Investigating mechanisms of transcriptional interference in *Schizosaccharomyces pombe*

A dissertation presented for the degree of D. Phil.

Beth Watts

Department of Biochemistry

Linacre College

University of Oxford

Trinity Term 2015

ABSTRACT

Investigating mechanisms of transcriptional interference in *Schizosaccharomyces pombe*

Eukaryotic cells transcribe a vast array of non-coding RNAs, most of which have not been assigned a functional role. The work presented here reveals a novel mechanism of transcriptional repression that is mediated by the non-coding RNA *prt* (*pho1*-repressing transcript). The *prt* transcript is shown to recruit a histone deacetylase, Clr3, to repress *pho1*. This gene encodes a secreted acid phosphatase essential for phosphate acquisition in fission yeast. In the presence of phosphate, *prt* is produced from an upstream promoter and leads to silencing of *pho1*. Thus far, this has been explained by *prt* transcription leading to deposition of repressive methylation over the locus. However, this explanation is known to be incomplete since deletion of the only known histone methyltransferase does not lead to *pho1* induction comparable to deletion of the *prt* promoter. This suggests that another mechanism must be involved in mediating transcriptional interference via non-coding transcription.

In the present study the putative ncRNA-binding protein Seb1, together with the chromatin modifying complex SHREC, is demonstrated to associate with *prt* to elicit silencing of *pho1* by a mechanism that is independent of H3K9 methylation and instead relies on deacetylase activity provided by the Clr3 component of SHREC. These data reveal a previously uncharacterised layer of ncRNA-mediated gene regulation and provide important conceptual advances in understanding the mechanisms governing the phenomenon known as transcriptional interference.

Beth Watts, Linacre College, Trinity Term 2015

ACKNOWLEDGEMENTS

First and foremost I wish to express my sincere thanks to my supervisor, Dr Lidia Vasiljeva. Her knowledge has been invaluable, her enthusiasm inspiring, and her guidance greatly valued throughout the course of this DPhil thesis.

I am grateful to my fellow labmates Cornelia Kilchert, Monica Passoni, Dong-Hyuk Heo, Sina Wittmann, Tea Kecman, and Sneha Shah for stimulating discussions (both scientific and otherwise), and for being great co-workers; it has been a true pleasure. I would also like to mention past lab members Jessica Teal and Luke Porter if only for their (thoroughly appreciated) comic relief. I am indebted to James Rhodes and Sina Wittmann for feedback on this thesis; I promise to repay the favour when your turn comes.

I am also thankful to Linacre College for providing a stimulating intellectual and social environment. And for accepting applications from so many students who would become such good friends. I must also mention LCBC for, not only getting me up in the mornings, but for maintaining my health and happiness these past four years.

To my parents, my grandparents, and my brothers - thank you for your love, support and unwavering (here's looking at you, Nan!) belief in me. We may not be the kind to say it often but I could not have wished for a more caring, genuine family. Thank you for everything.

"Yes, that's it! Said the Hatter with a sigh, it's always tea time."
- Lewis Carroll

To my parents

ABBREVIATIONS

bp/kb	base pairs/kilobases
ChIP	Chromatin Immunoprecipitation
CPA	Cleavage and Polyadenylation
CPF	Cleavage and Polyadenylation Factor
cRACE	circularized Rapid Amplification of cDNA Ends
CTD	C-terminal domain
DNA	Deoxyribonucleic acid
DTT	Dithiothreitol
EMMG	Edinburgh Minimal Media Glutamate
H3	Histone 3
hr	hour/hours
kDa	Kilo Daltons
lncRNA	long non-coding RNA
mRNA	messenger RNA
nc	Non-coding
ncRNA	Non-coding RNA
nt	nucleotides
pA	Polyadenylation and cleavage site
PCR	Polymerase Chain Reaction
PHO pathway	Phosphate-responsive signalling pathway
Pi	Inorganic phosphate
PNPP	para-Nitrophenylphosphate
Pol II	RNA Polymerase II
prt	*pho1*-repressing transcript

qRT-PCR	quantitative Reverse Transcription PCR
RDRC	RNA-dependent RNA Polymerase Complex
RISC	RNA-induced silencing complex
RITS	RNA-induced transcriptional silencing
RNA	Ribonucleic acid
RNAi	RNA interference
RNase	Ribonuclease
rRNA	ribosomal RNA
RT-PCR	Reverse Transcriptase PCR
S. cerevisiae	*Saccharomyces cerevisiae*
S. pombe	*Schizosaccharomyces pombe*
snoRNA	Small nucleolar RNA
snRNA	Small nuclear RNA
TAP	Tandem Affinity Purification
TI	Transcriptional Interference
tRNA	transfer RNA
TSS	Transcription Start Site
µM	micro molar
µg	micro gram
µl	micro litre
WT	Wild-type
YES	Yeast Extract media with Supplements

TABLE OF CONTENTS

Abstract **II**

Acknowledgements **III**

Abbreviations **V**

Table of contents **VII**

List of figures **XI**

List of tables **XIII**

Chapter 1: Introduction

1.1 The central dogma of molecular biology 1

1.2 Non-coding transcription 2

1.3 Non-coding RNAs and their regulation 6

1.4 Transcriptional regulation by lncRNAs 9

1.4.1 Regulation in *trans* 10

1.4.2 Regulation in *cis* 11

1.5 Heterochromatin 12

1.5.1 Constitutive heterochromatin - regulation by RNAi 13

1.5.2 Facultative heterochromatin - regulation of meiotic 16

1.5.3 Transient heterochromatin 17

1.6 Transcriptional gene silencing at heterochromatic loci 18

1.7 ncRNA-mediated regulation of the *pho1* gene in response to phosphate 20

1.7.1 The *S. cerevisiae* PHO response 20

1.7.2 The *S. pombe* PHO response 23

1.8 The RNA polymerase II C-terminal domain (CTD) 27

1.9 CTD-interactions 29

1.10 Nrd1 and non-coding transcription 31

1.11 Seb1 31

1.12 Objectives of this work 33

Chapter 2: Materials and Methods

2.1 Media and growth conditions 35
2.2 Yeast transformation 35
2.3 CTD-Peptide binding experiments 36
2.4 Immunoblotting 36
2.5 Purification of Seb1-TAP 37
2.6 Strain construction 38
2.7 Northern blotting 38
2.8 ChIP experiments 39
2.9 RNase ChIP 40

Chapter 3: A ncRNA regulates expression of *pho1* by an H3K9 methylation-independent mechanism

3.1 Introduction 46
3.2 Results 49
3.2.1 Clr3 HDAC contributes to *pho1* silencing 49
3.2.2 Clr3 is co-transcriptionally recruited to the *prt* region 51
3.2.3 Lack of Clr3 correlates with elevated Pol II occupancy and H3K14 acetylation 53
3.2.4 Seb1 is recruited to the *pho1* non-coding transcript and the 3' end of protein-coding genes 57
3.2.5 Clr3 and Clr4 function independently to silence *pho1* 58
3.3 Discussion 60

Chapter 4: A ncRNA is required for Seb1 and Clr3 recruitment to the *pho1* locus

4.1 Introduction 62
4.2 Results 65
4.2.1 Seb1 interactions 65
4.2.2 Identification of CTD-interacting proteins in *S. pombe* 67
4.2.3 Seb1 physically associates with the RNA polymerase II (Pol II) C-terminal domain (CTD) 72

4.2.4 Non-coding transcription is required for Seb1 and Clr3 recruitment 75
4.2.5 *prt* RNA is required for *pho1* silencing 80
4.3 Discussion 83

Chapter 5: Seb1-dependent silencing is a general mechanism of gene repression

5.1 Introduction 86
5.2 Results 88
5.2.1 Clr3 represses another protein-coding gene, *zrt1* 88
5.2.2 Seb1 and Clr3 are recruited to the *zrt1* locus 90
5.2.3 Recruitment of Clr3 to *zrt1* is RNA-dependent 90
5.2.4 Seb1 elicits *tgp1* silencing independently of SHREC 92
5.2.5 Seb1 elicits *tgp1* silencing independently of HDACs 95
5.2.6 Seb1 but not Clr3 is recruited to *tgp1* 95
5.3 Discussion 99

Chapter 6: General Discussion

6.1 A new mechanism of transcriptional interference 102
6.1.1 Clr3 and Seb1 contribute to transcriptional interference at *pho1* 102
6.1.2 Seb1 interacts with RNA and the CTD 106
6.1.2.1 RNA interactions 106
6.1.2.2 CTD interactions 107
6.1.3 Transcriptional repression through non-coding transcription/RNA 109
6.1.4 ncRNA-mediated TI functions at other loci 111
6.1.5 Are more players of ncRNA-mediated TI yet to be identified? 112
6.1.6 Biosynthetic pathways and non-coding transcription 114
6.2 The role of CID proteins in termination of non-coding and protein-coding transcripts 115
6.3 A dual role for Seb1 118

6.3.1 Seb1 recruitment to protein-coding vs. non-coding genes 118
6.3.2 Other Seb1 protein interactions and possible roles 120
6.3.3 Concluding remarks 122

References 123

LIST OF FIGURES

Figure 1-1: The RNAi pathway for heterochromatin assembly

Figure 1-2: The *S. cerevisiae* PHO pathway

Figure 1-3: lncRNA-mediated regulation of the *pho1* gene

Figure 1-4: The carboxyl-terminal domain (CTD) phosphorylation profile across *S. cerevisiae* protein-coding genes.

Figure 1-5: CTD-interacting domain (CID)

Figure 1-6: CID sequences

Figure 2-1: *pho1* is sensitive to the HDAC inhibitor TSA

Figure 2-2: Clr3 is the HDAC that contributes to *pho1* silencing

Figure 2-3: Clr3 is recruited to the *pho1* locus

Figure 2-4: Lack of Clr3 correlates with elevated Pol II occupancy and H3K14 acetylation levels at the *pho1* locus

Figure 2-5: Pol II occupancy and H3K14 acetylation levels are partially elevated in *seb1-1* mutant cells at the *pho1* locus

Figure 2-6: The contribution of SHREC on *pho1* silencing

Figure 2-7: Seb1 is recruited to the *pho1* non-coding transcript and the 3' end of protein-coding genes

Figure 2-8: Clr3 acts independently of Clr4 to promote *pho1* silencing

Figure 3-1: Seb1 purification

Figure 3-2: Verification of CTD-binding assay

Figure 3-3: Silver stain analysis of proteins purified from CTD peptides

Figure 3-4: Mass spectrometry analysis

Figure 3-5: Seb1 physically associates with the RNA polymerase II (Pol II) C-terminal domain (CTD)

Figure 3-6: Non-coding RNA is required for Seb1 and Clr3 recruitment

Figure 3-7: *prt* is required for Clr3 recruitment

Figure 3-8: Clr3 recruitment is phosphate-dependent

Figure 3-9: *prt* RNA is required for *pho1* silencing

Figure 4-1: Clr3 represses another protein-coding gene, *zrt1*

Figure 4-2: Seb1 is recruited to the *zrt1* locus

Figure 4-3: Recruitment of Clr3 to *zrt1* is RNA-dependent

Figure 4-4: Seb1 elicits *tgp1* silencing independently of SHREC

Figure 4-5: Seb1 elicits *tgp1* silencing independently of HDACs

Figure 4-6: Seb1, but not Clr3, is recruited to *tgp1*

Figure 4-7: Seb1 recruitment to *tgp1* is RNA-dependent

Figure 5-1: Proposed models for transcriptional interference

Figure 5-2: *prt* ncRNA mediates *pho1* repression via Seb1 and Clr3

LIST OF TABLES

Table 1: Strains used in this study

Table 2: Oligonucleotides used in this study

1. Introduction

1.1 The central dogma of molecular biology

The 19th century was a period marked by many profound and influential discoveries in biology. Indeed, the development of cell theory by Schleiden and Schwann, Pasteur's refutation of the doctrine of spontaneous generation, not to mention Darwin's theories of evolution (Darwin, 1859) and Mendel's discovery of the laws of heredity (Mendel, 1866) led to the emergence of various specialised biological disciplines.

Often overlooked as one of the seminal scientific advances of this age, however, is the work of one Friedrich Miescher who, in 1868, isolated from the nuclei of white blood cells a phosphate-rich chemical which he termed *nuclein* (Miescher, 1871). Unfortunately for Miescher eponyms do not usually apply to molecules but, more than half a century after his discovery, the importance of what are now called nucleic acids was becoming clear. In determining the structure of DNA Francis Crick and James Watson identified the carrier of genetic information and proposed a model by which this information could be propagated (Watson and Crick, 1953). The problem of how the genetic code is expressed remained a mystery.

In 1956, Francis Crick formulated the central dogma of molecular biology: that DNA makes RNA makes proteins (Crick, 1956). The process by which DNA is copied to RNA is now known as transcription and the 'reading' of RNA into a protein is translation. Transcription takes place in two broad steps. Firstly, a primary transcript is formed, involving RNA polymerase enzymes in a process that relies on Watson-

Crick base pairing. Secondly, the RNA can be edited by various processes including polyadenylation, capping, and splicing to produce the desired RNA molecule e.g. ribosomal RNA (rRNA), messenger RNA (mRNA) and transfer RNA (tRNA). Eukaryotes have multiple types of RNA polymerases, each responsible for synthesis of a distinct subset of RNA. The most studied of these is RNA polymerase II (Pol II) which synthesises precursors of mRNA. The next step in the flow of genetic information involves translation of mRNAs into proteins. By characterising an alanine tRNA, Robert W. Holley further revealed the link between DNA and protein synthesis (Holley et al., 1965).

While RNAs very often have coding potential there are some which do not. These are the non-coding RNAs (ncRNAs), many of which are also synthesised by Pol II. The development of technologies such as high resolution tiling arrays and next generation sequencing has revealed that eukaryotic genomes are in fact pervasively transcribed, and that a large percentage of this represents ncRNA.

1.2 Non-coding transcription

There has been growing interest in researching the putative functions of the fraction of the genome that has, historically, been referred to as "junk DNA". While it has been known for some time that pervasive transcription by Pol II is a general feature in all domains of life, the reasons behind Pol II's association with and transcription of intergenic regions are, for the most part, unknown. However, when such transcription overlaps with, or occurs in close proximity of, coding genes there is potential for so called 'transcriptional interference' (TI). This process describes a regulatory

phenomenon in which one transcriptional process has a direct and *in cis* oppressive effect on another.

First described by Prescott and Proudfoot (2002) who found, upon removal of gene terminators, a severe effect on the transcription processes of two convergent genes, TI appears to be widespread in biology; examples have been identified in bacteria, yeast, and higher eukaryotes, including *Drosophila*, mouse, and human (Martens et al., 2004; Wang et al., 1998; Boussadia et al., 1997; Whitelaw and Proudfoot, 1986). TI provides an important platform for gene regulation. Importantly, TI is a description of a process and does not provide an explanation for the governing mechanism(s), of which only a few have been identified and studied.

Those mechanisms which have been characterised come from both *in vivo* and *in vitro* work and are implemented either in the initiation or elongation phases of gene transcription. Depending on the arrangement of interfering promoters, TI can arise from (i) promoter competition (in which one promoter-bound Pol II blocks binding of another Pol II at a second promoter), (ii) promoter occlusion (whereby one Pol II (from an upstream promoter) passes over a promoter and in so doing precludes its occupation by a second Pol II), or (iii) collision (where two converging polymerases physically collide). Additionally, other less well characterised models which explain the possible mechanisms, called ‘roadblock’ and ‘sitting duck’ interference, have been identified (see Shearwin et al., 2005 for review). Each of these have been studied across a range of different organisms and genes (in both naturally occurring and artificially generated arrangements), some examples of which will now be considered.

Work performed in the Proudfoot laboratory has involved use of the tandemly arranged genes *GAL10* and *GAL7* as a model system for dissecting the mechanisms of TI in the budding yeast *Saccharomyces cerevisiae* (*S. cerevisiae*). Inefficient termination of the upstream *GAL10* gene perturbs transcription initiation of the downstream gene, *GAL7*. In this case, TI blocks access of the Gal4p transcription factor to the *GAL7* promoter region, thus inhibiting any transcriptional activation (Greger et al., 2000). TI can also occur when the orientation of the same two genes is artificially engineered such that they are in a convergent arrangement (Prescott and Proudfoot, 2001). Here, a transcriptional collision effect occurs whereby elongation complexes advancing from opposing sides either are stalled and prevented from moving forward, or are knocked off the template completely. In either case this effect can be alleviated by suppressing transcription on one of the two strands. Collisions between converging polymerases lead to untimely termination and this perhaps provides an explanation for the apparent evolutionary pressure against this gene orientation in yeast (Marin et al., 2004).

A well characterised natural example of where transcription of a ncRNA functions via TI to regulate gene expression is *SRG1* (*SER3* regulatory gene 1). *SRG1* regulates levels of the *S. cerevisiae* gene, *SER3*, which encodes a phosphoglycerate dehydrogenase involved in serine biosynthesis. Chromatin immunoprecipitation (ChIP) experiments initially performed in the Winston laboratory revealed that, even under repressive conditions (presence of serine), the 5' regulatory region upstream of the *SER3* gene promoter is bound by factors associated with active transcription (including TATA-binding protein (TBP), Pol II, and capping proteins) (Martens et al., 2004). Further work identified that the reason for the presence of these factors is that

the non-coding transcript, *SRG1*, is produced from this upstream region. Significantly, *SRG1* was shown to cause TI at *SER3* and have an essential role in repressing its expression. During *SRG1* transcription, two histone chaperones, Spt6 and Spt16, which associate with elongating Pol II, reposition nucleosomes over the *SER3* promoter (Thebault et al., 2011). The alteration in chromatin dynamics at this region occludes access of transcription factors and thus leads to *SER3* repression. On the other hand, under conditions of serine deficiency, *SRG1* transcription is down-regulated, leading to an active nucleosomal structure and hence *SER3* activation.

In addition to impacting on gene expression through the repositioning of nucleosomes, recent studies of the *GAL10* ncRNA in *S. cerevisiae* has revealed that ncRNA transcription can also modify the makeup of certain histone modifications (Houseley et al., 2008; Pinskaya et al., 2009). The *GAL10* ncRNA is transcribed antisense from the 3' end of the *GAL10* gene, extends across the bidirectional *GAL1-10* promoter, and through *GAL1* in the sense orientation to its 3' end. Under repressed and non-induced conditions, Reb1-driven transcription of *GAL10* ncRNA establishes Set2-mediated H3K36 trimethylation over the *GAL1-10* locus which, in turn, recruits the histone deacetylase (HDAC) complex Rpd3S. Rpd3S deacetylates the entire locus which leads to *GAL1* and *GAL10* attenuation. TI-directed chromatin re-modification therefore provides another platform for gene regulation.

Whilst *SRG1* transcription over the *SER3* promoter alters chromatin dynamics at that region to repress transcription, there are cases where the opposite is also true. The fission yeast *Schizosaccharomyces pombe* (*S. pombe*) gene *fbp1* is activated by glucose starvation. Here, upstream transcription of ncRNAs induces chromatin

remodelling at the mRNA transcription start site to allow access of Pol II and transcription factors. Insertion of a transcriptional terminator to block upstream ncRNA transcription prevents chromatin taking on an open conformation which, in turn, leads to loss of recruitment of transcription factors to the *fbp1* promoter, and thus attenuation of the mRNA (Hirota et al., 2008). This case highlights the need to consider TI on a case by case basis because regulation can mean either repression or activation and can be established in a whole host of different ways.

The examples described above (*GAL10/GAL7*, *SER3*, *GAL1-10*, and *fbp1*) represent a common theme in the regulation of genes by non-coding transcription and that is that these encode proteins involved in biosynthetic pathways. Orchestrating and fine-tuning gene expression patterns is essential if cells are to respond to nutrient fluctuations, and it is becoming clear that transcription of ncRNAs, which can impact expression in a multitude of ways, are essential for this rapid and efficient response to occur.

1.3 Non-coding RNAs and their regulation

ncRNAs can be grouped according to various aspects of their structure, function, and localisation within the cell. Some, which perhaps constitute the most conserved and essential of ncRNAs, are the highly abundant tRNAs, rRNAs, and small nucleolar RNAs (snoRNAs) all with well characterised roles in protein translation. Others have seemingly less obvious or understood roles to play and include microRNAs (miRNAs) (with functions in regulating gene expression by directing target mRNAs for degradation or translational repression), small interfering RNAs (siRNAs) (double

stranded RNAs of 20-25 bp involved in the RNAi pathway of gene silencing), piwi-interacting RNAs (piRNAs) (through interactions with Piwi proteins are involved in silencing of retrotransposons), enhancer RNAs (eRNAs) (possibly involved in transcriptional regulation) and the Pol II transcribed long ncRNAs (lncRNAs). One major emerging function of lncRNAs is in the regulation of gene expression (very often by chromatin re-modification - see section 1.4 below). However, the mechanistic details underlying this process have not been fully elucidated, not least because lncRNAs are typically very unstable. Such instability is, in part, due to their rapid degradation by the nuclear exosome complex.

The nuclear exosome is a multi-protein complex and major component of the RNA surveillance machinery that is involved in the regulation of all cellular RNAs. It functions in quality control by degrading many different aberrant RNA species, in addition to controlling mRNA levels via its 3'-5' exoribonuclease activity. The complex also has a role in the precise 3' processing of certain stable RNAs, such as the 5.8S rRNA, as well as sn/snoRNAs and tRNAs (Zanchin and Goldfarb, 1999; van Hoof et al., 2000; Allmang et al., 1999a; Gudipati et al., 2012)

The core of the yeast exosome is a barrel-like structure that consists of nine different subunits arranged into a six-membered ring and trimeric cap. None of these subunits possess nuclease activity; instead, this is derived from core-associated catalytic subunits Rrp44 (also known as Dis3) and Rrp6. Rrp44 possesses both endo- and 3'-5' exonuclease activities which, together, are thought to coordinate degradation of RNAs that have been fed through the pore of the inactive core (Schneider et al., 2009; Allmang et al., 1999b). Rrp6 is a 3'-5' exonuclease and binds directly to the core near

the top of the barrel (Wasmuth et al., 2014). In yeast, unlike Rrp44, Rrp6 is only connected to the nuclear exosome and can even function independently of the core in the 3' end processing of snoRNAs (Callahan and Butler, 2008).

1.4 Transcriptional regulation by lncRNAs

lncRNAs are ubiquitous across the taxa and, quite arbitrarily, are defined as Pol II transcripts larger than 200 bp. Despite their prevalence, the question of whether they are themselves functionally important or are merely the product of spurious transcriptional "noise" that is ultimately degraded by RNA surveillance machineries continues to be the subject of intense debate. Whilst the vast majority have yet to be fully characterised, work in various systems has uncovered an extensive repertoire of lncRNA roles including, but by no means limited to, roles in X-chromosome inactivation, high-order chromosome dynamics, nuclear organisation and cellular trafficking (Amaral and Mattick, 2008).

Most of the studied cases in which non-coding transcription has a regulatory role involve the process of transcription rather than the resulting RNA. Nevertheless there are a few examples where the RNA itself is important. The modes of RNA action can take several forms - firstly, they can function either in *cis* or in *trans*. The former describes a case where a lncRNA exerts its effects to the region from which it was transcribed, whereas the latter applies to a lncRNA which has the potential to affect expression of a gene(s) at some distance from the region where the lncRNA was produced, for example on a different chromosome. Furthermore, lncRNAs may negatively or positively regulate transcription. Some of these different modes of lncRNA-mediated regulation (specifically via the RNA product as opposed to its transcription) will now be considered in the context of specific examples.

1.4.1 Regulation in *trans*

Since regulation in *trans* applies to lncRNAs affecting expression of genes at a distance from their transcription site it therefore follows that they must act via their RNA product. lncRNAs can specifically target certain loci - as is the case with the HOX antisense intergenic RNA (HOTAIR). HOTAIR establishes epigenetic silencing of HOXD genes by acting as a scaffold for the repressive H3K27-trimethylase PRC2 (Polycomb repressive complex 2) and lysine-specific demethylase LSD1 (Rinn et al., 2007). lncRNA-mediated regulation can also have a more global effect on transcriptional output by preventing the binding or inhibiting the activity of transcription factors or Pol II. Examples include human 7SK RNA (Peterlin et al., 2012) and mouse B2 RNA (Espinoza et al., 2007).

One particularly well characterised example of ncRNA-mediated repression is at the budding yeast *PHO84* gene. *PHO84* antisense RNA transcripts (the production of which is stimulated by Set1 mediated H3K4 methylation) are stabilised upon loss of the exosome subunit Rrp6 - a phenotype which is also observed during chronological ageing when Rrp6 association with *PHO84* is impaired (Camblong et al., 2007). Transcriptional gene silencing of *PHO84* can actually be induced both in *cis* and in *trans* by these antisense RNAs. By comparison to that described for *cis* silencing (see below), in *trans* silencing is still very poorly understood but appears to function independently of Hda1 deacetylation; and consequently is distinct from *cis* silencing (Camblong et al. 2009).

Another example in *S. cerevisiae* is the antisense *TY1* CUT whose transcription is initiated from within the Ty1 retrotransposon (Berretta et al., 2008). The *TY1* CUT triggers repression of Ty1 in *trans* by acting on the de novo *TY1* RNA. This CUT-dependent silencing is, by a currently unknown mechanism, mediated by histone deacetylation and requires Set1-dependent histone methylation.

1.4.2 Regulation in *cis*

In contrast to regulation in *trans*, *cis*-action means that the lncRNA functions in the vicinity of its transcription site. Therefore, as well as acting through the RNA product, regulation can also involve the process of transcription. In *cis* silencing in particular has been described for various mammalian lncRNAs. *Air* (Antisense to *Igf2r* RNA) is a 108 kb lncRNA transcribed from mouse chromosome 17. The transcript is produced from an antisense promoter found within an intron of the *Igf2r* (insulin-like growth factor type 2 receptor) gene (Wutz et al., 1997). Another lncRNA, *Kcnq1ot1*, of approximately 91 kb is transcribed from an intron in the *Kcnq1* (Potassium voltage-gated channel subfamily KQT member 1) gene (Smilinich et al., 1999). Both *Air* and *Kcnq1ot1* are expressed exclusively from the paternal allele and are responsible for silencing the expression of a cluster of neighbouring imprinted genes. Repression of genes is partly achieved via recruitment of the H3K9 methyltransferase G9a which leads to remodelling of the chromatin structure (Nagano et al., 2008; Wagschal et al., 2008). In the case of *Kcnq1ot1*, PRC2 can also be recruited via its EZH2 component through an interaction with a highly conserved region at the 5' end of the lncRNA (Mohammad et al., 2008). This promotes trimethylation of H3K27 and thus silencing of specific genes.

With regards to functioning through the RNA product one of the best characterised examples is that of X inactivation by mammalian *Xist* (X-inactive specific transcript) RNA. One of the two copies of the X chromosomes of female mammals is coated by *Xist* which, in turn, recruits epigenetic modifiers leading to increases in DNA methylation, decreases in acetylation (the two classical hallmarks of repressive heterochromatin – see section 1.5 below) and hence inactivation of the chromosome (see for example Wutz et al., 2011). A recent report has identified the human RNA-binding protein SHARP as an interactor of the *Xist* lncRNA. SHARP is essential for silencing and helps exclude Pol II from the inactive X chromosome through activation of the histone deacetylase HDAC3 (McHugh et al., 2015).

As mentioned above, silencing of the *PHO84* gene can also be induced in *cis*. In this case, stabilisation of antisense RNAs is paralleled by H3K18 deacetylation by the class II HDAC, Hda1. Despite antisense RNAs spanning the whole of *PHO84* – from 3′ end to promoter - histone deacetylation is limited only to the promoter and 5′ end. This specifically localised histone deacetylation may be required for maintaining antisense transcription and repression of *PHO84* (Camblong et al. 2009).

cis-regulation by TI is another option by which lncRNAs can exert their effects, several examples of which (including *SER3*, *GAL1-10*, and *fbp1*) have already been considered.

1.5 Heterochromatin

As described, ncRNA-mediated repression very often correlates with the presence of heterochromatin - the extreme case being mammalian *Xist* lncRNA which serves to

silence the whole chromosome from which it is transcribed. In contrast to the usually highly transcribed, gene-rich euchromatin, heterochromatin is a very compact chromatin structure. It is often categorised as being of either the constitutive or facultative form. The former of these describes a case where heterochromatin cannot be reverted back to the active, euchromatic state and is found especially at highly repetitive DNA sequences such as centromeres, telomeres, and the mating-type locus. Conversely, facultative heterochromatin, which is less well studied, describes a highly regulated form of chromatin that is readily able to revert from the silent to active forms and is typically associated with developmentally regulated genes.

Heterochromatin of both forms is characterised by hypoacetylation of histones and hypermethylation of H3K9. This leads to recruitment of the HP1 proteins Swi6 and Chp2, both of which are required for heterochromatic gene silencing. Nucleation of heterochromatin can occur via the action of several different specificity determinants - through DNA-binding proteins, small RNAs (as in the case of RNAi), or lncRNAs acting as scaffolds for recruitment of chromatin modifiers.

1.5.1 Constitutive heterochromatin - regulation by RNAi

Repression of heterochromatic regions is mediated by RNAi (Fig. 1-1). This pathway is conserved from fungi through plants, both in functional terms and of having a common core of proteins. *S. pombe* has proven a powerful model for understanding the pathway, which is triggered by ncRNA transcripts that are converted into dsRNAs through the action of the RNA-directed RNA polymerase complex (RDRC). These dsRNAs are cleaved by the ribonuclease III Dicer (Dcr1) to produce small interfering

RNAs (siRNAs) of 21-23 nucleotides (Verdel et al., 2004; Buhler and Moazed 2007; Colmenares et al., 2007). siRNAs are incorporated, together with Argonaute (Ago1), the chromodomain protein Chp2, and Tas3 into the RNA-induced transcriptional silencing (RITS) complex whose role it is to promote transcriptional silencing by repressive histone methylation (Verdel et al., 2004). The siRNAs guide Ago1 and the associated RITS complex components to complementary nascent RNAs where it mediates recruitment of the Clr4-methyltransferase complex (CLRC) to methylate H3K9 (Hong et al., 2005; Gerace et al., 2010). The repressive H3K9me mark, in turn, acts as a landing platform for Swi6 and Chp2 proteins (Buhler and Moazed 2007). RITS and Clr4 are physically associated and operate interdependently in a self-reinforcing feedback loop which involves siRNAs produced by the RNAi machinery feeding back to facilitate additional recruitment of RITS.

An intriguing facet of RNAi-dependent heterochromatin assembly is that, somewhat paradoxically, in order for nucleation, propagation and maintenance of a repressive state, siRNAs originating from the region must first be transcribed. The antisilencing factor Epe1, a JmjC-domain protein, counters repressive heterochromatin and allows Pol II access to repeated regions. Intriguingly, Epe1 is recruited to heterochromatin by Swi6 (Zofall and Grewal, 2006).

