Fission yeast Dicer harbours a unique dsRBD that participates in the spatial organization of the RNAi pathway

Inauguraldissertation

zur
Erlangung der Würde eines Doktors der Philosophie
vorgelegt der
Philosophisch-Naturwissenschaftlichen Fakultät
der Universität Basel

von

Stephan Emmerth
aus Männedorf, ZH

Basel, 2011

Genehmigt von der Philosophisch-Naturwissenschaftlichen Fakultät auf Antrag von Dr. Javier Martinez, Dr. Helge Grosshans, and Prof. Dr. Marc Bühler.

Basel, den 21. Juni 2011

Prof. Dr. Martin Spiess
Dekan

creative commons

Namensnennung-Keine kommerzielle Nutzung-Keine Bearbeitung 2.5 Schweiz

---

Sie dürfen:

das Werk vervielfältigen, verbreiten und öffentlich zugänglich machen

Zu den folgenden Bedingungen:

Namensnennung. Sie müssen den Namen des Autors/Rechteinhabers in der von ihm festgelegten Weise nennen (wodurch aber nicht der Eindruck entstehen darf, Sie oder die Nutzung des Werkes durch Sie würden entlohnt).

Keine kommerzielle Nutzung. Dieses Werk darf nicht für kommerzielle Zwecke verwendet werden.

Keine Bearbeitung. Dieses Werk darf nicht bearbeitet oder in anderer Weise verändert werden.

- Im Falle einer Verbreitung müssen Sie anderen die Lizenzbedingungen, unter welche dieses Werk fällt, mitteilen. Am Einfachsten ist es, einen Link auf diese Seite einzubinden.
- Jede der vorgenannten Bedingungen kann aufgehoben werden, sofern Sie die Einwilligung des Rechteinhabers dazu erhalten.
- Diese Lizenz lässt die Urheberpersönlichkeitsrechte unberührt.

Quelle: http://creativecommons.org/licenses/by-nc-nd/2.5/ch/ Datum: 3.4.2009

Table of contents

Abbreviations and nomenclature

small RNA	sRNA
microRNA	miRNA
Caenorhabditis elegans	*C. elegans*
double-stranded RNA	dsRNA
RNA interference	RNAi
short interfering RNA	siRNA
nucleotides	nt
RNA-induced silencing complex	RISC
post-transcriptional gene silencing	PTGS
chromatin-dependent gene silencing	CDGS
Schizosaccharomyces pombe	*S. pombe*
PIWI-interacting RNA	piRNA
double-stranded RNA binding domain	dsRBD
Piwi Argonaute Zwilli	PAZ
P-element induced wimpy testis	PIWI
RNA polymerase II	RNA Pol II
primary miRNA	pri-miRNA
Dicer-like 1	Dcl1
untranslated region	UTR
RNA-directed DNA methylation	RdDM
base-pairs	bp
histone H3 lysine 9 methylation	H3K9me
Heterochromatin Protein 1	HP1
position effect variegation	PEV
Dicer	Dcr1
Argonaute	Ago1
RNA-directed RNA polymerase	Rdp1
RNA-induced transcriptional silencing	RITS
Argonaute siRNA chaperone	ARC

RNA-directed RNA polymerase complex	RDRC
transcriptional gene silencing	TGS
co-transcriptional gene silencing	CTGS
nuclear localization signal	nls
Schizosaccharomyces japonicus	*S. japonicus*
S. japonicus Dicer	jDcr1
S. japonicus Argonaute	jAgo1
S. japonicus RNA-directed RNA polymerase	jRdp1
multivesicular bodies	MVB
processing bodies	P-bodies
dicing bodies	D-bodies
small nucleolar RNA	snoRNA

protein names	first letter in capital form
gene names	italics

1. Summary

The term RNAi describes a set of conserved pathways found in most eukaryotes. RNAi is involved in various cellular processes, ranging from the control of gene expression to the establishment of heterochromatic structures. Common to all RNAi pathways is the association of small RNAs with members of the Argonaute family of proteins, forming the core component of a diverse set of protein-RNA complexes. The small RNAs guide these complexes via base-pairing interactions to homologous sequences, which usually results in reduced activity of these targets. The vast majority of small RNA molecules are generated by Dicer enzymes by endonucleolytically processing double-stranded RNAs.

The *Schizosaccharomyces pombe* RNAi pathway is required for the formation of centromeric heterochromatin, and this process has been biochemically characterized in great detail. However, our knowledge about the spatial organization of RNAi in *Schizosaccharomyces pombe* is very limited. The few experiments performed so far, which have mainly addressed the cellular localization of Dicer and Argonaute, have resulted in data conflicting with the biochemical observations.

In my PhD thesis, I have employed yeast genetics, biochemical and proteomics approaches to untangle these conflicting data with a major focus on the investigation of Dicer localization. I was able to demonstrate for the first time that Dicer is primarily localized to the nucleus where it associates with the nuclear periphery. Furthermore, I showed that nuclear retention of Dicer is essential for the formation of centromeric heterochromatin. These findings are consistent with the existing biochemical data and further support our model proposed for the formation of centromeric heterochromatin by the RNAi pathway.

My early work demonstrated that nuclear localization of Dicer depends on its C-terminus. In a subsequent collaborative effort, we have solved the solution structure of this C-terminus, which showed that it encodes for a unique type of dsRBD and revealed novel insights into the mechanisms of nuclear retention of Dicer. Importantly, I have found that binding of this domain to RNA is dispensable for RNAi. Rather, the dsRBD represents a novel regulatory module for RNAi, which can mediate nucleo-cytoplasmic shuttling of Dicer. This feature seems to be conserved in higher eukaryotes.

My work does also suggest a new function for RNAi in fission yeast, which is different from the well-established RNAi-mediated formation of heterochromatin at the centromeres and is likely to function in controlling environmentally regulated genes. Future studies in our laboratory will focus on the dissection of the mechanistic details of this novel mode of gene regulation.

2. Introduction

The central dogma of molecular biology established by Francis Crick in 1958 (Crick, 1958) states that information only flows in one direction, from DNA via RNA to protein. This view was challenged quite early with the discovery of reverse transcriptases, which convert RNA into DNA (Baltimore, 1970; Temin and Mizutani, 1970). Nevertheless, it was long believed that RNA only serves as a passive message transporter between the DNA and the active protein. The demonstration that RNA molecules, termed ribozymes, harbour an enzymatic activity (Kruger et al., 1982; Stark et al., 1978) refuted this opinion. And the discovery of RNAi by Fire and Mello in 1998 (Fire et al., 1998), which I believe can from a historical perspective be compared with determining the structure of the double helix by Watson and Crick (Watson and Crick, 1953), has shifted RNA molecules to the centre of eukaryotic gene regulation. Since the initial discovery in 1998, research on RNAi has truly exploded, our knowledge regarding the mechanisms of RNAi has been constantly increasing and RNAi pathways have been discovered in numerous organisms, ranging from fungi to human, and they are involved in a variety of cellular processes.

My PhD thesis deals with a very intriguing aspect of RNAi, namely its ability to influence chromatin modifications in *Schizosaccharomyces pombe*. In this introduction I will first summarize the concept of RNAi in general before focusing on the RNAi-mediated heterochromatin formation in *Schizosaccharomyces pombe.*

2.1. sRNA-mediated silencing: A historical perspective

Although not recognized as such, small RNA (sRNA-)mediated silencing was first observed in 1990 in plants, when overexpression of an enzyme related to pigment synthesis led to the repression of the transgene as well as the endogenous gene, a phenomenon the authors termed cosuppression (Napoli et al., 1990; van der Krol et al., 1990). In 1993, the first microRNA (miRNA), *lin-4*, was discovered (Lee et al., 1993), and two years later it was found that sense as well as antisense RNAs can disrupt gene expression in *Caenorhabditis elegans* (*C. elegans*) (Guo and Kemphues, 1995). However, it was not until 1998 that Fire and Mello showed that injection of a double-stranded RNA (dsRNA) can specifically interfere with expression of the homologous gene in *C. elegans*, a process the authors named RNA interference (RNAi) (Fire et al., 1998). Shortly after, several groups demonstrated that the dsRNA is converted into short interfering RNAs (siRNAs) ranging from 21 - 25 nucleotides (nt) in length in plants and *Drosophila melanogaster*. The siRNAs were found to be associated with a nuclease activity, the RNA-induced silencing complex (RISC) and to mediate post-transcriptional gene silencing (PTGS) of complementary mRNAs (Elbashir et al., 2001b; Hamilton and Baulcombe, 1999; Hammond et al., 2000; Zamore et al., 2000). Shortly after, it was demonstrated that mammalian genes can be efficiently silenced by transfecting siRNAs (Elbashir et al., 2001a). This was a breakthrough for RNAi in biomedical research, since transfection of long dsRNAs leads, via the interferon response, to a general inhibition of translation. Biochemical experiments revealed that the enzyme Dicer is required to process the dsRNA into siRNAs (Bernstein et al., 2001), as well as pre-miRNAs into miRNAs (Grishok et al., 2001; Hutvagner et al., 2001; Ketting et al., 2001). Furthermore, purified RISC in mammalian cells was shown to contain Argonaute 1 or Argonaute 2 (Martinez et al., 2002), with Argonaute 2 being the enzyme responsible for the nuclease activity (Liu et al., 2004; Song et al., 2004).

