio3 vector were destined for analysis of protein-protein interactions by BLI. These proteins required biotinylation to allow attachment to the streptavidin biosensor. The constructs were biotinylated in vivo by transformation into *E. coli* express strain BL21 (DE3)-R3-pRARE2 BirA, which contains a plasmid encoding BirA protein which transfers biotin to a biotinylation recognition sequence that is fused to the recombinant protein when expressed from a pNIC-Bio3 vector. Cultures are grown as described above but growth is supplemented

twice with 0.1 mM biotin, firstly at the point of induction and secondly, 1 hour prior to harvesting. The biotinylated protein is purified as described in section 2.18.

2.17.2 Baculovirus/insect cells

To amplify the virus in sufficient amount and titre for large scale expression, several rounds of passages were needed. 120 µL of P0 was added to 50 mL of *Sf9* cells and incubated at 27 ºC for 66-72 h. The cells were centrifuged at 900 x g for 20 min. The media was collected and became the P2 virus. 3 ml of P2 baculovirus was added to each 1 L flask of *Sf9* cells supplemented with 1% FBS. The cells were incubated at 27 °C for 48 h and harvested by centrifugation at 1500 x *g* for 20 min. Cell resuspension and purification was performed as described for *E. coli* (section 2.18).

2.18 Purification of recombinant CNOT subunits

2.18.1 Lysis

Cells grown from *E. coli* and insect cell expression were harvested and re-suspended in 2 x binding buffer. Typically, cell pellets were resuspended in a buffer volume equal to the weight of the pellet e.g. a pellet weighing 20 g would be resuspended in 20 ml of binding buffer. Cells were sonicated using a 750 Watt Ultrasonic Processor probe (Sonics Materials) (program 10 sec off, 5 sec on) and cell debris was removed by centrifugation at 30 000 x *g* for 20 min.

2.18.2 Nickel-affinity resin purification

The clarified supernatant containing the soluble cell fraction was incubated with 1 ml of 50% (v/v) Nickel Sepharose 6 FF (GE Life sciences) at 4 °C for 60 min in a 50 ml falcon tube while rotating. The resin was washed with 100 column volumes (CV) of 1 x binding

buffer, and then 100 CV of wash buffer to remove non-specifically bound contaminating proteins. Bound His_6-tagged protein was eluted in 5 fractions of 5 ml using elution buffer. Eluates were analysed by SDS-PAGE to ascertain solubility of the target protein.

2.18.3 Preparative gel filtration

Protein samples which had been captured by nickel affinity chromatography were further purified using preparative gel filtration to separate contaminating proteins based on size. Either S75 or S200 HiLoad 16/60 Superdex (GE Healthcare, CV = 120 ml) columns were used alongside an ÄKTA-Express system (GE Healthcare). Prior to applying the protein, the Superdex column was washed with 2 CV and equilibrated with 1.2 CV of gel filtration buffer. The partially purified protein was concentrated to <5 ml using an Amicon centrifugal device (MWCO 10 kDa), filtered by molecular weight, and loaded onto the column by injection. 1.5 ml fractions were collected using a flow-rate of 1 ml/min. Samples from peak fractions were visualised using SDS-PAGE and protein-containing fractions were pooled.

2.18.4 Analytical gel filtration

Analytical gel filtration was used either to characterise the oligomeric state of a protein or to explore the formation of protein complexes. Purified protein samples were run an analytical gel filtration columns (S200 10/300 GL, CV = 24 ml) at a flow rate of 0.3 ml/min where 0.2 ml fractions were collected. Samples from peak fractions were visualised by SDS-PAGE. An estimated molecular weight of protein and protein complexes were calculated based on the behaviour of the complex during gel filtration in comparison to molecular weight standards (Bio-Rad). The identity and molecular weight of the standards is detailed in **Table 2.4**.

Component	Molecular weight (Da)
Thyroglobulin (bovine)	670,000
γ-globulin (bovine)	158,000
Ovalbumin (chicken)	44,000
Myoglobin (horse)	17,000
Vitamin B12	1,350

Table 2.4 Identities and molecular weights of analytical gel filtration standards

Five molecular weight standards ranging from 670 kDa to 1.3 kDa were used to calculate molecular weight estimates for proteins and complexes.

These molecular weights gave some insight into the possible oligomeric state of proteins or protein complexes. The estimated molecular weight was calculated by comparing the log of elution volume to a standard graph where log of elution volume was plotted against K_{av}. K_{av} is calculated using the following equation:

$$K_{av} = \frac{(V_e - V_o)}{(V_t - V_o)}$$

where V_e is the elution volume, V_o is the void volume and V_t is the total bed volume of the column.

2.18.5 Cleavage of vector-incorporated fusion tags

All vectors were designed to fuse a His_6 tag with the recombinant protein to allow nickel affinity capture but this tag was often removed to ensure it did not disrupt possible crystal contacts important to crystallisation. Furthermore, cleavage of the tag provided an additional purification step, thus produced more homogeneous protein preparations as contaminations which specifically bind nickel resin were captured in the first step, would be removed from the preparation at this stage.

In order to remove the fusion tag from the protein of interest, TEV protease were added to purified protein (1: 20 mass ratio of enzyme to protein) and incubated overnight at 4 °C. The following morning, 0.5 ml 50% (v/v) Nickel Sepharose 6 FF was added at 4 °C for 60 min in a 50 ml falcon tube while rotating. The nickel resin acts to both re-bind the

cleaved tag and the TEV protease which is expressed recombinantly in *E. coli*, fused to a His_6 tag thereby allowing separation of the protease. As the CNOT protein is no longer tagged, it will pass through the column without interacting with the resin, thus it will be collected in the flow through. The resin was washed with 20 CV of gel filtration buffer, followed by 20 CV of 1 x binding buffer, and then 20 CV of wash buffer to remove the protein of interest which may still non-specifically bind the resin. The TEV protease which specifically interacts with the resin due to the His_6 tag was eluted using 5 CV of elution buffer. All samples were analysed by SDS-PAGE to ascertain in which fraction the protein eluted and protein-containing fractions were pooled.

2.19 Mass spectrometry

2.19.1 Denaturing intact mass spectrometry

After each purification, denaturing intact mass spectrometry (MS) was used to ensure the observed and expected masses of the protein are the same and therefore check for post-translational modifications, mutations etc. Intact protein mass determination was completed using 1 µl of protein sample resuspended in 59 µl of 0.1% formic acid which is then analysed using LC-QTOF instruments (Agilent). Protein mass was measured to within +/- 0.5 Da and compared to the expected mass derived from the construct sequence of the target protein. Where applicable, the masses observed from intact MS were mapped onto sequences of CNOT proteins to predict the amino acid boundaries of the protein.

2.19.2 Tryptic digest mass spectrometry analysis of SDS-PAGE gel bands

The identity of a protein can be established using SDS-PAGE, a technique which separates proteins based on size, by comparing the migration of the protein of interest

with that of molecular weight standards. However, the identity of the SDS-PAGE protein band may also be determined using tryptic digest mass spectrometry (tandem MS). Bands of interest from SDS-PAGE gels were excised from instant blue stained SDS-PAGE gels and analysed by tandem MS, where digestion and analysis were completed in collaboration with the MS facility at the SGC. The protein within the gel band was prepared for analysis by irreversibly breaking disulphide bonds via reduction using 100 mM dithiothrietol (DTT) and alkylation using saturated iodoacetamide. Following this, the protein was digested into peptide fragments overnight using 50 µl of 0.0032 mg/ml trypsin. The resulting peptide mixture was separated by reversed phase HPLC and passed directly to an ion trap system (Bruker). The peptide fragmentation data was used to search a protein sequence database using the Mascot search algorithm (Perkins et al., 1999). Mascot performs in-silico digestion of all proteins in the database and generates a table of theoretical fragment ion masses for each tryptic peptide. This allows identification of the proteins residing within the submitted gel band.

2.20 Crystallisation

Protein purified as above, validated for its homogeneity, was concentrated to >10 mg/ml for crystallisation trials, using the sitting drop vapour diffusion method.

2.20.1 Sitting drop vapour diffusion

Sitting drop vapour diffusion is a technique which relies on the equilibration of the protein-containing drop and the larger reservoir solution in a sealed environment. Water diffuses from the protein-containing drop to the reservoir solution until the precipitant concentration becomes equal. Diffusion from the drop causes shrinkage, increasing both the precipitant and protein concentration, allowing nucleation.

Initial crystallisation trials were performed in Swissci plates (Hampton Research) composed of 96 reservoirs each serving 3 sub-wells. Sampling of crystallisation conditions involves commercial random matrix screens that include Hampton crystal screen (HCS), Hampton Index screen (HIN) (Hampton Research) and Joint Centre for Structural Genomics (JCSG) (Molecular Dimensions). To set up the plate, 20 µl of each precipitant is dispensed into the reservoir followed by 50, 75 and 100 nl of protein in the 3 sub-wells using a Mosquito liquid handler (TTP Lab Tech). The mosquito is used to mix the reservoir precipitant with the dispensed protein in 3 protein:precipitant ratios (1:2, 1:1 and 2:1) with a final volume of 150 nl. The plates were incubated at 4 °C or 20 °C and monitored by regular images taken at timed intervals using a Minstrel crystallisation imager (Rigaku).

2.20.2 Optimisation of crystallisation trials

Initial observations were followed up using two approaches.

1. The precipitant composition found to produce initial observations was expanded upon in a follow-up experiment. 96 precipitants were designed whereby the concentration of the pH buffer and precipitant (e.g. polyethylene glycol (PEG)) were varied. For example, if an initial observation was found in 0.1 M $CdCl_2$ 0.1 M CH_3COONa, pH 4.6, 30% PEG 400, a follow-up screen may contain 24-32% PEG 400 at pH 3.5-6.

2. Seeding was used to optimise small or poorly formed crystals. The crystal seeds are generated by crushing the micro-crystals initially observed to form tiny fragments. The seeds are dispensed into crystallisation screens, along with the protein and precipitant where they act as nucleation sites to initiate crystal growth. Precipitant conditions which facilitate a metastable zone will allow the seeds to grow into full crystals. This removes

the need to find a condition which first allows initial nucleation, then growth thus crystals may grow more quickly.

Crystals of CNOT10 and CNOT11 were removed from the crystallisation plate using a pipette and resuspended in 100-500 µl of mother liquor, depending on density of crystals. A plastic bead (Hampton Research) was added an Eppendorf containing the crystals of CNOT10 and CNOT11 and the crystals were vortexed vigorously to produce the seed stock. 20 nl of the 100% seed stock was added to a crystallisation screen which was prepared as normal. The seeds were added to both the screen where the microcrystals were initially found and an optimisation screen to screen for conditions which facilitate a metastable zone, indicated by growth of seeds into crystals.

2.21 Interaction mapping using SPOT peptide arrays

As it was shown, using BLI, that motifs throughout the sequence of TTP were involved in binding both CNOT2 and CNOT9, and full length TTP was found difficult to produce, peptide arrays were used to analyse the regions within TTP that bind CNOT2 and CNOT9. Peptide arrays consisted of a cellulose membrane onto which 15 residue TTP peptides were printed in a grid format of spots. The membrane was exposed to the potential interacting protein and the location of any bound protein was detected by antibodies.

Cellulose-bound peptide arrays were prepared by S. Picaud (SGC) using a MultiPep-RSi-Spotter (INTAVIS) employing Fmoc solid phase peptide synthesis according to the SPOT synthesis method (Frank, 2002) and the manufacturer′s instructions. An array of 15-mer peptides covering the human TTP sequence (109 peptides with a 3 residues overlap throughout the sequence of TTP; list peptide sequences in **Appendix Table 7.15**) (UniProt accession code P26651) were synthesised on Amino-PEG500-UC540 sheets

optimised for use with the MultiPep instruments (INTAVIS). The presence of SPOTed peptides was confirmed using ultraviolet light (λ = 280 nM) to spot on the membrane.

For the SPOT interaction assay, membranes with SPOTed peptides were blocked with 5% BSA overnight and then incubated overnight with purified His_6-tagged recombinant CNOT2/ CNOT9 at 4 ˚C (1 μM). After washing with PBST buffer (3.2 mM Na_2HPO_4, 0.5 mM KH_2PO_4, 1.3 mM KCl, 135 mM NaCl and 0.1% Tween 20, pH 7.4) to remove any unbound material, bound protein was detected using a horseradish peroxidase (HPR)-conjugated antibody (Novagene) and the Pierce® ECL Western blotting substrate (Thermo Scientific). Chemiluminescence was detected with an image reader (Fujifilm LAS-4000 ver.2.0) typically using an incremental exposure time of 5 min for a total of 60 min. Peptide locations on the array and their sequences are given in **Appendix Table 7.15 and Table 7.16**.

2.22 BioLayer interferometry

BioLayer Interferometry (BLI) is an optical analytic technique used to measure protein-protein, protein-peptide and protein-small molecule interactions. In the context of this study, it was first utilised to determine which domain is TTP interacts with CNOT9 and CNOT2, and later used to characterise TTP and CNOT9 mutants to explore the binding mechanism of these two proteins.

BLI experiments measure the interference of white light reflected from two surfaces; a layer of biotinylated protein immobilised on a streptavidin-coated biosensor (**Figure 2.6**) and an internal reference. Any change in the layer of protein bound to the biosensor (e.g. when CNOT9 binds TTP in **Figure 2.6**), will cause a change in the wavelength which is measured in real-time (Shah and Duncan, 2014). BLI experiments were performed on a 16-channel ForteBio Octet RED384 instrument at 25 °C in 50 mM HEPES, pH 7.5 and 300 mM NaCl buffer. The technique requires one biotinylated element which, in the case of this project was TTP. TTP was biotinylated in vivo, as described in section 2.17.1. C-terminal biotinylated TTP protein (sub-cloned using pNIC-Bio3 vector) or C- or N-terminal TTP-derived peptides (synthesised by Lifetein) were attached to streptavidin coated biosensors (super-streptavidin) by incubating for 6 min at 3 μM concentration as shown in **Figure 2.6**.

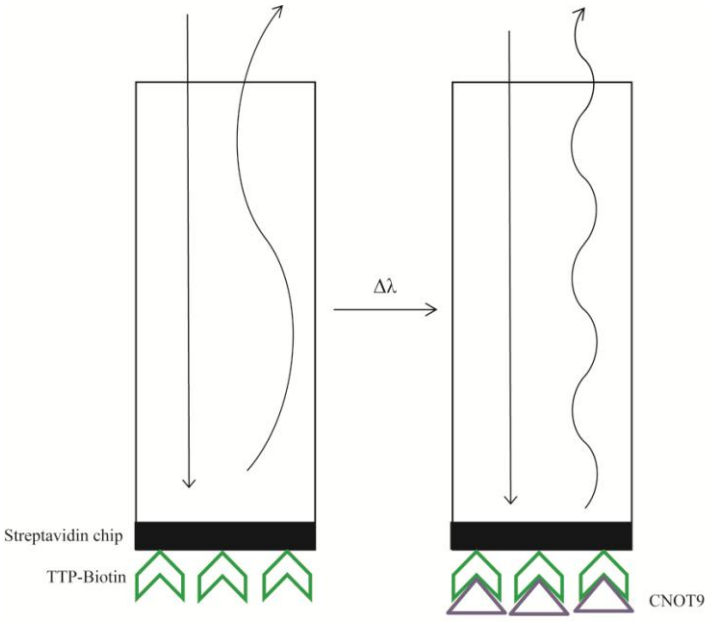


Figure 2.6 Schematic of BLI

BLI uses streptavidin-coated biosensors to capture biotinylated proteins (e.g. TTP). The wavelength of light reflected from inside the probe is taken as a measure of the thickness of protein bound to the biosensor. When the probe is exposed to an analyte that binds (e.g. CNOT9), the thickness of protein on the biosensor changes the reflected wavelength and can be quantified.

For measurements to determine the dissociation constant ($K_{d,}$ which represents the concentration of analyte at which 50% of the binding sites are occupied at equilibrium), His_6-CNOT9 sample was prepared in seven 2-fold dilutions starting from 20 or 80 μM. Measurements were performed using a 180 sec association step followed by a 180 sec

dissociation step on a black 384-well plate with tilted bottom (ForteBio). The baseline was stabilised for 120 sec prior to association and signal from the reference sensors was subtracted prior to K_d calculations.

The change in wavelength (nm) measured for each concentration of CNOT9 was plotted against CNOT9 concentration. This plot was used to derive the K_d of CNOT9 for each TTP construct using GraphPad prism and the following equation.

$$Y = \frac{Bmax\ X}{(Kd+X)} + NS*X + Background$$

Where Bmax is the maximum specific binding in the same units as Y. K_d is the equilibrium binding constant. NS is the slope of non-specific binding in Y units divided by X units. Background is the amount of non-specific binding with no added substrate.

2.23 Streptavidin-mediated purification

In order to determine the purity of biotinylated TTP proteins used in BLI experiments, where they will be further enriched using the streptavidin coated biosensor, a streptavidin-mediated purification was performed, based on the well-characterised interaction between streptavidin and biotin (Chivers et al., 2011). 100 µg of the bait protein (biotin-TTP) was incubated with 35 µl of prewashed 50% strepavidin agarose (Thermo Fisher) in a 1.5 ml Eppendorf tube that was rotated at 4 ^{o}C for 1 h. The reaction was washed three times with 1 ml wash buffer. 100 µg of target protein (control proteins/TTP) was added to the Eppendorf tube, and the reaction was incubated in the Glas-Col shaker at 4 ^{o}C overnight at 400 rpm. The reaction was washed three times with wash buffer. To determine if the prey protein has been captured by the bait, 20 µl of protein Novex sample buffer was added to the beads, boiled and visualised using SDS-

PAGE. The observation of two bands on SDS-PAGE gel indicates that the bait has captured the prey proteins. The identity of the prey can be confirmed by tandem MS.

2.24 Deadenylase activity assay

To determine if reconstituted CCR4-NOT sub-complexes were biologically active, the deadenylase activity assay was completed as in published literature (Maryati et al., 2014). 0.4 µM enzyme complex, expressed in insect cells and purified using nickel affinity chromatography, was incubated with 1 µM (final concentration) Flc-labelled RNA substrate (Eurogenetec) in a deadenylase reaction buffer (20 mM Tris-HCl pH 7.9, 50 mM NaCl, 2 mM $MgCl^{2+}$, 10% glycerol, 1 mM DTT) (10 µl final volume) for 60 min at 30 ^{o}C. The reaction was terminated by the addition of 12 µl NOVEX TBE-Urea RNA loading buffer (Thermo Fisher) and heated for 3 min at 85 ^{o}C. 8 µl of mixture was analysed by denaturing PAGE using 20% acrylamide:bisacrylamide (19:1) containing 50% (w/v) urea (Bio-Rad). Polyacrylamide gels were run for 1-1.5 h at 200 V and the labelled substrate was visualised using a Fujifilm LAS-400 imager (460 nm) based on appearance of bands.

2.25 Limited proteolysis

2.25.1 Small analytical scale

Limited proteolysis was performed on a purified sample of CNOT10:CNOT11 to determine the minimal interacting regions between the complex (chapter 3, section 3.8). The rationale was that regions involved in the interaction would likely be shielded from proteolysis, and thus protected. These protected domains may be identified by SDS-PAGE and intact MS. 20 µl of 15.3 mg/ml purified CNOT10:CNOT11 complex was diluted to 200 µl in gel filtration buffer and aliquoted into two fractions each of 100 µl

(final 150 mg per reaction) to investigate the proteolysis by both trypsin and chymotrypsin. To one fraction, trypsin was added and to the other chymotrypsin at an enzyme-to-protein ratio of 1:400. The proteolysis reaction was incubated at room temperature. A 15 µl sample for SDS-PAGE and a 1 µl sample for denaturing MS were taken prior to protease addition and at the following time points: 15 min, 30 min, 1 h, 2 h, 3 h. Protein bands separated by SDS-PAGE were extracted and tryptic digest provided a selection of peptides to assist analysis of amino acid boundaries.

2.25.2 Large preparative scale

Preparative proteolysis was used to generate the stable interacting fragments of CNOT10 and CNOT11 established at small scale in a greater quantity sufficient for protein crystallisation. To 75 µl of 15 mg/ml purified CNOT10:CNOT11 complex, 1:400 (enzyme-to-protein) trypsin was added and the reaction was incubated at room temperature for 4 h. Protease inhibitor was added at a 1:150 ratio of inhibitor to protein concentration (Merck) to stop the reaction. The protein was diluted to 300 µl with gel filtration buffer and loaded onto an S200 10/300 column for analytical gel filtration, to separate the complex of CNOT10:CNOT11 from the smaller proteolysed fragments which provided a more homogenous sample for crystallisation.

3. Characterising the interaction between CNOT10 and CNOT11 generated from high-throughput expression

3.1 Introduction

The CCR4-NOT complex is comprised of eight stably associated CNOT subunits where CNOT1 forms the platform onto which CNOT2-11 bind. Although the function of some CNOT subunits (e.g. CNOT1, CNOT9 and the deadenylase subunits) is partly understood, fundamental questions remain concerning certain subunits, primarily the recently identified CNOT10 and CNOT11. CNOT10 was recognised as a CNOT subunit by its discovery in a number of independent purifications of the human CCR4-NOT complex (Lau et al., 2009, Morita et al., 2007, Schwede et al., 2008). An interaction between human CNOT10 and C2ORF29 was identified by a yeast-2-hybrid (Y2H) screen, which prompted C2ORF29 to be named CNOT11. The Y2H screen revealed that the complex of CNOT10 and CNOT11 binds CNOT1 via an interaction with CNOT11 (Mauxion et al., 2013).

Beyond this finding, little information is known about the function and positioning of the CNOT10:CNOT11 module within the other subunits of the human CCR4-NOT complex. Inferences may be taken from the *T. brucei* (*Tb*) CCR4-NOT complex where the function of *Tb*CNOT10 has started to be characterised. However, as these proteins are only 31% similar at the sequence level, some differences in function may be discovered. The function of *Tb*CNOT10 was analysed using the migration of CNOT complexes through a glycerol gradient. When *Tb*CNOT10 was depleted from cells by siRNA knockdown, the isolated complexes also lacked *Tb*CNOT7. This observation was further explored using a Y2H screen which identified a direct interaction between *Tb*CNOT10 and *Tb*CNOT7

(Farber et al., 2013). Although an equivalent interaction has not been observed in the human CCR4-NOT complex, these results indicate that *Tb*CNOT10 may be integral to the formation of the *Tb*CCR4-NOT complex. While CNOT10 and CNOT11 are widely conserved across eukaryotes, no homologous protein is predicted in the yeast genome so vital structural information on this module is absent from the recent yeast CCR4-NOT cryo-EM models (Ukleja et al., 2016a, Nasertorabi et al., 2011).

Thus, this study aimed to characterise the individual CNOT10 and CNOT11 subunits alone and as part of the CNOT10:CNOT11 complex in the human system. Any structural or biochemical data focusing on CNOT10:CNOT11 will embellish our understanding of human CCR4-NOT complex. To achieve this, crystallisation trials, limited proteolysis and MS were enlisted to gain further insight into the motifs required to form the CNOT10:CNOT11 module and how it may be recruited to the CCR4-NOT complex.

3.2 Construct design

To study the interactions between CCR4-NOT subunits in vitro, soluble recombinant proteins were generated for use in crystallisation experiments, complex reconstitution and activity assays. The design of expression constructs was assisted firstly by amino acid boundaries detailed in the literature which produced soluble proteins. Secondly, bioinformatic analyses were used to predict secondary structure composition and disordered areas of the protein. Full length CNOT10 and an N-terminal truncation of CNOT11 (aa 61-510) were used in Y2H screening, so these boundaries were replicated. For bioinformatic analysis of CNOT10 and CNOT11, the psiPRED suite was chosen which incorporates tools to predict secondary structure and protein disorder (http://bioinf.cs.ucl.ac.uk/psipred, (Buchan et al., 2013)). CNOT10 is predicted to contain mainly helical regions (blue box) with two beta-strands (red arrows) between aa 421-425

and aa 432-438 (**Figure 3.1**). DISOpred, a disordered region prediction tool within the psiPRED suite, highlighted seven regions of disorder throughout the sequence of CNOT10 (purple box) (aa 1-37, aa 187-206, aa 338-343, aa 441-455, aa 478-497, aa 578-607, aa 687-717) (**Figure 3.1**). The DISOpred software is also trained to highlight disordered areas which are predicted to be involved in protein binding regions (Jones and Cozzetto, 2015). The significant stretches of disorder in the N- and C-termini (aa 1-37 and aa 687-717, pink text) are predicted by DISOpred to be protein binding. Hence, these regions may be of significance to the interaction with CNOT11 and were considered during construct design. A search using the PFAM database was completed (http://pfam.xfam.org/, (Finn et al., 2016)). The PFAM database contains information on

Figure 3.1 Output of secondary structure prediction tools for CNOT10

psiPRED and DISOpred were used to provide predictions for the composition of secondary structure elements within the CNOT10 sequence. The predictions were used to influence the design of CNOT10 truncations. Regions predicted to be helical have been boxed in blue and regions predicted to contain disordered stretches are boxed in purple. Two regions have been predicted to form beta strands and these are highlighted with a red arrow. The pink text indicates the disordered regions which are predicted to be involved with protein binding. The figure has been annotated with coloured triangles to indicate a selection of domain boundaries designed for CNOT11. Refer to **Figure 3.4** for full list of constructs and their respective expression levels.

a large number of protein families which can be probed for the discovery of sequences that have a similar sequence to known protein domains. Thus, hypothetical domains and possible protein functions can be postulated for unknown proteins. Unfortunately, the PFAM search did not retrieve any predicted domains for CNOT10.

psiPRED predicts CNOT11 to contain helical regions (blue box) with four regions predicted to be disordered (purple box) (aa 1-58, aa 164-172, aa 252-271, aa 499-510) (**Figure 3.2**). The N-terminal 60 residues annotated as disordered were not present in the CNOT11 construct which bound CNOT10 and CNOT1 in the Y2H screen (Mauxion et al., 2013), hence this region may not be important to facilitate these interactions. The middle region and C-terminal disordered regions (aa 164-172, aa 252-271, aa 499-510, pink text) are predicted to be involved in protein binding by DISOpred. These may be

Figure 3.2 Output of secondary structure prediction tools for CNOT11

psiPRED and DISOpred were used to provide predictions for the composition of secondary structure elements within the CNOT11 sequence. The predictions were used to influence the design of CNOT11 truncations. Regions predicted to be helical have been boxed in blue and regions predicted to contain disordered stretches are boxed in purple. PFAM reported a predicted domain of unknown function DUP2362 for CNOT11 (aa 391-475, green text). The pink text indicates the disordered regions which are predicted to be involved with protein binding. The figure has been annotated with coloured triangles to indicate a selection of domain boundaries designed for CNOT11. Refer to **Figure 3.4** for full list of constructs and their respective expression levels.

important for the interaction with CNOT10 and CNOT1 and this was considered during construct design. A search using the PFAM database revealed a domain of unknown function (DUF) within CNOT11 (DUF2363) (aa 391-475, green text) (**Figure 3.2**). There are 521 incidences of this domain across 382 different species. Interestingly, this domain is also found in the NOT2/3/5 family of CCR4-NOT subunits. This may indicate that these DUF2363 containing proteins interact with each other, meaning that CNOT10 may interact with the NOT module of CNOT2/CNOT3. Alternatively, this shared domain may indicate that DUF2363 containing proteins both interact with the same stretch of RNA, as the complex is a deadenylase, or protein partner. Phyre 2, a suite of tools used to predict protein structure, was adopted to analyse the DUF2363 domain (Kelley and Sternberg, 2009). The resulting predicted structure contains 6 helices and is modelled based on a 75% similarity to the 28s ribosomal protein s15. This close similarity to a solved structure highlights the DUF2362 domain as an ideal crystallisation candidate.

CNOT10 and CNOT11 were interesting targets alone as the individual function of these proteins has not been established, thus determining the type of fold by crystallography may provide insight into protein function. Additionally, the sub-complex of CNOT10 and CNOT11 was an interesting target to ascertain how these proteins bind and how they might dock onto the CCR4-NOT complex. Consequently, construct design focussed on producing recombinant proteins that were rigidly structured and highly expressed for crystallisation trials. Additionally, longer open reading frames were designed to include potential protein interacting regions for the study of CNOT10:CNOT11 complex formation. **Figure 3.3** shows the domain boundaries designed to produce CNOT10 and CNOT11 constructs.

3.3 High-throughput cloning and small scale expression testing of constructs from subunits of the CCR4-NOT complex

HTP cloning techniques were used to increase the number of potential constructs that can be assessed in a single experiment. When combined with bioinformatics-guided construct design, this could improve the likelihood of producing soluble proteins for the study of interactions within subunits of the CCR4-NOT complex. Gene fragments were generated from template DNA (**Appendix Table 7.4**) in a 96-well format by PCR, as described in section 2.7. Subsequently, the gene fragments were sub-cloned, using LIC, into expression vectors which incorporated a range of fusion tags. This approach also allows the vectors/recombinant tags to be easily inter-changed (Savitsky et al., 2010). All vectors provided a His_6 tag to allow capture by nickel affinity chromatography but fusion proteins to improve the solubility (SUMO, GST) or increase the functionality of the proteins (e.g. biotinylation recognition sequence for the use of protein in BLI) were also trialled. The vectors used in cloning of CCR4-NOT subunits are described in **Table 2.2**. As a result of bioinformatics-guided construct design, 489 constructs from 12 different proteins of or associated with the CCR4-NOT complex (CNOT1-11, TAB182 and TTP) were cloned successfully and subsequently verified at the DNA level. Construct design included nested truncations of amino acid boundaries sub-cloned into 13 different vectors encompassing both bacterial and baculovirus expression plasmids. The landscape of constructs cloned in the search for suitable crystallisable domains (e.g. short, rigidly structured domains) and constructs for studying protein-protein interactions (longer regions which may increase likelihood of containing the regions required for the interaction) is shown in **Table 3.1**.

Once successfully cloned, constructs were transformed into expression strain BL21-DE3-pRARE2 which contains tRNAs for seven rare codons to facilitate expression of human proteins in *E. coli*. HTP expression tests of 96 constructs, as described in chapter 2.21.1, used nickel affinity chromatography to capture recombinant proteins. The wide range of boundaries attempted allowed the discovery of many well expressed proteins (**Table 3.1, Appendix Figures 7.5-7.14**).

Of the 18 CNOT10 constructs designed, only four were expressed and soluble, indicated by protein visible in elutions fractions. Two of the four constructs were highly expressed, shown by a thick band on SDS-PAGE. An example of the expression test results is depicted in **Figure 3.3** and a summary of the expression levels of all *E. coli* expressed CNOT10 constructs is shown in **Figure 3.4**. Neither the full length (aa 1-717) nor near-full length (aa 66-717, aa 1-708) constructs were expressed and soluble. Only constructs

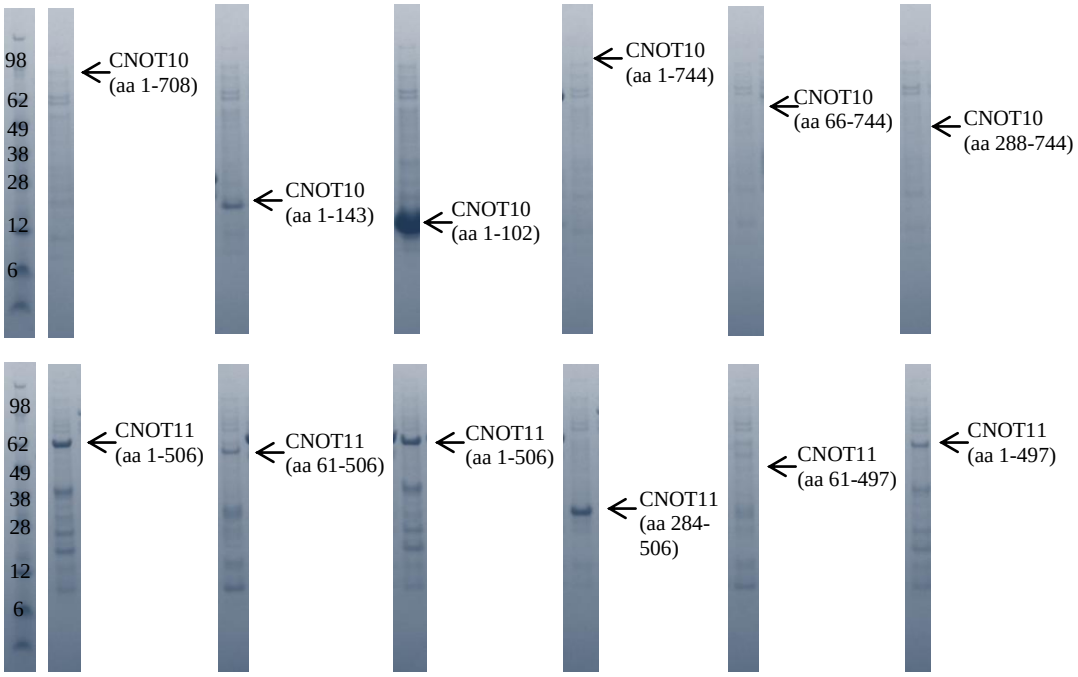


Figure 3.3 Example expression test of CNOT10 and CNOT11 constructs

Expression of CNOT10 and CNOT11 constructs were tested using 1 ml of *E. coli*. Protein was purified using nickel affinity chromatography and elutions were visualised by SDS-PAGE. The construct boundaries and position of band have been highlighted to the right of each lane. Molecular weight markers are indicated for size comparison (kDa).

spanning the first ~100 amino acids (aa 9-102, aa 1-102 and aa 1-143) were visible by SDS-PAGE, with well expressing constructs indicated by dark green bars (**Figure 3.4**). These short constructs were highly expressed and as such would make good candidates for crystallisation. The type of tag chosen did not appear to have an effect on the solubility level of the constructs for CNOT10 constructs. Conversely for CNOT11, 15/18 cloned constructs were soluble and expressed, and contained regions spanning almost the entire open reading frame (**Figure 3.4**). Two near-full length constructs (aa 1-506, aa 61-506) were soluble but expression levels were very low, indicated by the pale green bars. Constructs (aa 319-497, aa 284-506) which focussed around the DUF domain (aa 391-475, pink box) were well expressed. Again, the type of tag did not appear to affect expression or solubility levels for CNOT11 constructs. The identity of each protein in the expression test was confirmed by tandem MS, intact mass analysis by denaturing intact MS and its apparent molecular weight on SDS-PAGE. Across all CCR4-NOT subunits, solubility levels varied considerably between the different vectors (e.g. 26.7% for pNIC-Bio3 and 100% for pSUMO-LIC) (**Table 3.1**). The percentage difference, which is also protein-dependent, between successfully cloned and soluble constructs highlights the importance for a HTP screening technique as not every cloned construct will become a soluble protein. The best expressing and soluble constructs from the HTP approach (**Figure 3.4**, dark green, **Table 3.2**) were subsequently purified at 6 L scale to provide material to study them individually and also used to determine which domains were essential for interaction in the complex.

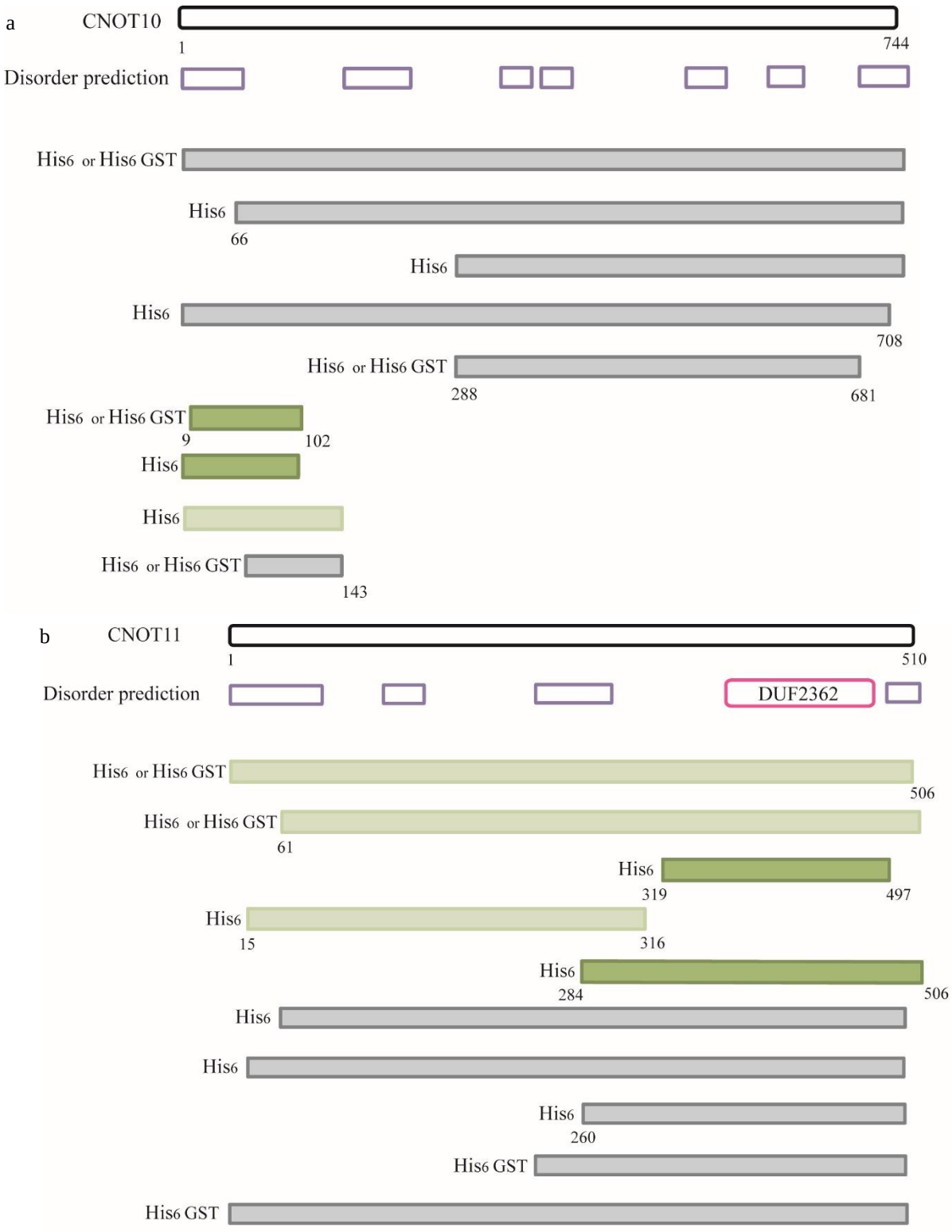


Figure 3.4 Depiction of the construct boundaries designed for CNOT10 and CNOT11

(a) Construct boundaries for CNOT10 where numbers represent amino acid numbers. The full length protein is depicted by the black bar. The purple boxes indicate the regions of protein disorder. The colour of bars beneath indicate the expression level observed; no expression (grey), medium expression (pale green) and high expression (dark green). (b) Construct boundaries for CNOT11. The full length protein is depicted by the black bar, where the predicted DUF from PFAM is indicated by a pink box. The purple boxes indicate the regions of protein disorder. The numbers represent amino acids in the sequence and the colour of the bar indicates the expression level observed; no expression (grey), medium expression (pale green) and high expression (dark green).

Subunit	Bacterial expression							Baculovirus expression								Total	Total soluble (%)
	pNIC28-Bsa4	pGTVL2	pNH-TrxT	pNIC-MBP-LIC	pNIC-SMPLX-MTLATH	pNIC-Bio2/3	pCDF-LIC	pSUMO-LIC	pFB-HGT-LIC	pFL-NT6H-LIC	pFL-CT10HSUII-LIC	pSPL-CTF-LIC	pUCDM-CTF-LIC	pFB-LIC-Bse	pFB-Bio5		
CNOT1	30 (20)	34 (23)					2 (2)	2 (2)		2 (2)	1 (1)			9 (5)		80 (55)	69
CNOT2	30 (20)	4 (4)	24 (14)	11 (5)	4 (2)	3 (1)							2 (2)			78 (48)	61
CNOT3	4 (4)	4 (4)										1 (1)	1 (1)			10 (10)	100
CNOT4	3 (3)	2 (1)														5 (4)	125
CNOT6	3 (0)	3 (1)														6 (1)	17
CNOT6L	4 (2)	4 (2)										1 (1)				9 (5)	56
CNOT7	3 (2)	2 (1)										1 (1)				6 (4)	67
CNOT8	4 (3)	4 (1)														8 (4)	50
CNOT9	45 (40)	3 (1)	5 (3)	1 (1)		4 (3)				1 (1)	1 (1)	1 (1)				51 (61)	84
CNOT10	12 (3)	4 (0)											1 (1)	1 (0)		18 (4)	22
CNOT11	12 (10)	4 (3)											1 (1)	1 (1)		18 (15)	83
TAB182	6 (2)	6 (4)										1 (1)				13 (7)	54
TTP	36 (16)	2 (1)	53 (18)		7 (1)	23 (4)		9 (9)	18 (1)					29 (5)	1 (0)	178 (55)	31
Total	272 (135)	76 (45)	82 (35)	12 (6)	11 (3)	30 (8)	2 (2)	11 (11)	18 (1)	3 (3)	2 (2)	5 (5)	5 (5)	40 (11)	1 (0)		
Total soluble (%)	50	60	42	50	27	27	100	100	6	100	100	100	100	28	0		

Table 3.1 Number of cloned and soluble constructs for each of the CCR4-NOT complex subunits

The number of cloned and soluble constructs (indicated by the brackets) was mined from internal databases. The percentage of soluble constructs (Total soluble (%)) represents the portion of total cloned constructs for all subunits where a corresponding protein band was observed by SDS-PAGE. The percentage of total soluble constructs has been calculated for individual vectors and for each subunit. A total soluble percentage of 52% was calculated which denotes the ratio of soluble constructs over total number of cloned constructs (for all subunits) attempted for various purposes, throughout the entire duration of the PhD study. It includes constructs with nested boundaries with the purpose of investigating solubility at the initial phase, but also includes constructs incorporating different fusion tags on a known soluble construct at the latter phase for biophysical arrays. Therefore, this percentage should not be interpreted as attrition rate for solubility of a certain target.

3.4 Characterisation of individual subunits

3.4.1 Large scale preparation of CNOT11 (aa 284-506)

Well expressed and soluble constructs were scaled up to 6 L of *E. coli* culture and protein purified for attempts at crystallisation. One such construct was CNOT11 (aa 284-506) which contains an N-terminal His_6 tag followed by a site cleavable by TEV protease. This construct extends around the predicted DUF2363 domain. **Figure 3.5** describes the typical purification steps undertaken to produce homogeneous protein for crystallisation. CNOT11 (aa 284-506) was first captured using nickel affinity chromatography (**Figure 3.5a**). The protein was highly expressed and approximately 80% pure after nickel affinity purification.

To achieve greater purity, the protein was loaded on an S75 16/60 gel filtration column to separate remaining contaminants (**Figure 3.5b**). An S75 column was chosen to purify CNOT11 (23 kDa) as the beads are designed to best separate proteins <80 kDa. The CNOT11 protein eluted from the column in two peaks (elution volume ~45 ml and 65 ml). The second peak, corresponding to the majority of the CNOT11 protein based on SDS-PAGE analysis, was pooled. TEV protease was used to cleave the His_6 tag at 4 ºC overnight. The following morning, the protein was passed over nickel resin to remove the cleaved tag and TEV protease. The resin was washed with buffers containing increasing imidazole concentration (**Figure 3.5d**). The sample which flowed over the column (FT) without binding is expected to contain the untagged target protein as it can no longer interact with the resin since the tag has been removed. The washes of increasing imidazole concentration contain non-specifically bound proteins and perhaps some of the untagged CNOT11. The final imidazole elution (containing 300 mM imidazole) contains the TEV protease, which is expressed with a His_6 tag. The CNOT11 protein, which

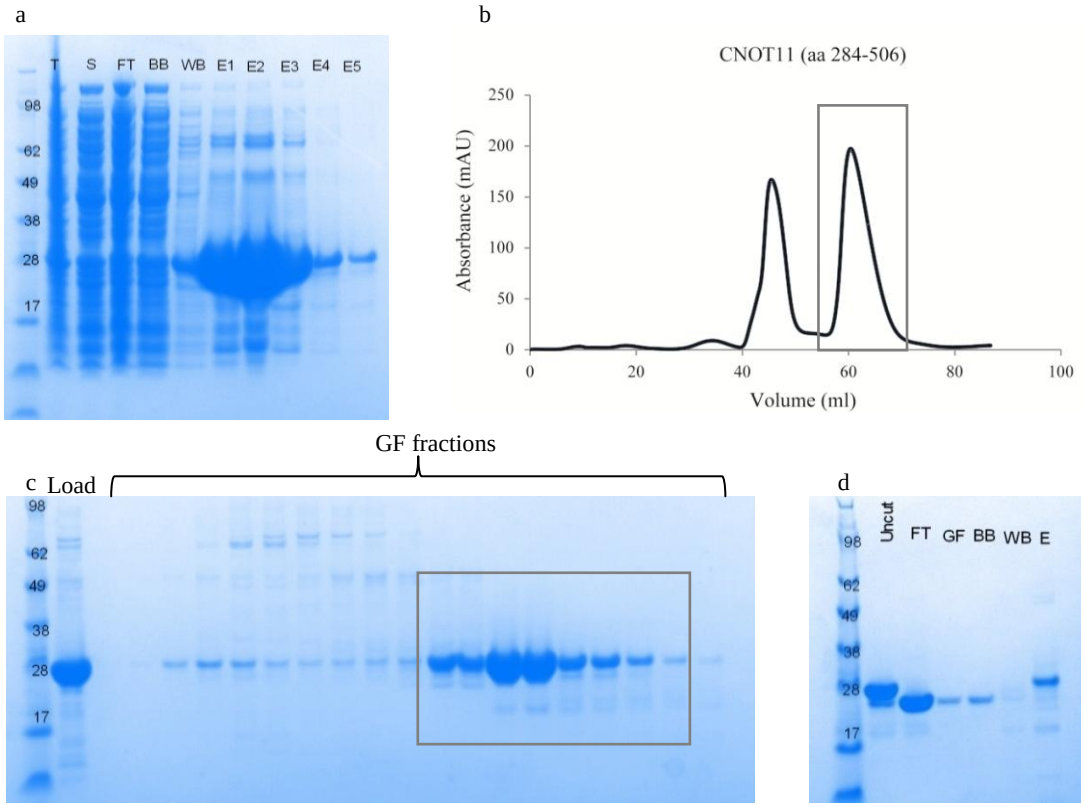


Figure 3.5 An example purification strategy of CNOT11 (aa 284-506)

(a) SDS-PAGE of a nickel affinity capture of CNOT11 (aa 284-506). Lanes represent (from left to right) the total protein expressed in cells (T), the soluble fraction from centrifugation (S), the fraction which flowed through nickel resin without binding (FT), the protein which eluted when the column was washed with binding buffer (BB), the protein which eluted when the column was washed with wash buffer (WB), the protein which eluted when the column was washed with elution buffer (E). E1-E2 was pooled for subsequent gel filtration. (b) Chromatogram of gel filtered CNOT11 (aa 284-506) using an S75 column showing two peaks. Fractions spanning both peaks were visualised by SDS-PAGE and the second peak was found to contain the majority of the CNOT11 protein. The box indicates the fractions which correspond to the box in figure c. (c) SDS-PAGE of fractions from gel filtration, where the box indicates fractions from the boxed peak in figure b (d) SDS-PAGE of TEV cleavage of His_6 tag and subsequent imidazole washes and elution of His_6-tagged TEV protease. Lanes represent (from left to right) the uncut protein, FT and imidazole washes of 0 mM (GF), 20 mM (BB), 40 mM (WB) and 300 mM (E). Molecular weight markers are indicated for size comparison (kDa).

eluted in the flow through (FT) and low imidazole (0-20 mM) washes was pooled. The resulting protein was of sufficient purity for crystallisation trials. Three constructs of CNOT10 (aa 1-102, 9-102 and 1-143) and two of CNOT11 (aa 1-506 and aa 284-506) (**Figure 3.4**) were scaled up in this manner (**Table 3.2, Figure 3.6**) using a similar purification regime.

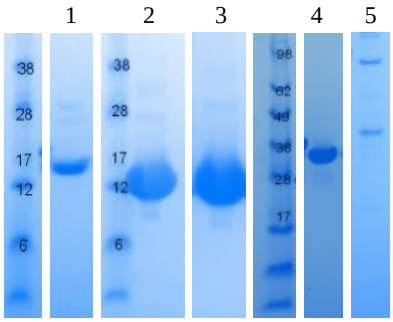


Figure 3.6 Example purifications of CNOT10 and CNOT11

SDS-PAGE showing the purification of three CNOT10 (1= aa 1-143, 2= aa 9-102 and 3= aa 1-102) and two CNOT11 (4= aa 284-506 and 5= aa 1-506) constructs.

	Construct boundary	Yield (mg)
CNOT10	aa 9-102	5.3
	aa 1-102	15.6
	aa 1-143	3.6
CNOT11	aa 284-506	5.6
	aa 1-506	1.2

Table 3.2 The yields of purifications of CNOT10 and CNOT11

The individual yields (mg) from 3 L of culture of different CNOT10 and CNOT11 constructs which were purified for crystallisation trials.

3.4.2 Crystallisation of CNOT10 and CNOT11 constructs

Purified proteins were concentrated to >15 mg/ml before being dispensed into crystallisation trials. Sitting drop crystallisation plates were used where the drop contained 150 nl mixture of protein and precipitant and the reservoir contained 20 µl of each precipitant solution from the random matrix crystallisation screen (e.g. HIN3, HCS and JCSG). Therefore, each crystallisation trial contained 96 different precipitant compositions with three protein:precipitant ratios (1:2, 1:1 and 2:1) which amounted to 288 conditions sampled per plate. Plates were incubated preferentially at 20 ºC to provide a warm enough temperature to encourage vapour diffusion but cool enough as to prevent the drop from drying out too quickly. If sufficient quantity of protein was purified, a replica set of plates were set up and incubated at 4 ºC to investigate crystallisation and drop behaviour at lower temperatures. Crystallisation plate ‘hotels’ (either Minstrel or Formulatrix) were used to incubate the plates and each drop was imaged at regular time intervals.

A number of conditions yielding microcrystals of the various purified proteins of CNOT10 (**Figure 3.7c**) and CNOT11 (**Figure 3.7a and b**) were observed. Microcrystals formed at neutral pH where high molecular weight polyethylene glycol (PEG) was the precipitant. HEPES buffer also featured in two of the three precipitant conditions (**Figure**

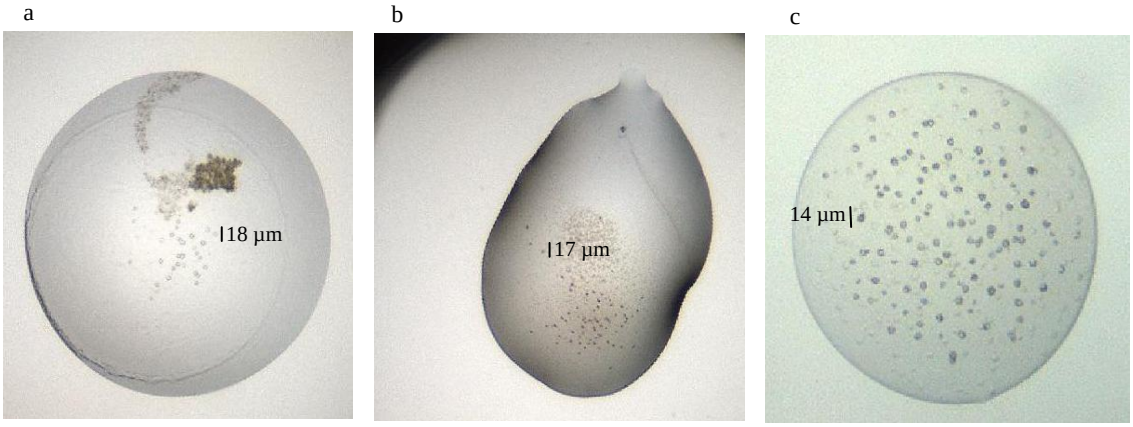


Figure 3.7 Crystalline elements observed in trials for CNOT10 and CNOT11

Three examples of drops where crystalline elements have been observed. Microcrystals were observed at either the 28 day or 56 day inspection. (a) CNOT11 (aa 284-506) microcrystals grew in 5% tacsimate, 10% PEG 5000 MME, 0.1 M HEPES pH 7. (b) CNOT11 (aa 1-506) microcrystals grew in 28% PEG (of varying molecular weights), 0.1 M Tris pH 8.5, 0.15 M ammonium acetate, 0.01 M $CdCl_2$. (c) CNOT10 (aa 9-102) microcrystals appeared in 1 M succinic acid, 1% PEG2000 MME, 0.1 M HEPES pH 7.

3.7a and c). Typically, these microcrystals appeared after one month and are suspected to be formed by protein due to the time they took to form, and the observation that they continued to grow over time. To optimise crystallisation conditions to produce larger crystals which may yield sufficient diffraction for structural determination, the individual precipitant conditions were expanded upon in the design of an optimisation screen, where the concentration of the pH buffer and precipitant (e.g. PEG) were varied. Additionally, a seeding approach was adopted for the CNOT10 microcrystals (**Figure 3.7c**). Crystal seeds are crushed fragments of initial crystal hits which require optimisation. The seeds act as nucleation sites to initiate crystal growth, which then grow in conditions which facilitate a metastable zone. The seeds were added to both the screen where the microcrystals were initially found and an optimisation screen to screen for conditions

which facilitate a metastable zone, indicated by growth of seeds into crystals. Unfortunately, no crystals of a sufficient size to be mounted, as judged by eye from images, were observed for either protein.

3.5 Interaction analysis between CNOT10 (aa 1-102) and CNOT11 (aa 284-506)

To study the interaction between CNOT10 and CNOT11, the complex was first reconstituted from individually purified proteins. Complex formation between well expressed *E. coli* constructs of CNOT10 (aa 1-102) and CNOT11 (aa 284-506) was determined by analytical gel filtration as each protein elutes at a specific volume (elution volume) depending on the size and shape of that protein. Thus, if a complex between CNOT10 and CNOT11 had formed, the resulting complex would be larger in size than the single components and pass more quickly through the column, causing a change in the elution volume when the chromatograms are overlaid.

Firstly, a sample of 350 µg of CNOT10 (aa 1-102) and 450 µg of CNOT11 (aa 284-506) were run separately on a Superdex S200 10/300 GL analytical gel filtration column as a control for complex formation. A single peak in the chromatogram was observed for both CNOT10 (aa 1-102) and CNOT11 (aa 284-506) separately (**Figure 3.8**). The corresponding SDS-PAGE for CNOT10 (aa 1-102; 13.8 kDa) also displays a single protein (**Figure 3.9a**). For CNOT11 (aa 284-506: 28.6 kDa), a faint band (13.8 kDa) is visible before the peak for CNOT11 (28.6 kDa) but is not visible in the pre-GF sample (**Figure 3.9b**). It was confirmed by tandem MS to be CNOT10 protein, likely carried over from the previous analytical gel filtration run of CNOT10 alone. To show if CNOT10 (aa 1-102) and CNOT11 (aa 284-506) form a complex, the two proteins (in

quantities as above) were mixed and incubated at room temperature for 1 h before running on analytical gel filtration (**Figure 3.8**).

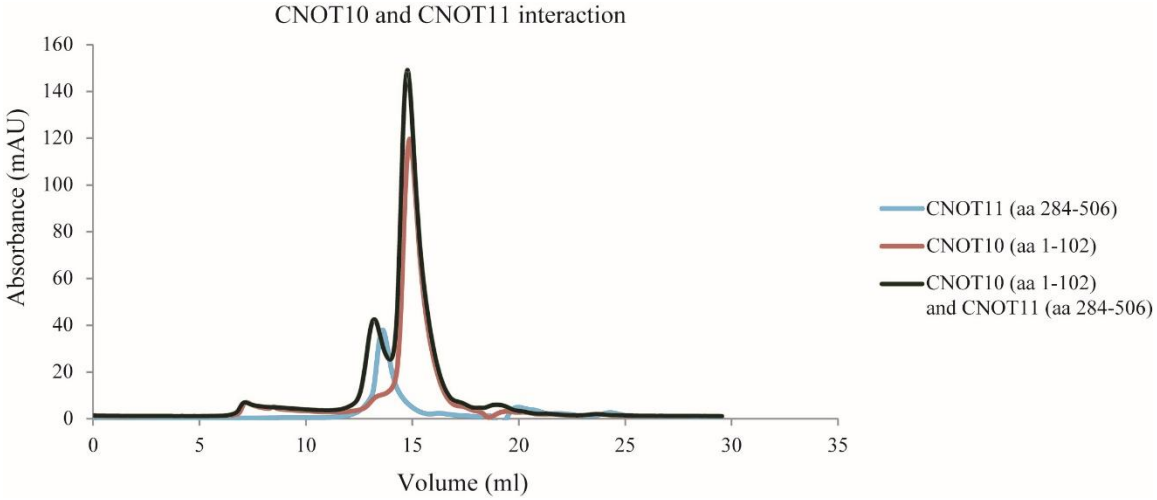


Figure 3.8 Overlaid chromatograms of CNOT10 (aa 1-102), CNOT11 (aa 284-506) and CNOT10:CNOT11

To determine if a complex forms between CNOT10 (aa 1-102) and CNOT11 (aa 284-506), CNOT10 (aa 1-102) (red), CNOT11 (aa 284-506) (blue) and a mix of CNOT10 (aa 1-102) and CNOT11 (aa 284-506) (black) were run on a S200 10/300 GL column. A shift in the elution volume of the sample of CNOT10 and CNOT11 (black) would indicate a complex has formed. No such shift was detected.