Heterochromatin is not confined to the region from which it is nucleated. Indeed, spreading of heterochromatin can occur through multiple mechanisms. Central to this process is Swi6 which, through a conformational change, switches from an auto-inhibited state to a spreading competent state. Swi6 monomers contain a sequence that is a histone-mimic and are able to block recognition of H3K9 methyl marks by other

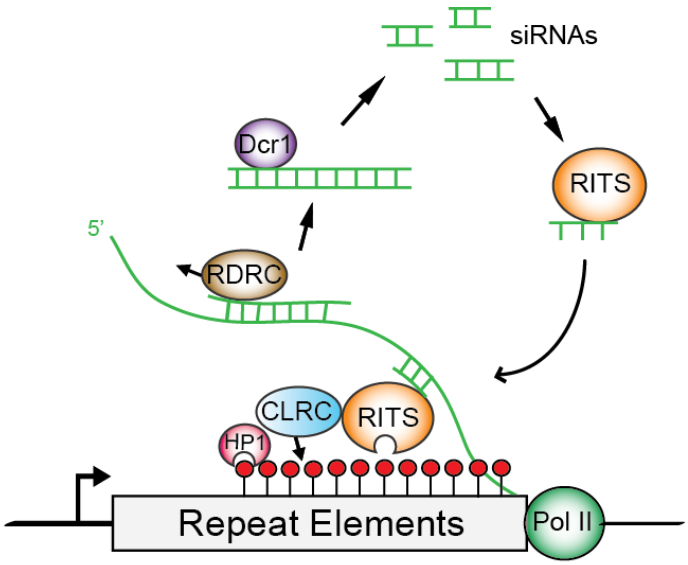


Figure 1-1: The RNAi pathway for heterochromatin assembly
The RNA-directed RNA polymerase (RDRC) promotes synthesis of dsRNAs which are subsequently cleaved by Dicer (Dcr1) to produce siRNAs. These guide the RNA-induced transcriptional silencing (RITS) complex to the nascent transcript where it mediates recruitment of the Clr4 methyltransferase (CLRC). Clr4 methylates H3K9 (red circles) which is bound by the HP1 proteins Swi6 and Chp2.

Swi6 monomers. This auto-inhibition is alleviated by recognition of both H3K9me and nucleosomal DNA, driving the spread of heterochromatin to surrounding regions (Canzio et al., 2013). The RITS component Tas3 is also implicated in heterochromatin propagation. Tas3 contains an alpha helical motif at its C terminus which is able to oligomerise into a polymer and establish RITS spreading (Haitao et al., 2009).

1.5.2 Facultative heterochromatin - regulation of meiotic genes

RNAi is the major pathway for heterochromatin assembly at most genomic loci. However, loss of the RNAi machinery does not lead to complete abolition of H3K9me (Reyes-Turcu et al., 2011), meaning that other, RNAi-independent, pathways for heterochromatin formation exist. One example of loci where H3K9me is not entirely abolished with the loss of RNAi components is meiotic genes. Unlike constitutive heterochromatin however, levels of H3K9 methylation at meiotic genes do not seem to correlate with restriction of Pol II access or expression levels (Egan et al., 2014). Instead RNA-binding proteins mediate repression at the post-transcriptional level.

Undergoing meiosis to ultimately generate four haploid cells involves the complex coordination of gene regulation at multiple levels, including transcription, RNA processing, and RNA stability. Upon nitrogen starvation, haploid *S. pombe* cells commit to meiosis whereby hundreds of transcripts become functionally expressed, including the master regulators Ste11 and Mei4 (Mata et al., 2002). An intriguing feature of these and other meiotic genes is that they are constitutively transcribed, even in mitotic growth. In these vegetatively growing cells the RNA-binding and

YTH-domain protein Mmi1 recognises the transcripts via their DSR (determinant of selective removal) motifs and, in association with Red1 and Pab2, mediates their degradation by the exosome (St-Andre et al., 2010; Yamanaka et al., 2010; Sugiyama and Sugioka-Sugiyama, 2011). Conversely, upon induction of meiosis, Mmi1 is sequestered to the Mei2 dot (a subnuclear structure of Mei2 protein and non-coding meiRNA) - allowing for stable expression of the meiotic transcripts (Watanabe and Yamamoto, 1994).

Recently, this Mmi1-mediated degradation has also been shown to cooperate with the RNAi pathway described above (Hiriart et al., 2012; Zofall et al., 2012). The RITS components Ago1, Chp2 and Tas3 have been shown to stably associate with *mei4* and various other meiotic genes and their (or Dcr1) deletion results in huge upregulation of *mei4* transcript levels. Importantly, the association of RITS with meiotic transcripts is lost when cells enter meiosis, that is, when Mmi1 is sequestered to the Mei2 dot. It appears that the role of RITS at meiotic transcripts is much the same as it is at constitutive heterochromatin - it mediates Clr4-dependent methylation (the role of which is unclear) as well as histone deacetylation by Clr3 (Hiriart et al., 2012; Zofall et al., 2012).

1.5.3 Transient heterochromatin

Finally, a transient, intermediate form of heterochromatin also exists. This kind of heterochromatin is induced by certain stresses, for example the presence or absence of particular nutrients. In contrast to the constitutive and facultative forms, transient heterochromatin is not dependent on the cell cycle stage and is often very short lived.

Only very few examples have been described and little is known about its role. One example is found in the flowering plant *Arabidopsis thaliana* and involves transient heterochromatin formation in response to heat-shock (Tittel-Elmer et al., 2010), another is the *pho1* gene in *S. pombe* (Shah et al., 2014), details of which are described on page 23.

1.6 Transcriptional gene silencing at heterochromatic loci

Whilst our understanding of heterochromatin formation has been dissected in quite some detail, the mechanisms which actually govern the transcriptional gene silencing (TGS) effect have been less thoroughly examined. What is clear is that the HP1 proteins Swi6 and Chp2 appear to be central to mediating this effect. Firstly, their oligomerisation leads to a highly condensed chromatin structure which blocks access of the transcription machinery, including Pol II (Canzio et al., 2011). Further, they provide a binding platform for the loading of various chromatin-modifying complexes, including SHREC (Snf2–HDAC [histone deacetylase] repressor complex) and the HDAC Clr6 (Yamada et al., 2005; Sugiyama et al. 2007; Garcia et al., 2010).
SHREC is composed of four factors – the class II HDAC Clr3, the ATPase and chromatin-remodeling enzyme Mit1, as well as two largely uncharacterised proteins, Clr1 and Clr2 (Sugiyama et al. 2007). Purification and mass spectrometry analysis has revealed that SHREC associates with Chp2 via the N-terminal region of Clr1 (Motamedi et al., 2008). This so called SHREC2 complex mediates H3K14 deacetylation at sites of H3K9me (where Chp2 is bound). ChIP experiments have shown that there is a comparable increase in levels of Pol II between SHREC mutants

and *chp2*Δ at centromeric regions, indicating that the SHREC2 complex can also restrict access of the polymerase to heterochromatin.

Swi6 also provides a recruiting platform for the Clr6 HDAC complex. Clr6 mediates silencing of heterochromatic repeats and suppression of antisense transcription through the removal of histone acetyl marks. It does this is concert with a histone chaperone complex composed of Asf1 and HIRA. Asf1/HIRA also functions alongside SHREC to promote nucleosome occupancy to assemble repressive chromatin architecture (Yamane et al., 2011).

1.7 ncRNA-mediated regulation of the *pho1* gene in response to phosphate

Inorganic phosphate (Pi) is essential for optimal growth and energy metabolism in all organisms. As such, it is vital to non-motile microorganisms that a steady internal supply of the nutrient is maintained. Sensing fluctuations in external Pi, as well as its acquisition and storage, is well characterised in the model organism *S. cerevisiae* (Tomar and Sinha, 2014), and to a lesser extent that of *S. pombe* (Henry et al., 2011), and is mediated by the phosphate signal transduction (PHO) pathway. In order to scavenge extracellular Pi when the internal supply is scarce, and/or under conditions of low adenine levels, a set of genes encoding phosphatases that hydrolyse Pi-containing compounds and Pi transporters are coordinately derepressed.

1.7.1 The *S. cerevisiae* PHO response

On account of its charged nature, the yeast plasma membrane is impermeable to Pi. In order to combat this, low-affinity and high-affinity transport systems function when cells are grown in high and limiting Pi conditions respectively (Fig. 1-2) (Persson et al., 1999). Under high Pi conditions (Fig. 1-2, left hand side), the PHO transcriptional activator Pho4 is phosphorylated by the Pho80-Pho85 complex. Phosphorylated Pho4 is exported to the cytoplasm where it is unable to activate high-affinity transporters (including Pho84) or secretory phosphatases (Pho5) - effectively turning the PHO pathway off. Conversely, under Pi limiting conditions (Fig. 1-2, right hand side) the Pho80-Pho85 complex is repressed, leading to hypo-phosphorylation and thus enhanced transcriptional activation activity of Pho4 (O'Neill et al., 1996; Kaffman et al., 1998; Wykoff et al., 2007). Pho4, cooperatively with the co-transcription activator

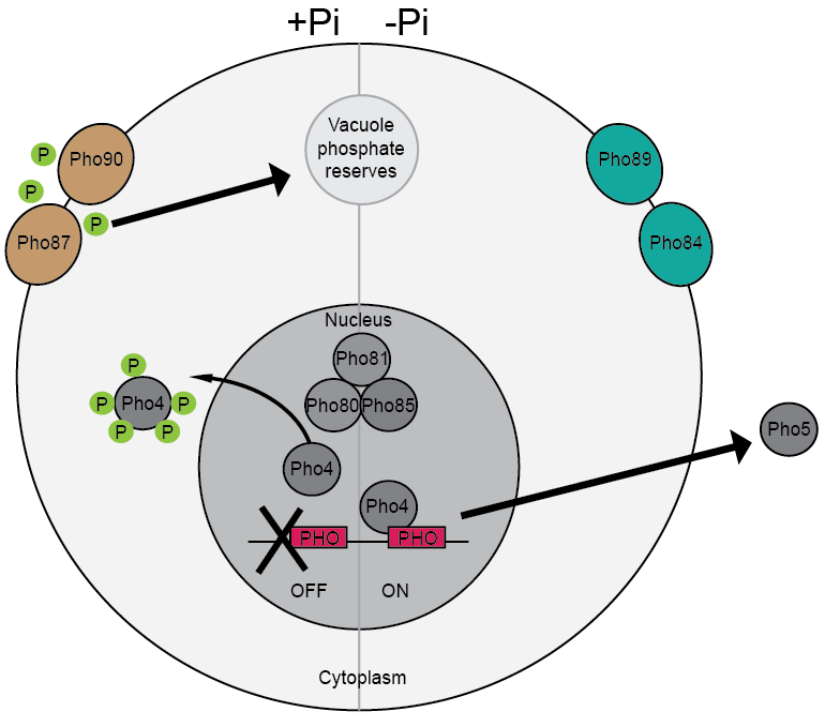


Figure 1-2: The *S. cerevisiae* PHO pathway
A schematic of the *S. cerevisiae* PHO pathway is shown under high extracellular phosphate (left hand side) and low extracellular phosphate concentrations (right hand side). In +Pi conditions the transcription factor Pho4 is phosphorylated by the Pho80-Pho81-Pho85 complex leading to its export to the cytoplasm. The low affinity transporters Pho87 and Pho90 take up phosphate for storage in vacuoles. In -Pi conditions Pho4 is no longer phosphorylated so remains localised in the nucleus where it induces phosphate responsive genes (including the phosphatase Pho5). High affinity transporters Pho84 and Pho89 are also upregulated (Figure adapted from Tomar and Sinha 2014).

Pho2 (not shown), binds to the promoters of PHO-responsive genes via the CACGTG/CACGTT motif and in so doing activates their transcription (Vogel et al., 1989; Oshima 1997).

Another way to maintain Pi homeostasis under starvation conditions is by mobilising internal reserves of the nutrient. Pi, in the form of long polyphosphate chains, is stored in vacuoles and is the first point of call when extracellular Pi becomes limiting. These reserves are built up from Pi that has been internalised by the low-affinity transporters Pho87 and Pho90. Mobilisation of these stores, through the action of polyphosphatases, precedes the up-regulation of PHO-responsive genes which comes only with harsh or persistent starvation (Kornberg et al., 1999).

It is becoming increasingly appreciated that various ncRNAs serve key regulatory roles in the PHO response pathway. In particular, their affect on the *PHO84* gene is well characterised, and has been addressed already in this chapter (see page 10).
Non-coding transcription also plays a role in the regulation of expression of the acid phosphatase-encoding *PHO5* gene. An intergenic, unstable transcript is produced antisense from the end of the *PHO5* ORF up to its promoter. Transcription of this 2.4 kb non-coding RNA through the *PHO5* promoter affects the remodelling kinetics of this region, enabling rapid nucleosome disassembly and Pol II recruitment to activate the gene (Uhler et al., 2007).

1.7.2 The *S. pombe* PHO response

Studies in *S. pombe* have revealed the existence of a non-orthologous PHO pathway to that of *S. cerevisiae*, which involves the common components *pho1* (Sc*PHO5*), encoding a major acid phosphatase, and *pho84* (Sc*PHO84*), encoding a Pi transporter. No orthologues of the budding yeast proteins Pho4, Pho2, or Pho81 have been identified (Henry et al., 2011). The fission yeast transcription factor Pho7 acts as a positive regulator of the pathway and, like Pho4, binds upstream of *pho1*. Unlike Pho4, however, Pho7 is constantly bound to its targets. Csk1 (CDK-activating kinase-activating kinase) prevents Pho7 from reaching its optimal activation activity under repressive conditions. Pi starvation relieves this Csk1 repression (and thus leads to *pho1* and other Pho7-dependent gene activation) by an unknown mechanism (Henry et al., 2011; Carter-O'Connell et al. 2012).

Non-coding RNAs have also been shown to play a significant part in the regulation of PHO genes in *S. pombe*. Of particular note is the *pho1* gene which is highly expressed in the absence of Pi and repressed in its presence. Recent work (Shah et al., 2014) has shown that, in the presence of Pi, a non-coding RNA termed *prt* (*pho1*-repressing transcript) is produced from a promoter ~1 kb upstream of *pho1* (Fig. 1-3). This RNA overlaps with and brings about repression of *pho1* by a novel mechanism that involves RNAi-dependent H3K9 methylation. The RNA-binding protein Mmi1 recognises DSR sequences present in the region unique to *prt* and associates with the exosomal subunit Rrp6 (which degrades the ncRNA), and the RNAi machinery, including the methyltransferase Clr4 which deposits repressive H3K9me2 over the *pho1* locus. Strikingly, H3K9me2 marks are lost upon Pi deficiency and subsequent *pho1*

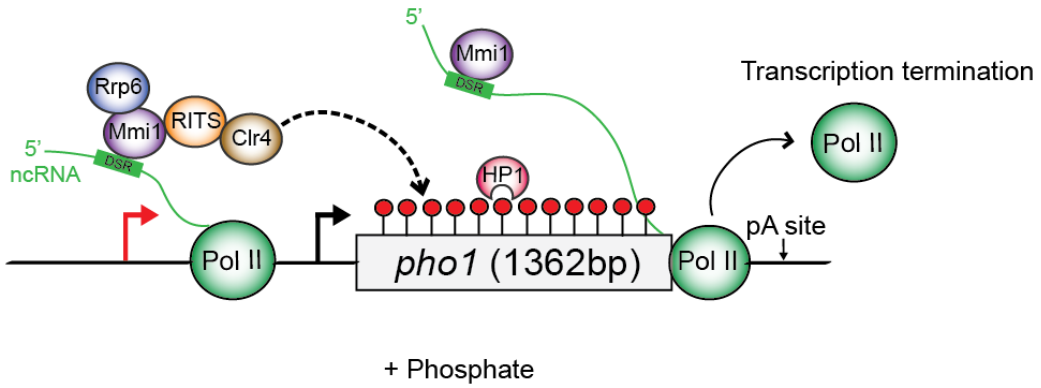


Figure 1-3: lncRNA-mediated regulation of the *pho1* gene
In the presence of phosphate, a lncRNA, *prt*, is transcribed from an upstream promoter and leads to repression of *pho1*. The RNA-binding protein Mmi1 binds to determinant of selective removal (DSR) motifs specific to *prt* where it recruits the Clr4 methyltransferase to place repressive H3K9me2 marks over the locus. *prt* RNA is subsequently degraded by the nuclear exosome. Mmi1 and the exosome also contribute to transcription termination (Shah et al., 2014).

induction. Under these conditions the *prt* RNA is no longer transcribed. It was also shown that, unlike ncRNAs that act in the assembly of constitutive heterochromatin, *prt* is not able to repress *pho1* mRNA levels in *trans*. Thus, this provides an example of transient, facultative heterochromatin.

In addition to *pho1*, recent work (Ard et al., 2014) has identified a repressive lncRNA adjacent to another Pi-responsive gene, *tgp1*, which encodes a transporter for glycerophosphodiester. *tgp1*, like *pho1*, is expressed in the absence of Pi. Here, transcription of the lncRNA, termed *nc-tgp1*, represses *tgp1* expression. Cells lacking *nc-tgp1* display hypersensitivity to a range of compounds including the microtubule destabilising drug thiabendazole (TBZ), the DNA replication inhibitor hydroxyurea (HU), and caffeine, which inhibits cyclic AMP phosphodiesterase, suggesting that transcription of *nc-tgp1* regulates drug tolerance.

A DSR motif is present in the *nc-tgp1* transcript and was shown by RNA IP (RIP) analysis to recruit Mmi1 (Ard et al., 2014). Moreover, there is a substantial increase in levels of the non-coding transcript when *rrp6* or *mmi1* are deleted. Taken together, *nc-tgp1* appears to be targeted for exosomal degradation by Mmi1. Interestingly, however, very low levels of H3K9me were detected over *nc-tgp1* or *tgp1* itself in wild-type cells, nor did these levels drop upon gene induction (Pi starvation). Further, deletion of components of the RNAi machinery (*ago1*Δ or *dcr1*Δ) and *clr4*Δ appears to have no effect on expression of *nc-tgp1* or *tgp1* by RT-qPCR. This suggests that RNAi-dependent heterochromatin formation does not mediate repression of *tgp1*.

Since transcription of *nc-tgp1* over the gene's promoter results in increased nucleosome density (as measured by histone H3 ChIP analysis) and transcription factor (Pho7) occlusion, the authors conclude that *tgp1* must be regulated by TI but do not stipulate a mechanism for this.

1.8 The RNA polymerase II C-terminal domain (CTD)

The largest subunit of RNA polymerase II contains an interesting feature called the CTD (C-terminal domain). This intrinsically unstructured yet evolutionarily conserved domain consists of a string of repeated heptads that follow the consensus sequence YSPTSPS. It has long been known that the CTD is subject to extensive modification, including glycosylation, ubiquitinylation, methylation, prolyl isomerisation and the most prevalent and best-studied of these, phosphorylation. Such phosphorylation can occur at all five of the hydroxylated amino acids within the consensus - those are Tyr1, Ser2, Thr4, Ser5, and Ser7. A plethora of possible CTD states can be created by the combinatorial action of various CTD kinases and phosphatases. These states form the "CTD code" in which is encrypted information regarding the polymerase's location in the transcription cycle. This phosphorylation profile has been particularly well characterised for budding yeast (Fig. 1-4). Importantly, the CTD serves as a flexible binding platform for the recruitment of various different transcription and processing factors, the binding of which is determined by the specific CTD phosphorylation pattern.

The number of the CTD tandem repeats varies from organism to organism (29 in *S. pombe*, 52 in mammals), with evidence pointing to complexity as a marker for length. Compared with those of other model organisms, the *S. pombe* CTD is comparatively homogenous - that is it contains few heptads that deviate from the consensus sequence. This renders it a useful model in studying the structural and functional aspects of the CTD. Of the 29 tandem repeats found in the *S. pombe* CTD, the 25 C-terminal heptads match perfectly - bar one alanine in place of proline in

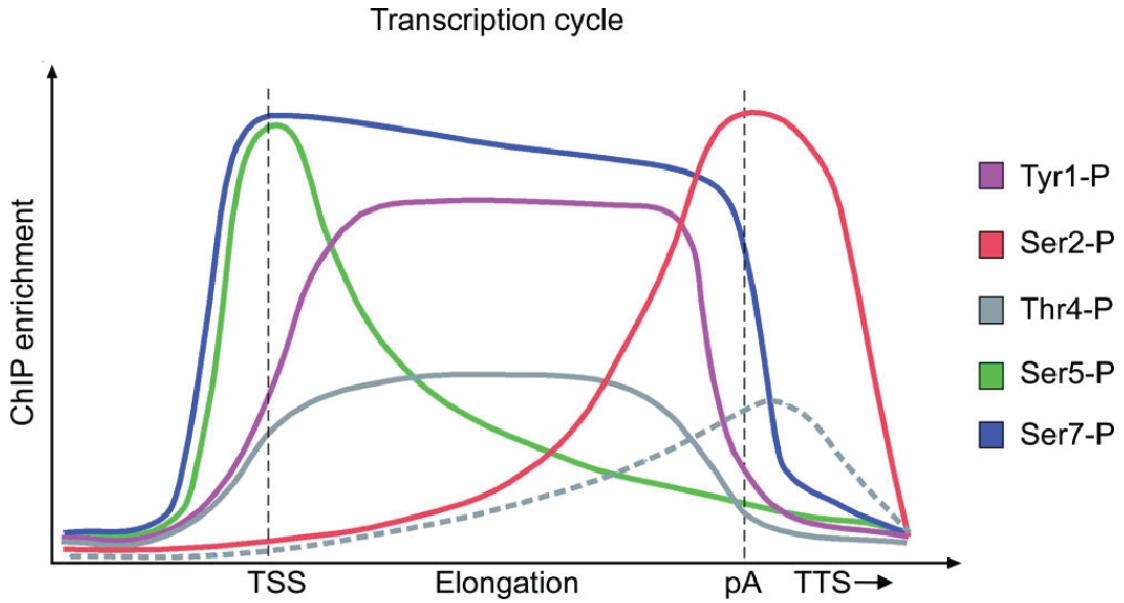


Figure 1-4: The carboxyl-terminal domain (CTD) phosphorylation profile across *S. cerevisiae* protein-coding genes
Average levels of Tyr1-(P), Ser2-(P), Thr4-(P), Ser5-(P), and Ser7-(P) modifications along a transcription cycle is shown by colour curved lines (dashed line represents Thr4-(P) profile for mammalian cells). In budding yeast, the CTD is phosphorylated by the cyclin-dependent kinases (CDKs), Srb10 (Ser2/Ser5), Kin28 (Ser5/Ser7), Ctk1 (Ser2) and Bur1 (Ser2). TSS (transcription start site), pA (polyadenylation site), TTS (transcription termination site) (Figure taken from Eich and Geyer 2013).

repeat nine - to the consensus YSPTSPS. The other four repeats differ in size and sequence from the consensus and lie at the N-terminal end; these have been referred to as the CTD "rump" (Schwer et al., 2012).

1.9 CTD-interactions

The role of the CTD in recruitment of RNA processing factors is well established and includes recruitment of capping enzyme (Cho et al., 1997), elongation factors (see for example: Hsin and Manley, 2012), and the termination factors Nrd1 (Vasiljeva et al., 2008), Pcf11 (Licatalosi et al., 2002), and Rtt103 (Kim et al., 2004b). Nrd1, Pcf11 and Rtt103 share a highly conserved common fold at their N-termini – the CTD interacting domain (CID) (Fig. 1-5).

Although WW domains, FF domains and the SRI domain are also dedicated to CTD binding, the CID is the best studied scaffold for recruitment. The CID is a right-handed superhelix composed of eight alpha-helices. CID-CTD structures have been solved for each of the *S. cerevisiae* factors mentioned, as well as the human RNA processing factor SCAF8 (a homologue of Nrd1) (Meinhart et al., 2004; Vasiljeva et al., 2008; Becker et al., 2008; Lunde et al., 2010). Pcf11 and Rtt103 bind the CTD in a manner which is enhanced by Ser2 phosphorylation (Barilla et al., 2001; Licatalosi et al., 2002; Meinhart and Cramer 2004; Kim et al., 2004b; Lunde et al., 2010), whereas Nrd1 binds with a preference for Ser5-(P) (Kubicek et al., 2012; Vasiljeva et al., 2008) (Fig. 1- 5). SCAF8 binds more promiscuously to the CTD but with a higher affinity to Ser2-(P) than Ser5-(P) (Becker et al., 2008).

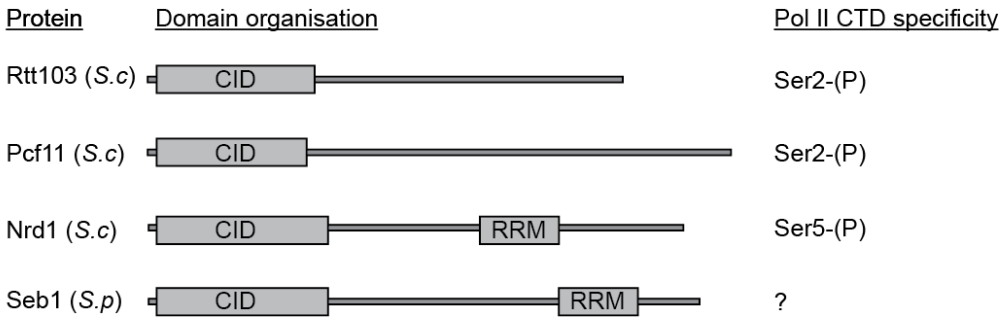


Figure 1-5: CTD-interacting domain (CID)
Domain organisation and CTD specificity of CID proteins from *S. cerevisiae* (*S.c*) and *S. pombe* (*S.p*). CID (CTD-interacting domain), RRM (RNA-recognition motif).

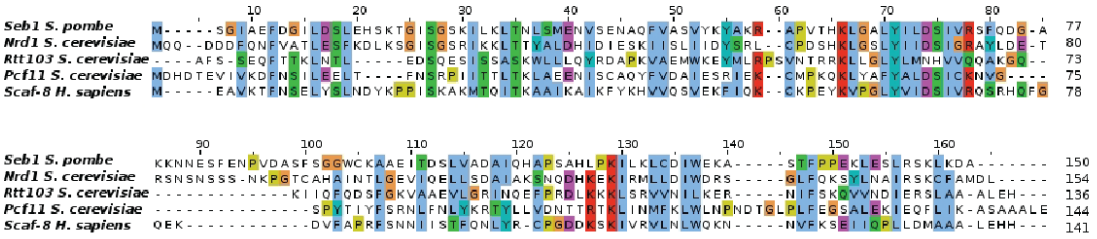


Figure 1-6: CID sequences
Sequence alignment of *S. pombe* Seb1, *S. cerevisiae* Nrd1, Rtt103, Pcf11, and human SCAF8 CIDs.

1.10 Nrd1 and non-coding transcription

Ser5-(P) marks are highest early in elongation, whereas Ser2-(P) follows an inverse pattern that peaks over the polyadenylation (pA) site of genes (Fig. 1-4). It thus fits that the termination factors for mRNA genes, Pcf11 and Rtt103, bind Ser2-(P) marks. Why, then, does the other budding yeast CID protein and termination factor Nrd1 bind preferentially to Ser5-(P)? Nrd1, together with an RNA-binding protein Nab3 and the Sen1 helicase, is essential for the termination of cryptic unstable transcripts (CUTS), snoRNAs and other short (fewer than 600 nt) Pol II transcripts (Vasiljeva and Buratowski, 2006; Steinmetz et al., 2001; Lykke-Andersen and Jensen, 2006). Interestingly, CID swapping experiments have revealed that the choice between the distinct termination pathways mediated by Nrd1 and Rtt103 is mediated through CTD:CID interactions (Heo et al., 2013). As well as its CID for Pol II binding, Nrd1 has a second domain - an RNA-recognition motif (RRM) - through which it can interact with RNA in a sequence-specific manner. Via these interactions, Nrd1 preferentially targets the termination complex to 5' ends where, in association with the nuclear exosome and TRAMP complex, it couples termination and 3' processing/degradation (Vasiljeva and Buratowski, 2006; Gudipati et al., 2008). Because of the strong cross-linking of Nrd1 at early transcribed regions, CUTs are limited in size to a maximum of ~500 bp.

1.11 Seb1

The *S. pombe* orthologue of Nrd1 is Seb1 (Mitsuzawa et al., 2003). Like Nrd1, Seb1 has a CID at its N-terminus (Fig. 1-6) and a conserved RRM (Fig. 1-5), suggesting

that it can bind both Pol II via the CTD and RNA. Another similarity between Nrd1 and Seb1 is a short region between the two domains that is rich in arginine-serine (RS) dipeptides; something that is typical of a family of metazoan proteins involved in RNA splicing (Steinmetz and Brow, 1996). Seb1 is essential for cell viability (Mitsuzawa et al., 2003) and has been shown to play a role in RNAi-independent heterochromatin assembly (Marina et al., 2013). Since Seb1 was shown to interact with *dg* and *dh* ncRNAs and the snoRNA *snR30*, it was proposed that recognition of ncRNAs by this protein is conserved between the two yeasts (Marina et al., 2013).

As one might expect, mutation of the RNAi component Dicer leads to a reduction in H3K9me levels at pericentromeric repeats. This is also true for a *seb1* mutant (Marina et al., 2013). However, a double mutant of these two genes has an additive effect, completely eliminating methylation, and consequently indicates that Seb1 functions independently of the RNAi machinery in silencing centromeric regions. Instead, Seb1 promotes H3K9me through recruitment of the SHREC complex. The SHREC component Clr3 is an H3K14 deacetylase and it was found that a *seb1* mutant results in increased acetylation at *dg* and *dh* repeats to the same degree as a *clr3* deletion strain. From this the authors conclude that SHREC, via recruitment by Seb1, promotes H3K9me by one of two mechanisms, either (i) Clr3 deacetylation of histone H3 permits, and is necessary for, Clr4 methylation, or (ii) Clr3 deacetylates some other non-histone substrate which leads to the increases seen in H3K9me.

1.12 Objectives of this work

Prior to the work described herein, the mechanism(s) governing repression of the *pho1* gene via the action of non-coding transcription was incomplete. It was shown that the non-coding RNA *prt* is able to recruit Clr4 to repress *pho1* mRNA by the deposition of H3K9me2 marks (Shah et al., 2014). However, in response to Pi, deletion of Clr4 results in partial induction of *pho1* mRNA that is not comparable to the level of increase observed upon elimination of non-coding transcription. This prompted an investigation into other means by which *prt* transcription is able to mediate *pho1* repression.

The initial objective of this work, then, was to examine how non-coding transcription at the *pho1* locus is able to regulate *pho1* expression, and in particular to determine whether this regulation is mediated by chromatin modifiers. As well as histone methylation, which is controlled by only one methyltransferase in *S. pombe* - Clr4 - the other common and widespread hallmark of heterochromatin is histone hypoacetylation. In the following chapters it is shown that deacetylation of histones over *pho1* is important for the gene's repression in the presence of Pi. Interestingly, the histone deacetylase Clr3 is identified as responsible for removal of marks of active transcription (H3K14ac), and allowing for this *pho1* repression.

The RNA-binding protein Seb1 appears to play a role in defining the location of Clr3 recruitment by interacting with the lncRNA *prt*. In chapter 4, Seb1 interactions are considered in the context of its two domains. As well as RNA, Seb1 is shown to preferentially bind serine-2 and serine-5 phosphorylated Pol II CTD via its CID

in vivo. Strikingly, the specificity of this Seb1:CTD interaction does not match that of its *S. cerevisiae* (Nrd1) homologue. Seb1 is also shown to strongly associate with proteins involved in 3' end processing and is enriched at the pA sites of protein-coding genes; raising the question of whether Seb1 has a dual role - one rather more specific role in mediating recruitment of chromatin modifiers to non-coding transcripts, and another, potentially more global role in 3' end processing and/or termination of protein-coding gene transcripts.

A mechanism of TI is presented that involves recognition of ncRNA by Seb1. Seb1, in turn, recruits Clr3 (probably as part of SHREC) to deacetylate histones and lead to *pho1* repression independently of Clr4-mediated methylation. In the final chapter it is demonstrated that the mechanism describing this ncRNA-mediated transcriptional repression is not unique to *pho1*. Another metabolic gene, *zrt1*, responsible for zinc acquisition, appears to be regulated in an identical manner. Finally, a case is described whereby Seb1 is shown to contribute to gene silencing (in this case the Pi-responsive gene *tgp1* which encodes transporter for glycerophosphodiester 1) independently of Clr3-mediated deacetylation. This highlights that the mode of TI described here is just one of a complex network of mechanisms which can potentially employ a host of different players to regulate gene expression.