In parallel, RNAi was also linked to chromatin-dependent gene silencing (CDGS). In plants, sRNAs were implicated in targeting *de novo* DNA methylation (Mette et al., 2000; Sijen et al., 2001b). Furthermore, RNAi in *Schizosaccharomyces pombe* (*S. pombe*) was found to be involved in the

formation of heterochromatin at centromeres (Volpe et al., 2002), and it was proposed that RNAi targets histone-modifying enzymes via siRNA-mediated base-pairing interactions to chromatin, which was corroborated by additional studies (Buhler et al., 2006; Verdel et al., 2004).

Besides these key findings, various groups have identified a myriad of factors involved in RNAi, many different families of sRNAs have been described, and RNAi-mediated silencing pathways have been found in most eukaryotic systems, ranging from fungi to mammals, and they have been linked to almost every aspect of eukaryotic gene regulation. Of note, a class of sRNAs termed PIWI-interacting RNAs (piRNAs) has been recently discovered that is generated independently of Dicer and is implicated in transposon silencing (Aravin et al., 2006; Brennecke et al., 2007; Girard et al., 2006; Lau et al., 2006).

2.2. sRNA-mediated silencing: An overview

Despite their mechanistic differences, all sRNA-mediated silencing pathways share the same basic concept: a sRNA termed the guide strand is loaded onto an Argonaute-containing RISC effector complex. Subsequently, RISC binds through base-pairing interactions mediated via the sRNA to a homologous sequence, which usually results in reduced activity of the target. The sRNAs themselves are predominantly generated by Dicer enzymes, which process a precursor RNA of double-stranded nature into sRNA duplexes. In some organisms, the sRNAs are additionally amplified by an RNA-directed RNA polymerase (see also Figure 1).

In this section, I will introduce Dicer and Argonaute, the two core enzymes involved in RNAi, describe quickly RNA-directed RNA polymerases, and briefly discuss miRNA, siRNA, and piRNA pathways.

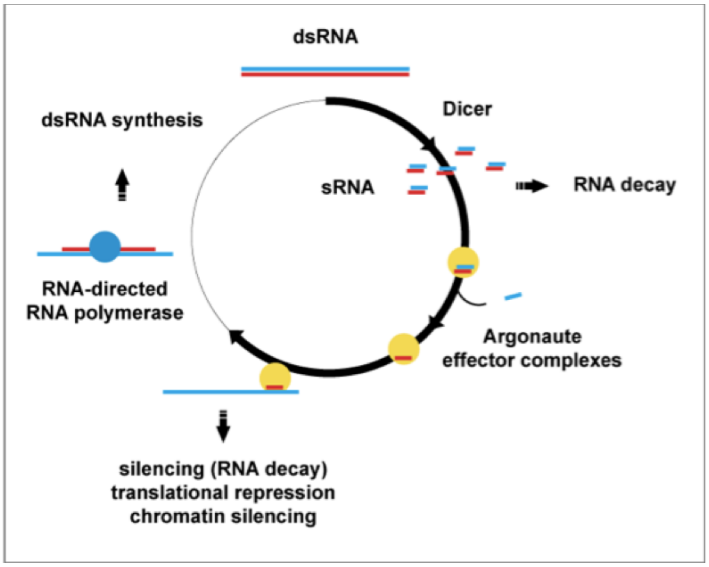


Figure 1: Introducing RNAi
A dsRNA is processed by Dicer into sRNAs, which are loaded onto Argonaute effector complexes and recruit them to homologous targets. This usually results in reduced activity of the target sequences. In some organisms, the RNAi response is amplified by an RNA-directed RNA polymerase creating additional dsRNAs, which can be processed by Dicer.

2.2.1. Dicer processes dsRNAs into sRNAs

sRNA generation involves an endonucleolytic cleavage reaction mediated by Dicer enzymes, converting a dsRNA into sRNAs, such as siRNAs and miRNAs, which are duplexes of 21 - 25 nt in size that have ~2 nt 3` overhangs and a 5` phosphate group. In mammals, *C. elegans*, and some other organisms, one single Dicer enzyme is responsible for the biogenesis of siRNAs and miRNAs (Tomari and Zamore, 2005). In other organisms, such as plants and *Drosophila melanogaster*, different Dicer enzymes process dsRNAs into either siRNAs or miRNAs (Lee et al., 2004b; Xie et al., 2004). Recently, a sRNA species that is generated independently of Dicer enzymes has been discovered (see section 2.2.6.). Of note, in many organisms Dicer enzymes associate with proteins that are important for full activity of the corresponding RNAi pathway (see section 2.3.7.3.).

Dicer enzymes belong to the family of RNase III enzymes, which can be divided into two classes: class I enzymes contain a single catalytic RNase III domain and function as homodimers whereas class II enzymes bear two catalytic RNase III domains and are active as monomers (Jaskiewicz and Filipowicz, 2008; Macrae et al., 2006; Zhang et al., 2004). Most Dicer proteins belong to class II and contain a double-stranded RNA binding domain (dsRBD) in the carboxy-terminus, which is preceded N-terminally by the two RNase III domains. The amino-terminal part usually contains a ATPase/helicase domain, a domain of unknown function (DUF283), and a Piwi Argonaute Zwilli (PAZ) domain, which

is also present in Argonaute proteins (Jinek and Doudna, 2009). Notably, in the budding yeast *Saccharomyces castelli*, a Dicer protein has recently been discovered that belongs to class I of the RNase III enzymes and contains a single RNase III domain followed by two dsRBDs. This suggests that Dicer in *Saccharomyces castelli* is active as a homodimer (Drinnenberg et al., 2009).

Dicer produces sRNAs primarily by excising them from the ends of a dsRNA. The PAZ domain preferentially binds dsRNA ends with ~2 nt 3` overhangs, and this positions the dsRNA in a cleft of an intramolecular dimer formed by the two RNase III domains. Subsequently, each of these two RNase III domains cleaves endonucleolytically one strand of the bound dsRNA leading to staggered duplex scission generating new ends with ~2 nt 3` overhangs. The mature sRNA duplex is released and loaded onto RISC. The distance between the PAZ domain and the RNase III domains is thought to function as a molecular ruler to produce sRNAs of characteristic lengths (Lingel et al., 2003, 2004; Ma et al., 2004; MacRae et al., 2007; Macrae et al., 2006; Song et al., 2003; Yan et al., 2003; Zhang et al., 2004). The role of the ATPase/helicase domain as well as that of the DUF283 domain for sRNA generation is not well understood. For example, not all Dicer enzymes require ATP for the endonucleolytic cleavage reaction (Tomari and Zamore, 2005). Interestingly, the dsRBD of *S. pombe* Dicer is not required to process a dsRNA *in vitro* (Colmenares et al., 2007), and deletion of the dsRBD of human Dicer has only minor effects on the processivity of the enzyme *in vitro* (Zhang et al., 2004). This shows that Dicer does not need all of its domains to be catalytically active. Accordingly, *Giardia intestinalis* Dicer consists solely of the PAZ domain followed by the two RNase III domains (Macrae et al., 2006) suggesting that *Giardia intestinalis* Dicer represents the minimal domain requirements of Dicer enzymes to cleave a dsRNA, whereas the other domains have evolved to fine tune Dicer activity *in vivo* and could potentially be involved in regulation of the enzyme (see also Figure 2).

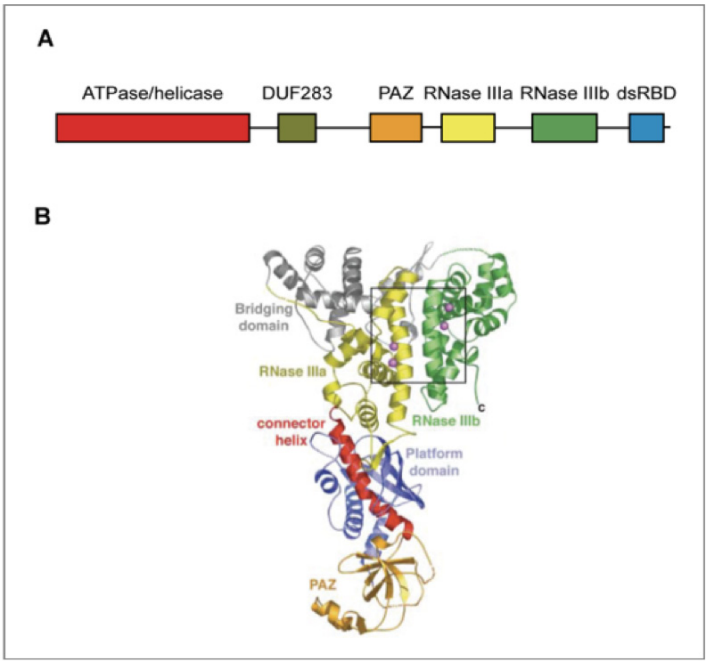


Figure 2: The schematic domain organization of human Dicer and the crystal structure of *Giardia intestinalis* Dicer
A: Human Dicer consists of an N-terminal ATPase/helicase domain followed by the DUF283, the PAZ, the two RNase III domains, and a C-terminal dsRBD. **B:** The crystal structure of *Giardia intestinalis* Dicer, which is composed of an N-terminal Platform domain (see section 3.3. for details) followed by the PAZ domain and the two C-terminal RNase III domains.
Adapted from (Macrae et al., 2006).