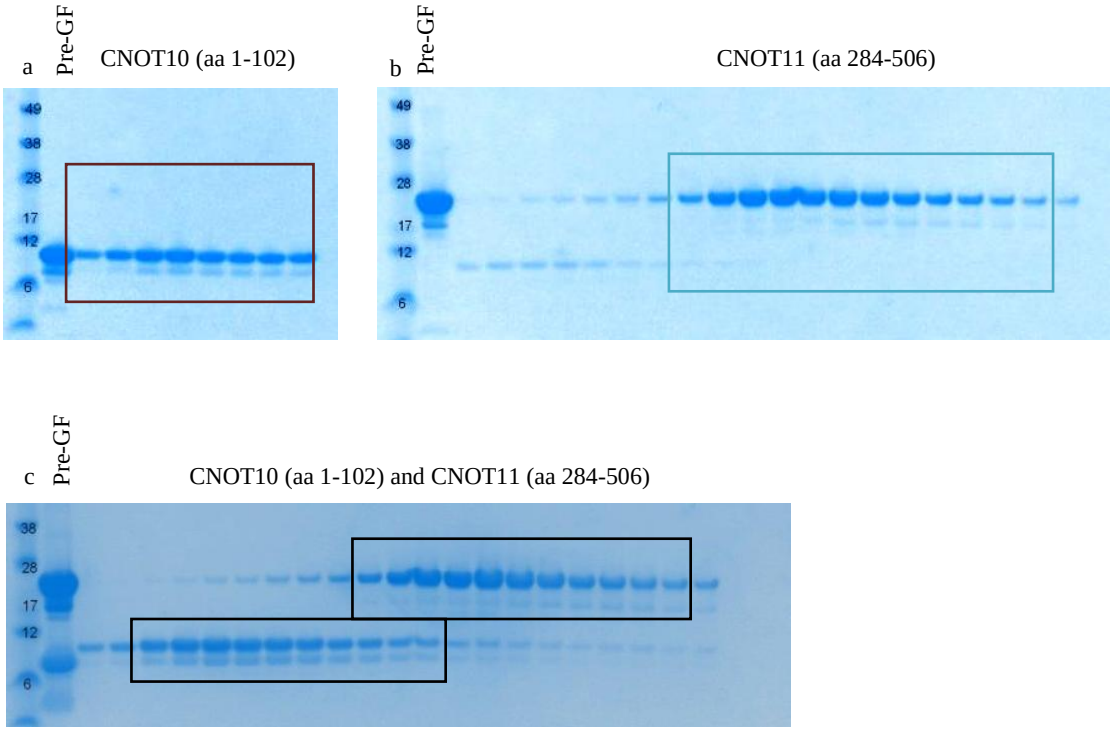


Figure 3.8 Gel filtration of CNOT10 (aa 1-102), CNOT11 (aa 284-506) and CNOT10:CNOT11

Complex formation between CNOT10 and CNOT11 was analysed by analytical gel filtration using an S200 10/300 GL column. Gel filtration samples of 350 µg of CNOT10 (aa 1-102), 240 µg of CNOT11 (aa 284-506) and a mix of 350 µg of CNOT10 (aa 1-102) and 450 µg CNOT11 (aa 284-506) were analysed on a SDS-PAGE gel. Molecular weight markers are indicated for size comparison (kDa).

Although the CNOT10:CNOT11 chromatogram (black) appears to display a change to elution volume in the first peak, when analysed in combination with SDS-PAGE gels (**Figure 3.9c**), it is clear that the two proteins are eluting separately, in elution volumes according to their respective unmixed sample. The results suggest that CNOT10 (aa 1-102) and CNOT11 (aa 284-506) do not form a complex in vitro. Although a negative result, it can be concluded that either the N-terminus of CNOT10 or the C-terminal half of CNOT11 is not sufficient to form the CNOT10:CNOT11 module, suggesting the interacting domains may lie within the C-terminus of CNOT10 and the N-terminus of CNOT11. Only the N-terminal ~100 amino acids of CNOT10 were explored due to solubility problems and a larger domain of CNOT10 and CNOT11 may be required for these proteins to bind. Longer constructs could not be expressed in *E. coli* and as a result, protein production was attempted in *Sf9* cells.

3.6 Expression of CNOT10 and CNOT11in *Sf9* insect cells

In the subsequent experiment, full length CNOT10 and CNOT11 were cloned into pFB-LIC-Bse, a vector for baculovirus-mediated insect cell expression which provides an N-terminal His_6 tag followed by a site cleavable by TEV protease. In the first instance, full length boundaries were chosen to ensure coverage of interacting regions, in order to confirm in vitro that CNOT10 and CNOT11 interact. The cloned vectors were transposed into DH10Bac *E. coli* to produce recombinant bacmids which were used to transfect *Sf9* cells. The resulting P0 baculoviruses were was used for small scale expression and purification tests. Results showed that CNOT10 alone was not soluble and although CNOT11 was not visible by eye, expression was confirmed by tandem MS (**Figure 3.10**). This suggests that CNOT10 and CNOT11, when expressed alone, are largely insoluble.

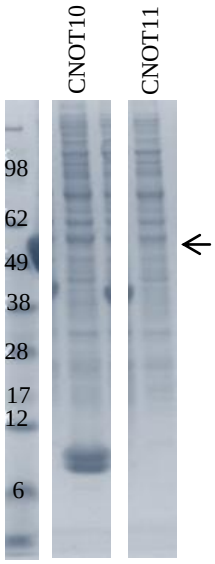


Figure 3.9 Expression test of CNOT10 and CNOT11 in Sf9 insect cells

A 3 ml expression test of CNOT10 and CNOT11 was completed. Elutions were run on a SDS-PAGE gel to determine solubility and expression levels of proteins. *Sf9* cells were infected with 120 µl of CNOT10 and CNOT11 P0 virus. The arrow indicates a band of CNOT11 protein which was confirmed by tandem MS. No soluble protein was observed for CNOT10. Molecular weight markers are indicated for size comparison (kDa).

3.6.1 Co-infection of CNOT10 and CNOT11

CNOT10 and CNOT11 have been shown to interact, as verified by Y2H screening, thus a co-expression test was attempted with 50 ml of *Sf9* cells in the hope of increasing solubility and stability of both partners. The *Sf9* cells were co-infected with a virus expressing full length CNOT10 and a virus expressing full length CNOT11. Co-infection of viruses enables expression of both CNOT10 and CNOT11 in the same cell, thus allowing the proteins to interact during the process of translation. The interactions made between CNOT10 and CNOT11 may then stabilise the proteins and protect them from degradation by proteases during purification of the complex (Gao et al., 1996, Desai et al., 2012). Expression tests of 50 ml were completed in parallel using 120 µl of P0 virus for single infections, and 100 µl of each virus for co-infection. Infected cells were incubated at 27 ºC for 66-72 h and the resulting protein was purified using nickel affinity chromatography. As observed in the 3 ml expression tests, purified protein from 50 ml of *Sf9* cells infected with virus expressing CNOT10 alone produced little to no soluble protein, and of CNOT11 alone showed a weak band by SDS-PAGE (**Figure 3.11a and b**). On the contrary, when viruses were co-infected, the expression of both proteins was

amplified (**Figure 3.11c**). The identities of the bands were confirmed by tandem MS. CNOT10 was only soluble when co-expressed with CNOT11, indicating that a complex has formed between the two proteins.

To further investigate the potential complex formation of CNOT10:CNOT11, the above experiment was repeated at litre scale. 3 L of *Sf9* cells were co-infected with CNOT10 and CNOT11 expressing viruses, and proteins were purified using nickel affinity chromatography. The N-terminal His_6 tag was removed from both proteins using TEV protease, and the proteins were gel filtered using an S200 16/60 column to determine if they co-migrated which would indicate complex formation was not tag-dependent.

The chromatogram of gel filtration displays four minor peaks and one major peak (**Figure 3.12a**). The first peak likely contains nucleic acids, as no protein is observed on the SDS-PAGE gel (**Figure 3.12b**). The second and largest peak of the chromatogram (indicated by the solid box) corresponds to the co-elution of CNOT10 and CNOT11, thus

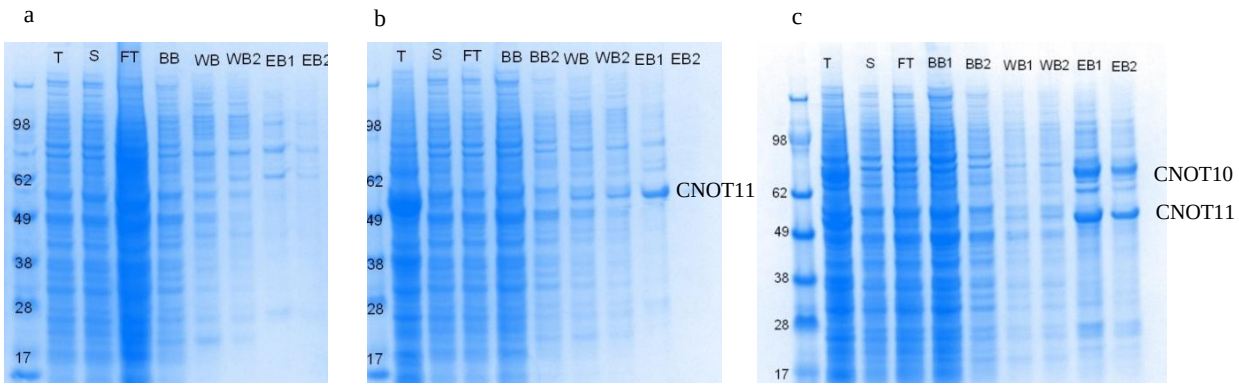


Figure 3.10 SDS-PAGE of nickel affinity capture of CNOT10, CNOT11 and CNOT10:CNOT11

Expression was tested using 50 ml of *Sf9* cells, which were either singularly or co-infected with virus expressing either full length CNOT10, CNOT11 or both. Protein was purified using nickel affinity chromatography and elutions were analysed by SDS-PAGE. *Sf9* cells were infected with 120 µl of (a) CNOT10, (b) CNOT11 and 100 µl of each of (c) CNOT10:CNOT11 virus. Lanes represent (from left to right) the total protein expressed in cells (T), the soluble protein (S), the fraction which flowed through nickel resin without binding (FT), the protein which eluted when the column was washed with binding buffer (BB), the protein which eluted when the column was washed with wash buffer (WB), the protein which eluted when the column was washed with elution buffer (EB). Molecular weight markers are indicated for size comparison (kDa).

it can be concluded they are a true complex. A small shoulder from the second peak can be detected (indicated by the dashed box). This corresponds to uncomplexed CNOT11 protein in the SDS-PAGE gel, which is stable alone. The third peak was TEV protease (~28 kDa) which was not removed before gel filtration. Subsequent peaks are likely small molecules as they elute close to the bed volume of the column.

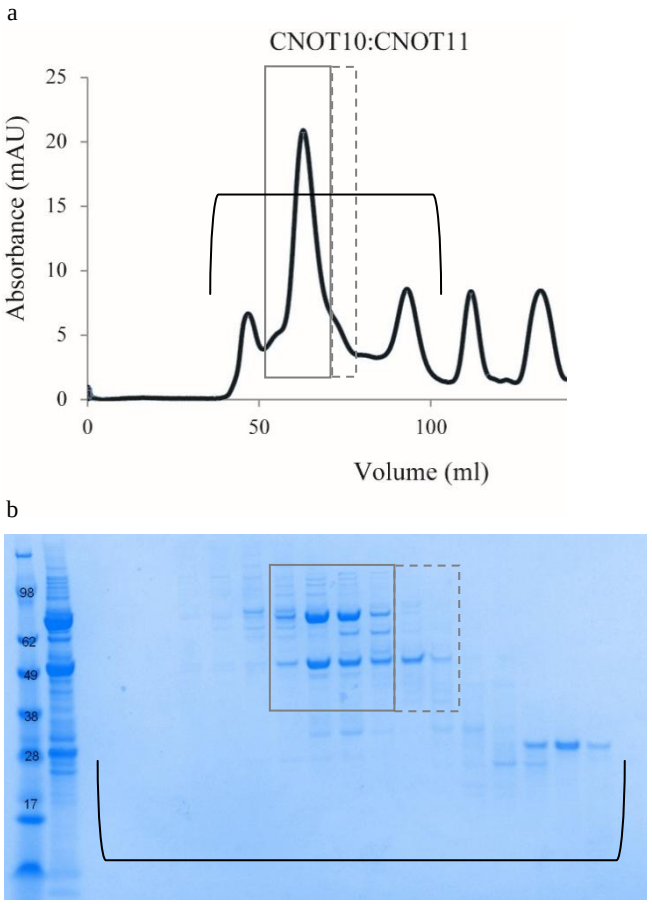


Figure 3.11 Gel filtration of CNOT10:CNOT11

CNOT10:CNOT11 proteins expressed from 3 L of insect cells were analysed by gel filtration to characterise complex formation. (a) Chromatogram of full length CNOT10:CNOT11 purified from a S200 16/60 gel filtration column. (b) SDS-PAGE of eluted fractions from gel filtration of CNOT10:CNOT11. The brace indicates the fractions from gel filtration which were visualised by SDS-PAGE. The second peak (boxed) corresponds to the elution of CNOT10:CNOT11. There is a slight shoulder to the second peak, which corresponds to CNOT11 alone, indicated by the dashed box. The third peak was TEV protease, observed by SDS-PAGE as 28 kDa band by SDS-PAGE. Molecular weight markers are indicated for size comparison (kDa).

3.7 Crystallisation of the complex of full length CNOT10:CNOT11

The above experiments showed that it is possible to form a complex of full length CNOT10 and CNOT11 when the proteins were expressed together using viral co-infection. Consequently, crystallisation attempts were then made using this complex. CNOT10:CNOT11 protein was purified as described above from 6 L of insect cells co-

infected with CNOT10 and CNOT11 viruses. Protein was pooled and concentrated to 9 mg/ml for crystallisation trials using JCSG and HIN3 random matrix screens. Small crystals were observed in the 9th inspection (+54 days) in a condition containing 0.1 M magnesium formate and 15% PEG3350 (**Figure 3.13**). The crystals only grew to a small size (7 µm) as there were many nucleation points in the drop. However, fungal growth was also present in the crystal drop. As it was observed that other drops containing a similar precipitant, without fungal growth observed, did not produce crystals, it was hypothesised that the crystal formation was due to the presence of the fungal growth. This phenomenon previously led to the structural determination of yeast CPSF-100, a polyadenylation specificity factor (Mandel et al., 2006). In this instance, a *Penicillium* infection is thought to have catalysed the removal of ~200 hydrophilic residues from CPSF-100 likely by secretion of a fungal protease. This proteolysis greatly reduced solubility, allowing crystal formation. Thus, the fungus observed in **Figure 3.13** may also have secreted proteases, removing the flexible regions of the CNOT10:CNOT11 complex and favouring crystallisation.

The crystals were screened in-house using a Bruker X8 Proteum with microstar x-ray home source to ascertain if they contained protein, based on the fact that protein diffraction and salt diffraction appear different. Unfortunately, no diffraction of either kind was observed thus the nature of these crystals cannot be concluded. However, the length of time which the crystals took to form (i.e. microcrystals appeared after 28 days) and the observation that they continued to grow over time (inspected between 28 and 54 days) suggest they are protein crystals. Whether the crystals contain both CNOT10 and CNOT11 proteins, and which domains are included in the crystals, were important details to establish in order to optimise these further.

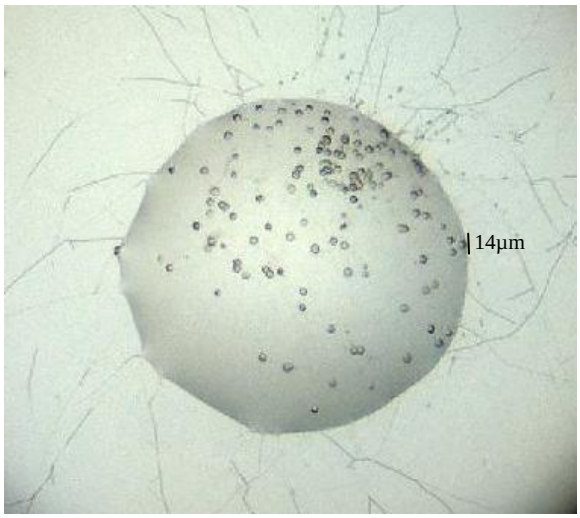


Figure 3.12 Crystals of CNOT10:CNOT11

Crystals of CNOT10:CNOT11 were observed after 54 days in a 150 nl crystallisation drop. The coarse screen condition contained 0.1 M magnesium formate and 15% PEG3350. A significant amount of fungal growth was visible in the drop. A scale bar is included to allow crystal size to be estimated.

An optimisation screen around the original condition was designed and set up with the same stock of protein, but this did not replicate the initial crystals. This again suggests that crystal formation was dependent on the removal of flexible domains by proteases secreted by fungal contamination. The remaining crystals from the original drop shown in **Figure 3.13** were removed to produce a crystal seed stock and used in both initial and optimisation screens. Unfortunately, seeding also did not succeed in replicating the crystals and an alternative means to truncate regions for crystallisation were sought. As a result, limited proteolysis was carried out on the purified complex protein, to both ascertain the interacting domains of CNOT10:CNOT11 and to assist crystallisation.

3.8 Limited proteolysis of CNOT10:CNOT11

3.8.1 Small scale proteolysis

To imitate the potential proteolysis of CNOT10:CNOT11 by the fungus, limited enzymatic proteolysis was used. Two samples of 150 µg of purified CNOT10:CNOT11 complex were treated with trypsin and chymotrypsin in a time-course experiment. Trypsin, a serine protease with specificity for lysine and arginine residues (Krieger et al., 1974), and chymotrypsin, a serine endopeptidase which cleaves peptide bonds between

aromatic residues (Blow, 1968), were added to samples of CNOT10:CNOT11 at a mass ratio of 1:400 (enzyme to protein). Samples for SDS-PAGE and denaturing MS analyses were taken at the following time points: 0 h, 15 min, 30 min, 1 h, 2 h and 3 h (**Figure 3.14**). The first lane containing the 0 h sample shows intact CNOT10 (78 kDa, asterisk) and CNOT11 (55 kDa, caret). Some degradation products (~30-35 kDa, underscore) were also detected. These represent degraded CNOT11 as the bands are consolidated to a single band at 30 kDa after 4 h determined to be CNOT11 by tandem MS. Over a period of 4 h, proteolysis by trypsin (**Figure 3.14a**) and chymotrypsin (**Figure 3.14b**) have depleted the full-length proteins, with the concomitant production of a stable fragment of 46 kDa, which tandem MS revealed to be CNOT10. Trypsin proteolysis further produced one band of ~30 kDa, found to be CNOT11 whereas chymotrypsin proteolysis produced two ~30 kDa bands, confirmed to be CNOT11. Denaturing MS was used to determine the exact mass of the stable fragments in order to correlate with what is observed by SDS-

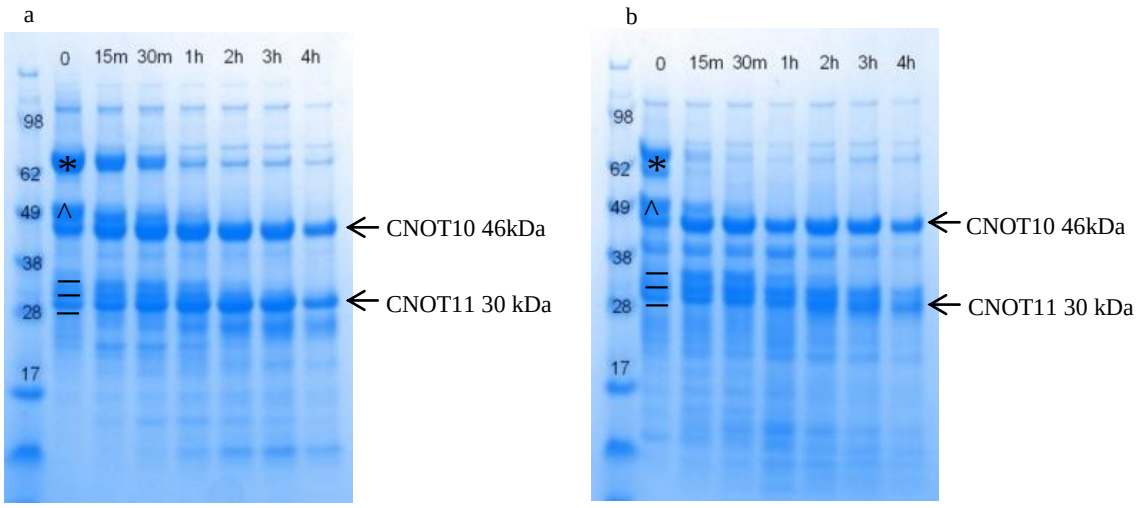


Figure 3.13 Limited proteolysis of CNOT10:CNOT11

Limited proteolysis samples taken at the following intervals: 0 min, 15 min, 30 min, 1 h, 2 h, 3 h, 4 h were visualised by SDS-PAGE. A ratio of 1:400 (enzyme:protein) was used of each protease and the samples were incubated at room temperature. (a) Limited proteolysis using trypsin. (b) Limited proteolysis using chymotrypsin. The asterisks indicates the position of intact CNOT10, and the caret indicates intact CNOT11. An underscore highlights some degradation products in the 0 h lane. The two bands in both 4 h lanes were identified by tandem MS to be CNOT10 (top) and CNOT11 (bottom). Molecular weight markers are indicated for size comparison (kDa).

PAGE. Denaturing MS samples of the time course were overlaid in **Figure 3.15**. In the 0 h spectra, the intact CNOT10 (78 kDa, asterisk) and CNOT11 (55 kDa, caret) can be observed. The overlaid spectra also highlight that proteolysis produces stable fragments of 46357 Da (CNOT10, boxed) and 30728 Da (CNOT11, boxed) which is consistent with SDS-PAGE analysis (**Figure 3.14**). The proteolysed fragments of CNOT10 and CNOT11 were next analysed using analytical gel filtration to determine if these fragments still form a complex in solution.

3.8.2 Validation of truncated complex

The same limited proteolysis experiment, where protein is incubated with 1:400 (enzyme to protein mass) trypsin for 4 h, was next repeated using a larger amount of protein (1.13

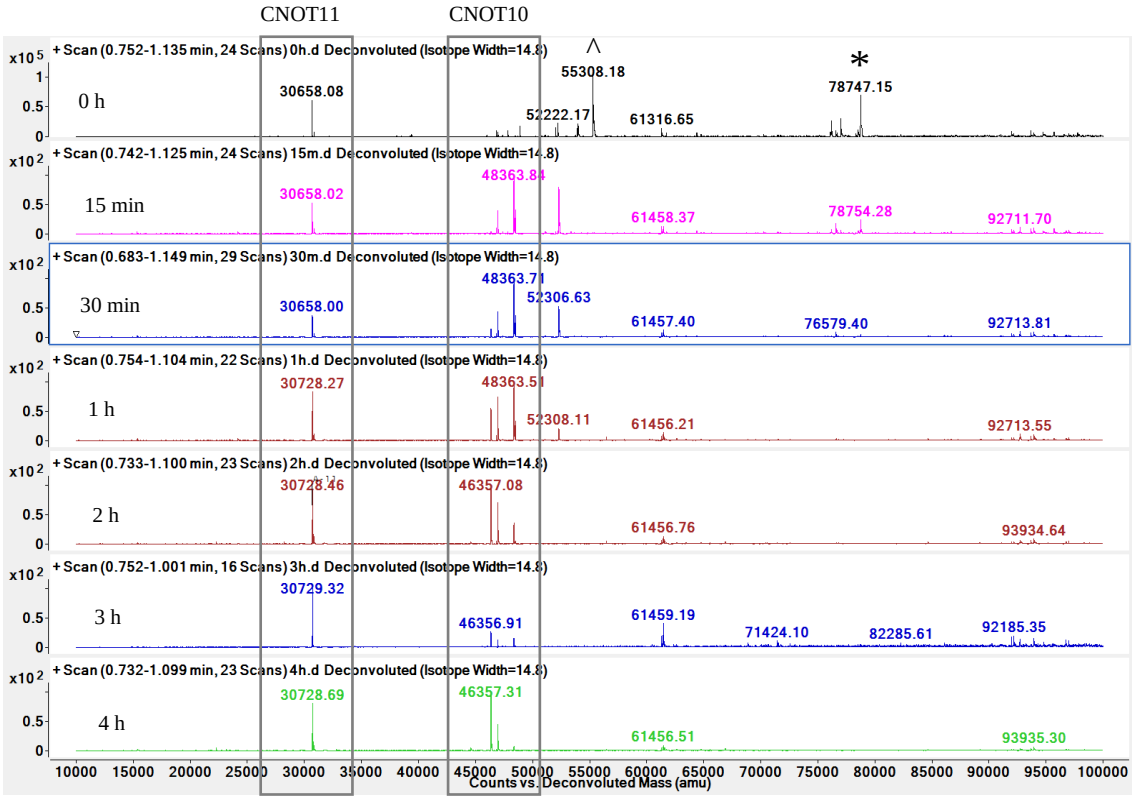


Figure 3.14 Stacked denaturing spectrum of trypsin limited proteolysis samples of CNOT10:CNOT11

Denaturing MS samples were taken at the following time intervals: 0 h, 15 min, 30 min, 1 h, 2 h, 3 h, 4 h. Samples were run on a QTOF (Agilent) and analysed with Mass Hunter (Agilent). Stacked spectra show the gradual degradation of the full length proteins (CNOT10, 78747 Da, asterisk; and CNOT11, 55308 Da, caret) and the production of the two stable domains observed by SDS-PAGE (30728 Da boxed and 46357 Da, boxed). MS samples for trypsin and chymotrypsin were collected but the trypsin samples produced a spectrum with less background and therefore chymotrypsin is not shown.

mg) to produce enough truncated CNOT10:CNOT11 for characterisation by analytical gel filtration.

Protease inhibitor was added at a 1:150 ratio (inhibitor:protein mass) to stop proteolysis and the reaction mixture was loaded onto an S200 10/300 GL column, resolving into two main peaks. The first corresponds to co-elution of the 46 kDa CNOT10 and 30 kDa CNOT11, confirmed by tandem MS (**Figure 3.16**). Therefore, co-migration of the two proteolysed fragments on gel filtration indicates that they still form a complex. To determine possible stoichiometry, the molecular mass of the truncated CNOT10:CNOT11 complex was calculated by the comparing elution volume of the complex to that of molecular weight standards (**Figure 3.15c**). The molecular mass of the truncated CNOT10:CNOT11 complex was estimated to be ~160 kDa which may represent a dimer of dimers of CNOT10 (46 kDa) and CNOT11 (30 kDa) fragments (~152 kDa). Further experiments such as analytical ultracentrifugation may be required to determine the true nature of the oligomerisation of the truncated CNOT10:CNOT11 complex.

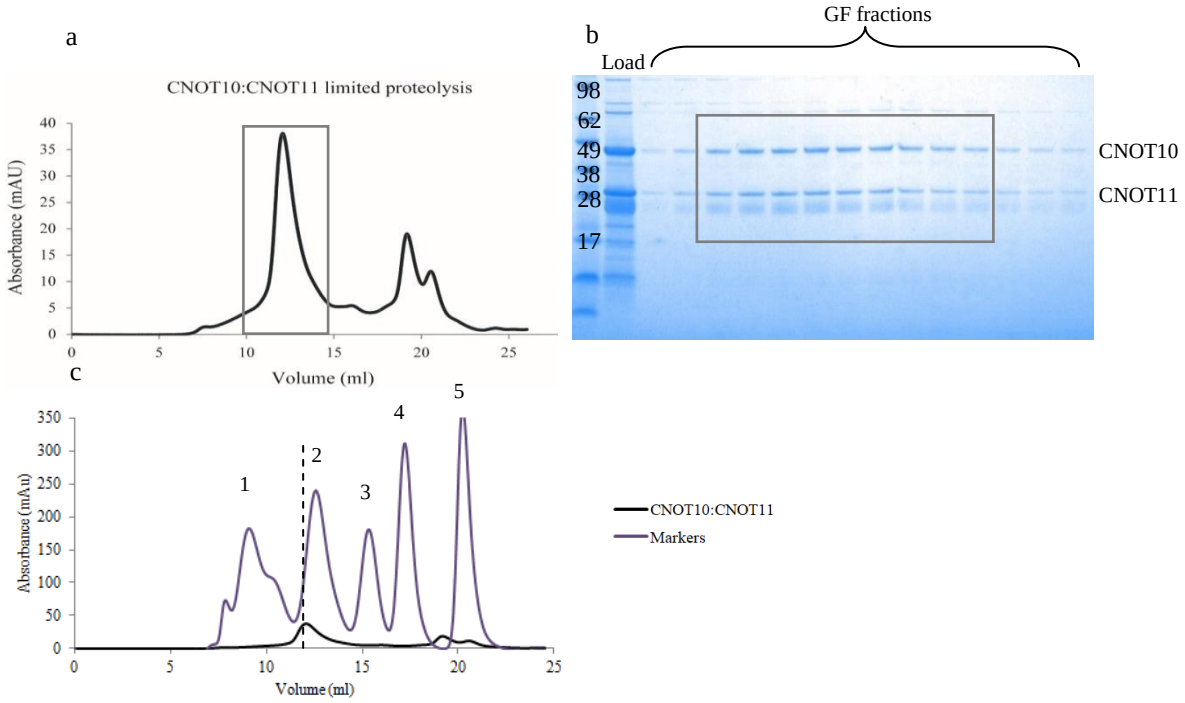


Figure 3.15 Gel filtration of proteolysed CNOT10:CNOT11

CNOT10:CNOT11 was proteolysed for 4 h with 1:400 trypsin and loaded onto an S200 10/300 GL column. (a) Chromatogram of gel filtration of proteolysed CNOT10:CNOT11. (b) SDS-PAGE of elutions from gel filtration of proteolysed CNOT10:CNOT11. Fragments elute in one peak indicating they are still complexed. Molecular weight markers are indicated for size comparison (kDa). (c) Overlay of the gel filtration chromatogram of proteolysed CNOT10:CNOT11 from (a) (black) and proteins of known size (molecular markers) (purple). Five proteins of known size were run (1=670 kDa, 2=158 kDa, 3= 44 kDa, 4=17 kDa and 5=13.5 kDa). The centre of the peak where CNOT10:CNOT11 elutes is highlighted by a dashed line to allow comparison with the molecular markers.

3.8.3 Limited proteolysis using full length CNOT11 alone

To determine that the cleavage patterns observed above are a direct consequence of complex formation, and not purely a result of secondary structure arrangement, a control experiment which explored the proteolytic cleavage of the single components was required. Since CNOT10 protein alone was not available due to solubility problems, the experiment was performed on CNOT11, with the rationale that the protection pattern from CNOT11 alone would be different from that of CNOT11 when in complex with CNOT10. Thus, CNOT11 was proteolysed using trypsin following the same conditions as with the CNOT10:CNOT11 complex.

Intact full length CNOT11 (~57 kDa, tagged) is visible in the 0 h lane, indicated by the caret. CNOT11 is degraded to a ~50 kDa band, indicated by the hash (**Figure 3.17**).

Accompanying MS analyses shows the presence of a 50433 Da protein in 0 h, 15 min, 3 h and 4 h spectra but it is not detected in 30 min, 1 h and 2 h spectra which may represent an error in sample preparation (**Figure 3.17**). This stable fragment is larger than the fragment of CNOT11 (**Figure 3.14**, ~30 kDa) which was produced during limited proteolysis of CNOT10:CNOT11. Hence, it can be concluded that the fragment of CNOT11 produced after proteolysis of CNOT10:CNOT11 is due to the interaction with CNOT10.

CNOT11 has been shown to interact with itself in a Y2H screen (Mauxion et al., 2013), therefore, the protected fragment from limited proteolysis of CNOT11 alone, observed in **Figure 3.18** (boxed), may represent a dimerisation domain of CNOT11. The exact mass of the stable fragment (50433 Da) has been used to map a region within the sequence of CNOT11 that likely represents this fragment. This 468-residue region spans from part way through the His_6 towards the C-terminus (aa 14-457). However, it is unknown if this dimer is biologically relevant and if it exists as such within the CCR4-NOT complex. Thus, further experiments are required to determine the nature of the interaction between CNOT11 homodimers.

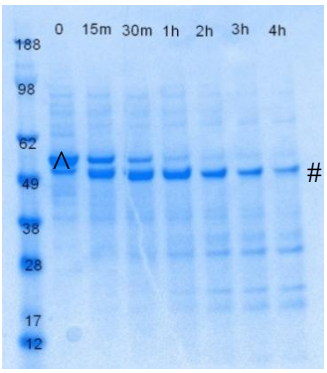


Figure 3.17 Limited proteolysis of CNOT11 alone

Limited trypsin proteolysis on CNOT11 alone. Trypsin was added to CNOT11 at a ratio of 1:400 and the sample was incubated for 4 h. Time course samples were taken. Intact CNOT11, indicated by the caret, was observed in the 0 lane. This was degraded to a ~50 kDa band, indicated by the hash. Samples were analysed by 4-12% SDS-PAGE. Molecular weight markers are indicated for size comparison (kDa).

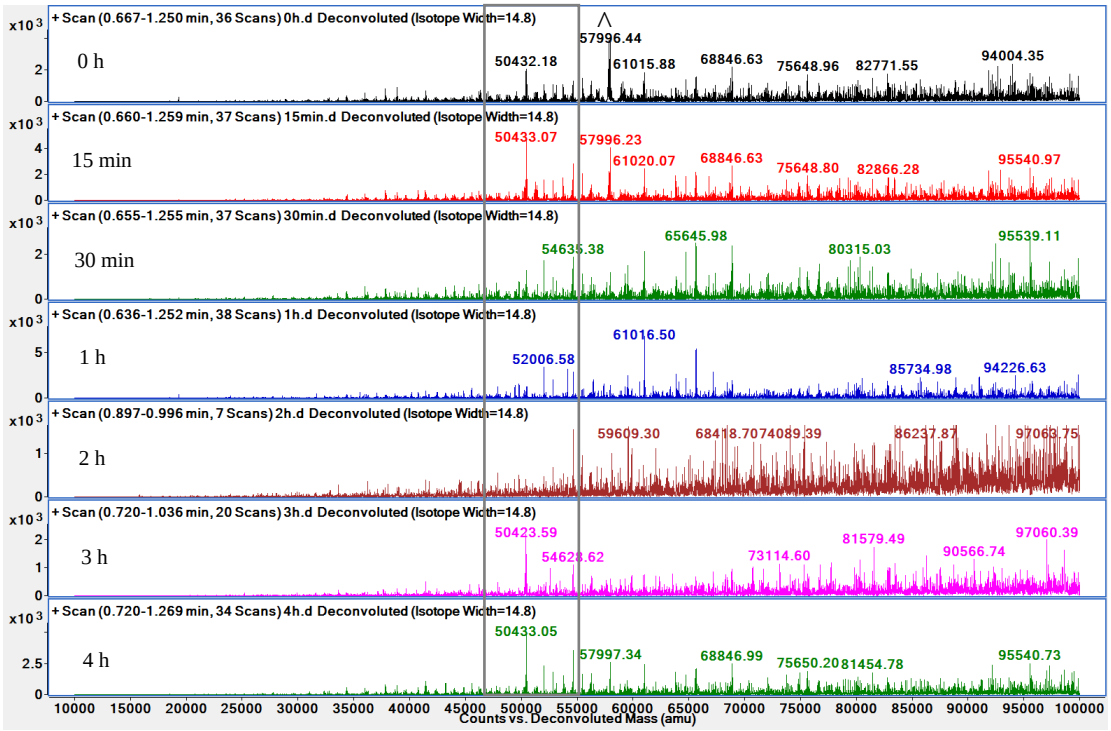


Figure 3.17 Stacked mass spectra of limited proteolysis of CNOT11

Denaturing MS samples were taken at the following time intervals: 0 h, 15 min, 30 min, 1 h, 2 h, 3 h, 4 h. Samples were run on a QTOF (Agilent) and analysed with Mass Hunter (Agilent). Stacked spectra show the gradual degradation of the full length CNOT11 (55308 Da, caret) and the production of the stable domain observed on SDS-PAGE (50433.05 Da, boxed).

3.9 Determining the interacting regions of CNOT10:CNOT11

The data above from analytical gel filtration (**Figure 3.16**) demonstrated that the proteolysed domains of CNOT10 and CNOT11 form a complex. To suggest possible boundaries of the truncated domains from limited proteolysis, denaturing MS and associated analysis software were utilised. The exact masses of fragments from limited proteolysis (46356.9 Da and 30729.3 Da, **Figure 3.16**) were mapped onto sequences of

CNOT10 and CNOT11 proteins to identify matching regions. A number of near matches (± 2 Da, underlined in different colours) of possible sequence regions were found for both CNOT10 and CNOT11, with the closest match for both proteins underlined in red, with triangles indicating the terminal residues (CNOT10, **Figure 3.19** and CNOT11, **Figure 3.20**). From this, the predicted boundary for CNOT11 (aa 214-486) encompasses the DUF2363 domain (aa 391-475) which may represent the minimal binding domain of CNOT11 for CNOT10. The predicted sequence for CNOT10 fragment (aa 33-444) would indicate a large C-terminal truncation has been made, thus the N-terminal domain may be important for the interaction with CNOT11.

As a tentative validation of these boundaries, the list of detected peptides used to assign identity during tandem MS (red text in **Figure 3.19** and **Figure 3.20**) were compared with the predicted matching sequences, with the rationale that these peptides should fall within the proposed sequences. The CNOT10 peptides (red text) detected from tandem MS fall within most of the predicted sequences (**Figure 3.19**) but the CNOT11 peptides (red text) do not fall within the predicted sequences (underlined in many colours) **(Figure 3.20)**. A peptide at the extreme C-terminus of CNOT11 (aa 496-510) was detected by tandem MS but is not encompassed in any of the predicted sequences. Hence, further work is required to clarify the minimal region on CNOT11 required to bind CNOT10.

Figure 3.18 MS analysis of possible sequences of proteolysed CNOT10

Mapped regions of possible sequences of truncated CNOT10 protein. Sequence of full length proteins (black text) is underlined to indicate the different proposed matches with the closest match underlined in red with triangles indicating the terminal residues. The sequences of peptides detected by tandem MS are highlighted in red.

The predicted sequences of CNOT10 and CNOT11 fragments (underlined in red, **Figure 3.19** and **Figure 3.20**) represent novel boundaries which have not been trialled for construct design and crystallisation in this study. Therefore, these amino acid boundaries were pursued in the hope of producing soluble interacting domains which may contain crystallisable regions.

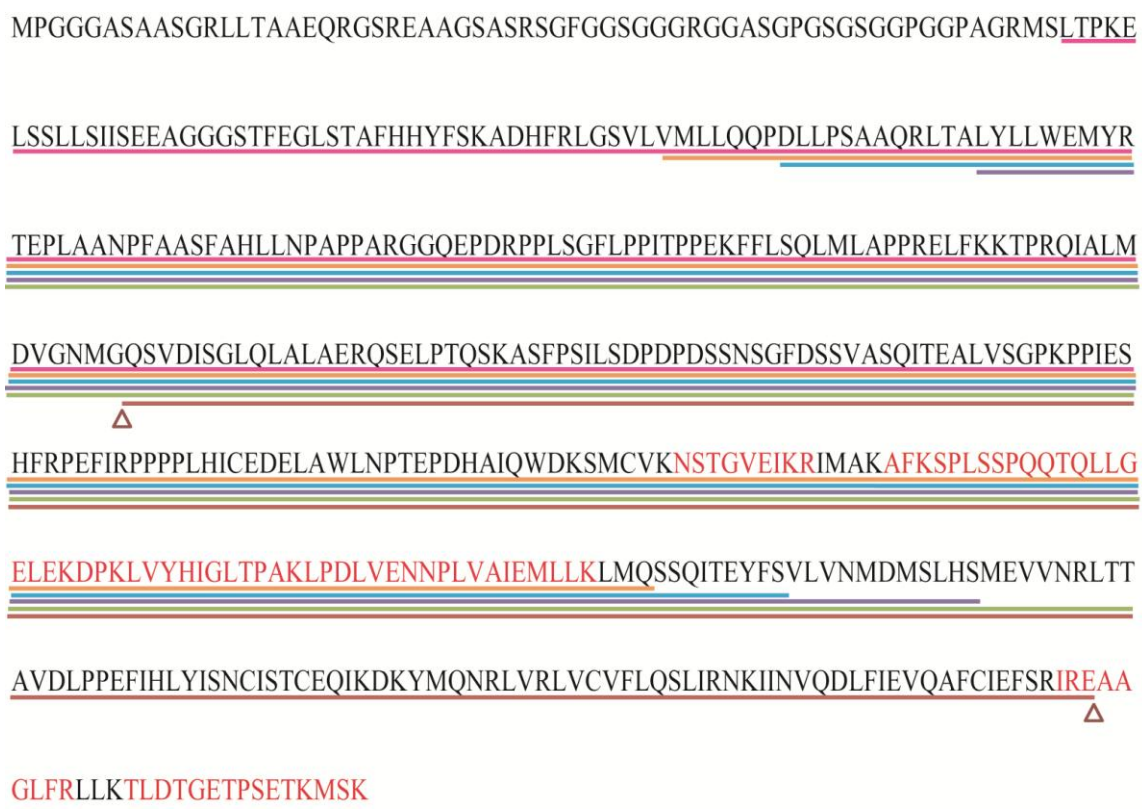


Figure 3.19 MS analysis of possible sequences of proteolysed CNOT11

Mapped regions of possible sequences of truncated CNOT11 protein. Sequence of full length proteins (black text) is underlined to indicate the different proposed matches with the closest match underlined in red with triangles indicating the terminal residues. The sequences of peptides detected by tandem MS are highlighted in red.

3.10 Generation of bicistronic constructs based on protein binding domains of CNOT10:CNOT11

Significant truncations were made to both CNOT10 and CNOT11 during the limited proteolysis used to mimic possible fungal degradation. Consequently, new *E. coli* constructs were designed to clone the sequence region suggested by MS data in **Table 3.3** and **Figure 3.19** and **Figure 3.20**. These constructs may contain the domains found in the original crystal hit of CNOT10:CNOT11. Re-cloning the truncated domains alone will allow preparations to be made for crystallisation without the need to proteolyse samples, as this may introduce variations between batches and reproducibility problems. The proteins were designed to be expressed bicistronically i.e. under one promoter in the *E. coli* expression system, thus allowing the CNOT10:CNOT11 complex to form post-

translationally (**Table 3.3**). It is the hope that as with the co-infection in the insect cell expression system, this will aid the solubility and stability of the respective proteins.

Construct	Gene 1 (His_6 tagged)	Gene 2 (untagged)
1	CNOT11 (aa 214-486)	CNOT10 (aa 33-444)
2	CNOT10 (aa 33-444)	CNOT11 (aa 214-486)

Table 3.3 Bicistronic constructs of CNOT10:CNOT11 for *E. coli* expression

New construct boundaries for CNOT10 and CNOT11 were designed based on information gained from the limited proteolysis experiments. New design will incorporate both genes under the same promoter to co-express the complex bicistronically.

3.11 Conclusion

This chapter describes the reconstitution, crystallisation and MS experiments employed to further our understanding of the newly identified CNOT10:CNOT11 module of the CCR4-NOT complex. The results presented here corroborate the interaction presented by Mauxion et al (Mauxion et al., 2013) but also demonstrate, for the first time, the reconstitution of the human CNOT10:CNOT11 sub-complex using recombinant proteins. The importance of the interaction to the solubility and expression levels has been demonstrated as co-infection of CNOT10 and CNOT11 viruses solubilises the otherwise insoluble CNOT10 and increases its protein yield.

Furthermore, the boundaries of the interacting domains of CNOT10 and CNOT11 were explored by exposing the recombinant complex to trypsin limited proteolysis, in order to characterise the nature of the interaction between the two subunits. Truncated regions of CNOT10 and CNOT11 were identified and subsequently confirmed to bind in solution. The sequences of truncated CNOT10 and CNOT11 have been tentatively mapped to CNOT10 (aa 34-444) and CNOT11 (aa 214-486). These regions may constitute previously unknown protein binding domains of CNOT10 and CNOT11. To validate that

these proposed sequences constitute the interaction domains of the respective proteins, additional experiments (e.g. BLI/surface plasmon resonance (SPR)/ITC) must be performed. Moreover, the interface between the domains should be studied by structural biology. Structural characterisation of CNOT10 and CNOT11 proteins, alone and in complex, has been attempted but these newly identified protein binding domains provides a new focus for crystallisation trials in the future.

Guided by previous reports that CNOT11 forms homodimers (Mauxion et al., 2013), limited proteolysis and MS have also been used to characterise the dimerisation of CNOT11, and preliminary results reveal a 468 residue region which is protected from proteolysis. However, this initial finding requires further exploration encompassing structural and binding data to determine the minimal region required for dimerisation. Firstly, solving the structure of CNOT11 will reveal whether a dimer is present in the crystal, involving the predicted interfacial region. Should this approach fail, pull-down experiments using overlapping fragments of recombinant CNOT11 proteins may also bring into focus the minimal domain for dimerisation.

This initial analysis of the CNOT11 dimerisation region also raises questions concerning the interaction between CNOT11 and CNOT10. The sequence protected from proteolysis of CNOT11 alone overlaps partly with the sequence of CNOT11 found to interact with CNOT10. Thus, if CNOT11 dimers are biologically relevant, and the interface for interaction is larger than that of the CNOT10:CNOT11 interaction, then it would be interesting to study what causes the transition between these different interactions. A similar transition of homo- to hetero-dimerisation is observed with another CCR4-NOT subunit, CNOT9. CNOT9 exists as a free dimer in solution but is recruited to the CNOT complex via an interaction with CNOT1 at the same interface as the CNOT9 dimer forms, as shown by crystallographic modelling (Garces et al., 2007, Mathys et al., 2014,

Chen et al., 2014). An interesting future experiment may include a competition assay with the CNOT11 dimer and excess CNOT10 (if it can be produced alone) to determine if CNOT10 binding would disrupt the CNOT11 dimerisation.

With regards to functional characterisation of CNOT10:CNOT11, much work remains to be done. Published studies of CNOT11 reveal interactions with both CNOT10 and CNOT1 and thus, it may function as a linker to recruit CNOT10 to the CCR4-NOT complex. *Tb*CNOT10 has been shown to interact with catalytic subunit *Tb*CNOT7 and be important for the stability of the complex, but no such interaction between human CNOT10 and CNOT7 has been established to date. Reconstitution of the recombinant CNOT10:CNOT11 complex, as demonstrated in this study, now provides the platform to determine if a similar role is found of human CNOT10. Moreover, an experiment where the protein levels of CNOT10/CNOT11 are knocked down in cells may also help to assign a possible function to this sub-complex.

One possible function of the CNOT10-CNOT11 sub-complex may be drawn from the large amount of nucleic acids which were often observed during CNOT10:CNOT11 purification, as this may indicate a function in binding DNA/RNA. Exploring this concept in future may provide some initial directions. It has been shown, by x-ray crystallography and in-gel DNA interaction assays, that CNOT9 contains a positively charged cleft which binds nucleic acids, though no specificity has been found for poly-(A) nucleotides (Garces et al., 2007). It may be possible that CNOT10 also binds nucleic acids. An electrophoretic mobility shift assay (EMSA) may be useful to validate this speculation in future.

4. Characterising the interaction between TTP and CCR4-NOT subunits

4.1 Introduction

The CCR4-NOT complex mediates targeted deadenylation of mRNA by interaction with RNBPs. For example, TTP binds to AREs in the 3′-UTR of inflammatory mRNAs using a central zinc-finger domain, and then recruits the CCR4-NOT complex to deadenylate the mRNA cargo. Previous work in Jonathan Dean's group used siRNAs targeted against each of the CNOT subunits to knock down protein levels, and determined its effect on recruitment of TTP to the complex (section 1.6, **Appendix Figure 7.11**). When CCR4-NOT subunits CNOT9 and CNOT2 were knocked down individually, less TTP was recruited to the complex (Bulbrook, 2015, DPhil thesis page 150). These data suggested that novel interactions between TTP and CNOT9/CNOT2 are important for the formation of the complex. Hence, this project focuses on validating the interactions between TTP and CNOT2/9 using recombinant proteins, but also characterising the nature of the interaction by establishing the domains or motifs required to form the TTP-CNOT2/9 sub-complexes.

This chapter describes the use of BLI, SPOT peptide arrays, and site-directed mutagenesis to resolve the mode of binding between TTP and CNOT2/9. Increasing our understanding of the seemingly multi-faceted interface between TTP and the subunits of the CCR4-NOT complex may provide spatial and temporal information on TTP recruitment. If additional subunits, beyond the published interaction with CNOT1, are required to bind TTP, determining the nature of the novel interactions with subunits may

be important to further explain the mechanism of targeted deadenylation in inflammatory mRNAs.

4.2 CNOT9-binding domain analysis of TTP

4.2.1 Construct design

TTP is a 326-residue protein which consists of three main domains: an N-terminal domain (aa 1-98), a C-terminal domain (aa 171-326), and a central zinc-finger domain (aa 99-170). To identify which domain(s) of TTP are responsible for binding CNOT9, four TTP constructs were generated encompassing (i) the N-terminal domain and zinc-finger domain (aa 14-171), (ii) zinc-finger and C-terminal domain (aa 98-326), (iii) zinc-finger only-domain (aa 98-171), and (iv) a near full-length (FL) protein (aa 14-326) (**Figure 4.1**). All constructs included the zinc-finger domain because without it, the recombinant proteins were unstable and easily degraded. The above boundaries were chosen to allow the contribution of each domain to be studied. Small scale expression/purification tests showed that the solubility of TTP in *E. coli* expression was

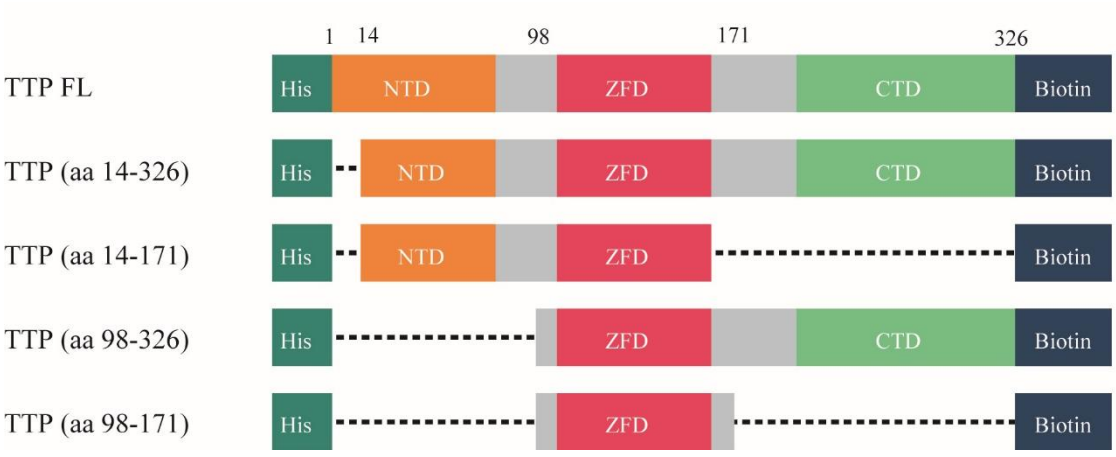


Figure 4.1 TTP domain constructs used for BLI analysis

TTP is composed of three main domains: a central zinc-finger domain (ZFD, red) and flexible N- and C-terminal domains (NTD, orange and CTD, light green). Constructs contained an N-terminal His_6 tag (dark green) to allow capture by nickel affinity and a C-terminal biotin recognition sequence (blue) to allow in vivo biotinylation during recombinant expression.

improved if the N-terminal 13 residues were removed (data not shown). Thus, this region (aa 1-13) was excluded from subsequent construct design.

4.2.2 Purification of biotinylated TTP constructs for interaction experiments

Constructs detailed in **Figure 4.1** were designated to be analysed using BLI for their ability to bind CNOT9, which required them to be biotinylated. In this study, proteins were biotinylated in vivo using the BL21-DE3-pRARE2-BirA *E. coli* strain which encodes a separate plasmid for a biotin transfer enzyme, BirA. Biotin was supplemented during bacterial growth, allowing BirA to add a biotin molecule to the C-terminal Avi tag provided by the pNIC-Bio3 vector. Incorporation of biotin was determined using intact MS which highlighted a mass increase of +244 Da. C-terminal biotinylation allows proteins to be attached to the BLI biosensors, for association and disassociation of CNOT9 to be measured.

Biotinylated TTP proteins were purified from 3 L of *E. coli* culture by nickel affinity chromatography (**Figure 4.2**). Only a single purification step was implemented for the biotinylated constructs as additional purity was achieved when the proteins were selectively bound to the streptavidin biosensor during a BLI experiment. The interaction between biotin and streptavidin is one of the strongest known in nature (K_d in the femtomolar range), and this ensured only the biotinylated TTP protein, and not the contaminants or unbiotinylated protein, bound the biosensor.

The expression level of two TTP constructs containing the C-terminus (aa 98-326 and aa 14-326) was low, and many degradation bands smaller in mass than the expected target protein band were observed (asterisks in **Figure 4.2**, right top and bottom). The flexible C-terminus of TTP is easily degraded thus purifications are rarely homogenous, an observation reported by other groups working in the TTP field (Taylor et al., 1995,

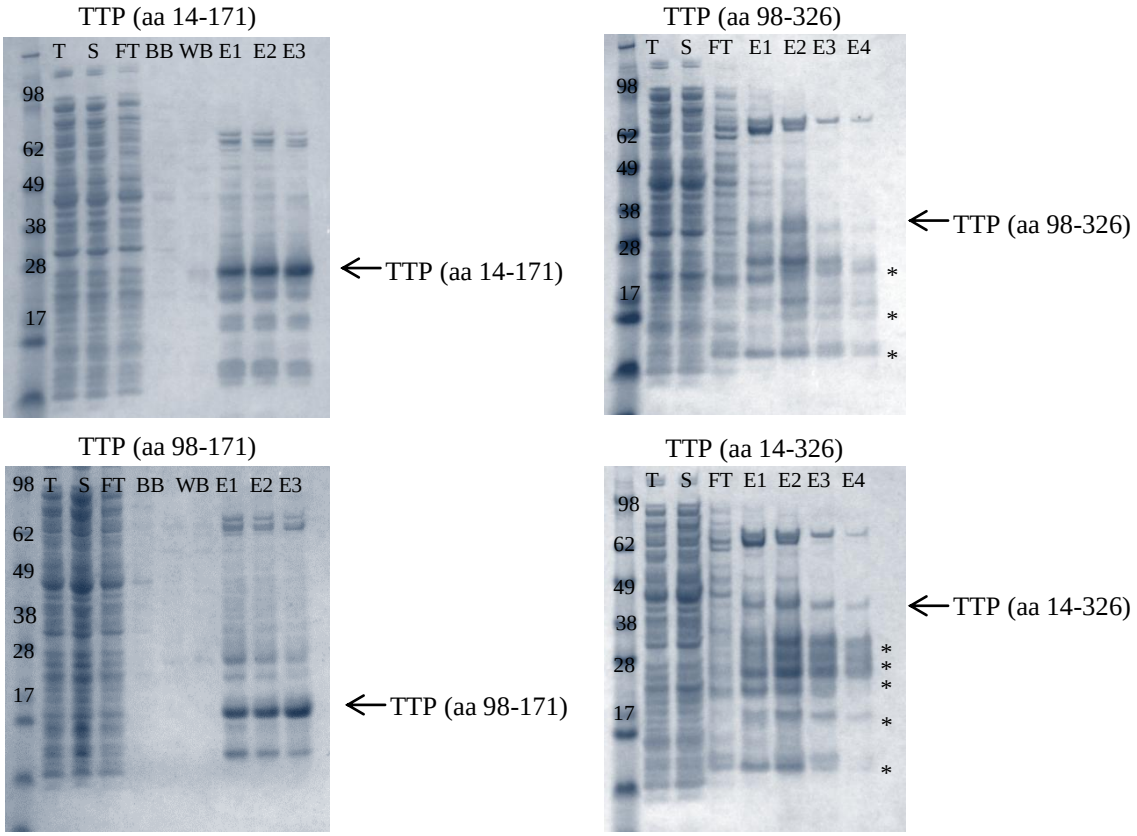


Figure 4.2 Nickel affinity purification of TTP domains for BLI

SDS-PAGE gels showing the purification of four biotinylated TTP domains (aa 98-171, aa 14-171, aa 98-326 and aa 14-326) purified from 3 L of *E. coli*. Fractions of total cell lysate (T), the soluble supernatant after centrifugation (S), the protein which did not bind nickel beads (FT), washes (BB, WB) and elutions (E1-E4) steps were visualised by SDS-PAGE to determine the purity of samples. Arrows indicate the position of the intact proteins and asterisks indicate TTP degradation products that were verified by tandem MS.