2. Materials and Methods

2.1 Media and growth conditions

Cells were grown either in standard YES medium or EMMG in the presence or absence of phosphate (10 mM KH_2PO_4) (Henry et al., 2011). Unless otherwise stated, cells were grown to OD (A_{600}) ~ 0.5 before harvesting. To inhibit HDAC activity, cells were grown in YES to exponential phase before the addition of TSA (Sigma, dissolved in ethanol) to a final concentration of 20 µg/ml. An equivalent volume of ethanol was added to untreated samples. Cells were grown in the presence of the inhibitor for two hours.

2.2 Yeast transformation

5ml of cells were grown in YES to exponential phase before harvesting and washing with sterile water and LiTE (0.1 M LiAc, 10 mM Tris-HCl (pH 7.5), and 1 mM EDTA). Pellets were resuspended in 100 µl LiTE, and 20 µg ssDNA and 1 µg of plasmid or product DNA were added. Following a 10 minute incubation at room temperature, 260 µl PLATE (40% PEG, 0.1 M LiAc, 10 mM Tris-HCl (pH 7.5), and 1 mM EDTA) was added. Cells were mixed by gently vortexing and subsequently incubated at 30°C for one hour. 43 µl DMSO was added and cells were heat shocked at 42°C for 5 minutes. Pelleted cells were washed with water, resuspended in 650 µl YES or EMMG and incubated at 30°C overnight. Transformants were spread onto selective plates.

2.3 CTD-Peptide binding experiments

CTD peptides (GL Biochem and peptides&elephants) were synthesised with four repeats of the consensus YSPTSPS with N-terminal biotinylation, and were either unphosphorylated, serine 2, 5, or 7 phosphorylated. Peptides were dissolved in TE buffer (10 mM Tris-HCl (pH 8.0), 1 mM EDTA). 0.6 mg streptavidin-coated magnetic beads (M280 Dynabeads) were suspended in 500 µl oligopeptide binding buffer (25 mM Tris-HCl (pH 8.0), 50 mM NaCl, 1 mM DTT, 5% glycerol, 0.03% Triton X-100) and incubated with 5 µg of peptide for 2 hours at 4°C. After three washes with binding buffer to remove unbound peptide, yeast whole cell extract (500 mg) was added and the mixture incubated overnight at 4°C. Beads were then washed five times with binding buffer (5 minutes, 4°C) and bound proteins analysed by western blot.

2.4 Immunoblotting

TAP-tagged Seb1 was purified via IgG sepharose beads as described below. TEV cleaved samples were separated on an SDS PAGE gel and blotted onto Immobilon-P PVDF membranes. Membranes were blocked and incubated with either peroxidase anti-peroxidase (PAP) (Sigma, P1291), α-HA (Abcam, ab16918), α-Rpb1 (Millipore, 8WG16), α-CTD repeat (phosphor S2) (Abcam, ab5095), α-CTD repeat (phosphor S5) (Abcam, ab5131), or α-CTD repeat (phosphor S7) (Abcam, ab126537). Following incubation with the appropriate horseradish peroxidase (HRP)-conjugated secondary antibody, Western blots were detected by enhanced chemiluminescence (ECL).

2.5 Purification of Seb1-TAP

TAP-tagged Seb1 was purified from 16 L of yeast culture grown in YES to OD (A_{600}) =0.9. Harvested cells were washed once with sterile water and once with TMN lysis buffer (20 mM Tris-HCl (pH 8.0), 5 mM $MgCl_2$, 150 mM NaCl, 10% glycerol) supplemented with 1 mM phenylmethylsulfonyl fluoride (PMSF), 1 mM benzamidine, 2.3 µM leupeptin (1 µg/mL), 1.5 µM pepstatin A (1 µg/mL), 81 µM bestatin (25 µg/mL), and 1.5 µM aprotinin (10 µg/mL). Cells were resuspended in TMN and frozen into beads by dropping into liquid nitrogen. To prepare extract, yeast beads were broken and ground into a fine powder using a pestle and mortar and subsequently vortexed with glass beads in TMN buffer. The extract was centrifuged at 2500 g for 7 minutes, to remove glass beads, followed by ultracentrifugation at 75,000 g for 1.5 hr.

Seb1-TAP was incubated with 1000 µl IgG sepharose (VWR) for 16 hrs. Beads were washed at 4°C for 2 x 5 minutes with TMN buffer (plus protease inhibitors as above), 3 x 15 minutes with TMN (without protease inhibitors), and once with TEV cleavage buffer (20 mM Tris-HCl (pH 8.0), 150 mM NaCl, 0.5 mM DTT, 0.05% NP-40, 5% glycerol) for 5 minutes. 20 µl AcTEV protease (Invitrogen) was then added for overnight cleavage. Seb1-TAP cleavage and purification was analysed by western blot and silver stain.

2.6 Strain construction

Strains were constructed by standard methods used for genetic crosses or transformation (Bähler et al., 1998; Gao et al., 2014; Moreno et al., 1991). The pop-in, pop-out method for allele replacement was used to introduce deletions of the *pho1* lncRNA (Gao et al., 2014). DNA fragments carrying the desired deletions (see Table 2 for oligonucleotides) were generated by 2-step PCR. Fragments were confirmed by sequencing, and cloned into pCR blunt II TOPO (Life Technologies), according to the manufacturer's instructions, before subcloning into pKS-URA4 (Bähler et al., 1998) by SpeI/NotI digestion and ligation. BseRI was used to linearise the plasmid before transformation into *S. pombe* cells lacking the *ura4* locus. Ura^+ colonies were selected on minimal media without uracil and genotyped by DNA sequence analysis of PCR products. Positive clones were plated onto 5-FOA to select for cells in which the $ura4^+$ gene had been 'popped-out'. Clones were verified by colony PCR and sequencing of products.

2.7 Northern blotting

8 μg of RNA was resolved on a formaldehyde-agarose gel (1.2%). The gel was submersed in 0.05 M NaOH for 20 minutes with shaking and the RNA blotted onto a nylon membrane by capillary transfer for 18-24 hours. RNA was immobilised onto the membrane by UV cross-linking. Following methylene blue (0.04% methylene blue in 0.5 M sodium acetate) staining, the membrane was pre-hybridised at 42°C for four hours in pre-hybridisation solution (50% formamide, 5% Denhardt's solution, 5% SSPE, 0.1% SDS, and 200 μg/ml single-stranded DNA). DNA fragments amplified

from genomic loci were used as templates for random oligo-labelling using Prime-it Random Primer Labelling Kit. Hybridisation with the radiolabelled probe was performed for 12 hours. Membranes were then washed with a solution of 2X SSC (300 mM NaCl, 30 mM sodium citrate (pH 7.0))/ 0.1% SDS at 42°C and exposed to X-ray film.

2.8 ChIP experiments

ChIP was performed essentially as described in (Keogh and Buratowski, 2004; Kim et al. 2004a, b). Briefly, 200 mL of exponentially growing cells were cross-linked with 11% formaldehyde solution for 20 minutes at room temperature. 30 mL of 3M glycine, 20 mM Tris was used to quench the reaction. Cells were pelleted and washed once with cold TBS and once with FA lysis buffer (50 mM Hepes-KOH (pH 7.5), 150 mM NaCl, 1 mM EDTA, 1% Triton X-100, 0.1% Na Deoxycholate)/ 0.1% SDS. To prepare chromatin, cells were resuspended in FA lysis buffer with 0.5% SDS and vortexed for 30 cycles of one minute vortexing and one minute on ice. The lysate was ultracentrifuged (150,000 g, 20 minutes) and the pellet crushed in lysis buffer. Samples were sheared for 80 minutes of sonication cycles 15 sec ON/ 45 sec OFF with a Biorupter sonicator, and ultracentrifuged (150,000 g, 20 minutes) to yield sheared chromatin in the supernatant. At this point the concentration of NaCl was adjusted to 275 mM. Immunoprecipitations (IPs) were conducted with rabbit IgG agarose (Sigma) or antibodies recognising H3 (Abcam, 1791), H3K9me2 (Abcam, 1220), H3K14ac (Millipore, 07-353), Rpb1 (Millipore, 8WG16), or c-Myc (Santa Cruz, SC-40) coupled to protein-G dynabeads (Life Technologies). After phenol-chloroform extracting DNA, the amount of IP DNA relative to an input sample was

determined by quantitative PCR analysis using SensiMix SYBR (Bioline). All values were normalised to a control gene (*fbp1*).

2.9 RNase ChIP

ChIP with RNase treatment was performed as usual, excepting the following: (i) Formaldehyde cross-linking time was reduced from 20 minutes to 5 minutes. (ii) Chromatin was prepared in FA lysis buffer with 0.05% SDS. (iii) Cross-linked chromatin was treated with 7.5 U of RNase A (Thermo Scientific) and 300 U of RNase T1 (Thermo Scientific). An equivalent volume of RNase storage buffer (50 mM Tris-HCl (pH 7.4) and 50% (v/v) glycerol) was added to untreated samples. After incubation at room temperature for 30 minutes, IPs were performed as described above.

Table 1: Strains used in this study

Name	Genotype	References
YP51 (wild-type)	*h+, ura4-Δ18, leu1-32, ade6-M216, imr1R(NcoI)::ura4+*	(Bühler et al., 2008)
YP52	*h+, ura4-Δ18, leu1-32, ade6-M216, imr1R(NcoI)::ura4+, rrp6Δ::natMX6*	(Bühler et al., 2008)
YP56	*h+, ura4-Δ18, leu1-32, ade6-M210, otr1R(Sph1)::ura4+, clr4Δ::natMX6*	(Gullerova et al., 2011)
YP144 (wild-type)	*h+, ura4-Δ18, leu1-32, ade6-M216, his3Δ::1*	(Lemieux et al., 2011)
YP223	*h-, ura4-Δ18, leu1-32, ade6-M210, pro1Δ(-346-(-1))*	(Shah et al., 2014)
YP225	*h-, ura4-Δ18, leu1-32, ade6-M216, pro2Δ(-1438-(-1204) (ATG=1)*	(Shah et al., 2014)
YP241	*h+, ura4-Δ18, leu1-32, ade6-M216, imr1R(NcoI)::ura4+, mmi1-his6-TEV-ProtA::Kan MX6*	(Shah et al. 2014)
YP291	*h+, ura4-Δ18, leu1-32, ade6-M216, imr1R(NcoI)::ura4+ seb1-3HA-TAP::kan MX6*	This study
YP292	*h+, ura4-Δ18, leu1-32, ade6-M216, imr1R(NcoI)::ura4+ rhn1-3HA-TAP::kan MX6*	This study
YP307 (wild-type)	*h-*	(Marina et al., 2013)
YP319	*h+, ura4-Δ18, leu1-32 ade6-M216, imr1 R(NcoI)::ura4+ oriI, chp2Δ:: kanR*	(Motamedi et al., 2008)
YP320	*h+, ura4-Δ18, leu1-32, ade6-216, imr1R(NcoI)::ura4+ oriI , swi6Δ:: kanR*	(Motamedi et al., 2008)
YP344	*h-, ura4-Δ18, leu1-32, clr3:13myc:ura4+*	National BioResource Project (NBRP), Japan
YP345	*h+, ura4-Δ18, leu1-32, clr3::kanR*	National BioResource Project (NBRP), Japan
YP351	*h-, seb1-1::natMX*	(Marina et al., 2013)
YP381	*leu1, his2, clr6-1*	National BioResource Project (NBRP), Japan

YP383	*h-, ura4-Δ18, ura4A-13, clr6-1*	National BioResource Project (NBRP), Japan
YP390	*h90, ura4-Δ18, leu1-32, ade6-M210, sir2::kanMX6*	National BioResource Project (NBRP), Japan
YP408	*h+, ura4-Δ18, leu1-32, ade6-M210, otr1R(Sph1)::ura4+, clr4Δ::natMX6, clr3Δ::kanR*	This study
YP443	*h+, ura4-Δ18, leu1-32, ade6-M216, imr1R(NcoI)::ura4+ clr3-3HA-TAP::kan MX6*	This study
YP444	*h-, ura4-Δ18, leu1-32, ade6-M216, pro2Δ(-1438-(-1204) (ATG=1), clr3-3HA-TAP::kan MX6*	This study
YP469	*h+, ura4DS/E, leu1-32, ade6-M210, otr1R::ura4+ mit1-myc::kanR*	National BioResource Project (NBRP), Japan
YP478	*h+, ura4-Δ18, leu1-32, ade6-M216, imr1R(NcoI)::ura4+ oriI clr1 Δ::TAP-kanR*	(Motamedi et al., 2008)
YP479	*h+, ura4DS/E, leu1-32, ade6-M210, otr1(SphI)::ura4+oriA kanR-TAP-clr2*	(Motamedi et al., 2008)
YP480	*h+, ura4-Δ18, leu1-32, ade6-M216, imr1R(NcoI)::ura4+ oriI clr2 Δ::TAP-kanR*	(Motamedi et al., 2008)
YP481	*h+, ura4-Δ18, leu1-32, ade6-M216, imr1R(NcoI)::ura4+ oriI mit1Δ:: kanR*	(Motamedi et al., 2008)
YP482	*h+, ura4-Δ18, leu1-32, ade6-M216, imr1R(NcoI)::ura4+ oriI clr1-TAP- kanR*	(Motamedi et al., 2008)
YP510	*h+, ura4-Δ18, leu1-32, ade6-M210, ura4::p(nmt1)-seb1-FLAG::NatMX*	This study
YP558	*h+, ura4-Δ18, leu1-32, ade6-M216, his3Δ::1, prt-1Δ(-1197-(-1008)) (ATG=1)*	This study
YP559	*h+, ura4-Δ18, leu1-32, ade6-M216, his3Δ::1, prt-2Δ(-1008-(-819)) (ATG=1)*	This study
YP560	*h+, ura4-Δ18, leu1-32, ade6-M216, his3Δ::1, prt-3Δ(-819-(-630)) (ATG=1)*	This study
YP561	*h+, ura4-Δ18, leu1-32, ade6-M216, his3Δ::1, prt-5Δ(-441-(-252)) (ATG=1)*	This study
YP562	*h+, ura4-Δ18, leu1-32, ade6-M216, his3Δ::1, prt-6Δ(-252-(-63)) (ATG=1)*	This study

Table 2: Oligonucleotides used in this study

No.	Name	Sequence (5'-3' direction)	Purpose
2463	Pho1-forward	TAAATTTGATTTCAAGGAACATTTG ACTTC	Northern
2464	Pho1-reverse	TTGTAAGTACTAGCATTAAAGAGCT CATGG	Northern
2469	Adh1-1 F	CGGAAGCTGGTGAGAAGAAC	ChIP
2470	Adh1-1 R	CGTTGGAATGCGGAGTAGAG	ChIP
2471	Adh1-2 Forward	CAACCTCCCATTTCCTCCTT	ChIP
2472	Adh1-2 Reverse	GTGGACACATTTCGGGAATC	ChIP
2473	Adh1-3 Forward	TCTCTCGCTTTCCTCATTCG	ChIP
2474	Adh1-3 Reverse	GCCAACTGCTTGTCAGGAAT	ChIP
2475	Adh1-4 Forward	GGTCCCGAGAACGTCAAGT	ChIP
2476	Adh1-4 Reverse	ACTTGACACCAACACGGTCA	ChIP
3046	Pho1 ChIP F1	TAGCTAACTAAGGTAGGCGTTGC	ChIP
3047	Pho1 ChIP R1	GGGGCTATGCTATATATACCTTTTTG	ChIP
3048	Pho1 ChIP F2	ACAATTATATCTTGGTCTGGGGAAC	ChIP
3049	Pho1 ChIP R2	ATCATTAAATTGTGAATATCGCAAGAC	ChIP
3050	Pho1 ChIP F3	ATGTTTGAGATTTACGGGAAGTG	ChIP
3051	Pho1 ChIP R3	TTTGTCCTAATTTTCCAAACAGC	ChIP
3052	Pho1 ChIP F4	TTTGTACCAACTTGGACTCCTG	ChIP
3053	Pho1 ChIP R4	GCGTCCCATGTCAAATAACTC	ChIP
3054	Pho1 ChIP F5	CTTCGCCTTTACTCATGATGC	ChIP
3055	Pho1 ChIP R5	TTGGTAGGAAGTAGGCAATGG	ChIP
2547	Pho1 F14	AGCAGAGGATAGTTTATGTAGGAGATAATG	ChIP
2548	Pho1 R14	TTTATATGGTGAGAGTATTGTCAAAGAAAC	ChIP
3129	Pho1 F17	AAAATTCTATGTTTCTATACATGCC TCTG	ChIP
2857	Pho1 R17	AAACTAAGTCTTGACAACTATAACG	ChIP

		AAACC	
3072	Fbp1f	ACTGCGATGAAGTCGAACG	ChIP
3032	Fbp1r	CAAGTGACGGCATAGGAACC	ChIP
4148	Zrt1 fwd	TCATCATGGTCCAACTCTTG	Northern; ChIP
4149	Zrt1 rev	CGAGAACCTAAACCACAACC	Northern; ChIP
4163	SPNCRNA.425 fwd	CTTGTTTGAGGCGGTAACTT	ChIP
4164	SPNCRNA.425 rev	TCGATCTGATAGAGCGACAA	ChIP
3571	1271.09-ORF-fwd	GACAACGGCGAAGATGAAGT	Northern
3572	1271.09-ORF-rev	ATTTAGGTGACACTATAGAAGCATACCAAT GTCCTGCAACG	Northern
4445	Tgp1 PP1f	TCGGTTGGAATGTTCTAATCAATAC	ChIP
4446	Tgp1 PP1r	AGACCGGTGATCAAACAATATTTAG	ChIP
4447	Tgp1 PP2f	TGAAGTAGTTAGACAGGTTAGCGA	ChIP
4448	Tgp1 PP2r	CTTGTCGTCCAACTTCTCTTCATC	ChIP
4449	Tgp1 PP3f	GGCAGTAAATCTATCTGTAGCGAGT	ChIP
4450	Tgp1 PP3r	TACACGGTAAATGTCAAGTCTGCTA	ChIP
4451	Tgp1 PP4f	CTGACAAACCAATTATCCCTACACG	ChIP
4452	Tgp1 PP4r	GTATTACGATTTGGCAACCTCATCC	ChIP
4453	Tgp1 PP5f	TTAAATGCTGCACTCACATACTGAC	ChIP
4454	Tgp1 PP5r	ACTCTCCCTTGGGTTCATTTGATTA	ChIP
4455	Tgp1 PP6f	ATACAGACGTGTGGATTGCAA	ChIP
4456	Tgp1 PP6r	CCTCTTCTATACGCAATCAATGTC	ChIP
4648	DncRNApart2f	CATGCGCGAATTTAAATTTATGGTACGAAG CATGTTGACTGAATTCTTTAAACCTTTT	Cloning
4649	DncRNApart2r	AAAAGGTTTAAAGAATTCAGTCAACATGCT TCGTACCATAAATTTAAATTCGCGCATG	Cloning
4650	DncRNApart3f	ATATCGAAGGTCAGCGTCTAAAATTTTTAT AAAAATCGTTAAAAATCTTAATTTACTA	Cloning

4651	DncRNApart3r	TAGTAAATTAAGATTTTTAACGATTTTTATA AAAATTTTAGACGCTGACCTTCGATAT	Cloning
4654	DncRNApart5f	TTAATACAGGGTTAGGGCATTTACAAGTCA ACAGGGGAATAGTCGCTGCTTGAAACAT	Cloning
4655	DncRNApart5r	ATGTTTCAAGCAGCGACTATTCCCCTGTTGA CTTGTAAATGCCCTAACCCTGTATTAA	Cloning
4656	DncRNApart6f	ACTCGGCATTACGTGGGACATGTTTGAACT CTGCTTCTTTGGTTACAACCAAACTTTT	Cloning
4657	DncRNApart6r	AAAAGTTTGGTTGTAACCAAAGAAGCAGAG TTCAAACATGTCCCACGTAATGCCGAGT	Cloning
4668	DncRNApart1f	ATAGCCCCCCTAATGAACTATTCTTTTAGTG TTTTCATTGATTAACTGTTTGTTGCTT	Cloning
4669	DncRNApart1r	AAGCAACAAACAGTTAATCAATGAAAACAC TAAAAGAATAGTTCATTAGGGGGGCTAT	Cloning

3. A ncRNA regulates expression of *pho1* by an H3K9 methylation-independent mechanism

3.1 Introduction

The genomes of many organisms - no less the fission yeast *S. pombe* model system used in this work - are extensively transcribed into long non-coding RNAs (lncRNAs), many of which are implicated in gene regulation. There is a growing number of recognised roles for lncRNAs. One recently uncovered is repression of the *pho1* gene by transcription of the upstream overlapping ncRNA termed *prt* (*pho1*-repressing transcript). In response to phosphate, *prt* transcription leads to the formation of transient heterochromatin over the *pho1* locus (Shah et al., 2014). This high-order chromatin assembly has been shown to be both RNAi- and Clr4-dependent and is characterised by H3K9me - a classical hallmark of heterochromatin.

Abolishing H3K9 methylation at the *pho1* locus by deleting the only known methyltransferase Clr4 results in an accumulation of the mRNA but not to a level comparable to when transcription of *prt* is prevented. This suggests, then, that *prt* can mediate *pho1* repression by other, methylation-independent, means. The initial aim of the following work was to identify and characterise how this repression is established.

As mentioned, methylation is one of the widespread hallmarks of heterochromatin (which has previously been shown to contribute to *pho1* repression), the other being histone deacetylation. Histone acetyl transferases (HATs) are the family of proteins

responsible for the acetylation of histone tails at several conserved lysine residues. These include K9 and K14 of histone H3, and K5, K8, K12, and K16 of histone H4 (Bjerling et al., 2002). The reverse reaction - that is, removal of the acetyl group - is carried out by the histone deacetylases (HDACs). Fission yeast has two protein families with deacetylase activity - the NAD^+-dependent Sir2 HDACs (of phylogenetic class III) and the 'classical' class I and II HDACs. The activity of the classical HDACs is known to be inhibited by trichostatin A (TSA) via their zinc-coordinated active sites, whereas the Sir2 HDACs are not affected by this inhibitor. Three classical HDACs - Clr3, Clr6, and Hda1 (also known as Phd1) - have been identified in *S. pombe*. Western blot analysis of bulk histones has revealed that Clr3 is highly specific for H3K14 whereas Clr6 has a very broad specificity, and the cytoplasmic HDAC Hda1 has low level activity (Bjerling et al., 2002). Clr3 is part of the multienzyme complex SHREC which functions to regulate nucleosome structure to form high-order chromatin states that are essential for both constitutive and facultative heterochromatin.

In this chapter the function of histone deacetylation, particularly via HDACs, at the *pho1* locus is examined. After first identifying a role for histone deacetylation, it is shown here that Clr3 is the specific HDAC which contributes to and is, in part, responsible for *pho1* repression. Furthermore, the way in which Clr3 actually brings about transcriptional gene silencing is explored in the context of this locus.

As described, Clr3 is one component of the SHREC effector complex. Therefore, whether the recruitment and function of Clr3 is dependent on the presence and activity of the other four components is examined. Finally, the relationship between histone

methylation by Clr4 and deacetylation by Clr3 is investigated. At odds with previous data for pericentromeric heterochromatin (Marina et al., 2013), it is shown here that, at *pho1*, ncRNA-mediated methylation and deacetylation contribute to repression in a mutually exclusive manner.

3.2 Results

3.2.1 Clr3 HDAC contributes to *pho1* silencing

It was previously shown that expression of the *pho1* gene is regulated by transcription of the overlapping, non-coding RNA *prt*. In response to phosphate, deletion of the methyltransferase Clr4 results in induction of *pho1* mRNA but this is not, however, to an extent comparable to when non-coding transcription is abolished. Another way by which *pho1* expression can be regulated via non-coding transcription was therefore sought. Since Clr4 is the only known methyltransferase present in *S. pombe*, the other hallmark of heterochromatin, namely, hypoacetylation of histones was considered.

The conserved active sites of HDACs coordinate a Zn^{2+} ion which can be blocked by the potent inhibitor TSA (Finnin et al., 1999). Therefore, in order to determine whether *pho1* expression is regulated via the HDAC-dependent removal of histone acetyl marks, cells were treated with TSA and examined for changes in the level of *pho1* mRNA by Northern blot (Fig. 2-1). Compared to the untreated wild-type (Fig. 2-1, lane 1), a substantial increase in *pho1* mRNA was observed when wild-type cells were transiently treated with TSA (Fig. 2-1, lane 2). At the same time, no change in the constant basal level of expression of the house-keeping gene *adh1* was observed upon TSA treatment (Fig. 2-1, bottom panel).

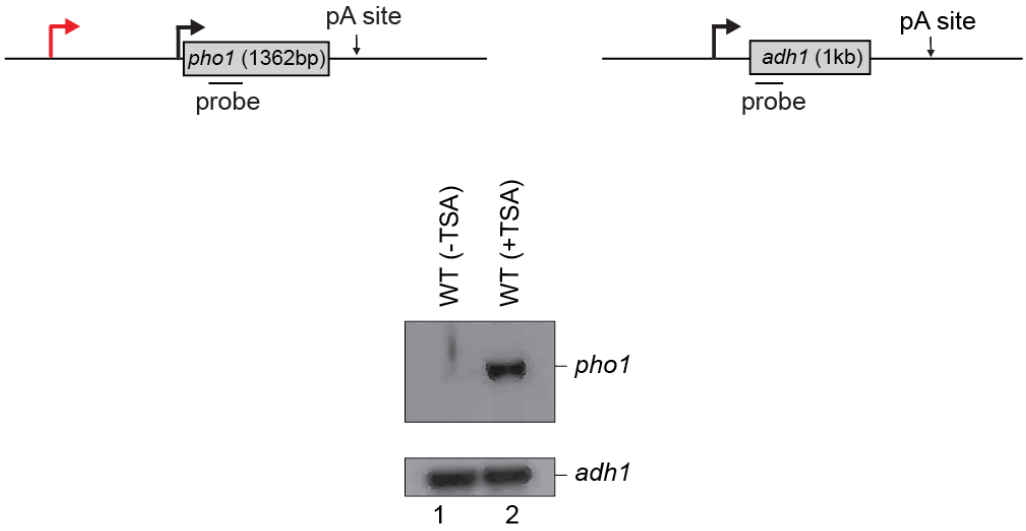


Figure 2-1: *pho1* is sensitive to the HDAC inhibitor TSA
(Top) Schematic of *pho1* and *adh1* genes with probes indicated by black bars. (Bottom) Northern blot analysis of RNA isolated from a wild-type (WT) strain, YP51, treated with (+TSA) or without (-TSA) 20 µg/mL of the HDAC inhibitor trichostatin A (TSA). The housekeeping gene *adh1* serves as loading control. *pho1* mRNA levels are shown.

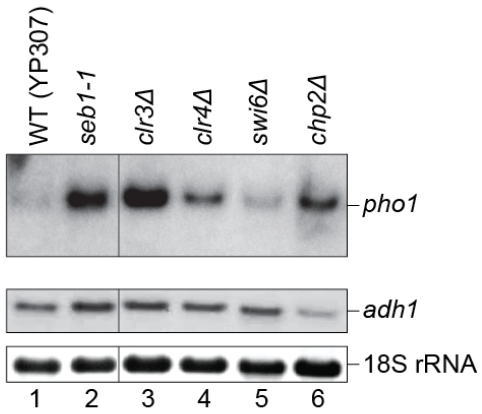


Figure 2-2: Clr3 is the HDAC that contributes to *pho1* silencing
Northern blot analysis of RNA isolated from the indicated strains grown in YES. The housekeeping gene *adh1* and 18S ribosomal RNA serve as loading controls. *pho1* probe is as used in (Fig. 2-1).

In order to identify which of the three classical HDACs, Clr6, Clr3, or Hda1, is responsible for repression of *pho1*, RNA from *clr6-1*, *clr3*Δ and *hda1*Δ mutant strains were examined by Northern blot (data not shown). Of these, only the *clr3* deletion had any effect on *pho1* expression levels (Fig. 2-2, compare lane 3 to lane 1). Encouragingly, the extent of derepression compared to wild-type RNA is similar to what was previously seen when cells were treated with TSA, indicating that Clr3 is likely to be the only HDAC contributing to *pho1* repression.

Since Clr3, as part of SHREC, has been shown to be recruited to pericentromeric heterochromatin by the RNA-binding protein Seb1 (Marina et al., 2013), it is reasonable to speculate that the same might be true at this euchromatic locus. Therefore, the influence of Seb1 on the expression of *pho1* was tested. Strikingly, Northern blot analysis revealed a substantial increase in *pho1* mRNA levels in a *seb1-1* mutant strain (Fig. 2-2, compare lanes 1 and 2). Furthermore, the *seb1-1* mutant gave a phenotype markedly similar to that of *clr3*Δ (Fig. 2-2, lane 3). In agreement with previously reported data (Shah et al., 2014), deletion of Clr4 and the HP1 proteins Swi6 and Chp2 resulted in an increase in *pho1* expression (Fig. 2-2, lanes 4-6).

3.2.2 Clr3 is co-transcriptionally recruited to the *prt* region

Previous studies have revealed Clr3 binding to centromeres, telomeres, rDNA and the *mat* locus (Wiren et al., 2005; Yamada et al., 2005). Moreover, as part of SHREC, genome-wide ChIP-chip analyses show that Clr3 appears to be able to target a specific

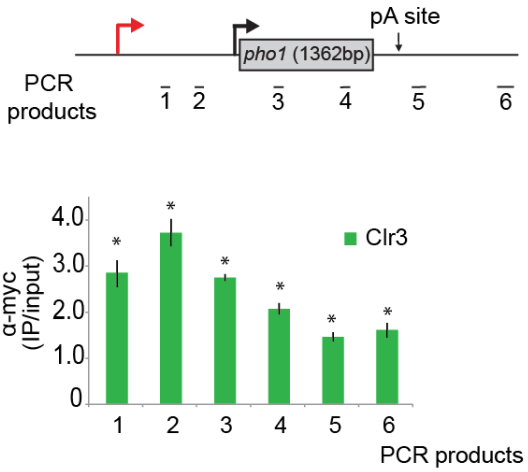


Figure 2-3: Clr3 is recruited to the *pho1* locus
(Top) Schematic of *pho1* gene with PCR products used for qPCR indicated by black bars. (Bottom) Clr3 ChIP was performed in a myc-tagged strain grown in YES. Quantification of results from three independent experiments is shown; error bars indicate the standard error. Statistical differences between tagged and untagged strains were analysed using the Student's t-test and evaluated using P-values. P<0.05 was considered to indicate a statistically significant difference between tagged and untagged strains and is represented by an asterisk (*).