2.2.2. Argonaute – the key component of all RNAi-pathways

Argonaute proteins are the key players in all RNAi-mediated silencing pathways. They bind to the sRNAs, which recruit them to the target and are also required for target silencing mediated by RISC. A few organisms, like the fission yeast *S. pombe*, only encode for one Argonaute protein (Volpe et al., 2002), but most species express multiple Argonautes. *Drosophila melanogaster* for example expresses 5 Argonaute proteins, whereas 8 exist in mammals, 10 are known in *Arabidopsis thaliana*, and an astonishing 27 Argonaute family members can be found in *C. elegans* (Hunter et al., 2003; Sasaki et al., 2003; Williams and Rubin, 2002; Yigit et al., 2006). Like Dicer enzymes, Argonaute proteins are often found in multisubunit complexes, which together form the mature RISC complex. However, only the *S. pombe* specific Argonaute complex will be introduced in greater detail in this thesis (see section 2.3.3.).

Argonaute proteins can be divided into three separate subgroups. The Ago clade is usually loaded with miRNAs or siRNAs (see sections 2.2.4. and 2.2.5.). The PIWI clade proteins associate with piRNAs and are required for transposon silencing in the germlines (see section 2.2.6.). The remaining Ago clade has so far only been identified in *C. elegans* (Carthew and Sontheimer, 2009; Yigit et al., 2006).

Argonaute proteins are composed of four domains: an amino-terminal N-domain followed by the PAZ domain, the Mid domain, and the carboxy-terminal P-element induced wimpy testis (PIWI) domain. The PAZ domain of Argonaute binds to the 3` end and the Mid domain to the 5` phosphate of the sRNA, whereas the backbone of the guide strand is bound by all four domains. The PIWI domain adopts an RNase H fold, which in many Argonaute proteins contains three conserved residues that catalyze guide strand-dependent cleavage of the target RNA. However, not all Argonautes are cleavage competent, and some are merely acting as scaffolds to recruit other factors necessary for target silencing (Carthew and Sontheimer, 2009; Lingel et al., 2003, 2004; Ma et al., 2004; Ma et al., 2005; Parker et al., 2005; Song et al., 2003; Song et al., 2004; Wang et al., 2008a; Wang et al., 2009; Wang et al., 2008b; Yan et al., 2003).

2.2.3. RNA-directed RNA polymerase

In some organisms, RNA-directed RNA polymerases are required for efficient RNAi-mediated silencing by amplifying the RNAi response. They are involved in the generation of additional, secondary sRNAs. RNA-directed RNA polymerases have so far been discovered in plants, fungi, and *C. elegans*. Notably, recently an RNA-directed RNA polymerase activity has also been described in human cells (Cogoni and Macino, 1999; Dalmay et al., 2000; Maida et al., 2009; Mourrain et al., 2000; Smardon et al., 2000).

Whereas only one RNA-directed RNA polymerase is found in *S. pombe*, four exist *in C. elegans*, and six are found in *Arabidopsis thaliana*. In *Arabidopsis thaliana* and *C. elegans*, the different RNA-directed RNA polymerases are involved in the generation of different families of secondary sRNAs that are loaded onto the various Argonaute proteins in these organisms. Briefly speaking, these secondary sRNAs amplify and sustain the RNAi response in *C. elegans* and *Arabidopsis thaliana* and can lead to systemic silencing that spreads throughout the body of the organism (Carthew and Sontheimer, 2009; Pak and Fire, 2007; Sijen et al., 2001a; Voinnet, 2008). However, the role of RNA-directed RNA polymerases in these organisms will not be introduced in greater detail in this thesis.

In *S. pombe*, secondary sRNAs are generated by recruiting RNA-directed RNA polymerases to a target RNA marked by a sRNA-loaded Argonaute. The polymerase synthesizes new dsRNA molecules that can be processed by Dicer enzymes into sRNAs, and they are required for the formation of centromeric heterochromatin (see section 2.3.3.).

2.2.4. miRNA-mediated silencing

With the characterization of *lin-4* in 1993, miRNAs were the first class of sRNA to be discovered (Lee et al., 1993). They are mainly involved in PTGS in a variety of cellular pathways and processes (Wienholds and Plasterk, 2005).

miRNA biogenesis involves transcription of miRNA genes by RNA polymerase II (RNA Pol II) generating long hairpin-containing primary miRNAs (pri-miRNA) of double-stranded nature, which are capped and polyadenylated (Cai et al., 2004; Lee et al., 2004a). In animals, these pri-miRNAs are trimmed by the class II RNase III enzyme Drosha and its partner Pasha (DGCR8) into pre-miRNAs, which are exported from the nucleus by Exportin-5. In the cytoplasm, Dicer enzymes process the pre-miRNAs into ~21 nt miRNAs duplexes (Denli et al., 2004; Gregory et al., 2004; Grishok et al., 2001; Han et al., 2004; Hutvagner et al., 2001; Ketting et al., 2001; Lee et al., 2003; Lund et al., 2004; Yi et al., 2003). The miRNA duplexes usually contain internal mismatches, which target them to miRNA specific Argonautes where one strand is selected as the guide strand leading to the mature RISC complex (see also Figure 3) (Forstemann et al., 2007; Jannot et al., 2008; Schwarz et al., 2003; Steiner et al., 2007; Tomari et al., 2007). Notably, in plants Dicer-like 1 (Dcl1) is involved in processing the pri-miRNAs as well as the pre-miRNAs, and both reactions are carried out in the nucleus (Kurihara and Watanabe, 2004; Papp et al., 2003; Park et al., 2005; Park et al., 2002; Reinhart et al., 2002; Xie et al., 2004).

RISC is recruited through base-pairing interactions of the guide strand with the 3` untranslated region (UTR) to a target mRNA. Importantly, most mRNAs have multiple target sites in their 3` UTRs and several different miRNAs can target the same mRNA (Didiano and Hobert, 2006; Lewis et al., 2005; Saetrom et al., 2007; Vella et al., 2004). In most cases, base-pairing between the

target mRNA and the miRNA is imperfect with only nucleotides 2 - 8 from the miRNA (seed region) showing perfect complementary (Brennecke et al., 2005; Doench and Sharp, 2004).

The recruitment of RISC to an mRNA usually leads to its inactivation. However, the exact mechanism of silencing still remains a matter of debate. Different modes of action have been proposed, ranging from inhibition of translational initiation or elongation to degradation, either mediated directly through RISC or via deadenylation-promoted degradation. Most likely, several silencing mechanisms act in parallel and the mode of silencing depends on the size of the seed sequence, on which Argonaute protein is involved, and on the target mRNA (Chendrimada et al., 2007; Ding and Grosshans, 2009; Doench and Sharp, 2004; Giraldez et al., 2006; Guo et al., 2010; Humphreys et al., 2005; Olsen and Ambros, 1999; Petersen et al., 2006; Pillai et al., 2005; Wu et al., 2006).

2.2.5. siRNA-mediated silencing

Unlike miRNA precursors, which are transcribed from miRNA genes by RNA Pol II, siRNA precursors in mammals usually consist of exogenously introduced dsRNAs.

The maturation of siRNAs on the other hand proceeds very similarly to miRNA biogenesis: Dicer processes a dsRNA into siRNAs, which are loaded onto Argonaute proteins. Only the guide stand is retained whereas the passenger strand is eliminated. Unlike miRNAs, siRNAs normally bind with perfect complementary to their targets (see also Figure 3). While mammalian genomes encode only one Dicer protein, separate Dicer enzymes generate miRNAs and siRNAs in *Drosophila melanogaster* and in plants (Carmell and Hannon, 2004; Hammond, 2005; Hutvagner et al., 2001; Ketting et al., 2001; Margis et al., 2006; Xie et al., 2004). In contrast, animals as well as plants have evolved different Argonaute proteins for miRNA- and siRNA-related silencing pathways (Carmell et al., 2002).