Mahtani et al., 2001). However, as the biotin tag is incorporated as a C-terminal fusion, only proteins with an intact C-terminus would bind to the biosensor, allowing the selective removal of some TTP degradation products. The expression and purity of two TTP constructs containing the N-terminus (aa 14-171 and aa 98-171; **Figure 4.2**, left top and bottom) were higher than those containing the C-terminus, with purity again improved by streptavidin capture during BLI. A band of approximately 62 kDa was observed in all purifications of TTP domains, and was identified to be an endogenous *E. coli* chaperone by tandem MS. It may be possible that the chaperone associated with TTP during recombinant expression due to its intrinsically disordered regions.

To simulate the potential purity of the three most heterogeneous TTP constructs (aa 14-171, aa 98-326 and aa 14-326) when TTP becomes bound to the biosensor during BLI, a streptavidin-mediated affinity purification was carried out (**Figure 4.3a**). The streptavidin affinity step has greatly improved the purity of TTP by enriching the TTP protein, which now represents the major species, and reducing the degradation products. Hence, this would represent more appropriately the quality of sample being subjected to interaction with partner proteins during a BLI experiment. As the TTP zinc-finger domain-only protein (aa 98-171) purified more homogeneously than other constructs, streptavidin purification was not required to check protein quality and a sample of the concentrated protein is shown in **Figure 4.3b**. In summary, the four domains of TTP were successfully biotinylated and purified to provide material for the interaction with CNOT9 to be studied.

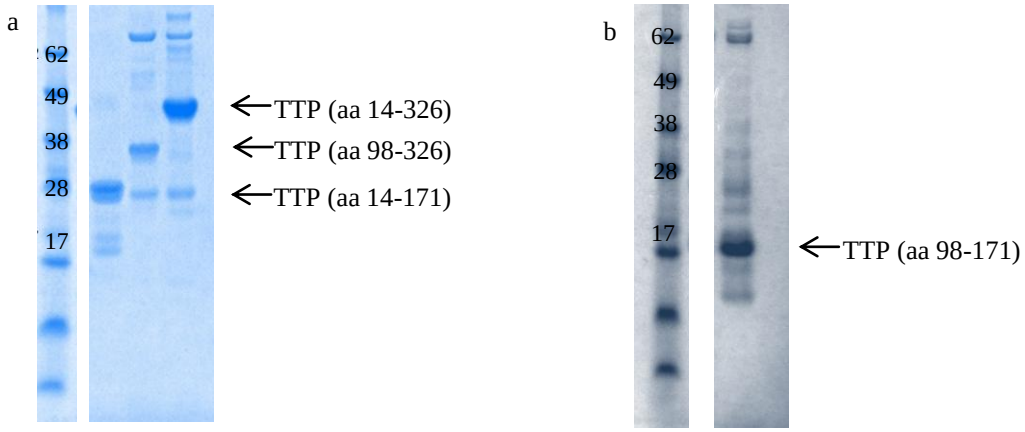


Figure 4.3 Purified TTP domain constructs used for BLI

(a) Streptavidin-mediated affinity purification of TTP domains (aa 14-171, aa 98-326 and aa 14-326). (b) Post-affinity purification sample of TTP zinc-finger domain-only protein (aa 98-171). Arrows indicate the position of the intact proteins.

In subsequent BLI experiments, the interaction between TTP and CNOT9 was measured, using full length His_6-tagged CNOT9 that was isolated from 6 L of *E. coli* culture by nickel affinity chromatography (**Figure 4.4a**). CNOT9 was highly expressed with few contaminating proteins. Eluted protein (EB fractions) was gel filtered using an S75 16/60 column to remove contaminating proteins (**Figure 4.4b and c**). Fractions containing CNOT9 were pooled and concentrated to 80 µM for BLI experiments (**Figure 4.4d**).

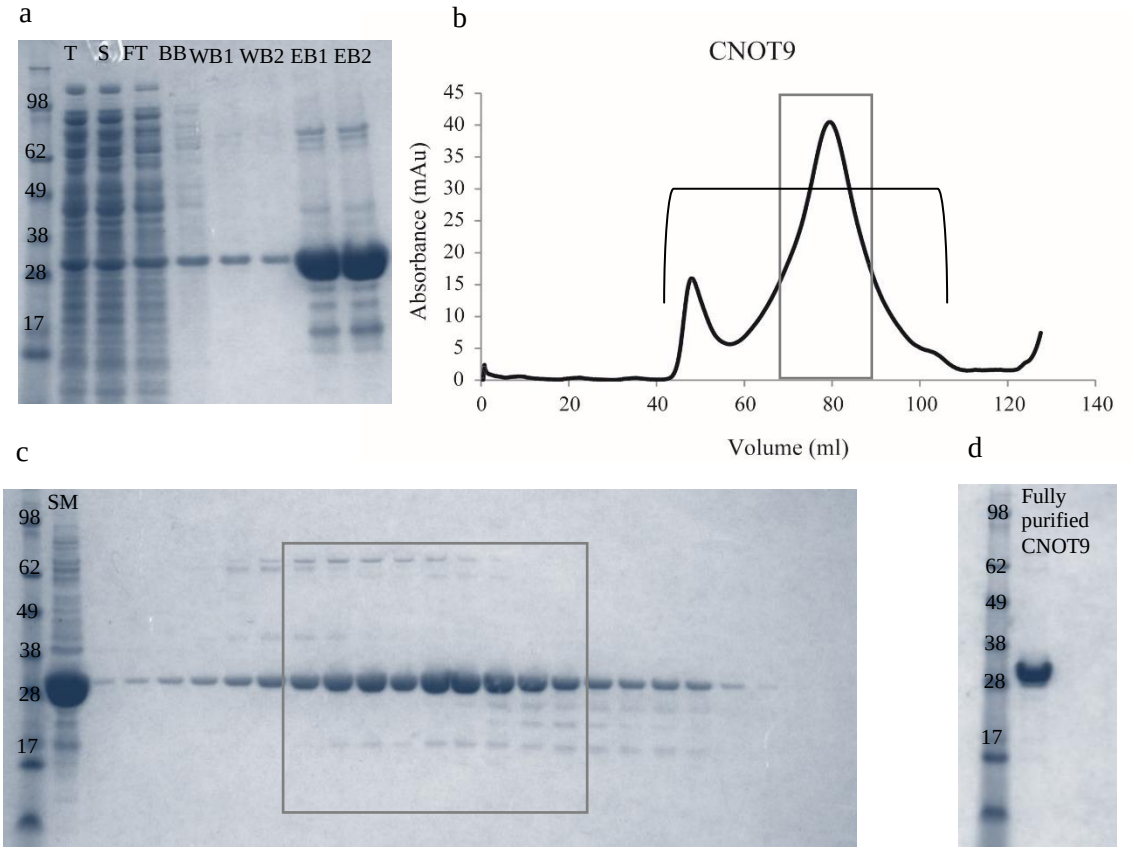


Figure 4.4 Purification of full length CNOT9 for TTP domain analysis

(a) SDS-PAGE showing fractions collected during nickel affinity purification of CNOT9 from 6 L of *E. coli*. (b) Chromatogram showing the gel filtration of fractions EB1 and EB2 pooled from the affinity step using an S75 16/60 column. The bracket indicates fractions which were visualised by SDS-PAGE in panel (c) and the box indicates fractions that were pooled. (c) SDS-PAGE of gel filtration fractions within the bracket indicated in (b). Starting material (SM) pooled from EB1 and EB2 is shown in lane 1. The box indicates the final pooled fractions. (d) SDS-PAGE gel showing concentrated fractions within the box in (c).

4.2.3 Analysis of TTP-CNOT9 interaction by BLI

BLI is an optical analytical technique used to measure protein-protein, protein-peptide and protein-small molecule interactions where dissociation constants (K_d) are between 1 mM-10 pM can be measured. Experiments were performed as described in section 2.22 using a biosensor coated with biotinylated TTP, exposed to varying concentrations of CNOT9 in solution. The change in wavelength (Δλ, nm), caused by the change in thickness of protein bound to the biosensor, was plotted against CNOT9 concentration (**Figure 4.5**).

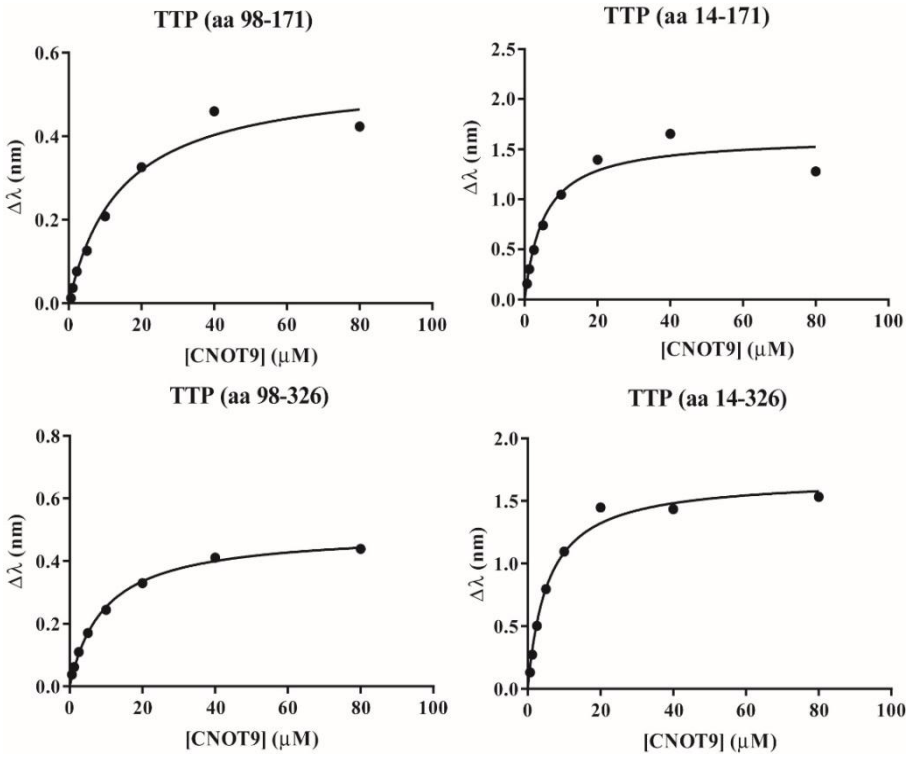


Figure 4.5 Plot of the change in wavelength (Δλ, nm) measured in response to CNOT9 binding TTP domains

The change in wavelength Δλ (nm) has been plotted against CNOT9 concentration (μM) and points fitted using a one-site total binding model. GraphPad Prism 7 has been used to determine dissociation constants for the TTP domains. Points are representative of an average of two independent technical replicates where protein from the same batch was attached to the biosensor and measurements taken against CNOT9.

The binding curves in **Figure 4.5** show an initial increase in the magnitude of the wavelength shift which is a direct measure of the number of bio-molecules bound to the

tip of the biosensor, thus, indicative of CNOT9 binding TTP. The increase in magnitude of wavelength is followed by a plateau, indicating that equilibrium between association and dissociation of CNOT9 has been reached. For a reliable fit, all data points should adhere to the line of best fit. However, in **Figure 4.5**, some data points do not follow the line of best fit as the data is only representative of two technical replicates. The reliability of the data collected may be improved by ≥3 repeats in future studies. In addition, to add greater weight to conclusions drawn from the data, significance could be determined using statistical tests (e.g. an unpaired t-test or one-way ANOVA) in future experiments where multiple data sets are collected. To indicate the variation from the mean, standard deviations could also be calculated in future studies, ideally with ≥3 repeats as standard deviations in data sets of n=2 typically underestimate population deviations.

The plots in **Figure 4.5** were used to calculate dissociation constants (K_d) of each of the TTP domains (see section 2.22 for equation) (**Table 4.1**). TTP proteins encompassing the N-terminal and zinc-finger domains (aa 14-171), the zinc-finger and C-terminal domains (aa 98-326), and the near-full length TTP (aa 14-326) bound CNOT9 to a similar micromolar affinity, whereas the zinc-finger domain-only protein (aa 98-171) bound CNOT9 two-fold weaker (**Table 4.1**). Therefore, all domains of TTP are involved in interacting with CNOT9, although dissociation constants suggest that the contribution of the zinc-finger domain to CNOT9 binding is less than that of the N- and C-terminal domains. Moreover, although dissociation constants for constructs containing the N- and C-terminal domains (aa 14-171, aa 98-326 and aa 14-326; **Table 4.1**) are somewhat similar, it could be interpreted that the N-terminus has a greater contribution to CNOT9 binding, as domains encompassing this region (aa 14-171 and aa 14-326) displayed moderately stronger binding (5.8 and 5.5 µM, respectively) than TTP protein containing the C-terminal domains (aa 98-326) (9.6 µM).

Construct	Dissociation constant (K_d) (μM)
TTP (aa 98-171)	17.3
TTP (aa 14-171)	5.8
TTP (aa 98-326)	9.6
TTP (aa 14-326)	5.5

Table 4.1 Dissociation constants of TTP for CNOT9

Dissociation constants (K_d) of TTP proteins measured against full length CNOT9. TTP proteins aa 14-171, aa 98-326 and aa 14-326 bind with similar affinities. The zinc-finger only-domain (aa 98-171) protein binds 2-fold weaker.

4.3 Identifying CNOT9-binding TTP motifs using SPOT peptide arrays

4.3.1 Probing a TTP SPOT peptide array with CNOT9

The dissociation constants presented above indicate that there may be multiple CNOT9-binding motifs located throughout the sequence of TTP, thus a method that could further explore these individual regions was required to evaluate their individual contributions. The SPOT peptide array method was enlisted to further probe the individual motifs involved in binding, as it is a time- and cost-effective way to map epitopes, but also bypasses the need for soluble expression of difficult-to-purify proteins e.g. TTP (Hilpert et al., 2007, Briant et al., 2009). As TTP has been shown to contain large regions of disorder, it can be assumed that each peptide is somewhat independent from others in terms of the 3D environment, thus mapping epitopes constitutes a highly relevant method to distinguish binding mechanisms. The array was formed of a nitrocellulose membrane onto which 15-residue TTP peptides were printed in a grid format. Each spot contains a single 15-mer peptide sequence and consecutive spots were staggered by three residues in order to cover the entire sequence of TTP (**Figure 4.6**). Therefore, for a 326-aa TTP protein, 109 spots are required to provide complete sequence coverage. The sequences of all the spot peptides and their locations on the membrane are listed in **Appendix Table 7.15.**

Full length His_6-tagged CNOT9 protein (1 µM, purified as described in Section 4.2.1 and **Figure 4.4**) was used as prey protein for binding to the TTP peptide array. TTP peptides which bound CNOT9 were identified by immunoblot using an anti-His antibody conjugated to a horseradish peroxidase (HRP) enzyme for chemiluminescence (**Figure 4.7**). Thus, the more intense the SPOT 'colour', the stronger the CNOT9 bound to that particular peptide sequence. The anti-His antibody bound strongly to the His_{10} positive control peptides (spots A1 and A2, **Figure 4.7**), to ensure that the antibody detection is functional.

Figure 4.7 Schematic depicting the overlap between consecutive peptides printed on the TTP peptide array

The array was printed with 15 residue TTP peptides. Each spot contains a single 15-mer peptide sequence and consecutive spots are staggered by 3 residues, shown in turquoise. Arrows indicate the position of these example peptides in context of the whole TTP sequence.

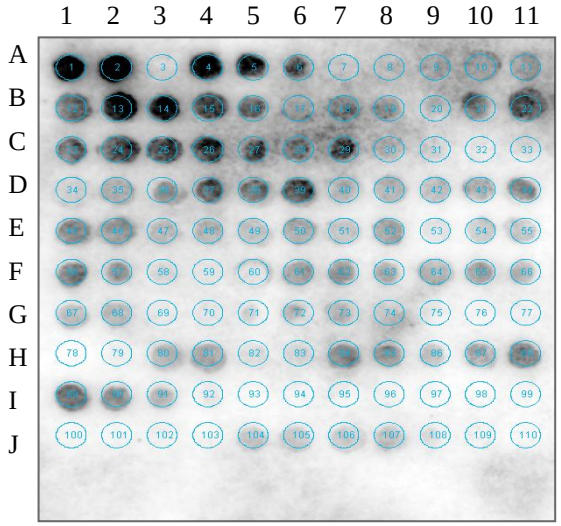


Figure 4.7 TTP peptide array probed with CNOT9

Intensities of immunoblotting when probing the TTP peptide membrane array for bound His_6-CNOT9 using an anti-His HRP conjugate were detected on a Gel Logic 200 Imaging System. Numbers and letters provide an identifying code for all peptides in the array, the sequence of each is given in **Appendix Table 7.15**. Positive controls (His_{10}) are shown in position A1 and A2, followed by a negative control (no peptide) in A3.

The intensity of each individual spot, relative to His_{10} control peptides which were taken as 100% intensity, was plotted on a bar graph to determine any CNOT9-binding hotspot regions of TTP (**Figure 4.8**). To be considered an area of interest, the intensity of adjacent spots should rise and fall in a manner similar to a normal distribution, indicating a motif that binds CNOT9. Five such motifs were observed; blue arrows indicate the summit of each intensity peak which is labelled based on the number/letter code in **Appendix Table 7.15**. Peptides corresponding to these motifs mostly fall within the N-

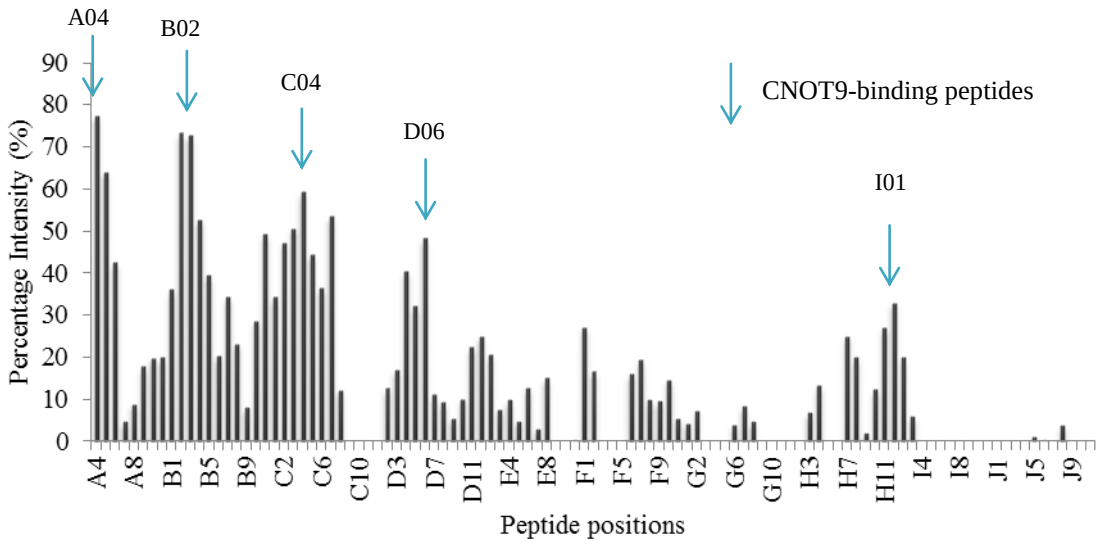


Figure 4.8 TTP peptide array intensities

The immunoblotting intensities for each of the peptides in the SPOT array calculated as a percentage relative to the control His_{10} peptides and plotted on a bar graph. Blue arrows highlight the central peptide in a motif identified to bind CNOT9. Refer to **Appendix Table 7.15** for the identity of each peptide.

terminus of the TTP polypeptide (A04, B02, C04, D06 in **Figure 4.8**), with one binding region in the C-terminus (I01 in **Figure 4.8**). This supports the initial finding from BLI using recombinant proteins (section 4.2.2) that regions in both the N- and C-termini of TTP contribute towards CNOT9 binding. It also supports the observation that TTP protein with an N-terminal domain binds CNOT9 with stronger affinity than those with a C-terminal domain (**Table 4.1)**.

The individual peptides (A04, B02, C04, D06 and I01) that bound CNOT9 in **Figure 4.8** were mapped onto the sequence of TTP (**Figure 4.9,** underlined). Three (B02, C04, I01) of the five peptides contain tryptophan residues, shown in blue. CNOT9 is known to interact with other binding partners (e.g. GYF and GW182) in a tryptophan-mediated manner where tryptophan residues in GYF and GW182 are recognised by tryptophan-binding pockets in CNOT9 (Chen et al., 2014, Mathys et al., 2014, Ajiro et al., 2009). Therefore, it was hypothesised that CNOT9 may also interact with TTP using the same tryptophan-binding pockets. To investigate this hypothesis, the individual contribution of the residues within the marked peptides (A04, B02, C04, D06, I01 in **Figure 4.9)** were investigated by alanine-scanning mutagenesis coupled with a peptide array.

Figure 4.9 TTP sequence annotated with the CNOT9-binding peptides

The amino acid sequence of human TTP (Uniprot accession: P26651) has been annotated with peptides from **Figure 4.8** shown to bind CNOT9. The sequence of each peptide has been underlined and tryptophan residues are indicated in blue.

4.3.2 Alanine scanning TTP peptide array probed with CNOT9

An alanine scanning peptide array was used to investigate the contribution towards CNOT9 binding of the individual residues within the TTP peptides. Alanine scanning is a process by which residues within a protein or peptide sequence are sequentially mutated to alanine, one by one, producing a library of proteins/peptides derived from the same

'parent' to assess the effect of each individual residue on a chosen outcome. In this instance, the peptides identified from **Figure 4.8** were sequentially mutated and the outcome measured was CNOT9 binding, as detected by immunoblotting with anti-His antibodies.

The SPOT-peptides were printed in a series of 16 spots, one series per peptide identified from **Figure 4.8** (A04, B02, C04, D06, I01). The first spot in each series of 16 contains the wild-type peptide and the subsequent 15 spots each had successive residues mutated to alanine, one by one, within the original 15-mer peptide sequence (**Figure 4.10**, shown in red). The peptide sequences and locations are detailed in **Appendix Table 7.16**. The intensity of the alanine mutant spots was measured (**Figure 4.10**, SPOT-peptides 2-16), and compared to that of the wild-type spot (**Figure 4.10**, SPOT-peptide 1) of each peptide. Decreased intensity of the mutant SPOT-peptide in comparison to the wild-type would indicate that the mutated residue was involved with binding CNOT9. The intensities of SPOT-peptides have been plotted as bar graphs to reveal any fluctuations in intensity (**Figure 4.11**).

Peptide D06 series (panel b, **Figure 4.11**) showed mostly increases in intensity, compared to the wild-type SPOT-peptide, when residues were mutated to alanine. This indicates that the alanine mutants bind CNOT9 more strongly than the wild-type peptide. Therefore, this motif may not be important for the binding of CNOT9 as no residue appears to negatively affect CNOT9 binding when mutated. On the contrary, series A04 (**Figure 4.11** panel a), B02 (panel c), C04 (panel e) and I01 (panel d) all showed some residues which may be important for CNOT9 binding: A04 (Tyr7, Leu10, Leu11), B02 (Trp32, Trp38), C04 (Arg65, Trp69, Phe76) and I01 (Trp262, Pro264). Within the sequence of the binding peptides, tryptophan residues Trp32, Trp38, Trp69 and Trp362 appeared to significantly reduce CNOT9 binding when mutated, suggesting that CNOT9

interacts with TTP in a tryptophan-mediated manner. Interestingly, single substitution of either Tyr7, Leu10 or Leu11 to alanine in peptide A04 decreased CNOT9 binding to < 20% of wild-type. The peptide A04 spans the first 15 residues of TTP, the deletion of which in construct design was found to increase expression of TTP constructs.

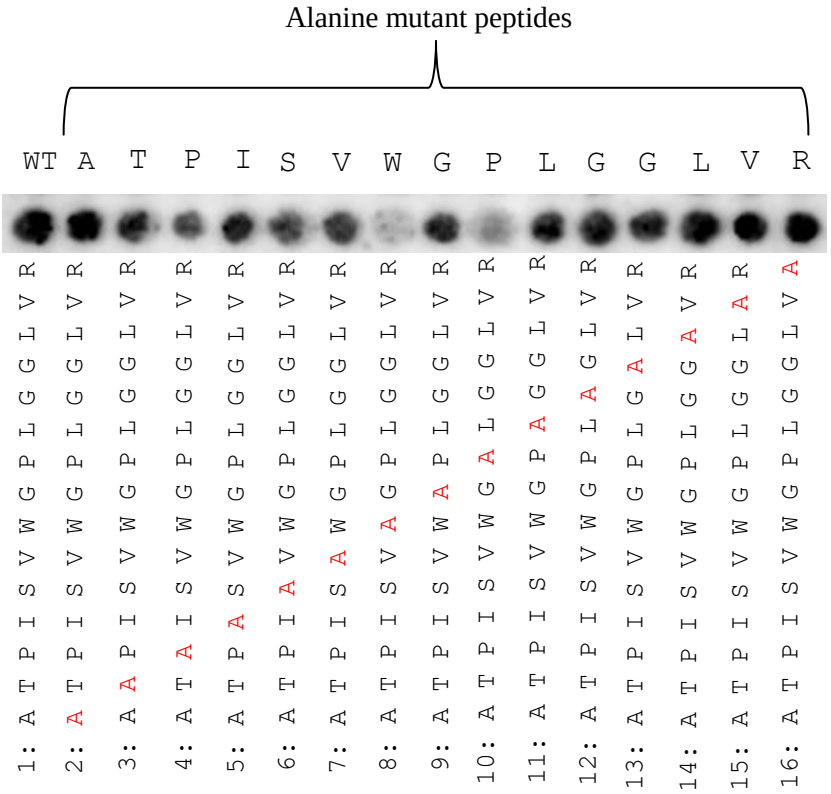


Figure 4.10 Schematic of sequential site-directed mutagenesis of SPOT-peptides printed for the alanine scanning TTP peptide array

A series of 16 SPOT-peptides were printed for each of the five identified peptides from **Figure 4.9**. In the series, the first peptide is wild-type (WT). Subsequent peptides (in brackets) have sequential residues mutated to alanine, shown in red. As an example, the figure shows the successive mutation of alanine in peptide I01. The numbers indicate the position of the SPOT-peptide within the sequence. The sequence of all SPOT-peptides is provided in **Appendix Table 7.16.**

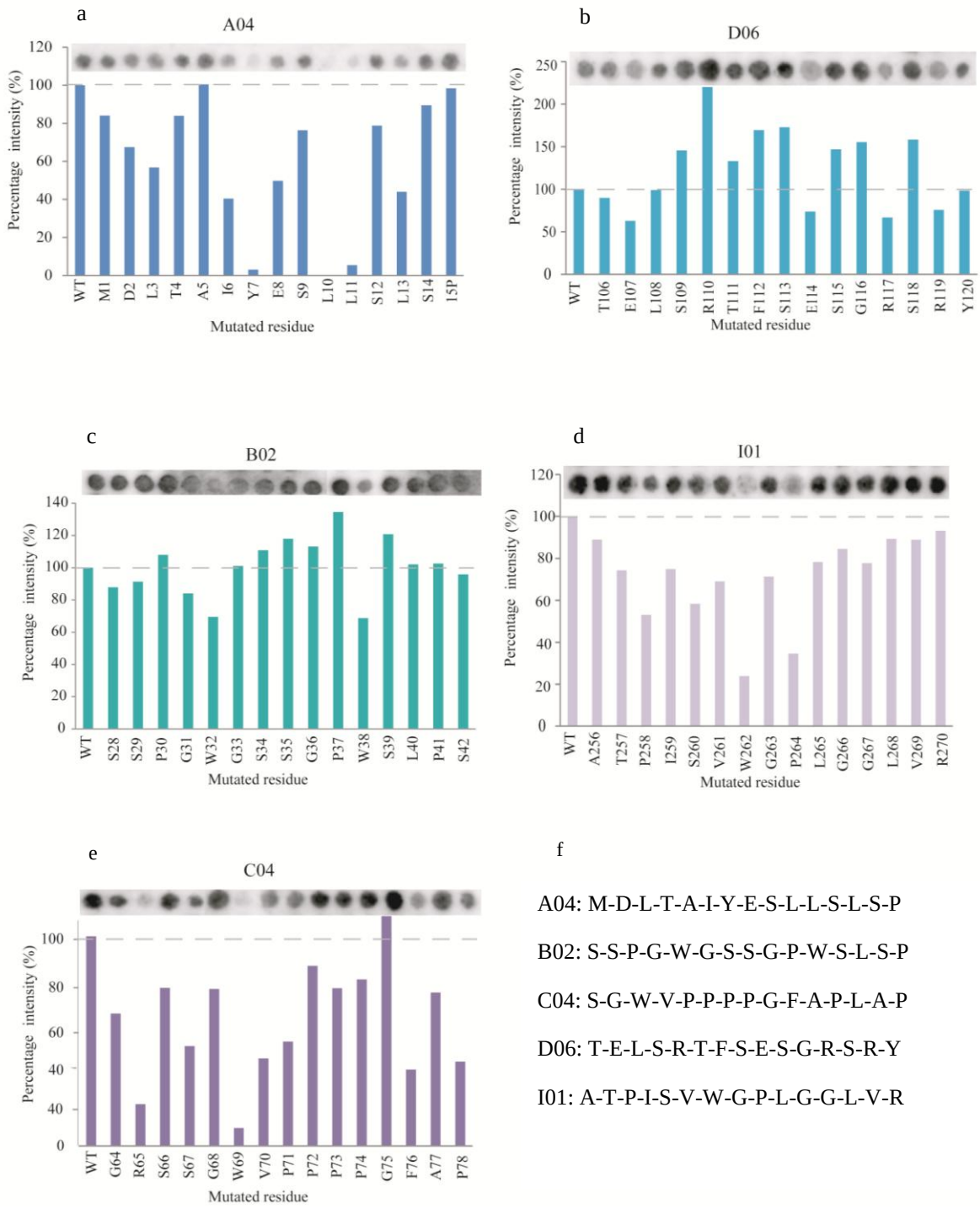


Figure 4.11 Alanine scanning TTP peptide array data

The intensity of the spot on the alanine scanning peptide array was measured and calculated as a percentage using the wild-type peptide in each series, which was taken as 100%, indicated by the dashed line. Each panel represents the alanine scan array of a different TTP peptide. Panel a represents A04, panel b represents D06, panel c presents B02, panel d represents I01, panel e represents C06 and panel f contains the wild-type sequences for each peptide. The probed arrays have been overlaid above the bar graphs. Each bar is labelled with the residue number which has been mutated to alanine within the particular 15-mer peptide.

Therefore, the binding of this motif to CNOT9 was not measured during the domain analysis of TTP constructs. Further investigation will be required, in order to ascertain whether this region contributes to CNOT9 binding along with other N-terminal motifs. As such, BLI was used to validate the binding of these peptides to CNOT9.

4.3.3 BLI analysis of CNOT9 interaction with TTP peptides

Although five TTP peptides have been verified by alanine-scanning arrays to be important for the interaction with CNOT9, an orthogonal method, in this case BLI, was desirable to further validate binding and calculate dissociation constants. Peptides A04, B02, C04, D06 and I01 were synthesised commercially (Lifetein) incorporating a C-terminal biotin tag. The biotin-tagged TTP peptides were attached to the streptavidin chip and binding was measured against full length CNOT9. Binding could not be detected for three of the peptides (A04, C04 and D06). This may suggest that these peptides represent false positives from the original SPOT peptide arrays. However, an alternate explanation for the lack of observable binding may be that the steric hindrance of a biotin tag may have affected the interaction with CNOT9. Peptide C04 was re-synthesised with an N-terminal biotin tag to test this possibility; though still no CNOT9 binding was detected.

Peptides B02 and I01 did show a response during BLI experiments in the presence of CNOT9, and dissociation constants were calculated (**Figure 4.12, Table 4.2**). Peptide B02 binds CNOT9 with a K_d of 56.5 μM, whereas I01 binds CNOT9 with a K_d of 9.3 μM. The 6-fold stronger binding towards CNOT9 of peptide I01, compared to B02, would indicate that the motif covered in peptide I01 has a greater contribution to CNOT9 binding than peptide B02.

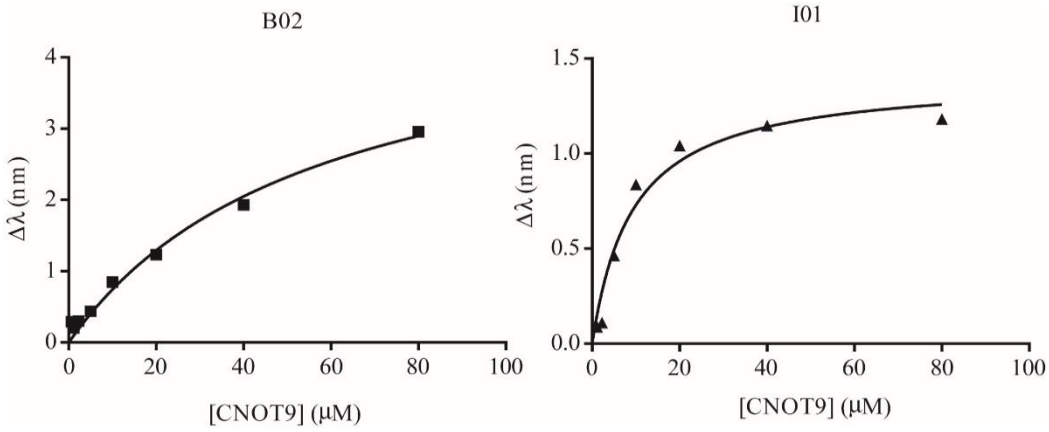


Figure 4.12 Plot of change in wavelength measured in response of CNOT9 binding to TTP peptides

Plotted points were fitted using a one-site total binding model. GraphPad Prism 7 has been used to determine dissociation constants for the TTP peptides. The experiment represents a technical replicate of n=2.

Peptide	Dissociation constant (K_d) (μM)
TTP B02	56.5
TTP I01	9.3

Table 4.2 Dissociation constants of TTP peptides for CNOT9

Dissociation constants of TTP peptides were measured using BLI. Errors are based on technical replicates of n=2.

The K_d value calculated for the peptide I01 is similar to that of the C-terminal domain of TTP (aa 98-326, K_d 9.6 μM) in **Table 4.1**. Thus, the residues within I01 (aa 256-270) may represent the motif within the C-terminal domain that is required to bind CNOT9. However, the K_d calculated for peptide B02 is approximately 10-fold higher than the N-terminal domains of TTP (aa 14-171, K_d ~5 μM). One explanation for the difference in dissociation constants may be that additional motifs in the N-terminus, beyond the sequence of B02, could contribute to CNOT9 binding. These additional residues may be the one other tryptophan residue in this region (Trp69), as tryptophan-containing peptides were found to be important for CNOT9-binding, as established by peptide arrays. Another explanation for the difference in dissociation constants may be that the tryptophans could bind in a co-operative manner, with peptide I01 being recruited first and B02 second, in which case the affinity of the second peptide is affected by the first.

To explore both of these possibilities, the effect of the tryptophans and tryptophan-binding pockets on the interaction between TTP and CNOT9 was next measured in BLI using recombinant full length proteins coupled with site-directed mutagenesis.

4.4 Exploring TTP and CNOT9 binding site using site-directed mutagenesis

4.4.1 Mutating the CNOT9 tryptophan-binding pockets

CNOT9 has previously been known to interact with a number of binding partners in a tryptophan-mediated manner (section 1.3.2). Crystal structures of human CNOT9, determined in complex with CNOT9-binding domain of CNOT1, have identified two tryptophan-binding pockets in the convex surface of CNOT9 (**Figure 1.8**) (Chen et al., 2014, Mathys et al., 2014). These two pockets (P1 and P2), shown to each bind a tryptophan ligand in the published structure, could represent physiologically relevant motifs used to bind CNOT9 partner proteins, including TTP identified here.

To explore the hypothesis that TTP binding to CNOT9 also engages with the tryptophan residues in a similar manner, amino acid substitutions in the CNOT9 tryptophan pocket were generated. Residues for site-directed mutagenesis were chosen to replicate published data which found that mutation of specific residues in tryptophan-binding pockets of CNOT9 successfully abolished interactions with TNRC6 (Mathys et al., 2014). These substitutions included: pocket 1 mutant (P1) (R205D, H208D), pocket 2 mutant (P2) (R244E, A248L) and a double pocket mutant (P1+P2) (R205D, H208D, R244E, A248L) (**Figure 4.13**).

Mutant and wild-type CNOT9 proteins were purified from 3 L of *E. coli* culture, using the same protocol as described for wild-type CNOT9 in **Figure 4.4.** Purified proteins were of a similar level of purity to wild-type (**Figure 4.13a**), thus any additional

contaminants present are similar among the mutant proteins and would unlikely cause any difference in binding between them. Gel filtration was used to demonstrate that the elution profile of the wild-type and mutant proteins were similar (**Figure 4.13b**), suggesting that the overall structural integrity of the proteins is unaffected by the mutations.

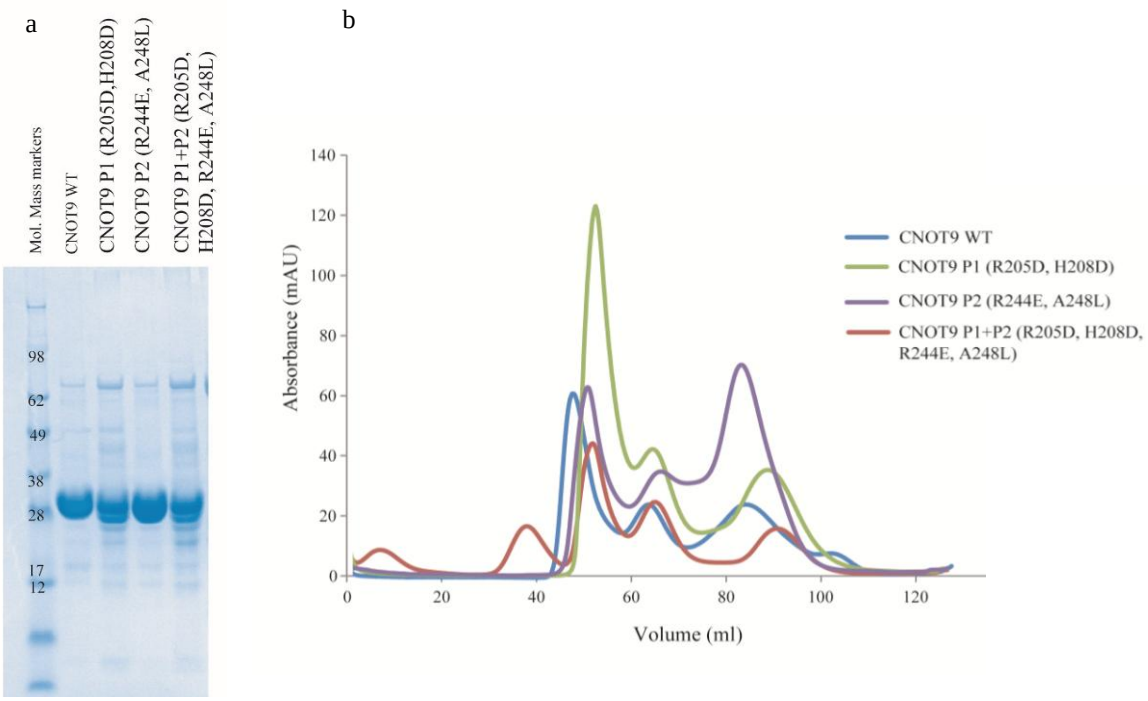


Figure 4.13 CNOT9 tryptophan pocket mutants used for BLI analysis

(a) SDS-PAGE gel displaying the purity of CNOT9 mutants used for BLI analysis. Three mutants were constructed to disrupt the tryptophan binding pockets in order to investigate contribution to binding. Each lane represents the protein sample after nickel affinity chromatography and gel filtration. (b) Chromatogram showing the gel filtration of nickel affinity elution fractions of CNOT9 WT, CNOT1 P1, CNOT9 P2 and CNOT9 P1+P2. Protein was gel filtered using an S75 16/60 column.

Binding of the purified CNOT9 protein (WT and pocket mutants) to biotinylated near-full length TTP (aa 14-326) was demonstrated by BLI in a single-concentration, time-course experiment (**Figure 4.14)**. The CNOT9 single pocket mutants (pocket 1, purple and pocket 2, green) bound 2-fold weaker to TTP than CNOT9 WT (blue), as indicated by the reduced $\Delta\lambda$ on the BLI biosensor for the mutants. As mutation of the single pockets only reduces rather than completely abolishes binding, this indicates that both

pockets are involved in binding TTP. Consistent with this, the double pocket mutant (red) displayed 2- to 3-fold weaker binding than the single pocket mutants.

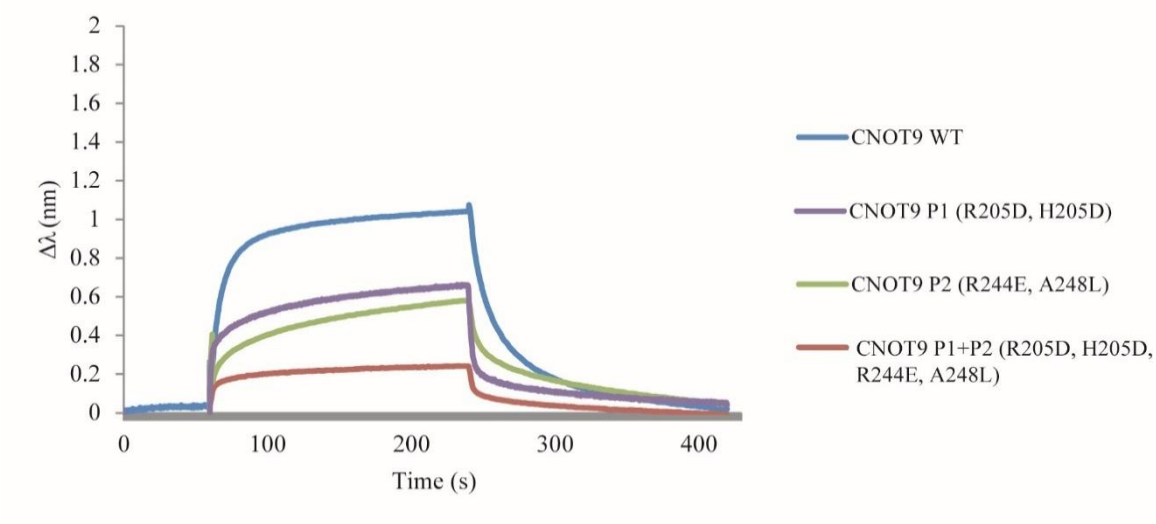


Figure 4.14 BLI association/dissociation trace of CNOT9 mutants

Binding curves for CNOT9 WT (blue), P1 mutant (purple), P2 mutant (green) and P1+P2 (red). A 60s baseline was established when the TTP-bound biosensor was exposed to gel filtration buffer, followed by an association step (180 sec) where the biosensor was exposed to CNOT9, allowing CNOT9 to bind. After association, the biosensor was moved to a well containing gel filtration buffer and dissociation was measured for 180 sec.

The dissociation constants for each mutant, calculated by titrating CNOT9 (**Figure 4.15**) (**Table 4.3**), reflect the trends observed in the association/dissociation trace in **Figure 4.14**, showing that both tryptophan pockets are directly involved in binding TTP. This result lends further support to the concept that CNOT9 interacts with TTP, as it does with other binding partners, in a tryptophan-mediated manner. This concept was next tested using mutants of TTP to determine if indeed the tryptophan residues are responsible for binding their cognate pockets in CNOT9.

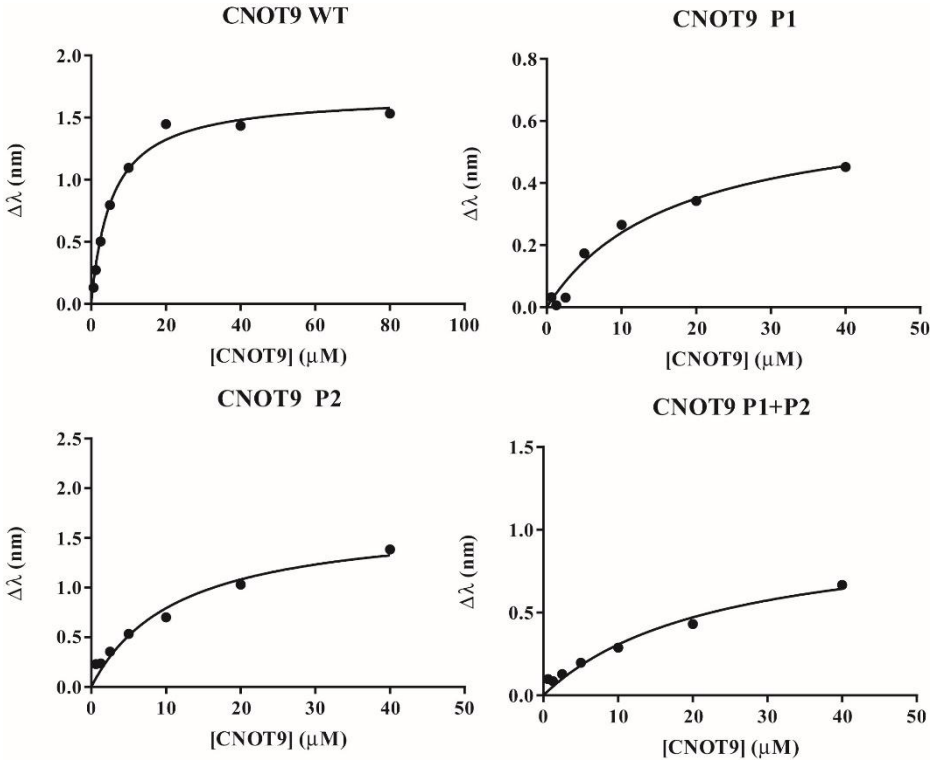


Figure 4.15 Plot of the change in wavelength (Δλ, nm) caused in response to TTP binding CNOT9 pocket mutants

The change in wavelength (nm), as measured by BLI, has been plotted against CNOT9 concentration and points fitted using a one-site total binding model. GraphPad Prism 7 has been used to determine binding affinities for the TTP domains. Points are representative of an average of two technical replicates.

Construct	Dissociation constant (K_d) (μM)
CNOT9 WT	8.3
CNOT9 P1 (R205D, H208D)	18.2
CNOT9 P2 (R244E, A248L)	18.3
CNOT9 P1 + P2 (R205, H208D, R244E, A248L)	23.1

Table 4.3 Dissociation constants of CNOT9 mutants

Dissociation constants were calculated from the curves in **Figure 4.15** for each CNOT9 protein (WT, P1, P2 and P1+P2).

4.4.2 Mutation of TTP tryptophan residues

Four tryptophan residues in TTP present in the peptides that were identified using the SPOT peptide array (Trp32 (peptide B02), Trp38 (peptide B02), Trp69 (peptide C04) and Trp262 (peptide I01)) were mutated to alanine, either individually (W32A, W38A, W69A, W262A) or together as a quadruple mutant ('4WA'), and binding was assessed by BLI using wild-type full length CNOT9. All TTP mutant proteins were biotinylated in vivo during expression, and purified from 3 L of *E. coli* culture. Although successfully cloned, TTP protein containing W69A degraded easily during nickel purification, and could not be used for the binding analysis. Trp69, located within peptide C04, was shown to bind CNOT9 by peptide arrays, but this peptide was not found to bind in BLI experiments (section 4.3.3).

Again, a streptavidin-mediated purification was implemented to indicate the purity of TTP proteins once immobilised to the BLI biosensor (**Figure 4.16**). TTP mutants were of similar purity, but TTP 4WA appeared to migrate with reduced mobility in SDS-PAGE. All constructs were verified by DNA sequencing, and their encoded proteins were confirmed as having correct molecular weights by denaturing MS. Hence, it was unclear why the quadruple mutation caused this change in electrophoretic mobility.

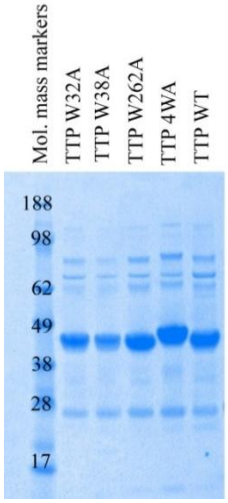


Figure 4.16 TTP mutants used for BLI analysis

SDS-PAGE gel displaying the Strep-Tactin-mediated purification of TTP mutants to be used for BLI analysis. Each lane represents sample eluted from Strep-Tactin using 10 mM biotin buffer.

The binding of biotinylated TTP mutants to CNOT9 and the underlying dissociation constants K_d were determined by BLI (**Table 4.4, Figure 4.18**). All single tryptophan TTP mutants bound CNOT9 with similar association and dissociation traces to wild-type (**Figure 4.17**, compare blue with red, green and turquoise) indicating that mutation of any single tryptophan residue has no observable effect on CNOT9 binding. When all four tryptophans in TTP were mutated, CNOT9 binding was greatly reduced (orange). This reduction in binding of the quadruple mutant (4WA) is reflected in an 8-fold higher K_d than the other proteins (**Table 4.4**).

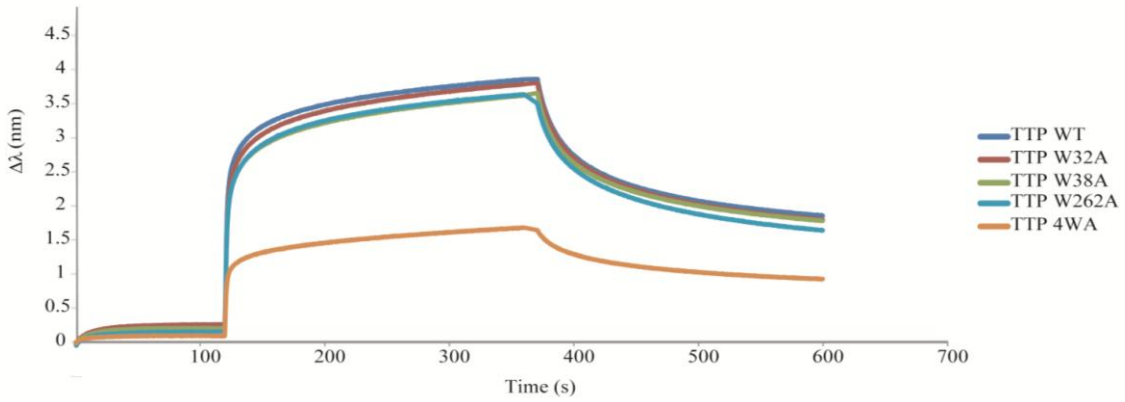


Figure 4.17 BLI association/dissociation trace of TTP mutants

Binding of TTP mutants (WT (blue), W32A (red), W38A (green), W262A (turquoise) and 4WA (orange) was measured using BLI to determine how mutation of tryptophan residues affects ability to bind CNOT9.

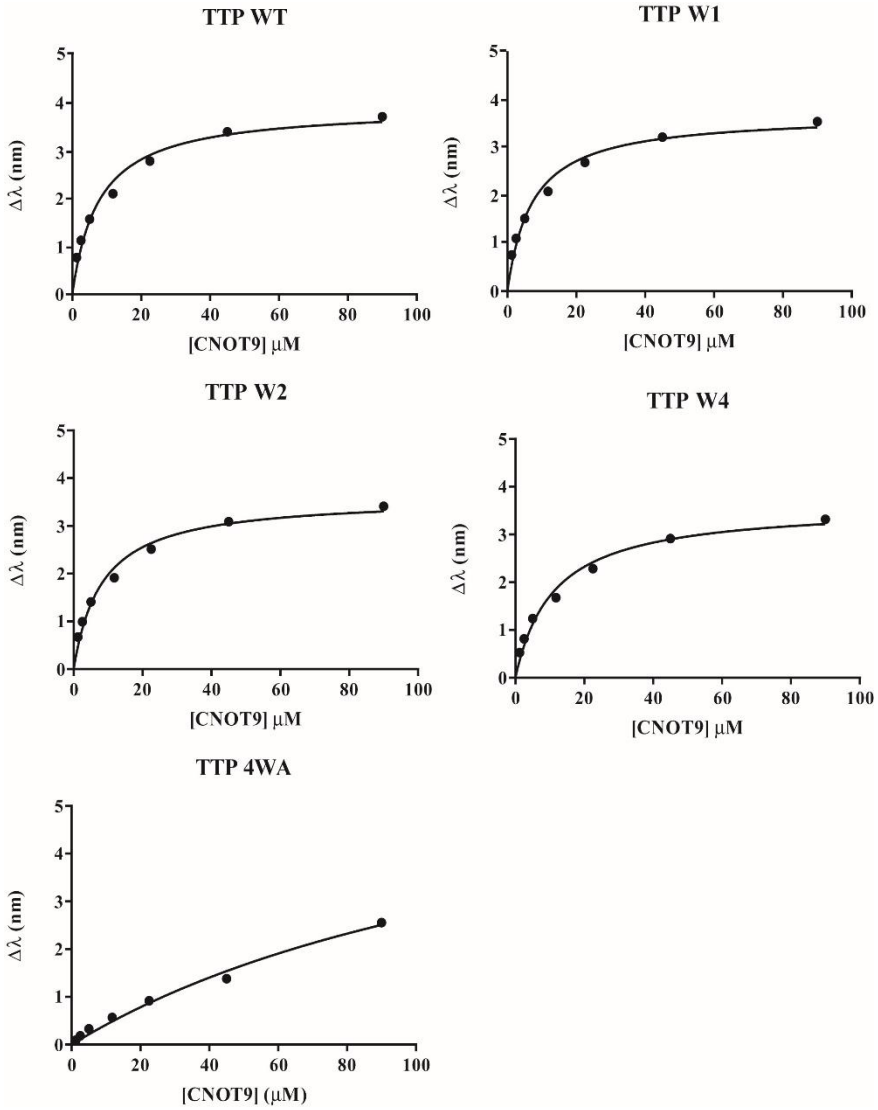


Figure 4.18 Binding curves generated from BLI data for TTP mutant proteins titrated with CNOT9

The change in wavelength (Δλ, nm) is plotted against CNOT9 concentration and points fitted using a one-site total binding model. GraphPad Prism 7 has been used to determine dissociation constants for the TTP mutants. Points are representative of an average of two technical replicates.

Construct	Dissociation constant (K_d) (μM)
TTP WT	10.3
TTP W32A	9.3
TTP W38A	10.9
TTP W262A	11.4
TTP 4WA	81.4

Table 4.4 Dissociation constants of TTP mutants

Dissociation constant of TTP mutants against CNOT9, as measured by BLI.

Together, these mutational studies involving recombinant proteins confirm that TTP interacts with CNOT9 using tryptophan residues in TTP that bind to tryptophan pockets in CNOT9. Published crystal structures indicate that the two CNOT9 tryptophan-binding pockets each bind a single tryptophan residue. Considering that the amino acid sequence of TTP contains four tryptophan residues, all shown to be involved in binding CNOT9 in this study, it remains to be determined how all four tryptophan residues in TTP would interact with two pockets in CNOT9, or if only two tryptophans in TTP are required for binding at any one time.

4.5 Characterisation of the interaction between CNOT2 and TTP using BLI and SPOT array

The siRNA screen (section 1.6, **Appendix Figure 7.11**) that was used to identify CNOT9 as an interactor of TTP, also demonstrated an interaction between TTP and another CCR4-NOT subunit, CNOT2. Thus, a second aim of the study was to validate if indeed CNOT2 forms part of the interface between TTP and the CCR4-NOT complex in vitro.

CNOT2 is 59 kDa protein of 540 residues, and contains a C-terminal NOT box (aa 429-540) that has been shown to interact with NOT-box domains within CNOT1 and CNOT3 (Boland et al., 2013). A range of CNOT2 constructs, including full length protein and N- and C-terminal truncations, were cloned during HTP tests described in chapter 3 (**Figure 4.19**).