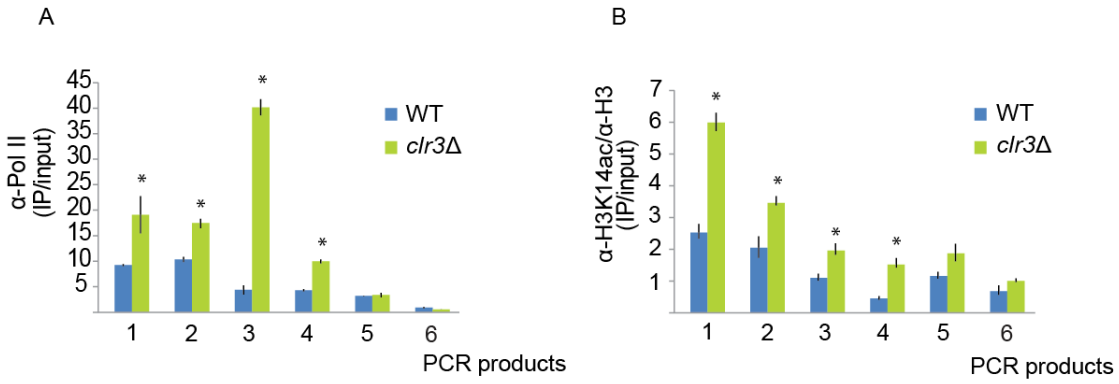


Figure 2-4: Lack of Clr3 correlates with elevated Pol II occupancy and H3K14 acetylation levels at the *pho1* locus
(A) Pol II ChIP analysis performed in *clr3*Δ strain compared to wild-type (WT) levels. (B) ChIP analysis of H3K14ac performed in *clr3*Δ compared to wild-type (WT) levels. Quantification of results from three independent experiments is shown; error bars indicate the standard error. Statistical differences between WT and *clr3*Δ strains were analysed using the Student's t-test and evaluated using P-values. An asterisk (*) indicates a significant difference (P<0.05) compared to WT. See (Fig. 2-3) for position of PCR products.

selection of protein-coding genes and non-coding RNAs (Sugiyama et al., 2007). However, whether its presence has any functional consequence was not addressed.

To further understand the role of Clr3 at the *pho1* euchromatic locus, chromatin immunoprecipitation (ChIP) was performed to determine its binding profile. A strain expressing carboxy-terminally myc-tagged Clr3 was used for this purpose. Interestingly, ChIP analysis across the *pho1* locus revealed that Clr3 specifically localises to the non-coding region (Fig. 2-3). Together with the observations that TSA treatment and loss of Clr3 result in derepression of *pho1*, these data suggest that Clr3 is, by an as yet unknown mechanism, recruited to the upstream ncRNA *prt* to elicit *pho1* silencing.

3.2.3 Lack of Clr3 correlates with elevated Pol II occupancy and H3K14 acetylation

The obvious question arises, then, of how Clr3 exerts its silencing effect on *pho1*. Previous studies at the *mat* locus have revealed that Clr3 is able to block access of Pol II to these naturally silenced sequences (Yamada et al., 2005). So, to further examine the role of Clr3 in repressing expression of *pho1* mRNA, the occupancy of Pol II at the locus was determined by ChIP analysis (Fig. 2-4A). Interestingly, compared to wild-type cells, ablation of *clr3* leads to increased Pol II levels upstream of the *pho1* promoter and particularly across the gene body. This was also found to be the case, though to a lesser extent, with the *seb1-1* mutant strain (Fig. 2-5A). These data suggest that Clr3 is able to restrict access of the transcription machinery to, and therefore repress expression of, *pho1*.

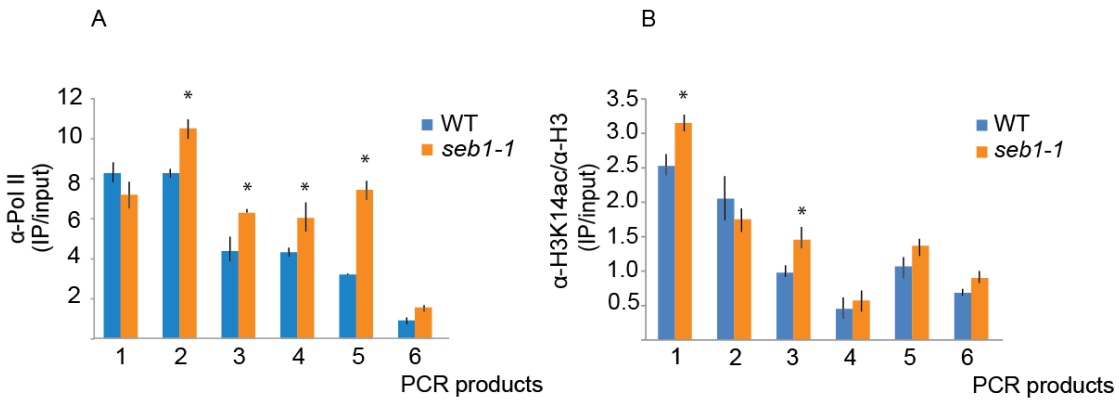


Figure 2-5: Pol II occupancy and H3K14 acetylation levels are partially elevated in *seb1-1* mutant cells at the *pho1* locus
(A) Pol II ChIP analysis performed in *seb1-1* strain compared to wild-type (WT) levels. (B) ChIP analysis of H3K14ac performed in *seb1-1* compared to wild-type (WT) levels. *seb1-1* was grown at permissive temperature (30ºC) and shifted to restrictive temperature (37ºC) for two hours. Quantification of results from three independent experiments is shown; error bars indicate the standard error. Statistical differences between WT and *seb1-1* strains were analysed using the Student's t-test and evaluated using P-values. An asterisk (*) indicates a significant difference ($P<0.05$) compared to WT. See (Fig. 2-3) for position of PCR products.

Since Clr3 possesses histone deacetylase activity, one might surmise that the silencing defect found at *pho1* in *clr3*Δ (Fig. 2-2) is a result of a loss of Clr3's ability to remove this mark of active histones. To test this, levels of H3K14 acetylation (H3K14ac) were analysed by ChIP. A reproducible increase in the levels of H3K14ac was observed in the *clr3*Δ mutant (Fig. 2-4B); similar to what has been shown previously at pericentromeric heterochromatin (Sugiyama et al., 2007). Concomitantly, the biggest increase in H3K14ac occurred over the *prt* non-coding region. Additionally an, albeit milder, increase in H3K14ac was also detected in *seb1-1* (Fig. 2-5B). The salient implication of these data is that Clr3 functions to silence *pho1* mRNA by a mechanism that depends on its HDAC activity.

Clr3 has previously been shown to function as part of the multienzyme effector complex SHREC. As well as Clr3, SHREC is composed of Clr1, Clr2, and Mit1. Whilst the specific functions of Clr1 and Clr2 remain unknown, it has recently been reported that Mit1 can, in an ATP-dependent manner, mobilise histone octamers to remodel nucleosomes and silence transcription (Creamer et al., 2014). Clr3 can, however, localise to several sites independently of all, or part of, the other SHREC components. In order to determine whether Clr3 functions at *pho1* as part of SHREC or independently of the complex, the effect of deletion of the *clr1*, *clr2*, and *mit1* components on RNA levels was analysed by Northern blot (Fig. 2-6). Compared to wild-type *pho1* mRNA levels (Fig. 2-6, lane 1), a slight accumulation was detected for both *clr1*Δ and *clr2*Δ strains (Fig. 2-6, lanes 2 and 3). This minor effect may reflect the fact that neither of these proteins have been shown to have any enzymatic activity.

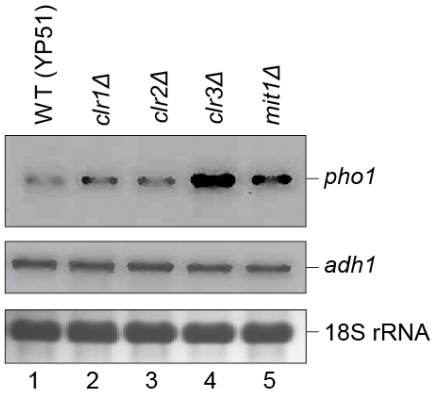


Figure 2-6: The contribution of SHREC to *pho1* silencing
Northern blot analysis of RNA isolated from strains harbouring SHREC component deletions (*clr1*, *clr2*, *clr3*, and *mit1*) compared to wild-type (WT) on *pho1* expression. The housekeeping gene *adh1* and 18S ribosomal RNA serve as loading controls. *pho1* probe is as used in (Fig. 2-1).

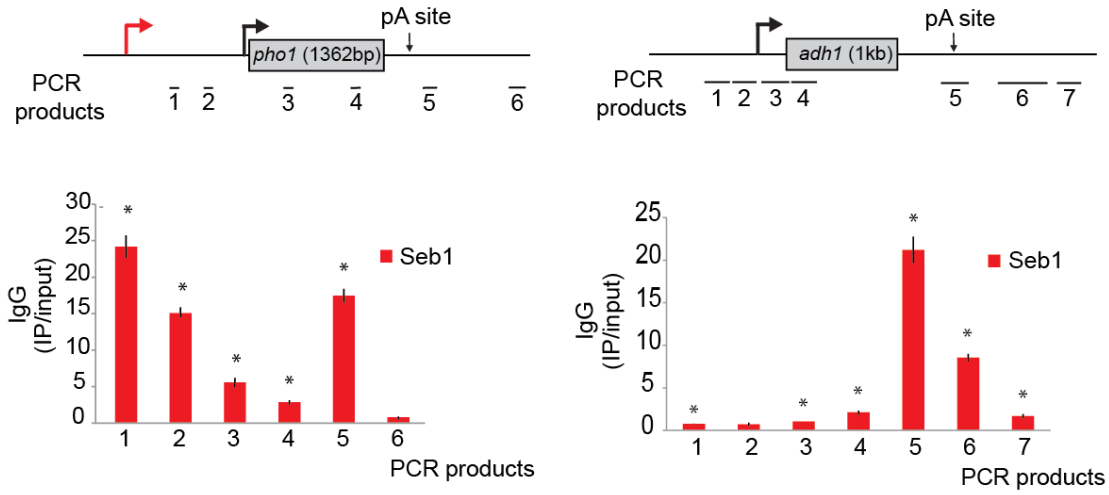


Figure 2-7: Seb1 is recruited to the *pho1* non-coding transcript and the 3'end of protein-coding genes
Schematic representation of *pho1* (left) and *adh1* (right) genes are shown. Black bars indicate locations of PCR products. Seb1 ChIP experiments were performed in a TAP-tagged Seb1 strain grown in YES. Quantification of results from three independent experiments is shown; error bars indicate the standard error. Statistical differences between tagged and untagged strains were analysed using the Student's t-test and evaluated using P-values. An asterisk (*) indicates a significant difference ($P<0.05$) compared to the untagged strain.

On the other hand, *mit1*Δ does evince a more pronounced silencing defect on *pho1* levels (Fig. 2-6, lane 5) - somewhere intermediate between the single *clr1*Δ/*clr2*Δ deletions and *clr3*Δ. As already seen (Fig. 2-2) *clr3*Δ leads to a strong derepression of *pho1* (Fig. 2-6, lane 4) and, perhaps not surprisingly considering its enyzmatic activity, has the strongest effect of all SHREC components.

Somewhat perplexingly considering their effects on *pho1* mRNA levels (Fig. 2-6), the other SHREC subunits Clr1, Clr2, or Mit1 were not detected by ChIP at the locus (data not shown). Moreover, no signal was detected for any of these three SHREC components at *dg* and *dh* repeats (where they have previously been shown to associate (Marina et al., 2013)). It cannot be excluded that the lack of ChIP qPCR signal at *pho1* may be because the large TAP tag interferes with binding of these proteins. Whilst there may be a lack of evidence to conclusively assert that Clr3 functions as part of SHREC in silencing *pho1*, it is quite likely, based on the changes in mRNA levels observed in deletion strains (Fig. 2-6) that this is the case.

3.2.4 Seb1 is recruited to the *pho1* non-coding transcript and the 3'end of protein-coding genes

Since *seb1-1* and *clr3*Δ mutants show a strikingly similar effect on *pho1* mRNA levels (Fig. 2-2), and the two have been shown to physically associate *in vivo* (Marina et al., 2013), the question was asked whether Seb1, like Clr3, is recruited to the *pho1* locus. Indeed, Seb1 is enriched across the *prt* region upstream of *pho1* (Fig. 2-7), and, surprisingly, is also present at the pA site. This is in contrast to other protein-coding genes that were tested, including the highly expressed house-keeping gene *adh1* (Fig.

2-7) and the gene encoding ribosomal protein S15a, *rps2202* (data not shown), where Seb1 recruitment is restricted only to the pA site.

Taken together with the effect of *seb1-1* on expression (Fig. 2-2), these data suggest that, under conditions of phosphate availability, Seb1 is recruited to the *pho1* locus to elicit silencing; possibly through mediating recruitment of the SHREC component and HDAC Clr3 to the non-coding transcript.

3.2.5 Clr3 and Clr4 function independently to silence *pho1*

Since it has previously been suggested that Seb1 recruitment of SHREC at pericentromeric heterochromatin promotes H3K9me2 (Marina et al., 2013) - indicating that SHREC (and therefore Clr3) functions in the same pathway as the Clr4 methyltransferase - it was desirable to test whether this is also the case at the *pho1* locus. In order to address this, the double mutant *clr3*Δ*clr4*Δ was generated. Remarkably, Northern blot revealed that, whilst *clr3*Δ and *clr4*Δ single mutants show a modest increase in *pho1* mRNA, their respective double mutant results in an additive accumulation (Fig. 2-8A, compare lanes 2 and 3 to lane 4). Indeed, this is also reflected by the finding that the double mutant has a slow growth phenotype considerably more severe than either of the single mutants, *clr3*Δ or *clr4*Δ (Fig. 2-8B). These data demonstrate that Clr3 and Clr4 act independently of each other to promote silencing of *pho1*. Importantly, this is in striking contrast to data which has previously been presented indicating that deacetylation is linked to Clr4 methylation by a linear pathway (Marina et al., 2013).

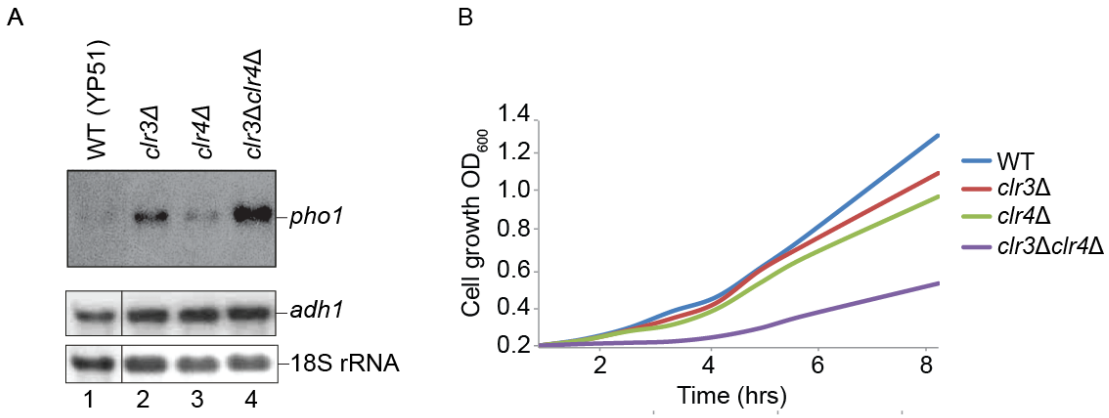


Figure 2-8: Clr3 acts independently of Clr4 to promote *pho1* silencing
(A) Northern blot analysis of *pho1* mRNA levels in the indicated strains grown in YES. The housekeeping gene *adh1* and 18S ribosomal RNA serve as loading controls. *pho1* probe is as used in (Fig. 2-1). (B) Growth profile of indicated strains grown in YES media.

3.3 Discussion

The classical hallmarks of heterochromatin are characterised by high levels of histone methylation and low levels of histone acetylation. Since deletion of the only fission yeast methyltransferase does not lead to complete derepression of the *pho1* gene it follows that there must exist another contributing factor. Here it is shown that deacetylation by the HDAC Clr3 is, at least part of, this missing link.

As suspected, deletion of the other SHREC subunits Clr1, Clr2, and Mit1 resulted in some (varied) changes to *pho1* expression (Fig. 2-6). However, as noted only the Clr3 component can be detected at the locus by ChIP (Fig. 2-3). This may not necessarily mean Clr3 is acting completely independently of SHREC, but could instead merely reflect the dynamic nature of the complex. Thus, whilst it seems probable based on these data and others (which show more or less consistent co-localisation of the four SHREC subunits to most genomic locations) (Sugiyama et al., 2007), further verification is needed to support the notion that Clr3 functions as part of the multienzyme SHREC to elicit *pho1* silencing.

ChIP analyses show that, at all protein-coding genes tested, Seb1 associates to pA sites. This is an interesting facet given the role of the *S. cerevisiae* homologue Nrd1 as part of the NNS complex in terminating transcription at ncRNA genes. Is Seb1 also involved in transcription termination? And how does this differ to the situation in budding yeast where there are two distinct pathways for termination of mRNA-coding and ncRNA genes? Already it can be seen that Seb1's recruitment to these mRNA gene pA sites is distinct from Nrd1 binding. Further, if Seb1 is indeed

responsible for recruitment of Clr3 why is the HDAC only present at the region encoding *prt* and not the *pho1* pA site - how does Seb1 distinguish between the two? These are questions that will be considered in more detail in the following chapter.

Unlike the case at pericentromeric heterochromatin, where it has been shown that Seb1 functions in the same pathway as SHREC to promote H3K9me, deletion of both *clr3* and *clr4* resulted in an additive accumulation of *pho1* mRNA compared to the respective single mutants *clr3*Δ and *clr4*Δ. This indicates that deacetylation of H3K14 and methylation of H3K9 are not entirely interdependent in their influence on *pho1* repression. Interestingly however, the double mutant *clr3*Δ*clr4*Δ does not lead to as strong a silencing defect as deletion of *prt* non-coding transcription. Whether this reflects the involvement of yet another ncRNA-mediated silencing mechanism is something which requires further work. What is clear is that the *pho1* locus provides an excellent model for dissecting the diverse mode of action of non-coding RNAs and their transcription.

4. A ncRNA is required for Seb1 and Clr3 recruitment to the *pho1* locus

4.1 Introduction

Thus far it has been shown that Seb1 is important for transcriptional gene silencing of *pho1* mRNA. This is likely due to the recruitment of the HDAC Clr3 to remove active histone marks and block access of the transcription machinery. Seb1 has two domains which allow it to select its target genes - a CID for Pol II binding and an RRM for RNA interactions. This of course raises the question of whether Seb1, in associating with the *pho1* locus in particular, interacts with RNA or the Pol II CTD, or indeed both.

The role of the CTD in recruitment of RNA processing factors is well established in budding yeast and includes recruitment of capping enzyme (Cho et al., 1997), elongation factors (see Hsin and Manley, 2012), as well as the termination factors Nrd1 (Vasiljeva et al., 2008), Pcf11 (Licatalosi et al., 2002), and Rtt103 (Kim et al., 2004). Each of these termination factors share the highly conserved common fold known as the CID. However, some other known CTD-interacting proteins, including capping enzyme and various elongation factors, do not possess this conventional domain. This makes it difficult to bioinformatically predict CTD-interacting proteins based on amino acid compostion alone and, therefore, protein interactions need to be experimentally tested.

Despite progress in many model systems, the role of the CTD in gene expression in *S. pombe* has not been addressed. This system is ideal since it can be readily grown and manipulated in the laboratory using a variety of sophisticated molecular and genetic methodologies. More importantly it shares many molecular, genetic, and biochemical features with humans that are not present in the well-studied budding yeast *S. cerevisiae*, rendering it a particularly useful experimental model in, for example, studying the function and expression of genes related to various human diseases. In addition, the *S. pombe* CTD contains few heptads that deviate from the consensus sequence, meaning it is easier to assess its structural and functional aspects than the larger, more complex CTDs of some other organisms.

For these reasons *S. pombe* CTD interactors were investigated, in particular those proteins which specifically interact with hyperphosphorylated (and therefore transcriptionally active) polymerase. Here, it is shown that various proteins with known or predicted roles in transcription and/or RNA processing associate with the CTD. Importantly, the Nrd1 homologue Seb1 is identified and verified as an interactor of CTD phosphorylated at serine 2 and serine 5. It is highly likely that this interaction is mediated by the Seb1 CID. Furthermore, these studies reveal Seb1 as associating with several key transcriptional regulators, predominantly 3' end processing factors.

Additionally in this chapter the role of Seb1 at the *pho1* locus is considered in more detail, in particular how it recognises the gene as a target for transcriptional repression. The data presented here point to non-coding transcription as being essential for Seb1 to exert its silencing function. Lastly, binding of Seb1 to the *prt* ncRNA is addressed. It is shown that two particular regions of *prt* are important for

repression of *pho1*.

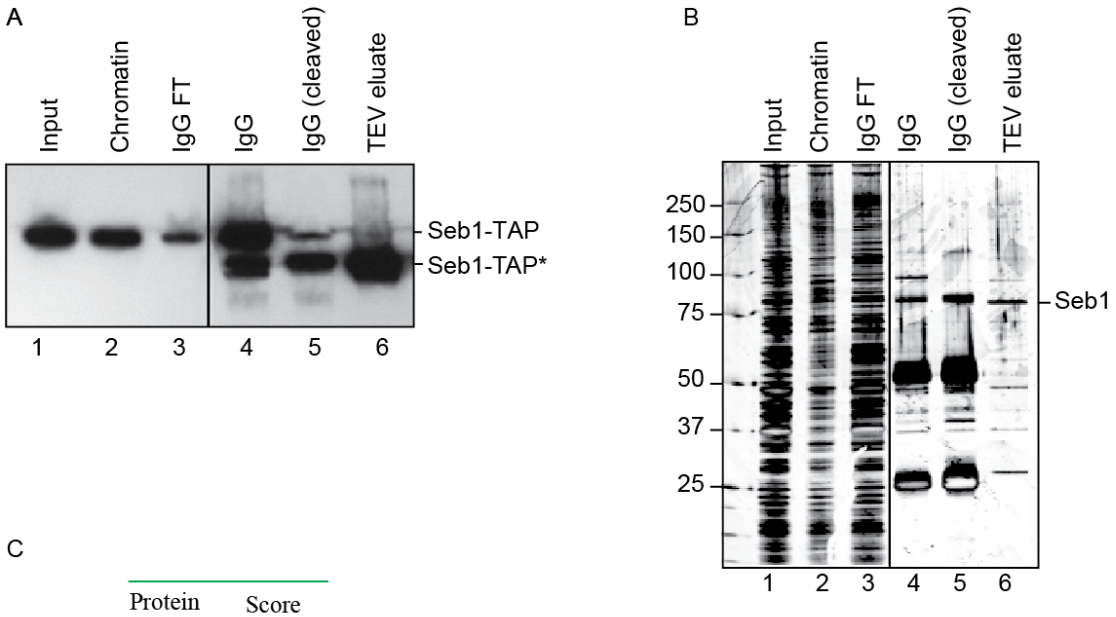
4.2 Results

4.2.1 Seb1 interactions

It has previously been shown by yeast two-hybrid experiments that Seb1 is able to interact with Pol II via the Rpb7 subunit (Mitsuzawa et al., 2003). In addition, co-immunoprecipitation experiments have revealed that Seb1 is able to physically associate with each of the SHREC components Clr1, Clr2, Clr3, and Mit1, and, at least for Clr3 this interaction is not mediated by RNA (Marina et al., 2013).

To glean further information regarding Seb1's role in the regulation of RNA transcription, the protein was purified via a TAP-tag that was engineered to be expressed at its C-terminal end. Western blot analysis (Fig. 3-1A) tracking the purification process revealed that Seb1 is enriched on the IgG beads used for purification and a shift in band size coincides with cleavage of the protein A moiety of the TAP-tag by TEV protease (Fig. 3-1A, lanes 4 and 5). Silver stain analysis of the Seb1-TAP purified fraction (Fig. 3-1B, lane 6) detected various polypeptides (most likely representing products of partial degradation), together with a prominent band at ~ 75 kDa. Mass spectrometry analysis confirmed enrichment for Seb1.

As expected, several subunits of Pol II were recovered, of which the top three are listed (Fig. 3-1C). Of these, the largest subunit of Pol II, Rpb1, had the highest Mascot score - perhaps indicating an interaction between the Rpb1 CTD and Seb1's CID. In agreement with previously published data (Mitsuzawa et al., 2003), Seb1 also co-immunoprecipitated with the Rpb7 subunit of Pol II (not shown).



C

Protein	Score	
Pab1	961	3'end processing factors
Rna14	733	
Cft1	387	
Cft2	233	
Ysh1	207	
Pta1	169	
Spp42	518	Splicing factors
Prp39	506	
Cwf8	319	
Sn114	314	
Utp10	235	
Brr2	213	
Elf1	565	Export factors
Uap56	345	
Mex67	46	
Mlo3	24	

Protein	Score	
Nap1	111	Chromatin modifiers
Ssr2	68	
Ssr1	22	
Rnc1	417	RNA-binding proteins
Puf6	407	
SPAC17H9.04c	299	
SPCC1827.05c	279	
Mmi1	258	
SPAC644.16	159	
Rpb1	346	Other
Rpb3	180	
Rpb2	135	
Mrt4	133	

Figure 3-1: Seb1 purification
(A) Western blot analysis tracking Seb1-TAP purification procedure. The chromatin fraction includes cell debris that was pelleted from the input sample. The asterisk represents Seb1-TAP which has been cleaved of its protein A moiety by TEV protease. (B) Silver stain analysis of proteins purified from Seb1-TAP strain. (C) Table showing a subset of proteins and their Mascot score that were associated with purified Seb1-TAP as identified by LC-MS/MS analysis. Criteria for a positive hit included identification of at least two peptides and a Mascot score >20.

In addition to Pab1 and Rna14 (with scores of 961 and 733 respectively), at least four other proteins with previously described roles in 3' end processing were recovered (Fig. 3-1C). Interestingly, this fits with Seb1 interacting with the 3' ends of protein-coding genes (Fig. 2-7) and suggests the protein may play an important role in cleavage/polyadenylation and/or transcription termination.

Another strongly enriched protein was the RNA binding and YTH domain protein Mmi1. Convincingly, the reciprocal purification of Mmi1-HTP reproducibly pulls down Seb1 (personal communication, Heo Dong-Hyuk). HTP is based on the TAP tag except the calmodulin binding peptide is replaced with six histidines. Surprisingly, Clr3 was the only component of SHREC that was identified to co-purify with Seb1 (not shown), and given the low number of peptides detected does not represent one of the more credible hits. A significant number of proteins with roles in splicing were also pulled down (see Fig. 3-1C for top hits). Finally, various RNA binding proteins were recovered in the purification. It is unclear whether these represent direct interactions or are perhaps bridged by RNA since Seb1 also has an RNA recognition motif (RRM). This is something that purifications with RNase and DNase treatment could divulge. The purification also contained various ribosomal proteins that were ignored in these analyses as they are the common false positives associated with TAP purifications (Gavin et al., 2002 and Ho et al., 2002).

4.2.2 Identification of CTD-interacting proteins in *S. pombe*

From the data described above (Fig. 3-1) it appears that Seb1 is able to interact with the Pol II Rpb1 subunit. Since Seb1 possesses at its extreme N-terminus a CID, one

might presume that the protein can associate with the CTD. In order to investigate whether there is truth behind this assumption, and indeed to explore what other fission yeast CTD interacting proteins exist, an *in vitro* system was established. A series of commercially synthesised and biotinylated CTD peptides consisting of four heptapeptide repeats with different phosphorylation states (unphosphorylated, Ser2-(P), Ser5-(P), and Ser7-(P)) were used. It has previously been shown that four repeats are sufficient for CTD binding by Nrd1, Pcf11, Rtt103 and capping enzyme in *S. cerevisiae* (Vasiljeva et al., 2008; Licatalosi et al., 2002; Kim et al., 2004; Ho et al., 1999).

In order to gauge the functionality of this system a known CTD interacting protein from budding yeast was used as a control to validate the approach. Unphosphorylated and Ser2-(P) peptides were immobilised onto streptavidin beads and used to isolate Rtt103, a protein involved in transcription termination, from *S. cerevisiae* whole cell extracts (Fig. 3-2A). Rtt103, which has a CID, has previously been shown to specifically interact with Ser2-(P) CTD (Kim et al., 2004). Extracts were prepared from HA-tagged Rtt103 and incubated with the indicated peptide-bound beads. Bound Rtt103 was monitored by western blot with α-HA antibody and, as predicted, the protein specifically binds Ser2-(P) CTD (Fig. 3-2A, lane 5), but not unphosphorylated CTD (Fig. 3-2A, lane 4), or beads with no bound peptide (Fig. 3-2A, lane 3). Depletion of Rtt103 is also observed in the flow-through fraction of Ser2-(P) (Fig. 3-2A, compare lanes 6 and 7 to lane 8). Additionally, the *S. pombe* homologue Rhn1, which also contains a CID (Sugiyama et al., 2012), was found, by this same approach, to associate with Ser2-(P) CTD (Fig. 3-2B, lane 4). It appears, therefore, that this system can be used to pull down CTD interacting proteins. Accordingly, the system

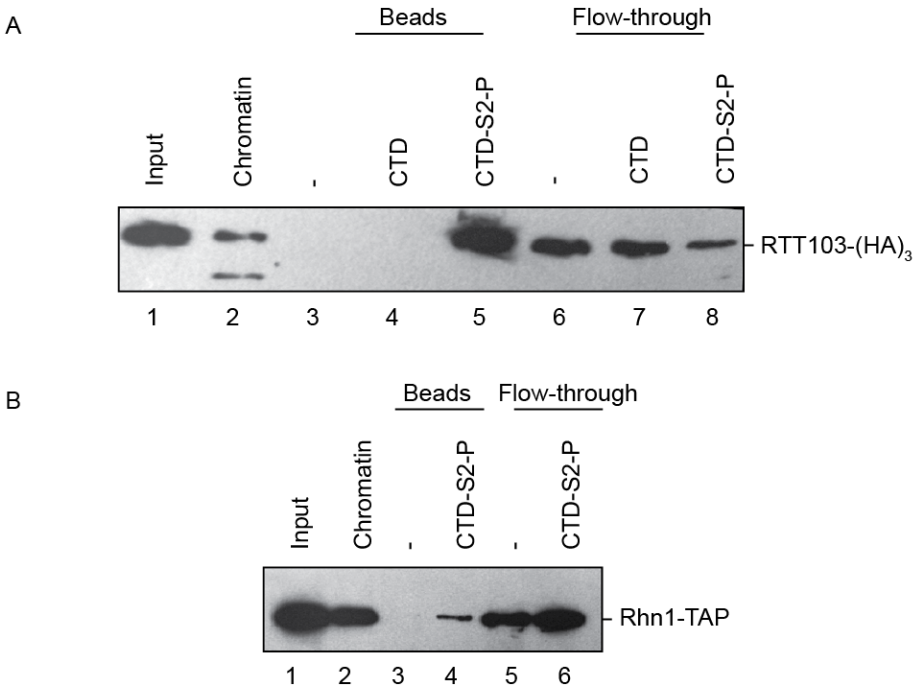


Figure 3-2: Verification of CTD-binding assay
(A) *S. cerevisiae* extract from a triple haemagglutinin $(HA)_3$-tagged Rtt103 strain and (B) *S. pombe* extract from a TAP-tagged Rhn1 strain were incubated with streptavidin-coated dynabeads pre-bound to CTD peptides of the indicated phosphorylations. Bound Rtt103 was visualised with α-HA antibody (12CA5) and Rhn1 with PAP antibody. Chromatin fractions includes cell debris that was pelleted from the input sample.

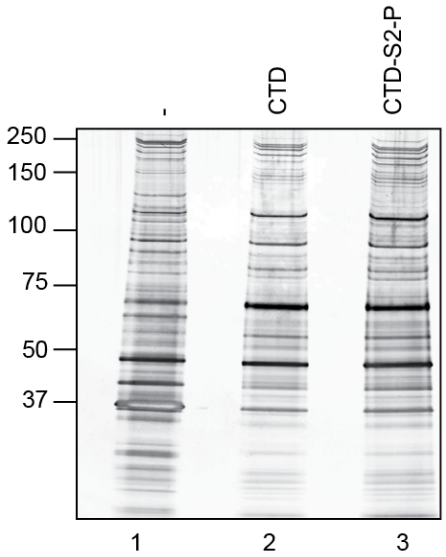


Figure 3-3: Silver stain analysis of proteins purified from CTD peptides
Streptavidin-coated dynabeads pre-bound to CTD peptides of the indicated phosphorylations (- indicates no peptide) were used to purify proteins from wild-type *S. pombe* extracts.

was applied to *S. pombe* to identify unphosphorylated and Ser2-(P) CTD binding proteins by mass spectrometry and, more specifically, to establish whether Seb1 is one of these.