Endogenously encoded siRNAs that regulate target genes have long been known to exist in plants and *C. elegans*, but until recently they were not found in other animal species (see below) (Allen et al., 2005; Brodersen and Voinnet,

2006; Pak and Fire, 2007; Ruby et al., 2006; Sijen et al., 2007; Vazquez et al., 2004). In plants, endogenously encoded siRNAs, besides being involved in PTGS, also participate in CDGS. Studies in *Arabidopsis thaliana* have shown that siRNAs are required for RNA-directed DNA methylation (RdDM) as well as histone methylation (Henderson and Jacobsen, 2007; Henderson et al., 2006; Lippman et al., 2003; Mette et al., 2000; Sijen et al., 2001b; Zilberman et al., 2003). siRNAs directing histone methylation are also found in *Tetrahymena thermophila*, and this methylation marks chromatin regions for programmed DNA elimination (Mochizuki et al., 2002; Taverna et al., 2002). Similarly, endogenously encoded siRNAs implicated in CDGS have also been sequenced in *S. pombe* (Reinhart and Bartel, 2002) where they participate in the formation of heterochromatin (see section 2.3.3.). Broadly speaking, siRNA-mediated CDGS proceeds through base-pairing mediated recruitment of Argonaute proteins to chromatin, which in turn guide histone-modifying enzymes to chromatin.

With the exception of *C. elegans*, it was long believed that siRNA-mediated silencing pathways in animals mainly function as self-defence mechanisms against invasive nucleic acids, such as viruses. For example flies in which RNAi has been selectively inactivated are hypersensitive to viruses (Ding and Voinnet, 2007; van Rij et al., 2006). This view has however been challenged by the recent discovery of endogenously encoded siRNAs in mice and *Drosophila melanogaster*. They arise from dsRNAs that are derived from transposable elements, probably through base-pairing of transcripts from bidirectional transcription or by transcription of long palindromic sequences (see also Figure 3). It is proposed that these endogenously encoded siRNAs are involved in the repression of transposon activity by degrading transposon RNAs via cleavage-dependent PTGS mechanisms. However, direct evidence for their involvement in the repression of transposition is still lacking (Chung et al., 2008; Czech et al., 2008; Ghildiyal et al., 2008; Kawamura et al., 2008; Okamura et al., 2008a; Okamura et al., 2008b; Tam et al., 2008; Watanabe et al., 2008). Furthermore, in contrast to the findings in plants, fungi and *Tetrahymena thermophila*, no satisfying evidence has so far been presented that would allow the conclusion that endogenously encoded siRNAs also mediate CDGS in mammals, and this topic remains a matter of debate.

2.2.6. piRNA-mediated silencing

piRNAs are a sRNA species of 24 – 30 nt in length that are exclusively bound by Argonaute proteins of the PIWI clade. They have so far only been discovered in animals and are mainly active in the germline (Aravin et al., 2006; Girard et al., 2006; Grivna et al., 2006; Lau et al., 2006; Saito et al., 2006; Vagin et al., 2006; Watanabe et al., 2006).

piRNAs are generated independently of Dicer by two distinct pathways. The primary piRNA pathway involves endonucleolytic cleavage of a long single-stranded RNA, which is transcribed from transposons and/or piRNA clusters, into primary piRNAs by an as yet unknown nuclease (Li et al., 2009; Malone et al., 2009), followed by or coupled to the association of piRNAs with Argonaute proteins. The primary piRNA pathway also initiates the secondary piRNA pathway, which proceeds through a ping-pong cycle. This cycle depends on at least two different Argonaute proteins and involves sense piRNA-directed cleavage and processing of a target RNA by Argonaute, producing anti-sense piRNAs, which again induce the Argonaute-mediated production of sense piRNAs (see also Figure 3). Alternatively, the single-stranded products of the unknown nuclease can also directly enter the ping-pong amplification cycle (Brennecke et al., 2007; Gunawardane et al., 2007).

The piRNA pathway is mainly active in the germline where it is proposed to be required to silence transposons. Whereas endogenously encoded siRNAs affect transposon RNA levels via PTGS, the situation is less clear for piRNAs. It has been shown that Argonaute-associated piRNAs can recognize and cleave transposon RNAs *in vitro* or promote their deadenylation (Brennecke et al., 2007; Gunawardane et al., 2007; Nishida et al., 2007; Rouget et al., 2010). It has also been suggested that piRNAs directly influence chromatin modifications, and mechanisms similar to CDGS in plants or *S. pombe* have been proposed (Aravin et al., 2007; Brower-Toland et al., 2007; Klattenhoff et al., 2009; Kuramochi-Miyagawa et al., 2008; Saito et al., 2010). Future studies will show whether this view holds true.

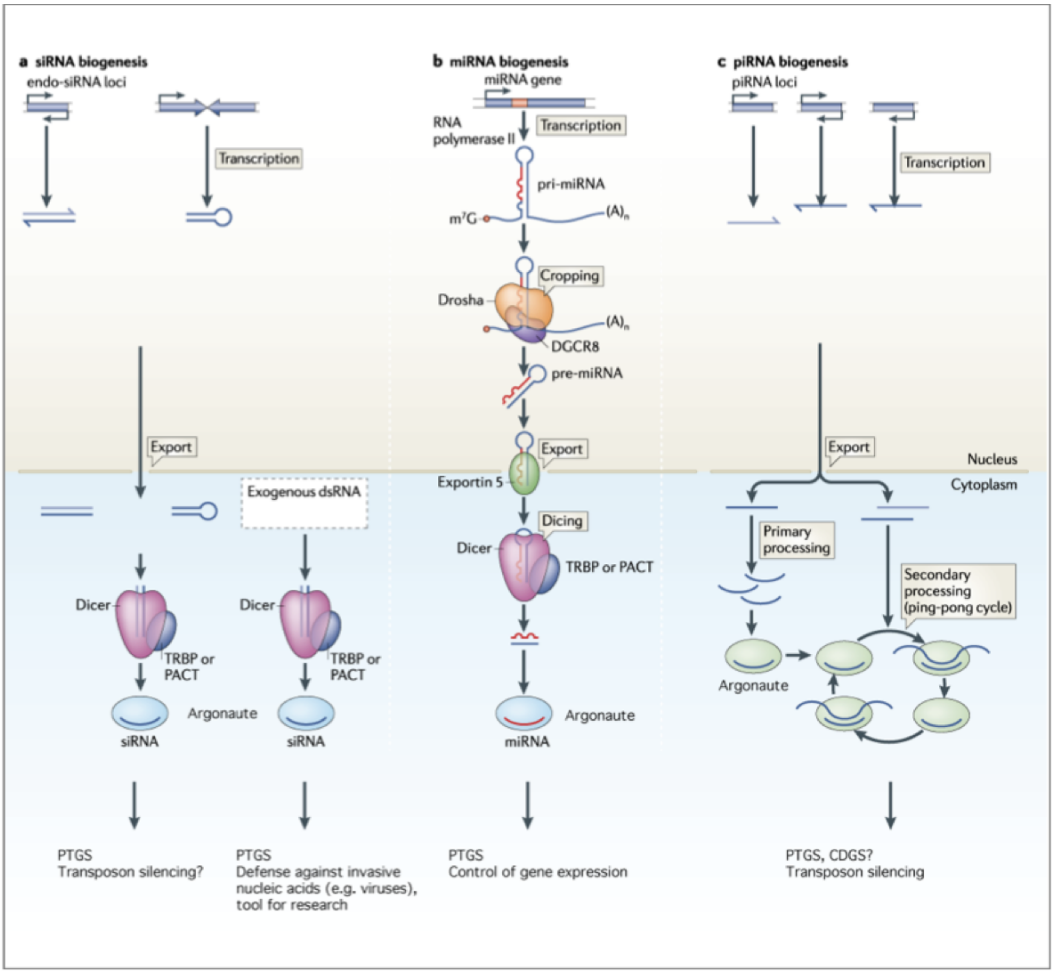


Figure 3: Overview of the three main RNAi pathways in mammals
a siRNA pathway: dsRNAs arising either from bidirectional transcription or transcription of long palindromic sequences and exogenously introduced dsRNAs are processed by Dicer enzymes into siRNAs, which are loaded onto Argonaute proteins and mediate PTGS. For details, see the main text (see section 2.2.5.). **b miRNA pathway:** RNA polymerase II transcription of miRNA genes generates hairpin-containing pri-miRNAs, which are processed into pre-miRNAs by the Drosha complex, exported to the nucleus and processed by Dicer enzymes into miRNAs. They are loaded onto Argonaute proteins and mediate PTGS. For details, see the main text (see section 2.2.4.). **c piRNA pathway:** Long single-stranded RNAs are processed by an unknown nuclease and are amplified via the ping-pong cycle. Alternatively, they are processed via Argonaute proteins into primary piRNAs, which are amplified via the ping-pong cycle. piRNAs are involved in transposon silencing via PTGS, and also CDGS mechanisms have been proposed. For details, see the main text (see section 2.2.6.).
Argonaute proteins of the Ago clade are shown in blue and those of the PIWI clade are shown in green. Modified from (Siomi et al., 2011).

2.3. siRNAs in *S. pombe* mediate heterochromatin formation

The concept of RNAi-mediated CDGS, which proceeds via sRNA-induced recruitment of histone-modifying enzymes to chromatin provides an explanation of how epigenetic marks are set at defined sites in the genome and how they are propagated. This concept has been extensively studied in *S. pombe* where we know in great detail how sRNAs are involved in mediating CDGS.

In this section I will first introduce the principle of higher order genome organization after which I will describe how RNAi contributes to the formation of heterochromatin in *S. pombe*.