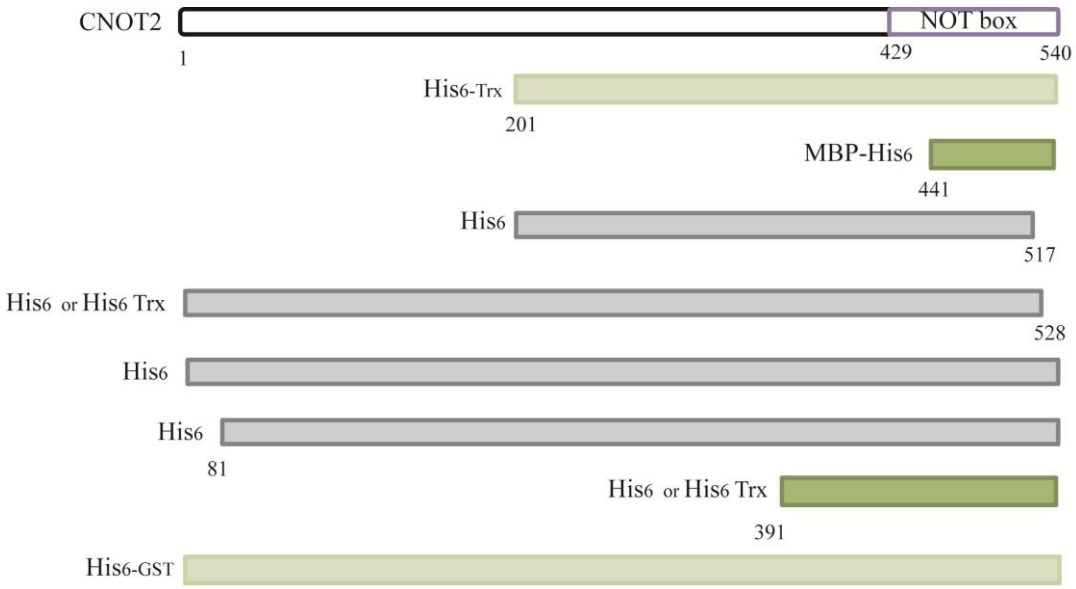


Figure 4.19 Construct diagram of CNOT2

Schematic of a representative set of CNOT2 constructs cloned by LIC. Full length CNOT2 (aa 1-540) (black box) and the NOT box (purple box) are shown. Solid coloured bars indicate cloned constructs of CNOT2 where the numbers indicate the amino acid boundaries. The colour of the bar indicates expression levels determined at small scale; high (dark green), medium (pale green), no expression (grey). A range of tags were included to improve solubility and allow affinity capture: His_6, His-Thioredoxin (Trx), Maltose-Binding protein (MBP)-His_6, and His_6-Glutathione-S-Transferase (GST).

Although a Glutathione S-Transferase (GST) fusion of the full length protein (aa 1-540) can be expressed, it was not expressed to a high level and degraded during purification. However, shorter constructs spanning the NOT box were found to be well expressed and soluble (aa 201-540, aa 391-540 and aa 441-540, **Figure 4.19**). These constructs were used to both demonstrate a direct interaction between TTP and CNOT2 in BLI, and to reveal any CNOT2-binding motifs with the TTP sequence by SPOT array.

A CNOT2 construct (aa 441-540) N-terminally tagged with Maltose Binding Protein, (MBP)-His_6, was used to analyse the interaction with TTP. The MBP fusion protein was included to improve solubility (Kapust and Waugh, 1999), and the His_6 tag allowed nickel affinity capture. MBP-His_6-CNOT2 (aa 441-540) was purified by nickel affinity chromatography from 6 L of *E. coli* culture and gel filtered to remove contaminants. The MBP tag spontaneously cleaved during nickel purification for reasons which are unclear,

leaving His_6-CNOT2 (aa 441-540) alone after gel filtration and a second nickel affinity step. The purified protein (**Figure 4.20)** was used in the subsequent experiments.

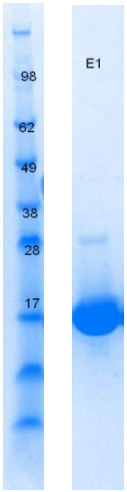


Figure 4.20 Purification of CNOT2 (aa 441-540)

His_6-tagged CNOT2 (aa 441-540) was purified from 6 L of *E. coli* by nickel affinity capture. Eluted samples were gel filtered using an S75 16/60 column and pooled fractions were purified a second time using nickel affinity chromatography to remove the MBP tag which spontaneously cleaved. Fractions eluted from the second nickel affinity purification (E1) were concentrated and used for BLI experiments.

The four biotinylated proteins of TTP harbouring different domain combinations (aa 98-171, aa 14-171, aa 98-326 and aa 14-326, **Figure 4.2**) were used in BLI experiments to determine which domains are able to bind CNOT2(aa 441-540). All TTP proteins were found to bind CNOT2 (aa 441-540) (**Figure 4.21, Table 4.5**). The dissociation constants measured show that the two N-terminal domain-containing proteins (aa 14-171 and aa 14-326) bound with a similarly weak affinity (194.1 and 198.8 µM), and the protein containing C-terminal and zinc-finger domains bound marginally tighter (156.1 µM). The zinc-finger domain-only protein, by contrast, displayed the tightest binding to TTP (62.4 µM), which was still approximately 6-fold weaker when compared to the other TTP-binding subunit, CNOT9. Therefore, as with CNOT9 binding, multiple motifs throughout the sequence of TTP are required to bind CNOT2 (aa 441-540), but unlike for CNOT9, the main contribution to CNOT2 (aa 441-540) binding is provided by the central zinc-finger domain in TTP.

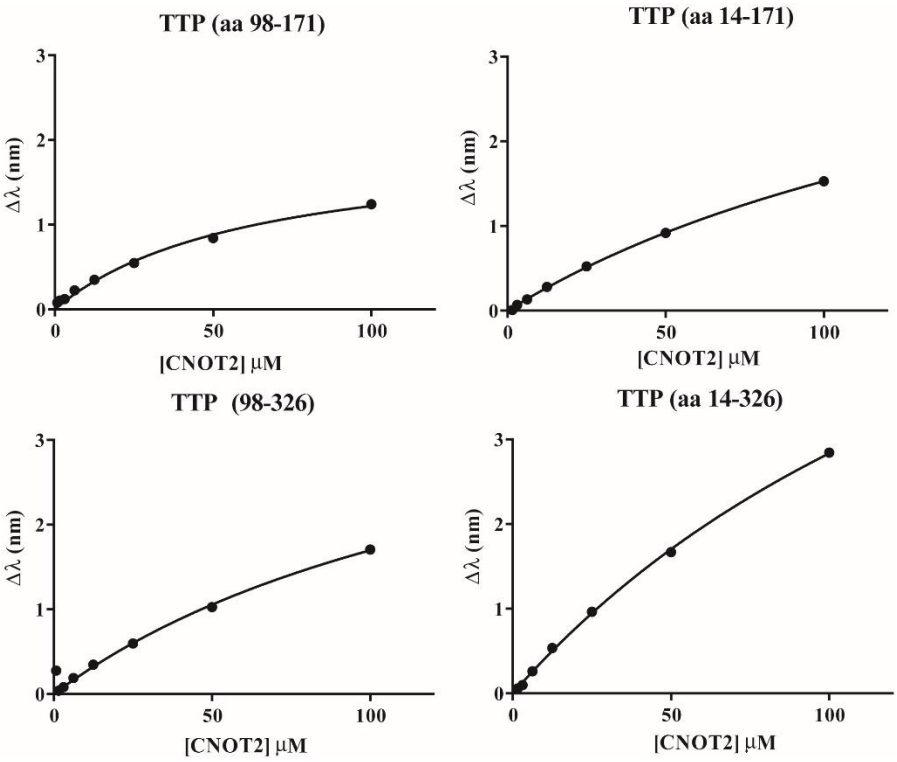


Figure 4.21 Binding of TTP and CNOT2 (aa 441-540) proteins measured by BLI

The change in wavelength (Δλ, nm) is plotted against CNOT2 (aa 441-540) concentration and points fitted using a one-site total binding model. GraphPad Prism 7 has been used to determine dissociation constants for the TTP mutants. Points are representative of an average of two technical replicates.

Construct	Dissociation constant (K_d) (μM)
TTP (aa 98-171)	62.4
TTP (aa 14-171)	194.1
TTP (aa 98-326)	156.1
TTP (aa 14-326)	198.8

Table 4.5 Dissociation constants of TTP domains measured against CNOT2 (441-540)

The dissociation constants of TTP mutants against CNOT2 (aa 441-540), as measured by BLI.

To investigate this hypothesis, a TTP peptide array was once again used for identification of the possible CNOT2-binding motifs within the polypeptide sequence of TTP (**Figure 4.8** and **Appendix Table 7.15**). As dissociation constants measured by BLI (**Figure 4.21, Table 4.5**) were weak in the high micromolar range for CNOT2 (aa 441-540), a longer construct of CNOT2 was used to probe the peptide array to ascertain if including more of the polypeptide of CNOT2 would improve its binding to TTP.

The TTP peptide array was probed with 1 µM of CNOT2 construct (aa 201-540) N-terminally tagged with His_6-Thioredoxin (Trx). The Trx fusion protein was included to improve solubility (Yasukawa et al., 1995), and the His_6 tag allows nickel affinity capture. His_6-Trx-CNOT2 (aa 201-540) was purified by nickel affinity chromatography from 6 L of *E. coli* and gel filtered to remove contaminants. The purified protein (**Figure 4.22)** was used in the subsequent experiments.

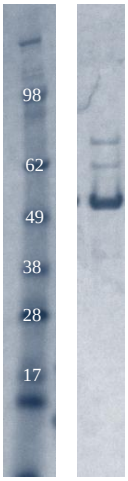


Figure 4.22 Purification of CNOT2 (aa 201-540)

SDS-PAGE analysis of a sample of His_6-Trx-CNOT2 (aa 201-540), taken after nickel affinity capture and gel filtration using an S75 16/60 column.

As shown previously (Section 4.3.1), the intensity of the array spots after immunoblotting was indicative of the amount/level of protein binding to the particular SPOT peptide (**Figure 4.23**). In comparison to the CNOT9-probed TTP peptide array, the CNOT2 (aa 201-540)-probed membrane showed more peaks in intensity (9 peaks for CNOT2 (aa 201-540) vs 5 for CNOT9) which are distributed across the TTP sequence as illustrated in the **Figure 4.24** bar graph.

A larger number of intensity peaks observed for CNOT2 (aa 201-540), compared to CNOT9, would indicate more interaction regions between TTP and CNOT2 (aa 201-540), than between TTP and CNOT9.

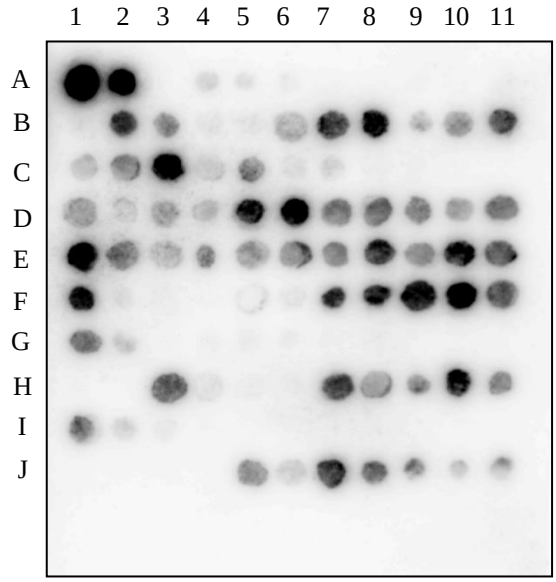


Figure 4.23 TTP peptide array probed with CNOT2 (aa 201-540)

The array was printed with 15 residue TTP peptides, where consecutive peptides overlapped by 3 residues. The position of bound His-CNOT2 (aa 201-540) was detected using an anti-His HRP conjugate. Image captured the intensities recorded when the array was probed with His_6-CNOT2 (aa 201-540). Two His_{10} control peptides were printed in A1 and A3 with a negative control (no peptide) in A3.

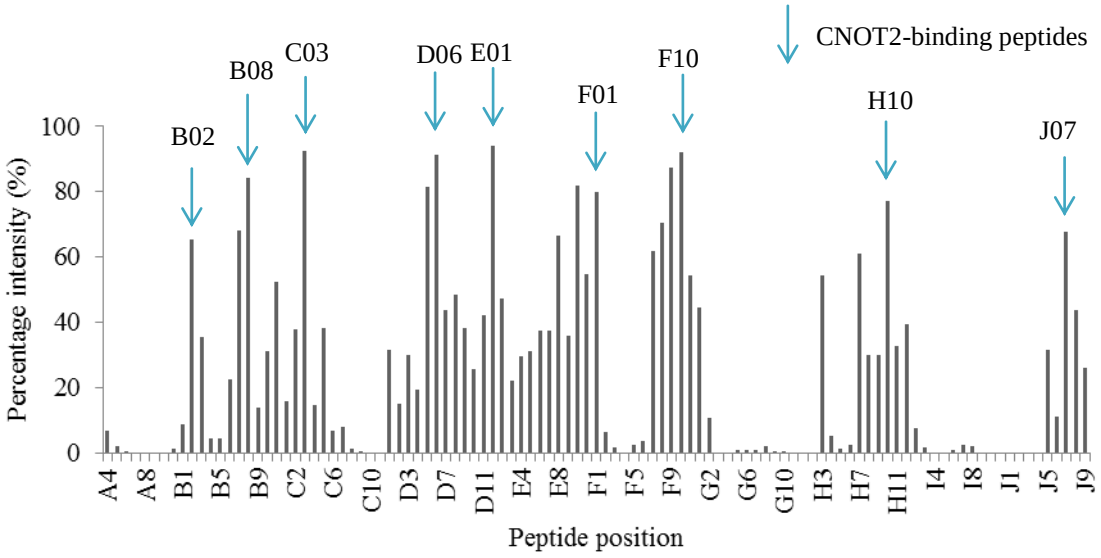


Figure 4.23 TTP peptide array intensities probed by CNOT2 (aa 201-540)

Intensities of each spot from the peptide array in **Figure 4.23** were measured and analysed using Gel Logic 200 Imaging System software. The intensities were calculated as a percentage relative to the control His_{10} peptides and plotted on a bar graph. Blue arrows indicate the peptide at the summit of each peak in intensity which identifies a CNOT2-binding motif.

Figure 4.25 highlights the positions of the peptides found to bind CNOT9 (labelled grey), as demonstrated by BLI, CNOT2 (aa 201-540) (labelled black), as highlighted by the peptide array, and those peptides found to bind both (labelled red). Two peptides (B02, D06) were found to bind both CNOT2 (aa 201-540) and CNOT9, and two regions represented by C03/C04 and H10/I01 peptides also appeared important to bind both proteins. In addition, five motifs unique to CNOT2 (aa 201-540)-binding were identified (B08, D06, E01, E08, F10, J07). It appears that CNOT2 (aa 201-540) interacts with TTP using similar motifs as CNOT9. Moreover, peptide J07 overlaps with the 10-residue

Figure 4.24 TTP sequence annotated with the CNOT2 (aa 201-540)- and CNOT9-binding peptides

The amino acid sequence of human TTP (Uniprot accession: P26651) has been annotated with peptides from **Figure 4.9** shown to bind CNOT9, and with peptides from **Figure 4.23 and Figure 4.24** shown to bind CNOT2. The sequence of each peptide has been underlined and tryptophan residues are coloured blue. Peptide identifiers (e.g. A04) labelled in grey are CNOT9-binding motifs, with black indicating CNOT2-binding motifs and red motifs which can bind both CNOT subunits.

peptide found to interact with CNOT1.

Collating the weak binding affinities established by BLI and the multiple CNOT2-binding motifs highlighted from the TTP peptide array experiment, the data in this study indicate that CNOT2 (aa 201-540) interacts directly with TTP in vitro, albeit weakly, and at many motifs throughout the sequence. Future work is necessary to characterise these

possible motifs further, using BLI/ITC/SPR, to determine their contribution on the interaction between TTP and CNOT2.

4.6 Conclusion

Targeted mRNA deadenylation by the CCR4-NOT complex is a process which requires the interaction of a number of specific RNBPs, one of which is TTP. TTP regulates the half-life of inflammatory mRNAs as a result of upstream signalling by the p38 MAPK pathway. It is currently understood that the C-terminus of TTP is recruited to the CCR4-NOT complex via an interaction with an N-terminal domain of CNOT1 (Fabian et al., 2013). An unpublished siRNA screen in our group identified further interactions between TTP and CCR4-NOT complex subunits, CNOT2 and CNOT9 (**Appendix Figure 7.11**). This formed the basis of this study aimed to characterise the interactions between TTP and CNOT9/CNOT2 in vitro to further map the interface between CCR4-NOT and TTP, in order to better understand the mechanism of regulation of ARE-containing mRNAs.

The CNOT9:TTP interaction was a main focus of the investigation described in this study. Firstly, initial domain analysis highlighted multiple regions throughout the sequence of TTP which are involved in binding CNOT9. Peptide arrays were later used to explore these individual motifs, and tryptophan residues were found to be important for the TTP:CNOT9 interaction, a mechanism common to other CNOT9-binding proteins namely GW182 from the RISC complex. Therefore, this result adds TTP to the growing list of proteins (GW182 and GIGFY2) which interact with CNOT9 in a tryptophan-mediated manner.

A second aim of the study was to characterise the binding of TTP to CNOT2 using similar methods. Dissociation constants were established for the binding of each TTP domain to CNOT2 and nine motifs were identified from the peptide array which may be

responsible for the interaction. Given the multiple, weak affinity binding motifs measured, one explanation may follow that CNOT2 could be important for initially capturing TTP using the whole length of the TTP polypeptide, thus providing a wide net for interactions. Once bound by CNOT2, interactions with TTP domains may be out-competed by stronger interactions with other CNOT subunits i.e. the N-terminal domain of CNOT1 which binds a C-terminal peptide of TTP tightly (~2 μM (Fabian et al., 2013)). Future work should focus on validating these motifs using orthogonal biophysical methods to further investigate the binding of TTP peptides to CNOT2 e.g. SPR/ITC may assist in determining how CNOT2 contributes to the interface between TTP and the CCR4-NOT complex.

To complement the biophysical study of the interaction between TTP and CNOT9, work was performed by a former DPhil student in the group (Bulbrook, 2015), to determine if the quadruple tryptophan mutant (4WA) of TTP affected ARE-mediated mRNA decay in vivo. To this end, TTP (4WA) was transiently transfected into a HeLa cell line with a plasmid expressing a reporter mRNA encoding rabbit β-globin, where the 44 nucleotides of the TNF-α ARE (a target of TTP) were inserted into the 3′UTR. Quantitative reverse-transcriptase PCR (qRT-PCR) was used to determine the levels of the β–globin mRNA substrate in the TTP WT and 4WA conditions as a measure of TTP recruitment of the CCR4-NOT complex. WT TTP targetted the β–globin mRNA substrate for degradation, showing a 69% reduction in mRNA levels, whereas the TTP (4WA) only caused a 43% reduction in β–globin mRNA levels (**Appendix Figure 7.12**) (manuscript in preparation) (Bulbrook, 2015, DPhil thesis, page 150). Thus, the quadruple tryptophan mutations in TTP sufficiently disrupt interactions with the CCR4-NOT complex, and have a direct, functionally relevant effect on its cellular function i.e. mRNA decay.

In summary, this study has identified and characterised a novel interaction between TTP and CNOT9. Future work is required to analyse the contribution and order of interactions to increase our knowledge of TTP-directed mRNA decay, a concept further discussed in chapter 6. Additional work is also warranted to characterise the interaction between TTP and CNOT2, to a similar extent as CNOT9.

5. Construction and characterisation of the CCR4-NOT complex using the MultiBac system

5.1 Introduction

Although a detailed picture of the yeast CCR4-NOT complex is forming, facilitated by two cryo-EM structures, the human complex remains more of a mystery due to the lack of structural information beyond that of some individual subunits. Thus, solving the structure of the human CCR4-NOT complex is an important goal, especially in light of the additional subunits of CNOT10 and CNOT11, which are recruited to the N-terminus of CNOT1. These additional subunits may cause the human complex to adopt a different architecture from the L-shape of the yeast complex, revealing more information about the purpose of the CNOT10:CNOT11 module, and the complex as a whole. A greater understanding of the architecture of the CCR4-NOT complex may also highlight how the CNOT subunits and associated proteins interact to regulate deadenylation, and how this system could be modulated as part of a therapeutic intervention to control overexpression of mRNA in certain diseases.

To study the CCR4-NOT for characterisation of the complex as a whole, it must first be isolated. The complexes studied by cryo-EM were purified endogenously from yeast strains by tagging one CNOT subunit. Although this approach proved successful in providing sample for structural studies, purification of endogenous protein has certain drawbacks in comparison with recombinant expression. Firstly, there are typically lower yields with greater heterogeneity which may reduce the quality and resolution of the data collected. Moreover, purification of endogenous protein does not allow the precise composition of subunits or interacting proteins to be dictated, or the introduction of any

protein engineering (e.g. disease-associated mutations). The latter could be of importance as certain CNOT3 mutations have been associated with human cancers thus, it may be of future interest to determine the effect of these mutations on CNOT3 and how it is recruited to the complex (Zheng et al., 2012). For these reasons, a recombinant method was sought to reconstitute the entire human CNOT complex.

Great success has been found by using the MultiBac expression system (Bieniossek et al., 2008, Stowell et al., 2016), an adaption to traditional baculovirus-mediated expression where viruses known to infect insect cells were adapted for recombinant protein expression. Viral DNA was modified to facilitate the site-specific transposition of a baculovirus-mediated expression plasmid into the viral DNA, which is termed a bacmid. The recombinant bacmid is used to transfect insect cells which generate baculoviruses that are harvested and used to infect insect cells enabling expression of the protein of interest.

The MultiBac system builds on these techniques and facilitates protein complexes to be expressed using a single baculovirus. The system uses a range of vectors, into which the genes of interest are cloned. Each vector contains a LoxP site which can be cleaved and re-joined, using Cre recombinase, with another LoxP site to fuse two vectors. This strategy enables the generation of a vector containing many genes, which can then be transposed and transfected to produce a virus which expresses the protein complex of interest in *Sf9* cells. In this study, the MultiBac system was adopted to create multi-protein viruses for the expression of the CCR4-NOT complex. As a result, a number of sub-complexes of CCR4-NOT were successfully generated with sufficient protein yield for future structural studies. A fluorescence-based deadenylase assay was used to characterise the activity of complexes as a means of ensuring recombinant production

generates active enzymes, but also to explore how activity may be influenced by various parameters (e.g. subunit composition, enzyme concentration).

5.2 Construction of individual MultiBac viruses

5.2.1 Cloning the CCR4-NOT subunits into the MultiBac vectors

Production of MultiBac viruses required a number of cloning steps and PCR screens to ensure that all genes of interest were successfully cloned into a selection of vectors for the production of multigene-expressing viruses. These vectors each contain a different antibiotic resistance gene: pFL (acceptor, ampicillin), pKL (acceptor, kanamycin), pSPL (donor, spectinomycin) and pUCDM (donor, chloramphenicol). Two of the vectors are known as donor vectors and two as acceptor vectors. Multiple donor vectors can be interchangeably utilised in one recombination but only in combination with a single acceptor, as the acceptor contains the transposition sites required to transpose the recombined plasmid into the bacmid. Two highly transcribed late promoters, polH and p10, are present on each MultiBac vector allowing the expression of two genes per plasmid. The multiple cloning site used to introduce genes under polH promotion was modified by colleagues at the SGC to facilitate LIC cloning and a range of recombinant tags were introduced e.g. N-terminal His_6 and C-terminal His_{10} with a twin Strep-Tactin tag.

The human CCR4-NOT complex contains eight stably associated proteins where two mutually exclusive isoforms of each of the two catalytic subunits can be found, meaning that complexes contain either CNOT6 or CNOT6L, and either CNOT7 or CNOT8 (Lau et al., 2009). When deciding which deadenylase isoforms to include in the design of the CCR4-NOT complex for the MultiBac system, CNOT6L and CNOT7 were chosen due to research regarding the nature of the interactions between the catalytic subunits; Lau et

al utilised stable HeLa cell lines, each expressing a different tagged catalytic subunit, to pull out the CCR4-NOT complex. After which, the composition of the subunits in each case was analysed by MS. It was found that greater levels of CNOT6 and CNOT6L were found associated with CNOT7, and of these, greater CNOT6L level were discovered. These data may suggest the interaction between CNOT6L and CNOT7 is more stable (Lau et al., 2009).

The CNOT and TTP genes were amplified by PCR, and cloned by LIC or restriction cloning into the MultiBac vectors. The subunits were cloned into particular pairings in the MultiBac vectors, based on known interacting subunits within the CCR4-NOT complex, i.e. CNOT10 and CNOT11, CNOT1 and CNOT9, CNOT2 and CNOT3, CNOT6L and CNOT7, while TTP was cloned alone (**Table 5.1**). CNOT2, CNOT3, CNOT6L, CNOT7, CNOT10 and CNOT11 were cloned into donor vectors (**Table 5.1**). CNOT1 and CNOT9 were chosen for sub-cloning into the acceptor vector, as a single acceptor is essential in every recombination, and being a scaffold protein CNOT1 will be required in every sub-complex formed during the study.

Fusion tags, provided by the LIC site in each vector, were incorporated to explore how this affects complex reconstitution and heterogeneity of the purified sample. Fusions include either an N-terminal His_6 tag (vectors 1, 11, 12, 14, 16) to allow nickel affinity capture of the complexes or a C-terminal His_{10}StII tag (vectors 3 and 4) to allow both nickel and Strep-Tactin affinity purification steps, which may provide additional purity. Only one subunit in each complex was tagged thus it is expected that the untagged subunits which are captured from nickel purifications have formed a complex with the tagged protein. A description of all cloned donor and acceptor vectors is detailed in **Table 5.1**.

Vector ID	Gene	Start	End	Vector		Site/ promoter	Tag
Vector 1	CNOT1	Met1	Ser2376	pFL-NT6H-LIC	Acceptor	LIC polH	NT His_6
	CNOT9	Met1	Gln299			RE (XhoI/ NcoI) p10	No tag
Vector 2	CNOT1	Met1	Ser2376	pFL-CT10HF-LIC	Acceptor	LIC polH	CT His_{10}flag
	CNOT9	Met1	Gln299			RE (XhoI/ NcoI) p10	No tag
Vector 3	CNOT1	Met1	Ser2376	pFL-CT10HStII-LIC	Acceptor	LIC polH	CT His_{10}StII
	CNOT9	Met1	Gln299			RE (XhoI/ NcoI) p10	No tag
Vector 4	CNOT1	Met1	Ser2376	pFL-CT10HStII-LIC	Acceptor	RE (XhoI/ PmeI) p10	No tag
	CNOT9	Met1	Gln299			LIC polH	CT His_{10}StII
Vector 5	CNOT2	Met1	Phe540	pKL-CTF-LIC	Donor	RE (XmaI/NheI) p10	No tag
	CNOT3	Met1	Gln753			LIC polH	No tag
Vector 6	CNOT6L	Met1	Arg555	pUCDM-CTF-LIC	Donor	RE (XhoI/NheI) p10	No tag
	CNOT7	Met1	Ser246			LIC polH	No tag
Vector 7	CNOT10	Met1	Lys717	pSPL-CTF-LIC	Donor	RE (NcoI/NheI) p10	No tag
	CNOT11	Met1	Lys510			LIC polH	No tag
Vector 8	TAB182	Met1	Val1729	pKL-CTF-LIC	Donor	LIC polH	No tag
						Empty	
Vector 9	CNOT2	Met1	Phe540	pSPL-CTF-LIC	Donor	RE (XmaI/NheI) p10	No tag
	CNOT3	Met1	Gln753			LIC polH	No tag
Vector 10	TTP	Met1	Glu326	pUCDM-CTF-LIC	Donor	LIC polH	No tag
						Empty	
Vector 11	TTP	Met1	Glu326	pFL-NT6H-LIC	Donor	LIC polH	NT His_6
						Empty	
Vector 12	CNOT2	Ser391	Phe540	pFL-NT6H-LIC	Acceptor	LIC polH	NT His_6
						Empty	
Vector 13	CNOT3	Met1	Gln753	pUCDM-CTF-LIC	Donor	LIC polH	No tag
						Empty	
Vector 14	CNOT1	Leu800	Ser2376	pFL-NT6H-LIC	Acceptor	LIC polH	NT His_6
						Empty	
Vector 15	CNOT9	Met1	Gln299	pUCDM-CTF-LIC	Donor	LIC polH	No tag
						Empty	
Vector 16	CNOT1	Glu1090	Ser2376	pFL-NT6H-LIC	Acceptor	LIC polH	NT His_6
						Empty	

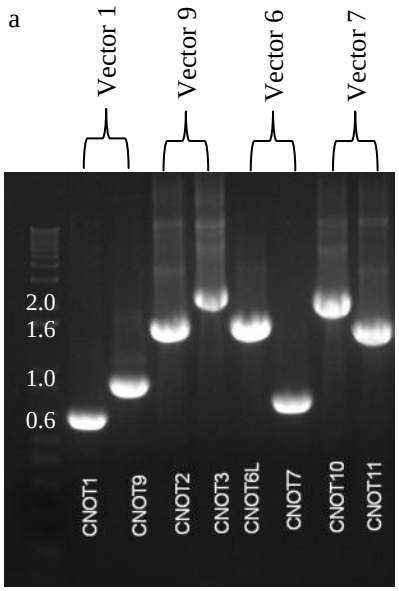
Table 5.1 Library of donor and acceptor constructs made for the CCR4-NOT MultiBac system

The table provides the details of 16 MultiBac constructs generated for the reconstitution of the CCR4-NOT complex. Each vector was sub-cloned to contain two CNOT genes, one in the LIC site and one in the restriction enzyme (RE) site. The exceptions are TAB182 (vector 8) and TTP (vectors 11 and 12) which were cloned as single components into the LIC site, leaving the RE site empty. The start and end amino acid residues have been listed for each construct. CT indicates the C-terminal fusion tag and NT indicates an N-terminal fusion tag. The brackets indicate the restriction enzymes used to insert the gene into the multiple cloning site under p10 promotion.

Full length (FL) open reading frames of each protein were used in these vectors to ensure all possible interacting regions between subunits were covered. A number of truncations at the N- and C-termini were also screened for CNOT2 and CNOT1. CNOT1 is a 251 kDa protein and it was unclear if a protein of this size would be successfully expressed and soluble in insect cells. Thus, full length CNOT1 was trialled but also CNOT1 was truncated into domains shown to interact with the catalytic subunits (aa 1090-1842, vector 16), and both the catalytic subunits and the NOT module (aa 800-2373, vector 14). It was noted that CNOT2 constructs encompassing the C-terminal NOT box domain were the most stable and highly expressed, and constructs which expanded towards the N-terminus were degraded easily. Hence, both full length CNOT2 (vectors 5 and 9) and a NOT box truncation (vector 12) were trialled (**Table 5.1**).

The cloned MultiBac constructs described above were transformed into MACH1 *E. coli* and screened by PCR using gene specific primers. For CNOT2-11, primers which amplify the entire gene were used. However, for CNOT1, primers that only amplify the N-terminal region were used, as amplification of the full length CNOT1 gene (~7 kb) proved difficult to reliably reproduce. PCR products were analysed by agarose gel electrophoresis (**Figure 5.1a**), which indicated that the lengths of the PCR products matched the expected size for all CNOT subunits (**Figure 5.1b**). This was corroborated by DNA sequencing of each plasmid, which confirmed that the CNOT genes had been successfully sub-cloned with the expected nucleotide sequences. Thus, a library of 16 CNOT-containing MultiBac vectors (9 donors, 7 acceptors) was generated.

a



b

Subunit	Size (kb)
CNOT1	0.6*
CNOT9	0.9
CNOT2	1.6
CNOT3	2.3
CNOT6L	1.7
CNOT7	0.7
CNOT10	2.2
CNOT11	1.5

Figure 5.1 PCR screening of individual subunits

(a) Two PCR screens were used per vector to identify the presence of both genes in each plasmid. The size of the PCR product, as determined by agarose gel electrophoresis, was used to confirm if the correct clone had been made. CNOT1 was amplified with primers which anneal to an N-terminal region only, indicated by the asterisk, and produce a 0.6 kb fragment. The other CNOT subunits were screened with primers which amplify the entire gene. (b) Expected sizes for each subunit are listed in the accompanying table.

5.2.2 Recombination, screening and transfection of MultiBac complexes

Once the individual plasmids in **Table 5.1** were successfully cloned, they were recombined to form one multi-gene vector using the Cre-LoxP system (Abremski et al., 1986). Each MultiBac vector contains a LoxP site, comprised of two inverted repeats of ATAACTTCGTATA separated by a spacer sequence (**Figure 5.2a**). During recombination, Cre recombinase binds to each inverted repeat and dimerises. The Cre makes a single strand break in the spacer region of the LoxP site, after which a cross-over event occurs and opposing strands are re-ligated, creating the multigene vector (Van Duyne, 2001) (**Figure 5.2b**). To ensure the selection of a multi-gene vector which contains all intended donors and acceptor, the products of recombination are transformed into MACH1 *E. coli* and selected for based upon the specific antibiotic resistance genes encoded by each vector. For example, a recombination of pFL, pUCDM and pSPL would be selected for resistance towards ampicillin, chloramphenicol and spectinomycin.

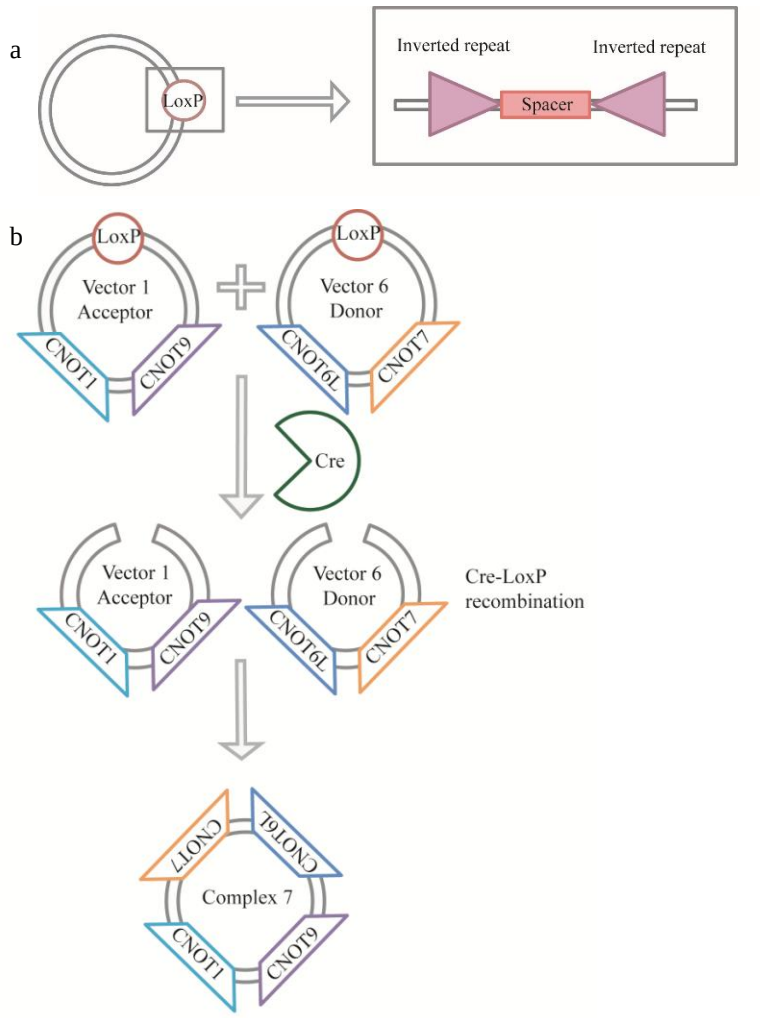


Figure 5.2 The Cre-LoxP recombination process

(a) Each individual MultiBac vector has a LoxP site which is comprised of two inverted repeats (triangle) separated by a spacer sequence. (b) Schematic shows the recombination of an acceptor (vector 1) and a donor (vector 6), each containing two CNOT genes. Cre recombinase (green) binds to the inverted repeat and dimerises. Cre recombinase makes a single strand break in the spacer region and a cross over event occurs. The strands are repaired which produces a multi-gene vector (complex 7).

In total, 26 multi-gene vectors for the expression of CCR4-NOT complexes were created by Cre-LoxP recombination (**Table 5.2**). Complexes range from containing two to eight subunits, which encompass the different genes of the CCR4-NOT complex. For complexes comprising more than two donors (complexes 12 and 13), two steps of recombination were required to incorporate all genes. The range of sub-complexes were designed to allow the study of different organisational aspects of the CCR4-NOT complex e.g. complexes containing CNOT1:CNOT9:CNOT10:CNOT11 (complex 8)

Complex ID	Acceptor	Donor 1	Donor 2	Donor 3
Complex 1	Vector 1	Vector 7		
Complex 2	Vector 1	Vector 5		
Complex 3	Vector 1	Vector 11		
Complex 4	Vector 16	Vector 11		
Complex 5	Vector 1	Vector 11		
Complex 6	Vector 1	Vector 5		
Complex 7	Vector 1	Vector 6		
Complex 8	Vector 1	Vector 7		
Complex 9	Vector 1	Vector 5	Vector 6	
Complex 10	Vector 1	Vector 6	Vector 7	
Complex 11	Vector 1	Vector 5	Vector 7	
Complex 12	Vector 1	Vector 7	Vector 5	Vector 6
Complex 13	Vector 1	Vector 5	Vector 6	Vector 11
Complex 14	Vector 1	Vector 5	Vector 11	
Complex 15	Vector 1	Vector 5	Vector 11	
Complex 16	Vector 1	Vector 6	Vector 11	
Complex 17	Vector 16	Vector 6	Vector 11	
Complex 18	Vector 16	Vector 6		
Complex 19	Vector 11 (S60A, S323A)			
Complex 20	Vector 11 (S60A, S186A)			
Complex 21	Vector 11 (S186A, S323A)			
Complex 22	Vector 11 (S323A)			
Complex 23	Vector 11 (WT)			
Complex 24	Vector 12	Vector 13		
Complex 25	Vector 16	Vector 15		
Complex 26	Vector 14	Vector 15		

Table 5.2 Subunits within MultiBac viruses used for expression testing

The table provides details of which vectors from **Table 5.1** were recombined to form each of the 26 MultiBac complexes. A set of 4 TTP phosphorylation mutants were generated (complexes 19, 20, 21, 22) where serine residues in TTP known to be phosphorylated by MK2 were mutated to alanine.

may give insight into the role of the CNOT10:CNOT11 and structural studies may indicate how CNOT10:CNOT11 bind CNOT1.

Furthermore, the full complex has been recombined with and without CNOT10:CNOT11 (**Table 5.2** complex 9 and 12) so the effect of these poorly characterised subunits on deadenylase activity can be studied. Another recombined complex included CNOT1, the scaffold subunit, CNOT9, the mediator of protein-protein interactions, and the deadenylase subunits CNOT6L and CNOT7 (**Table 5.2,** complex 7).

DNA prepared from colonies selected for the corresponding antibiotic resistance were screened by PCR using gene-specific primers. Again, the presence of each gene was confirmed by a band of the correct size (**Figure 5.3 and Figure 5.2b**). **Figure 5.3** depicts the screening of two complexes. Genes encoding CNOT1, CNOT9, CNOT2, CNOT3, CNOT10 and CNOT11 were confirmed in complex 11, and genes comprising the entire complex were confirmed in complex 12 (**Table 5.1** and **Table 5.2**).

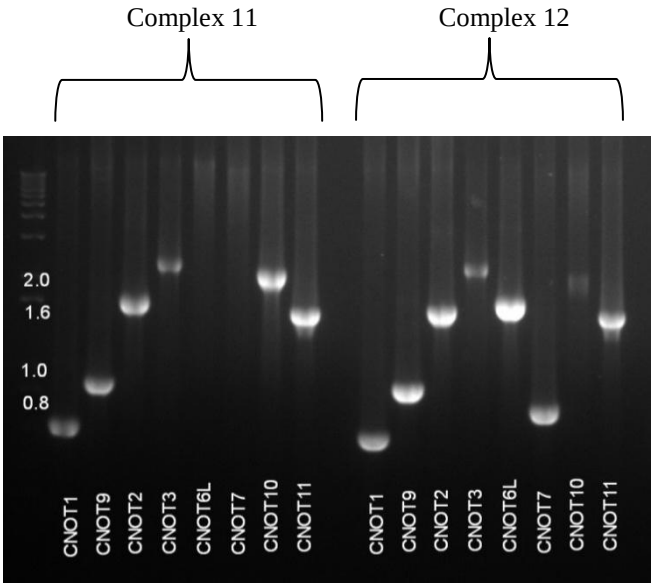


Figure 5.3 Screening of recombined plasmids

Recombined plasmids were screened for the presence of all the genes from the individual vectors. PCR screening used gene specific primers and the products were visualised on a 1.5% agarose gel to check that the correct size for each gene has been produced. The gel shows two recombined plasmids, namely complex 11, which contains CNOT1:CNOT9:CNOT2:CNOT3:CNOT10:CNOT11, and complex 12 which contains CNOT1:CNOT9:CNOT2:CNOT3:CNOT10:CNOT11:CNOT6L:CNOT7.

The 26 confirmed constructs were transformed into DH10Bac *E. coli* to allow transposition of the multi-gene plasmid into the bacmid DNA. This strain contains bacmid DNA which has a *lacZ* gene situated between two transposons and a separate plasmid encoding a transposase enzyme. The recombined vector also contains transposable Tn7 sites from the acceptor. Together, these protein and DNA elements

allow transposition which disrupts the *lacZ* gene within the transposition sites of the bacmid and prevents the gene product, β-galactosidase enzyme from converting X-gal, to a blue product. Thus, colonies where the expression plasmid has successfully transposed appear white.

The 26 CNOT-containing bacmids were purified from the white colonies and used to transfect 1 ml of *Sf9* cells. After 6 days of incubation at 27 ºC, the medium containing virus, termed P0, was harvested. P0 virus was amplified in subsequent passages, termed P1, P2 etc., by infecting further *Sf9* cells to produce a high titre (concentration of virus particles) stock. These passages were used to determine expression levels of the complexes generated, by small scale expression and purification tests.

5.3 Expression characterisation of CCR4-NOT MultiBac complexes

5.3.1 Small scale expression/purification tests

The 26 CNOT-containing baculoviruses isolated from blue/white screening were tested in small scale to ascertain the expression levels of each complex. 3 ml of *Sf9* cells were infected with 120 µL of P0 virus for each of the 26 complexes, and incubated for 66-72 h before harvesting. The medium was collected for viral amplification (P1, P2 etc.) and complexes were purified using nickel affinity chromatography, as CNOT1 was fused with a His_6 tag. Only one subunit in each complex was tagged so it was expected that co-purified proteins have formed a complex with CNOT1, but further analysis of complex formation (e.g. size exclusion chromatography) was explored when complexes were purified at larger scale. SDS-PAGE was used to visualise protein which eluted during small scale expression/purification tests of the first 23 CCR4-NOT sub-complexes from **Table 5.2** (**Figure 5.4**). A control protein known to express well, Activin A receptor type 1 (ACVR1) was included to ensure the transfection and purification protocols were

effective. Full length CNOT1 was expressed to a medium-high level (complexes 1, 2, 3, 6 and 16), with the expectation that the entire CNOT1 polypeptide would cover all the necessary regions to interact with different CNOT subunits and accessory proteins, thus allowing all possible sub- and full complexes to be generated. Truncations of CNOT1 (aa 1090-1842, complexes 4, 19 and 26) were similarly well expressed. The complex containing CNOT1 (aa 800-2371):CNOT9 (complex 26) also shows soluble expression of CNOT1 but at a lower level, and accordingly, less CNOT9 has also co-purified. It is possible that these truncation boundaries produced a less stable protein, leading to lower expression levels. All subunits in **Figure 5.4** were consistent with the apparent molecular weight of migrating bands, and were confirmed by tandem MS and intact MS.

Beyond CNOT1 and CNOT9, expression was also detected for CNOT6L, and its binding partner, CNOT7 (complexes 7 and 9). Successful expression/purification of these subunits allowed the deadenylase activity of complexes to be characterised. However, some subunits were not detected e.g. CNOT2 and CNOT3. Although the CNOT2 and CNOT3 genes were confirmed by PCR screening during virus generation, their protein expression were not detected by tandem MS, intact MS or SDS-PAGE (complex 2, 6 and 9-15). Thus, it cannot be confirmed that these proteins are expressed and soluble.

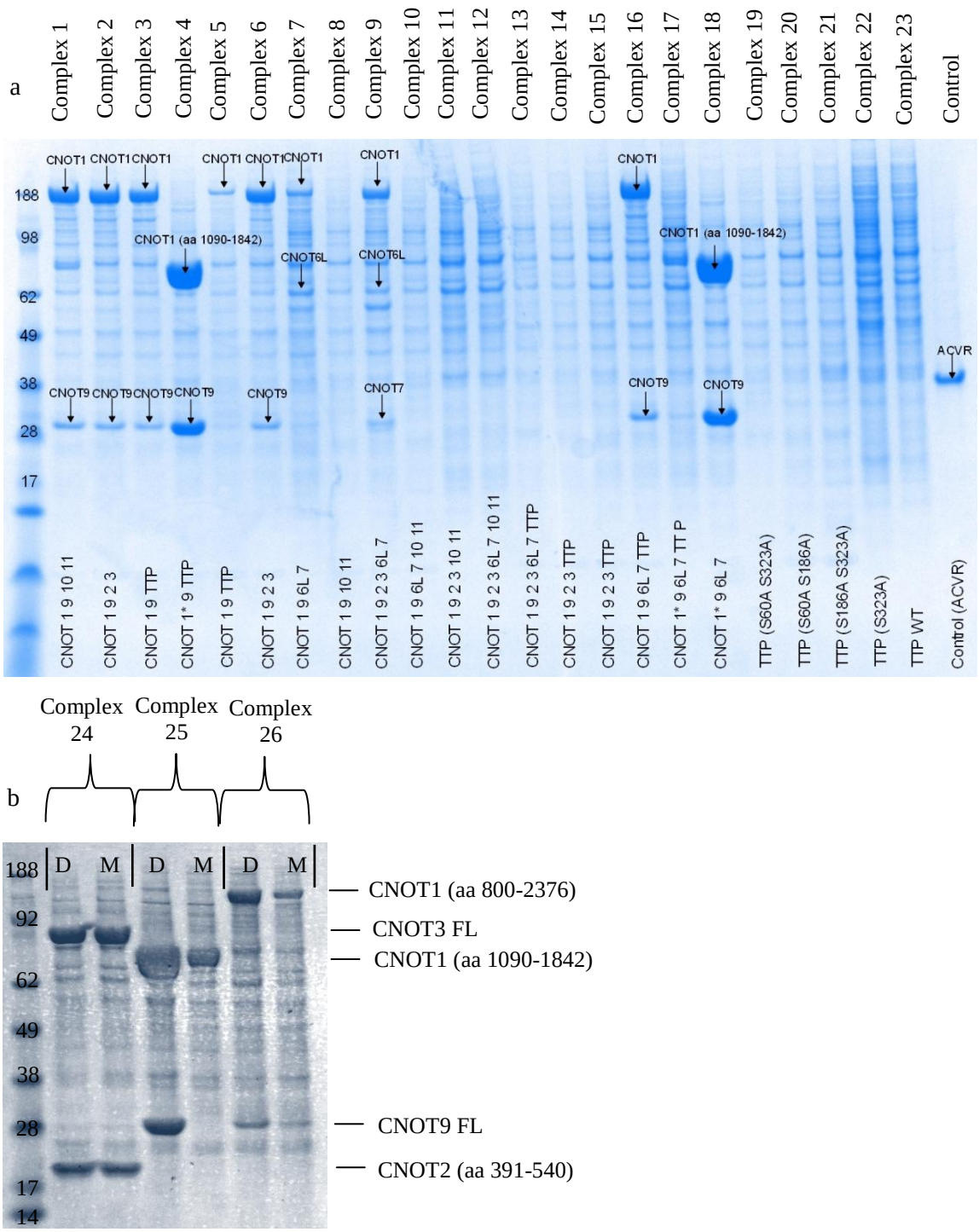


Figure 5.4 Expression/ purification test of CNOT protein complexes

(a) SDS-PAGE of elutions from expression/purification tests of complexes 1-23. The labels along the bottom indicate the subunits included in the different recombinations. CNOT1 was the only tagged component in each expression and was fused with an N-terminal His_6 tag. (b) SDS-PAGE of elutions from expression/purification tests of complexes 24-26. Two different strains of *E. coli* DH10Bac used for transposition were compared (D=DH10Bac, M=DH10MBac). The positions of CNOT subunits, as determined by tandem MS, are labelled. Activin A receptor type 1 (ACVR1) was included as complex 24 as a control as it was shown to express previously.

However, expression of CNOT2 (aa 391-540) and CNOT3 in complex 24 were detected. Therefore, it is unclear why CNOT2 and CNOT3 were not observed in complexes 1-23. Similarly, expression of CNOT10 and CNOT11 was not detected by SDS-PAGE, tandem MS or denaturing MS despite previous PCR confirmation. A co-infection strategy has been demonstrated in chapter 3 to produce a sub-complex of full length CNOT10 and CNOT11, thus it is known at this stage of the project that these proteins are expressed and soluble in insect cells. Therefore, it remains unclear why expression of CNOT10 and CNOT11 were not detected for the larger complexes in **Figure 5.4a**. However, expression of CNOT2 (aa 391-540) and CNOT3 in complex 24 and of CNOT10 and CNOT11 in chapter 3 provide the opportunity for co-infection to reconstitute all CCR4-NOT subunits. As such, a co-infection of CCR4-NOT MultiBac viruses to express CNOT1, CNOT9, CNOT6L, CNOT7, CNOT2 (aa 391-540) and CNOT3 is described in section 5.4.1.

When comparing the expression profiles of viruses, many more protein bands were observed in lanes which contained CCR4-NOT complexes compared to the control (ACVR **Figure 5.4a**). The majority of these bands were degradation products of CNOT1, as confirmed by tandem MS. As CNOT1 is the only His_6-tagged component, protein which may have degraded from the C-terminus, perhaps due to the lack of CNOT2:CNOT3 which bind to this region, would still contain the N-terminal His_6 tag thus could be captured during nickel affinity chromatography. For larger scale purification, a gel filtration step was implemented to attempt to separate the CNOT1 degradation products from the main complex.

Expression tests and associated MS confirm the expression of 12 CCR4-NOT sub-complexes that have been successfully generated using the MultiBac system. Now that

the composition of complexes has been confirmed, the expression can be optimised, purifications scaled-up and the deadenylase activity characterised.

5.3.2 Optimisation of conditions of baculovirus-mediated expression

To ensure optimal expression of recombinant CNOT complexes from the baculovirus-infected cultures, a time course experiment was initiated using different viral volumes. Viruses expressing CNOT2 (aa 391-540):CNOT3 (**Table 5.2**, complex 24) and CNOT1 (aa 1090-1842):CNOT9 (complex 25) were used due to the consistent expression of subunits which allowed visual inspection by SDS-PAGE of any changes in expression caused by the optimisation. The amount of virus used for infection did not affect expression, and thus the minimal volume tested (30 µl per 50 ml *Sf9* cells) will be used for future infections to conserve viral stocks (**Figure 5.5**). Moreover, protein levels in the 48 h and 72 h timeframe appeared consistent, suggesting protein expression has reached an optimal level after 48 h. In subsequent expressions, infections were harvested between 48-72 h which reduces the duration of viral incubation, thus may decrease the likelihood of cell lysis, the final phase of viral infection.

This simple optimisation experiment ensured that the amount of virus and the length of infection trialled in this study produced the maximum protein quantity. However, future studies may also benefit from calculation of viral titres to allow infection of *Sf9* cells with the same amount of virus each time, ensuring reproducible protein quality from each infection.

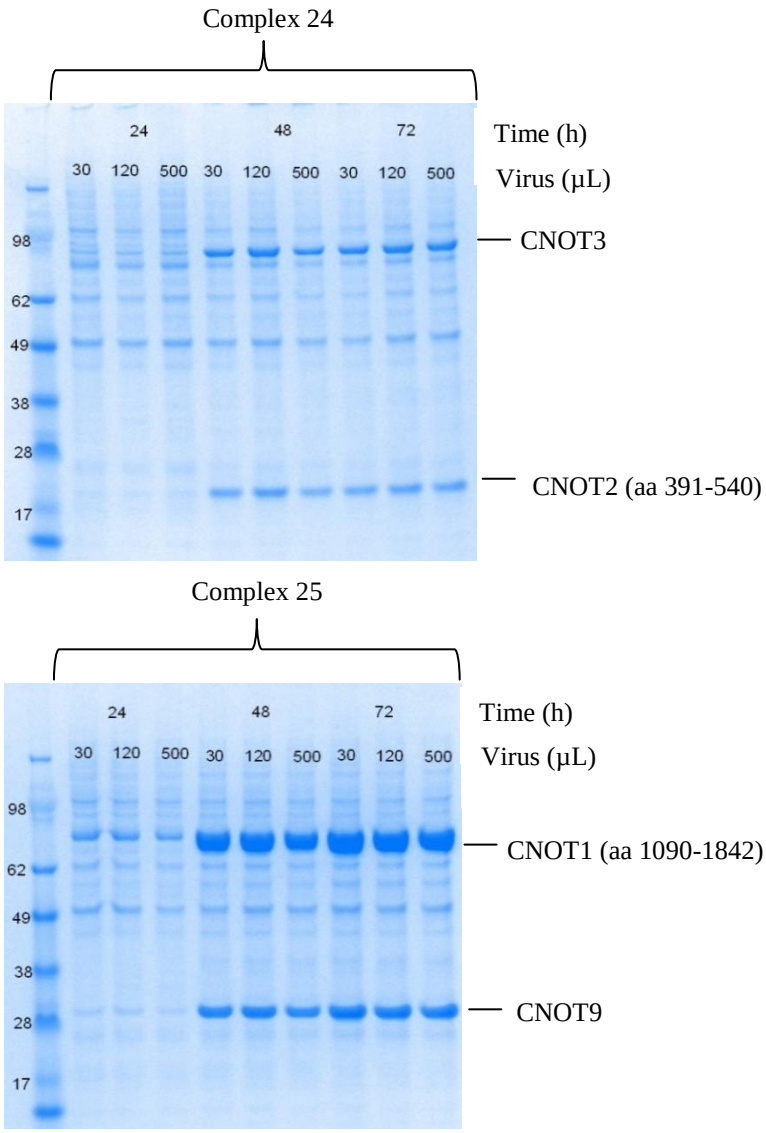


Figure 5.5 Optimisation of conditions of baculovirus-mediated expression

3 ml of *Sf9* cells were infected with either 30, 120 or 500 µl virus expressing either CNOT2 (aa 391-540) and CNOT3 (complex 24) or CNOT1 (aa 1090-1842) and CNOT9 (complex 25). Infections were incubated for either 24, 48 or 72 h. The positions of CNOT subunits, as determined by tandem MS, are labelled. All samples were purified in parallel using nickel affinity chromatography.

5.4 Large scale purifications of CNOT complexes

5.4.1 Purification of CNOT1:CNOT9:CNOT6L:CNOT7(:CNOT2:CNOT3)

The small scale expression and purification tests in **Figure 5.4** highlighted a number of sub-complexes of CCR4-NOT subunits, one of which (complex 9) was confirmed during PCR screening to contain CNOT1, CNOT9, CNOT2, CNOT3, CNOT6L and CNOT7 but expression could only be detected by SDS-PAGE or MS for CNOT1, CNOT9, CNOT6L and CNOT7. Therefore, it cannot be confirmed that CNOT2 and CNOT3 are present in the baculovirus, hence they are thereafter indicated in parentheses. Nevertheless, both

deadenylase subunits were successfully confirmed to be expressed. Therefore, large scale production of this complex was carried out to enable the characterisation of the deadenylase subunits.