Yeast whole cell extracts were prepared from wild-type (YP51) *S. pombe*. Extracts were incubated with peptide-bound beads and after washes were analysed by silver staining. A large number of polypeptides were detected for the unphosphorylated and Ser2-(P) peptide pull-downs (Fig. 3-3, lanes 2 and 3). Somewhat disconcertingly many of the same bands were also present in the negative control (that is, beads with no bound peptide) (Fig. 3-3, lane 1) indicating unspecific binding. Interestingly, it appears that more of the larger polypeptides (those between ~125-250 kDa) were present in the unphosphorylated CTD pull-down, and particularly the Ser2-(P) pull-down. Polypeptides from the Ser2-(P) CTD purification were identified by mass spectrometry, the results of which are summarised (Fig. 3-4). Importantly, only those proteins which were completely absent from the negative control, or at least three times enriched in the CTD pull-down are presented. It must also be stressed that, whilst some insights into the *S. pombe* CTD and its interactors may be gleaned from these experiments, these are only preliminary data which must be viewed as such until further replications can allow for complete confidence. Nevertheless, as might be anticipated, various proteins with roles in splicing, transcription, processing and export were detected for both pull-downs.

Consistent with their CTD specificity in *S. cerevisiae* both Pcf11 and Rhn1 (ScRTT103) were found to specifically co-elute with Ser2-(P). Surprisingly, (though consistent with Rpb1 purifying with Seb1 (Fig. 3-1)), Seb1 was found to interact with

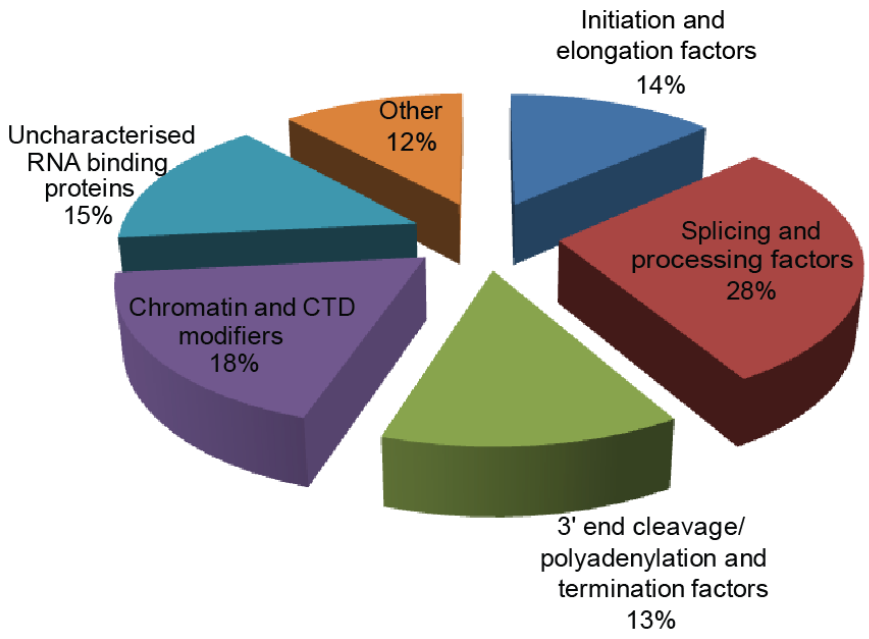


Figure 3-4: Mass spectrometry analysis
Pie chart showing the functional roles and percentages of proteins pulled down with Ser2-(P) CTD peptide.

Ser2-(P) CTD but not unphosphorylated CTD (not shown). This is particularly striking when one considers that Seb1's *S. cerevisiae* homologue Nrd1 is specific to Ser5-(P) and does not interact with Ser2-(P) (Vasiljeva et al., 2008).

4.2.3 Seb1 physically associates with the RNA polymerase II (Pol II) C-terminal domain (CTD)

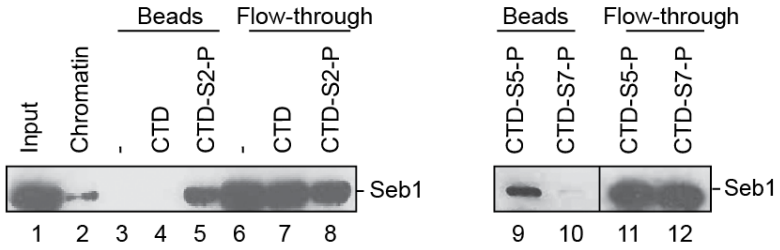
In order to investigate the interaction of Seb1 with the different phosphoisoforms of the Pol II CTD (Figs. 3-1 and 3-4), two experiments were performed (Fig. 3-5). In the first, Seb1-TAP extracts were incubated with streptavidin-coated dynabeads pre-bound to CTD peptides (Fig. 3-5A), the peptides being either unphosphorylated or phosphorylated on one of the three serines. To rule out the possibility that Seb1 is able to non-specifically interact with the beads alone, extracts were also incubated with beads in which no peptide had been pre-bound (Fig. 3-5A, lane 3). After incubation and washes, Seb1 binding to beads was monitored by western blot with peroxidase anti-peroxidase (PAP) antibody. A prominent band was detected for both the Ser2-(P) and Ser5-(P) CTD peptides (Fig. 3-5A, lanes 5 and 9), indicating that Seb1 specifically interacts with these forms of the CTD. A slight depletion of Seb1 was also detected in the flow-through fraction of Ser2-(P) (Fig. 3-5A, compare lane 8 to lane 6). A very faint band was also discernible for Ser7-(P) (Fig. 3-5A, lane 10), whereas no signal was detected for the unphosphorylated peptide or indeed the no-peptide control, as expected (Fig. 3-5, lanes 3 and 4).

To further confirm the interaction of Seb1 with the CTD, co-immunoprecipitations were performed (Fig. 3-5B). This was deemed necessary since, whilst CTD peptide experiments can be informative, they are limited in that the peptides are much shorter

and more uniform than the endogenous CTD, and therefore do not entirely simulate its conformational flexibility and complexity. Further, phospho-CTD specificities may also require the transcriptional context.

In an alternative approach, the different Pol II phosphoisoforms which associate with Seb1 were analysed. Seb1-TAP extract was co-immunoprecipitated with IgG agarose beads and eluted by TEV protease cleavage of the TAP-tag. Using specific antibodies, the resulting eluate was tested for the presence of the various phosphorylated forms of the CTD (Fig. 3-5B, top panel). Levels of Seb1-TAP were visualised with α-HA antibody and serve as a loading control (Fig. 3-5B, bottom panel). Consistent with peptide binding (Fig. 3-5A), Rpb1 was detected in the eluted fraction using antibodies recognising both Ser2-(P) and Ser5-(P) (Fig. 3-5B, lanes 4 and 6). However, no Rpb1 band was detectable in the eluted fraction when using either antibodies recognising unphosphorylated CTD (8WG16) or α-Ser7-(P) (Fig. 3-5B, lanes 2 and 8). Taken together these data demonstrate that Seb1 is able to physically associate with the Pol II CTD *in vivo* when it is phosphorylated at positions Ser2 and Ser5.

A



B

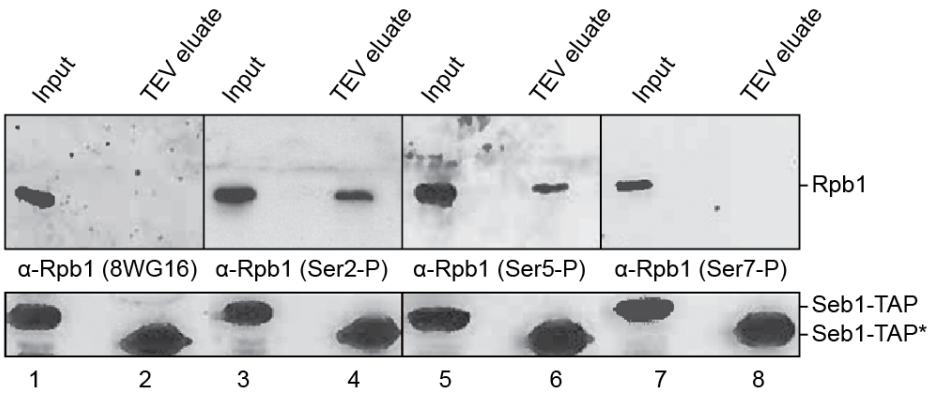


Figure 3-5: Seb1 physically associates with the RNA polymerase II (Pol II) C-terminal domain (CTD) (A) Extracts prepared from a strain expressing endogenously tagged Seb1-TAP were incubated with streptavidin-coated dynabeads pre-bound to CTD peptides of the indicated phosphorylations. After washes, bound Seb1 was visualised with PAP antibody. The chromatin fraction includes cell debris that was pelleted from the input sample. (B) Co-immunoprecipitations of Seb1 with Rpb1. A strain expressing Seb1-TAP was subjected to IgG-agarose immunoprecipitation. The input and immunoprecipitated (TEV eluate) samples were detected with α-Rpb1 antibodies 8WG16, Ser2-(P), Ser5-(P), or Ser7-(P). Seb1-TAP was detected with α-HA (bottom panel) and serves as a loading control. Note that the shift in size between Seb1-TAP and Seb1-TAP* reflects cleavage of the protein A moiety by TEV protease.

4.2.4 Non-coding transcription is required for Seb1 and Clr3 recruitment

As well as a CID, Seb1 also has an RRM, indicative of the capacity to bind RNA. Considering ChIP experiments at *pho1* (Fig. 2-7) which shows two peaks of Seb1 recruitment - one at the region encoding *prt* ncRNA and another at the pA site - it was speculated that Seb1 might distinguish between the two sites by binding separately via its RRM and CID domains. Given also that *prt* has also been implicated in the recruitment of Mmi1 to *pho1* (Shah et al. 2014 and Fig. 3-6C), it is possible that Seb1 and Clr3 recruitment is also RNA-dependent.

In order to test whether Seb1 recruitment to chromatin is mediated, at least in part, via RNA, ChIP was performed on chromatin extracts that were either treated with or without RNase (Fig. 3-6). The known *prt*-binding protein Mmi1 was used as a positive control for loss of binding, together with Pol II at *pho1* and Pol II and Seb1 at the constitutively expressed gene *adh1* as negative controls (Fig. 3-6C, D, E, F). The data presented (Fig. 3-6A) indicate that Clr3 binding to chromatin is at least partially dependent, whereas Seb1 recruitment (Fig. 3-6B) seems to be entirely dependent on RNA at both the non-coding region and the pA site. To rule out the possibility that RNase treatment inadvertently impairs Seb1 binding to DNA, Seb1 ChIP was performed at *adh1* where its recruitment was not found to be affected in RNase treated versus untreated samples (Fig. 3-6F). This suggests that Seb1 is recruited to the non-coding transcript and 3' end of protein-coding genes through different mechanisms.

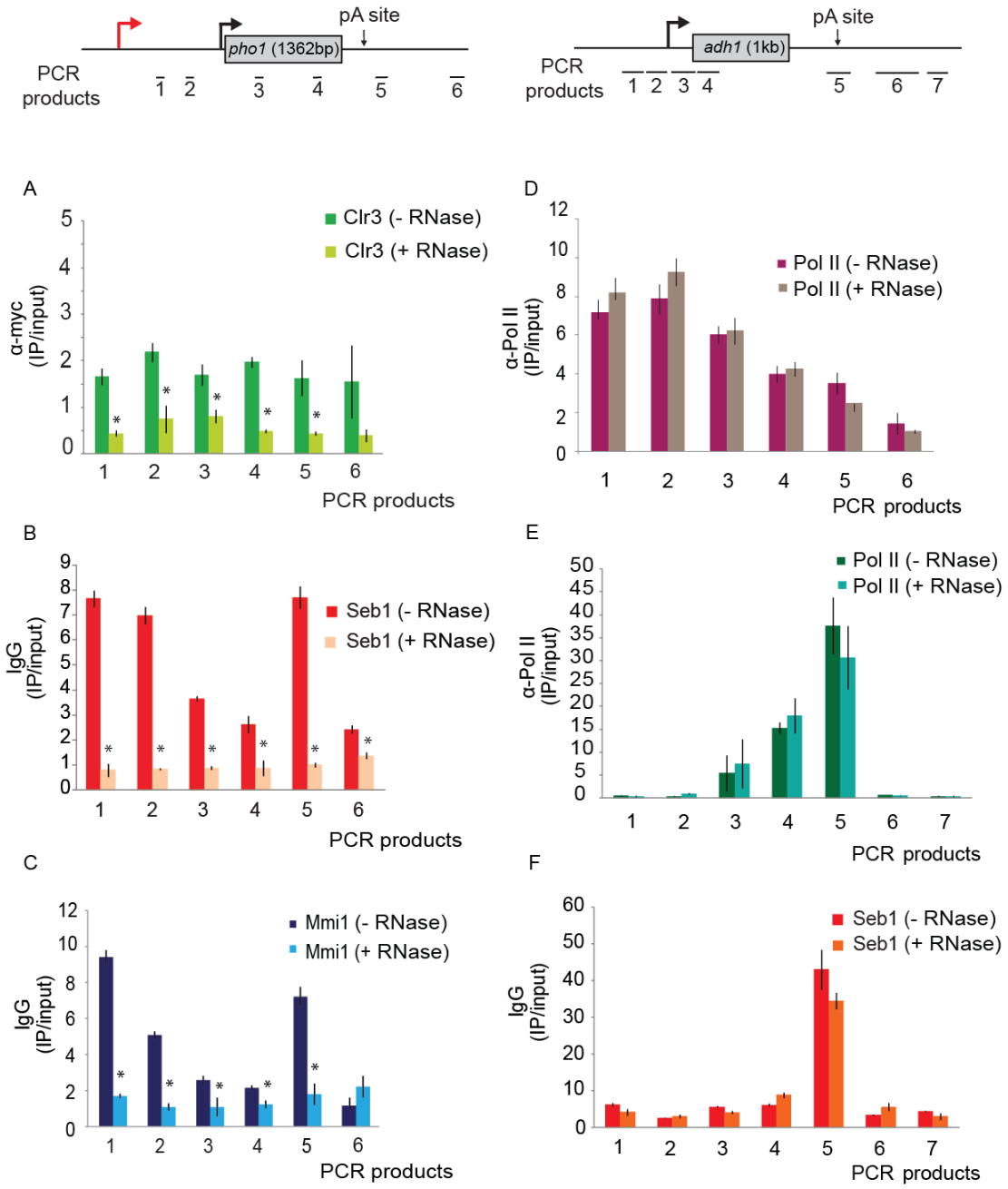


Figure 3-6: Non-coding RNA is required for Seb1 and Clr3 recruitment
(Top) Schematics of *pho1* and *adh1* genes with PCR products used for qPCR indicated by black bars. (A, B, C, D) ChIP analysis of myc-tagged Clr3, TAP-tagged Seb1, Mmi1-HTP (tag based on TAP except 6xhis replaces CBP) and Pol II over the *pho1* locus either untreated (-) or treated (+) with RNase. (E, F) ChIP analysis of Pol II and Seb1-TAP over the *adh1* locus either untreated (-) or treated (+) with RNase. Quantification of results from three independent experiments is shown for all ChIP analysis; error bars indicate the standard error. Statistical differences between treated and untreated samples were analysed using the Student's t-test and evaluated using P-values. An asterisk (*) indicates a significant difference ($P<0.05$) compared to untreated sample.

In addition to RNase treatment it was found that deletion of the *prt* promoter (which results in loss of non-coding transcription (Shah et al. 2014)) leads to the simultaneous loss of Clr3 from this region (Fig. 3-7A). This also coincides with increased levels of H3K14ac - comparable to that seen in *clr3*Δ (Fig. 2-4B and Fig. 3-7B). Moreover, since Clr3 recruitment seems to depend on the production of the ncRNA one would expect it to operate downstream in the pathway for silencing *pho1* expression. To test this, Northern blot analysis was performed to compare the effect of the single deletions *clr3*Δ and *ncpro*Δ (in which the *prt* promoter is deleted) to the double mutant *clr3*Δ*ncpro*Δ (Fig. 3-7C). Combining *clr3*Δ with *ncpro*Δ has no obvious additive effect, indicating that Clr3 does indeed operate downstream of *prt* transcription. *ncpro*Δ compared to *clr3*Δ has a more pronounced influence on *pho1* derepression (Fig. 3-7C, compare lanes 2 and 3). This is likely due to the fact that *prt* RNA can also function to repress *pho1* via the Mmi1/Clr4 pathway (Shah et al., 2014).

Levels of *pho1* RNA and protein are induced in the absence of phosphate while transcription of *prt* is no longer detected (Shah et al., 2014). With this in mind one might hypothesise that, upon depletion of phosphate, Clr3 recruitment will be lost. Importantly, loss of Clr3 was detected at the region between the promoters upon phosphate depletion (Fig. 3-8A); a pattern very much similar to that described above (Fig. 3-6A and Fig. 3-7A). At the same time, no change in Pol II occupancy was observed at *adh1* caused by the presence or absence of phosphate (Fig. 3-8C). As expected Pol II levels over the *prt* region are low in the absence of phosphate (Fig. 3-8B, primer pairs 1 and 2) but increase over the mRNA region relative to phosphate availability (Fig. 3-8B, primer pairs 3 and 4). In the presence

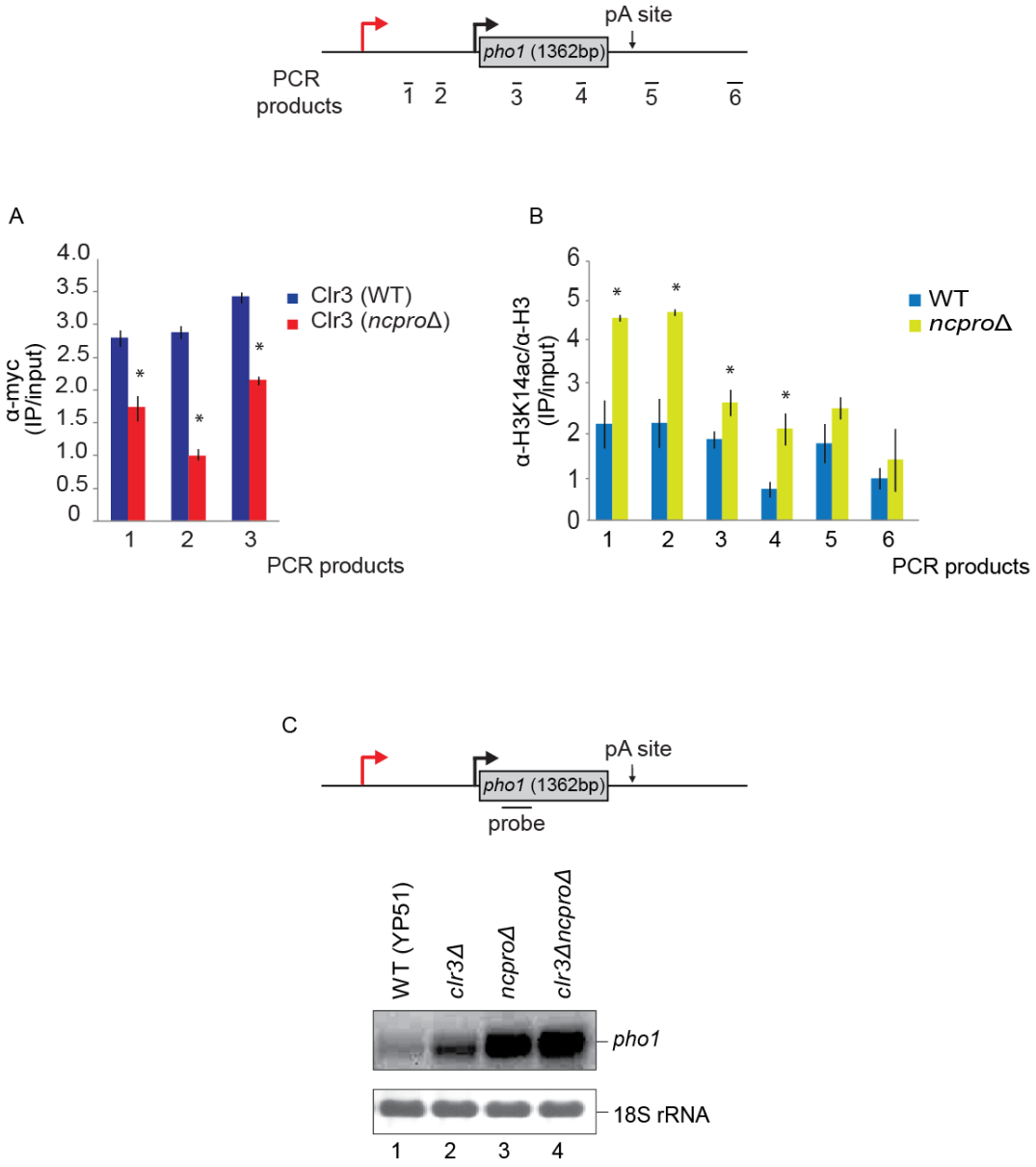


Figure 3-7: *prt* is required for Clr3 recruitment
(A) Clr3 recruitment ChIP performed in cells lacking the *prt* promoter (*ncpro*Δ). (B) H3K14ac ChIP analysis performed in a strain lacking the *prt* promoter (*ncpro*Δ) compared to wild-type (WT) levels. Quantification of results from three independent experiments is shown for all ChIP analysis; error bars indicate the standard error. Statistical differences between WT and *ncpro*Δ strains were analysed using the Student's t-test and evaluated using P-values. An asterisk (*) indicates a significant difference (P<0.05) compared to WT. (C) Northern blot analysis of *pho1* from RNA isolated from the indicated strains. 18S ribosomal RNA represents a loading control. The probe used is indicated by a black bar on the schematic above the blot.

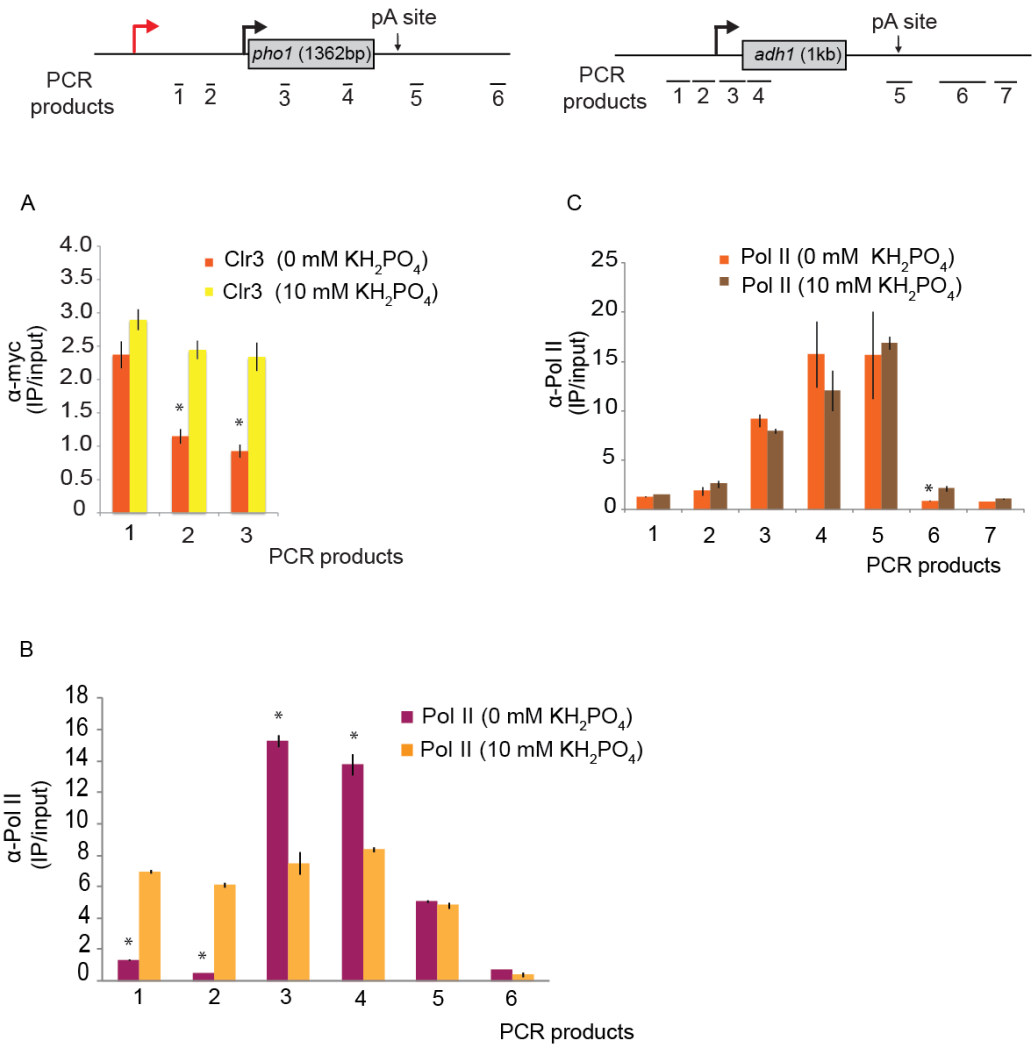


Figure 3-8: Clr3 recruitment is phosphate-dependent
(Top) Schematics of *pho1* and *adh1* genes with PCR products used for qPCR indicated by black bars (A) Clr3 ChIP performed in a strain grown in EMMG in the presence or absence of 10 mM KH_2PO_4 over *pho1*. (B, C) Pol II ChIP performed in a strain grown in EMMG in the presence or absence of 10 mM KH_2PO_4 over *pho1* and *adh1*. Quantification of results from three independent experiments is shown for all ChIP analysis; error bars indicate the standard error. Statistical differences between samples were analysed using the Student's t-test and evaluated using P-values. An asterisk (*) indicates a significant difference ($P<0.05$) between samples grown in the absence and presence of KH_2PO_4.

of phosphate, Pol II levels increase more than five-fold over the *prt* region as the ncRNA is transcribed.

Taken together, these data provide strong evidence that recruitment of both Seb1 and Clr3 to the *pho1* locus is dependent on the production of the non-coding RNA *prt* under conditions of phosphate availability.

4.2.5 *prt* RNA is required for *pho1* silencing

Seb1 recruitment is dependent on *prt* but whether this interaction is direct, like Mmi1 binding to DSRs, is something that still needs to be explored. In order to identify any putative Seb1 binding site, five different 189 bp regions of *prt* were deleted (Fig. 3-9). These deletions span the entire non-coding region excepting the region (-630-(-441)) (ATG=1) which could not be generated. The deleted regions are therefore as given: *prt-1*Δ (-1197- (-1008)), *prt-2*Δ (-1008- (-819)), *prt-3*Δ (-819- (-630)), *prt-5*Δ (-441- (-252)), and *prt-6*Δ (-252- (-63)). To determine whether and which of the deletions affects *pho1* expression, Northern analysis was performed. RNA was extracted from wild-type and each deletion strain and examined (Fig. 3-9). Compared to wild-type levels no change in *pho1* expression could be detected for either *prt-5*Δ (-441- (-252)) or *prt-6*Δ (-252- (-63)) (Fig. 3-9, compare lanes 7 and 6 to 1). Whilst it is tempting to conclude that this is because neither deletion impacts on *prt* and its silencing potential, it should not escape one's attention, especially considering the proximity to the promoter, that the deletions themselves may prevent transcription of *pho1* mRNA.

An increase in *prt* levels was detected in the deletion spanning the region that contains the DSR motifs (*prt-3*Δ (-819- (-630))) (Fig. 3-9, lane 6) - an effect that is reminiscent of a *mmi1*Δ deletion strain (Shah et al., 2014). Preliminary data based on RNA immunoprecipitation (RIP) analysis and electrophoretic mobility shift assays (EMSA) (personal communication, Sina Wittmann) suggest that the Seb1 binding site on *prt* lies upstream of the Mmi1 site. Interestingly, the deletion spanning the very start of the ncRNA, *prt-1*Δ (-1197- (-1008)), resulted in an accumulation of both *pho1* mRNA and *prt* levels (Fig. 3-9, lanes 2 and 3), whereas *prt-2*Δ seemed to have little or no effect on expression (Fig. 3-9, lanes 4 and 5).

Since it is already known that the region deleted in the strain *prt-3*Δ is required for Mmi1 recruitment, the phenotype observed for this strain is not surprising. However, the region deleted in *prt-1*Δ has not previously been shown to be essential for *pho1* regulation. Based on the similarity between *prt-1*Δ, *seb1-1* and *clr3*Δ strains on *pho1* levels, it is possible that this portion contains a motif(s) essential for Seb1 binding.

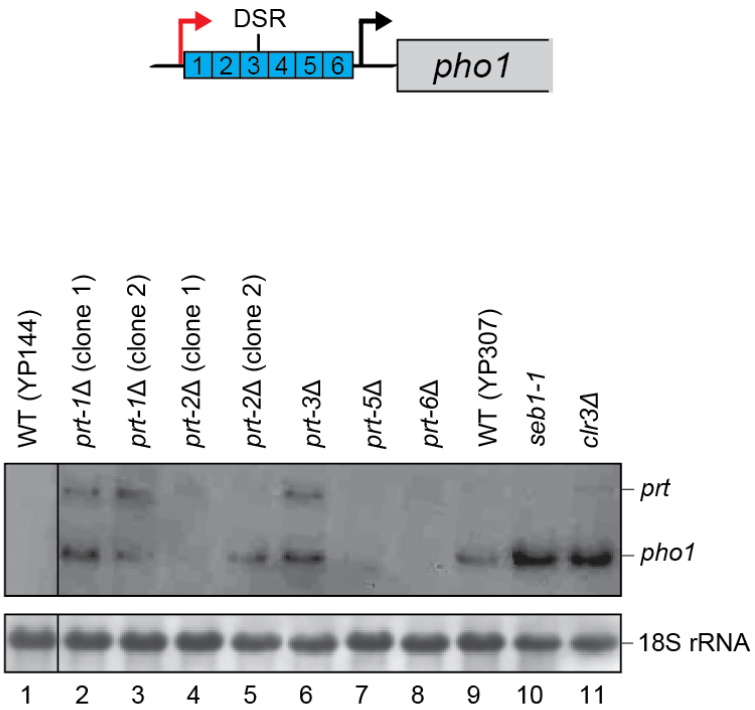


Figure 3-9: *prt* RNA is required for *pho1* silencing
(Top) Schematic of *pho1* locus with numbered deletions of *prt* indicated by blue boxes. Location of DSR motifs is indicated (Bottom) Northern blot analysis of the effect of *prt* deletion strains on *pho1* mRNA and *prt* expression levels. All strains were grown in the presence of 10 mM KH_2PO_4. 18S ribosomal RNA serves as a loading control. *pho1* probe is as used in (Fig. 3-7). N.B. Two clones for strains *prt-1*Δ and *prt-2*Δ were tested.

4.3 Discussion

Seb1 is the *S. pombe* orthologue of the well-characterised budding yeast protein Nrd1. The two proteins have many features in common; for example, like Nrd1, Seb1 is an essential nuclear protein that has both a CID at its N-terminus and, separated by a small region that is rich in arginine-serine (RS) dipeptides, an RRM at its C-terminal end. From this study it is beginning to emerge that, as well as some additional similarities, there are also a number of key differences between the two proteins.