2.3.1. Genomes are organized into different chromatin domains

Eukaryotic genomes assemble into a structure termed chromatin, which is required to handle the increased size of the genomes as compared to prokaryotes. Furthermore, chromatin can be chemically modified providing eukaryotic systems with an additional layer of genome regulation.

The fundamental building block of chromatin is the nucleosome, around which approximately 150 base-pairs (bp) of DNA is wrapped. Nucleosomes are organized into higher order structures, the chromatin fibres, which again assemble in a complex manner and form the chromosome arms. Nucleosomes themselves are composed of 4 different histone proteins, which are always present in two copies: H2A, H2B, H3, and H4 assemble into the globular nucleosome core structure, around which the DNA is wrapped. Importantly, the amino-terminal tails of histone proteins protrude out of the nucleosome core and are subjected to post translational modifications, such as acetylation, methylation, and many others. These modifications divide chromatin into euchromatic and heterochromatic regions. Euchromatic regions are usually associated with histone H4 acetylation and histone H3 methylation at lysine 4. The hallmark of heterochromatic domains is the methylation of histone H3 on lysine 9 (H3K9me) by histone methyltransferases. H3K9me recruits Heterochromatin Protein 1 (HP1), which in turn can recruit histone methyltransferases resulting in the spreading of heterochromatic domains. Generally speaking, euchromatic domains are usually associated with gene

expression whereas heterochromatic structures are linked to silencing, although this view does not always hold true (Felsenfeld and Groudine, 2003; Grewal and Jia, 2007; Jenuwein and Allis, 2001).

2.3.2. Heterochromatin formation in *S. pombe*

In *S. pombe,* heterochromatin is mainly associated with telomeres, the silent mating-type region, and the centromeres of the three chromosomes. A common feature shared by these loci is the presence of repetitive DNA elements, the *dg* and *dh* repeats, which are sufficient to promote heterochromatin formation and spreading into adjacent regions. Spreading can lead to silencing of nearby reporter genes, a process reminiscent of position effect variegation (PEV) initially described in *Drosophila melanogaster* (Cam et al., 2005; Grewal and Klar, 1997; Hall et al., 2002; Kanoh et al., 2005; Müller, 1930; Partridge et al., 2002).

Heterochromatin assembly involves an orchestrated array of conserved chromatin modification events: deacetylation of the histone H3 amino terminus by the class I and II histone deacetylases Clr3 and Clr6, as well as the class III NAD-dependent deacetylase Sir2, is followed by H3K9me catalyzed by a complex containing the methyltransferase Clr4. This creates a binding site for the chromodomain proteins Swi6, Chp1, and Chp2, which are homologous to mammalian HP1 (Bannister et al., 2001; Bjerling et al., 2002; Grewal et al., 1998; Motamedi et al., 2008; Nakayama et al., 2001; Partridge et al., 2000; Rea et al., 2000; Shankaranarayana et al., 2003; Sugiyama et al., 2007). H3K9me has been proposed to spread along the chromatin fibre through sequential cycles of methylation that are coupled to oligomerization of Swi6 (Canzio et al., 2011; Cowieson et al., 2000; Nakayama et al., 2001; Partridge et al., 2002; Sadaie et al., 2008). At the centromeres, heterochromatin formation also requires components of the RNAi pathway (Volpe et al., 2002).

2.3.3. RNAi-mediated heterochromatin formation at the centromeres

S. pombe contains a single gene each for Dicer (Dcr1), Argonaute (Ago1), and RNA-directed RNA polymerase (Rdp1). Deletion of any of these genes leads to a drastic reduction in H3K9me and Swi6 localization at the centromeres and is

accompanied by lagging chromosomes and chromosome loss during mitosis (Hall et al., 2003; Provost et al., 2002; Volpe et al., 2002). Moreover, siRNAs ranging in size from 20 - 26 nt that match to the *S. pombe* centromeric repeats have been found (Buhler et al., 2008; Cam et al., 2005; Reinhart and Bartel, 2002), which suggested a direct involvement of the RNAi pathway in formation of heterochromatin at the centromeres. Subsequent biochemical studies have revealed that the RNAi proteins reside in three main complexes, which are all essential for the formation of heterochromatin at centromeres. Ago1 assembles into the RNA-induced transcriptional silencing (RITS) complex, together with the chromodomain protein Chp1, Tas3, and a single-stranded siRNA (Verdel et al., 2004). Additionally, Ago1 is also found in a complex termed Argonaute siRNA chaperone (ARC) complex, together with two uncharacterized proteins, Arb1 and Arb2, and a double-stranded siRNA. It has been proposed that ARC functions as a chaperone, which receives the siRNA duplexes directly from Dcr1 and hands Ago1 over to RITS once the passenger siRNA strand is released (Buker et al., 2007). Rdp1 is found in a complex termed RNA-directed RNA polymerase complex (RDRC), together with two conserved proteins, Cid12, a member of the polyA polymerase family, and Hrr1, an RNA helicase. RDRC has RNA-directed RNA polymerase activity *in vitro*, and this activity is required for heterochromatin formation (Motamedi et al., 2004).

Subunits of both, RITS and RDRC, interact with centromeric transcripts as well as with each other. Importantly, they can also be crosslinked to chromatin. Notably, all these interactions depend on heterochromatin and RNAi proteins suggesting that they happen on chromatin and require a functional RNAi pathway. Furthermore, Dcr1 also associates with centromeric sequences and RITS can interact with the methyltransferase Clr4 (Bayne et al., 2010; Cam et al., 2005; Motamedi et al., 2004; Sugiyama et al., 2005; Woolcock et al., 2011). Intriguingly, two different subunits of RNA Pol II are essential for RNAi-mediated heterochromatin assembly in *S. pombe* demonstrating the importance of transcription for heterochromatin formation (Djupedal et al., 2005; Kato et al., 2005). These data, together with the observations that siRNA-mediated heterochromatin formation is usually *cis* restricted, and the slicer activity of Ago1 is important for heterochromatin formation (Buhler et al., 2006; Irvine et

al., 2006; Sigova et al., 2004), have led to the nascent transcript model for the formation of heterochromatin at centromeres (see Figure 4).