Virus expressing CNOT1:CNOT9:CNOT6L:CNOT7(:CNOT2:CNOT3) was used to infect 1 L of *Sf9* for 48-66 hours, and the resulting protein was purified using nickel affinity chromatography (**Figure 5.6a**). CNOT protein complexes were eluted in the first wash (E1) with elution buffer (300 mM imidazole) but very faint bands were present. The sample was concentrated and again visualised by SDS-PAGE (**Figure 5.6b**). As previously, CNOT2 and CNOT3 were not detected by SDS-PAGE or MS, but CNOT1, CNOT9, CNOT6 and CNOT7 were detected by intact MS. Approximately 1.2 mg of CNOT complex was isolated from 1 L of *Sf9* after one step of affinity purification. Thus,

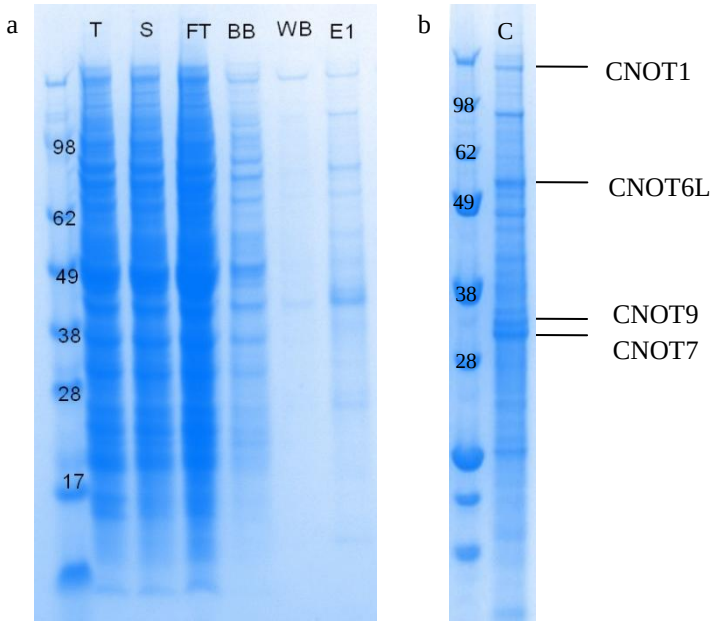


Figure 5.6 Nickel affinity purification of CNOT1:CNOT9:CNOT6L:CNOT7:(CNOT2:CNOT3) (complex 9)

(a) Protein produced from infection of *Sf9* cells using a virus expressing CNOT1:CNOT9:CNOT6L:CNOT7:(CNOT2:CNOT3) (complex 9) was captured using nickel affinity chromatography. Samples for the total cell expression (T), soluble protein fraction (S), proteins which flowed through the nickel beads (FT), proteins eluted with binding buffer (BB), proteins eluted with wash buffer (WB) and proteins eluted with elution buffer (E1) were run on a SDS-PAGE gel. (b) E1 was concentrated (C) and visualised by SDS-PAGE. Labels are based on tandem MS result from small scale expression test (**Figure 5.4a,** complex 9).

a sufficient yield of CNOT1:CNOT9:CNOT6L:CNOT7(:CNOT2:CNOT3) was achieved which provided enough material to characterise the deadenylase activity. For future structural experiments, the yield could be increased by infecting a larger volume of *Sf9* cells and optimising viral titres.

5.4.2 Coinfection of two viruses, CNOT1:CNOT9:CNOT6L:CNOT7:(CNOT2:CNOT3) and CNOT2 (aa 391-540):CNOT3

As the presence of CNOT2 and CNOT3 in complex 9 could not be confirmed by small scale (**Figure 5.4**) or large scale expressions (**Figure 5.8**), attempts were made to co-infect two CNOT-expressing MultiBac viruses in an effort to reconstitute a larger sub-complex. Consequently, 2 L of *Sf9* cells were co-infected with 2 ml of virus expressing CNOT1:CNOT9:CNOT6L:CNOT7:(CNOT2:CNOT3) (complex 9) and 2 ml of virus containing CNOT2 (aa 391-540):CNOT3 (complex 24). As cells were co-infected with two separate complexes, two elements of the final reconstituted complex were His_6-tagged: CNOT1 in complex 9 and CNOT2 in complex 24. The protein complex was purified using nickel affinity chromatography and elutions were analysed by SDS-PAGE (**Figure 5.7**). Many protein bands can be observed in the eluted fraction (lane EB1). The bands are consistent with the expected masses of the subunits and tandem MS was used to identify CNOT1, CNOT2, CNOT3, CNOT6L and CNOT7. However, CNOT9 was not identified despite its confirmation during PCR screening. It is possible that CNOT9 is present but the band was too weak to be distinguishable by SDS-PAGE.

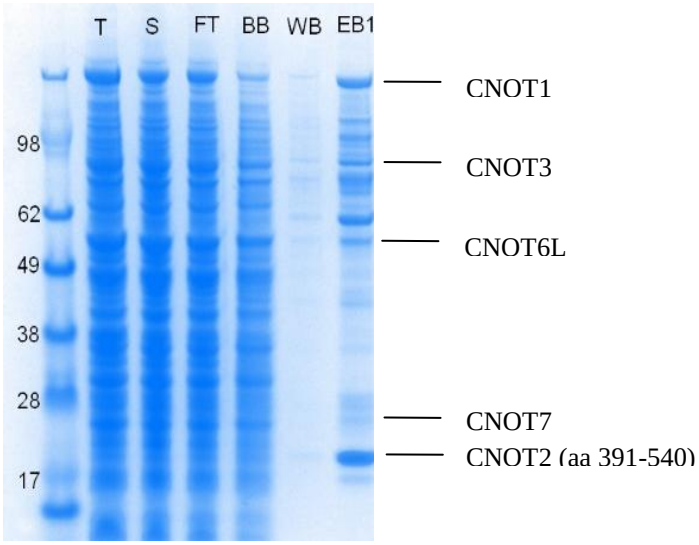


Figure 5.7 Nickel affinity purification of a co-infection of CNOT1:CNOT9:CNOT6L:CNOT7:(CNOT2:CNOT3) (complex 9) with truncated CNOT2 (aa 391-540):CNOT3 (complex 24)

The protein complex was captured using nickel affinity chromatography. Samples were run on a SDS-PAGE gel. Labels are based on tandem MS results.

Although the CNOT proteins are captured together during affinity purification, this does not necessarily indicate that a complex has formed for all the expressed subunits, as both CNOT1 and CNOT2 (aa 391-540) are His_6-tagged. Only the complex between CNOT2 (aa 391-540) and CNOT3, and the complex between CNOT1, CNOT9, CNOT6L and CNOT7 can be assumed to have formed. To achieve greater purity and determine if all subunits co-purify as one complex, the eluted sample was gel filtered using a Superdex S6 16/60 column (**Figure 5.8**).

Multiple peaks are shown on the chromatogram, which is expected when purifying a large complex because of excess un-complexed samples. Peak 1 likely represents nucleic acids as no protein is observed in this region by SDS-PAGE (**Figure 5.8**). Peak 2 contains two sub-species by SDS-PAGE (**Figure 5.9**), where the first two fractions show elution of a complex containing CNOT1, CNOT9, CNOT6L, CNOT3 and CNOT2 (aa 391-540), as identified by tandem MS, whereas the second two fractions additionally contain CNOT7. This finding is odd given that CNOT6L associates with the complex via an interaction with CNOT7, thus would be unlikely to co-migrate with the complex in the

absence of CNOT7. It is probable that CNOT7 is present within the first two fractions of peak 2, but at a too low concentration to be observed visually by SDS-PAGE. Peak 3 contains many more degradation products of CNOT1, as confirmed by tandem MS. Peak 4 contains mostly excess CNOT2. This is because CNOT2 was a His_6-tagged component, thus, an excess of un-complexed CNOT2 protein would be captured during nickel affinity chromatography. However, due to its small size, it has been successfully separated by gel filtration from the CCR4-NOT complex in peak 2. **Figure 5.10** shows four lanes from peak 2 in **Figure 5.9** in isolation to more clearly indicate the position of migrating subunits and show how the purity of the complex has greatly been improved by gel filtration. It is clear that the two co-infected sub-complexes are indeed co-migrating during gel filtration which indicates that the complex of CNOT1:CNOT9:CNOT6L:CNOT7:CNOT2 (aa 391-540):CNOT3 has formed. Additionally, at this stage of purification, CNOT9 was successfully detected by tandem MS which suggests that previous analyses may have missed its detection.

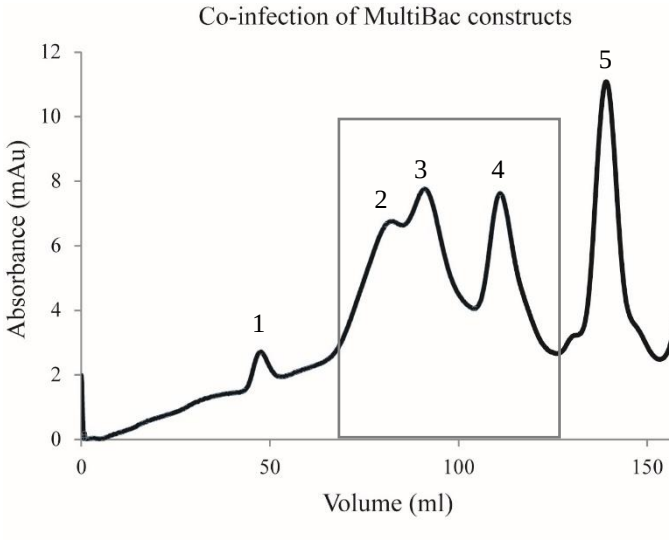


Figure 5.9 Gel filtration of protein complex purified from co-infection of CNOT1:CNOT9: CNOT6L:CNOT7: (CNOT2:CNOT3) with truncated CNOT2 (aa 391-540):CNOT3

EB1 fraction from affinity purification (**Figure 5.7**) was run on an S6 16/60 gel filtration column, producing a chromatogram with five peaks (labelled 1-5). The box highlights fractions which were analysed by SDS-PAGE in **Figure 5.9**.

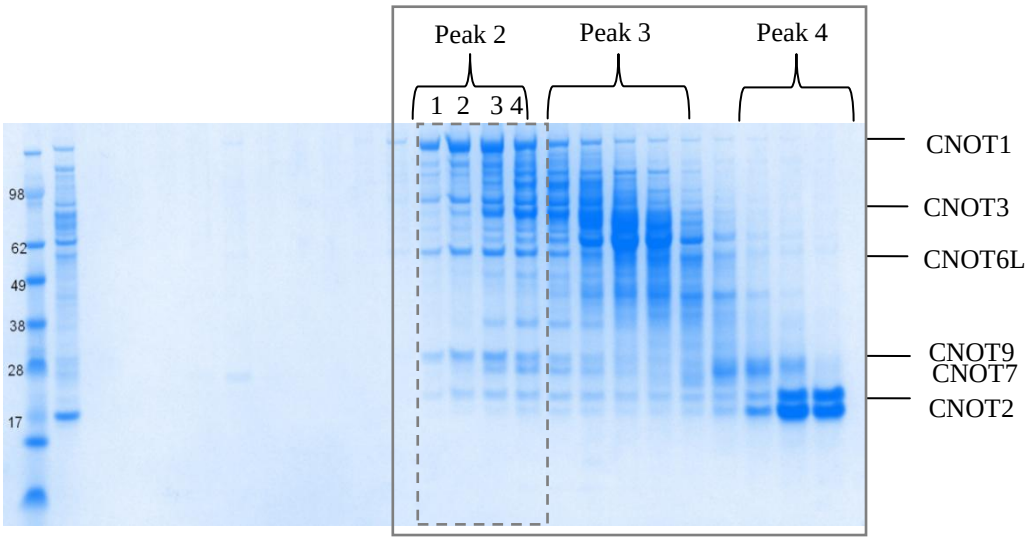


Figure 5.9 Gel filtration of co-infection of CNOT1:CNOT9:CNOT6L:CNOT7:(CNOT2:CNOT3) with truncated CNOT2:CNOT3

Fractions from gel filtration of co-infection of CNOT1:CNOT9:CNOT6L:CNOT7:(CNOT2:CNOT3) with CNOT2 (aa 391-540):CNOT3 were visualised on a SDS-PAGE gel. The boxed fractions are those within the box in **Figure 5.8** and contain the complex of CNOT subunits. The fractions within the dashed box have been enlarged in **Figure 5.10**, and labelled with CNOT subunits, as identified by tandem MS.

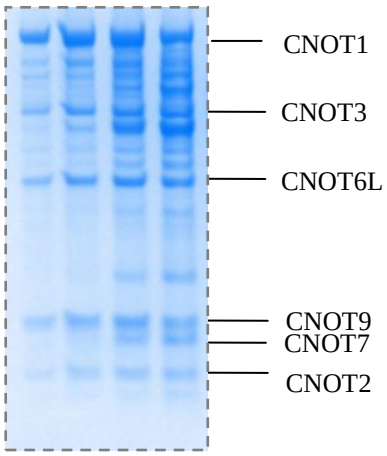


Figure 5.9 Insert from gel filtration fractions of co-infection of CNOT1:CNOT9:CNOT6L: CNOT7:(CNOT2:CNOT3) with truncated CNOT2:CNOT3

Fractions within the dashed box in **Figure 5.9** have been enlarged and labelled with the identity of proteins. These fractions represent the gel filtration of co-infection of CNOT1:CNOT9:CNOT6L :CNOT7:(CNOT2:CNOT3) with CNOT2 (aa 391-540):CNOT3 were visualised on a SDS-PAGE gel. Molecular weight markers are indicated for size comparison (kDa).

5.5 Characterisation of CNOT complexes using a fluorescence-based deadenylase assay

The final screening step for the reconstitution of CCR4-NOT complexes was to determine if any complexes expressed by MultiBac co-expression and isolated by chromatography purification were enzymatically active. This step served to both ensure complexes were isolated in a near-native state and thus were active, but also to provide a means whereby complexes of different subunits could be easily compared to assess the effect of subunit composition on deadenylase activity.

The two deadenylase subunits of the CCR4-NOT complex, CNOT6L and CNOT7, remove the poly-(A) tail of mRNAs, and an assay based on a published procedure was employed in order to probe this activity (Maryati et al., 2014). The assay uses a poly-(A) substrate where the 5′ terminus is fused with a fluorescence molecule, fluorescein (Flc). The substrate nucleotide sequence (Flc-CCUUUCCAAAAAAAAA) was designed to contain a 9 nucleotide poly-(A) segment shown to be removed by CCR4-NOT deadenylases. The substrate was incubated with the CCR4-NOT complex (complex 9)

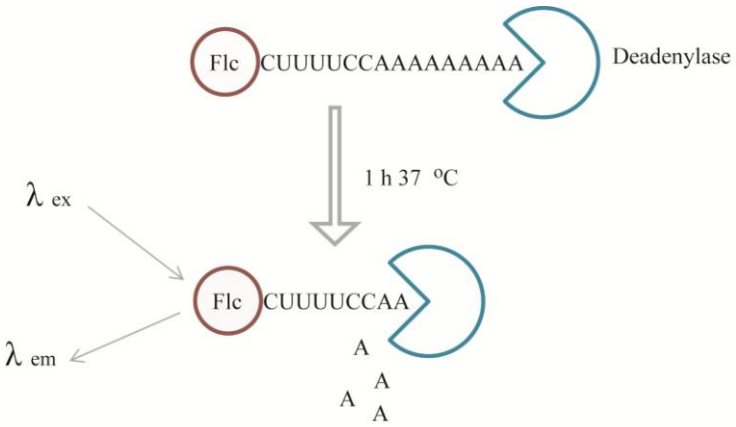


Figure 5.10 Schematic of method to measure deadenylase activity using a fluorescence-based assay

In the deadenylase assay, the deadenylase complex (blue line) is incubated with the fluorescein-poly-(A) substrate for 1 h at 37 ºC. Afterwards the degradation products of the reaction are visualised on a 20% TBE:Urea SDS-PAGE gel, by exciting the fluorescein at 490 nm (λ ex) and measuring emission of fluorescence at 525 nm (λ ex) using a Bio-Rad gel doc imager.

for 1 h at 37 ºC (**Figure 5.11**) and hydrolysis products were visualised on a 20% TBE-urea SDS-PAGE gel.

During the reaction, if complexes are active, the deadenylase subunits therein will hydrolyse the poly-(A) tail of the mRNA substrate, producing products with varying lengths of poly-(A) tail which are separated using urea-PAGE. The extent to which the poly-(A) tail has been degraded and hence the extent of migration by urea-PAGE will be detected by measuring the emitted fluorescence at 525 nm of the Flc-tag using a Bio-Rad gel doc.

A complex of CNOT1:CNOT9:CNOT6L:CNOT7:(CNOT2:CNOT3) (complex 9), where the purification is shown in section 5.4.1 and **Figure 5.6**, was characterised using the fluorescence-based deadenylase assay. The complex (1.5 µM final concentration) was incubated with 1 µM substrate at 37 ºC for 1 h. A control reaction without any protein sample was included to highlight the background level of mRNA degradation. Moreover, a reaction containing CNOT2 (0.4 µM final concentration), a CCR4-NOT subunit not known to have deadenylase activity, was purified in the same buffer and included as a negative control. Additionally, a positive control of *E. coli* Exonuclease T (ExoT), a non-specific exonuclease that cleaves free 3′ termini, was included to ensure the substrate can be degraded in the provided conditions. After 1 h incubation, the sample was visualised by urea-PAGE (**Figure 5.12a**).

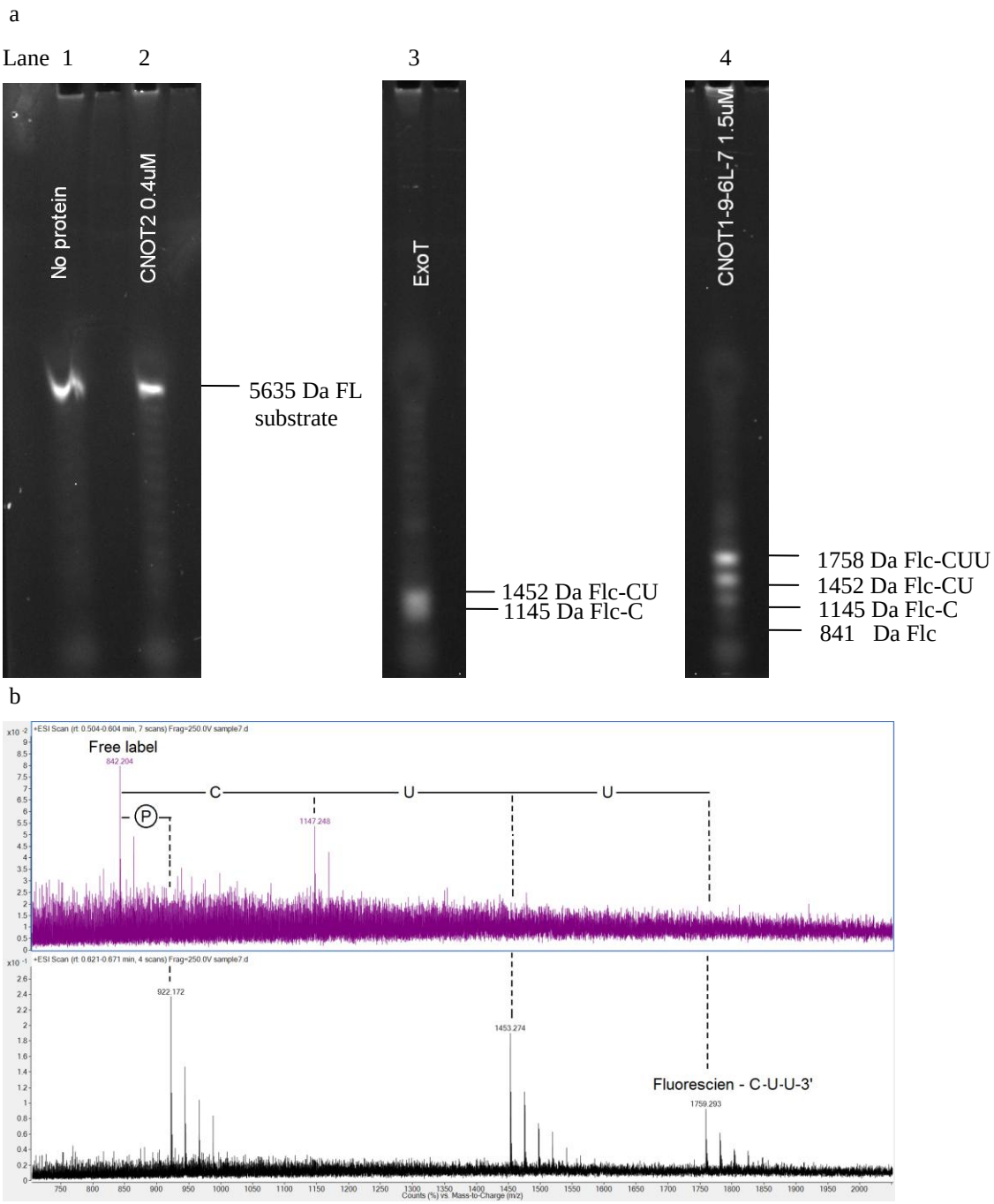


Figure 5.11 Deadenylase assay using a complex of CNOT1:CNOT9:CNOT6L:CNOT7

(a) Results of deadenylase assay using a complex of CNOT1:CNOT9:CNOT6L:CNOT7. Negative controls include a no-protein (lane 1) and CNOT2-alone (lane 2) sample. A positive control of Exonuclease T (ExoT) was included (lane 3). Lane 4 represents the concentrated sample of CNOT1:CNOT9:CNOT6L:CNOT7 (**Figure 5.10)**. Samples were run on a 20% TBE-Urea gel where the products can be visualised by exciting the fluorescein at 490 nm and measuring emission of fluorescence at 525 nm using a Bio-Rad gel doc imager. The samples were also analysed by MS to determine the identity of the degradation products, and results are indicated on the right hand side of the gel lanes (in Da). Full length substrate (5635 Da) was detected in the negative controls. Fragments of the substrate (Flc-CU, 1452 Da and Flc-C, 1145 Da) were detected in the ExoT positive control. Substrate incubated with the CNOT complex was degraded producing fragments Flc-CUU, Flc-CU, Flc-C and Flc. (b) Spectra collected when substrate which had been degraded by the CNOT complex was analysed by MS. Four peaks are detected where the mass corresponds to Flc-CUU (1759 Da), Flc-CU (1453 Da), Flc-C (1147 Da) and the free label Flc (842 Da).

It was anticipated that the no-protein and CNOT2 negative controls would have little effect on the substrate, whereas the positive control of ExoT would non-specifically degrade the entire substrate. As for the CNOT complex, it was expected that the substrate would be degraded such that a ladder of degradation products, representing the sequential removal of nine adenosines, would be visible on the gel as reported in literature (Maryati et al., 2014)

As expected, full length substrate (5635 Da, corresponding to the mass of the RNA substrate with 5′ fluorescein modification) is observed in the no-protein and CNOT2-alone control conditions (**Figure 5.12a**, lanes 1 and 2), which was further confirmed by intact MS. Faint degradation bands are observed in the no-protein condition (lane 1), indicating the low level basal activity of nucleases in the environment. Slightly more pronounced banding can be seen in the CNOT2-alone condition (lane 2), likely due to nucleases introduced during protein expression and purification. The ExoT positive control protein degraded the substrate extensively (lane 3), as indicated by the disappearance of the band corresponding to full length substrate and appearance of smaller bands that were analysed by denaturing MS to correspond to Flu-C and Flu-CU.

The CNOT complex 9 (lane 4) has degraded the substrate such that no full length substrate remains and when lane 4 is compared to the negative (lane 1 and 2) and positive controls (lane 3), it is clear that the catalytic subunits are actively degrading the poly-(A) substrate. Thus, the reconstituted complex of CNOT1:CNOT9:CNOT6L:CNOT7 is enzymatically active with regards to deadenylase activity. MS analysis indicates the entire poly-(A) tail has been removed by the CNOT complex as products with a mass which corresponds to Flc-CUU, Flc-CU, Flc-C and Flc were identified (**Figure 5.12b**).

Detection of fragments Flc-CUU, Flc-CU, Flc-C and Flc indicates that the CNOT complex has degraded the poly-(A) tail of the substrate, and also continued to non-specifically hydrolyse other nucleotides. This phenomenon of non-specific nucleotide hydrolysis by deadenylase enzymes has previously been observed when the activity of yeast deadenylase Pop2 (a homolog of CNOT7) was measured against poly-(A), poly-(U), poly-(C) and poly-(G) substrates, showing activity against all substrates except poly-(G), but with a preference for poly-(A)-containing substrates (Thore et al., 2003).

Hence, it has been shown here, and elsewhere (Niinuma et al., 2016), that CNOT deadenylase subunits have the capacity to degrade non-poly-(A) substrates. Future experiments may be improved by the inclusion of a lane for length markers, consisting of Flc-tagged nucleic acids of known length. This may help to ease the analysis of the assay by eye.

In the above reaction assay, an almost equimolar enzyme and substrate concentration was used in order to determine if the complex contained deadenylase activity. This may have provided an excess of enzyme such that once the poly-(A) tail was cleaved specifically, the enzyme hydrolysed further nucleotides.

To investigate the progression of hydrolysis and allow intermediate degradation products to be identified, the assay was repeated using a serial dilution of the CNOT complex, where the concentration of enzyme ranged between 0.0625 and 1.5 µM (**Figure 5.13**). A fixed substrate concentration of 1 µM was used in all conditions. The negative controls

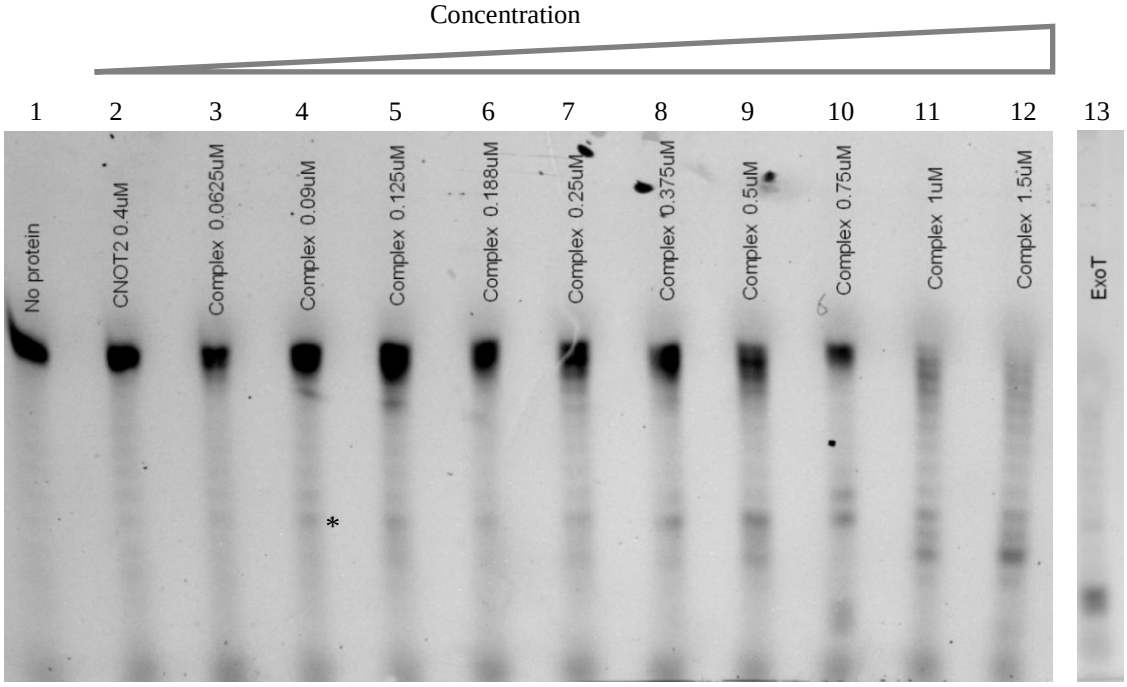


Figure 5.12 Time course characterisation of a complex of CNOT1:CNOT9:CNOT6L:CNOT7 using the deadenylase assay

A time course assay showing the intermediates produced during degradation of the poly-(A) RNA substrate by the CNOT complex. Positive controls included a no-protein condition and a CNOT2-alone condition. Exonuclease T (ExoT) was included to demonstrate non-specific degradation of the RNA substrate. A serial dilution of the complex was made ranging from 0.0625 µM to 1.5 µM and each concentration was assayed using the fluorescein tagged substrate (1 µM). The reaction was incubated for 1 h at 37 ºC and the degradation products were visualised on a 20% TBE-Urea SDS-PAGE gel. The products can be visualised by exciting the fluorescein at 490 nm and measuring emission of fluorescence at 525 nm using a Bio-Rad gel dock imager.

containing no-protein and CNOT2-alone (lanes 1 and 2), and the positive control containing ExoT (lane 13) were again included for comparison, and their observations were as expected. The main body of the gel (lanes 3-12) shows that as the concentration of the CNOT complex increases, the band representing full length substrate decreases and with the concomitant appearance of products of digestion. In conditions with greater than 1 µM complex (1:1 ratio to substrate) (lanes 11 and 12), the full length substrate is

completely degraded. A band (indicated by the asterisk in lane 4) migrating approximately halfway between the full length substrate and the completely degraded substrate (Flc) likely represents the deadenylated substrate (Flc-CCUUUCC), which started to appear in reactions containing >0.09 μM of complex 9 and became more intense as CNOT concentration increased.

Together the results from **Figures 5.12 and 5.13** demonstrate that the CCR4-NOT complexes generated by expression using the MultiBac system are catalytically active with the ability to degrade the poly-(A) tail of an mRNA substrate.

5.6 Conclusion

This chapter described the recombinant expression of the CCR4-NOT complex using the MultiBac system which successfully produced a flexible library of 16 plasmids which can interchangeably be recombined to produce multiple CNOT-containing complexes. As a result of this technique, 26 sub-complexes have been generated, one nearing the complete CCR4-NOT complex. Of these 26 complexes, a sub-complex containing CNOT1, CNOT9, CNOT6L and CNOT7 was characterised using a fluorescence-based deadenylase activity assay and was found to successfully degrade a poly-(A) substrate, demonstrating that it was biologically active. This achievement represents the first published effort to reconstitute the human CCR4-NOT complex and as such, has contributed a valuable reagent to the field for further study of the CCR4-NOT complex. The successful reconstitution of CNOT subunits, as shown here, may facilitate an increased understanding of the function and composition of the human complex through the use of structural and biochemical techniques in future studies.

Characterisation of CNOT1:CNOT9:CNOT6L:CNOT7 indicates that the deadenylase activity of CCR4-NOT complex in vitro does not require the presence of all subunits,

since the isolated complex assayed for activity does not contain CNOT2, CNOT3, CNOT10 or CNOT11. Unfortunately, attempts to use the MultiBac technique to reconstitute the entire CCR4-NOT complex using a single baculovirus were unsuccessful at this time. Although the presence of all CNOT genes were maintained throughout cloning, as determined by PCR screening, their expression was not uniformly detected for all subunits when the entire complex was expressed from a single virus. However, the expression of all subunits has been confirmed when expressed from separate viruses, thus it can be concluded that each CNOT subunit is soluble in *Sf9* cells. One can postulate that either the metabolic burden of expressing an eight-subunit complex was too great for the *Sf9* cells, or that the proteins were expressed at a level too low to be detected by SDS-PAGE and MS at the 3 ml expression test. Alternatively, recombination of the DNA during viral replication may have removed some of the genes. Amplification of the CNOT genes, using specific primers and the viral DNA, may help to clarify if all genes are present in the baculoviruses during infection. Future experiments confirming the subunits within the virus, using expression of these baculoviruses at larger scale and viral titre optimisation may prove beneficial in reconstituting all CNOT subunits from a single virus.

Despite unsuccessful reconstitution of the entire complex, the sub-complexes generated within this study provide a great resource for future study. Firstly, these smaller complexes may provide an alternative strategy to assist in determining the relative position of subunits within cryo-EM models. Two cryo-EM models have been proposed for the yeast CCR4-NOT and the positions of CNOT subunits were assigned using immunomicroscopy and computational modelling. If cryo-EM data could be collected on a range of human CCR4-NOT sub-complexes where the subunit composition is known, the models may be fitted within each other to help identify subunit positioning.

In addition, the smaller sub-complexes present the opportunity for full reconstitution of the CCR4-NOT complex in the future, by the co-expression of two or more sub-complexes. For example, the sub-complex of CNOT10 and CNOT11, successfully reconstituted for the first time in this study (chapter 3), may be co-expressed with the co-infection of CNOT1, CNOT9, CNOT6L and CNOT9 (complex 9), and with CNOT2 (aa 391-540) and CNOT3 (complex 25). This combinatorial strategy may lead to the reconstitution of the whole human CCR4-NOT complex, as an alternative to the use of a single all-gene virus. This full reconstitution may facilitate the most challenging yet potentially highly informative experiment of using cryo-EM to determine the structure of the human CCR4-NOT complex.

The novel findings presented in this chapter detail a strategy of how the CCR4-NOT complex can be reconstituted to facilitate both structural and functional studies. Furthermore, this tool may be used to investigate the increasing number of functions and interaction partners attributed to the subunits of the CCR4-NOT, which have expanded from a role in transcriptional regulation over 30 years ago when it was first discovered, to possible functions in ubiquitination through a transient interaction with CNOT4, and looking to the future where further functions of CNOT subunits or associated proteins may be discovered.

6. Discussion

The CCR4-NOT complex is one of the major deadenylases in humans and has been shown to participate in protein-targeted deadenylation, miRNA-targeted deadenylation and ubiquitination, among other functions (Albert et al., 2002, Chen et al., 2014, Mathys et al., 2014, Mahtani et al., 2001). Deadenylation is the rate-limiting step in mRNA decay therefore; the CCR4-NOT complex is a global regulator of gene expression. Precise regulation of gene expression is vital to preserve normal cellular function as dysregulation of gene expression causes a number of chronic inflammatory diseases including RA. Studying the mechanism of mRNA decay and how it is regulated by interactors of the CCR4-NOT complex is important for understanding how to re-balance chronic inflammation in disease. However, many questions still remain regarding the mechanism of interaction and function of both the stably and transiently associated subunits of the CCR4-NOT complex. Thus, this project sought to characterise a number of different interactions, both within the CCR4-NOT complex and with other protein binding partners, to expand our knowledge of deadenylase-mediated decay.

6.1 TTP recruits the CCR4-NOT complex via tryptophan-mediated interactions with CNOT9

TTP is an RNBP which interacts with the CCR4-NOT complex to facilitate targeted deadenylation of inflammatory mRNAs in response to p38 MAPK signalling (Marchese et al., 2010, Mahtani et al., 2001). Although an interaction between a C-terminal peptide of TTP and the N-terminal HEAT domains of CNOT1 (aa 800-1015) has been established (Fabian et al., 2013), the N-terminus of TTP was found to be important for deadenylation-mediated decay of mRNAs (Lykke-Andersen and Wagner, 2005). Thus, it

was suspected that additional interactions with CNOT subunits were involved in recruitment of TTP to the CCR4-NOT complex. Prior to this DPhil project, researchers at the Kennedy institute have used siRNA knock-down experiments to reveal interactions between TTP and both CNOT9 and CNOT2 (**Appendix Figure 7.14**). Hence, this study aimed to characterise the interface between the CCR4-NOT complex and TTP by validating the interaction of TTP with CNOT2/9 in vitro, using peptide arrays, biophysical techniques and site-directed mutagenesis (chapter 4). Firstly, these techniques confirmed that TTP indeed interacts with both CNOT9 and CNOT2 directly. Secondly, these techniques were utilised to identify the molecular determinants by which TTP binds CNOT9. When either the tryptophan residues in TTP or residues within the tryptophan-binding pockets in CNOT9 were mutated, interactions between TTP and CNOT9 were reduced. Therefore, it was concluded that TTP interacts with CNOT9 in a tryptophan-mediated manner, a mechanism common to other CNOT9-interacting partners (Chen et al., 2014, Mathys et al., 2014). Additionally, this finding provides evidence regarding a role for CNOT9 in the recruitment of RNBPs which interact with the CCR4-NOT complex. Other such examples include the Roquin family of RNBPs which bind CNOT9 via a C-terminal binding motif that recruits the CCR4-NOT complex to deadenylate mRNAs bound by Roquin (Sgromo et al., 2017). Together, this information presents a common function of CNOT9 in recruiting RNBPs.

Although the results presented here validate the interaction between TTP and CNOT9 and present a possible binding mechanism, some questions remain concerning which tryptophans in TTP are bound by which pockets in CNOT9. To rationalise how the binding event could take place at a molecular level, it remains to be determined if CNOT9 interacts with TTP as a dimer, such that four tryptophan-binding pockets would bind all four tryptophans within TTP. There is precedence for CNOT9 dimer formation

when it is not associated with other subunits of the complex (Garces et al., 2007). The CNOT9 homodimer is out-competed by an interaction with CNOT1 as the CNOT9:CNOT1 dimer interface is larger than CNOT9:CNOT9 and thus likely more energetically favourable (Garces et al., 2007). Perhaps a more plausible explanation for why CNOT9 only contains two tryptophan-binding pockets would be that only two of the tryptophans in TTP are important for the interaction with CNOT9. Therefore, further research is required to dissect the effect of combination mutants of TTP, which may shed some light on which tryptophans in TTP associate with which pockets in CNOT9.

One method to reach this conclusion would be to determine the crystal structure of the complex of TTP protein/peptide and CNOT9. Although there are a number of CNOT9 structures published in different complexes, only a structure of a short peptide of TTP is known, as TTP is an intrinsically disordered protein which likely hampers crystallisation attempts. Future crystallisation attempts using short peptides spanning the tryptophan regions may facilitate structural determination. Alternatively, BLI may again be utilised for measuring the binding of the TTP peptides spanning the tryptophans to CNOT9 pocket mutants, to establish which tryptophan binds most strongly to which pocket.

For the TTP-CNOT2 interaction, a peptide array identified nine potential CNOT2-binding peptides on the TTP sequence, where four overlap with those which also were highlighted to bind CNOT9. The stage is now set for future experiments such as ITC, SPR and BLI to validate these interaction motifs. Establishing an interaction between TTP and both CNOT9 and CNOT2 has transformed our knowledge of the interface between TTP and the CCR4-NOT complex and further complicated the concept of TTP recruitment. However, these additional interactions may harbour mechanisms for orientation of the TTP and mRNA substrate and explain why many non-catalytic subunits are required. An important focus for future studies would be to characterise the

sub-complex of CNOT1:CNOT9:CNOT2:TTP to determine how all TTP-CNOT interactions co-operate to mediate TTP recruitment. Important aspects to explore include whether the function of all CNOT-TTP interactions are purely to increase proximity, or if any may be involved in the orientation of the TTP and its mRNA cargo. It is also important to establish if the interactions form simultaneously or act co-operatively. To explore the nature of the binding events, the dissociation constants of TTP bound to either CNOT1-CNOT2, CNOT1-CNOT9 and CNOT1-CNOT9-CNOT2 would need to been calculated (for example using ITC, BLI or SPR). Comparison of the measured dissociation constants may reveal how the subunit composition effects TTP recruitment i.e. if the dissociation constant is smaller (lower K_d, tighter binding) for CNOT1-CNOT9-CNOT2 then this would suggest the presence of all TTP-binding subunits provides a more efficient recruitment of TTP.

The extreme C-terminus of TTP has been shown to bind CNOT1, and tryptophan residues in the N- and C-terminus of TTP bind CNOT9, as presented in chapter 4. Thus, a hypothetical co-operative binding mechanism may follow that TTP associates with the CCR4-NOT complex in a zipper-type motion. In this zipper-type model, supported by dissociation constant determination of TTP-CNOT2/9 interactions, firstly the C-terminus of TTP could be bound by CNOT1 (K_d TTP (aa 312-326) peptide-CNOT1, 2 µM) (**Figure 6.1**, panel b), followed by binding of tryptophans at the TTP C-terminus (Trp 262) to CNOT9 (K_d TTP I01 peptide-CNOT9, 9 µM) (**Figure 6.1**, panel c), and lastly binding of tryptophans at the TTP N-terminus (Trp32, Trp38 and possibly Trp69) to CNOT9 (K_d TTP B02 peptide-CNOT9, 56 µM) (**Figure 6.1**, panel d). While an interesting notion, this theory requires further validation.

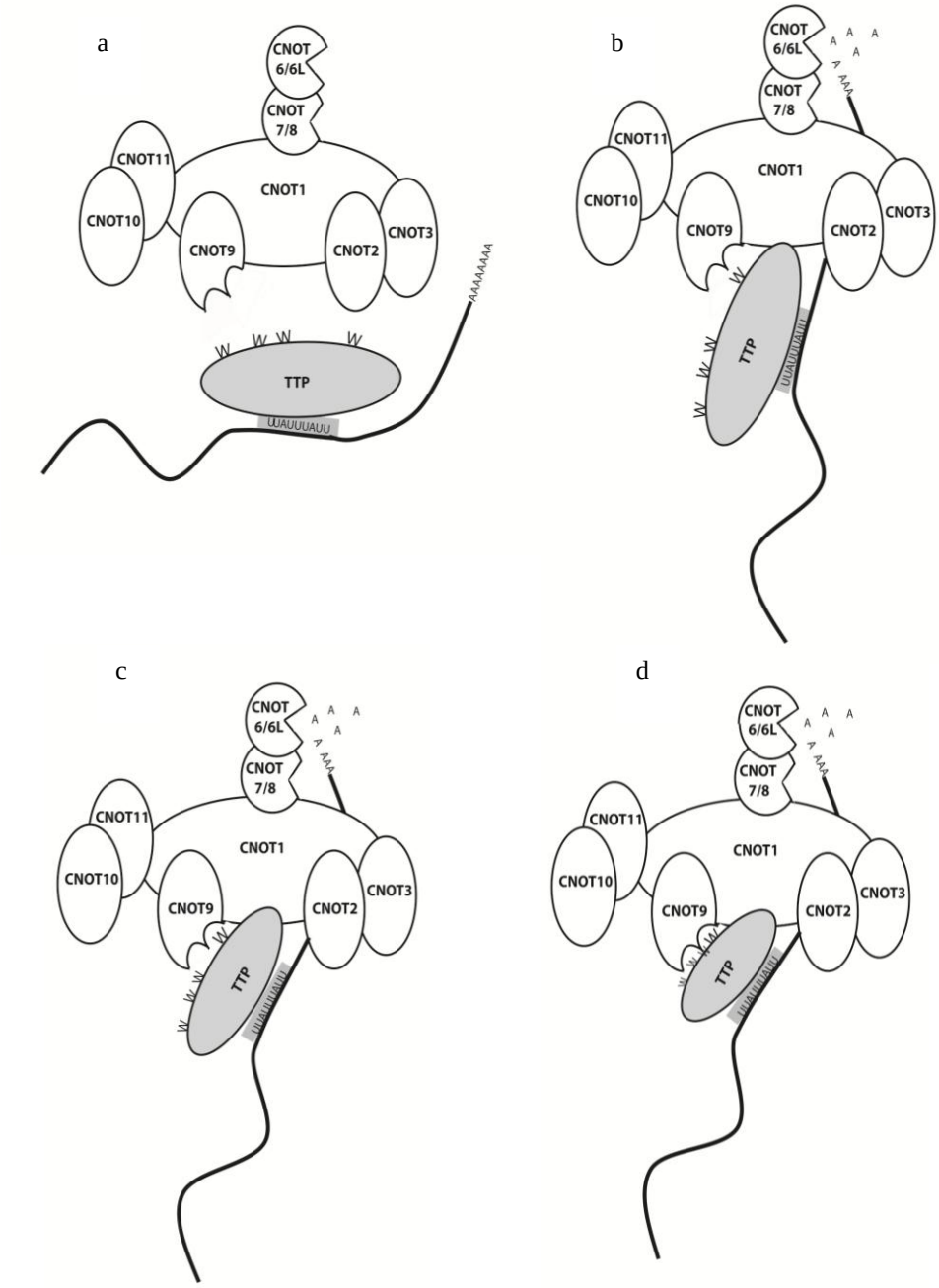


Figure 6.1 Possible zipper-type mechanism of recruitment of TTP by the CCR4-NOT complex

Dissociation constants measured during this project may suggest a possible zipper-type mechanism of TTP recruitment whereby the TTP is initially bound by HEAT repeats in CNOT1 (panel b), followed by C-terminal tryptophan residue binding to CNOT9 (panel c) and lastly the N-terminal tryptophans binding to CNOT9 (panel d). In addition, it is possible that the multiple weak interactions established here between TTP and CNOT2 may act to initially capture TTP, which is then passed to CNOT9 in order to orientate TTP and its mRNA cargo. However, as these interactions are yet to be fully validates, they have not been included in the proposed mechanism.

To explore the finer details of how TTP is recruited by multiple CCR4-NOT subunits, it may be assumed that interactions to increase proximity between the complex and the TTP-mRNA will be made first, followed by interactions to orientate the TTP-mRNA. It may be possible to glean some proximity-related information using fluorescence resonance energy transfer (FRET) which is a technique that measures the excitation of a fluorescence acceptor when in proximity with a fluorescence donor which together act as a spectroscopic ruler. If each of the CNOT subunits shown to interact with TTP (CNOT1, CNOT2 and CNOT9) were produced in separate sub-complexes as a FRET acceptor and donor pair with TTP and the proximity of the interaction measured, the collated data could indicate which FRET pair was excited first. Thus, the pattern of FRET pair excitation will show the sequence of binding events. If the FRET pair consisting of the C-terminus of TTP and CNOT1 is excited first then this may suggest the binding mechanism depicted in **Figure 6.1** may be important to recruit TTP.

Unravelling the mechanism of TTP recruitment could present opportunities to design potential therapeutics for the treatment of chronic inflammatory diseases. Potential treatment strategies would aim to degrade inflammatory mRNAs where the overexpression of associated proteins has shown to contribute to the disease. Therefore, increasing the frequency of interaction between the inflammatory mRNA and the CCR4-NOT complex could up-regulate mRNA deadenylation and degradation. One possible therapeutic modality to facilitate the interaction with CCR4-NOT complex may be to design an RNA aptamer. RNA aptamers are nucleic acids which are selected in vitro to bind to specific targets e.g. a nucleic acid, protein or small molecule (Keefe et al., 2010). Aptamers act as DNA antibodies as they are highly specific (K_d= <1 nM for an aptamer targeting HIV-1 REV) and can be crafted to bind multiple targets (Giver et al., 1993). However, unlike antibodies, they are cheaper and simpler to produce because an

expression system is not required. One initial drawback of aptamers was the potential for degradation by circulating nucleases (unmodified aptamers can have half-lives of 2 mins (Keefe et al., 2010)) but chemical modifications such as PEG fusions have increased the half-life of aptamers to 10 days in humans. Thus, this technology has been explored for the treatment of certain human diseases, such as neo-vascular age-related macular degeneration where an RNA aptamer-based treatment (pegaptanib sodium) has been designed to inhibit vascular endothelial growth factor receptor (VEGF-R) (Ng et al., 2006).

An RNA aptamer designed for the treatment of inflammatory diseases would be structured such that it both specifically binds the target mRNA and recruits the CCR4-NOT complex. Leppek et al describe the design of an aptamer which binds both streptavidin and the ARE from TNF-α. The use of this aptamer has been shown to be a successful method for identifying novel RNBPs (Leppek and Stoecklin, 2014). The design of this aptamer may be modified as to replace the streptavidin-binding section with a module which can bind and recruit the CCR4-NOT complex. Such a module may be influenced by the findings revealed in this study and hence could be designed to bind the tryptophan-binding pockets of CNOT9, shown to recruit TTP. Thus, the aptamer will bind both the ARE in inflammatory mRNAs and recruit the CCR4-NOT complex via interaction with CNOT9. This strategy may prove effective for the treatment of chronic inflammatory diseases. Moreover, as aptamer design is simple, the ARE-recognising portion may be adapted to recruit other over-expressed mRNAs e.g. HER2 implicated in human breast carcinoma (Menard et al., 2000). Hence, this strategy may provide a therapeutic application beyond the treatment of chronic inflammatory diseases.

6.2 Co-infection and limited proteolysis generate a complex of CNOT10 and CNOT11 and identify interacting domains

In addition to the TTP-CNOT2/9 interaction, another sub-complex which remains poorly understood is the recently identified sub-complex of CNOT10 and CNOT11, as identified by Y2H screening. Chapter 3 described a study of CNOT10 and CNOT11 which firstly confirmed, using recombinant proteins, that these subunits form a complex. Soluble expression of full length CNOT10 or CNOT11 alone in *Sf9* cells was so far limited. This observation was not wholly unexpected as disordered regions are predicted within the sequence of both CNOT10 and CNOT11 which may be easily degraded in cells. However, as these two proteins were identified to be interaction partners, a co-infection of viruses expressing CNOT10 and CNOT11 was attempted. If the disordered regions were the cause of poor soluble expression, it was hoped that co-expression would allow these proteins to bind, providing protection and reducing degradation. This strategy successfully produced expression of both CNOT10 and CNOT11 and showed, for the first time, the reconstitution of this module of the CCR4-NOT complex.

CNOT10 and CNOT11 were trialled for crystallisation alone with little success. The predicted disordered regions, thought to be the cause of the limited solubility of the proteins alone, may also be responsible for hampering crystallisation attempts as the formation of an ordered lattice may have been deterred by flexible, disordered regions. However, the CNOT10:CNOT11 complex was then used for crystallisation trials where small crystals were observed in a condition with fungal growth. The observation that crystals were not obtained in drops without fungal growth suggests the fungus is in some way the determining factor in CNOT10:CNOT11 crystallisation. The fungus may have

assisted in crystal formation by the secretion of proteases, cleaving flexible regions, thus facilitating crystal packing, as has been shown elsewhere (Mandel et al., 2006).

As a method to both replicate similar crystallisation conditions and isolate protected domains which may represent the interacting regions of the proteins, limited proteolysis was used. A stable domain of each protein was identified and it was established, using gel filtration, that these protected domains remain bound together. Therefore, the protected domains, tentatively mapped to CNOT10 (aa 33-444) and CNOT11 (aa 214-486) sequences may represent novel and minimal interaction regions. These findings represent the first attempts to characterise the domain architecture of CNOT10 and CNOT11, suggesting previously unknown information that a C-terminal domain of CNOT10 is bound by a region encompassing the domain of unknown function (DUF) of CNOT11.

Moreover, these interaction domains represent excellent crystallisation candidates for future structural studies because of their reduced molecular mass (30 kDa for CNOT11 and 46 kDa for CNOT10) in comparison to the full length proteins, and the removal of disordered regions. Determining the structure of the CNOT10:CNOT11 complex will provide an atomic view of the interface between two proteins, and may reveal additional folded domains which could be indicative of a certain function that may be relevant to the role of these two subunits within the CCR4-NOT complex (e.g. zinc-finger containing proteins often bind nucleic acids and leucine zipper domains are known to mediate protein-protein interaction) (Buchanan and Gay, 1996, Laity et al., 2001). If a characteristic domain such as this was revealed in the structure of the complex, it may help to direct future functional exploration.

Beyond structural analysis, comparison with the CCR4-NOT complex from other organisms may also reveal possible functions for the module of CNOT10 and CNOT11. In *T. brucei*, *Tb*CNOT10 may ensure the stability of the complex by forming an interaction with *Tb*CNOT7 because when *Tb*CNOT10 is knocked-down, *Tb*CNOT7 no longer migrates with the complex during sedimentation. This result is surprising as it would indicate that the mammalian complex evolved additional subunits to aid stability, whereas the yeast CCR4-NOT forms a stable complex without CNOT10:CNOT11, as a orthologous module to CNOT10 and CNOT11 has not been found. Although a Y2H screen using human CNOT10 and CNOT7 did not indicate these subunits form a direct interaction, this finding may be validated using the recombinant complex generated here in a simple pull-down experiment. Furthermore, determining the stability of the human CCR4-NOT complex with and without CNOT10:CNOT11 (e.g. by measuring the melting temperature using differential scanning fluorimetry (DSF)) may also help to ascertain if indeed stabilisation of CNOT7 is the role of CNOT10.

Further to a possible role in binding CNOT7, another potential avenue to explore for future studies should include assessing the DNA/RNA-binding capabilities of CNOT10 and CNOT11, as it was noted that a large amount of nucleic acids were often co-purified with these proteins in this study. An EMSA may be useful to establish nucleic acid binding activity where the CNOT10:CNOT11 complex could be incubated with RNA strands of a specific sequence to either represent the mRNA body or the tail. The incubated sample would then be analysed by non-denaturing polyacrylamide or agarose gel electrophoresis where the binding of nucleic acids impairs migration, thus the degree of impairment may be taken as an indicator of binding. CNOT10 and CNOT11 would be analysed separately, assuming they were able to be purified, and then as a complex to distinguish if a single subunit or the complex contained the nucleic acid-binding activity.

Another CNOT subunit, CNOT9, has been shown to contain a positively charged cleft which binds nucleic acids, although no preference for poly-(A) has been determined indicating CNOT9 may have a role in binding the mRNA body (Garces et al., 2007). If CNOT10:CNOT11 is found to bind nucleic acids, it would need to be further established if a preference for poly-(A) is observed and if perhaps CNOT10:CNOT11 co-operates with CNOT9 to stabilise/orientate the mRNA to facilitate deadenylation. DSF could be used to determine the stability by calculating the melting temperature of a complex of TTP-mRNA-CNOT9 and TTP-mRNA-CNOT9-CNOT10-CNOT11. If it is found that the melting temperature of the TTP-mRNA-CNOT9-CNOT10-CNOT11 is higher, then it would indicate that CNOT10 and CNOT11 help to stabilise the complex by binding RNA.

The CNOT10:CNOT11 sub-complex has been shown to be recruited by the N-terminal domain of CNOT1, a region which also binds vertebrate conserved DND1, an RNBP affecting the mRNA levels of genes associated with regulating pluripotent stem cells and cancer (Yamaji et al., 2017, Kedde et al., 2007). As some evidence suggests CNOT10:CNOT11 only evolved in higher eukaryotes to improve the stability of the CCR4-NOT complex and DND1 is a vertebrate-only protein, a speculative function of CNOT10:CNOT11 may be to facilitate the interaction of higher eukaryotic RNBP such as DND1, or other transient interactors of the CCR4-NOT complex, with the N-terminus of CNOT1. A similar mechanism of facilitating the binding of RNBPs is observed elsewhere in the CCR4-NOT complex. TTP is recruited to the CCR4-NOT complex by interactions made with the N-terminal region of CNOT1 and also, as shown in this study, with CNOT9. Moreover, CNOT1 and CNOT9 also form a platform for the interaction with other binding partners e.g. GW182, a component of the RISC complex. Thus, it may

be that CNOT10:CNOT11 assist CNOT1 in forming an interaction surface which can provide an interface for the interaction of higher eukaryote-specific interacting proteins.

6.3 Enzymatically active sub-complexes of the CCR4-NOT were generated using the MultiBac system

Beyond the two-subunit sub-complexes described above, the final focus of this project directed the reconstitution of the CCR4-NOT complex using the MultiBac system for baculoviral expression. As such, each CCR4-NOT subunit was cloned, in pairs, into the various MultiBac donor and acceptor vectors. The 16 single vectors were recombined to successfully produce a library of 26 CNOT-expressing multi-gene viruses where the resulting protein expression levels within infected cells were tested in a HTP manner. Twelve well-expressed sub-complexes were generated, with one particular containing full length CNOT1 and CNOT9 along with the catalytic subunits, CNOT6L and CNOT7. This reconstituted sub-complex was assayed for deadenylase function and confirmed to effectively degrade a poly-(A) substrate. It is of note that until now, there has been no report of the recombinant production of the human CCR4-NOT complex, or sub-complex therein. Hence, the successful production of biologically active complexes and sub-complexes has now provided an important tool for future experiments such as cryo-EM.

The large size of the CCR4-NOT complex (>1.0 MDa) and the structural flexibility of subunit domains pose problems for x-ray crystallography. Hence, other structural techniques may be required to shed light on the CCR4-NOT complex. One such technique is cryo-EM, which has gained a growing following over the last few years due to impressive advances in technology and understanding leading to previously unseen resolutions (e.g. atomic resolution of viruses) and visualisation of particles sizes beyond the reach of x-ray crystallography (Russo and Passmore, 2016, Liu et al., 2010). Two

previous structures of the endogenous yeast CCR4-NOT complex have been solved by cryo-EM with resolutions of ~20 Å (Nasertorabi et al., 2011, Ukleja et al., 2016b). A module similar to that of CNOT10:CNOT11 has yet to be found in yeast, thus generating a high resolution model of the human CCR4-NOT complex remains an important goal for the near future to reveal insight regarding its function and how the subunits co-operate to facilitate deadenylation.

One method which may facilitate the generation and increase the resolution of a human CCR4-NOT cryo-EM model is to use recombinant expression techniques to reconstitute the complex, as described in chapter 5. Recombinant expression may improve upon the yield but also the purity in comparison to endogenous purifications, as fusion tags can be incorporated which provide a means to capture complexes and remove endogenous contaminating proteins. Thus, providing a more homogeneous sample for structural studies and improving the quality of data collection by reducing background noise. This may help to provide answers to remaining questions.

One such unanswered question relating to the activity of the complex concerns the deadenylase subunits of which there are two, each with two mutually exclusive isoforms of each, CNOT6/6L and CNOT7/8. It remains to be determined why two catalytic subunits are required when both have been shown to be biologically active alone, in vitro (Lau et al., 2009, Wang et al., 2010, Maryati et al., 2014). One explanation may be that each enzyme is active under different conditions e.g. cellular location as Lau et al suggests, evidenced by gene ontology analysis of interactors, that CCR4-NOT complexes containing CNOT8 have a role in mRNA splicing in the nucleus whereas complexes containing CNOT7 function in splicing as well as transport and localisation of mRNA in the cytoplasm. An alternate explanation is that each enzyme possesses different binding capabilities for auxiliary proteins as it has been shown, that PUF RNBPs bind more

strongly to CNOT7 than CNOT8 (Van Etten et al., 2012). Moreover, denaturing MS analysis indicates that CNOT6/6L protein purified from HeLa cells associates more strongly with CNOT7 than CNOT8 in vitro (Lau et al., 2009). Thus, it is apparent that certain differences exist between the subunits and the isoforms, but the biological relevance of the variation in subunit incorporation is as yet unclear. Therefore, future studies, exemplified below, could aim to determine the functional niche of each subunit.