Firstly, unlike Nrd1, it is shown here and elsewhere that Seb1 is able to recruit silencing factors to constitutive heterochromatin (Marina et al., 2013) as well as euchromatic gene loci (see chapter 3). Second, despite attempts to find one, no role for Seb1 in snRNA processing could be identified (data not shown). In this study Seb1, resembling Nrd1, is shown to interact with the Pol II CTD. However, the phospho-CTD specificity of the two differ somewhat. Nrd1 is highly specific to Ser5-(P), whereas Seb1 appears to have an equal preference for the singly phosphorylated Ser2 and Ser5 CTDs. Binding to the Ser2-(P) CTD typically indicates recruitment to gene 3' ends; indeed this is what is seen for Seb1 at several protein-coding genes (Fig. 2-7). Binding to Ser5-(P) on the other hand is usually indicative of recruitment in the early elongation phase of the transcription cycle. It is unclear from these data alone if Seb1 has a real preference for either of these two phospho-marks and whether the bipartite binding profile seen at *pho1* (Fig. 2-7) is a result of both of these interactions, or whether RNA binding alone is responsible for recruitment to the 5’ end.

Of course, more in-depth biochemical and quantitative studies are needed to fully characterise and understand these binding patterns including, for example, binding affinity (K_d) measurements. It is interesting to speculate whether the difference in CTD binding can explain, at least in part, the versatility and diverse roles that are attributed to Nrd1 and Seb1.

From affinity purification and mass spectrometry analysis, Seb1 is found to strongly co-purify with Mmi1 (and vice versa). To establish whether or not it is mediated by RNA, it would be interesting to see if this interaction is affected by RNase A. In either case, an interaction between these two proteins may be important at loci such as *pho1* in the coupling of histone methylation and acetylation. Several proteins with previously reported roles in 3' end processing and termination were also recovered from this purification. Since ChIP analysis (Fig. 2-7) reveals its enrichment at the pA sites of all protein-coding genes tested it may be that Seb1 plays a global role here. This is something that needs to be taken into consideration when dissecting the role of Seb1 at genes where it is also responsible for recruitment of silencing factors.

Unexpectedly, the complete SHREC complex was not detected in the Seb1 purification peptide spectra; only the Clr3 subunit was present. There could be several reasons for this: (i) the large C-terminal TAP tag used for Seb1 purification may interfere with SHREC interactions, (ii) it is unclear how strong the interaction between SHREC and Seb1 is. It is possible that the binding is too weak or that a stronger interacting protein (possibly Mmi1) masks SHREC binding, (iii) the conditions used for purification are suboptimal for SHREC: Seb1 interactions, or (iv) the interaction is transient and/or indirect.

Finally, association of Seb1 with peptides from several proteins involved in splicing was also detected. SCAF8 is the human orthologue of Seb1. This protein is part of the SCAF family of proteins which both interact with the CTD and resemble the SR family of splicing factors. As the only such group of proteins to combine a role for splicing with CTD binding they may represent an essential factor in the coupling of transcription and splicing. SCAF8 has been shown to interact with the CTD when it is phosphorylated at serine residues 2 and 5, suggesting that it associates with the elongating transcription machinery (Patturajan et al., 1998). This more closely resembles the Seb1 CTD specificity than the Nrd1 specificity. Considering also that Seb1 co-purifies with splicing factors it is interesting to speculate a potential role for the *S. pombe* protein that is analogous to SCAF8. Whilst this is not something that has been explored in this study it certainly represents an interesting avenue for future work.

In addition to these CTD and protein interactions, ChIP analysis shows that recruitment of Seb1 to the *pho1* gene is dependent on RNA, phosphate, and the *prt* promoter (Figs. 3-6, 3-7, 3-8). Taken together these data demonstrate that binding of Seb1 requires transcription of the ncRNA *prt* under conditions of phosphate availability. Since Seb1 also interacts with the Pol II CTD, it is likely that it employs both its protein domains (that is, its CID and RRM) in binding at the *pho1* locus. However, more detailed biochemical experiments are required to conclusively show this and determine the exact details of RNA versus CTD binding.

5. Seb1-dependent silencing is a general mechanism of gene repression

5.1 Introduction

An emerging theme stemming from work in various model systems is that lncRNAs have a role in the complex regulation of genes involved in biosynthetic pathways. This is particularly evident in examples from budding yeast, exemplified by, but by no means limited to, the regulation of serine and galactose homeostasis by lncRNAs affecting *SER3* and *GAL1* gene expression. Similarly, it has transpired that multiple lncRNAs from fission yeast can mediate regulation of such genes - examples being *fbp1* and *pho1*, which regulate glucose and phosphate levels respectively.

The regulation of the *S. pombe pho1* gene is mediated by transcription of an upstream lncRNA, *prt*. It appears that there are at least two mechanisms by which *prt* can exert its repressive effects on the production of the mRNA. One previously described mechanism (Shah et al., 2014) is mediated by Mmi1 and involves methylation of H3K9 by Clr4. The other, described in the preceding chapters, appears to be mediated by Seb1 and involves deacetylation of H3K14 by Clr3. Thus the question remains: is this mechanism unique to the *pho1* gene locus or does it function at other (conceivably metabolic) protein-coding genes?

This chapter explores the role of Seb1 and SHREC/Clr3 at two particular gene loci; *zrt1* and *tgp1*. Both these genes are involved in metabolic homeostasis - *zrt1* is

essential for zinc acquisition and *tgp1* (which is homologous to *S. cerevisiae GIT1*) for glycerophosphoinositol uptake. It is demonstrated here that the regulation of *zrt1* expression appears to follow the same principles as those that have been described for *pho1*. Seb1 and Clr3 interact with a nearby non-coding region and function to silence *zrt1* expression under repressive (zinc-replete) conditions.

Conversely, whilst it is clear that Seb1 contributes to *tgp1* silencing, Clr3 appears to have little or no effect on its expression. Other histone deacetylases tested also appear to play no part, indicating that another mechanism of TI operates to repress this gene. Therefore, the data presented in this chapter shed light on the diverse array of mechanisms by which TI is mediated.

5.2 Results

5.2.1 Clr3 represses another protein-coding gene, *zrt1*

Previous work (Sugiyama et al., 2007) has revealed that Clr3 localises to a number of euchromatic sites. However, whether Clr3 functions at these loci to regulate gene expression by modulating histone acetylation levels is not clear. One of these loci (see Fig. 4-1 for schematic) includes the region encoding the annotated ncRNAs *SPNCRNA.425* (on the forward strand) and *SPNCRNA.1622* (on the reverse strand). Interestingly, reminiscent of the *pho1* locus, expression of the downstream gene is known to be dependent on a particular metabolite (Dainty et al., 2008). This gene, *zrt1*, encodes a Zrt/IRT-like protein (ZIP) zinc transporter whose expression is increased in response to zinc deprivation very much similar to *pho1*'s response to phosphate. Because of these similarities it was hypothesised that Seb1 and Clr3 may function at *zrt1* in the same way that they do to repress *pho1*. Therefore, the *zrt1* locus was exploited in order to explore the generality of the mechanism of repression already described for *pho1*.

To examine this hypothesis, *zrt1* levels were analysed in *seb1-1* and *clr3*Δ by Northern blot analysis (Fig. 4-1). These strains revealed elevated levels of *zrt1* mRNA as compared to wild-type (Fig. 4-1, compare lanes 2 and 3 to lane 1). This effect was also observed, though to a milder degree, in *clr4*Δ (Fig. 4-1, lane 4). Moreover, as previously seen with *pho1* (Fig. 2-8), an additive accumulation of mRNA levels was detected in the double mutant *clr3*Δ*clr4*Δ compared to the respective single mutants *clr3*Δ and *clr4*Δ (Fig. 4-1, compare lane 5 to lanes 3 and 4). Thus, in the presence of a

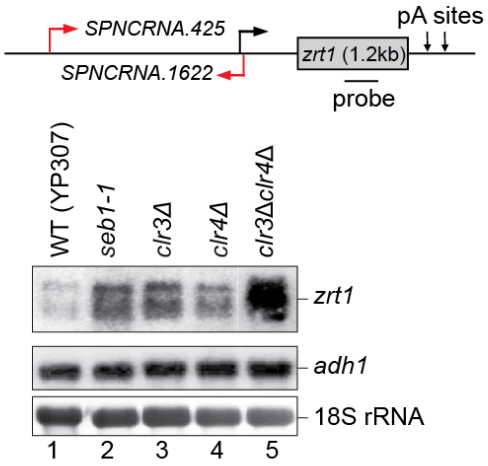


Figure 4-1: Clr3 represses another protein-coding gene, *zrt1*
(Top) Schematic of *zrt1* locus. The position of the probe on *zrt1* is indicated by a black bar. Two clusters of polyadenylation sequence elements are shown. (Bottom) Northern blot analysis of *zrt1* mRNA levels in the indicated strains grown in YES. The housekeeping gene *adh1* and 18S ribosomal RNA serve as loading controls.

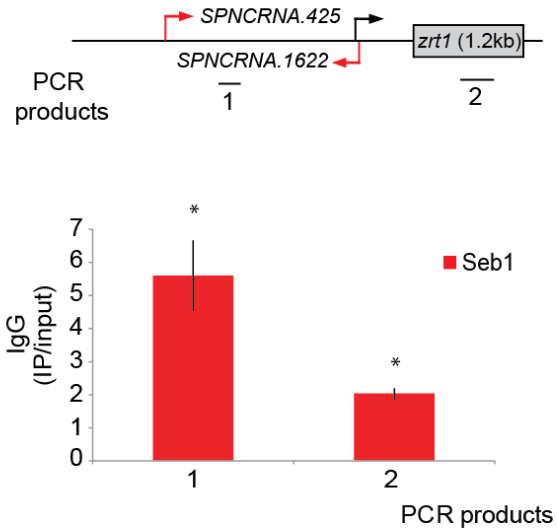


Figure 4-2: Seb1 is recruited to the *zrt1* locus
(Top) Schematic with black bars indicating locations of PCR products at the *zrt1* gene and upstream ncRNAs *SPNCRNA.425/ SPNCRNA.1622*. (Bottom) Seb1 ChIP experiments were performed in an endogenously TAP-tagged Seb1 strain grown in YES. Quantification of results from three independent experiments is shown; error bars indicate the standard error. Statistical differences between samples were analysed using the Student's t-test and evaluated using P-values. An asterisk (*) indicates a significant difference ($P<0.05$) compared to the untagged strain.

zinc source, repression of *zrt1* is dependent on all three proteins Seb1, Clr3, and Clr4.

5.2.2 Seb1 and Clr3 are recruited to the *zrt1* locus

Prompted by this initial finding of a *zrt1* silencing defect as observed by Northern blot (Fig. 4-1), one might suppose that Seb1, in order to initiate this repression, is present over the *zrt1* locus and upstream non-coding region. To test this, Seb1 recruitment was examined over the non-coding RNAs (*SPNCRNA.425* and *SPNCRNA.1622*) and the *zrt1* gene body by ChIP analysis (Fig. 4-2). Encouragingly, Seb1 was found to be enriched at both the upstream non-coding region and also, to a lesser extent, the downstream gene *zrt1*, suggesting that Seb1 is likely to act via the non-coding transcript(s) to repress *zrt1*.

Concordant with Seb1's association, Clr3 (conceivably via recruitment of the former) was found to be enriched over the non-coding region and the *zrt1* gene body (Fig. 4-3). This is consistent with and confirms genome-wide data showing Clr3 enrichment over the upstream region encoding *SPNCRNA.425*/*SPNCRNA.1622* (Sugiyama et al., 2007). Although the level of their recruitment is not directly comparable, based on ChIP signals it appears that the stoichiometry of Seb1 and Clr3 differs at the *zrt1* locus from that at *pho1*. In the case of *pho1*, ChIP analysis (Fig. 2-3 and 2-7) is indicative of Clr3 being substoichiometric to Seb1, whereas the reverse is observed here.

5.2.3 Recruitment of Clr3 to *zrt1* is RNA-dependent

Recruitment of both Seb1 and Clr3 to *pho1* is dependent on the production of its

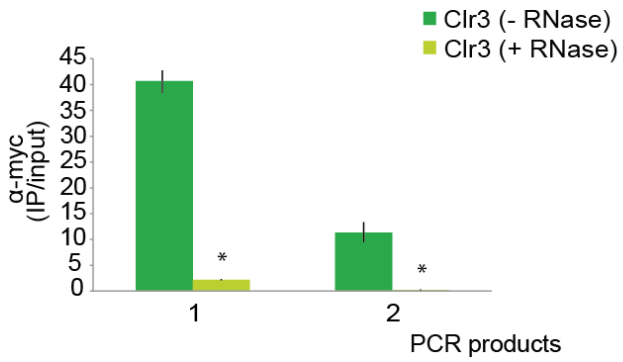


Figure 4-3: Recruitment of Clr3 to *zrt1* is RNA-dependent
ChIP analysis of myc-tagged Clr3 over the *zrt1* locus either untreated (-) or treated (+) with RNase. PCR products are the same as in (Fig. 4-2). Quantification of results from three independent experiments is shown for all ChIP analysis; error bars indicate the standard error. Statistical differences between samples were analysed using the Student's t-test and evaluated using P-values. An asterisk (*) indicates a significant difference ($P<0.05$) compared to untreated sample.

upstream, overlapping RNA, *prt*. To establish whether this is also the case at *zrt1* which, although flanked by ncRNAs, is not overlapped, RNase ChIP was performed (Fig. 4-3). Clr3 binding to the convergent, overlapping ncRNAs *SPNCRNA.425*/*SPNCRNA.1622* and the *zrt1* ORF was entirely lost when cells were treated with RNase A and T1. To rule out the trivial possibility that treatment with RNases impairs protein cross-linking to DNA, Pol II and Seb1 ChIP was performed at the *adh1* locus (Fig. 3-6) where no change in levels between treated and untreated samples were observed.

Taken in all, these data provide evidence that Seb1, likely via interaction with either or both ncRNAs *SPNCRNA.425* and *SPNCRNA.1622*, recruits Clr3 to elicit repression of the downstream gene *zrt1* under conditions of zinc availability.

5.2.4 Seb1 elicits *tgp1* silencing independently of SHREC

Recent work (Ard et al., 2014) has revealed that transcription of the lncRNA termed *nc-tgp1* represses expression of the adjacent phosphate-responsive permease gene *transporter for glycerophosphodiester 1* (*tgp1*) (see Fig. 4-4B for gene schematic). Like *pho1* this gene is involved in the *S. pombe* phosphate response, therefore the locus was utilised in order to further explore the generality of the pathway defined in the preceding chapters. In view of the fact that repression of *tgp1* has been attributed to an, as yet, undefined mechanism of TI, the question of whether the Seb1/Clr3 silencing pathway has any implication here was also investigated.

Remarkably, Northern analysis reveals that *tgp1* accumulates in the *seb1-1* mutant compared to wild-type (Fig. 4-4B, compare lanes 1 and 2). Similarly, placing the *seb1* gene under the control of the thiamine regulated *nmt* promoter, where transcription is rapidly repressed upon switch from medium lacking to medium containing thiamine, coincident with loss of Seb1 (Fig. 4-4A), *tgp1* mRNA levels start to accrue (Fig. 4-4B, lane 13). An even more pronounced derepression of *tgp1* mRNA was detected for cells which were grown in the absence of phosphate (Fig. 4-4B, compare lanes 10 and 11 to lane 9), suggesting that a Seb1-independent silencing pathway operates to repress *tgp1*. Interestingly, since no further accumulation is detected after 24 hours, complete derepression of *tgp1* appears to be reached at or before 12 hours growth in the absence of phosphate (0 mM KH_2PO_4) (Fig. 4-4B, compare lanes 10 and 11).

Somewhat unexpectedly, considering the case at *pho1* and *zrt1,* deletion of *clr3* (Fig. 4-4, lane 3) has no discernible induction effect on *tgp1*. Similarly, *clr4*Δ (Fig. 4-4, lane 4), or indeed the double mutant *clr3*Δ*clr4*Δ (Fig. 4-4, lane 5), revealed no change in mRNA expression levels. These results signify that the requirement for Seb1 to repress *tgp1* expression is not reliant on Clr3's HDAC activity (as it is at *pho1* and *zrt1*), nor Clr4's methyltransferase activity (as it has been reported to be at pericentromeric heterochromatin (Marina et al., 2013)).

To query the possibility that Seb1 recruits another HDAC to elicit silencing, the effect of *sir2*Δ and two strains harbouring a *clr6* mutation was analysed. As before, no detectable increase in either *tgp1* or *nc-tgp1* was observed (Fig. 4-4, lanes 6-8), suggesting that these HDACs do not play a major part in *tgp1* silencing.

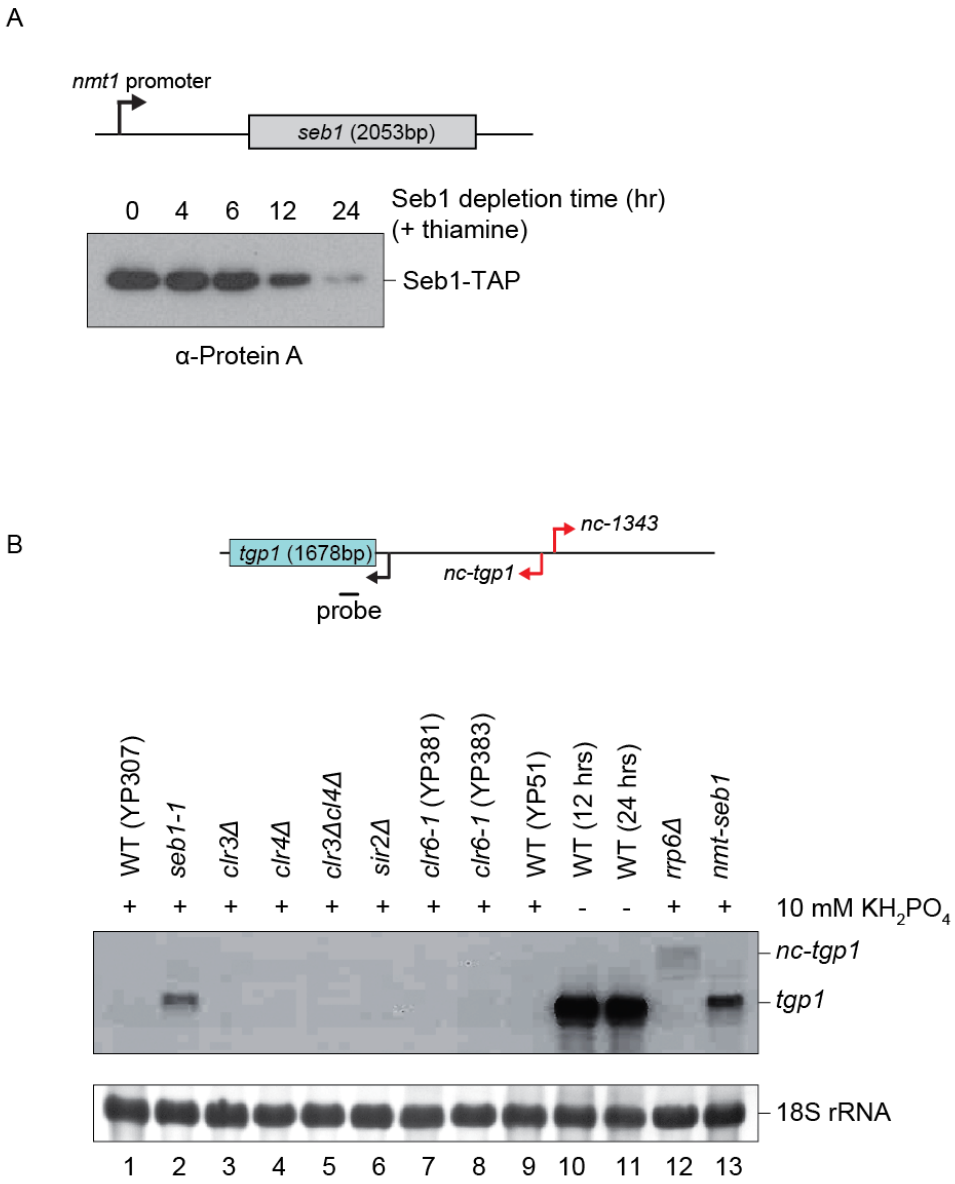


Figure 4-4: Seb1 elicits *tgp1* silencing independently of SHREC

(A) Western blot analysis of Seb1-TAP depletion in thiamine (B) (Top) Schematic of *tgp1* locus with probe used for Northern blot indicated by a black bar. (Bottom) Northern blot analysis of *tgp1* levels from RNA isolated from strains indicated in the presence (+) or absence (-) of 10 mM KH_2PO_4. *nmt-seb1* was grown in the presence of thiamine for 24 hours before RNA extraction. 18S ribosomal RNA serves as a loading control.

5.2.5 Seb1 elicits *tgp1* silencing independently of HDACs

None of the HDACs Clr3, Clr6 or Sir2 appears to have any effect on expression of *tgp1* (Fig. 4-4B). This does not, however, rule out the possibility that the HDACs may function in concert to elicit silencing or perhaps a separate, as yet unidentified, HDAC is involved. In order, then, to completely rule out any HDAC-dependent activity, cells were treated with the inhibitor TSA (Fig. 4-5). Northern blot confirmed previous data from Ard et al., (2014) that show *tgp1* accumulates in cells grown in the absence of phosphate (Fig. 4-5, lane 1), but is not detected in its presence (10 mM KH_2PO_4) (Fig. 4-5, lane 2). Upon treatment with TSA (Fig. 4-5, lane 3) no accumulation of *tgp1* was observed. This is in contrast to *pho1*, where it was seen that TSA treatment results in a patent accumulation of mRNA levels (Fig. 3-1). Thus, it appears that HDACs, or at least those sensitive to the potent HDAC inhibitor TSA, are not implicated in the repression of *tgp1* expression.

5.2.6 Seb1, but not Clr3, is recruited to *tgp1*

As described, levels of *tgp1* mRNA were found to accumulate in *seb1-1* and *nmt::seb1* strains (Fig. 4-5, lanes 2 and 13). One might expect, then, that Seb1 might be present at the *tgp1* locus to elicit this silencing effect. Hence Seb1's recruitment to the *tgp1* gene locus was examined by ChIP analysis. As can be seen, Seb1 is enriched over *nc-tgp1* (Fig. 4-6, probes 2-5) and drops off either side of the non-coding transcript (Fig. 4-6, probes 1 and 6). Consistent with a *clr3* deletion strain having no effect on *tgp1* mRNA levels (Fig. 4-4, lane 3), recruitment of the protein, or indeed any other SHREC subunit, could not be detected at the locus by ChIP analysis (data

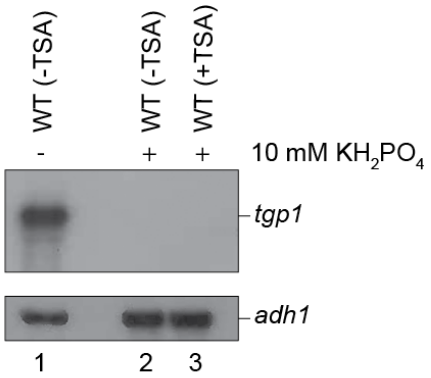


Figure 4-5: Seb1 elicits *tgp1* silencing independently of HDACs

Northern blot analysis of RNA isolated from a wild-type (WT) strain, YP51, treated with (+TSA) or without (-TSA) 20 µg/mL of the HDAC inhibitor trichostatin A (TSA) in the presence (+) or absence (-) of 10 mM KH_2PO_4. *tgp1* mRNA levels are shown. The housekeeping gene *adh1* serves as loading control.The *tgp1* probe used is as in Fig. 4-4.

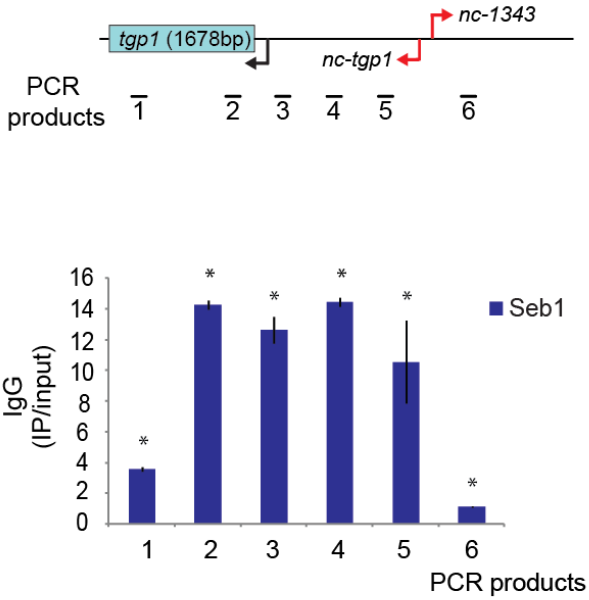


Figure 4-6: Seb1, but not Clr3, is recruited to *tgp1*
(Top) Schematic of *tgp1* locus with PCR products used for qPCR indicated. (Bottom) Seb1 ChIP experiments were performed in an endogenously TAP-tagged Seb1 strain grown in YES. Quantification of results from three independent experiments is shown; error bars indicate the standard error. Statistical differences between samples were analysed using the Student's t-test and evaluated using P-values. An asterisk (*) indicates a significant difference ($P<0.05$) compared to the untagged strain.

not shown).

To test the hypothesis that Seb1 enrichment at probes spanning *nc-tgp1* (Fig. 4-6, probes 2-5) reflects binding of Seb1 to the RNA, RNase ChIP was performed. Seb1 binding was completely lost when cells were treated with RNase A and T1 (Fig. 4-7). This result argues that Seb1 is recruited by *nc-tgp1* to prevent induction of *tgp1*. How the protein effectuates this repression remains to be determined though.

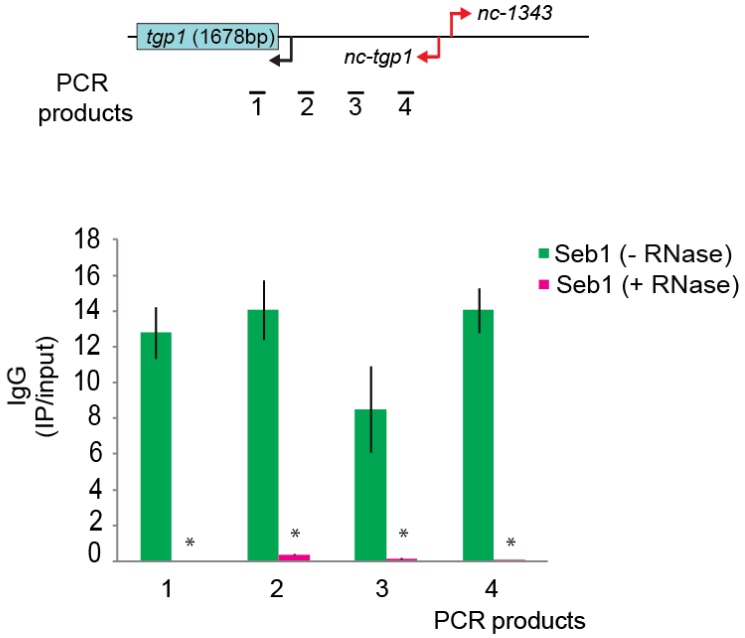


Figure 4-7: Seb1 recruitment to *tgp1* is RNA-dependent
(Top) Schematic of *tgp1* locus with PCR products used for qPCR indicated. (Bottom) ChIP analysis of TAP-tagged Seb1 over the *tgp1* locus either untreated (-) or treated (+) with RNase. Quantification of results from three independent experiments is shown for all ChIP analysis; error bars indicate the standard error. Statistical differences between samples were analysed using the Student's t-test and evaluated using P-values. An asterisk (*) indicates a significant difference ($P<0.05$) compared to the untreated sample.

5.3 Discussion

Genes encoding proteins involved in biosynthetic pathways need to respond rapidly to fluctuations in external and/or internal metabolite concentrations. As such, the regulation of their expression is closely monitored and it is becoming well established that lncRNAs are central to orchestrating and controlling genetic regulatory outputs. Facultative heterochromatin, like that found at *pho1*, is an example of where a lncRNA - in this case *prt* - can serve as a platform for the recruitment of effector complexes that are able to modify chromatin architecture in such a way as to initiate an appropriate and efficient response to environmental cues.

The number of annotated lncRNAs is increasing at a fast pace. So too is their acceptance as key gene regulators. Despite this, the scope of lncRNA involvement in regulatory circuits and the actual mechanisms by which they operate remain poorly understood. Here, the effects of two ncRNA-encoding regions on their adjacent protein-coding genes have been examined. In the first case, and expanding on earlier findings implicating *prt* ncRNA in recruitment of Seb1 (Fig. 3-6), expression of the zinc-dependent *zrt1* gene is shown to be dependent on both Seb1 and Clr3 (Fig. 4-1), whose recruitment relies on the upstream non-coding transcription of *SPNCRNA.425* and/or *SPNCRNA.1622* (Figs. 4-2 and 4-3). The salient finding here is that the mechanism of repression described at *pho1*, whereby ncRNA, through recruitment of the Pol II and RNA-interacting protein Seb1, silences the downstream gene, is not unique to that locus, and instead operates to repress other protein-coding genes outside of the phosphate response.

As can be seen (Fig. 4-1) two *zrt1* transcripts are detected with the probe used for Northern blot analysis. Based on band size and *zrt1*'s polyadenylation sequence elements it is possible that this represents alternative polyadenylation, however this is not something that has been explored in great detail. Since the expression levels of *zrt1* have been shown to respond to zinc concentrations (Dainty et al., 2008) it would be interesting to gauge the effect of zinc-replete versus zinc starvation on Seb1 and Clr3 recruitment. Presumably, as with *pho1* (Fig. 3-8), their association would be lost upon depletion of $ZnSO_4$ from the media. Additionally, to fully elucidate the sequence of events that leads to *zrt1* silencing it would be important to conclusively show that Seb1 (i) binds specifically to the ncRNA(s) and (ii) is essential for the subsequent recruitment of Clr3 to bring about histone deacetylation.

Concordant with the case at *pho1*, Clr3 and Clr4 appear to operate independently of one another to repress *zrt1* transcription (as assessed by Northern blot - Fig. 4-1). As noted previously, this data is in disagreement with those of Marina et al., (2013) where Clr3, as part of SHREC, is reported to function to promote H3K9 methylation by Clr4. Simply put, the data presented here for repression of *pho1* and *zrt1* contrast with that described at pericentromeric heterochromatin in that Clr3 and Clr4 appear to operate via separate pathways. Thus, in these two cases H3K14 deacetylation is, in and of itself, sufficient to bring about transcriptional repression without the need for Clr4-mediated H3K9 methylation.

To extend the role of Seb1-dependent transcriptional repression even further, another gene was considered: *tgp1*. Like *pho1* this gene encodes a protein involved in the phosphate response. *tgp1* is partly overlapped by *nc-tgp1* and the data presented

herein suggest that Seb1, via binding to the ncRNA, represses *tgp1* expression (Figs. 4-4, 4-6, and 4-7). However, Seb1 does not appear to be recruiting Clr3 to this locus since it could neither be detected by ChIP, nor did a *clr3*Δ strain have any discernible effect on mRNA levels. Indeed, it seems likely that HDACs are not involved in repression at all (Fig. 4-5).

These data add to the recent study of Ard et al., (2014) regarding the mechanism(s) of TI that operates to repress *tgp1* expression. Although the mechanism(s) has not been identified here, it is clear that Seb1, likely via a direct interaction with *nc-tgp1*, is part of the silencing machinery that is set in place to prevent *tgp1* induction under phosphate-replete conditions.