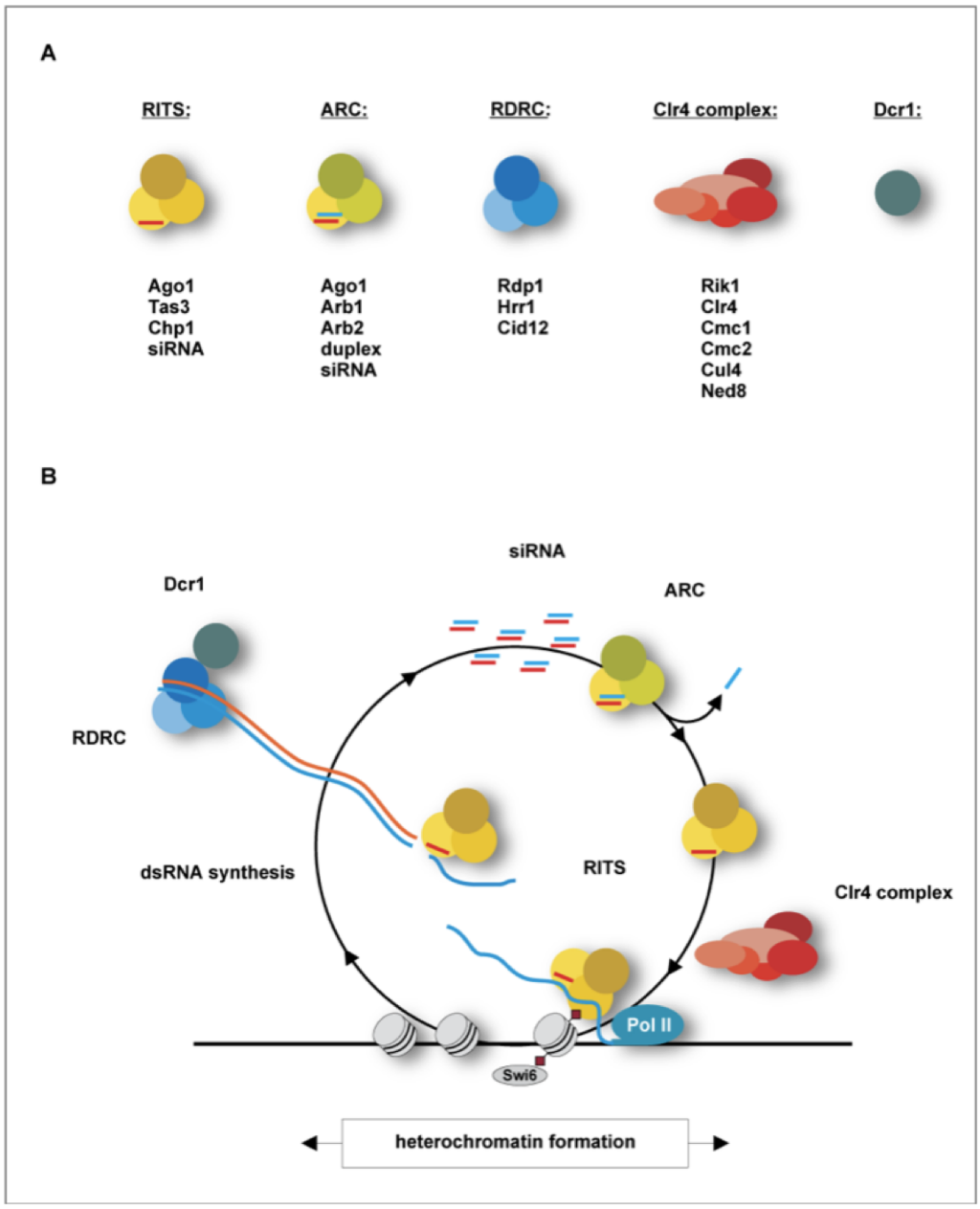


Figure 4: RNAi-mediated heterochromatin formation at the centromeres
A: Composition of the complexes involved in RNAi-mediated heterochromatin formation at the centromeres. For details, see the main text. The Clr4 complex will not be introduced in greater detail in this thesis. **B:** Model for RNAi-mediated centromeric heterochromatin formation. The RITS complex associates with nascent transcripts at centromeric repeats via siRNA-mediated base-pairing and by Chp1 binding to H3K9me. dsRNA is produced through subsequent RITS-mediated RDRC recruitment. These transcripts are rapidly processed by Dcr1 into siRNAs and loaded onto the ARC complex. The ARC complex cleaves off the passenger strand followed by loading of the guide strand onto RITS. Together with recruitment of the Clr4 complex to centromeric heterochromatin, this leads to a self-amplification loop resulting in RNAi-mediated heterochromatin formation at centromeres.

This model is also supported by experiments demonstrating that artificial tethering of RITS to a euchromatic gene can initiate heterochromatin formation at this site and leads to the production of siRNAs matching the targeted gene. This depends both on RNAi and heterochromatin components (Buhler et al., 2006). Accordingly, tethering of Clr4 to a euchromatic locus leads to RNAi-independent assembly of heterochromatin, further corroborating the model that siRNAs are involved in targeting chromatin-modifying enzymes to the centromeres (Kagansky et al., 2009).

2.3.4. Heterochromatin formation at the mating-type region and the telomeres

Whereas RNAi is absolutely essential for the assembly and maintenance of heterochromatin at the centromeres, this is not the case for the silent mating-type region and the telomeres.

At the silent mating-type region, RNAi as well as the transcription factors Atf1 and Pcr1, which bind via DNA elements specifically to the silent mating-type region, act redundantly to maintain heterochromatin. Deleting either the transcription factors or RNAi proteins does not greatly interfere with H3K9me or Swi6 binding. H3K9me is only lost when both, the transcription factors and RNAi proteins, are deleted. However, the RNAi machinery as well as the transcription factors are essential for the *de novo* assembly of heterochromatin and its subsequent spreading across the silent mating-type region: When H3K9me marks are erased by using chemical compounds or by deleting Clr4, heterochromatin establishment and spreading after the reintroduction of Clr4 or after the removal of the compounds only occurs when both, RNAi proteins and the transcription factors, are present (Hall et al., 2002; Jia et al., 2004; Kim et al., 2004; Yamada et al., 2005).

At telomeres, RNAi is required for the formation of certain heterochromatic structures. Whereas the telomeric repeats themselves are associated with the telomeric repeat binding protein Taz1 and recruit additional telomere-specific proteins, H3K9me and Swi6 are not present. At subtelomeric sequences however, the situation seems to be different: here, Taz1, localized to

the telomeric repeats, is required to promote H3K9me and Swi6 recruitment. Furthermore, at certain *dg*- and *dh*-containing repeats in the subtelomeric sequences, Taz1 and RNAi act redundantly to recruit Swi6 similar to the situation at the silent mating-type region. However, this process is not studied in great detail and additional, as yet unidentified factors might also be involved (Cam et al., 2005; Hansen et al., 2006; Kanoh et al., 2005).

2.3.5. Heterochromatin and transcription: co-transcriptional gene silencing

Based on early work showing that heterochromatin is less accessible to the recombination machinery and to the action of enzymatic probes, such as a bacterial DNA methyltransferase or restriction endonucleases (Gottschling, 1992; Loo and Rine, 1994; Singh and Klar, 1992), the generally accepted view has been that heterochromatin is an inaccessible structure that prevents transcription of its underlying DNA sequence. The nascent transcript model now demonstrates that heterochromatin formation and maintenance is intrinsically dependent on transcription and suggests that heterochromatic structures are not, or at least not exclusively, repressed via transcriptional gene silencing (TGS). Accordingly, deletion of components of the SHREC complex, which, via deacetylation of histone H3, impairs access of RNA Pol II to heterochromatin and thus limits transcription (Bjerling et al., 2002; Sugiyama et al., 2007), leads only to a partial derepression of heterochromatic sequences at centromeres (Motamedi et al., 2008). Moreover, at *S. pombe* centromeric heterochromatin, centromeric siRNA levels increase upon deletion of SHREC components, suggesting that RNAi is compensating for the loss of TGS at centromeric repeats in these mutants (Sugiyama et al., 2007). Furthermore, RNAi or heterochromatin mutants of strains, in which a transgene inserted within centromeric heterochromatin is derepressed, do not show changes in RNA Pol II occupancy at the transgene locus (Buhler et al., 2006). Therefore, RNA degradation by the RNAi pathway seems to play a prominent role in backing up and supporting TGS (Buhler et al., 2007; Noma et al., 2004; Volpe et al., 2002). Unlike PTGS however, where siRNA-loaded RISC acts *in trans* to induce the degradation of cytoplasmic target RNAs (Elbashir et al., 2001b), RNAi-mediated heterochromatin formation

is *cis* restricted (Buhler et al., 2006). These observations have lead to the concept of co-transcriptional gene silencing (CTGS), where transcription by RNA Pol II is coupled to the RNAi-mediated degradation of the corresponding transcript (Buhler et al., 2006) and thereby assists TGS in keeping heterochromatin silent. Moreover, the findings that RNAi can silence retrotransposon long terminal repeats via CTGS independently of H3K9me demonstrate that this concept is not limited to heterochromatic domains but could be a general RNA-decay pathway happening on chromatin (Woolcock et al., 2011) (see section 2.3.6.).

2.3.6. Other roles for RNAi in *S. pombe*

Besides siRNAs matching to the centromeres, the telomeres and the mating-type region, siRNAs aligning to rDNA repeats have also been sequenced in *S. pombe* (Buhler et al., 2008; Cam et al., 2005). Furthermore, RITS binds to rDNA repeats and deleting RNAi or heterochromatin genes results in loss of H3K9me, loss of Swi6 localization, and expression of a reporter gene integrated at the rDNA repeats. These effects are accompanied by a marked increase in mitotic recombination of the repeats (Cam et al., 2005), indicating that RNAi is involved in rDNA silencing.

Furthermore, Dcr1 and Rdp1 also associate with retrotransposon long terminal repeats, independently of H3K9me. Localization of Dcr1 as well as Rdp1 to the long terminal repeats has been shown to correspond with silencing of these sequences, since the long terminal repeats are upregulated in cells having RNAi genes deleted (Woolcock et al., 2011).

Moreover, RNAi-dependent transient heterochromatin is formed at convergent gene pairs during G1/S: readthrough transcription of the convergent gene pairs produces dsRNAs, which leads to the RNAi-mediated establishment of heterochromatin in G1/S. This results in the recruitment of Cohesin between these convergent gene pairs in G2 and to the loss of readthrough transcription (Gullerova and Proudfoot, 2008).

Together, these data show that RNAi in *S. pombe* fulfils a variety of tasks, which are however all related to CTGS and heterochromatin formation. Surprisingly, RNAi-mediated PTGS very likely does not occur in *S. pombe* since

the introduction of dsRNA hairpins corresponding to a GFP transgene has only a minor influence on GFP expression (Sigova et al., 2004), and since siRNAs are not capable of acting *in trans* (Buhler et al., 2006).

2.3.7. Major questions related to RNAi-mediated formation of centromeric heterochromatin

Many things are still unclear about the functions and mechanisms of RNAi in *S. pombe*, especially the roles of RNAi at telomeres, at convergent gene pairs, and also other loci, such as the centromeres. However, since my PhD thesis focuses on RNAi-mediated heterochromatin formation at centromeres, I will introduce in the following only the main open questions concerning this process.

2.3.7.1. What is the event that triggers heterochromatin assembly at the centromeres?

The formation of heterochromatin at the centromeres involves an RNAi-mediated self-amplification loop. However, it is unclear what events trigger this loop. Two models have been proposed in which either low levels of H3K9me or primary siRNAs trigger heterochromatin assembly at centromeres.