In vivo UV-crosslinking and immunoprecipitation (iCLIP) may be used to physically link the targeted mRNAs with the deadenylase subunits to determine the types of mRNA targeted by different deadenylase isoforms. iCLIP is a technique where UV irradiation is used to form covalent bonds between protein and its bound RNA inside cells; after which, the cells are lysed and the RNA is partially digested. Immunoprecipitation captures the RNA-protein complexes which are purified by SDS-PAGE. Subsequently, reverse transcription and PCR amplification of the cross-linked RNA reveals the identity of the RNA. If this technique could be used with different cell types, expressing different subunit compositions, a pattern of which types of mRNA are targeted by which subunit may become clear.

Furthermore, in an effort to study how the combination of deadenylase subunits affects the type of mRNA degraded, the reconstitution of human CNOT subunits in sub-complexes, as described in this work, could be implemented to generate all permutations of the deadenylase module in order to measure the activity against a range of different mRNA substrates. This may reveal which permutation is most effective at deadenylating which type of mRNA, thus explaining the requirement for multiple enzymes within one complex. Ultimately, unravelling why the CCR4-NOT has evolved such a deadenylase module will help us understand the complexity of deadenylation and how it is modified within and across cell types.

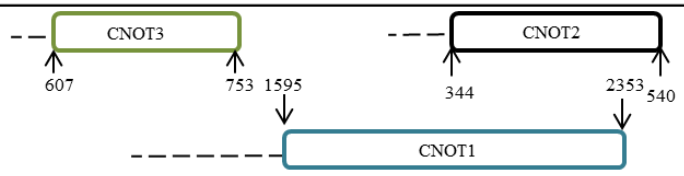
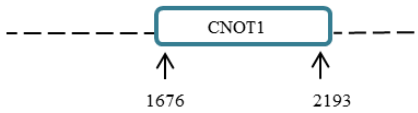
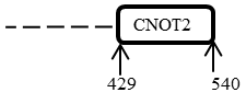
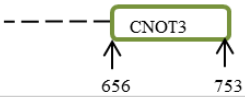
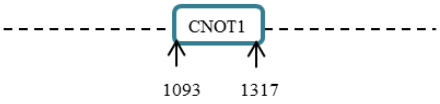
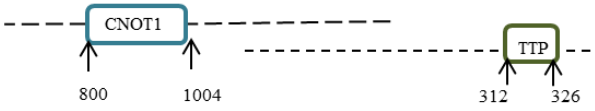
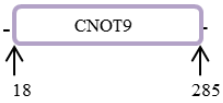
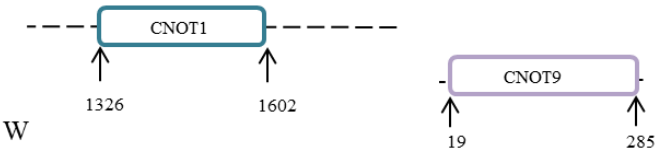
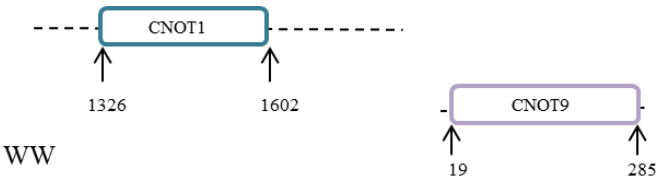
In summary, this study has embellished the current literature and increased our knowledge concerning the CCR4-NOT complex and its interaction partners. These findings firstly provide a set of reagents and tools for reconstituting the human CCR4-NOT complex, and secondly highlight a starting point for future research by identifying interacting domains and novel binding partners. This research can now be elaborated to fully understand the recruitment of TTP and the function of CNOT10:CNOT11. Although a number of functions have been established for the CCR4-NOT complex, our current knowledge may have only scratched the surface in what is an exciting research field concerning a hugely important protein complex. Thus, we can look forward to great advancements and the answers to the remaining puzzles.

7. Appendix

7.1 List of appendix figures and tables

Table 7.1 Solved structures of CCR4-NOT subunits

19 structures have been solved of the CCR4-NOT subunits where the PDB codes, authors, amino acids boundaries and a domain schematic are detailed.

Author	Structure summary	Schematic structure summary	PDB code
(Boland, Chen et al. 2013)	CNOT1 (aa 1565-2371) and CNOT2 (aa 344-540) and CNOT3 (aa 607-753)		4C0D
(Boland, Chen et al. 2013)	CNOT1 (aa 1676-2193) from *C.thermophilus*		4C0E
(Boland, Chen et al. 2013)	CNOT2 (aa 429-540)		4C0F
(Boland, Chen et al. 2013)	CNOT3 (aa 656-753)		4C0G
(Petit, Wohlbold et al. 2012)	CNOT1 (aa 1093-1317) and CNOT7 (aa 1-285)		4GMJ
(Petit, Wohlbold et al. 2012)	CNOT1 (aa 1093-1317)		4GML
(Fabian, Frank et al. 2013)	CNOT1 (aa 800-1004) and TTP (aa 312-326)		4J8S
(Garces, Gillon et al. 2007)	CNOT9 (aa 18-285)		2FV2
(Chen, Boland et al. 2014)	CNOT1 (aa 1326-1607) and CNOT9 (aa 19-285) with 1 tryptophan		4CRU
(Chen, Boland et al. 2014)	CNOT1 (aa 1326-1607) and CNOT9 (aa 19-285) with 2 tryptophan		4CRV

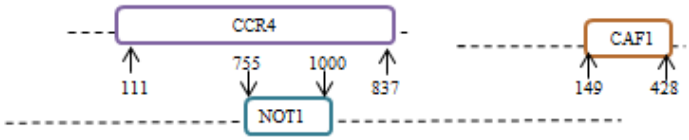
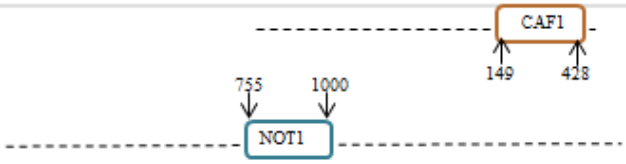
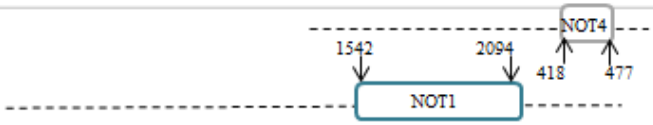
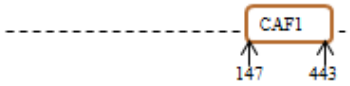
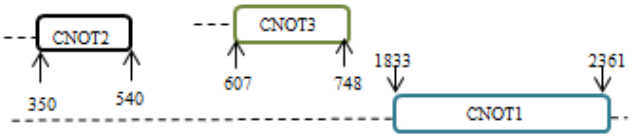
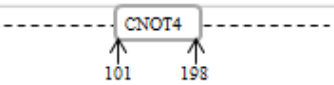
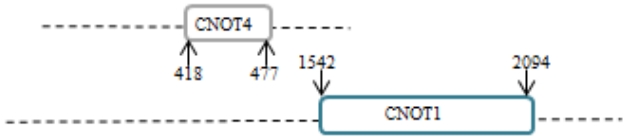
Author	Structure summary	Schematic structure summary	PDB code
(Basquin et al. 2012)	CAF1 (aa 149-428), CCR4 (111-837) and NOT1 (755-1000) from yeast		4B8C
(Basquin et al. 2012)	NOT1 (aa 755-1000) and CAF1 (aa 151-433) from yeast		4B8A
(Basquin et al. 2012)	NOT1 (aa 151-753) from yeast		4B8B
(Bhaskar et al. 2015)	NOT1 (aa 1542-2094) and NOT4 (aa 418-477) from yeast		4AJD
(Thore et al. 2003)	CAF1 (aa 147-443) from yeast		1U0C
(Wang et al. 2010)	CNOT6L (aa 158-555)		3NGN
(Raisch, Bhandari et al. 2016)	CNOT1 (aa 1833-2361 (H2344E, C2345E, A2346E)), CNOT2 (aa 350-540) and CNOT3 (aa 607-748)		5FU6
Nagata, T. (unpublished)	CNOT4 (aa 101-198) from mouse		2CPI
(Bhaskar et al. 2015)	NOT1 (aa1542-2094) and NOT4 (aa 418-477) from yeast		5AJD

Table 7.2 Forward primers used during LIC cloning

Forward primers, along with reverse primers listed in **Table 7.3**, were used to LIC clone all the constructs in **Table 7.5-7.14.**

Forward primer ID	Forward primer sequence
TTPA-f000	tacttccaatccatggatctgactgccatctacgagagc
TTPA-f001	tacttccaatccatgagccctgacgtgcccgtgccat
TTPA-f002	tacttccaatccatgtccagcccaggctggggctcct
TTPA-f003	tacttccaatccatgtccaccacccccctcgcgctaca
TTPA-f004	tacttccaatccatggatctgactgccatctacgag
TTPA-f013	tacttccaatccatgagccctgacgtgcccgtgcca
TTPA-f014	tacttccaatccatgtccaccacccccctcgcgct
TTPA-f015	tacttccaatccatgatggatctgactgccatctacgagagcc
TTPA-f016	tacttccaatccatgtccaccacccccctcgcgctacaagactgagc
TTPA-f017	tacttccaatccatgagccctgacgtgcccgtgccatccgacc
TTPA-f008	tacttccaatccatgtccaccacccccctcgcgc
TTPA-f007	tacttccaatccatgagccctgacgtgcccgtg
TTPA-f020	accaggagcaaacgggaggtgagggccgcagctgtggc
TTPA-f006	tacttccaatccatggatctgactgccatcta
TTPA-f019	accaggagcaaacgggaggttccaccacccccctcgcgc
TTPA-f018	accaggagcaaacgggaggtatggatctgactgccatc
TTPA-f005	tacttccaatccatggatctgactgccatctacg
TTPA-f011	tacttccaatccatgccaggctggggctcctcg
TTPA-f010	tacttccaatccatgagccctgacgtgcccgtg
TTPA-f009	tacttccaatccatggatctgactgccatcta
TTPA-f023	tacttccaatccatgtccaccacccccctcgcgc
TTPA-f012	actaattcttctatgtccaccacccccctcgcgctaca
TTPA-f022	cagaccaacgcagcagcttccaccacccccctcgcgc
CNOT1B-f002	tacttccaatccatggaaaatatccaggagaaaattgcttttatttttc
CNOT1B-f003	tacttccaatccatgcagtacagctaccacgacatc
CNOT1B-f006	tacttccaatccatgggaggcccaaactttatgatgc
CNOT1B-f005	tacttccaatccatggcttgggcaacagatgatgt
CNOT1B-f004	tacttccaatccatggacccaaagcagttggctgt
CNOT1B-f007	tacttccaatccatgtatgatgaccctccaggcct
CNOT1B-f001	tacttccaatccatggagcccccagaaaatatccag
CNOT10A-f002	tacttccaatccatgcagggagcagagaaacatga
CNOT11A-f001	tacttccaatccatgcccggcggaggggcgag
CNOT11A-f003	tacttccaatccatgagcttgaccccgaagga
CNOT1A-f001	tacttccaatccatgaatcttgactcgctctcgct
CNOT1A-f003	tacttccaatccatgcagcagacggacttgtct
CNOT11A-f005	tacttccaatccatgaagaatagcactggtgtggaga
CNOT1B-f008	tacttccaatccatgaatcttgactcgctctcgct
CNOT1B-f013	tacttccaatccatgcagcagtcaaaaatgaagccg

CNOT1B-f014	tacttccaatccatggagcccccagaaaatatcca
CNOT1A-f002	tacttccaatccatgaataacgacccttttgtacagag
CNOT1B-f016	tacttccaatccatggacccaaagcagttggctg
CNOT1B-f017	tacttccaatccatggcttgggcaacagatgatg
CNOT10A-f001	tacttccaatccatggctgcagacaagcctgc
CNOT1B-f010	tacttccaatccatgaataacgacccttttgtacagag
CNOT10A-f004	tacttccaatccatgggtgaatgcttgagatgcatg
CNOT10A-f003	tacttccaatccatgaatacagcagtagctgagtttttaaaag
CNOT1B-f015	tacttccaatccatgcagtacagctaccacgaca
CNOT11A-f002	tacttccaatccatgaccgccgcggagcaaagag
CNOT11A-f004	tacttccaatccatggttgcctctcagatcacaga
CNOT1B-f009	ttaagaaggagatatactatgaataacgacccttttgtacagagga
CNOT1B-f011	tacttccaatccatgaataacgacccttttgtacagagg
CNOT1B-f018	tacttccaatccatggagcccccagaaaatatcca
CNOT1B-f020	aaataattttatttataaggagatataccatgaataacgacccttttgtacaga
CNOT1B-f019	accaggagcaaacgggaggtaataacgacccttttgtac
CNOT10C-f102	aaataattttatttataaggagatataccatggctgcagacaagcctgc
CNOT10A-f102	aaataattttatttataaggagatataccatggctgcagacaagcctgc
CNOT11A-f101	aaataattttatttataaggagatataccatgcccggcggaggggcgag
CNOT11A-f102	tacttccaatccatgcccggcggaggggcgag
CNOT10C-f101	tacttccaatccatggctgcagacaagcctgc
CNOT10A-f101	tacttccaatccatggctgcagacaagcctgc
CNOT11A-f007	tacttccaatccatgtttattcgtccaccgcctcca
CNOT11A-f006	tacttccaatccatgcccggcggaggggcgagc
CNOT2A-f001	tacttccaatccatggtgaggactgatggacatacattat
CNOT2A-f003	tacttccaatccatggcatttggaatgaataactcc
CNOT2A-f002	tacttccaatccatggcactaggccttccaatgagg
CNOT2A-f004	tacttccaatccatggtgacggaccaatttggaatg
CNOT2A-f005	tacttccaatccatgtctcctgaaaatctctacccc
CNOT2A-f006	tacttccaatccatggacttccatgttccatctga
CNOT2A-f007	tacttccaatccatggtgaggactgatggacataca
CNOT2A-f008	tacttccaatccatgactaacattcctcaagggatggtg
CNOT2A-f009	tacttccaatccatgctgaactctcctgaaaatctctacc
CNOT2A-f013	tacttccaatccatgtctcctgaaaatctctaccccaaatttgc
CNOT2A-f014	tacttccaatccatgaacattcctcaagggatggtgacggacc
CNOT2A-f015	aaataattttatttataaggagatataccatggcatttggaatgaataactcct
CNOT2A-f016	aaataattttatttataaggagatataccatgtctcctgaaaatctctacccc
CNOT2A-f012	tacttccaatccatggcatttggaatgaataactccttatc
CNOT2A-f011	ttaagaaggagatatactatgtctcctgaaaatctctaccccaaat
CNOT2A-f010	tacttccaatccatgtctcctgaaaatctctaccccaaatttg
CNOT2A-f020	tacttccaatccatggaagaccttctcttctatctctattacatgaatgg
CNOT2A-f019	tacttccaatccatggataagctggctgcaataaaacttgg
CNOT2A-f018	tacttccaatccatggtgacggaccaatttggaatg
CNOT2A-f017	tacttccaatccatgactaacattcctcaagggatggtg
CNOT2A-f024	tacttccaatccatggacttccatgttccatctga
CNOT2A-f023	tacttccaatccatgtctcctgaaaatctctacccc

CNOT2A-f022	tacttccaatccatggcatttggaatgaataactcct
CNOT2A-f021	tacttccaatccatggtgaggactgatggaca
CNOT3A-f001	tacttccaatccatggcggacaagcgcaaact
CNOT3A-f002	tacttccaatccatgctgagcagtacatcagcacc
CNOT3A-f003	tacttccaatccatgctcaccaaggagcagctcta
CNOT3A-f004	tacttccaatccatgcactcggacactgtggaattc
CNOT3A-f005	ttaagaaggagatatactatggcggacaagcgcaaactc
CNOT4B-f001	tacttccaatccatgtctcgcagtcctgatgcg
CNOT4B-f002	tacttccaatccatggctagtgtacgtgtcgtaca
CNOT4B-f003	tacttccaatccatgaaagtgacaccactgcagagc
CNOT6LA-f001	tacttccaatccatgagactaatagggatgccaaagga
CNOT6LA-f003	tacttccaatccatgcttgacaatctcgcagttcatc
CNOT6LA-f004	tacttccaatccatgcttcctccgaggccatgga
CNOT6LA-f002	tacttccaatccatgttgacacacttgacagcgc
CNOT6A-f001	tacttccaatccatgcccaaagaaaaatacgagccc
CNOT6A-f002	tacttccaatccatgctaactcacctgacagctttgc
CNOT6A-f003	tacttccaatccatgttgcttgataatttgtcaggt
CNOT7A-f001	tacttccaatccatgccagcggcaactgtaga
CNOT7A-f002	tacttccaatccatgagaatttgtgaagtttgggcttg
CNOT8A-f001	tacttccaatccatgcctgcagcacttgtgga
CNOT8A-f003	tacttccaatccatggccagtaatctagaagaagagatg
CNOT8A-f002	tacttccaatccatgcaggttatctgtgaagtgtggg
RQCD1C-f001	tacttccaatccatgcacagcctggcgacggctg
RQCD1C-f002	tacttccaatccatggatagagaaaagatctatca
RQCD1C-f004	tacttccaatccatgcacagcctggcgacggctgcgcctgtgc
RQCD1C-f005	tacttccaatccatggatagagaaaagatctatcagtggatc
RQCD1C-f007	aaataattttatttataaggagatataccatggatagagaaaagatctatcagtgga
RQCD1C-f006	aaataattttatttataaggagatataccatgcacagcctggcgacg
RQCD1C-f003	ttaagaaggagatatactatgcacagcctggcgacgg
RQCD1C-f008	tacttccaatccatggatagagaaaagatctatcagtggatcaatgagc
RQCD1C-f009	tacttccaatccatgagagaaaagatctatcagtggatcaatgagc
FBAC1	tattcataccgtcccacca

Table 7.3 Reverse primers used for LIC cloning

Reverse primers, along with forward primers listed in **Table 7.2**, were used to LIC clone all the constructs in **Table 7.5-7.14.**

Reverse primer ID	Reverse primer sequence
TTPA-r000	tatccacctttactgtcactcagaaacagagatgcgattg
TTPA-r002	tatccacctttactgtcagctgatgctctggcgaagcaca
TTPA-r003	tatccacctttactgtcagtcttcgctagggttgtggatg
TTPA-r001	tatccacctttactgtcactggggtggtgcaaaaactccc
TTPA-r004	tatccacctttactgtcactcagaaacagagatgcgattg

TTPA-r008	tatccacctttactgtcactcagaaacttcgatgcgattg
TTPA-r007	tatccacctttactgctctcagaaacagagatgcgattg
TTPA-r014	tatccacctttactgtcactggggtggtgcaaaaact
TTPA-r016	tatccacctttactgtcagtcttcgctagggttgtgg
TTPA-r015	tatccacctttactgtcagctgatgctctggcgaagc
TTPA-r010	tatccacctttactgtcactcagaaacggcgatgcga
TTPA-r017	tatccacctttactgctcaggtcttcgctagggttgtggatgaagtggc
TTPA-r018	tatccacctttactgcttcactcagaaacagagatgcgattgaagatg
TTPA-r006	ttataaataaaattatttgtcactcagaaacagagatgcgattga
TTPA-r009	ttataaataaaattatttgtcagctgatgctctggcgaag
TTPA-r021	cagaccgccaccgctgccaggtcttcgctagggttgtgg
TTPA-r011	ttataaataaaattatttgtcagtcttcgctagggttgtgg
TTPA-r019	cagaccgccaccgctctcagaaacagagatgcg
TTP-r022	cagaccgccaccgctgcgcgagggggtggtgg
TTPA-r005	tatccacctttactgtcactcagaaacagagatgcgattg
TTPA-r013	cacatcactactggcgggcccccagacgctgat
TTPA-r012	cacatcactactggcctcagaaacagagatgcga
TTPA-r020	cacatcactactggcgtcttcgctagggttgtg
TTPA-r022	cacatcactactggccacccagccacagctgcg
TTPA-r024	cacatcactactggcggaggggctcaggctcca
TTPA-r023	tatccacctttactgctctcagaaacagagatgcgattgaagatg
TTPA-r027	accaccaccactactgccgtcttcgctagggttgtg
TTPA-r026	cccaccagaaggtttcagtcttcgctagggttgtg
TTPA-r033	tatccacctttactgtcactcactcagaaacagcgatgc
CNOT1B-r004	tatccacctttactgtcataaaaatcccgtgggctgac
CNOT1B-r005	tatccacctttactgtcatggctgcttgacatctttct
CNOT1B-r002	tatccacctttactgtcatcccatgcagcactgtgcga
CNOT1B-r001	tatccacctttactgtcaactggcacctgtcccttcca
CNOT1B-r006	tatccacctttactgtcaagagagttgctcatctaaattcttcagg
CNOT1B-r003	tatccacctttactgtcaatactctgaggcttgagagatcc
CNOT10A-r004	tatccacctttactgtcaatccatttcttcaacagctgagtg
CNOT11A-r001	tatccacctttactgtcattttgacatttcggtctcagaag
CNOT1A-r006	tatccacctttactgtcatctagactgctgtccatactcaatatac
CNOT1A-r004	tatccacctttactgtcaacaattgtcctgagctaattgagc
CNOT11A-r002	tatccacctttactgtcacaatgtcttcaacaaccgga
CNOT1B-r019	tatccacctttactgtcaattcactgcagggagtcca
CNOT1B-r016	tatccacctttactgtcatctagactgctgtccatactcaatatac
CNOT1A-r007	tatccacctttactgtcactggaatgtaggctgattcagt
CNOT1A-r003	tatccacctttactgtcaatactctgaggcttgagagatcc
CNOT1B-r007	tatccacctttactgtcaactggcacctgtcccttcc
CNOT1A-r002	tatccacctttactgtcattccattacttgctgggcc
CNOT1B-r014	tatccacctttactgtcactccacaattctctcagtttgatc
CNOT1B-r011	tatccacctttactgtcaatactctgaggcttgagagatcc
CNOT1B-r009	tatccacctttactgtcattccattacttgctgggcc
CNOT10A-r001	tatccacctttactgtcactttctctgcacagtggtg
CNOT10A-r002	tatccacctttactgtcacagctgattccttttgatgatc

CNOT10A-r003	tatccacctttactgtcaaaaaggctctatgaactgataaagt
CNOT11A-r003	tatccacctttactgtcacatcgatttatcccactggatcg
CNOT1A-r009	tatccacctttactgtcacaggatcctgtccattagaatgt
CNOT1A-r001	tatccacctttactgtcaactggcacctgtcccttcc
CNOT1A-r008	tatccacctttactgtcaattcactgcagggagtcca
CNOT1A-r005	tatccacctttactgtcactccacaattctctcagtttgatct
CNOT1B-r008	gattggaagtagaggttctctgctcaactggcacctgtcccttcc
CNOT1B-r012	gattggaagtagaggttctctgctcaatactctgaggcttgagag
CNOT1B-r010	tatccacctttactgtcaactggcacctgtcccttcc
CNOT1B-r015	tatccacctttactgtcaatactctgaggcttgagagatc
CNOT1B-r020	tatccacctttactgcttggctgcttgacatcttt
CNOT1B-r023	tatccacctttactgcttctagactgctgtccatact
CNOT1B-r022	tatccacctttactgtcatctagactgctgtccatact
CNOT1-r022	cagaccgccaccgcttttcacaggaggatctctagactgctg
CNOT1-r021	cagaccgccaccgcttctagactgctgtccatactca
CNOT1B-r017	tatccacctttactgtcatggctgcttgacatcttt
CNOT11A-r005	tatccacctttactgtcattttgacatttcggtctcag
CNOT10A-r006	tatccacctttactgtcactttctctgcacagtggtga
CNOT1B-r021	tatccacctttactgtcacaggatcctgtccattagaatgt
CNOT10C-r102	tatccacctttactgtcactttctctgcacagtggtg
CNOT10A-r102	tatccacctttactgtcactttctctgcacagtggtg
CNOT1B-r026	tatccacctttactgtcatgcctgtgctccatcac
CNOT11A-r101	tatccacctttactgtcattttgacatttcggtctcagaag
CNOT11A-r102	ttataaataaaattatttgtcattttgacatttcggtctcagaag
CNOT10C-r101	ttataaataaaattatttgtcactttctctgcacagtggtg
CNOT10A-r101	ttataaataaaattatttgtcactttctctgcacagtggtg
CNOT11A-r006	tatccacctttactgtcattttgacatttcggtctcag
CNOT2A-r001	tatccacctttactgtcagaaggcttgctgagcagggttgt
CNOT2A-r002	tatccacctttactgtcatggcaggtgaggccgttcttctaa
CNOT2A-r003	tatccacctttactgtcattccagatggaactccttag
CNOT2A-r004	tatccacctttactgtcagaaggcttgctgagcagggt
CNOT2A-r007	tatccacctttactgctgaaggcttgctgagcagggttgtagttg
CNOT2A-r009	tatccacctttactgctgaaggcttgctgagcagg
CNOT2A-r008	tatccacctttactgtcagaaggcttgctgagcagg
CNOT2A-r006	gattggaagtagaggttctctgctcagaaggcttgctgagcaggg
CNOT2A-r005	tatccacctttactgtcagaaggcttgctgagcaggg
CNOT2A-r010	tatccacctttactgtcagaaggcttgctgagcagg
CNOT2A-r011	cacatcactactggcgaaggcttgctgagcagg
CNOT3A-r001	tatccacctttactgtcactggaggtcccggtcctcca
CNOT3A-r002	gattggaagtagaggttctctgctcactggaggtcccggtcctcc
CNOT4B-r001	tatccacctttactgtcaggccacagtagtgtggaga
CNOT4B-r003	tatccacctttactgtcagtaactgcagtattttgttg
CNOT4B-r002	tatccacctttactgtcaagaggttgacagtggcatg
CNOT6LA-r001	tatccacctttactgtcacctccgattaggcaagtga
CNOT6LA-r002	tatccacctttactgtcagagtggagggtggagttca
CNOT6LA-r003	tatccacctttactgtcatgaagggatgtgagggtgt

CNOT6A-r001	tatccacctttactgtcacctcctgccaggaaggtgg
CNOT6A-r002	tatccacctttactgtcacaggaaaggcagtaagagc
CNOT7A-r001	tatccacctttactgtcatgactgcttgttggcttcc
CNOT8A-r001	tatccacctttactgtcactgctgcatgttgttgataatcg
CNOT8A-r004	tatccacctttactgtcactcaaaaaacaactctttcatcctaaag
CNOT8A-r002	tatccacctttactgtcacttctcctgggcagagtcc
CNOT8A-r003	tatccacctttactgtcaatcctcattctgcttctggg
RQCD1C-r001	tatccacctttactgtcactgaggggggcaggggggatac
RQCD1C-r002	tatccacctttactgtcactcttgcaggttcttcacca
RQCD1C-r004	tatccacctttactgctctgaggggggcaggggggataccccgggg
RQCD1C-r005	tatccacctttactgctgccctcttgcaggttcttcaccagttgtgc
RQCD1C-r009	tatccacctttactgtcactcttgcaggttcttcacc
RQCD1C-r006	tatccacctttactgtcactgaggggggcaggggggat
RQCD1C-r008	tatccacctttactgctctcttgcaggttcttcacc
RQCD1C-r007	tatccacctttactgctctgaggggggcaggggggat
RQCD1C-r003	gattggaagtagaggttctctgctcactgaggggggcaggggggata
FBAC2	gggaggttttttaaagcaagtaaa
RQCD1C-r010	tatccacctttactgtcactcttgcaggttcttcacc

Table 7.4 All entry clones (DNA templates) of CNOT genes and associated proteins

The table lists the source of each entry clone (DNA template sequence) used to generate the constructs in **Table 7.5-7.14.**

CNOT subunit	Entry clone sequence
CNOT1	Gene ID: 23019
CNOT2	IMAGE:3143289
CNOT3	IMAGE:4181383
CNOT4	IMAGE:5576434
CNOT6	Addgene plasmid: 37368
CNOT6L	Addgene plasmid: 37369
CNOT7	IMAGE:2989585
CNOT8	IMAGE:3506054
CNOT9	Gene ID: 9125
CNOT10	IMAGE:3954507
CNOT11	IMAGE:5752134
TTP	IMAGE:3900312

Table 7.5 Constructs cloned within Biotech-Plate-4

Constructs were cloned by LIC to facilitate the study of the CCR4-NOT complex and associated proteins. The construct ID, start and stop residues, forward and reverse primer IDs are listed for each construct. The forward and reverse primer IDs are listed with the primer sequence in **Tables 7.2 and 7.3**, respectively.

TAB182	Gene ID: 85456

Construct ID	Start residue	Stop residue	Forward Primer	Reverse Primer	Construct Plate & Well
CNOT2A-c001	M1	F540	CNOT2A-f001	CNOT2A-r001	Biotech-Plate-4:A01
CNOT2A-c002	M1	P528	CNOT2A-f001	CNOT2A-r002	Biotech-Plate-4:A02
CNOT2A-c003	M1	P528	CNOT2A-f001	CNOT2A-r003	Biotech-Plate-4:A03
CNOT2A-c004	A81	F540	CNOT2A-f002	CNOT2A-r001	Biotech-Plate-4:A04
CNOT2A-c005	A81	P528	CNOT2A-f002	CNOT2A-r002	Biotech-Plate-4:A05
CNOT2A-c006	A81	E517	CNOT2A-f002	CNOT2A-r003	Biotech-Plate-4:A06
CNOT2A-c007	A201	F540	CNOT2A-f003	CNOT2A-r001	Biotech-Plate-4:A07
CNOT2A-c008	A201	P528	CNOT2A-f003	CNOT2A-r002	Biotech-Plate-4:A08
CNOT2A-c009	A201	E517	CNOT2A-f003	CNOT2A-r003	Biotech-Plate-4:A09
CNOT2A-c010	M350	F540	CNOT2A-f004	CNOT2A-r001	Biotech-Plate-4:A10
CNOT2A-c011	M350	P528	CNOT2A-f004	CNOT2A-r002	Biotech-Plate-4:A11
CNOT2A-c012	M350	E517	CNOT2A-f004	CNOT2A-r003	Biotech-Plate-4:A12
CNOT2A-c013	S391	F540	CNOT2A-f005	CNOT2A-r001	Biotech-Plate-4:B01
CNOT2A-c014	S391	P528	CNOT2A-f005	CNOT2A-r002	Biotech-Plate-4:B02
CNOT2A-c015	S391	E517	CNOT2A-f005	CNOT2A-r003	Biotech-Plate-4:B03
CNOT2A-c016	D414	F540	CNOT2A-f006	CNOT2A-r001	Biotech-Plate-4:B04
CNOT2A-c017	D414	P528	CNOT2A-f006	CNOT2A-r002	Biotech-Plate-4:B05
CNOT2A-c018	D414	E517	CNOT2A-f006	CNOT2A-r003	Biotech-Plate-4:B06
RQCD1C-c001	M1	Q299	RQCD1C-f001	RQCD1C-r001	Biotech-Plate-4:B07
RQCD1C-c002	D18	Q299	RQCD1C-f002	RQCD1C-r001	Biotech-Plate-4:B08
RQCD1C-c003	M1	E285	RQCD1C-f001	RQCD1C-r002	Biotech-Plate-4:B09
RQCD1C-c004	D18	E285	RQCD1C-f002	RQCD1C-r002	Biotech-Plate-4:B10
CNOT2A-c019	M1	F540	CNOT2A-f001	CNOT2A-r001	Biotech-Plate-4:B11
CNOT2A-c020	M1	P528	CNOT2A-f001	CNOT2A-r002	Biotech-Plate-4:B12
CNOT2A-c021	M1	P528	CNOT2A-f001	CNOT2A-r003	Biotech-Plate-4:C01
CNOT2A-c022	A81	F540	CNOT2A-f002	CNOT2A-r001	Biotech-Plate-4:C02
CNOT2A-c023	A81	P528	CNOT2A-f002	CNOT2A-r002	Biotech-Plate-4:C03
CNOT2A-c024	A81	E517	CNOT2A-f002	CNOT2A-r003	Biotech-Plate-4:C04
CNOT2A-c025	A201	F540	CNOT2A-f003	CNOT2A-r001	Biotech-Plate-4:C05
CNOT2A-c026	A201	P528	CNOT2A-f003	CNOT2A-r002	Biotech-Plate-4:C06
CNOT2A-c027	A201	E517	CNOT2A-f003	CNOT2A-r003	Biotech-Plate-4:C07
CNOT2A-c028	M350	F540	CNOT2A-f004	CNOT2A-r001	Biotech-Plate-4:C08
CNOT2A-c029	M350	P528	CNOT2A-f004	CNOT2A-r002	Biotech-Plate-4:C09
CNOT2A-c030	M350	E517	CNOT2A-f004	CNOT2A-r003	Biotech-Plate-4:C10
CNOT2A-c031	S391	F540	CNOT2A-f005	CNOT2A-r001	Biotech-Plate-4:C11
CNOT2A-c032	S391	P528	CNOT2A-f005	CNOT2A-r002	Biotech-Plate-4:C12
CNOT2A-c033	S391	E517	CNOT2A-f005	CNOT2A-r003	Biotech-Plate-4:D01
CNOT2A-c034	D414	F540	CNOT2A-f006	CNOT2A-r001	Biotech-Plate-4:D02
CNOT2A-c035	D414	P528	CNOT2A-f006	CNOT2A-r002	Biotech-Plate-4:D03
CNOT2A-c036	D414	E517	CNOT2A-f006	CNOT2A-r003	Biotech-Plate-4:D04
RQCD1C-c005	M1	Q299	RQCD1C-f001	RQCD1C-r001	Biotech-Plate-4:D05
RQCD1C-c006	D18	Q299	RQCD1C-f002	RQCD1C-r001	Biotech-Plate-4:D06
RQCD1C-c007	M1	E285	RQCD1C-f001	RQCD1C-r002	Biotech-Plate-4:D07
RQCD1C-c008	D18	E285	RQCD1C-f002	RQCD1C-r002	Biotech-Plate-4:D08

CNOT2A-c037	M1	F540	CNOT2A-f001	CNOT2A-r001	Biotech-Plate-4:D09
RQCD1C-c009	M1	Q299	RQCD1C-f001	RQCD1C-r001	Biotech-Plate-4:D10
TTPA-c133	M1	E326	TTPA-f000	TTPA-r000	Biotech-Plate-4:D11
TTPA-c134	S98	S186	TTPA-f003	TTPA-r002	Biotech-Plate-4:D12

Figure 7.1 Expression/purification test of Biotech-Plate-4

Expression of constructs cloned as part of Biotech-Plate-4 plate was tested using 1 ml of *E. coli*. Protein was purified using nickel affinity chromatography and elutions were visualised by SDS-PAGE. Each lane is labelled with a well ID which corresponds to information provided in **Table 7.5**. Molecular weight markers are indicated for size comparison (kDa).

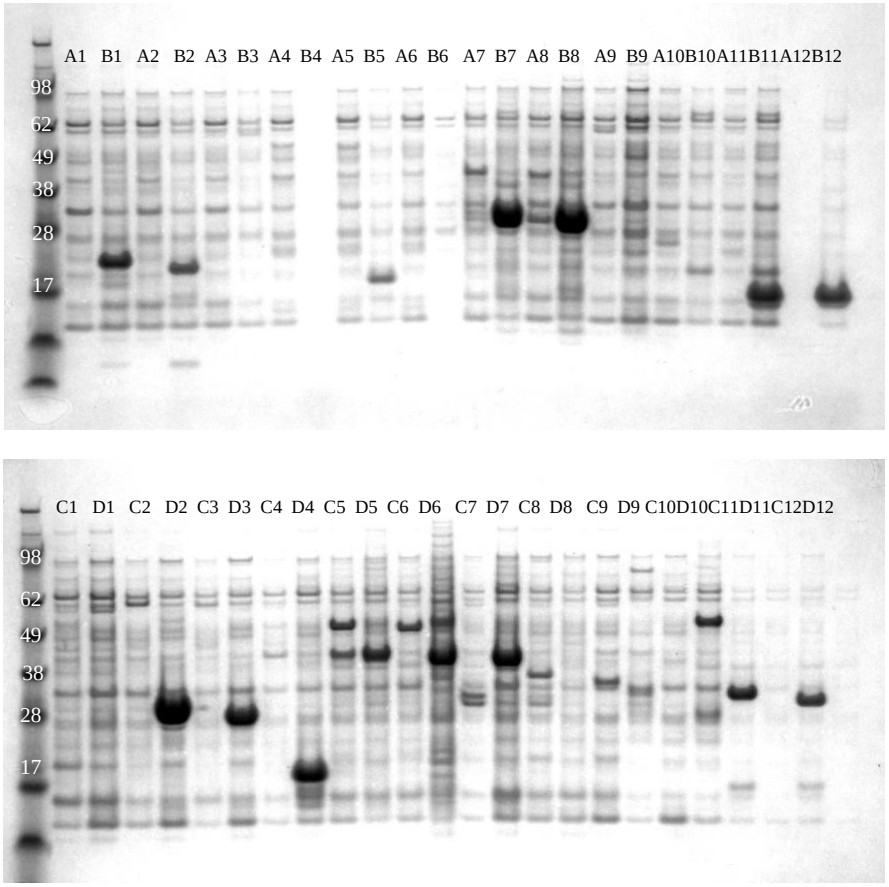


Table 7.6 Constructs cloned withi Biotech-Plate-6

Constructs were cloned by LIC to facilitate the study of the CCR4-NOT complex and associated proteins. The construct ID, start and stop residues, forward and reverse primer IDs are listed for each construct. The forward and reverse primer IDs are listed with the primer sequence in **Tables 7.2 and 7.3**, respectively.

Construct ID	Start residue	Stop residue	Forward Primer	Reverse Primer	Construct Plate & Well

CNOT1B-c001	E1090	S2376	CNOT1B-f001	CNOT1B-r001	Biotech-Plate-6:A01
CNOT1B-c002	E1090	G2362	CNOT1B-f001	CNOT1B-r002	Biotech-Plate-6:A02
CNOT1B-c003	E1090	Y1842	CNOT1B-f001	CNOT1B-r003	Biotech-Plate-6:A03
CNOT1B-c004	E1090	L1595	CNOT1B-f001	CNOT1B-r004	Biotech-Plate-6:A04
CNOT1B-c005	E1090	P1326	CNOT1B-f001	CNOT1B-r005	Biotech-Plate-6:A05
CNOT1B-c006	E1090	S1317	CNOT1B-f001	CNOT1B-r006	Biotech-Plate-6:A06
CNOT1B-c007	E1093	S2376	CNOT1B-f002	CNOT1B-r001	Biotech-Plate-6:A07
CNOT1B-c008	E1093	G2362	CNOT1B-f002	CNOT1B-r002	Biotech-Plate-6:A08
CNOT1B-c009	E1093	Y1842	CNOT1B-f002	CNOT1B-r003	Biotech-Plate-6:A09
CNOT1B-c010	E1093	L1595	CNOT1B-f002	CNOT1B-r004	Biotech-Plate-6:A10
CNOT1B-c011	E1093	P1326	CNOT1B-f002	CNOT1B-r005	Biotech-Plate-6:A11
CNOT1B-c012	E1093	S1317	CNOT1B-f002	CNOT1B-r006	Biotech-Plate-6:A12
CNOT1B-c013	Q1356	S2376	CNOT1B-f003	CNOT1B-r001	Biotech-Plate-6:B01
CNOT1B-c014	Q1356	G2362	CNOT1B-f003	CNOT1B-r002	Biotech-Plate-6:B02
CNOT1B-c015	D1565	S2376	CNOT1B-f004	CNOT1B-r001	Biotech-Plate-6:B03
CNOT1B-c016	D1565	G2362	CNOT1B-f004	CNOT1B-r002	Biotech-Plate-6:B04
CNOT1B-c017	A1602	S2376	CNOT1B-f005	CNOT1B-r001	Biotech-Plate-6:B05
CNOT1B-c018	A1602	G2362	CNOT1B-f005	CNOT1B-r002	Biotech-Plate-6:B06
CNOT1B-c019	G1826	S2376	CNOT1B-f006	CNOT1B-r001	Biotech-Plate-6:B07
CNOT1B-c020	G1826	G2362	CNOT1B-f006	CNOT1B-r002	Biotech-Plate-6:B08
CNOT1B-c021	Y1842	S2376	CNOT1B-f007	CNOT1B-r001	Biotech-Plate-6:B09
CNOT1B-c022	Y1842	G2362	CNOT1B-f007	CNOT1B-r002	Biotech-Plate-6:B10
CNOT2A-c038	M1	F540	CNOT2A-f007	CNOT2A-r004	Biotech-Plate-6:B11
CNOT2A-c039	T344	F540	CNOT2A-f008	CNOT2A-r004	Biotech-Plate-6:B12
CNOT2A-c040	L389	F540	CNOT2A-f009	CNOT2A-r004	Biotech-Plate-6:C01
CNOT3A-c001	M1	Q753	CNOT3A-f001	CNOT3A-r001	Biotech-Plate-6:C02
CNOT3A-c002	L572	Q753	CNOT3A-f002	CNOT3A-r001	Biotech-Plate-6:C03
CNOT3A-c003	L607	Q753	CNOT3A-f003	CNOT3A-r001	Biotech-Plate-6:C04
CNOT3A-c004	H656	Q753	CNOT3A-f004	CNOT3A-r001	Biotech-Plate-6:C05
CNOT1B-c023	E1090	S2376	CNOT1B-f001	CNOT1B-r001	Biotech-Plate-6:C08
CNOT1B-c024	E1090	G2362	CNOT1B-f001	CNOT1B-r002	Biotech-Plate-6:C09
CNOT1B-c025	E1090	Y1842	CNOT1B-f001	CNOT1B-r003	Biotech-Plate-6:C10
CNOT1B-c026	E1090	L1595	CNOT1B-f001	CNOT1B-r004	Biotech-Plate-6:C11
CNOT1B-c027	E1090	P1326	CNOT1B-f001	CNOT1B-r005	Biotech-Plate-6:C12
CNOT1B-c028	E1090	S1317	CNOT1B-f001	CNOT1B-r006	Biotech-Plate-6:D01
CNOT1B-c029	E1093	S2376	CNOT1B-f002	CNOT1B-r001	Biotech-Plate-6:D02
CNOT1B-c030	E1093	G2362	CNOT1B-f002	CNOT1B-r002	Biotech-Plate-6:D03
CNOT1B-c031	E1093	Y1842	CNOT1B-f002	CNOT1B-r003	Biotech-Plate-6:D04
CNOT1B-c032	E1093	L1595	CNOT1B-f002	CNOT1B-r004	Biotech-Plate-6:D05
CNOT1B-c033	E1093	P1326	CNOT1B-f002	CNOT1B-r005	Biotech-Plate-6:D06
CNOT1B-c034	E1093	S1317	CNOT1B-f002	CNOT1B-r006	Biotech-Plate-6:D07
CNOT1B-c035	Q1356	S2376	CNOT1B-f003	CNOT1B-r001	Biotech-Plate-6:D08
CNOT1B-c036	Q1356	G2362	CNOT1B-f003	CNOT1B-r002	Biotech-Plate-6:D09
CNOT1B-c037	D1565	S2376	CNOT1B-f004	CNOT1B-r001	Biotech-Plate-6:D10
CNOT1B-c038	D1565	G2362	CNOT1B-f004	CNOT1B-r002	Biotech-Plate-6:D11
CNOT1B-c039	A1602	S2376	CNOT1B-f005	CNOT1B-r001	Biotech-Plate-6:D12

CNOT1B-c040	A1602	G2362	CNOT1B-f005	CNOT1B-r002	Biotech-Plate-6:E01
CNOT1B-c041	G1826	S2376	CNOT1B-f006	CNOT1B-r001	Biotech-Plate-6:E02
CNOT1B-c042	G1826	G2362	CNOT1B-f006	CNOT1B-r002	Biotech-Plate-6:E03
CNOT1B-c043	Y1842	S2376	CNOT1B-f007	CNOT1B-r001	Biotech-Plate-6:E04
CNOT1B-c044	Y1842	G2362	CNOT1B-f007	CNOT1B-r002	Biotech-Plate-6:E05
CNOT2A-c041	M1	F540	CNOT2A-f007	CNOT2A-r004	Biotech-Plate-6:E06
CNOT2A-c042	T344	F540	CNOT2A-f008	CNOT2A-r004	Biotech-Plate-6:E07
CNOT2A-c043	L389	F540	CNOT2A-f009	CNOT2A-r004	Biotech-Plate-6:E08
CNOT3A-c005	M1	Q753	CNOT3A-f001	CNOT3A-r001	Biotech-Plate-6:E09
CNOT3A-c006	L572	Q753	CNOT3A-f002	CNOT3A-r001	Biotech-Plate-6:E10
CNOT3A-c007	L607	Q753	CNOT3A-f003	CNOT3A-r001	Biotech-Plate-6:E11
CNOT3A-c008	H656	Q753	CNOT3A-f004	CNOT3A-r001	Biotech-Plate-6:E12

Figure 7.2 Expression/purification test of Biotech-Plate-6

Expression of constructs cloned as part of Biotech-Plate-6 plate was tested using 1 ml of *E. coli*. Protein was purified using nickel affinity chromatography and elutions were visualised by SDS-PAGE. Each lane is labelled with a well ID which corresponds to information provided in **Table 7.6**. Molecular weight markers are indicated for size comparison (kDa).

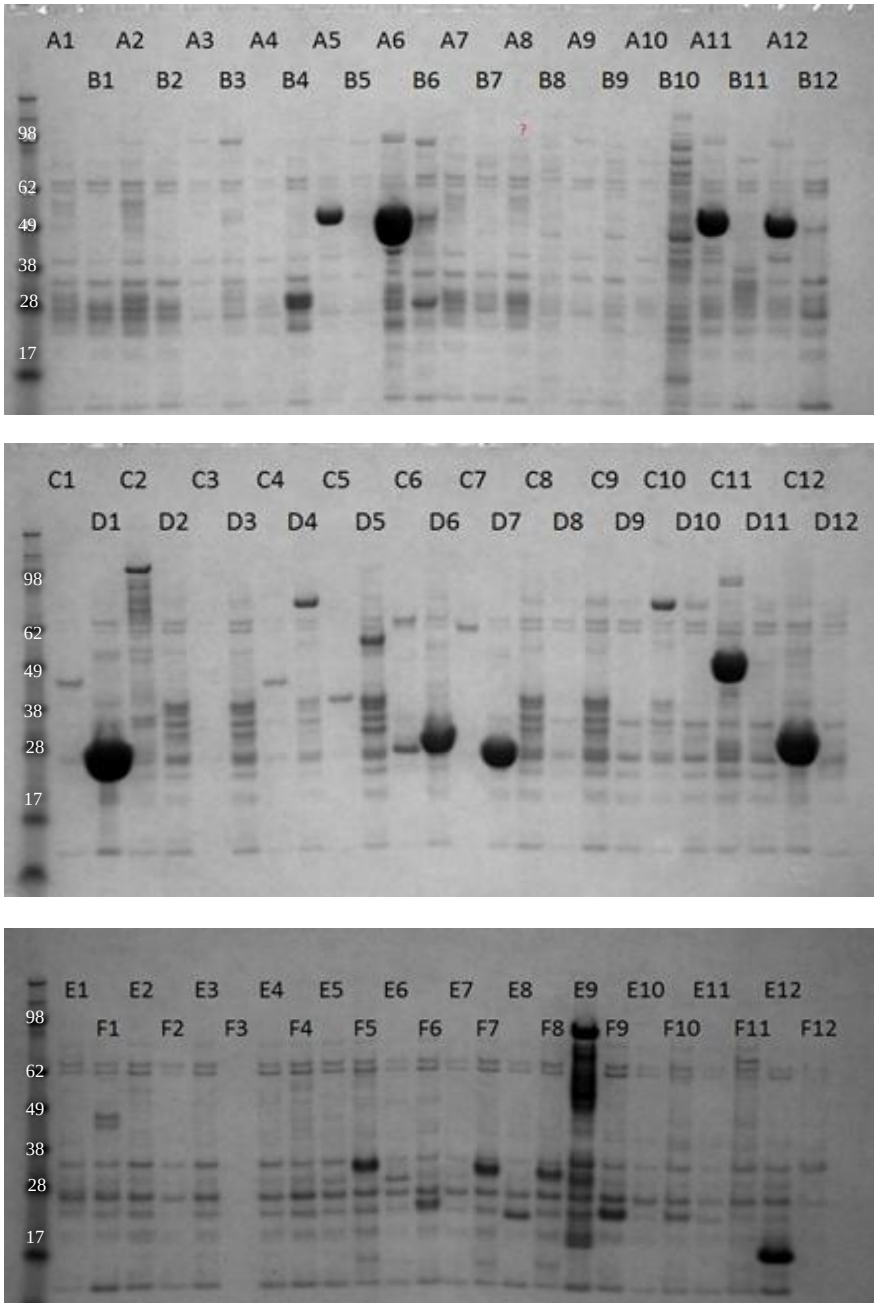


Table 7.7 Constructs cloned within Biotech-Plate-7

Constructs were cloned by LIC to facilitate the study of the CCR4-NOT complex and associated proteins. The construct ID, start and stop residues, forward and reverse primer IDs are listed for each construct. The forward and reverse primer IDs are listed with the primer sequence in **Tables 7.2 and 7.3**, respectively.

Construct ID	Start residue	Stop residue	Forward Primer	Reverse Primer	Construct Plate & Well
TTPA-c153	S14	E326	TTPA-f017	TTPA-r018	Biotech-Plate-7:G02
TTPA-c152	S98	E326	TTPA-f016	TTPA-r018	Biotech-Plate-7:G01
TTPA-c019	M1	E326	TTPA-f004	TTPA-r004	Biotech-Plate-7:A01
TTPA-c024	M1	Q308	TTPA-f004	TTPA-r014	Biotech-Plate-7:A02
TTPA-c025	M1	D171	TTPA-f004	TTPA-r016	Biotech-Plate-7:A03
TTPA-c026	S14	E326	TTPA-f013	TTPA-r004	Biotech-Plate-7:A04
TTPA-c027	S14	D171	TTPA-f013	TTPA-r016	Biotech-Plate-7:A05
TTPA-c028	S98	E326	TTPA-f014	TTPA-r004	Biotech-Plate-7:A06
TTPA-c029	S98	Q308	TTPA-f014	TTPA-r014	Biotech-Plate-7:A07
TTPA-c034	S98	S186	TTPA-f014	TTPA-r015	Biotech-Plate-7:A08
TTPA-c035	M1	E326	TTPA-f004	TTPA-r004	Biotech-Plate-7:A09
TTPA-c036	M1	E326	TTPA-f004	TTPA-r004	Biotech-Plate-7:A10
TTPA-c037	M1	I322	TTPA-f004	TTPA-r010	Biotech-Plate-7:A11
TTPA-c038	M1	E326	TTPA-f004	TTPA-r004	Biotech-Plate-7:A12
TTPA-c039	M1	E326	TTPA-f004	TTPA-r004	Biotech-Plate-7:B01
TTPA-c040	M1	E326	TTPA-f004	TTPA-r004	Biotech-Plate-7:B02
TTPA-c041	M1	E326	TTPA-f004	TTPA-r004	Biotech-Plate-7:B03
TTPA-c042	M1	I322	TTPA-f004	TTPA-r008	Biotech-Plate-7:B04
TTPA-c043	M1	E326	TTPA-f004	TTPA-r004	Biotech-Plate-7:B05
TTPA-c044	M1	Q308	TTPA-f004	TTPA-r014	Biotech-Plate-7:B06
TTPA-c045	M1	D171	TTPA-f004	TTPA-r016	Biotech-Plate-7:B07
TTPA-c046	S14	E326	TTPA-f013	TTPA-r004	Biotech-Plate-7:B08
TTPA-c047	S14	D171	TTPA-f013	TTPA-r016	Biotech-Plate-7:B09
TTPA-c048	S98	E326	TTPA-f014	TTPA-r004	Biotech-Plate-7:B10
TTPA-c049	S98	Q308	TTPA-f014	TTPA-r014	Biotech-Plate-7:B11
TTPA-c050	S98	S186	TTPA-f014	TTPA-r015	Biotech-Plate-7:B12
TTPA-c051	M1	E326	TTPA-f004	TTPA-r004	Biotech-Plate-7:C01
TTPA-c052	M1	E326	TTPA-f004	TTPA-r004	Biotech-Plate-7:C02
TTPA-c053	M1	I322	TTPA-f004	TTPA-r010	Biotech-Plate-7:C03
TTPA-c054	M1	E326	TTPA-f004	TTPA-r004	Biotech-Plate-7:C04
TTPA-c055	M1	E326	TTPA-f004	TTPA-r004	Biotech-Plate-7:C05
TTPA-c136	M1	L172	TTPA-f015	TTPA-r017	Biotech-Plate-7:D01
TTPA-c137	M1	E326	TTPA-f015	TTPA-r018	Biotech-Plate-7:D02
TTPA-c138	S98	L172	TTPA-f016	TTPA-r017	Biotech-Plate-7:D03
TTPA-c139	S98	E326	TTPA-f016	TTPA-r018	Biotech-Plate-7:D04
TTPA-c140	S14	L172	TTPA-f017	TTPA-r017	Biotech-Plate-7:D05
TTPA-c141	S14	E326	TTPA-f017	TTPA-r018	Biotech-Plate-7:D06

CNOT2A-c047	A201	F540	CNOT2A-f012	CNOT2A-r007	Biotech-Plate-7:D07
CNOT2A-c048	S391	F540	CNOT2A-f013	CNOT2A-r007	Biotech-Plate-7:D08
CNOT1B-c069	N800	K1004	CNOT1B-f019	CNOT1-r022	Biotech-Plate-7:F12
CNOT1B-c068	N800	R999	CNOT1B-f019	CNOT1-r021	Biotech-Plate-7:F11
TTPA-c056	M1	E326	TTPA-f004	TTPA-r004	Biotech-Plate-7:C06
TTPA-c057	M1	E326	TTPA-f004	TTPA-r004	Biotech-Plate-7:C07
TTPA-c058	M1	I322	TTPA-f004	TTPA-r008	Biotech-Plate-7:C08
TTPA-c059	M1	E326	TTPA-f004	TTPA-r007	Biotech-Plate-7:C09
TTPA-c151	E63	A174	TTPA-f020	TTPA-r021	Biotech-Plate-7:F10
TTPA-c150	E63	D171	TTPA-f020	TTPA-f020	Biotech-Plate-7:F09
TTPA-c149	E63	E326	TTPA-f020	TTPA-r019	Biotech-Plate-7:F08
TTPA-c148	S98	A174	TTPA-f019	TTPA-r021	Biotech-Plate-7:F07
TTPA-c147	S98	D171	TTPA-f019	TTPA-f020	Biotech-Plate-7:F06
TTPA-c146	S98	E326	TTPA-f019	TTPA-r019	Biotech-Plate-7:F05
TTPA-c145	M1	R103	TTPA-f018	TTP-r022	Biotech-Plate-7:F04
TTPA-c144	M1	A174	TTPA-f018	TTPA-r021	Biotech-Plate-7:F03
TTPA-c143	M1	D171	TTPA-f018	TTPA-f020	Biotech-Plate-7:F02
TTPA-c142	M1	E326	TTPA-f018	TTPA-r019	Biotech-Plate-7:F01
RQCD1C-c014	M1	Q299	RQCD1C-f004	RQCD1C-r004	Biotech-Plate-7:D10
RQCD1C-c015	D18	G286	RQCD1C-f005	RQCD1C-r005	Biotech-Plate-7:D11
RQCD1C-c016	D18	Q299	RQCD1C-f005	RQCD1C-r004	Biotech-Plate-7:D12
RQCD1C-c017	M1	G286	RQCD1C-f004	RQCD1C-r005	Biotech-Plate-7:E01

Figure 7.3 Expression/purification test of Biotech-Plate-7

Expression of constructs cloned as part of Biotech-Plate-7 plate was tested using 1 ml of *E. coli*. Protein was purified using nickel affinity chromatography and elutions were visualised by SDS-PAGE. Each lane is labelled with a well ID which corresponds to information provided in **Table 7.7**. Molecular weight markers are indicated for size comparison (kDa).