To date, very few eukaryotic genes have been shown to be regulated by a ncRNA transcription-driven mechanism of TI. However, considering how highly compact the yeast genome is, and taken with the observation that protein-coding genes have overlapping transcription from intergenic or intragenic sense and antisense promoters, there appears to be great potential for transcriptional regulation by TI. As an example, over 6,736 lncRNA genes have been identified in human and, of these, 63% overlap with or are in a close proximity (<10 kb) of known protein-coding genes (Jia et al., 2010), suggesting that TI and the mechanisms by which it operates are potentially widespread throughout biology.

6. General Discussion

6.1 A new mechanism of transcriptional interference

6.1.1 Clr3 and Seb1 contribute to transcriptional interference at *pho1*

The eukaryotic genome is pervasively transcribed, resulting in a complex network of ncRNAs. Frequently, non-coding transcription of intergenic regions overlaps with protein-coding genes and therefore it has great regulatory potential within the cell. One way in which this regulatory role can be achieved is by transcriptional interference (TI), which describes the suppressive influence of one transcriptional process on another. Although TI has been known for some time, and despite reports of the phenomenon spanning various different organisms, very little is known about the details of the causal mechanisms. The work presented in this thesis sheds some light on this by identifying and characterising a new mechanism by which TI can operate.

The few examples of models of TI which have previously been proposed are well outlined in reports by Shearwin et al., (2005) and Palmer et al., (2011). The arrangement of promoters - be they convergent, divergent, or tandem - defines which mechanisms TI can operate by. For the case of tandem promoters (for example the *pho1*/*prt* locus), the possible paths to repression include promoter occlusion, where the passage of a polymerase over a promoter (or transcription factor (TF) binding site) blocks recruitment of another polymerase or TF to that site. Another involves dislodgement of TFs; TI can occur at promoters which depend on the binding of certain TFs when a passing polymerase effectively knocks them off the DNA.

Occlusion and dislodgment are represented in the schematic (Fig. 5-1 (i)), and, as can be seen, the central idea here is that the polymerase itself is responsible for the TI effect.

Based on the data presented in this thesis a new model for TI is now proposed. In contrast to the models which have previously been described (which involve Pol II negatively affecting transcription, as in (Fig 5-1 (i))), key to this new mechanism are factors which are able to *associate with* the transcribing Pol II and/or a nascent RNA transcript (Fig. 5-1 (ii) and (iii)). Here, a repressor can either bind to an elongating Pol II (Fig. 5-1 (ii)) transcribing from an upstream promoter or with the nascent ncRNA it produces (Fig 5-1 (iii)). Or, conceivably, association of the repressor with both Pol II and RNA could occur simultaneously. This bound repressor is then able to block access of a TF/ Pol II to the second (downstream) promoter by modifying the chromatin environment such that it takes on a repressive state.

This new mechanism is primarily described in the context of *pho1*, and involves Seb1 bound to the overlapping ncRNA (*prt*) via its RRM, as well as to transcribing Pol II via its CTD. Seb1 recruits the HDAC Clr3 to deacetylate lysine 14 of histone H3 and thus lead to repression of the gene (*pho1*) (Fig. 5-2). Furthermore, it is likely that the identity of the ncRNA, in this case and others, facilitates the specificity of Seb1 to particular genomic loci. A novel model is therefore posited whereby the Seb1-dependent and Clr3-mediated mechanism of repression described herein contributes to the phenomenon of TI.

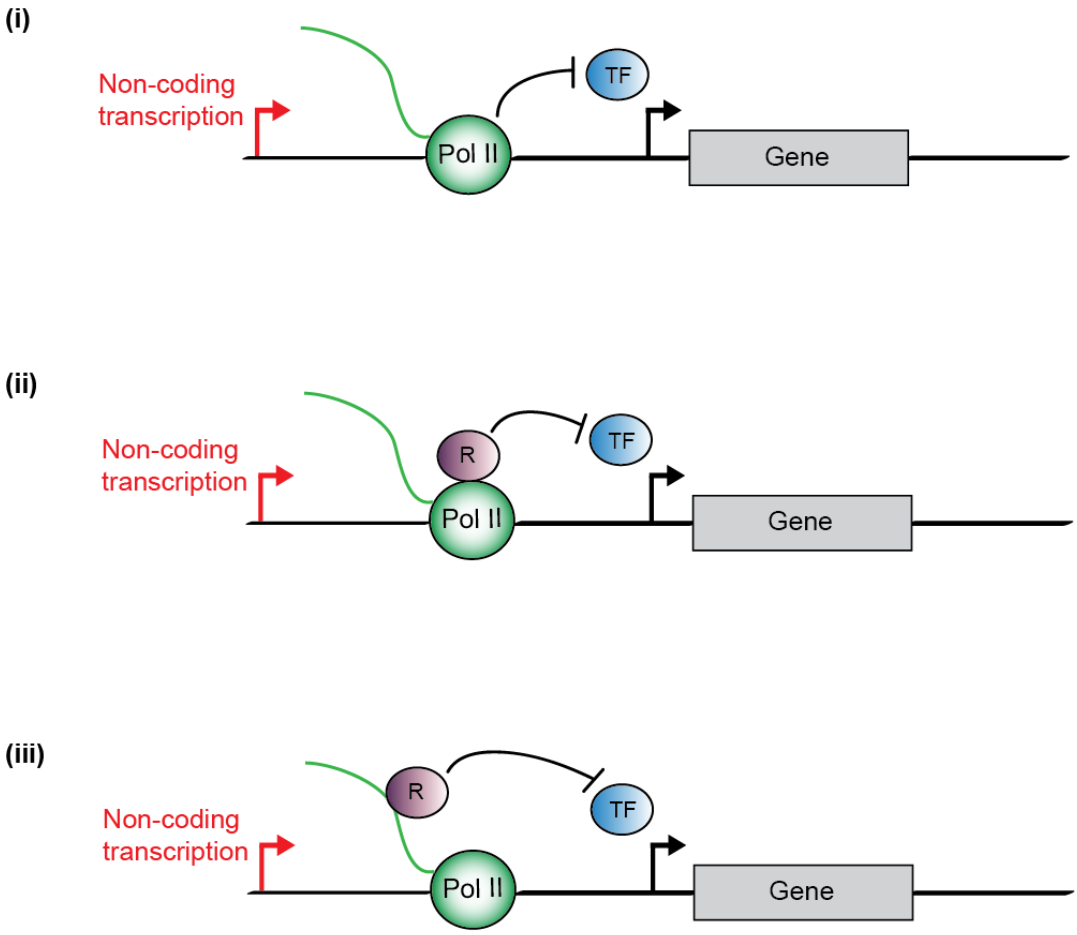


Fig. 5-1: Proposed models for transcriptional interference
Possible models for how transcriptional interference can occur include (i) displacement or blockage of a transcription factor (TF) or second Pol II (together represented by a blue circle) by a Pol II transcribing from an external promoter; (ii) a Pol II-associated repressor (R) blocks TF/ Pol II binding or (iii) the repressor associates with nascent ncRNA and precludes TF/ Pol II association with a second promoter. It is noted that there is a potential for a combination of all these mechanisms.

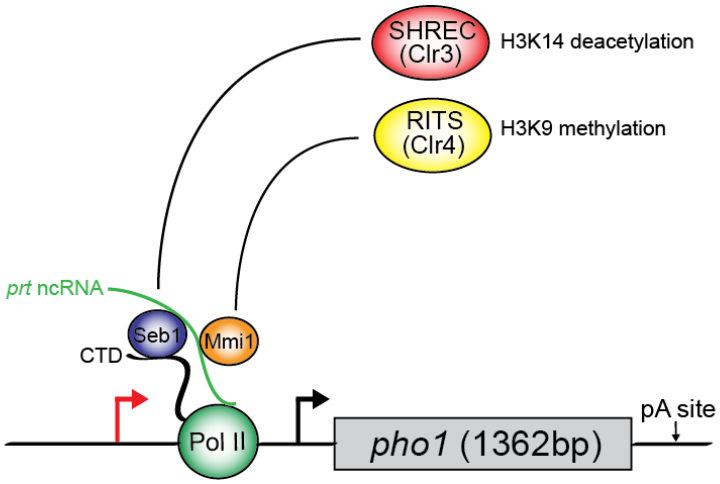


Figure 5-2: *prt* ncRNA mediates *pho1* repression via Seb1 and Clr3
In the presence of phosphate *prt* is transcribed from an upstream promoter and leads to repression of *pho1*. As previously described (Shah et al., 2014), Mmi1 binds *prt* and establishes repressive H3K9me2 through recruitment of the Clr4 methyltransferase. In addition to this, and in the scenario described herein, Seb1 also binds *prt* at a region likely to be upstream of the DSR motifs. Here it recruits the HDAC Clr3 (as part of SHREC) to bring about deacetylation of H3K14 and consequently silence the mRNA in a methylation-independent manner.

6.1.2 Seb1 interacts with RNA and the CTD

Based on the observation that Seb1 contributes to transcriptional repression at the *pho1*, *zrt1*, and *tgp1* loci, one must ask how it is able to mediate this effect. In particular, how exactly does it select its target genes? On the basis of its domain organisation (consisting of CID and RRM - Fig. 1-5) it could be presumed that Seb1 can interact with the CTD and RNA. Indeed, the experimental evidence now attests to this being the case.

6.1.2.1 RNA interactions

While additional work is required to conclusively define the RNA binding motif of Seb1, several lines of evidence presented in the preceding chapters point to an interaction with the ncRNA *prt*. Firstly, ChIP analysis examining Seb1 (and Clr3) recruitment across *pho1*/*prt* was performed in the presence of RNase T1 (which specifically degrades single-stranded RNA) and RNase A (which degrades both single- and double-stranded RNA). Under these conditions the association of Seb1 (and consequently Clr3) at this locus was completely abolished. Secondly, deletion of the *prt* promoter leads to loss of Clr3 recruitment and H3K14 deacetylation, and, in the same way, phosphate-starved cells (in which *prt* is not transcribed) results in a reduction in binding of the HDAC. Consistent with these data pointing to an association between Seb1 and RNA, cross-linking and RNA immunoprecipitation (RIP) analysis has also been used to show that Seb1 interacts with *dg* and *dh* ncRNAs, as well as the snoRNA snR30 (Marina et al., 2013). Taken together, all of these data suggest that Seb1 interacts with RNA but further work is required to investigate

whether this interaction is direct (and likely mediated by the RRM), and where the specificity of binding comes from.

Based on deletions of regions of *prt* it is tempting to speculate that Seb1 binding occurs upstream of Mmi1 binding - specifically somewhere in the region deleted in the strain *prt-1*Δ (-1197- (-1008)). Clearly then it would be informative to examine whether, in the *prt-1*Δ strain, Seb1 recruitment is lost. This could easily be verified by ChIP-qPCR analysis. To complement this, preliminary RIP data and EMSAs using the Seb1 RRM domain (personal communication, Sina Wittmann) suggest that the region deleted in strain *prt-1*Δ does indeed contain the Seb1 binding site. Assuming these data are correct and that Seb1 recruitment is indeed compromised in the *prt-1*Δ strain, a major aim for the future would be to identify if there is a precise RNA sequence for Seb1 binding and how this compares to the known Nrd1 binding motif (see section 6.3). Of note in this regard is the sequence identity of the Nrd1 and Seb1 RRM motifs which is 41.97%, as compared 29.21% for the whole protein.

6.1.2.2 CTD interactions

Structures from several CID-containing proteins bound to peptides have revealed that CIDs interact with the consensus sequence via a highly conserved binding pocket. In spite of this, CID proteins have very specific binding preferences for the different CTD phosphoisoforms. How such specificity is achieved is something of a mystery. In fact, the Pcf11 CID does not even make direct contacts with the phospho mark it is specific for (Ser2-(P) CTD) (Meinhart and Cramer, 2004). It may be that a CID's specificity can be achieved in multiple ways; through direct CTD contacts,

cooperative interactions (Lunde et al., 2010), enhanced CTD affinity through protein-protein and protein-RNA interactions, or a combination of any or all these factors. In a related manner a recent report (Baejen et al., 2014) investigating how RNA-binding proteins achieve such high target specificity supports this concept. Determining the Seb1 CID specificity and how it achieves this is, therefore, a difficult task.

The canonical CID fold consists of an eight helix bundle where helices 2, 4 and 7 form a groove to bind the CTD. The Seb1 CID has 37.93% sequence identity to the Nrd1 CID and, since with other CIDs structure appears to be more conserved than sequence, it is likely to have a similar fold. Based on sequence alignments to the *S. cerevisiae* Nrd1, Rtt103, Pcf11, and human SCAF8 CIDs, of which the structure has been determined, Seb1 has a conserved serine at position 22, known to be important for Ser5-(P) recognition, together with a basic lysine residue at position 25 (where the corresponding Nrd1 amino acid is arginine) (see Fig. 1-6). In terms of Ser2-(P) recognition, Seb1 has, like the other Ser2-(P) interactors Rtt103 and SCAF8, a basic amino acid (lysine) at position 124 (personal communication, Sina Wittmann). Hence, these observations are in good agreement with the results presented in this thesis (Fig. 3-4 and 3-5) where Seb1 is seen to interact with both Ser2-(P) and Ser5(P) CTD.

In further support of the Seb1 CTD specificity observed in chapter 4, binding of the Seb1 CID domain to CTD peptides has been partially characterised by fluorescence anisotropy (FA) experiments (personal communication, Sina Wittmann). Peptides of two CTD consensus repeats were labelled with FAM. These measurements reveal that the Seb1 CID binding affinity to Ser2-(P) peptide (K_d of 43.8 ± 1.4 µM) is

considerably higher than to Ser5-(P) peptide (K_d of 125.8 ± 27.6 µM). This may explain why ChIP analysis at genes *adh1* (Fig. 2-7), *rps2202* and *dbp2* (data not shown) reveals only one peak of Seb1 binding - at the 3' ends - where Ser2-(P) CTD levels are highest. Whereas, at *pho1* (Fig. 2-7), it may well be that the second peak of Seb1 enrichment at the gene's 5' end represents Seb1 association with *prt* ncRNA rather than to Ser5-(P) CTD. That said, as with Nrd1, it is unclear whether Seb1 is able to interact with both RNA and the Pol II CTD simultaneously. Since the affinity to Ser5-(P) is much weaker than Ser2-(P) it is possible that there is some cooperativity of Seb1 binding to CTD and RNA that underlies the protein's recruitment to the non-coding region of *pho1*.

From these binding data, Seb1 most likely mediates TI by interactions with ncRNA, whereas recruitment of the protein to gene pA sites probably correlates with Ser2-(P) CTD binding and a role in transcription termination.

6.1.3 Transcriptional repression through non-coding transcription/RNA

It has previously been proposed that Seb1 contributes to pericentromeric heterochromatin formation by recruiting SHREC to promote H3K9me (Marina et al., 2013). The mechanistic details of how SHREC achieves this are unclear but it has been suggested that either H3 deacetylation by Clr3 is necessary for Clr4-mediated methylation to occur, or that Clr3 is able to deacetylate a non-histone substrate which, in turn, functions to promote H3K9me. In light of the finding that Clr3 and Clr4 function independently of one another (Figs. 2-7 and 4-1) at the *pho1* and *zrt1* loci, it is unlikely that removal of acetyl marks by Clr3 contributes to

heterochromatic silencing by merely 'freeing up' space for Clr4's histone methyltransferase activity. Hence, H3K14 deacetylation is sufficient *per se* to bring about repression of these genes without dependence on subsequent H3K9 methylation.

Somewhat contradictory to this, however, is a previous report (Motamedi et al., 2008) showing that SHREC is recruited to regions of constitutive heterochromatin by the HP1 protein Chp2. Association of Chp2 with SHREC (SHREC2) couples H3K14 deacetylation to H3K9me binding via Chp2 and limits Pol II access to heterochromatin. In agreement with this finding, compared to the other HP1 protein Swi6, it was found that deletion of *chp2* leads to an increase in accumulation of *pho1* mRNA (Fig. 2-2). It is unclear whether this is due to an inability of SHREC2 to form but, since Clr4-dependent H3K9me creates the binding site for Chp2, this would imply at least some interdependence between Clr4 methylation and Clr3 deacetylation.

Clr3 and Clr4 each contribute significantly to *pho1* repression. However, it must be noted that combining the loss of both these proteins in the *clr3*Δ*clr4*Δ double mutant does not result in so pronounced a silencing defect as is detected in a strain deleted of the *prt* promoter. One might assume then that this means *prt* can mediate repression by another, Cl3/Clr4-independent mechanism. However, caution must be exercised when interpreting these data because whilst Clr4 can be seen as the final effector in the Mmi1-mediated pathway for repression, it is likely SHREC, rather than Clr3 in isolation, which is the endpoint of Seb1-mediated repression. The salient point here is that deletion of *clr3* in the double mutant *clr3*Δ*clr4*Δ does not necessarily remove the

other enzymatic activities of SHREC - most notably nucleosome remodelling by the Mit1 subunit. Since single deletions of all SHREC components - even Clr1 and Clr2 whose specific function remains unknown - has some effect on *pho1* repression, and each subunit has been reported to coimmunoprecipitate or purify with Seb1 (Marina et al., 2013; Motamedi et al. 2008), the activity of Clr3 cannot be considered the sole determinant of *pho1* repression.

Hence, it would be important to generate a complete SHREC deletion strain and combine this with *clr4*Δ. This would allow for a true comparison with the strain lacking the *prt* promoter. At present though it can by no means be excluded that the *prt* nascent transcript, via *cis*-acting elements, is able to recruit other chromatin modifiers or repressors of *pho1* expression.

6.1.4 ncRNA-mediated TI functions at other loci

Like the PHO pathway for phosphate, cells possess a zinc uptake system. Zinc is a structural component of many proteins and, as such, a host of essential cellular processes are dependent on its efficient acquisition. Indeed, at least 2.5% of all *S. pombe* genes are regulated in response to zinc deprivation (Dainty et al., 2008). The gene *zrt1* encodes a high-affinity transporter which is located in the plasma membrane and whose role is to scavenge extracellular zinc. In response to zinc starvation, *zrt1* mRNA levels are significantly increased whereas the gene is expressed at very low levels in the presence of zinc.

The data shown in chapter 5 indicate that expression of *zrt1* is also repressed by Seb1-mediated Clr3 deacetylation. Thus this mechanism of repression may be utilised quite generally by genes whose expression needs to be closely monitored and varied according to the supply and demand of specific metabolites. Although the mechanisms may differ, this is consistent with the regulation of biosynthetic pathways in other organisms - including galactose, serine and phosphate-response pathways in budding yeast - which also depend on non-coding transcription.

It is interesting to speculate that the low-level expression of *zrt1* found under repressive conditions is mediated by Clr3 since the active site of this enzyme is itself dependent on a Zn^{2+} ion (Finnin et al., 1999). Therefore, it is possible that in the presence of a zinc supply Clr3 is able to coordinate a Zn^{2+} ion and thus functions to repress *zrt1* expression. Conversely, in response to zinc deficiency it may be that Clr3 is unable to function (at least at its full activity) in deacetylating and silencing *zrt1* expression. This is something that could be determined experimentally by comparing Clr3 recruitment and activity (i.e. the extent of H3K14 deacetylation) between zinc-replete and zinc-starved states.

6.1.5 Are more players of ncRNA-mediated TI yet to be identified?

Whilst investigating the potential of the Seb1-dependent, Clr3-induced mechanism of transcriptional repression identified at *pho1* (Fig. 5-2), the phosphate-responsive *tgp1* gene was considered. It has previously been suggested (Ard et al. 2014) that, since *pho1* and *tgp1* mRNA levels remain unchanged in RNAi-defective cells these two genes must instead be repressed by TI. This may represent an observable fact but it

does not address the means by which the effect of TI is realised.

It is shown here that, like *pho1*, regulation of *tgp1* mRNA levels is dependent on Seb1 (Fig. 4-4). However, unlike the former, expression of the latter is not reliant on Clr3. Furthermore, no changes in *tgp1* transcript levels could be detected when cells were treated with the potent HDAC inhibitor TSA, seemingly ruling out histone deacetylation as a regulator of *tgp1* expression. A major unresolved question then is how Seb1 exerts its silencing effect on *tgp1*. It is possible that Seb1 is able to mediate recruitment of another player to induce silencing, though methylation by Clr4 has also been ruled out. Conversely, it is also feasible that Seb1 binding to *nc-tgp1* and/or Pol II prevents the recruitment of an activator of *tgp1* transcription. For example, the Pho7 transcription factor only stably engages with the *tgp1* promoter in phosphate-depleted cells when *nc-tgp1* transcription is reduced (Ard et al., 2014) and so perhaps, with the help of Seb1, lncRNA transcription occludes Pho7 binding.

In terms of identifying other players, a Seb1 purification (Fig. 3-1) yielded several proteins with known roles in chromatin modification. These include the SWI/SNF and RSC complex subunits Ssr1 and Ssr2, as well as the nucleosome assembly protein Nap1. The co-purification of these, and other modifiers, with Seb1 gives rise to the possibility that the repression effect seen at *tgp1*, and mediated by Seb1, is in some way related to recruitment of these proteins. This, however, is something that warrants further investigation.

Unfortunately, and as noted in a recent report (McHugh et al., 2015), a major challenge and limitation in the understanding of how lncRNAs regulate transcriptional

silencing is that it is difficult to systematically identify which proteins make direct lncRNA interactions within the cell. Interestingly, this report identifies the human protein SHARP as an interactor of the *Xist* lncRNA. In keeping with the mode of repression described in this thesis, SHARP is essential for silencing and helps exclude Pol II from the inactive X chromosome via activation of HDAC3 and histone deacetylation. Thus, while the specific proteins and lncRNAs themselves may differ, the unifying pattern that is emerging involves a role for a sequence specific protein in mediating the repressive function of non-coding transcription.

6.1.6 Biosynthetic pathways and non-coding transcription

S. pombe pho1 and now also *tgp1* are not the only phosphate-responsive genes that are known to be regulated by ncRNAs; in *S. cerevisiae*, expression of both *PHO5* and *PHO84* genes is regulated by chromatin modifications controlled by the action of ncRNAs. It is becoming something of a theme, then, that genes involved in the phosphate response in both fission and budding yeasts are heavily reliant on transcriptional regulation mediated by ncRNAs. In this thesis, this dependence is extended to other *S. pombe* metabolic pathways, namely regulation of zinc levels by *zrt1*. Considering the importance of phosphate and other metabolites within the cell, regulation and maintenance of their constant levels is of great consequence to survival. Thus, *pho1*, together with the other genes addressed in this work, provide useful models for dissecting the diverse mode of action of ncRNAs and their transcription.

6.2 The role of CID proteins in termination of non-coding and protein-coding transcripts

As previously mentioned, termination of budding yeast ncRNAs is dependent on the Nrd1 complex, and is connected to subsequent 3′ end processing and RNA degradation by the Trf4/5-Air1/2-Mtr4 polyadenylation (TRAMP) and exosome complexes. In a similar vein, ongoing work into the function of the homologue in fission yeast, Seb1, is beginning to intimate a role in transcription termination. This fits with its observed recruitment to protein-coding gene pA sites and co-purification with the polyA/cleavage machinery (Fig. 2-7 and Fig. 3-1). Whether Seb1 has a function analogous to Nrd1, however, is something which remains to be seen. Certainly, based on the CTD binding and ChIP analysis presented in this thesis, the two proteins do not play identical roles within the cell but there remains potential for some functional cross-over.

RNA extracted from *seb1-1* and *nmt::seb1* strains which are probed for the *pho1* and *tgp1* ORFs reveals the presence of several unidentified extended species (data not shown). Considering transcription termination could represent a global role for Seb1, and also potentially relates to the inviability of a *seb1*Δ deletion strain, it would be interesting to ascertain the origin of these extended species and determine whether *pho1* and *tgp1* read-through transcripts are made in Seb1-depleted and mutant strains. Hence, it may be that where Nrd1 and Seb1 differ is the type of transcripts they terminate - Nrd1 functioning as we know at ncRNAs and, potentially, Seb1 at protein-coding RNAs.

As well as transcription termination, it is possible that other roles may be shared between the two proteins. In fact, like Seb1, a role for the Nrd1 complex in regulating genes which are involved in the cellular response to environmental nutrient fluctuations has been established (Darby et al., 2012). In this report a set of mRNAs which are targeted by Nrd1 were identified. These mRNAs were found to be rapidly repressed in response to glucose starvation, concomitant with Nrd1 and Nab3 localisation to nuclear speckles. So, although the mechanisms may differ, the ability of Seb1 and Nrd1 to regulate metabolic genes seems to have been retained, or at the very least convergently evolved, over the 400 million or so years that separate the two species (Sipiczki 2000). Clearly there is at least some functional overlap between these two proteins, though exactly how closely these functions resemble one another requires further study.

Whereas its CTD binding has been fairly well characterised, to date, very little is known about the mode of action of the human CID protein SCAF8. In view of the fact that SCAF8 has a higher affinity for Ser2-(P) CTD than Ser5-(P) CTD, it is likely that the highest density of the protein is found at the 3' ends of genes (Becker et al., 2008). The ability of SCAF8 to associate with the elongating polymerase, together with its presence in the nuclear matrix and resemblance to the SR proteins previously implicated in splicing, imply a potential role in the coupling of transcription and pre-mRNA processing (Patturajan et al., 1998). However, the exact molecular details of how SCAF8 fulfills any splicing-related or regulatory role awaits further investigation. Despite this it is interesting to speculate, especially considering Seb1's CTD binding (which is perhaps more similar to SCAF8 than Nrd1) and its association

with various splicing proteins (Fig. 3-1), of where Seb1 fits on the evolutionary scale that separates budding yeast Nrd1 and human SCAF8.

6.3 A dual role for Seb1

6.3.1 Seb1 recruitment to protein-coding vs. non-coding genes

An interesting facet of Seb1 binding is how it is able to discriminate between non-coding and protein-coding genes. Specifically, the question one must ask is how does Seb1 inherently 'know' where (and, equally important, where *not*) to recruit Clr3 and its repressive effects. It is feasible that Seb1 binding to a specific sequence on a nascent non-coding transcript such as *prt1* brings about a conformational change in Seb1 that allows for a Clr3 interaction. Alternatively, interactions at the pA site may result in Seb1 taking on a conformation that is inhibitory to Clr3 recruitment. Or, perhaps the pA machinery and Clr3 interact with Seb1 at the same site - meaning their interactions are mutually exclusive. It may also be that there are factors that have yet to be identified which play a role in this discrimination. Undoubtedly this is a question which requires further study but structural analysis of the Clr3 and Seb1 association would likely be very informative.

For all protein-coding genes tested (including *pho1* and *adh1*) (see Fig. 2-7) Seb1 is seen by ChIP analysis to peak at the pA sites. A Seb1 purification also indicates that the protein interacts with the cleavage and 3' end processing machinery (Fig. 3-1). It is therefore not unreasonable to presume that Seb1 has some - most likely global - role in either 3' end cleavage/ processing or transcription termination of protein-coding genes. The potential for this role has been addressed in some detail in section 6.2.1. The second role for Seb1 concerns the establishment of TI and relates to its association with both Pol II and ncRNA. Of these two determinants, likely ncRNA

has the greatest potential for providing the specificity of Seb1 recruitment. This is because, unlike Pol II, which (aside from changes to CTD modifications) is quite constant from one transcription unit to another, a particular ncRNA sequence can provide a highly specific binding platform. Moreover, if indeed Pol II was essential for Seb1's recruitment to targets of transcriptional repression then the complete loss of binding observed in RNase-treated ChIP samples would not be expected. In further support of this idea, budding yeast Nrd1 functions exactly in this fashion - it selects its ncRNA targets via a specific sequence.

The GUA[A/G] consensus motif corresponds to the Nrd1-binding site (Steinmetz and Brow 1998; Carroll et al., 2004). Mutations that alter this sequence lead to read-through transcription of non-polyadenylated transcripts, indicating that recognition by Nrd1 is essential for efficient termination (Schulz et al., 2013). Recently however, it has been elegantly demonstrated (Bacikova et al., 2014) that Nrd1 binding to some transcripts *in vitro* does not necessarily require this motif, and the protein can instead recognise several other G- and AU-rich motifs. This additional specificity comes from interactions between the RRM and the N- and C-terminal regions of the RRM which fold together into a helix-loop bundle. This broad spectrum of Nrd1 targets also corroborates *in vivo* data (Wlotzka et al., 2011; Jamonnak et al., 2011; Creamer et al., 2011). Although the consensus motif mentioned above is not present in *prt*, if Seb1 is similar to Nrd1 in this regard, it may not necessitate the exact sequence. There is, for example, a GUA present in the region believed (from *prt* deletion experiments - see above) to represent the Seb1 binding site. Most likely though, structural and biochemical studies will help elucidate this.

It would doubtless be edifying to perform genome-wide analysis of Seb1 binding. This would be highly informative as to the types of genes (both non-coding and protein-coding) the protein is associated with, and would perhaps help pinpoint a Seb1 binding motif(s). It could also help to glean information regarding the apparent bifunctional role of the protein - in Clr3-dependent repression and transcription termination. At present though it remains to be seen how exactly Seb1 finds its target loci and whether it, like Nrd1, binds a specific motif within nascent RNA.

6.3.2 Other Seb1 protein interactions and possible roles

As well as the Rpb1 subunit, via its CTD region, a Seb1 purification revealed an interaction with the Pol II Rpb7 subunit. This is consistent with a previous report showing a direct interaction *in vitro* (Mitsuzawa et al., 2003). Similarly, this interaction is also made with the budding yeast homologue Nrd1. Because the CTD is highly flexible, the interaction between Seb1 and Rpb7 may be a prerequisite for anchoring the former to the polymerase. On account of the close proximity of Rpb7 to the emerging nascent RNA (see Armache et al., 2003; Bushnell and Kornberg, 2003 for structure) it is reasonable to suppose that this subunit is able to arbitrate between Seb1's ability to bind both CTD and RNA. Thus, Seb1 recruitment to the transcription machinery is not necessarily solely restricted to binding to the CTD. This should therefore be borne in mind when analysing Seb1 recruitment over genes; clearly there are many more layers to consider than simply comparing Seb1 binding profiles to phospho-CTD patterns.

Seb1 was also found to co-purify with the SHREC component Clr3. This is consistent with co-IP experiments showing a direct interaction between the two (Marina et al., 2013). It would be interesting to determine the region of Seb1 that is responsible for this interaction. In budding yeast, Nrd1 interacts with the RNA-binding protein Nab3 via an interaction domain which proceeds directly after the CID. However, the Nab3 homologue could not be found in a Seb1 purification, nor was a *seb1-1* mutation found to affect snoRNA processing (data not shown). Therefore it seems unlikely that Seb1 interacts with the Nab3 homologue or shares a role analogous to Nrd1's function in the termination of snoRNAs and other short Pol II transcripts. Instead, one might surmise that the region between Seb1's CID and RRM domains may be responsible for Clr3 association. Concurrently, protein binding sites are predicted in this region (Disorder Prediction Metaserver (DisMeta)) (Sina Wittmann, personal communication).

6.3.3 Concluding remarks

Non-coding transcription is widespread in all domains of life. The intergenic regions from which ncRNAs are transcribed were previously referred to by the term "junk DNA". However, more and more examples are being found of ncRNAs with important biological functions; in particular they appear to have huge regulatory potential. Non-coding transcription can play a regulatory role via TI and it is shown here for the first time that this can be mediated by repressors associated with transcribing Pol II and/or nascent ncRNA (Fig. 5-1). This represents a mechanism by which TI can operate which has not previously been envisioned.

To conclude then, a new mechanism of TI mediated by Seb1 has been identified at *pho1*, and potentially operates quite generally at various genes involved in biosynthetic pathways. This can involve recruitment of Clr3 as a repressor (via its deacetylation activity). However, initial work at the *tgp1* gene indicates that Seb1 can also mediate transcriptional repression in others ways. Clearly then, in terms of identifying and characterising the seemingly diverse and manifold mechanisms by which TI operates, the surface is only just beginning to be scratched.