On the one hand, it has been shown that cells in which Ago1 is separated from RITS can still maintain heterochromatin at the centromeres. However, they fail to establish heterochromatin *de-novo* (Partridge et al., 2007). Furthermore, when the chromodomain protein Chp1 is separated from RITS, cells completely lose centromeric heterochromatin (Debeauchamp et al., 2008). Moreover, high affinity binding of Chp1 to H3K9me is critical for the establishment but not the maintenance of heterochromatin at centromeres (Schalch et al., 2009). Together, these data suggest that low levels of H3K9me are present at the centromeres, which initially leads to RNAi-independent recruitment of RITS via Chp1. This interaction recruits Dcr1 and RDRC resulting in the RNAi-mediated self-amplification loop, which is required to fully assemble and spread heterochromatin at centromeres.

On the other hand, it has been shown that so-called primary siRNAs, which are generated independently of Dcr1, are loaded onto Ago1 and

subsequently recruit RITS to the centromeres. This would trigger heterochromatin assembly by recruiting Clr4, leading to low levels of H3K9me, which result in RNAi-mediated formation of heterochromatin (Halic and Moazed, 2010). In favour of this model, it has been demonstrated that H3K9me is completely abolished in Ago1 mutants, whereas some residual H3K9me still persists in Dcr1 or Rdp1 mutants. However, these findings could not be confirmed (Shanker et al., 2010). Furthermore, it has also been demonstrated that Dcr1 can associate independently of H3K9me with chromatin, but this does not trigger heterochromatin formation (Woolcock et al., 2011), questioning the model of primary siRNA-mediated heterochromatin formation.

2.3.7.2. What is the implication of G1/S-specific heterochromatin formation?

Recently, it has been proposed that RNAi-mediated heterochromatin formation occurs in G1/S phase of the cell cycle. Heterochromatic repeats at centromeres are transcribed by RNA Pol II during G1/S leading to the formation of dsRNAs, Dcr1-mediated generation of siRNAs, and heterochromatin establishment in G1/S, which persists in G2. During mitosis, phosphorylation of serine 10 on histone H3 results in loss of Swi6 binding, which again leads to transcription at centromeric repeats during G1/S. This suggests that heterochromatic regions are only transcribed during heterochromatin formation but transcription is not maintained once heterochromatin is established in G2, and CTGS is not active. However, none of these studies have shown that RNAi is not required for heterochromatin maintenance specifically in G2 (Chen et al., 2008; Kloc et al., 2008).

2.3.7.3. How does Dcr1 localize and with which proteins does it interact?

Ago1 is an integral part of the RITS as well as of the ARC complex, and Rdp1 is part of RDRC. For Dcr1 however, only a weak interaction with RITS and RDRC has been shown by co-immunoprecipitation experiments (Colmenares et al., 2007). Despite extensive biochemical work, no Dcr1-specific complex has been identified in *S. pombe*. On the other hand, experiments in other organisms have shown that Dicer enzymes interact with a variety of proteins, the most

important being dsRBD-containing proteins, which seem to be required for substrate specificity of Dicer enzymes. Human Dicer for example is associated with TRBP and with PACT-1 whose depletions impair RNAi-mediated silencing and interfere with the accumulation of mature miRNAs, respectively (Chendrimada et al., 2005; Haase et al., 2005; Lee et al., 2006). It also been shown that Dicer/PACT-1/TRBP complexes interact with Argonaute proteins and this interaction is thought to be required for the transfer of the sRNAs to RISC (Chendrimada et al., 2005; Hock et al., 2007; Lee et al., 2006; Tahbaz et al., 2004). In *Drosophila melanogaster*, Loquacious and R2D2 are the dsRBD-containing proteins associated with Dicer-1 and Dicer-2, respectively. The interaction of Dicer-1 with Loquacious is required to specifically target Dicer-1 to the miRNA pathway, whereas the association of Dicer-2 with R2D2 directs strand-specific incorporation of siRNAs into RISC (Forstemann et al., 2005; Liu et al., 2003; Saito et al., 2005). Surprisingly, it has been shown that Dicer-2 can also interact with Loquacious and this interaction is important for the accumulation of endogenously encoded siRNAs (Chung et al., 2008; Czech et al., 2008; Okamura et al., 2008a; Okamura et al., 2008b). A recent study in *C. elegans* pursuing a comprehensive analysis of Dicer interactors has identified a total of 108 candidate proteins, many of which are required for correct maturation of the diverse sRNA species found in *C. elegans* (Duchaine et al., 2006). Together, these data show that Dicer enzymes are generally found in complex(es) with other proteins and raise the question as to whether a hitherto unidentified complex would also exist for *S. pombe* Dcr1.

Additionally, the localization of Dcr1 in *S. pombe* has been a matter of debate. When I started my thesis project, only RITS and RDRC components could be crosslinked to the centromeres, but not Dcr1 (Verdel et al., 2004; Volpe et al., 2002). Furthermore, immunofluorescence microscopy has shown that the RITS components localize in nuclear foci whereas RDRC components localize diffusely in the nucleus (Motamedi et al., 2004; Noma et al., 2004). However, for Dcr1 it has been proposed that the enzyme localizes to the cytoplasm (Carmichael et al., 2006), which is difficult to reconcile with the current model of RNAi-mediated heterochromatin assembly.

2.4. Aim of this thesis

The concept of CTGS predicts that Dcr1 localizes to the nucleus and physically associates with chromatin. Inconsistent with this prediction fluorescence microscopy has shown that Dcr1 localizes to the cytoplasm. I set out to untangle these conflicting observations by employing yeast genetics, biochemical and proteomics approaches.

3. Results

3.1. Manuscript I: The C-terminus of Dcr1 participates in the spatial organization of nuclear RNAi

Published in Developmental Cell, January 2010, (see appendix).

3.1.1. Summary

Our knowledge about the biochemical properties of the fission yeast RNAi pathway is rather advanced. However, we are still lacking in depth insights into the spatial organization of RNAi in *S. pombe*. To gain a more detailed understanding of this, I have analyzed the subcellular localization of Dcr1. I demonstrate that Dcr1, in contrast to human Dicer, is a predominantly nuclear protein, which is enriched at the nuclear periphery in association with pores. Importantly, I find that Dicers in general might shuttle between the nuclear and cytoplasmic compartment and that this is mediated by their dsRBDs. In *S. pombe*, this shuttling is under the control of a yeast-specific motif, which I refer to as C33. In the absence of C33, Dcr1 accumulates in the cytoplasm, the formation of centromeric heterochromatin is impaired, and the protein acts promiscuously. In accordance with its clearly nuclear function in *S. pombe*, nuclear retention of Dcr1 is thus a prerequisite for proper assembly of heterochromatin at centromeric repeats (see also Figure 5).

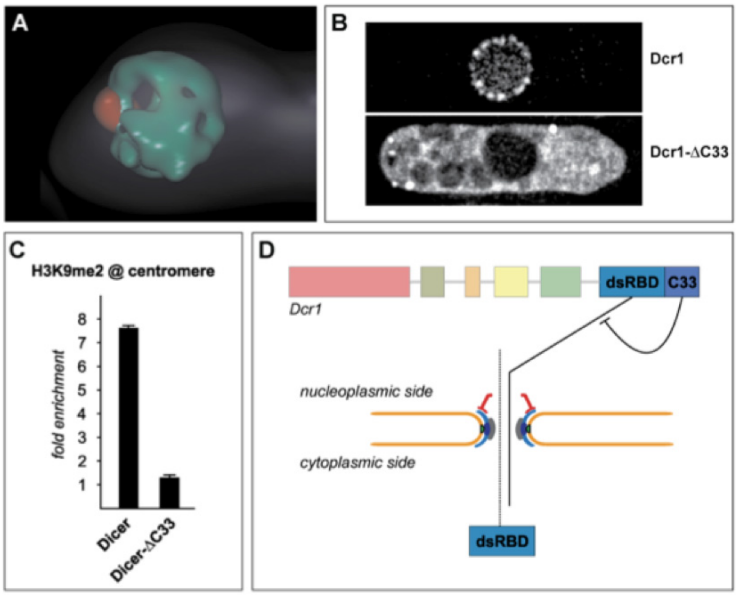


Figure 5: Manuscript I
A: Imaris surface rendering of live fission yeast cells shows that Dcr1 accumulates at the nuclear periphery (green, Dcr1; red, heterochromatin). **B:** Nuclear accumulation of Dcr1 is lost upon deletion of the C33 motif. **C:** ChIP experiment demonstrating that proper heterochromatin assembly depends on nuclear accumulation of Dcr1. **D:** Schematic illustration of the regulation of nucleo-cytoplasmic shuttling. The export-promoting activity of the dsRBD is inhibited by C33.

3.2. Manuscript II: Solution structure of the Dcr1 C-terminus

In press, The EMBO Journal (see appendix for the manuscript).