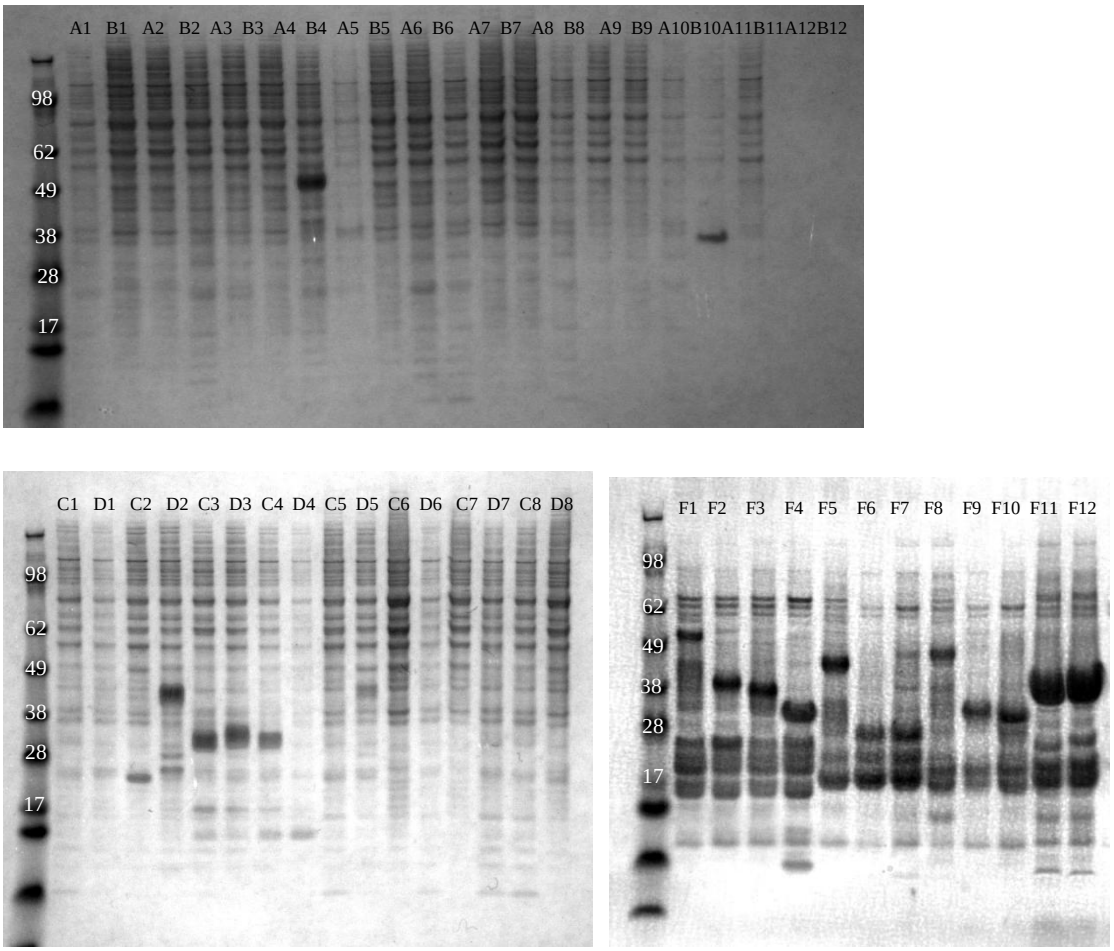


Table 7.8 Constructs cloned within Biotech-Plate-9

Constructs were cloned by LIC to facilitate the study of the CCR4-NOT complex and associated proteins. The construct ID, start and stop residues, forward and reverse primer IDs are listed for each construct. The forward and reverse primer IDs are listed with the primer sequence in **Tables 7.2 and 7.3**, respectively.

Construct ID	Start residue	Stop residue	Forward Primer	Reverse Primer	Construct Plate & Well
CNOT4B-c001	M1	A572	CNOT4B-f001	CNOT4B-r001	Biotech-Plate-9:A01
CNOT4B-c002	A101	Y198	CNOT4B-f002	CNOT4B-r003	Biotech-Plate-9:A02
CNOT4B-c003	K265	S554	CNOT4B-f003	CNOT4B-r002	Biotech-Plate-9:A03
CNOT6LA-c001	M1	R555	CNOT6LA-f001	CNOT6LA-r001	Biotech-Plate-9:A04
CNOT6LA-c002	M158	R555	CNOT6LA-f003	CNOT6LA-r001	Biotech-Plate-9:A05
CNOT6LA-c003	L169	L542	CNOT6LA-f004	CNOT6LA-r002	Biotech-Plate-9:A06
CNOT6LA-c004	L55	S527	CNOT6LA-f002	CNOT6LA-r003	Biotech-Plate-9:A07

CNOT6A-c001	M1	R557	CNOT6A-f001	CNOT6A-r001	Biotech-Plate-9:A08
CNOT6A-c002	L50	L545	CNOT6A-f002	CNOT6A-r002	Biotech-Plate-9:A09
CNOT6A-c003	L153	R557	CNOT6A-f003	CNOT6A-r001	Biotech-Plate-9:A10
CNOT7A-c001	M1	G235	CNOT7A-f001	CNOT7A-r001	Biotech-Plate-9:A11
CNOT7A-c002	R11	G235	CNOT7A-f002	CNOT7A-r001	Biotech-Plate-9:A12
CNOT8A-c001	M1	Q292	CNOT8A-f001	CNOT8A-r001	Biotech-Plate-9:B01
CNOT8A-c002	A17	E247	CNOT8A-f003	CNOT8A-r004	Biotech-Plate-9:B02
CNOT8A-c003	M1	K280	CNOT8A-f001	CNOT8A-r002	Biotech-Plate-9:B03
CNOT8A-c004	Q10	D273	CNOT8A-f002	CNOT8A-r003	Biotech-Plate-9:B04
CNOT10A-c001	M1	M708	CNOT10A-f001	CNOT10A-r001	Biotech-Plate-9:B12
CNOT10A-c002	G288	L681	CNOT10A-f004	CNOT10A-r002	Biotech-Plate-9:C01
CNOT10A-c003	N66	F143	CNOT10A-f003	CNOT10A-r003	Biotech-Plate-9:C02
CNOT10A-c004	Q9	D102	CNOT10A-f002	CNOT10A-r004	Biotech-Plate-9:C03
CNOT11A-c001	M1	T506	CNOT11A-f001	CNOT11A-r001	Biotech-Plate-9:C04
CNOT11A-c002	M61	T506	CNOT11A-f003	CNOT11A-r001	Biotech-Plate-9:C05
CNOT11A-c003	K319	L497	CNOT11A-f005	CNOT11A-r002	Biotech-Plate-9:C06
CNOT11A-c004	T15	M316	CNOT11A-f002	CNOT11A-r003	Biotech-Plate-9:C07
CNOT1A-c001	Q820	R994	CNOT1A-f003	CNOT1A-r006	Biotech-Plate-9:C08
CNOT1A-c002	M1	L240	CNOT1A-f001	CNOT1A-r009	Biotech-Plate-9:C09
CNOT1A-c003	M1	N800	CNOT1A-f001	CNOT1A-r008	Biotech-Plate-9:C10
CNOT1A-c004	M1	R994	CNOT1A-f001	CNOT1A-r006	Biotech-Plate-9:C11
CNOT1B-c049	Q820	R999	CNOT1B-f013	CNOT1B-r016	Biotech-Plate-9:C12
CNOT1B-c050	M1	N800	CNOT1B-f008	CNOT1B-r019	Biotech-Plate-9:D01
CNOT1B-c051	M1	R999	CNOT1B-f008	CNOT1B-r016	Biotech-Plate-9:D02
CNOT4B-c004	M1	A572	CNOT4B-f001	CNOT4B-r001	Biotech-Plate-9:D03
CNOT4B-c005	K265	S554	CNOT4B-f003	CNOT4B-r002	Biotech-Plate-9:D04
CNOT6LA-c005	M1	R555	CNOT6LA-f001	CNOT6LA-r001	Biotech-Plate-9:D05
CNOT6LA-c006	M158	R555	CNOT6LA-f003	CNOT6LA-r001	Biotech-Plate-9:D06
CNOT6LA-c007	L169	L542	CNOT6LA-f004	CNOT6LA-r002	Biotech-Plate-9:D07
CNOT6LA-c008	L55	S527	CNOT6LA-f002	CNOT6LA-r003	Biotech-Plate-9:D08
CNOT6A-c004	M1	R557	CNOT6A-f001	CNOT6A-r001	Biotech-Plate-9:D09
CNOT6A-c005	L50	L545	CNOT6A-f002	CNOT6A-r002	Biotech-Plate-9:D10
CNOT6A-c006	L153	R557	CNOT6A-f003	CNOT6A-r001	Biotech-Plate-9:D11
CNOT7A-c003	M1	G235	CNOT7A-f001	CNOT7A-r001	Biotech-Plate-9:D12
CNOT7A-c004	R11	G235	CNOT7A-f002	CNOT7A-r001	Biotech-Plate-9:E01
CNOT8A-c005	M1	Q292	CNOT8A-f001	CNOT8A-r001	Biotech-Plate-9:E02
CNOT8A-c006	A17	E247	CNOT8A-f003	CNOT8A-r004	Biotech-Plate-9:E03
CNOT8A-c007	M1	K280	CNOT8A-f001	CNOT8A-r002	Biotech-Plate-9:E04
CNOT8A-c008	Q10	D273	CNOT8A-f002	CNOT8A-r003	Biotech-Plate-9:E05
CNOT10A-c005	M1	M708	CNOT10A-f001	CNOT10A-r001	Biotech-Plate-9:F01
CNOT10A-c006	G288	L681	CNOT10A-f004	CNOT10A-r002	Biotech-Plate-9:F02
CNOT10A-c007	N66	F143	CNOT10A-f003	CNOT10A-r003	Biotech-Plate-9:F03
CNOT10A-c008	Q9	D102	CNOT10A-f002	CNOT10A-r004	Biotech-Plate-9:F04
CNOT11A-c005	M1	T506	CNOT11A-f001	CNOT11A-r001	Biotech-Plate-9:F05
CNOT11A-c006	M61	T506	CNOT11A-f003	CNOT11A-r001	Biotech-Plate-9:F06
CNOT11A-c007	V260	L497	CNOT11A-f004	CNOT11A-r002	Biotech-Plate-9:F07

CNOT11A-c008	T15	M316	CNOT11A-f002	CNOT11A-r003	Biotech-Plate-9:F08
CNOT1A-c005	M1	N800	CNOT1A-f001	CNOT1A-r008	Biotech-Plate-9:F09
CNOT1A-c006	M1	R994	CNOT1A-f001	CNOT1A-r006	Biotech-Plate-9:F10
CNOT1B-c054	Q820	R999	CNOT1B-f013	CNOT1B-r016	Biotech-Plate-9:F11
CNOT1B-c055	M1	N800	CNOT1B-f008	CNOT1B-r019	Biotech-Plate-9:F12
CNOT1B-c056	M1	R999	CNOT1B-f008	CNOT1B-r016	Biotech-Plate-9:G01
CNOT1B-c057	M1	S2376	CNOT1B-f008	CNOT1B-r007	Biotech-Plate-9:G02
CNOT1B-c058	E1090	S2376	CNOT1B-f014	CNOT1B-r007	Biotech-Plate-9:G03
CNOT1B-c059	M1	E1090	CNOT1B-f008	CNOT1B-r014	Biotech-Plate-9:G04
CNOT1B-c060	M1	Y1842	CNOT1B-f008	CNOT1B-r011	Biotech-Plate-9:G05
CNOT1B-c061	D1565	S2376	CNOT1B-f016	CNOT1B-r007	Biotech-Plate-9:G06
CNOT1B-c062	Q820	Y1842	CNOT1B-f013	CNOT1B-r011	Biotech-Plate-9:G07
CNOT1B-c063	A1602	E2371	CNOT1B-f017	CNOT1B-r009	Biotech-Plate-9:G08
CNOT1B-c064	N800	E2371	CNOT1B-f010	CNOT1B-r009	Biotech-Plate-9:G09
CNOT1B-c065	Q820	E2371	CNOT1B-f013	CNOT1B-r009	Biotech-Plate-9:G10
CNOT1B-c066	N800	Y1842	CNOT1B-f010	CNOT1B-r011	Biotech-Plate-9:G11
CNOT1B-c067	Q1356	E2371	CNOT1B-f015	CNOT1B-r009	Biotech-Plate-9:G12
CNOT1A-c007	Q820	S2371	CNOT1A-f003	CNOT1A-r001	Biotech-Plate-9:H01
CNOT1A-c008	M1	E1085	CNOT1A-f001	CNOT1A-r005	Biotech-Plate-9:H02
CNOT1A-c009	Q820	C1495	CNOT1A-f003	CNOT1A-r004	Biotech-Plate-9:H03
CNOT1A-c010	M1	C1495	CNOT1A-f001	CNOT1A-r004	Biotech-Plate-9:H04
CNOT1A-c011	M1	Q820	CNOT1A-f001	CNOT1A-r007	Biotech-Plate-9:H05
CNOT1A-c012	Q820	Y1837	CNOT1A-f003	CNOT1A-r003	Biotech-Plate-9:H06
CNOT1A-c013	N800	E2366	CNOT1A-f002	CNOT1A-r002	Biotech-Plate-9:H07
CNOT1A-c014	Q820	E2366	CNOT1A-f003	CNOT1A-r002	Biotech-Plate-9:H08
CNOT1A-c015	N800	Y1837	CNOT1A-f002	CNOT1A-r003	Biotech-Plate-9:H09

Figure 7.4 Expression/purification test of Biotech-Plate-9

Expression of constructs cloned as part of Biotech-Plate-9 plate was tested using 1 ml of *E. coli*. Protein was purified using nickel affinity chromatography and elutions were visualised by SDS-PAGE. Each lane is labelled with a well ID which corresponds to information provided in **Table 7.8**. Molecular weight markers are indicated for size comparison (kDa).

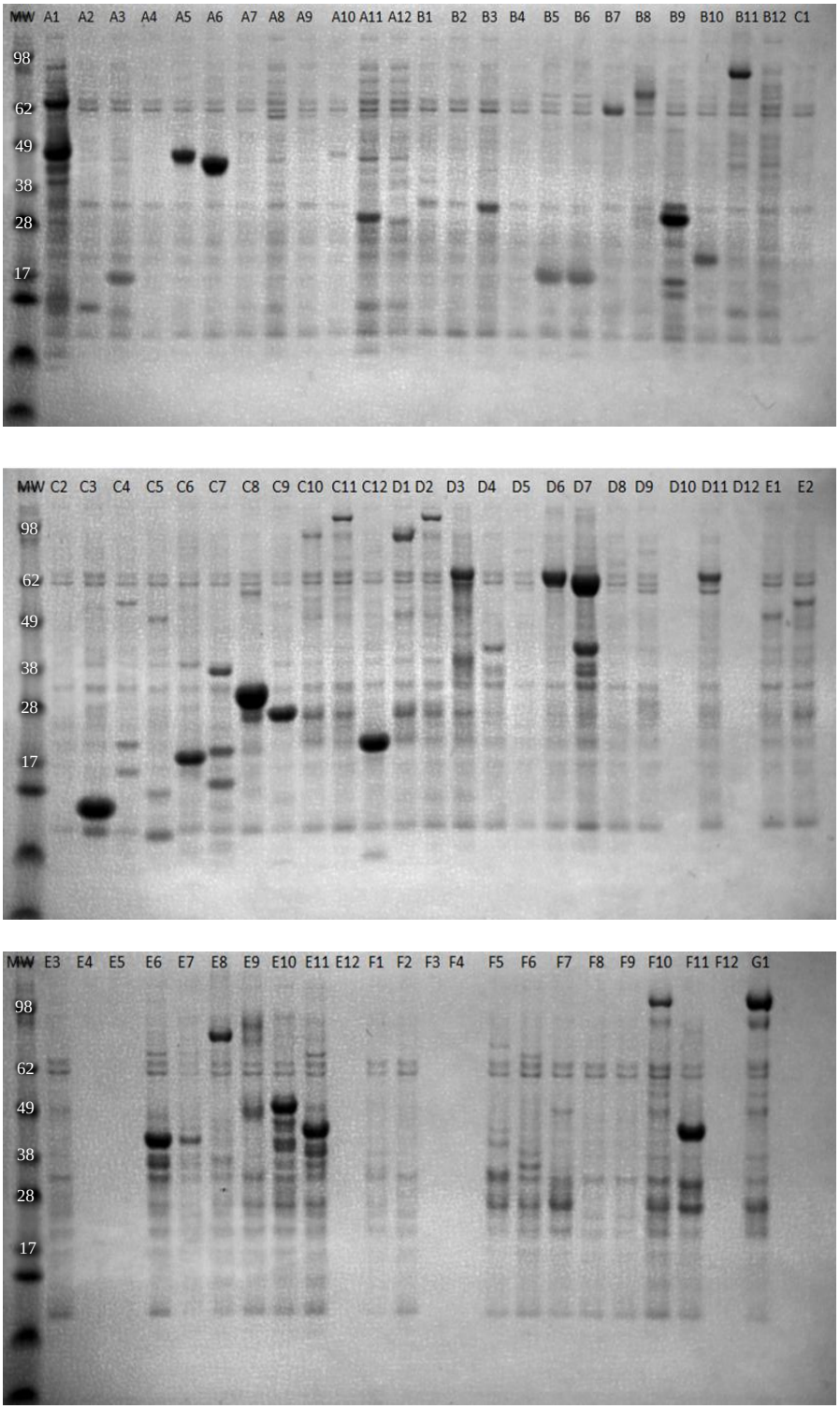


Table 7.9 Constructs cloned within Biotech-Plate-10

Constructs were cloned by LIC to facilitate the study of the CCR4-NOT complex and associated proteins. The construct ID, start and stop residues, forward and reverse primer IDs are listed for each construct. The forward and reverse primer IDs are listed with the primer sequence in **Tables 7.2 and 7.3**, respectively.

Construct ID	Start residue	Stop residue	Forward Primer	Reverse Primer	Construct Plate & Well
TTPA-c063	M1	D171	TTPA-f006	TTPA-r011	Biotech-Plate-10:A01
TTPA-c064	M1	D171	TTPA-f006	TTPA-r011	Biotech-Plate-10:A02
TTPA-c065	M1	S186	TTPA-f006	TTPA-r009	Biotech-Plate-10:A03
TTPA-c066	M1	S186	TTPA-f006	TTPA-r009	Biotech-Plate-10:A04
TTPA-c067	M1	S186	TTPA-f006	TTPA-r009	Biotech-Plate-10:A05
TTPA-c068	S14	S186	TTPA-f007	TTPA-r009	Biotech-Plate-10:A06
TTPA-c069	S14	S186	TTPA-f007	TTPA-r009	Biotech-Plate-10:A07
TTPA-c070	S14	S186	TTPA-f007	TTPA-r009	Biotech-Plate-10:A08
TTPA-c071	S98	E326	TTPA-f008	TTPA-r006	Biotech-Plate-10:A09
TTPA-c072	S98	E326	TTPA-f008	TTPA-r006	Biotech-Plate-10:A10

Figure 7.5 Expression/purification test of Biotech-Plate-10

Expression of constructs cloned as part of Biotech-Plate-10 plate was tested using 1 ml of *E. coli*. Protein was purified using nickel affinity chromatography and elutions were visualised by SDS-PAGE. Each lane is labelled with a well ID which corresponds to information provided in **Table 7.9**. Molecular weight markers are indicated for size comparison (kDa).

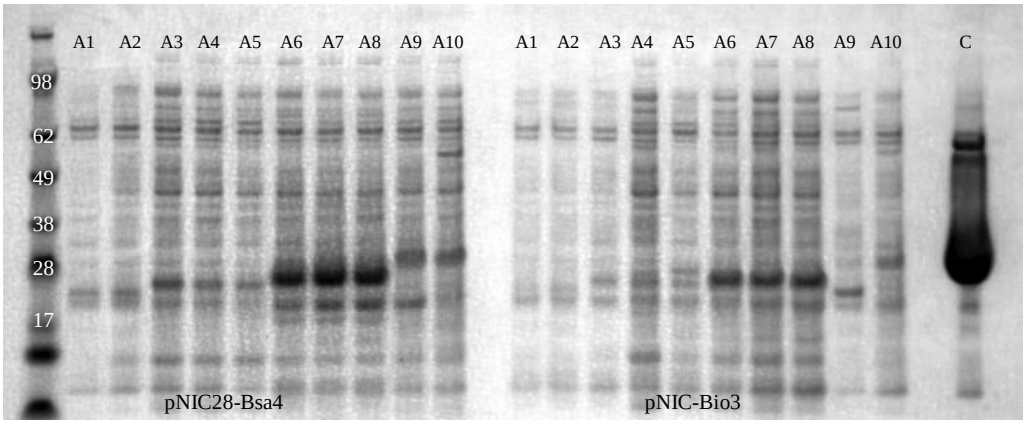


Table 7.10 Constructs cloned within OG-70-var

Constructs were cloned by LIC to facilitate the study of the CCR4-NOT complex and associated proteins. The construct ID, start and stop residues, forward and reverse primer IDs are listed for each construct. The forward and reverse primer IDs are listed with the primer sequence in **Tables 7.2 and 7.3**, respectively.

Construct ID	Start residue	Stop residue	Forward Primer	Reverse Primer	Construct Plate & Well
TTPA-c000	M1	E326	TTPA-f000	TTPA-r000	OG-70-var:A01
TTPA-c001	M1	Q308	TTPA-f000	TTPA-r001	OG-70-var:A02
TTPA-c002	M1	S186	TTPA-f000	TTPA-r002	OG-70-var:A03
TTPA-c003	M1	D171	TTPA-f000	TTPA-r003	OG-70-var:A04
TTPA-c010	S14	E326	TTPA-f001	TTPA-r000	OG-70-var:A05
TTPA-c011	S14	Q308	TTPA-f001	TTPA-r001	OG-70-var:A06
TTPA-c012	S14	S186	TTPA-f001	TTPA-r002	OG-70-var:A07
TTPA-c013	S14	D171	TTPA-f001	TTPA-r003	OG-70-var:A08
TTPA-c020	S28	E326	TTPA-f002	TTPA-r000	OG-70-var:A09
TTPA-c021	S28	Q308	TTPA-f002	TTPA-r001	OG-70-var:A10
TTPA-c022	S28	S186	TTPA-f002	TTPA-r002	OG-70-var:A11
TTPA-c023	S28	D171	TTPA-f002	TTPA-r003	OG-70-var:A12
TTPA-c030	S98	E326	TTPA-f003	TTPA-r000	OG-70-var:B01
TTPA-c031	S98	Q308	TTPA-f003	TTPA-r001	OG-70-var:B02
TTPA-c032	S98	S186	TTPA-f003	TTPA-r002	OG-70-var:B03
TTPA-c033	S98	D171	TTPA-f003	TTPA-r003	OG-70-var:B04
TTPA-c100	M1	E326	TTPA-f000	TTPA-r000	OG-70-var:F08
TTPA-c102	M1	S186	TTPA-f000	TTPA-r002	OG-70-var:F09
TTPA-c111	S14	Q308	TTPA-f001	TTPA-r001	OG-70-var:F10
TTPA-c113	S14	D171	TTPA-f001	TTPA-r003	OG-70-var:F11
TTPA-c120	S28	E326	TTPA-f002	TTPA-r000	OG-70-var:F12
TTPA-c122	S28	S186	TTPA-f002	TTPA-r002	OG-70-var:G01
TTPA-c130	S98	E326	TTPA-f003	TTPA-r000	OG-70-var:G02
TTPA-c132	S98	S186	TTPA-f003	TTPA-r002	OG-70-var:G03

Figure 7.6 Expression/purification test of OG-70-var

Expression of constructs cloned as part of OG-70-var plate was tested using 1 ml of *E. coli*. Protein was purified using nickel affinity chromatography and elutions were visualised by SDS-PAGE. Each lane is labelled with a well ID which corresponds to information provided in **Table 7.10**. Molecular weight markers are indicated for size comparison (kDa).

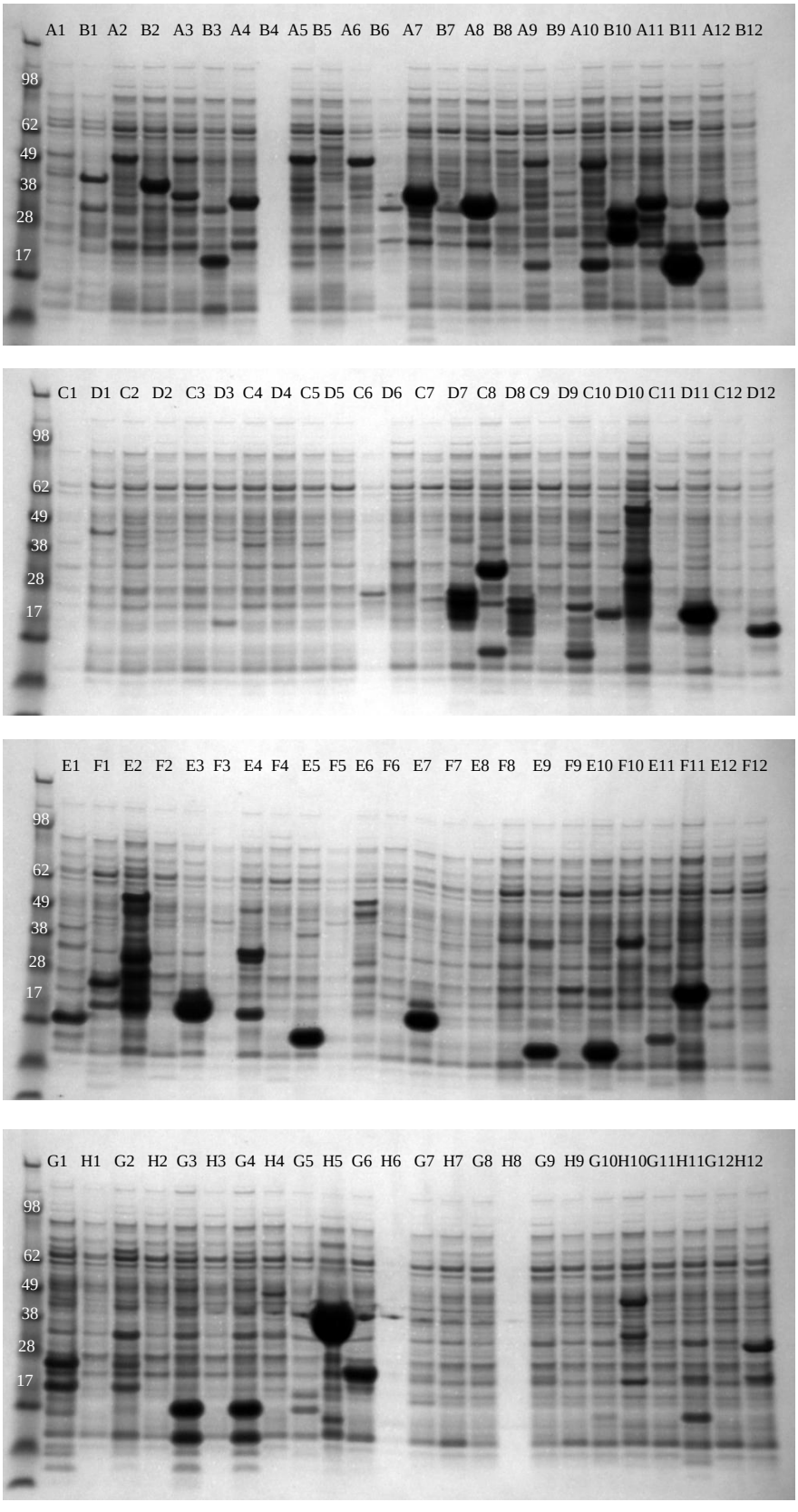


Table 7.11 Constructs cloned within WY-24-var

Constructs were cloned by LIC to facilitate the study of the CCR4-NOT complex and associated proteins. The construct ID, start and stop residues, forward and reverse primer IDs are listed for each construct. The forward and reverse primer IDs are listed with the primer sequence in **Tables 7.2 and 7.3**, respectively.

Construct ID	Start residue	Stop residue	Forward Primer	Reverse Primer	Construct Plate & Well
RQCD1C-c026	D18	E285	RQCD1C-f008	RQCD1C-r010	WY-24-var:E07
RQCD1C-c027	R19	E285	RQCD1C-f009	RQCD1C-r010	WY-24-var:E08
CNOT2A-c058	T344	F540	CNOT2A-f017	CNOT2A-r010	WY-24-var:H03
CNOT2A-c059	M350	F540	CNOT2A-f018	CNOT2A-r010	WY-24-var:H04
CNOT2A-c060	D429	F540	CNOT2A-f019	CNOT2A-r010	WY-24-var:H05
CNOT2A-c061	E441	F540	CNOT2A-f020	CNOT2A-r010	WY-24-var:H06

Figure 7.7 Expression/purification test of WY-24-var

Expression of constructs cloned as part of WY-24-var plate was tested using 1 ml of *E. coli*. Protein was purified using nickel affinity chromatography and elutions were visualised by SDS-PAGE. Each lane is labelled with a well ID which corresponds to information provided in **Table 7.11**. Molecular weight markers are indicated for size comparison (kDa).

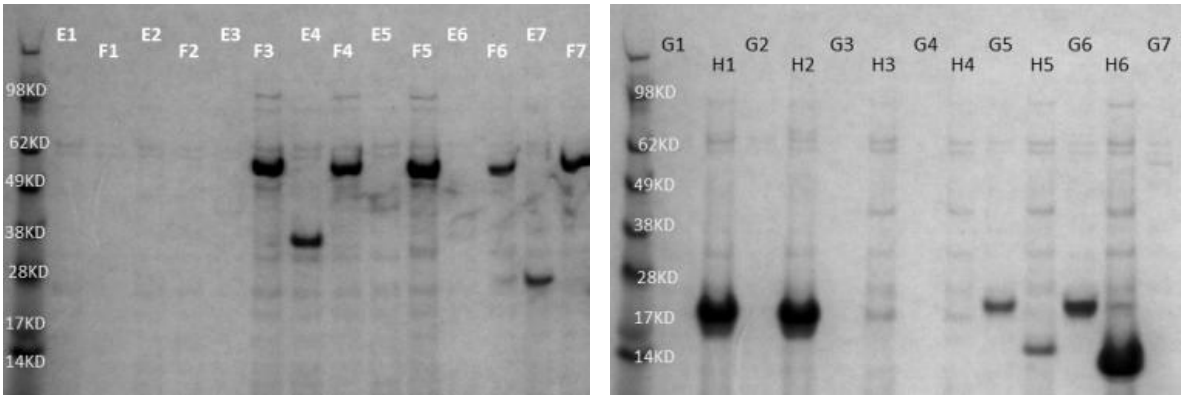


Table 7.12 Constructs cloned within WY-29-baculo-mam

Constructs were cloned by LIC to facilitate the study of the CCR4-NOT complex and associated proteins. The construct ID, start and stop residues, forward and reverse primer IDs are listed for each construct. The forward and reverse primer IDs are listed with the primer sequence in **Tables 7.2 and 7.3**, respectively.

Construct ID	Start residue	Stop residue	Forward Primer	Reverse Primer	Construct Plate & Well
CNOT11A-c009	M1	T506	CNOT11A-f001	CNOT11A-r005	WY-29-baculo-mam:E04
CNOT10A-c009	M1	M708	CNOT10A-f001	CNOT10A-r006	WY-29-baculo-mam:E03

Figure 7.12 Expression/purification test of WY-29-baculo-mam

Expression of constructs cloned as part of WY-29-baculo-mam plate was tested using 1 ml of *E. coli*. Protein was purified using nickel affinity chromatography and elutions were visualised by SDS-PAGE. Each lane is labelled with a well ID which corresponds to information provided in **Table 7.2**. Molecular weight markers are indicated for size comparison (kDa).

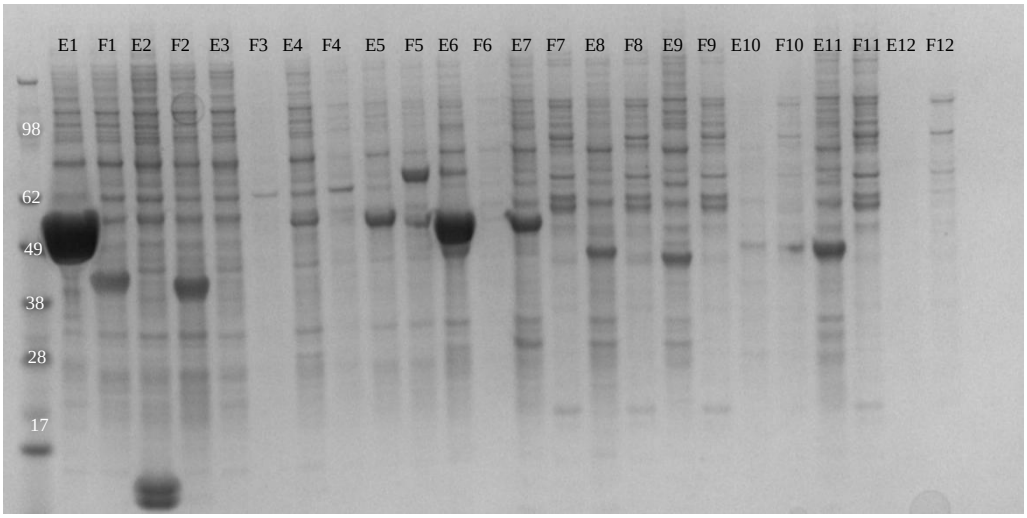


Table 7.13 Constructs cloned within WY-30-bacteria

Constructs were cloned by LIC to facilitate the study of the CCR4-NOT complex and associated proteins. The construct ID, start and stop residues, forward and reverse primer IDs are listed for each construct. The forward and reverse primer IDs are listed with the primer sequence in **Tables 7.2 and 7.3**, respectively.

Construct ID	Start residue	Stop residue	Forward Primer	Reverse Primer	Construct Plate & Well
CNOT11A-c010	M1	T506	CNOT11A-f001	CNOT11A-r006	WY-30-bacteria:C05
CNOT11A-c011	M61	T506	CNOT11A-f003	CNOT11A-r006	WY-30-bacteria:C06
CNOT11A-c012	M1	T506	CNOT11A-f006	CNOT11A-r006	WY-30-bacteria:C07
CNOT11A-c013	F284	T506	CNOT11A-f007	CNOT11A-r006	WY-30-bacteria:C08
CNOT11A-c014	M61	L497	CNOT11A-f003	CNOT11A-r002	WY-30-bacteria:C09
CNOT11A-c015	M1	L497	CNOT11A-f006	CNOT11A-r002	WY-30-bacteria:C10
CNOT11A-c016	F284	L497	CNOT11A-f007	CNOT11A-r002	WY-30-bacteria:C11
CNOT10C-c001	M1	L708	CNOT10A-f001	CNOT10A-r002	WY-30-bacteria:E01
CNOT10C-c002	M1	F143	CNOT10A-f001	CNOT10A-r003	WY-30-bacteria:E02
CNOT10C-c003	M1	D102	CNOT10A-f001	CNOT10A-r004	WY-30-bacteria:E03
CNOT10C-c004	M1	K744	CNOT10A-f001	CNOT10A-r001	WY-30-bacteria:E04
CNOT10C-c005	N66	K744	CNOT10A-f003	CNOT10A-r001	WY-30-bacteria:E05
CNOT10C-c006	G288	K744	CNOT10A-f004	CNOT10A-r001	WY-30-bacteria:E06
CNOT1B-c073	M1	A333	CNOT1B-f008	CNOT1B-r026	WY-30-bacteria:E08
CNOT1B-c074	M1	L240	CNOT1B-f008	CNOT1B-r021	WY-30-bacteria:E09
CNOT11A-c101	M1	T506	CNOT11A-f101	CNOT11A-r101	WY-30-bacteria:H08

CNOT10A-c101	M1	M708	CNOT10A-f101	CNOT10A-r101	WY-30-bacteria:H08
CNOT10C-c101	M1	K744	CNOT10C-f101	CNOT10C-r101	WY-30-bacteria:H09
CNOT10A-c102	M1	M708	CNOT10A-f102	CNOT10A-r102	WY-30-bacteria:H10
CNOT11A-c102	M1	T506	CNOT11A-f102	CNOT11A-r102	WY-30-bacteria:H10
CNOT10C-c102	M1	K744	CNOT10C-f102	CNOT10C-r102	WY-30-bacteria:H11

Figure 7.9 Expression/purification test of WY-30-bacteria

Expression of constructs cloned as part of WY-30-bacteria plate was tested using 1 ml of *E. coli*. Protein was purified using nickel affinity chromatography and elutions were visualised by SDS-PAGE. Each lane is labelled with a well ID which corresponds to information provided in **Table 7.13**. Molecular weight markers are indicated for size comparison (kDa).

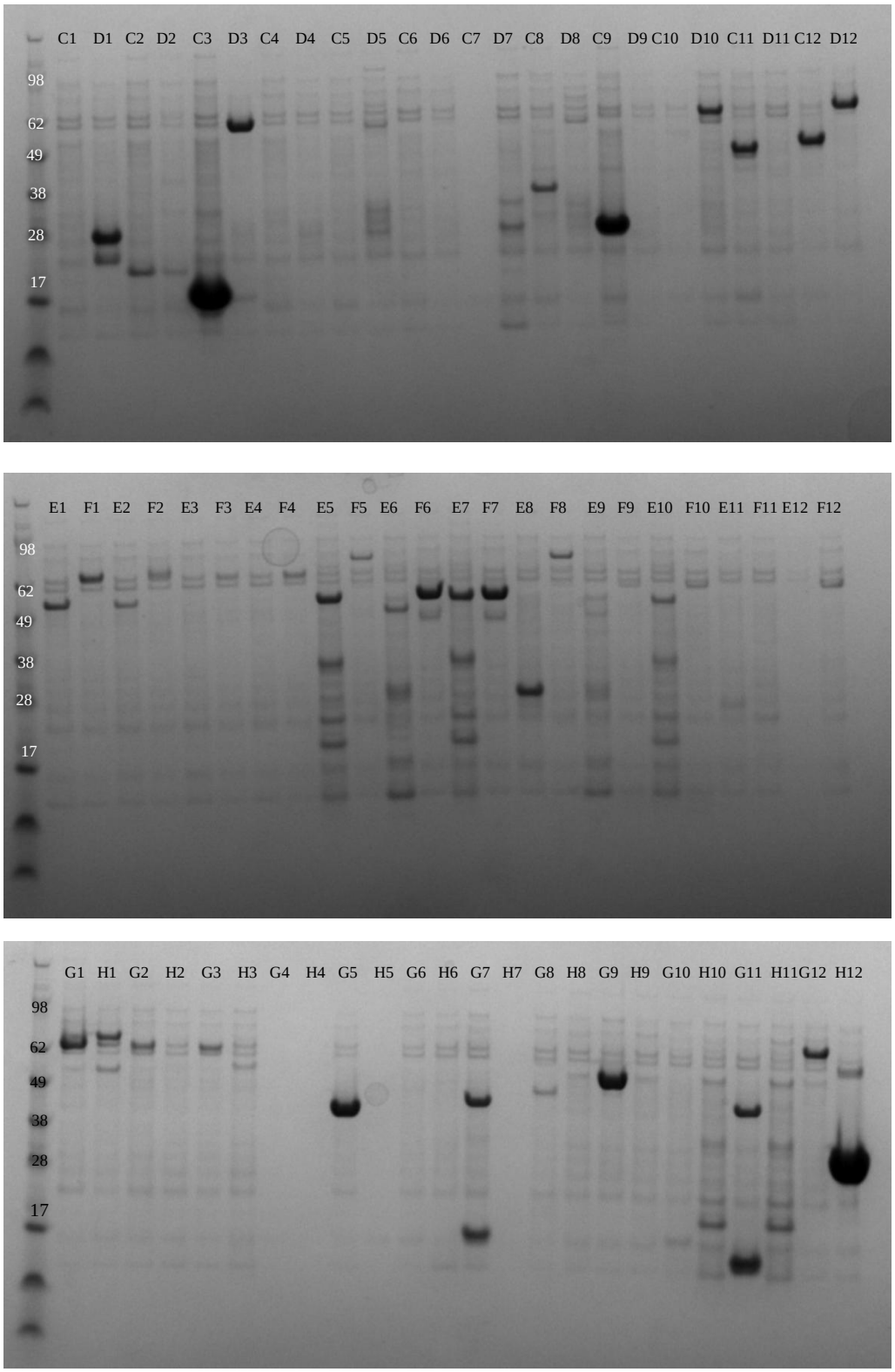


Table 7.14 Constructs cloned within WY-31-bacteria

Constructs were cloned by LIC to facilitate the study of the CCR4-NOT complex and associated proteins. The construct ID, start and stop residues, forward and reverse primer IDs are listed for each construct. The forward and reverse primer IDs are listed with the primer sequence in **Tables 7.2 and 7.3**, respectively.

Construct ID	Start residue	Stop residue	Forward Primer	Reverse Primer	Construct Plate & Well
CNOT10A-c111	K58	R466	CNOT10A-f011	CNOT10A-r111	WY-31-bacteria:F08
CNOT11A-c111	A238	T506	CNOT11A-f011	CNOT11A-r111	WY-31-bacteria:F09
CNOT10A-c113	N33	V444	CNOT10A-f012	CNOT10A-r112	WY-31-bacteria:F10
CNOT11A-c113	Q214	E486	CNOT11A-f012	CNOT11A-r112	WY-31-bacteria:F11

Figure 7.10 Expression/purification test of WY-31-bacteria

Expression of constructs cloned as part of WY-31-bacteria plate was tested using 1 ml of *E. coli*. Protein was purified using nickel affinity chromatography and elutions were visualised by SDS-PAGE. Each lane is labelled with a well ID which corresponds to information provided in **Table 7.14**. Molecular weight markers are indicated for size comparison (kDa).

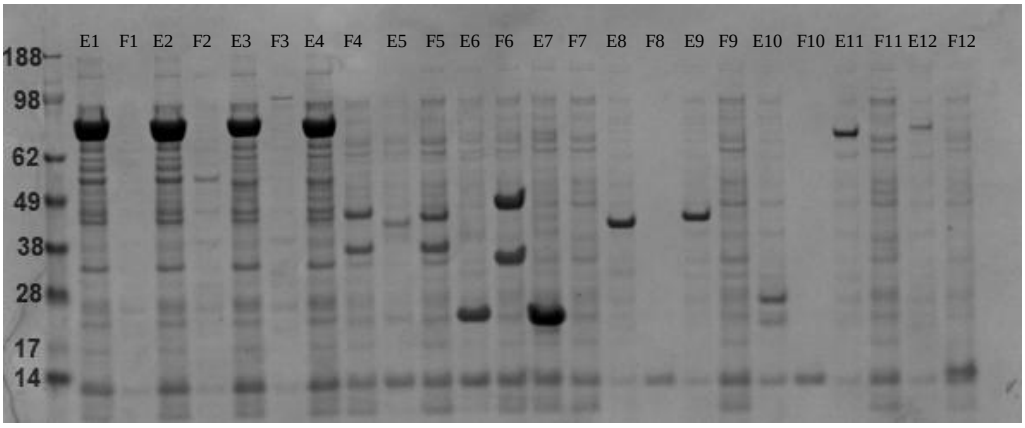


Table 7.15 TTP peptides printed on each spot of the peptide array

List of 109 peptide sequences and number/letter locator for each15-residue TTP peptide printed for the TTP peptide array. The starting residue of each peptide was staggered by 3 residues with the previous peptide.

Peptide position	Peptide Sequence
A4	M-D-L-T-A-I-Y-E-S-L-L-S-L-S-P
A5	T-A-I-Y-E-S-L-L-S-L-S-P-D-V-P
A6	Y-E-S-L-L-S-L-S-P-D-V-P-V-P-S
A7	L-L-S-L-S-P-D-V-P-V-P-S-D-H-G

A8	L-S-P-D-V-P-V-P-S-D-H-G-G-T-E
A9	D-V-P-V-P-S-D-H-G-G-T-E-S-S-P
A10	V-P-S-D-H-G-G-T-E-S-S-P-G-W-G
A11	D-H-G-G-T-E-S-S-P-G-W-G-S-S-G
B1	G-T-E-S-S-P-G-W-G-S-S-G-P-W-S
B2	S-S-P-G-W-G-S-S-G-P-W-S-L-S-P
B3	G-W-G-S-S-G-P-W-S-L-S-P-S-D-S
B4	S-S-G-P-W-S-L-S-P-S-D-S-S-P-S
B5	P-W-S-L-S-P-S-D-S-S-P-S-G-V-T
B6	L-S-P-S-D-S-S-P-S-G-V-T-S-R-L
B7	S-D-S-S-P-S-G-V-T-S-R-L-P-G-R
B8	S-P-S-G-V-T-S-R-L-P-G-R-S-T-S
B9	G-V-T-S-R-L-P-G-R-S-T-S-L-V-E
B10	S-R-L-P-G-R-S-T-S-L-V-E-G-R-S
B11	P-G-R-S-T-S-L-V-E-G-R-S-S-G-W
C1	S-T-S-L-V-E-G-R-S-S-G-W-V-P-P
C2	L-V-E-G-R-S-S-G-W-V-P-P-P-P-G
C3	G-R-S-S-G-W-V-P-P-P-P-G-F-A-P
C4	S-G-W-V-P-P-P-P-G-F-A-P-L-A-P
C5	V-P-P-P-P-G-F-A-P-L-A-P-R-L-G
C6	P-P-G-F-A-P-L-A-P-R-L-G-P-E-L
C7	F-A-P-L-A-P-R-L-G-P-E-L-S-P-S
C8	L-A-P-R-L-G-P-E-L-S-P-S-P-T-S
C9	R-L-G-P-E-L-S-P-S-P-T-S-P-T-A
C10	P-E-L-S-P-S-P-T-S-P-T-A-T-S-T
C11	S-P-S-P-T-S-P-T-A-T-S-T-T-P-S
D1	P-T-S-P-T-A-T-S-T-T-P-S-R-Y-K
D2	P-T-A-T-S-T-T-P-S-R-Y-K-T-E-L
D3	T-S-T-T-P-S-R-Y-K-T-E-L-S-R-T
D4	T-P-S-R-Y-K-T-E-L-S-R-T-F-S-E
D5	R-Y-K-T-E-L-S-R-T-F-S-E-S-G-R
D6	T-E-L-S-R-T-F-S-E-S-G-R-S-R-Y
D7	S-R-T-F-S-E-S-G-R-S-R-Y-G-A-K
D8	F-S-E-S-G-R-S-R-Y-G-A-K-S-Q-F
D9	S-G-R-S-R-Y-G-A-K-S-Q-F-A-H-G
D10	S-R-Y-G-A-K-S-Q-F-A-H-G-L-G-E
D11	G-A-K-S-Q-F-A-H-G-L-G-E-L-R-Q
E1	S-Q-F-A-H-G-L-G-E-L-R-Q-A-N-R
E2	A-H-G-L-G-E-L-R-Q-A-N-R-H-P-K
E3	L-G-E-L-R-Q-A-N-R-H-P-K-Y-K-T
E4	L-R-Q-A-N-R-H-P-K-Y-K-T-E-L-S
E5	A-N-R-H-P-K-Y-K-T-E-L-S-H-K-F
E6	H-P-K-Y-K-T-E-L-S-H-K-F-Y-L-Q
E7	Y-K-T-E-L-S-H-K-F-Y-L-Q-G-R-S
E8	E-L-S-H-K-F-Y-L-Q-G-R-S-P-Y-G
E9	H-K-F-Y-L-Q-G-R-S-P-Y-G-S-R-S

E10	Y-L-Q-G-R-S-P-Y-G-S-R-S-H-F-I
E11	G-R-S-P-Y-G-S-R-S-H-F-I-H-N-P
F1	P-Y-G-S-R-S-H-F-I-H-N-P-S-E-D
F2	S-R-S-H-F-I-H-N-P-S-E-D-L-A-A
F3	H-F-I-H-N-P-S-E-D-L-A-A-P-G-H
F4	H-N-P-S-E-D-L-A-A-P-G-H-P-P-V
F5	S-E-D-L-A-A-P-G-H-P-P-V-L-R-Q
F6	L-A-A-P-G-H-P-P-V-L-R-Q-S-I-S
F7	P-G-H-P-P-V-L-R-Q-S-I-S-F-S-G
F8	P-P-V-L-R-Q-S-I-S-F-S-G-L-P-S
F9	L-R-Q-S-I-S-F-S-G-L-P-S-G-R-R
F10	S-I-S-F-S-G-L-P-S-G-R-R-T-S-P
F11	F-S-G-L-P-S-G-R-R-T-S-P-P-P-P
G1	L-P-S-G-R-R-T-S-P-P-P-P-G-L-A
G2	G-R-R-T-S-P-P-P-P-G-L-A-G-P-S
G3	T-S-P-P-P-P-G-L-A-G-P-S-L-S-S
G4	P-P-P-G-L-A-G-P-S-L-S-S-S-S-F
G5	G-L-A-G-P-S-L-S-S-S-S-F-S-P-S
G6	G-P-S-L-S-S-S-S-F-S-P-S-S-S-P
G7	L-S-S-S-S-F-S-P-S-S-S-P-P-P-P
G8	S-S-F-S-P-S-S-S-P-P-P-P-G-D-L
G9	S-P-S-S-S-P-P-P-P-G-D-L-P-L-S
G10	S-S-P-P-P-P-G-D-L-P-L-S-P-S-A
G11	P-P-P-G-D-L-P-L-S-P-S-A-F-S-A
H1	G-D-L-P-L-S-P-S-A-F-S-A-A-P-G
H2	P-L-S-P-S-A-F-S-A-A-P-G-T-P-L
H3	P-S-A-F-S-A-A-P-G-T-P-L-A-R-R
H4	F-S-A-A-P-G-T-P-L-A-R-R-D-P-T
H5	A-P-G-T-P-L-A-R-R-D-P-T-P-V-S
H6	T-P-L-A-R-R-D-P-T-P-V-S-S-P-S
H7	A-R-R-D-P-T-P-V-S-S-P-S-S-R-R
H8	D-P-T-P-V-S-S-P-S-S-R-R-A-T-P
H9	P-V-S-S-P-S-S-R-R-A-T-P-I-S-V
H10	S-P-S-S-R-R-A-T-P-I-S-V-W-G-P
H11	S-R-R-A-T-P-I-S-V-W-G-P-L-G-G
I1	A-T-P-I-S-V-W-G-P-L-G-G-L-V-R
I2	I-S-V-W-G-P-L-G-G-L-V-R-T-P-S
I3	W-G-P-L-G-G-L-V-R-T-P-S-V-Q-S
I4	L-G-G-L-V-R-T-P-S-V-Q-S-L-G-S
I5	L-V-R-T-P-S-V-Q-S-L-G-S-D-P-D
I6	T-P-S-V-Q-S-L-G-S-D-P-D-E-Y-A
I7	V-Q-S-L-G-S-D-P-D-E-Y-A-S-S-G
I8	L-G-S-D-P-D-E-Y-A-S-S-G-S-S-L
I9	D-P-D-E-Y-A-S-S-G-S-S-L-G-G-S
I10	E-Y-A-S-S-G-S-S-L-G-G-S-D-S-P
I11	S-S-G-S-S-L-G-G-S-D-S-P-V-F-E

J1	S-S-L-G-G-S-D-S-P-V-F-E-A-G-V
J2	G-G-S-D-S-P-V-F-E-A-G-V-F-A-P
J3	D-S-P-V-F-E-A-G-V-F-A-P-P-Q-P
J4	V-F-E-A-G-V-F-A-P-P-Q-P-V-A-A
J5	A-G-V-F-A-P-P-Q-P-V-A-A-P-R-R
J6	F-A-P-P-Q-P-V-A-A-P-R-R-L-P-I
J7	P-Q-P-V-A-A-P-R-R-L-P-I-F-N-R
J8	V-A-A-P-R-R-L-P-I-F-N-R-I-S-V
J9	A-P-R-R-L-P-I-F-N-R-I-S-V-S-E

Table 7.16 Alanine scan of TTP peptides found to bind CNOT9

List of 15-residue TTP peptide sequences and number/letter locator for each peptide printed on the TTP peptide alanine scan array. Each residue of selected peptides was sequentially mutated, one by one, to determine the contribution of each residue towards CNOT9 binding.