References

Allmang, C., Petfalski, E., Podtelejnikov, A., Mann, M., Tollervey, D., & Mitchell, P. (1999a). The yeast exosome and human PM-Scl are related complexes of 3' →5' exonucleases. *Genes and Development*, *13*(16), 2148–2158.

Allmang, C., Kufel, J., Chanfreau, G., Mitchell, P., Petfalski, E., & Tollervey, D. (1999b). Functions of the exosome in rRNA, snoRNA and snRNA synthesis. *EMBO Journal*, *18*(19), 5399–5410.

Amaral, P. P., & Mattick, J. S. (2008). Noncoding RNA in development. *Mammalian Genome : Official Journal of the International Mammalian Genome Society*, *19*(7-8), 454–92.

Armache, K.-J., Kettenberger, H., & Cramer, P. (2003). Architecture of initiation-competent 12-subunit RNA polymerase II. *Proceedings of the National Academy of Sciences of the United States of America*, *100*(12), 6964–6968.

Aygün, O., Mehta, S., & Grewal, S. I. S. (2013). HDAC-mediated suppression of histone turnover promotes epigenetic stability of heterochromatin. *Nature Structural & Molecular Biology*, *20*(5), 547–54.

Bacikova, V., Pasulka, J., Kubicek, K., & Stefl, R. (2014). Structure and semi-sequence-specific RNA binding of Nrd1. *Nucleic Acids Research*, *42*(12), 8024–8038.

Baejen, C., Torkler, P., Gressel, S., Essig, K., Sö, J., & Cramer, P. (2014). Transcriptome Maps of mRNP Biogenesis Factors Define Pre-mRNA Recognition. *Molecular Cell*, *55*, 745–757.

Bahler, R. G. (1998). Heterologous Modules for Efficient and Versatile PCR-based Gene Targeting in Schizosaccharomyces pombe, *951*(February), 943–951.

Becker, R., Loll, B., & Meinhart, A. (2008). Snapshots of the RNA processing factor SCAF8 bound to different phosphorylated forms of the carboxyl-terminal domain of RNA polymerase II. *The Journal of Biological Chemistry*, *283*(33), 22659–69.

Berretta, J., Pinskaya, M., & Morillon, A. (2008). A cryptic unstable transcript mediates transcriptional trans-silencing of the Ty1 retrotransposon in S. cerevisiae. *Genes & Development*, *22*(5), 615–26.

Boussadia, O., Amiot, F., Cases, S., Triqueneaux, G., Jacquemin-Sablon, H., & Dautry, F. (1997). Transcription of unr (pstream of -ras) down-modulates N-ras expression in vivo. *FEBS Letters*, *420*(1), 20–24.

Bühler, M., & Moazed, D. (2007). Transcription and RNAi in heterochromatic gene silencing. *Nature Structural & Molecular Biology*, *14*(11), 1041–8.

Bühler, M., Spies, N., Bartel, D. P., & Moazed, D. (2008). TRAMP-mediated RNA surveillance prevents spurious entry of RNAs into the Schizosaccharomyces pombe siRNA pathway. *Nature Structural & Molecular Biology*, *15*(10), 1015–23.

Bushnell, D. A., & Kornberg, R. D. (2003). Complete, 12-subunit RNA polymerase II at 4.1-A resolution: implications for the initiation of transcription. *Proceedings of the National Academy of Sciences of the United States of America*, *100*(12), 6969–73.

Callahan, K. P., & Butler, J. S. (2008). Evidence for core exosome independent function of the nuclear exoribonuclease Rrp6p. *Nucleic Acids Research*, *36*(21), 6645–6655.

Canzio, D., Chang, E. Y., Shankar, S., Kuchenbecker, K. M., Simon, M. D., Madhani, H. D., et al. (2011). Chromodomain-mediated oligomerization of HP1 suggests a nucleosome-bridging mechanism for heterochromatin assembly. *Molecular Cell*, *41*(1), 67–81.

Canzio, D., Liao, M., Naber, N., Pate, E., Larson, A., Wu, S., et al. (2013). A conformational switch in HP1 releases auto-inhibition to drive heterochromatin assembly. *Nature*, *496*(7445), 377–81.

Carroll, K. L., Carroll, K. L., Pradhan, D. A., Pradhan, D. A., Granek, J. A., Granek, J. A., et al. (2004). Identication of cis Elements Directing Termination of Yeast Nonpolyadenylated snoRNA Transcripts. *Society*, *24*(14), 6241–6252.

Carter-O'Connell, I., Peel, M. T., Wykoff, D. D., & O'Shea, E. K. (2012). Genome-wide characterization of the phosphate starvation response in Schizosaccharomyces pombe. *BMC Genomics*, *13*(1), 697.

Castelnuovo, M., Rahman, S., Guffanti, E., Infantino, V., Stutz, F., & Zenklusen, D. (2013). Bimodal expression of PHO84 is modulated by early termination of antisense transcription. *Nature Structural & Molecular Biology*, *20*(7), 851–8.

Cho, E. J., Takagi, T., Moore, C. R., & Buratowski, S. (1997). mRNA capping enzyme is recruited to the transcription complex by phosphorylation of the RNA polymerase II carboxy-terminal domain. *Genes & Development*, *11*(24), 3319–3326.

Ciais, D., Bohnsack, M. T., & Tollervey, D. (2008). The mRNA encoding the yeast ARE-binding protein Cth2 is generated by a novel 3' processing pathway. *Nucleic Acids Research*, *36*(9), 3075–84.

Colmenares, S. U., Buker, S. M., Buhler, M., Dlakić, M., & Moazed, D. (2007). Coupling of double-stranded RNA synthesis and siRNA generation in fission yeast RNAi. *Molecular Cell*, *27*(3), 449–61.

Creamer, T. J., Darby, M. M., Jamonnak, N., Schaughency, P., Hao, H., Wheelan, S. J., & Corden, J. L. (2011). Transcriptome-wide binding sites for components of

the Saccharomyces cerevisiae non-poly(A) termination pathway: Nrd1, Nab3, and Sen1. *PLoS Genetics*, *7*(10), e1002329.

Creamer, K. M., Job, G., Shanker, S., Neale, G. a, Lin, Y.-C., Bartholomew, B., & Partridge, J. F. (2014). The Mi-2 Homolog Mit1 Actively Positions Nucleosomes within Heterochromatin To Suppress Transcription. *Molecular and Cellular Biology*, *34*(11), 2046–61.

Crick, F.H.C. (1956) On Protein Synthesis. *Symp. Soc. Exp. Biol. XII:* 139-163.

Dainty, S. J., Kennedy, C. a, Watt, S., Bähler, J., & Whitehall, S. K. (2008). Response of Schizosaccharomyces pombe to zinc deficiency. *Eukaryotic Cell*, *7*(3), 454–64.

Darby, M. M., Serebreni, L., Pan, X., Boeke, J. D., & Corden, J. L. (2012). The Saccharomyces cerevisiae Nrd1-Nab3 Transcription Termination Pathway Acts in Opposition to Ras Signaling and Mediates Response to Nutrient Depletion. *Molecular and Cellular Biology*, *32*(10), 1762–1775.

Darwin, C. R. (1859). On the Origin of the Species by Means of Natural Selection, or, the Preservation of Favoured Races in the Struggle for Life. *John Murray, London* .

Drinnenberg, I. A., Weinberg, D. E., Xie, K. T., Mower, J. P., Wolfe, K. H., Fink, G. R., & Bartel, D. P. (2009). RNAi in budding yeast. *Science (New York, N.Y.)*, *326*(5952), 544–550.

Egan, E. D., Braun, C. R., Gygi, S. P., & Moazed, D. (2014). Post-transcriptional regulation of meiotic genes by a nuclear RNA silencing complex. *RNA (New York, N.Y.)*, *20*(6), 867–81.

Eick, D., & Geyer, M. (2013). The RNA polymerase II carboxy-terminal domain (CTD) code. *Chem Rev*, *113*(11), 8456–8490.

Espinoza, C. A., Goodrich, J. A., & Kugel, J. F. (2007). Characterization of the structure, function, and mechanism of B2 RNA, an ncRNA repressor of RNA polymerase II transcription. *RNA (New York, N.Y.)*, *13*(4), 583–96.

Fazzio, T. G., & Rando, O. J. (2012). NURDs are required for diversity. *The EMBO Journal*, *31*(14), 3036–7.

Finnin, M. S., Donigian, J. R., Cohen, A., Richon, V. M., Rifkind, R. A., Marks, P. A., et al. (1999). Structures of a histone deacetylase homologue bound to the TSA and SAHA inhibitors. *Nature*.

Gao, J., Kan, F., Wagnon, J. L., Storey, A. J., Protacio, R. U., Davidson, M. K., & Wahls, W. P. (2014). Rapid, efficient and precise allele replacement in the fission yeast Schizosaccharomyces pombe. *Current Genetics*, *60*(2), 109–119.

Garcia, J. F., Dumesic, P. A., Hartley, P. D., El-Samad, H., & Madhani, H. D. (2010). Combinatorial, site-specific requirement for heterochromatic silencing factors in the elimination of nucleosome-free regions. *Genes and Development*, *24*(16), 1758–1771.

Gavin, A.-C., Bösche, M., Krause, R., Grandi, P., Marzioch, M., Bauer, A., et al. (2002). Functional organization of the yeast proteome by systematic analysis of protein complexes. *Nature*, *415*(6868), 141–147.

Gerace, EL., Halic, M., Moazed, D. (2010). The Methyltransferase Activity of Clr4Suv39h triggers RNAi Independently of Histone H3K9 Methylation. *Molecular Cell*, *39*(3), 360–372.

Greger, I. H., Aranda, A., & Proudfoot, N. (2000). Balancing transcriptional interference and initiation on the GAL7 promoter of Saccharomyces cerevisiae. *Proceedings of the National Academy of Sciences of the United States of America*, *97*(15), 8415–8420.

Gudipati, R. K., Villa, T., Boulay, J., & Libri, D. (2008). Phosphorylation of the RNA polymerase II C-terminal domain dictates transcription termination choice. *Nature Structural & Molecular Biology*, *15*(8), 786–794.

Gudipati, R. K., Xu, Z., Lebreton, A., Séraphin, B., Steinmetz, L. M., Jacquier, A., & Libri, D. (2012). Extensive Degradation of RNA Precursors by the Exosome in Wild-Type Cells. *Molecular Cell*, *48*(3), 409–421.

Gullerova, M., Moazed, D., & Proudfoot, N. J. (2011). Autoregulation of convergent RNAi genes in fission yeast. *Genes & Development*, *25*(6), 556–68.

Harigaya, Y., Tanaka, H., Yamanaka, S., Tanaka, K., Watanabe, Y., Tsutsumi, C., et al. (2006). Selective elimination of messenger RNA prevents an incidence of untimely meiosis. *Nature*, *442*(7098), 45–50.

Henry, T. C., Power, J. E., Kerwin, C. L., Mohammed, A., Weissman, J. S., Cameron, D. M., & Wykoff, D. D. (2011). Systematic screen of schizosaccharomyces pombe deletion collection uncovers parallel evolution of the phosphate signal transduction pathway in yeasts. *Eukaryotic Cell*, *10*(2), 198–206.

Heo, D. H., Yoo, I., Kong, J., Lidschreiber, M., Mayer, A., Choi, B. Y., et al. (2013). The RNA polymerase II C-terminal domain-interacting domain of yeast Nrd1 contributes to the choice of termination pathway and couples to RNA processing by the nuclear exosome. *J Biol Chem*, *288*(51), 36676–36690.

Hiriart, E., Vavasseur, A., Touat-Todeschini, L., Yamashita, A., Gilquin, B., Lambert, E., et al. (2012). Mmi1 RNA surveillance machinery directs RNAi complex RITS to specific meiotic genes in fission yeast. *The EMBO Journal*, *31*(10), 2296–308.

Hirota, K., Miyoshi, T., Kugou, K., Hoffman, C. S., Shibata, T., & Ohta, K. (2008). Stepwise chromatin remodelling by a cascade of transcription initiation of non-coding RNAs. *Nature*, *456*(7218), 130–134.

Ho, Y., Gruhler, A., Heilbut, A., Bader, G. D., Moore, L., Adams, S.-L., et al. (2002). Systematic identification of protein complexes in Saccharomyces cerevisiae by mass spectrometry. *Nature*, *415*(6868), 180–183.

Holley RW, Apgar J, Everett GA, Madison JT, Marquisee M, Merrill SH, Penswick JR, Zamir A. (1965). Structure of a Ribonucleic Acid. *Science*, *147*(3664), 1462–5.

Hong, E. E., Villén, J., Gerace, E. L., Gygi, S. P., & Moazed, D. (2005). A cullin E3 ubiquitin ligase complex associates with Rik1 and the Clr4 histone H3-K9 methyltransferase and is required for RNAi-mediated heterochromatin formation. *RNA Biology*, *2*(3), 106–111.

Houseley, J., Rubbi, L., Grunstein, M., Tollervey, D., & Vogelauer, M. (2008). A ncRNA modulates histone modification and mRNA induction in the yeast GAL gene cluster. *Molecular Cell*, *32*(5), 685–695.

Hsin, J.-P., & Manley, J. L. (2012). The RNA polymerase II CTD coordinates transcription and RNA processing. *Genes & Development*, *26*(19), 2119–37.

Jamonnak, N., Creamer, T. J., Darby, M. M., Schaughency, P., Wheelan, S. J., & Corden, J. L. (2011). Yeast Nrd1, Nab3, and Sen1 transcriptome-wide binding maps suggest multiple roles in post-transcriptional RNA processing. *RNA (New York, N.Y.)*, *17*(11), 2011–25.

Kaffman, A., Rank, N. M., O'Neill, E. M., Huang, L. S., & O'Shea, E. K. (1998). The receptor Msn5 exports the phosphorylated transcription factor Pho4 out of the nucleus. *Nature*, *396*(December), 482–486.

Keogh, M.-C., & Buratowski, S. (2004). Using chromatin immunoprecipitation to map cotranscriptional mRNA processing in Saccharomyces cerevisiae. *Methods in Molecular Biology (Clifton, N.J.)*, *257*, 1–16.

Kim, M., Ahn, S.-H., Krogan, N. J., Greenblatt, J. F., & Buratowski, S. (2004a). Transitions in RNA polymerase II elongation complexes at the 3' ends of genes. *The EMBO Journal*, *23*(2), 354–64.

Kim, M., Krogan, N. J., Vasiljeva, L., Rando, O. J., Nedea, E., Greenblatt, J. F., & Buratowski, S. (2004b). The yeast Rat1 exonuclease promotes transcription termination by RNA polymerase II. *Nature*, *432*(7016), 517–22.

Kornberg, A., Rao, N. N., & Ault-Riché, D. (1999). Inorganic polyphosphate: a molecule of many functions. *Annual Review of Biochemistry*. *68*, 89-125.

Kubicek, K., Cerna, H., Holub, P., Pasulka, J., Hrossova, D., Loehr, F., et al. (2012). Serine phosphorylation and proline isomerization in RNAP II CTD control recruitment of Nrd1. *Genes & Development*, *26*(17), 1891–6.

Lemieux, C., Marguerat, S., Lafontaine, J., Barbezier, N., Bähler, J., & Bachand, F. (2011). A Pre-mRNA degradation pathway that selectively targets intron-containing genes requires the nuclear poly(A)-binding protein. *Molecular Cell*, *44*(1), 108–119.

Li, H., Motamedi, M. R., Yip, C. K., Wang, Z., Walz, T., Patel, D. J., & Moazed, D. (2009). An Alpha Motif at Tas3 C Terminus Mediates RITS cis Spreading and Promotes Heterochromatic Gene Silencing. *Molecular Cell*, *34*(2), 155–167.

Licatalosi, D. D., Geiger, G., Minet, M., Schroeder, S., Cilli, K., McNeil, J. B., & Bentley, D. L. (2002). Functional Interaction of Yeast Pre-mRNA 3′ End Processing Factors with RNA Polymerase II. *Molecular Cell*, *9*(5), 1101–1111.

Lunde, B. M., Reichow, S. L., Kim, M., Suh, H., Leeper, T. C., Yang, F., et al. (2010). Cooperative interaction of transcription termination factors with the RNA polymerase II C-terminal domain. *Nature Structural & Molecular Biology*, *17*(10), 1195–1201.

Lykke-Andersen, S., & Jensen, T. H. (2006). CUT it out: silencing of noise in the transcriptome. *Nature Structural & Molecular Biology*, *13*(10), 860–1.

Marín, A., Wang, M., & Gutiérrez, G. (2004). Short-range compositional correlation in the yeast genome depends on transcriptional orientation. *Gene*, *333*, 151–5.

Marina, D. B., Shankar, S., Natarajan, P., Finn, K. J., & Madhani, H. D. (2013). A conserved ncRNA-binding protein recruits silencing factors to heterochromatin through an RNAi-independent mechanism. *Genes & Development*, *27*(17), 1851–6.

Martens, J. A., Laprade, L., & Winston, F. (2004). Intergenic transcription is required to repress the Saccharomyces cerevisiae SER3 gene. *Nature*, *429*(6991), 571–574.

Mata, J., Lyne, R., Burns, G., & Bähler, J. (2002). The transcriptional program of meiosis and sporulation in fission yeast. *Nature Genetics*, *32*(1), 143–147.

McHugh, C. A., Chen, C.-K., Chow, A., Surka, C. F., Tran, C., McDonel, P., et al. (2015). The Xist lncRNA interacts directly with SHARP to silence transcription through HDAC3. *Nature*, *521*(7551), 232–236.

Meinhart, A., & Cramer, P. (2004). Recognition of RNA polymerase II carboxy-terminal domain by 3'-RNA-processing factors. *Nature*, *430*(6996), 223–226.

Mendel, G. (1866). Versuche über Pflanzenhybriden. *Verh. Nat.forsch. Ver. Brünn*, *4*, 3–47.

Miescher, F. (1871). Ueber die chemische Zusammensetzung der Eiterzellen. *Medicinisch-Chemische Untersuchungen*, *4*, 441–460.

Mitsuzawa, H., Kanda, E., & Ishihama, A. (2003). Rpb7 subunit of RNA polymerase II interacts with an RNA-binding protein involved in processing of transcripts. *Nucleic Acids Research*, *31*(16), 4696–4701.

Mohammad, F., Pandey, R. R., Nagano, T., Chakalova, L., Mondal, T., Fraser, P., & Kanduri, C. (2008). Kcnq1ot1/Lit1 Noncoding RNA Mediates Transcriptional Silencing by Targeting to the Perinucleolar Region. *Molecular and Cellular Biology*, *28*(11), 3713–3728.

Moreno, S., Klar, A., Nurse, P. (1991). Molecular genetic analysis of fission yeast Schizosaccharomyces pombe. *Methods Enzymol.*, *194*, 795–823.

Motamedi, M. R., Hong, E. E., Li, X., Gerber, S., Gygi, S., & Moazed, D. (2008). HP1 Proteins Form Distinct Complexes and Mediate Heterochromatic Gene Silencing by Non-Overlapping Mechanisms. *Molecular cell*, *32*(6), 778–790.

Nagano. T., Mitchell, J.A., Sanz, L.A., Pauler, F.M., Ferguson-Smith, A.C., Feil, R., Fraser, P. (2008) The Air noncoding RNA epigenetically silences transcription by targeting G9a to chromatin. *Science*, *322*(5908), 1717–1720.

Nakagawa, S., Ip, J. Y., Shioi, G., Tripathi, V., Zong, X., Hirose, T., & Prasanth, K. V. (2012). Malat1 is not an essential component of nuclear speckles in mice. *Rna*, *18*(8), 1487–1499.

O'Neill, E. M., Kaffman, A., Jolly, E. R., & O'Shea, E. K. (1996). Regulation of PHO4 nuclear localization by the PHO80-PHO85 cyclin-CDK complex. *Science (New York, N.Y.)*, *271*(5246), 209–212.

Oshima, Y. (1997). The phosphatase system in Saccharomyces cerevisiae. *Genes & Genetic Systems*, *72*(6), 323–34.

Patturajan, M., Wei, X., Berezney, R., & Corden, J. L. (1998). A nuclear matrix protein interacts with the phosphorylated C-terminal domain of RNA polymerase II. *Molecular and Cellular Biology*, *18*(4), 2406–15.

Persson, B. L., Petersson, J., Fristedt, U., Weinander, R., Berhe, A., & Pattison, J. (1999). Phosphate permeases of Saccharomyces cerevisiae: Structure, function and regulation. *Biochimica et Biophysica Acta - Reviews on Biomembranes*, *1422*(3), 255–272.

Peterlin, B. M., Brogie, J. E., & Price, D. H. (2012). 7SK snRNA: A noncoding RNA that plays a major role in regulating eukaryotic transcription. *Wiley Interdisciplinary Reviews: RNA*, *3*(1), 92–103.

Pinskaya, M., Gourvennec, S., & Morillon, A. (2009). H3 lysine 4 di- and tri-methylation deposited by cryptic transcription attenuates promoter activation. *The EMBO Journal*, *28*(12), 1697–1707.

Prescott, E. M., & Proudfoot, N. J. (2002). Transcriptional collision between convergent genes in budding yeast. *Proceedings of the National Academy of Sciences of the United States of America*, *99*(13), 8796–8801.

Rinn, J. L., Kertesz, M., Wang, J. K., Squazzo, S. L., Xu, X., Brugmann, S. A., et al. (2007). Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. *Cell*, *129*(7), 1311–23.

Schneider, C., Leung, E., Brown, J., & Tollervey, D. (2009). The N-terminal PIN domain of the exosome subunit Rrp44 harbors endonuclease activity and tethers Rrp44 to the yeast core exosome. *Nucleic Acids Research*, *37*(4), 1127–40.

Schorderet, P., & Duboule, D. (2011). Structural and functional differences in the long non-coding RNA hotair in mouse and human. *PLoS Genetics*, *7*(5), e1002071.

Schulz, D., Schwalb, B., Kiesel, A., Baejen, C., Torkler, P., Gagneur, J., et al. (2013). Transcriptome surveillance by selective termination of noncoding RNA synthesis. *Cell*, *155*(5), 1075–1087.

Schwer, B., Bitton, D. A., Sanchez, A. M., Bähler, J., & Shuman, S. (2014). Individual letters of the RNA polymerase II CTD code govern distinct gene expression programs in fission yeast. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(11), 4185–90.

Shah, S., Wittmann, S., Kilchert, C., & Vasiljeva, L. (2014). lncRNA recruits RNAi and the exosome to dynamically regulate pho1 expression in response to phosphate levels in fission yeast. *Genes & Development*, *28*(3), 231–44.

Sipiczki, M. (2000). Where does fission yeast sit on the tree of life? *Genome Biology*, *1*(2).

Smilinich, N.J., Day, C.D., Fitzpatrick, G.V., Caldwell, G.M., Lossie, A.C., et al. (1999). A maternally methylated CpG island in KvLQT1 is associated with an antisense paternal transcript and loss of imprinting in Beckwith-Wiedemann syndrome. *Proceedings of the National Academy of Sciences of the United States of America*. 96(6), 8064–8069.

St-André, O., Lemieux, C., Perreault, A., Lackner, D. H., Bähler, J., & Bachand, F. (2010). Negative regulation of meiotic gene expression by the nuclear poly(a)-binding protein in fission yeast. *The Journal of Biological Chemistry*, *285*(36), 27859–68.

Steinmetz, E. J., & Brow, D. A. (1996). Repression of gene expression by an exogenous sequence element acting in concert with a heterogeneous nuclear ribonucleoprotein-like protein, Nrd1, and the putative helicase Sen1. *Molecular and Cellular Biology*, *16*(12), 6993–7003.

Steinmetz, E. J., & Brow, D. A. (1998). Control of pre-mRNA accumulation by the essential yeast protein Nrd1 requires high-affinity transcript binding and a

domain implicated in RNA polymerase II association. *Proceedings of the National Academy of Sciences of the United States of America*, *95*(12), 6699–704.

Steinmetz, E. J., Conrad, N. K., Brow, D. A., & Corden, J. L. (2001). RNA-binding protein Nrd1 directs poly(A)-independent 3'-end formation of RNA polymerase II transcripts. *Nature*, *413*(6853), 327–331.

Sugiyama, T., Cam, H. P., Sugiyama, R., Noma, K. I., Zofall, M., Kobayashi, R., & Grewal, S. I. S. (2007). SHREC, an Effector Complex for Heterochromatic Transcriptional Silencing. *Cell*, *128*(3), 491–504.

Sugiyama, T., & Sugioka-Sugiyama, R. (2011). Red1 promotes the elimination of meiosis-specific mRNAs in vegetatively growing fission yeast. *The EMBO Journal*, *30*(6), 1027–39.

Sugiyama, T., Sugioka-Sugiyama, R., Hada, K., & Niwa, R. (2012). Rhn1, a nuclear protein, is required for suppression of meiotic mRNAs in mitotically dividing fission yeast. *PloS One*, *7*(8), e42962.

Tashiro, S., Asano, T., Kanoh, J., & Ishikawa, F. (2013). Transcription-induced chromatin association of RNA surveillance factors mediates facultative heterochromatin formation in fission yeast. *Genes to Cells : Devoted to Molecular & Cellular Mechanisms*, *18*(4), 327–39.

Thebault, P., Boutin, G., Bhat, W., Rufiange, A., Martens, J., & Nourani, A. (2011). Transcription regulation by the noncoding RNA SRG1 requires Spt2-dependent chromatin deposition in the wake of RNA polymerase II. *Molecular and Cellular Biology*, *31*(6), 1288–1300.

Tittel-Elmer, M., Bucher, E., Broger, L., Mathieu, O., Paszkowski, J., & Vaillant, I. (2010). Stress-induced activation of heterochromatic transcription. *PLoS Genetics*, *6*(10), 1–11.

Tomar, P., & Sinha, H. (2014). Conservation of PHO pathway in ascomycetes and the role of Pho84. *Journal of Biosciences*, *39*(3), 525–536.

Van Hoof, A., Lennertz, P., & Parker, R. (2000). Yeast exosome mutants accumulate 3'-extended polyadenylated forms of U4 small nuclear RNA and small nucleolar RNAs. *Molecular and Cellular Biology*, *20*(2), 441–452.

Vasiljeva, L., & Buratowski, S. (2006). Nrd1 interacts with the nuclear exosome for 3' processing of RNA polymerase II transcripts. *Molecular Cell*, *21*(2), 239–48.

Vasiljeva, L., Kim, M., Mutschler, H., Buratowski, S., & Meinhart, A. (2008). The Nrd1-Nab3-Sen1 termination complex interacts with the Ser5-phosphorylated RNA polymerase II C-terminal domain. *Nature Structural & Molecular Biology*, *15*(8), 795–804.

Verdel, A., Jia, S., Gerber, S., Sugiyama, T., Gygi, S., Grewal, S. I. S., & Moazed, D. (2004). RNAi-mediated targeting of heterochromatin by the RITS complex. *Science (New York, N.Y.)*, *303*(5658), 672–6.

Vogel, K., Hörz, W., & Hinnen, A. (1989). The two positively acting regulatory proteins PHO2 and PHO4 physically interact with PHO5 upstream activation regions. *Molecular and Cellular Biology*, *9*(5), 2050–7.

Wagschal, A., Sutherland, H. G., Woodfine, K., Henckel, A., Chebli, K., Schulz, R., et al. (2008). G9a Histone Methyltransferase Contributes to Imprinting in the Mouse Placenta □. *Molecular and Cellular Biology*, *28*(3), 1104–1113.

Wang, P., Yang, J., Ishihama, A., & Pittard, A. J. (1998). Demonstration that the TyrR protein and RNA polymerase complex formed at the divergent P3 promoter inhibits binding of RNA polymerase to the major promoter, P1, of the aroP gene of Escherichia coli. *Journal of Bacteriology*, *180*(20), 5466–5472.

Wasmuth, E. V, Januszyk, K., & Lima, C. D. (2014). Structure of an Rrp6-RNA exosome complex bound to poly(A) RNA. *Nature*, *511*(7510), 435–9.

Watanabe, Y., & Yamamoto, M. (1994). S. pombe mei2+ encodes an RNA-binding protein essential for premeiotic DNA synthesis and meiosis I, which cooperates with a novel RNA species meiRNA. *Cell*, *78*(3), 487–498.

Watson, J.D., & Crick, F.H.C. (1953) A Structure for Deoxyribose Nucleic Acid. *Nature 171*, 737-738

Whitelaw, E., & Proudfoot, N. (1986). Alpha-thalassaemia caused by a poly(A) site mutation reveals that transcriptional termination is linked to 3' end processing in the human alpha 2 globin gene. *The EMBO Journal*, *5*(11), 2915–2922.

Wlotzka, W., Kudla, G., Granneman, S., & Tollervey, D. (2011). The nuclear RNA polymerase II surveillance system targets polymerase III transcripts. *The EMBO Journal*, *30*, 1790–1803.

Wutz, A. (2011). Gene silencing in X-chromosome inactivation: advances in understanding facultative heterochromatin formation. *Nature Reviews. Genetics*, *12*(8), 542–553.

Wutz, A., Smrzka, O.W., Schweifer, N., Schellander, K., Wagner, E.F., Barlow, D.P. (1997) Imprinted expression of the Igf2r gene depends on an intronic CpG island. *Nature*, *389*(6652):745–749.

Wyers, F., Rougemaille, M., Badis, G., Rousselle, J. C., Dufour, M. E., Boulay, J., et al. (2005). Cryptic Pol II transcripts are degraded by a nuclear quality control pathway involving a new poly(A) polymerase. *Cell*, *121*(5), 725–737.

Wykoff, D. D., Rizvi, A. H., Raser, J. M., Margolin, B., & O'Shea, E. K. (2007). Positive Feedback Regulates Switching of Phosphate Transporters in S. cerevisiae. *Molecular Cell*, *27*(6), 1005–1013.

Yamada, T., Fischle, W., Sugiyama, T., Allis, C. D., & Grewal, S. I. S. (2005). The nucleation and maintenance of heterochromatin by a histone deacetylase in fission yeast. *Molecular Cell*, *20*(2), 173–85.

Yamanaka, S., Yamashita, A., Harigaya, Y., Iwata, R., & Yamamoto, M. (2010). Importance of polyadenylation in the selective elimination of meiotic mRNAs in growing S. pombe cells. *The EMBO Journal*, *29*(13), 2173–2181.

Yamane, K., Mizuguchi, T., Cui, B., Zofall, M., Noma, K. I., & Grewal, S. I. S. (2011). Asf1/HIRA facilitate global histone deacetylation and associate with HP1 to promote nucleosome occupancy at heterochromatic loci. *Molecular Cell*, *41*(1), 56–66.

Zanchin, N. I. T., & Goldfarb, D. S. (1999). The exosome subunit Rrp43p is required for the efficient maturation of 5.8S, 18S and 25S rRNA. *Nucleic Acids Research*, *27*(5), 1283–1288.

Zhang, B., Arun, G., Mao, Y. S., Lazar, Z., Hung, G., Bhattacharjee, G., et al. (2012). The lncRNA malat1 is dispensable for mouse development but its transcription plays a cis-regulatory role in the adult. *Cell Reports*, *2*(1), 111–123.

Zofall, M., & Grewal, S. I. S. (2006). Swi6/HP1 Recruits a JmjC Domain Protein to Facilitate Transcription of Heterochromatic Repeats. *Molecular Cell*, *22*(5), 681–692.

Zofall, M., Yamanaka, S., Reyes-Turcu, F. E., Zhang, K., Rubin, C., & Grewal, S. I. S. (2012). RNA elimination machinery targeting meiotic mRNAs promotes facultative heterochromatin formation. *Science (New York, N.Y.)*, *335*(6064), 96–100.