Peptide position	Peptide Sequence
A 14	M-D-L-T-A-I-Y-E-S-L-L-S-L-S-P
A 15	A-D-L-T-A-I-Y-E-S-L-L-S-L-S-P
A 16	M-A-L-T-A-I-Y-E-S-L-L-S-L-S-P
A 17	M-D-A-T-A-I-Y-E-S-L-L-S-L-S-P
A 18	M-D-L-A-A-I-Y-E-S-L-L-S-L-S-P
A 19	M-D-L-T-A-I-Y-E-S-L-L-S-L-S-P
A 20	M-D-L-T-A-A-Y-E-S-L-L-S-L-S-P
B 11	M-D-L-T-A-I-A-E-S-L-L-S-L-S-P
B 12	M-D-L-T-A-I-Y-A-S-L-L-S-L-S-P
B 13	M-D-L-T-A-I-Y-E-A-L-L-S-L-S-P
B 14	M-D-L-T-A-I-Y-E-S-A-L-S-L-S-P
B 15	M-D-L-T-A-I-Y-E-S-L-A-S-L-S-P
B 16	M-D-L-T-A-I-Y-E-S-L-L-A-L-S-P
B 17	M-D-L-T-A-I-Y-E-S-L-L-S-A-S-P
B 18	M-D-L-T-A-I-Y-E-S-L-L-S-L-A-P
B 19	M-D-L-T-A-I-Y-E-S-L-L-S-L-S-A
B 20	(empty space)
C 11	S-S-P-G-W-G-S-S-G-P-W-S-L-S-P
C 12	A-S-P-G-W-G-S-S-G-P-W-S-L-S-P
C 13	S-A-P-G-W-G-S-S-G-P-W-S-L-S-P
C 14	S-S-A-G-W-G-S-S-G-P-W-S-L-S-P
C 15	S-S-P-A-W-G-S-S-G-P-W-S-L-S-P
C 16	S-S-P-G-A-G-S-S-G-P-W-S-L-S-P
C 17	S-S-P-G-W-A-S-S-G-P-W-S-L-S-P
C 18	S-S-P-G-W-G-A-S-G-P-W-S-L-S-P
C 19	S-S-P-G-W-G-S-A-G-P-W-S-L-S-P
C 20	S-S-P-G-W-G-S-S-A-P-W-S-L-S-P

D 11	S-S-P-G-W-G-S-S-G-A-W-S-L-S-P
D 12	S-S-P-G-W-G-S-S-G-P-A-S-L-S-P
D 13	S-S-P-G-W-G-S-S-G-P-W-A-L-S-P
D 14	S-S-P-G-W-G-S-S-G-P-W-S-A-S-P
D 15	S-S-P-G-W-G-S-S-G-P-W-S-L-A-P
D 16	S-S-P-G-W-G-S-S-G-P-W-S-L-S-A
D 17	(empty space)
D 18	S-D-S-S-P-S-G-V-T-S-R-L-P-G-R
D 19	A-D-S-S-P-S-G-V-T-S-R-L-P-G-R
D 20	S-A-S-S-P-S-G-V-T-S-R-L-P-G-R
E 11	S-D-A-S-P-S-G-V-T-S-R-L-P-G-R
E 12	S-D-S-A-P-S-G-V-T-S-R-L-P-G-R
E 13	S-D-S-S-A-S-G-V-T-S-R-L-P-G-R
E 14	S-D-S-S-P-A-G-V-T-S-R-L-P-G-R
E 15	S-D-S-S-P-S-A-V-T-S-R-L-P-G-R
E 16	S-D-S-S-P-S-G-A-T-S-R-L-P-G-R
E 17	S-D-S-S-P-S-G-V-A-S-R-L-P-G-R
E 18	S-D-S-S-P-S-G-V-T-A-R-L-P-G-R
E 19	S-D-S-S-P-S-G-V-T-S-A-L-P-G-R
E 20	S-D-S-S-P-S-G-V-T-S-R-A-P-G-R
F 11	S-D-S-S-P-S-G-V-T-S-R-L-A-G-R
F 12	S-D-S-S-P-S-G-V-T-S-R-L-P-A-R
F 13	S-D-S-S-P-S-G-V-T-S-R-L-P-G-A
F 14	(empty space)
F 15	S-P-S-G-V-T-S-R-L-P-G-R-S-T-S
F 16	A-P-S-G-V-T-S-R-L-P-G-R-S-T-S
F 17	S-A-S-G-V-T-S-R-L-P-G-R-S-T-S
F 18	S-P-A-G-V-T-S-R-L-P-G-R-S-T-S
F 19	S-P-S-A-V-T-S-R-L-P-G-R-S-T-S
F 20	S-P-S-G-A-T-S-R-L-P-G-R-S-T-S
G 11	S-P-S-G-V-A-S-R-L-P-G-R-S-T-S
G 12	S-P-S-G-V-T-A-R-L-P-G-R-S-T-S
G 13	S-P-S-G-V-T-S-A-L-P-G-R-S-T-S
G 14	S-P-S-G-V-T-S-R-A-P-G-R-S-T-S
G 15	S-P-S-G-V-T-S-R-L-A-G-R-S-T-S
G 16	S-P-S-G-V-T-S-R-L-P-A-R-S-T-S
G 17	S-P-S-G-V-T-S-R-L-P-G-A-S-T-S
G 18	S-P-S-G-V-T-S-R-L-P-G-R-A-T-S
G 19	S-P-S-G-V-T-S-R-L-P-G-R-S-A-S
G 20	S-P-S-G-V-T-S-R-L-P-G-R-S-T-A
H 11	(empty space)
H 12	G-R-S-S-G-W-V-P-P-P-P-G-F-A-P
H 13	A-R-S-S-G-W-V-P-P-P-P-G-F-A-P
H 14	G-A-S-S-G-W-V-P-P-P-P-G-F-A-P
H 15	G-R-A-S-G-W-V-P-P-P-P-G-F-A-P
H 16	G-R-S-A-G-W-V-P-P-P-P-G-F-A-P

H 17	G-R-S-S-A-W-V-P-P-P-P-G-F-A-P
H 18	G-R-S-S-G-A-V-P-P-P-P-G-F-A-P
H 19	G-R-S-S-G-W-A-P-P-P-P-G-F-A-P
H 20	G-R-S-S-G-W-V-A-P-P-P-G-F-A-P
I 11	G-R-S-S-G-W-V-P-A-P-P-G-F-A-P
I 12	G-R-S-S-G-W-V-P-P-A-P-G-F-A-P
I 13	G-R-S-S-G-W-V-P-P-P-A-G-F-A-P
I 14	G-R-S-S-G-W-V-P-P-P-P-A-F-A-P
I 15	G-R-S-S-G-W-V-P-P-P-P-G-A-A-P
I 16	G-R-S-S-G-W-V-P-P-P-P-G-F-A-P
I 17	G-R-S-S-G-W-V-P-P-P-P-G-F-A-A
I 18	(empty space)
I 19	T-E-L-S-R-T-F-S-E-S-G-R-S-R-Y
I 20	A-E-L-S-R-T-F-S-E-S-G-R-S-R-Y
J 11	T-A-L-S-R-T-F-S-E-S-G-R-S-R-Y
J 12	T-E-A-S-R-T-F-S-E-S-G-R-S-R-Y
J 13	T-E-L-A-R-T-F-S-E-S-G-R-S-R-Y
J 14	T-E-L-S-A-T-F-S-E-S-G-R-S-R-Y
J 15	T-E-L-S-R-A-F-S-E-S-G-R-S-R-Y
J 16	T-E-L-S-R-T-A-S-E-S-G-R-S-R-Y
J 17	T-E-L-S-R-T-F-A-E-S-G-R-S-R-Y
J 18	T-E-L-S-R-T-F-S-A-S-G-R-S-R-Y
J 19	T-E-L-S-R-T-F-S-E-A-G-R-S-R-Y
J 20	T-E-L-S-R-T-F-S-E-S-A-R-S-R-Y
K 11	T-E-L-S-R-T-F-S-E-S-G-A-S-R-Y
K 12	T-E-L-S-R-T-F-S-E-S-G-R-A-R-Y
K 13	T-E-L-S-R-T-F-S-E-S-G-R-S-A-Y
K 14	T-E-L-S-R-T-F-S-E-S-G-R-S-R-A
K 15	(empty space)
K 16	S-Q-F-A-H-G-L-G-E-L-R-Q-A-N-R
K 17	A-Q-F-A-H-G-L-G-E-L-R-Q-A-N-R
K 18	S-A-F-A-H-G-L-G-E-L-R-Q-A-N-R
K 19	S-Q-A-A-H-G-L-G-E-L-R-Q-A-N-R
K 20	S-Q-F-A-H-G-L-G-E-L-R-Q-A-N-R
L 11	S-Q-F-A-A-G-L-G-E-L-R-Q-A-N-R
L 12	S-Q-F-A-H-A-L-G-E-L-R-Q-A-N-R
L 13	S-Q-F-A-H-G-A-G-E-L-R-Q-A-N-R
L 14	S-Q-F-A-H-G-L-A-E-L-R-Q-A-N-R
L 15	S-Q-F-A-H-G-L-G-A-L-R-Q-A-N-R
L 16	S-Q-F-A-H-G-L-G-E-A-R-Q-A-N-R
L 17	S-Q-F-A-H-G-L-G-E-L-A-Q-A-N-R
L 18	S-Q-F-A-H-G-L-G-E-L-R-A-A-N-R
L 19	S-Q-F-A-H-G-L-G-E-L-R-Q-A-N-R
L 20	S-Q-F-A-H-G-L-G-E-L-R-Q-A-A-R
M 11	S-Q-F-A-H-G-L-G-E-L-R-Q-A-N-A
M 12	(empty space)

M 13	E-L-S-H-K-F-Y-L-Q-G-R-S-P-Y-G
M 14	A-L-S-H-K-F-Y-L-Q-G-R-S-P-Y-G
M 15	E-A-S-H-K-F-Y-L-Q-G-R-S-P-Y-G
M 16	E-L-A-H-K-F-Y-L-Q-G-R-S-P-Y-G
M 17	E-L-S-A-K-F-Y-L-Q-G-R-S-P-Y-G
M 18	E-L-S-H-A-F-Y-L-Q-G-R-S-P-Y-G
M 19	E-L-S-H-K-A-Y-L-Q-G-R-S-P-Y-G
M 20	E-L-S-H-K-F-A-L-Q-G-R-S-P-Y-G
N 11	E-L-S-H-K-F-Y-A-Q-G-R-S-P-Y-G
N 12	E-L-S-H-K-F-Y-L-A-G-R-S-P-Y-G
N 13	E-L-S-H-K-F-Y-L-Q-A-R-S-P-Y-G
N 14	E-L-S-H-K-F-Y-L-Q-G-A-S-P-Y-G
N 15	E-L-S-H-K-F-Y-L-Q-G-R-A-P-Y-G
N 16	E-L-S-H-K-F-Y-L-Q-G-R-S-A-Y-G
N 17	E-L-S-H-K-F-Y-L-Q-G-R-S-P-A-G
N 18	E-L-S-H-K-F-Y-L-Q-G-R-S-P-Y-A
N 19	(empty space)
N 20	L-R-Q-S-I-S-F-S-G-L-P-S-G-R-R
O 11	A-R-Q-S-I-S-F-S-G-L-P-S-G-R-R
O 12	L-A-Q-S-I-S-F-S-G-L-P-S-G-R-R
O 13	L-R-A-S-I-S-F-S-G-L-P-S-G-R-R
O 14	L-R-Q-A-I-S-F-S-G-L-P-S-G-R-R
O 15	L-R-Q-S-A-S-F-S-G-L-P-S-G-R-R
O 16	L-R-Q-S-I-A-F-S-G-L-P-S-G-R-R
O 17	L-R-Q-S-I-S-A-S-G-L-P-S-G-R-R
O 18	L-R-Q-S-I-S-F-A-G-L-P-S-G-R-R
O 19	L-R-Q-S-I-S-F-S-A-L-P-S-G-R-R
O 20	L-R-Q-S-I-S-F-S-G-A-P-S-G-R-R
P 11	L-R-Q-S-I-S-F-S-G-L-A-S-G-R-R
P 12	L-R-Q-S-I-S-F-S-G-L-P-A-G-R-R
P 13	L-R-Q-S-I-S-F-S-G-L-P-S-A-R-R
P 14	L-R-Q-S-I-S-F-S-G-L-P-S-G-A-R
P 15	L-R-Q-S-I-S-F-S-G-L-P-S-G-R-A
P 16	(empty space)
P 17	S-P-S-S-R-R-A-T-P-I-S-V-W-G-P
P 18	A-P-S-S-R-R-A-T-P-I-S-V-W-G-P
P 19	S-A-S-S-R-R-A-T-P-I-S-V-W-G-P
P 20	S-P-A-S-R-R-A-T-P-I-S-V-W-G-P
Q 11	S-P-S-A-R-R-A-T-P-I-S-V-W-G-P
Q 12	S-P-S-S-A-R-A-T-P-I-S-V-W-G-P
Q 13	S-P-S-S-R-A-A-T-P-I-S-V-W-G-P
Q 14	S-P-S-S-R-R-A-T-P-I-S-V-W-G-P
Q 15	S-P-S-S-R-R-A-A-P-I-S-V-W-G-P
Q 16	S-P-S-S-R-R-A-T-A-I-S-V-W-G-P
Q 17	S-P-S-S-R-R-A-T-P-A-S-V-W-G-P
Q 18	S-P-S-S-R-R-A-T-P-I-A-V-W-G-P

Q 19	S-P-S-S-R-R-A-T-P-I-S-A-W-G-P
Q 20	S-P-S-S-R-R-A-T-P-I-S-V-A-G-P
R 11	S-P-S-S-R-R-A-T-P-I-S-V-W-A-P
R 12	S-P-S-S-R-R-A-T-P-I-S-V-W-G-A
R 13	(empty space)
R 14	A-T-P-I-S-V-W-G-P-L-G-G-L-V-R
R 15	A-T-P-I-S-V-W-G-P-L-G-G-L-V-R
R 16	A-A-P-I-S-V-W-G-P-L-G-G-L-V-R
R 17	A-T-A-I-S-V-W-G-P-L-G-G-L-V-R
R 18	A-T-P-A-S-V-W-G-P-L-G-G-L-V-R
R 19	A-T-P-I-A-V-W-G-P-L-G-G-L-V-R
R 20	A-T-P-I-S-A-W-G-P-L-G-G-L-V-R
S 11	A-T-P-I-S-V-A-G-P-L-G-G-L-V-R
S 12	A-T-P-I-S-V-W-A-P-L-G-G-L-V-R
S 13	A-T-P-I-S-V-W-G-A-L-G-G-L-V-R
S 14	A-T-P-I-S-V-W-G-P-A-G-G-L-V-R
S 15	A-T-P-I-S-V-W-G-P-L-A-G-L-V-R
S 16	A-T-P-I-S-V-W-G-P-L-G-A-L-V-R
S 17	A-T-P-I-S-V-W-G-P-L-G-G-A-V-R
S 18	A-T-P-I-S-V-W-G-P-L-G-G-L-A-R
S 19	A-T-P-I-S-V-W-G-P-L-G-G-L-V-A
S 20	(empty space)
T 11	P-Q-P-V-A-A-P-R-R-L-P-I-F-N-R
T 12	A-Q-P-V-A-A-P-R-R-L-P-I-F-N-R
T 13	P-A-P-V-A-A-P-R-R-L-P-I-F-N-R
T 14	P-Q-A-V-A-A-P-R-R-L-P-I-F-N-R
T 15	P-Q-P-A-A-A-P-R-R-L-P-I-F-N-R
T 16	P-Q-P-V-A-A-P-R-R-L-P-I-F-N-R
T 17	P-Q-P-V-A-A-P-R-R-L-P-I-F-N-R
T 18	P-Q-P-V-A-A-A-R-R-L-P-I-F-N-R
T 19	P-Q-P-V-A-A-P-A-R-L-P-I-F-N-R
T 20	P-Q-P-V-A-A-P-R-A-L-P-I-F-N-R
U 11	P-Q-P-V-A-A-P-R-R-A-P-I-F-N-R
U 12	P-Q-P-V-A-A-P-R-R-L-A-I-F-N-R
U 13	P-Q-P-V-A-A-P-R-R-L-P-A-F-N-R
U 14	P-Q-P-V-A-A-P-R-R-L-P-I-A-N-R
U 15	P-Q-P-V-A-A-P-R-R-L-P-I-F-A-R
U 16	P-Q-P-V-A-A-P-R-R-L-P-I-F-N-A
U 17	(empty space)
U 18	(empty space)
U 19	H-H-H-H-H-H-H-H-H-H-X10
U 20	H-H-H-H-H-H-H-H-H-H-X10

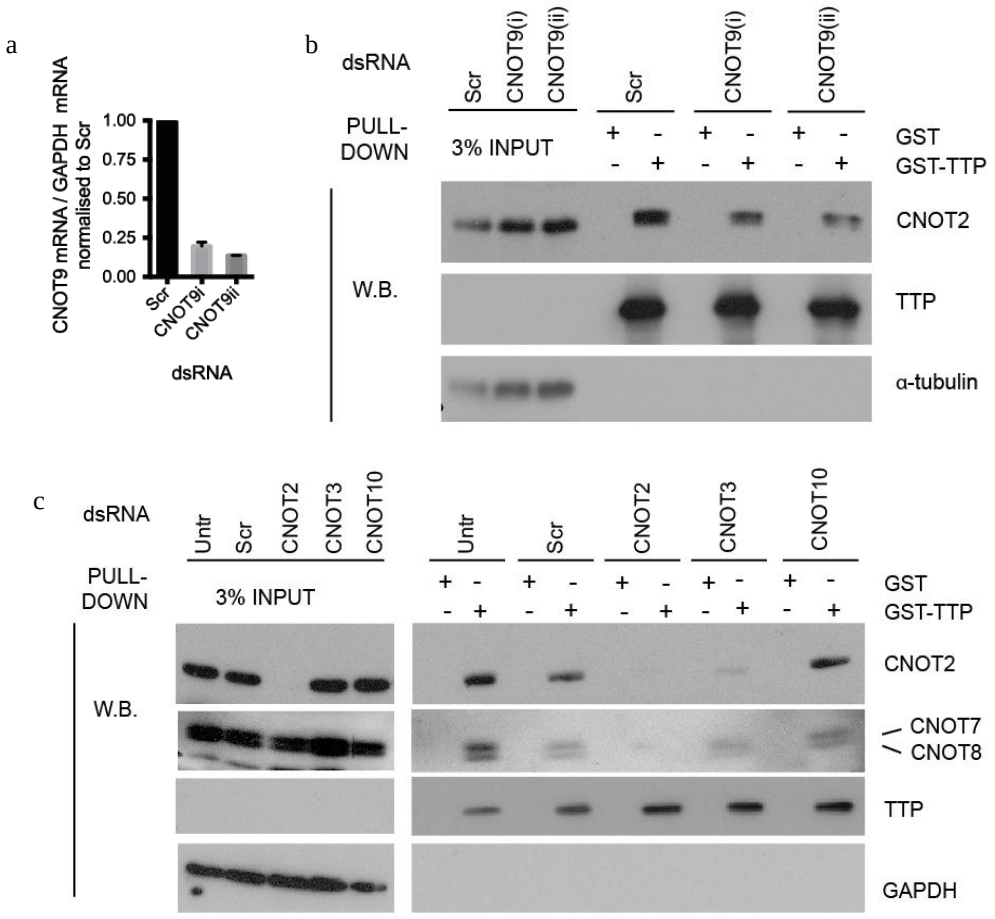


Figure 7.11 CNOT9 and CNOT2 protein knock-down inhibits the TTP: CCR4-NOT interaction.

An siRNA screen was used to determine if TTP interacts with any additional CNOT subunits beyond CNOT1. (a) HeLa cells were transfected with Scr, separately with two different siRNAs targeting CNOT9 or left untransfected. mRNA knockdown efficiencies were determined by qRT-PCR. (b) CNOT9 knockdown lysates were assessed by GST-pulldown. Pulled down proteins were resolved by SDS-PAGE and western blotted for CNOT2, TTP and α-tubulin. Western blots representative of three experiments (n=3). (c) HeLa cells were transfected separately with a single siRNA against either CNOT2, CNOT3, CNOT10 or a Scr control dsRNA. Cells were lysed and pull-downs performed with GST or GST-mTTP protein. Bound proteins were analysed by western blot for CCR4-CAF1 subunits using anti-CNOT2 antibody, an antibody recognising both CNOT7 and CNOT8, anti-TTP and anti-GAPDH. Inputs and pull-down panels from the same gels are shown. CNOT10 and GAPDH mRNAs were quantified by qRT-PCR for duplicate wells of cells used in pull-down assays. Western blots representative of three experiments (n=3). These experiments identified that TTP also interacts with CNOT2 and CNOT9.

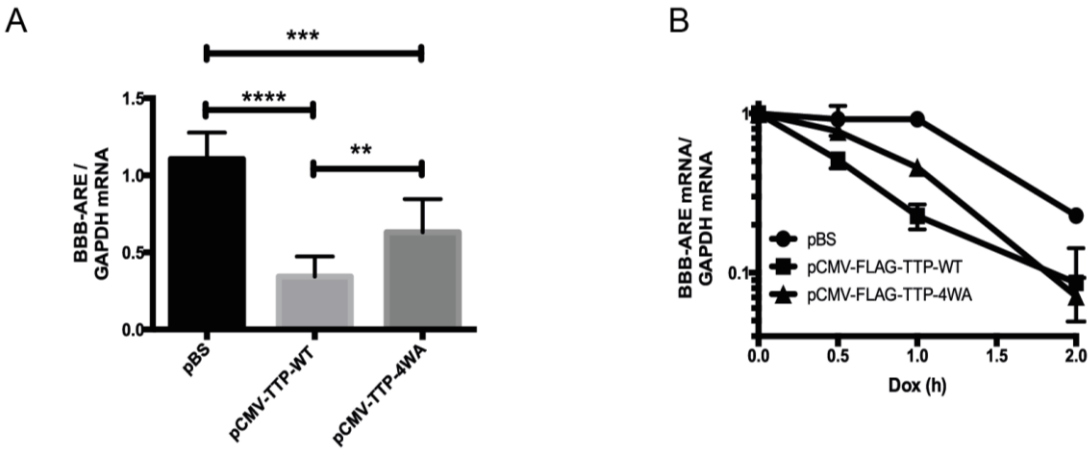


Figure 7.12 Mutation of Trp residues in TTP increases the stability of BBB-TNF-ARE mRNA levels

As it was shown in vitro that mutating the tryptophan residues in TTP disrupts the interaction with CNOT9, the interaction was explored in vitro. All four tryptophans in TTP were mutated (4WA) and the ability of TTP to recruit the CCR4-NOT complex was studied by measuring the amount of mRNA degradation. HeLa tet-off cells were transfected with 50 ng pCMV-FLAG-TTP WT, pCMV-FLAG-TTP 4WA or pBS only with carrier pBS DNA to a total of 1 µg plasmid DNA. Transcription was inhibited by addition of culture medium containing 500 ng/ml doxycycline, and cells were then lysed at the appropriate time points. BBB-TNF-ARE and GAPDH mRNA levels in the lysates were determined by qRT-PCR using the appropriate Taqman primer sets, and cDNA that had been generated from DNase I-treated RNA preparations. A – Bar graph showing mean steady state BBB-TNF-ARE mRNA levels (t=0) normalised against GAPDH mRNA. B - The semi-log graph shows the mean rate of decay of BBB-TNF-ARE mRNA normalised against GAPDH mRNA (+SD) from three independent experiments (n=3). Wild-type TTP degraded the β–globin mRNA substrate, showing a 69% reduction in mRNA levels, whereas the TTP (4WA) only caused a 43% reduction in β–globin mRNA levels. Thus, the quadruple tryptophan mutations in TTP sufficiently disrupt interactions with the CCR4-NOT complex, and have a direct, functionally relevant effect on its cellular function i.e. mRNA decay.

8. References

ABREMSKI, K., WIERZBICKI, A., FROMMER, B. & HOESS, R. H. 1986. Bacteriophage P1 Cre-loxP site-specific recombination. Site-specific DNA topoisomerase activity of the Cre recombination protein. *J Biol Chem,* 261**,** 391-6.

AJIRO, M., KATAGIRI, T., UEDA, K., NAKAGAWA, H., FUKUKAWA, C., LIN, M. L., PARK, J. H., NISHIDATE, T., DAIGO, Y. & NAKAMURA, Y. 2009. Involvement of RQCD1 overexpression, a novel cancer-testis antigen, in the Akt pathway in breast cancer cells. *Int J Oncol,* 35**,** 673-81.

ALBERT, T. K., HANZAWA, H., LEGTENBERG, Y. I., DE RUWE, M. J., VAN DEN HEUVEL, F. A., COLLART, M. A., BOELENS, R. & TIMMERS, H. T. 2002. Identification of a ubiquitin-protein ligase subunit within the CCR4-NOT transcription repressor complex. *EMBO J,* 21**,** 355-64.

ALBERT, T. K., LEMAIRE, M., VAN BERKUM, N. L., GENTZ, R., COLLART, M. A. & TIMMERS, H. T. 2000. Isolation and characterization of human orthologs of yeast CCR4-NOT complex subunits. *Nucleic Acids Res,* 28**,** 809-17.

ANDRADE, M. A., PETOSA, C., O'DONOGHUE, S. I., MULLER, C. W. & BORK, P. 2001. Comparison of ARM and HEAT protein repeats. *J Mol Biol,* 309**,** 1-18.

BAI, Y., SALVADORE, C., CHIANG, Y. C., COLLART, M. A., LIU, H. Y. & DENIS, C. L. 1999. The CCR4 and CAF1 proteins of the CCR4-NOT complex are physically and functionally separated from NOT2, NOT4, and NOT5. *Mol Cell Biol,* 19**,** 6642-51.

BASQUIN, J., ROUDKO, V. V., RODE, M., BASQUIN, C., SERAPHIN, B. & CONTI, E. 2012. Architecture of the nuclease module of the yeast Ccr4-not complex: the Not1-Caf1-Ccr4 interaction. *Mol Cell,* 48**,** 207-18.

BIENIOSSEK, C., RICHMOND, T. J. & BERGER, I. 2008. MultiBac: multigene baculovirus-based eukaryotic protein complex production. *Curr Protoc Protein Sci,* Chapter 5**,** Unit 5 20.

BLOW, D. M. 1968. Chymotrypsin: tertiary structure and enzymic activity. *Biochem J,* 110**,** 2P.

BOLAND, A., CHEN, Y., RAISCH, T., JONAS, S., KUZUOGLU-OZTURK, D., WOHLBOLD, L., WEICHENRIEDER, O. & IZAURRALDE, E. 2013. Structure and assembly of the NOT module of the human CCR4-NOT complex. *Nat Struct Mol Biol,* 20**,** 1289-97.

BRENNAN, S. E., KUWANO, Y., ALKHAROUF, N., BLACKSHEAR, P. J., GOROSPE, M. & WILSON, G. M. 2009. The mRNA-destabilizing protein tristetraprolin is suppressed in many cancers, altering tumorigenic phenotypes and patient prognosis. *Cancer Res,* 69**,** 5168-76.

BRIANT, D. J., MURPHY, J. M., LEUNG, G. C. & SICHERI, F. 2009. Rapid identification of linear protein domain binding motifs using peptide SPOT arrays. *Methods Mol Biol,* 570**,** 175-85.

BROUWER, R., ALLMANG, C., RAIJMAKERS, R., VAN AARSSEN, Y., EGBERTS, W. V., PETFALSKI, E., VAN VENROOIJ, W. J., TOLLERVEY, D. & PRUIJN, G. J. 2001. Three novel components of the human exosome. *J Biol Chem,* 276**,** 6177-84.

BUCHAN, D. W., MINNECI, F., NUGENT, T. C., BRYSON, K. & JONES, D. T. 2013. Scalable web services for the PSIPRED Protein Analysis Workbench. *Nucleic Acids Res,* 41**,** W349-57.

BUCHANAN, S. G. & GAY, N. J. 1996. Structural and functional diversity in the leucine-rich repeat family of proteins. *Prog Biophys Mol Biol,* 65**,** 1-44.

BULBROOK, D. 2015. Regulation of the interaction between the RNA-binding protein, tristetraprolin, and the CCR4-CAF1 deadenylase complex for the control of inflammation.

CAO, H., DETERDING, L. J. & BLACKSHEAR, P. J. 2007. Phosphorylation site analysis of the anti-inflammatory and mRNA-destabilizing protein tristetraprolin. *Expert Rev Proteomics,* 4**,** 711-26.

CARBALLO, E., LAI, W. S. & BLACKSHEAR, P. J. 1998. Feedback inhibition of macrophage tumor necrosis factor-alpha production by tristetraprolin. *Science,* 281**,** 1001-5.

CHANG, H., LIM, J., HA, M. & KIM, V. N. 2014. TAIL-seq: genome-wide determination of poly(A) tail length and 3' end modifications. *Mol Cell,* 53**,** 1044-52.

CHEKANOVA, J. A., SHAW, R. J., WILLS, M. A. & BELOSTOTSKY, D. A. 2000. Poly(A) tail-dependent exonuclease AtRrp41p from Arabidopsis thaliana rescues 5.8 S rRNA processing and mRNA decay defects of the yeast ski6 mutant and is found in an exosome-sized complex in plant and yeast cells. *J Biol Chem,* 275**,** 33158-66.

CHEN, C. Y. & SHYU, A. B. 1995. AU-rich elements: characterization and importance in mRNA degradation. *Trends Biochem Sci,* 20**,** 465-70.

CHEN, Y., BOLAND, A., KUZUOGLU-OZTURK, D., BAWANKAR, P., LOH, B., CHANG, C. T., WEICHENRIEDER, O. & IZAURRALDE, E. 2014. A DDX6-CNOT1 complex and W-binding pockets in CNOT9 reveal direct links between miRNA target recognition and silencing. *Mol Cell,* 54**,** 737-50.

CHERNOKALSKAYA, E., DUBELL, A. N., CUNNINGHAM, K. S., HANSON, M. N., DOMPENCIEL, R. E. & SCHOENBERG, D. R. 1998. A polysomal ribonuclease involved in the destabilization of albumin mRNA is a novel member of the peroxidase gene family. *RNA,* 4**,** 1537-48.

CHIVERS, C. E., KONER, A. L., LOWE, E. D. & HOWARTH, M. 2011. How the biotin-streptavidin interaction was made even stronger: investigation via crystallography and a chimaeric tetramer. *Biochem J,* 435**,** 55-63.

CHOY, E. H., KAVANAUGH, A. F. & JONES, S. A. 2013. The problem of choice: current biologic agents and future prospects in RA. *Nat Rev Rheumatol,* 9**,** 154-63.

CHRESTENSEN, C. A., SCHROEDER, M. J., SHABANOWITZ, J., HUNT, D. F., PELO, J. W., WORTHINGTON, M. T. & STURGILL, T. W. 2004. MAPKAP kinase 2 phosphorylates tristetraprolin on in vivo sites including Ser178, a site required for 14-3-3 binding. *J Biol Chem,* 279**,** 10176-84.

CLEMENT 2011. Phosphorylation of Tristetraprolin by MK2 Impairs AU-Rich Element mRNA Decay by Preventing Deadenylase Recruitment. *Molecular and Cellular Biology,* 33**,** 256-266.

DE BOER, J. F., DIKKERS, A., JURDZINSKI, A., VON FELDEN, J., GAESTEL, M., BAVENDIEK, U. & TIETGE, U. J. 2014. Mitogen-activated protein kinase-activated protein kinase 2 deficiency reduces insulin sensitivity in high-fat diet-fed mice. *PLoS One,* 9**,** e106300.

DEAN, J. L., BROOK, M., CLARK, A. R. & SAKLATVALA, J. 1999. p38 mitogen-activated protein kinase regulates cyclooxygenase-2 mRNA stability and transcription in lipopolysaccharide-treated human monocytes. *J Biol Chem,* 274**,** 264-9.

DESAI, P. J., PRYCE, E. N., HENSON, B. W., LUITWEILER, E. M. & COTHRAN, J. 2012. Reconstitution of the Kaposi's sarcoma-associated herpesvirus nuclear egress complex and formation of nuclear membrane vesicles by coexpression of ORF67 and ORF69 gene products. *J Virol,* 86**,** 594-8.

DOMINGUEZ, C., POWERS, D. A. & TAMAYO, N. 2005. p38 MAP kinase inhibitors: many are made, but few are chosen. *Curr Opin Drug Discov Devel,* 8**,** 421-30.

ESTEVEZ, A. M., KEMPF, T. & CLAYTON, C. 2001. The exosome of Trypanosoma brucei. *EMBO J,* 20**,** 3831-9.

FABIAN, M. R., FRANK, F., ROUYA, C., SIDDIQUI, N., LAI, W. S., KARETNIKOV, A., BLACKSHEAR, P. J., NAGAR, B. & SONENBERG, N. 2013. Structural basis for the recruitment of the human CCR4-NOT deadenylase complex by tristetraprolin. *Nat Struct Mol Biol,* 20**,** 735-9.

FARBER, V., ERBEN, E., SHARMA, S., STOECKLIN, G. & CLAYTON, C. 2013. Trypanosome CNOT10 is essential for the integrity of the NOT deadenylase complex and for degradation of many mRNAs. *Nucleic Acids Res,* 41**,** 1211-22.

FINN, R. D., COGGILL, P., EBERHARDT, R. Y., EDDY, S. R., MISTRY, J., MITCHELL, A. L., POTTER, S. C., PUNTA, M., QURESHI, M., SANGRADOR-VEGAS, A., SALAZAR, G. A., TATE, J. & BATEMAN, A. 2016. The Pfam protein families database: towards a more sustainable future. *Nucleic Acids Res,* 44**,** D279-85.

FISHMAN. P. 2010. Rheumatoid Arthritis: Background History, Molecular

Mechanisms and Therapeutic Applications. *Adenosine Receptors from Cell Biology to Pharmacology*

and Therapeutics. P.A Borea.

FITZGERALD, D. J., BERGER, P., SCHAFFITZEL, C., YAMADA, K., RICHMOND, T. J. & BERGER, I. 2006. Protein complex expression by using multigene baculoviral vectors. *Nat Methods,* 3**,** 1021-32.

FORBES, S. A., BINDAL, N., BAMFORD, S., COLE, C., KOK, C. Y., BEARE, D., JIA, M., SHEPHERD, R., LEUNG, K., MENZIES, A., TEAGUE, J. W., CAMPBELL, P. J., STRATTON, M. R. & FUTREAL, P. A. 2011. COSMIC: mining complete cancer genomes in the Catalogue of Somatic Mutations in Cancer. *Nucleic Acids Res,* 39**,** D945-50.

FRANK, R. 2002. The SPOT-synthesis technique. Synthetic peptide arrays on membrane supports--principles and applications. *J Immunol Methods,* 267**,** 13-26.

GAFFO, A., SAAG, K. G. & CURTIS, J. R. 2006. Treatment of rheumatoid arthritis. *Am J Health Syst Pharm,* 63**,** 2451-65.

GALLIE, D. R. 1998. A tale of two termini: a functional interaction between the termini of an mRNA is a prerequisite for efficient translation initiation. *Gene,* 216**,** 1-11.

GAO, M., LOE, D. W., GRANT, C. E., COLE, S. P. & DEELEY, R. G. 1996. Reconstitution of ATP-dependent leukotriene C4 transport by Co-expression of both half-molecules of human multidrug resistance protein in insect cells. *J Biol Chem,* 271**,** 27782-7.

GARCES, R. G., GILLON, W. & PAI, E. F. 2007. Atomic model of human Rcd-1 reveals an armadillo-like-repeat protein with in vitro nucleic acid binding properties. *Protein Sci,* 16**,** 176-88.

GARNEAU, N. L., WILUSZ, J. & WILUSZ, C. J. 2007. The highways and byways of mRNA decay. *Nat Rev Mol Cell Biol,* 8**,** 113-26.

GILEADI, O., BURGESS-BROWN, N. A., COLEBROOK, S. M., BERRIDGE, G., SAVITSKY, P., SMEE, C. E., LOPPNAU, P., JOHANSSON, C., SALAH, E. & PANTIC, N. H. 2008. High throughput production of recombinant human proteins for crystallography. *Methods Mol Biol,* 426**,** 221-46.

GIVER, L., BARTEL, D. P., ZAPP, M. L., GREEN, M. R. & ELLINGTON, A. D. 1993. Selection and design of high-affinity RNA ligands for HIV-1 Rev. *Gene,* 137**,** 19-24.

GOLDSTROHM, A. C. & WICKENS, M. 2008. Multifunctional deadenylase complexes diversify mRNA control. *Nat Rev Mol Cell Biol,* 9**,** 337-44.

HAMMAKER, D. & FIRESTEIN, G. S. 2010. "Go upstream, young man": lessons learned from the p38 saga. *Ann Rheum Dis,* 69 Suppl 1**,** i77-82.

HILPERT, K., WINKLER, D. F. & HANCOCK, R. E. 2007. Peptide arrays on cellulose support: SPOT synthesis, a time and cost efficient method for synthesis of large numbers of peptides in a parallel and addressable fashion. *Nat Protoc,* 2**,** 1333-49.

HITTI, E., IAKOVLEVA, T., BROOK, M., DEPPENMEIER, S., GRUBER, A. D., RADZIOCH, D., CLARK, A. R., BLACKSHEAR, P. J., KOTLYAROV, A. & GAESTEL, M. 2006. Mitogen-activated protein kinase-activated protein kinase 2 regulates tumor necrosis factor mRNA stability and translation mainly by altering tristetraprolin expression, stability, and binding to adenine/uridine-rich element. *Mol Cell Biol,* 26**,** 2399-407.

HUARTE, J., STUTZ, A., O'CONNELL, M. L., GUBLER, P., BELIN, D., DARROW, A. L., STRICKLAND, S. & VASSALLI, J. D. 1992. Transient translational silencing by reversible mRNA deadenylation. *Cell,* 69**,** 1021-30.

HUDSON, B. P., MARTINEZ-YAMOUT, M. A., DYSON, H. J. & WRIGHT, P. E. 2004. Recognition of the mRNA AU-rich element by the zinc finger domain of TIS11d. *Nat Struct Mol Biol,* 11**,** 257-64.

IWAKAWA, H. O. & TOMARI, Y. 2015. The Functions of MicroRNAs: mRNA Decay and Translational Repression. *Trends Cell Biol,* 25**,** 651-65.
JONES, D. T. & COZZETTO, D. 2015. DISOPRED3: precise disordered region predictions with annotated protein-binding activity. *Bioinformatics,* 31**,** 857-63.
JUERS, D. H., MATTHEWS, B. W. & HUBER, R. E. 2012. LacZ beta-galactosidase: structure and function of an enzyme of historical and molecular biological importance. *Protein Sci,* 21**,** 1792-807.
KAHVEJIAN, A., SVITKIN, Y. V., SUKARIEH, R., M'BOUTCHOU, M. N. & SONENBERG, N. 2005. Mammalian poly(A)-binding protein is a eukaryotic translation initiation factor, which acts via multiple mechanisms. *Genes Dev,* 19**,** 104-13.
KAPITANY, A., ZILAHI, E., SZANTO, S., SZUCS, G., SZABO, Z., VEGVARI, A., RASS, P., SIPKA, S., SZEGEDI, G. & SZEKANECZ, Z. 2005. Association of rheumatoid arthritis with HLA-DR1 and HLA-DR4 in Hungary. *Ann N Y Acad Sci,* 1051**,** 263-70.
KAPUST, R. B. & WAUGH, D. S. 1999. Escherichia coli maltose-binding protein is uncommonly effective at promoting the solubility of polypeptides to which it is fused. *Protein Sci,* 8**,** 1668-74.
KEDDE, M., STRASSER, M. J., BOLDAJIPOUR, B., OUDE VRIELINK, J. A., SLANCHEV, K., LE SAGE, C., NAGEL, R., VOORHOEVE, P. M., VAN DUIJSE, J., OROM, U. A., LUND, A. H., PERRAKIS, A., RAZ, E. & AGAMI, R. 2007. RNA-binding protein Dnd1 inhibits microRNA access to target mRNA. *Cell,* 131**,** 1273-86.
KEEFE, A. D., PAI, S. & ELLINGTON, A. 2010. Aptamers as therapeutics. *Nat Rev Drug Discov,* 9**,** 537-50.
KELLEY, L. A. & STERNBERG, M. J. 2009. Protein structure prediction on the Web: a case study using the Phyre server. *Nat Protoc,* 4**,** 363-71.
KONTOYIANNIS, D., PASPARAKIS, M., PIZARRO, T. T., COMINELLI, F. & KOLLIAS, G. 1999. Impaired on/off regulation of TNF biosynthesis in mice lacking TNF AU-rich elements: implications for joint and gut-associated immunopathologies. *Immunity,* 10**,** 387-98.
KOOPMAN, W. 2003. Clinical Primer of Rheumatology.
KRIEGER, M., KAY, L. M. & STROUD, R. M. 1974. Structure and specific binding of trypsin: comparison of inhibited derivatives and a model for substrate binding. *J Mol Biol,* 83**,** 209-30.
LABNO, A., TOMECKI, R. & DZIEMBOWSKI, A. 2016. Cytoplasmic RNA decay pathways - Enzymes and mechanisms. *Biochim Biophys Acta,* 1863**,** 3125-3147.
LAI, W. S., STUMPO, D. J. & BLACKSHEAR, P. J. 1990. Rapid insulin-stimulated accumulation of an mRNA encoding a proline-rich protein. *J Biol Chem,* 265**,** 16556-63.
LAITY, J. H., LEE, B. M. & WRIGHT, P. E. 2001. Zinc finger proteins: new insights into structural and functional diversity. *Curr Opin Struct Biol,* 11**,** 39-46.
LAU, N. C., KOLKMAN, A., VAN SCHAIK, F. M., MULDER, K. W., PIJNAPPEL, W. W., HECK, A. J. & TIMMERS, H. T. 2009. Human Ccr4-Not complexes contain variable deadenylase subunits. *Biochem J,* 422**,** 443-53.
LEE, J. C., LAYDON, J. T., MCDONNELL, P. C., GALLAGHER, T. F., KUMAR, S., GREEN, D., MCNULTY, D., BLUMENTHAL, M. J., HEYS, J. R., LANDVATTER, S. W. & ET AL. 1994. A protein kinase involved in the regulation of inflammatory cytokine biosynthesis. *Nature,* 372**,** 739-46.
LEPPEK, K. & STOECKLIN, G. 2014. An optimized streptavidin-binding RNA aptamer for purification of ribonucleoprotein complexes identifies novel ARE-binding proteins. *Nucleic Acids Res,* 42**,** e13.
LING, M. M. & ROBINSON, B. H. 1997. Approaches to DNA mutagenesis: an overview. *Anal Biochem,* 254**,** 157-78.
LIU, H., JIN, L., KOH, S. B., ATANASOV, I., SCHEIN, S., WU, L. & ZHOU, Z. H. 2010. Atomic structure of human adenovirus by cryo-EM reveals interactions among protein networks. *Science,* 329**,** 1038-43.

LUI, H.-Y. 1998. The NOT proteins are part of the CCR4 transcriptional complex and affect gene expression both positively and negatively. *The EMBO Journal* 1096-1106.
LYKKE-ANDERSEN, J. & WAGNER, E. 2005. Recruitment and activation of mRNA decay enzymes by two ARE-mediated decay activation domains in the proteins TTP and BRF-1. *Genes Dev,* 19**,** 351-61.
MAHTANI, K. R., BROOK, M., DEAN, J. L., SULLY, G., SAKLATVALA, J. & CLARK, A. R. 2001. Mitogen-activated protein kinase p38 controls the expression and posttranslational modification of tristetraprolin, a regulator of tumor necrosis factor alpha mRNA stability. *Mol Cell Biol,* 21**,** 6461-9.
MAILLET, L., TU, C., HONG, Y. K., SHUSTER, E. O. & COLLART, M. A. 2000. The essential function of Not1 lies within the Ccr4-Not complex. *J Mol Biol,* 303**,** 131-43.
MALVAR, T., BIRON, R. W., KABACK, D. B. & DENIS, C. L. 1992. The CCR4 protein from Saccharomyces cerevisiae contains a leucine-rich repeat region which is required for its control of ADH2 gene expression. *Genetics,* 132**,** 951-62.
MANDEL, C. R., GEBAUER, D., ZHANG, H. & TONG, L. 2006. A serendipitous discovery that in situ proteolysis is essential for the crystallization of yeast CPSF-100 (Ydh1p). *Acta Crystallogr Sect F Struct Biol Cryst Commun,* 62**,** 1041-5.
MARCHESE, F. P., AUBAREDA, A., TUDOR, C., SAKLATVALA, J., CLARK, A. R. & DEAN, J. L. 2010. MAPKAP kinase 2 blocks tristetraprolin-directed mRNA decay by inhibiting CAF1 deadenylase recruitment. *J Biol Chem,* 285**,** 27590-600.
MARTENS, G. J., PIOSIK, P. A. & DANEN, E. H. 1992. Evolutionary conservation of the 14-3-3 protein. *Biochem Biophys Res Commun,* 184**,** 1456-9.
MARYATI, M., KAUR, I., JADHAV, G. P., OLOTU-UMOREN, L., OVEH, B., HASHMI, L., FISCHER, P. M. & WINKLER, G. S. 2014. A fluorescence-based assay suitable for quantitative analysis of deadenylase enzyme activity. *Nucleic Acids Res,* 42**,** e30.
MATHYS, H., BASQUIN, J., OZGUR, S., CZARNOCKI-CIECIURA, M., BONNEAU, F., AARTSE, A., DZIEMBOWSKI, A., NOWOTNY, M., CONTI, E. & FILIPOWICZ, W. 2014. Structural and biochemical insights to the role of the CCR4-NOT complex and DDX6 ATPase in microRNA repression. *Mol Cell,* 54**,** 751-65.
MAUXION, F., PREVE, B. & SERAPHIN, B. 2013. C2ORF29/CNOT11 and CNOT10 form a new module of the CCR4-NOT complex. *RNA Biol,* 10**,** 267-76.
MELNIKOVA, I. & GOLDEN, J. 2004. Targeting protein kinases. *Nat Rev Drug Discov,* 3**,** 993-4.
MENARD, S., TAGLIABUE, E., CAMPIGLIO, M. & PUPA, S. M. 2000. Role of HER2 gene overexpression in breast carcinoma. *J Cell Physiol,* 182**,** 150-62.
MITCHELL, P., PETFALSKI, E., SHEVCHENKO, A., MANN, M. & TOLLERVEY, D. 1997. The exosome: a conserved eukaryotic RNA processing complex containing multiple 3'-->5' exoribonucleases. *Cell,* 91**,** 457-66.
MOHLER, K. M., TORRANCE, D. S., SMITH, C. A., GOODWIN, R. G., STREMLER, K. E., FUNG, V. P., MADANI, H. & WIDMER, M. B. 1993. Soluble tumor necrosis factor (TNF) receptors are effective therapeutic agents in lethal endotoxemia and function simultaneously as both TNF carriers and TNF antagonists. *J Immunol,* 151**,** 1548-61.
MORITA, M., SUZUKI, T., NAKAMURA, T., YOKOYAMA, K., MIYASAKA, T. & YAMAMOTO, T. 2007. Depletion of mammalian CCR4b deadenylase triggers elevation of the p27Kip1 mRNA level and impairs cell growth. *Mol Cell Biol,* 27**,** 4980-90.
MUGRIDGE, J. S., ZIEMNIAK, M., JEMIELITY, J. & GROSS, J. D. 2016. Structural basis of mRNA-cap recognition by Dcp1-Dcp2. *Nat Struct Mol Biol,* 23**,** 987-994.
NASERTORABI, F., BATISSE, C., DIEPHOLZ, M., SUCK, D. & BOTTCHER, B. 2011. Insights into the structure of the CCR4-NOT complex by electron microscopy. *FEBS Lett,* 585**,** 2182-6.
NEWTON, K. & DIXIT, V. M. 2012. Signaling in innate immunity and inflammation. *Cold Spring Harb Perspect Biol,* 4.

NG, E. W., SHIMA, D. T., CALIAS, P., CUNNINGHAM, E. T., JR., GUYER, D. R. & ADAMIS, A. P. 2006. Pegaptanib, a targeted anti-VEGF aptamer for ocular vascular disease. *Nat Rev Drug Discov,* 5**,** 123-32.

NGOC, L. V., WAUQUIER, C., SOIN, R., BOUSBATA, S., TWYFFELS, L., KRUYS, V. & GUEYDAN, C. 2014. Rapid proteasomal degradation of posttranscriptional regulators of the TIS11/tristetraprolin family is induced by an intrinsically unstructured region independently of ubiquitination. *Mol Cell Biol,* 34**,** 4315-28.

NIINUMA, S., FUKAYA, T. & TOMARI, Y. 2016. CCR4 and CAF1 deadenylases have an intrinsic activity to remove the post-poly(A) sequence. *RNA,* 22**,** 1550-9.

PERKINS, D. N., PAPPIN, D. J., CREASY, D. M. & COTTRELL, J. S. 1999. Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis,* 20**,** 3551-67.

PETIT, A. P., WOHLBOLD, L., BAWANKAR, P., HUNTZINGER, E., SCHMIDT, S., IZAURRALDE, E. & WEICHENRIEDER, O. 2012. The structural basis for the interaction between the CAF1 nuclease and the NOT1 scaffold of the human CCR4-NOT deadenylase complex. *Nucleic Acids Res,* 40**,** 11058-72.

PFAFF, J., HENNIG, J., HERZOG, F., AEBERSOLD, R., SATTLER, M., NIESSING, D. & MEISTER, G. 2013. Structural features of Argonaute-GW182 protein interactions. *Proc Natl Acad Sci U S A,* 110**,** E3770-9.

QIU, L. Q., STUMPO, D. J. & BLACKSHEAR, P. J. 2012. Myeloid-specific tristetraprolin deficiency in mice results in extreme lipopolysaccharide sensitivity in an otherwise minimal phenotype. *J Immunol,* 188**,** 5150-9.

REED, S. I. 1980. The selection of S. cerevisiae mutants defective in the start event of cell division. *Genetics,* 95**,** 561-77.

REESE, J. C. 2013. The control of elongation by the yeast Ccr4-not complex. *Biochim Biophys Acta,* 1829**,** 127-33.

REMMERS, E. F., PLENGE, R. M., LEE, A. T., GRAHAM, R. R., HOM, G., BEHRENS, T. W., DE BAKKER, P. I., LE, J. M., LEE, H. S., BATLIWALLA, F., LI, W., MASTERS, S. L., BOOTY, M. G., CARULLI, J. P., PADYUKOV, L., ALFREDSSON, L., KLARESKOG, L., CHEN, W. V., AMOS, C. I., CRISWELL, L. A., SELDIN, M. F., KASTNER, D. L. & GREGERSEN, P. K. 2007. STAT4 and the risk of rheumatoid arthritis and systemic lupus erythematosus. *N Engl J Med,* 357**,** 977-86.

RICHTER, J. D. 1999. Cytoplasmic polyadenylation in development and beyond. *Microbiol Mol Biol Rev,* 63**,** 446-56.

ROMAO, V. C., CANHAO, H. & FONSECA, J. E. 2013. Old drugs, old problems: where do we stand in prediction of rheumatoid arthritis responsiveness to methotrexate and other synthetic DMARDs? *BMC Med,* 11**,** 17.

ROSS, E. A., SMALLIE, T., DING, Q., O'NEIL, J. D., CUNLIFFE, H. E., TANG, T., ROSNER, D. R., KLEVERNIC, I., MORRICE, N. A., MONACO, C., CUNNINGHAM, A. F., BUCKLEY, C. D., SAKLATVALA, J., DEAN, J. L. & CLARK, A. R. 2015. Dominant Suppression of Inflammation via Targeted Mutation of the mRNA Destabilizing Protein Tristetraprolin. *J Immunol,* 195**,** 265-76.

RUSSO, C. J. & PASSMORE, L. A. 2016. Ultrastable gold substrates: Properties of a support for high-resolution electron cryomicroscopy of biological specimens. *J Struct Biol,* 193**,** 33-44.

SAKURAI, S., OHTO, U. & SHIMIZU, T. 2015. Structure of human Roquin-2 and its complex with constitutive-decay element RNA. *Acta Crystallogr F Struct Biol Commun,* 71**,** 1048-54.

SANDUJA, S., KAZA, V. & DIXON, D. A. 2009. The mRNA decay factor tristetraprolin (TTP) induces senescence in human papillomavirus-transformed cervical cancer cells by targeting E6-AP ubiquitin ligase. *Aging (Albany NY),* 1**,** 803-17.

SAVITSKY, P., BRAY, J., COOPER, C. D., MARSDEN, B. D., MAHAJAN, P., BURGESS-BROWN, N. A. & GILEADI, O. 2010. High-throughput production of human proteins for crystallization: the SGC experience. *J Struct Biol,* 172**,** 3-13.

SCHWEDE, A., ELLIS, L., LUTHER, J., CARRINGTON, M., STOECKLIN, G. & CLAYTON, C. 2008. A role for Caf1 in mRNA deadenylation and decay in trypanosomes and human cells. *Nucleic Acids Res,* 36**,** 3374-88.

SEYMOUR, H. E., WORSLEY, A., SMITH, J. M. & THOMAS, S. H. 2001. Anti-TNF agents for rheumatoid arthritis. *Br J Clin Pharmacol,* 51**,** 201-8.

SGROMO, A., RAISCH, T., BAWANKAR, P., BHANDARI, D., CHEN, Y., KUZUOGLU-OZTURK, D., WEICHENRIEDER, O. & IZAURRALDE, E. 2017. A CAF40-binding motif facilitates recruitment of the CCR4-NOT complex to mRNAs targeted by Drosophila Roquin. *Nat Commun,* 8**,** 14307.

SHAH, N. B. & DUNCAN, T. M. 2014. Bio-layer interferometry for measuring kinetics of protein-protein interactions and allosteric ligand effects. *J Vis Exp***,** e51383.

SKEPARNIAS, I., ALPHANASTASAKIS, D., SHAUKAT, A. N., GRAFANAKI, K. & STATHOPOULOS, C. 2017. Expanding the repertoire of deadenylases. *RNA Biol***,** 1-6.

STOECKLIN, G., GROSS, B., MING, X. F. & MORONI, C. 2003. A novel mechanism of tumor suppression by destabilizing AU-rich growth factor mRNA. *Oncogene,* 22**,** 3554-61.

STOECKLIN, G., STUBBS, T., KEDERSHA, N., WAX, S., RIGBY, W. F., BLACKWELL, T. K. & ANDERSON, P. 2004. MK2-induced tristetraprolin:14-3-3 complexes prevent stress granule association and ARE-mRNA decay. *EMBO J,* 23**,** 1313-24.

STOECKLIN, G., TENENBAUM, S. A., MAYO, T., CHITTUR, S. V., GEORGE, A. D., BARONI, T. E., BLACKSHEAR, P. J. & ANDERSON, P. 2008. Genome-wide analysis identifies interleukin-10 mRNA as target of tristetraprolin. *J Biol Chem,* 283**,** 11689-99.

STOWELL, J. A., WEBSTER, M. W., KOGEL, A., WOLF, J., SHELLEY, K. L. & PASSMORE, L. A. 2016. Reconstitution of Targeted Deadenylation by the Ccr4-Not Complex and the YTH Domain Protein Mmi1. *Cell Rep,* 17**,** 1978-1989.

TANG, D., KANG, R., COYNE, C. B., ZEH, H. J. & LOTZE, M. T. 2012. PAMPs and DAMPs: signal 0s that spur autophagy and immunity. *Immunol Rev,* 249**,** 158-75.

TAYLOR, G. A., CARBALLO, E., LEE, D. M., LAI, W. S., THOMPSON, M. J., PATEL, D. D., SCHENKMAN, D. I., GILKESON, G. S., BROXMEYER, H. E., HAYNES, B. F. & BLACKSHEAR, P. J. 1996a. A pathogenetic role for TNF alpha in the syndrome of cachexia, arthritis, and autoimmunity resulting from tristetraprolin (TTP) deficiency. *Immunity,* 4**,** 445-54.

TAYLOR, G. A., THOMPSON, M. J., LAI, W. S. & BLACKSHEAR, P. J. 1995. Phosphorylation of tristetraprolin, a potential zinc finger transcription factor, by mitogen stimulation in intact cells and by mitogen-activated protein kinase in vitro. *J Biol Chem,* 270**,** 13341-7.

TAYLOR, G. A., THOMPSON, M. J., LAI, W. S. & BLACKSHEAR, P. J. 1996b. Mitogens stimulate the rapid nuclear to cytosolic translocation of tristetraprolin, a potential zinc-finger transcription factor. *Mol Endocrinol,* 10**,** 140-6.

THORE, S., MAUXION, F., SERAPHIN, B. & SUCK, D. 2003. X-ray structure and activity of the yeast Pop2 protein: a nuclease subunit of the mRNA deadenylase complex. *EMBO Rep,* 4**,** 1150-5.

TIEDJE, C., DIAZ-MUNOZ, M. D., TRULLEY, P., AHLFORS, H., LAASS, K., BLACKSHEAR, P. J., TURNER, M. & GAESTEL, M. 2016. The RNA-binding protein TTP is a global post-transcriptional regulator of feedback control in inflammation. *Nucleic Acids Res,* 44**,** 7418-40.

TOLIA, N. H. & JOSHUA-TOR, L. 2007. Slicer and the argonautes. *Nat Chem Biol,* 3**,** 36-43.

UKLEJA, M., CUELLAR, J., SIWASZEK, A., KASPRZAK, J. M., CZARNOCKI-CIECIURA, M., BUJNICKI, J. M., DZIEMBOWSKI, A. & M. VALPUESTA, J. 2016a. The architecture of the Schizosaccharomyces pombe CCR4-NOT complex. *Nat Commun,* 7.

UKLEJA, M., CUELLAR, J., SIWASZEK, A., KASPRZAK, J. M., CZARNOCKI-CIECIURA, M., BUJNICKI, J. M., DZIEMBOWSKI, A. & VALPUESTA, J. M. 2016b. The architecture of the Schizosaccharomyces pombe CCR4-NOT complex. *Nat Commun,* 7**,** 10433.

VAN DUYNE, G. D. 2001. A structural view of cre-loxp site-specific recombination. *Annu Rev Biophys Biomol Struct,* 30**,** 87-104.

VAN ETTEN, J., SCHAGAT, T. L., HRIT, J., WEIDMANN, C. A., BRUMBAUGH, J., COON, J. J. & GOLDSTROHM, A. C. 2012. Human Pumilio proteins recruit multiple deadenylases to efficiently repress messenger RNAs. *J Biol Chem,* 287**,** 36370-83.

WANG, H., MORITA, M., YANG, X., SUZUKI, T., YANG, W., WANG, J., ITO, K., WANG, Q., ZHAO, C., BARTLAM, M., YAMAMOTO, T. & RAO, Z. 2010. Crystal structure of the human CNOT6L nuclease domain reveals strict poly(A) substrate specificity. *EMBO J,* 29**,** 2566-76.

WEI, Z. R., LIANG, C., FENG, D., CHENG, Y. J., WANG, W. M., YANG, D. J., WANG, Y. X. & CAI, Q. P. 2016. Low tristetraprolin expression promotes cell proliferation and predicts poor patients outcome in pancreatic cancer. *Oncotarget,* 7**,** 17737-50.

WELLS, S. E., HILLNER, P. E., VALE, R. D. & SACHS, A. B. 1998. Circularization of mRNA by eukaryotic translation initiation factors. *Mol Cell,* 2**,** 135-40.

WIEDERHOLD, K. & PASSMORE, L. A. 2010. Cytoplasmic deadenylation: regulation of mRNA fate. *Biochem Soc Trans,* 38**,** 1531-6.

WILLIAMS, R. O. 2012. What Have We Learned about the Pathogenesis of Rheumatoid Arthritis from TNF-Targeted Therapy? *ISRN Immunology,* 2012**,** 15.

WOLF, J., VALKOV, E., ALLEN, M. D., MEINEKE, B., GORDIYENKO, Y., MCLAUGHLIN, S. H., OLSEN, T. M., ROBINSON, C. V., BYCROFT, M., STEWART, M. & PASSMORE, L. A. 2014. Structural basis for Pan3 binding to Pan2 and its function in mRNA recruitment and deadenylation. *EMBO J,* 33**,** 1514-26.

YAMAJI, M., JISHAGE, M., MEYER, C., SURYAWANSHI, H., DER, E., YAMAJI, M., GARZIA, A., MOROZOV, P., MANICKAVEL, S., MCFARLAND, H. L., ROEDER, R. G., HAFNER, M. & TUSCHL, T. 2017. DND1 maintains germline stem cells via recruitment of the CCR4-NOT complex to target mRNAs. *Nature,* 543**,** 568-572.

YASUKAWA, T., KANEI-ISHII, C., MAEKAWA, T., FUJIMOTO, J., YAMAMOTO, T. & ISHII, S. 1995. Increase of solubility of foreign proteins in Escherichia coli by coproduction of the bacterial thioredoxin. *J Biol Chem,* 270**,** 25328-31.

ZHAO, W., LIU, M., D'SILVA, N. J. & KIRKWOOD, K. L. 2011. Tristetraprolin regulates interleukin-6 expression through p38 MAPK-dependent affinity changes with mRNA 3' untranslated region. *J Interferon Cytokine Res,* 31**,** 629-37.

ZHENG, X., DUMITRU, R., LACKFORD, B. L., FREUDENBERG, J. M., SINGH, A. P., ARCHER, T. K., JOTHI, R. & HU, G. 2012. Cnot1, Cnot2, and Cnot3 maintain mouse and human ESC identity and inhibit extraembryonic differentiation. *Stem Cells,* 30**,** 910-22.

ZUBIAGA, A. M., BELASCO, J. G. & GREENBERG, M. E. 1995. The nonamer UUAUUUAUU is the key AU-rich sequence motif that mediates mRNA degradation. *Mol Cell Biol,* 15**,** 2219-30.