3.2.1. Summary

I have shown that the dsRBD of Dcr1 can mediate nucleo-cytoplasmic shuttling of Dcr1. To gain mechanistic insights into this process, I have collaborated with Pierre Barraud and Frédéric Allain from the ETH Zürich to solve the solution structure of the C-terminus of Dcr1, which contains the dsRBD. The structure reveals an unusual dsRBD fold embedding a novel zinc-binding motif that is conserved among Dicers in yeast, many of which are pathogenic. The zinc-binding motif is required for proper folding of the dsRBD, which creates an exposed protein-protein interaction surface that most likely binds to a nuclear retention factor to prevent nuclear export mediated by the dsRBD. Mutating the zinc-binding motif renders Dcr1 mainly cytoplasmic and leads to promiscuous activity of the protein, establishing the motif as a potential target for therapeutic intervention in fungal diseases. Although the dsRBD does bind to RNA, I show that this property is dispensable for the formation of centromeric heterochromatin. This further supports my proposal that the dsRBDs of Dicer proteins in general are not required for RNA binding but could be involved in controlling the subcellular localization of Dicer enzymes (see also Figure 6).

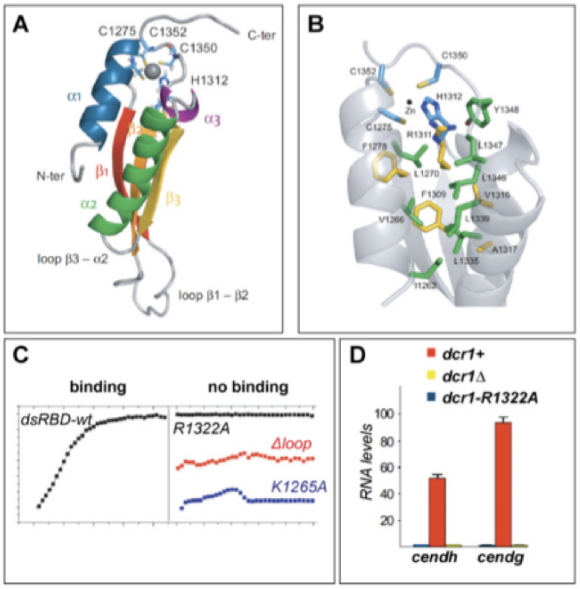


Figure 6: Manuscript II
A: Solution structure of the C-terminus of Dcr1 revealing an unusual dsRBD fold. **B:** The dsRBD contains a novel zinc-binding motif. **C:** ITC experiments demonstrating that the indicated mutants of the dsRBD do not bind to dsRNA *in vitro*. **D:** dsRNA binding of the dsRBD is dispensable for the formation of centromeric heterochromatin.

3.3. Manuscript III: Mapping and purifying the perinuclear localization domain of Dcr1

Unpublished results.

3.3.1. Introduction

Whereas Dicer enzymes in animals mainly localize to the cytoplasm, I have shown that *S. pombe* Dcr1 localizes to the nucleus and that nuclear localization is essential for the formation of centromeric heterochromatin. Moreover, I have found that Dcr1 localizes to the nuclear periphery (Figure 7A and manuscript I, (Emmerth et al., 2010)). However, it is not known which domain of Dcr1 recruits the enzyme to the nuclear periphery nor are the proteins, which do so. Therefore, I set out to map this domain and to purify it in order to identify proteins that would recruit Dcr1 to the nuclear periphery. These results would allow me to address the question as to whether association of Dcr1 with the nuclear periphery is required for the formation of centromeric heterochromatin.

3.3.2. The Platform domain is necessary and sufficient to recruit Dcr1 to the nuclear periphery

To identify the domain that recruits Dcr1 to the nuclear periphery, I cloned three different fragments of Dcr1 C-terminally to a GFP-LacZ reporter as described previously (manuscript I, (Emmerth et al., 2010)) and analyzed the localization of these fusion proteins by fluorescence microscopy. The three fragments encoded for the ATPase/helicase domain, the DUF283/PAZ-like domains, and the two RNase III domains. I did not include the C-terminus of Dcr1, which contains the dsRBD and C33 in these mapping experiments since I have previously shown that these sequences do not recruit GFP-LacZ to the nuclear periphery (manuscript I, (Emmerth et al., 2010)). Analysis of the fusion proteins by fluorescence microscopy revealed that only the fragment encoding the DUF283/PAZ-like domains recruited GFP-LacZ to the nuclear periphery (Figure 7B). Therefore, I divided this sequence into three fragments and cloned them again C-terminally to the GFP-LacZ reporter. This experiment showed that

a domain situated between DUF283 and PAZ-like specifically recruits the reporter to the nuclear periphery (Figure 7C). Since the domain shows sequence similarities to the Platform domain of *Giardia intestinalis*, which is proposed to mediate contacts with the posphodiester backbone of the dsRNA (Macrae et al., 2006), I will commonly refer to it as the Platform domain. In order to find residues in the Platform domain that are responsible for peripheral association, I subcloned the sequence by dividing it into two fragments of equal length C-terminally to the GFP-LacZ reporter. Surprisingly, both sequences recruited GFP-LacZ to the nuclear periphery (Figure 8A). However, when I divided the Platform into 4 fragments of equal length and cloned these C-terminally to the GFP-LacZ reporter, none of the sequences resulted in peripheral association of the reporter (Figure 8B).

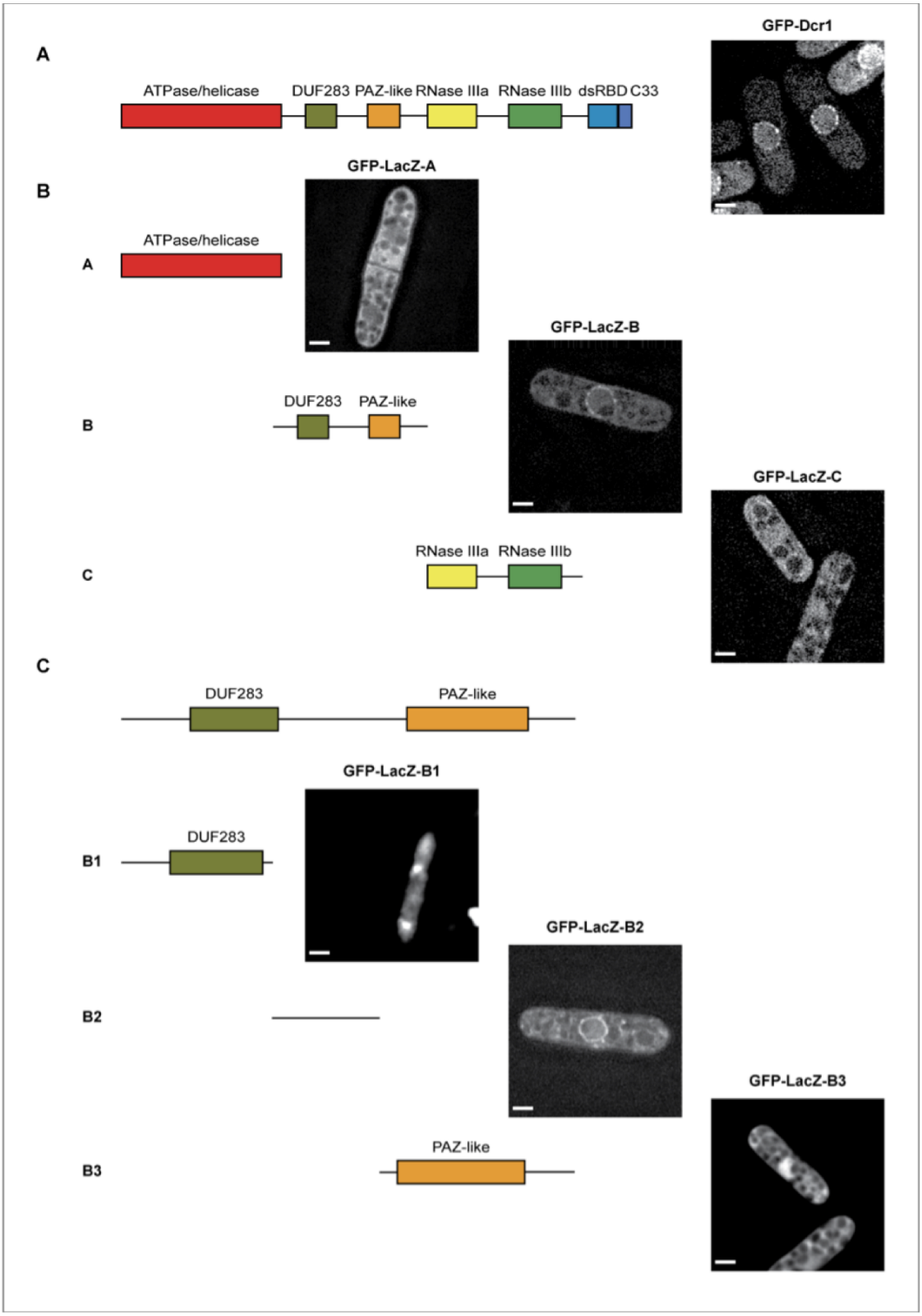


Figure 7: Mapping of the domain required to recruit Dcr1 to the nuclear periphery
A: Domain organization of Dcr1 and live-cell imaging of GFP-Dcr1 expressing cells. **B:** Live-cell imaging of wild-type strains transformed with the indicated plasmids. **C:** Live-cell imaging of wild-type strains transformed with the indicated plasmids. Scale bars = 2 μm.

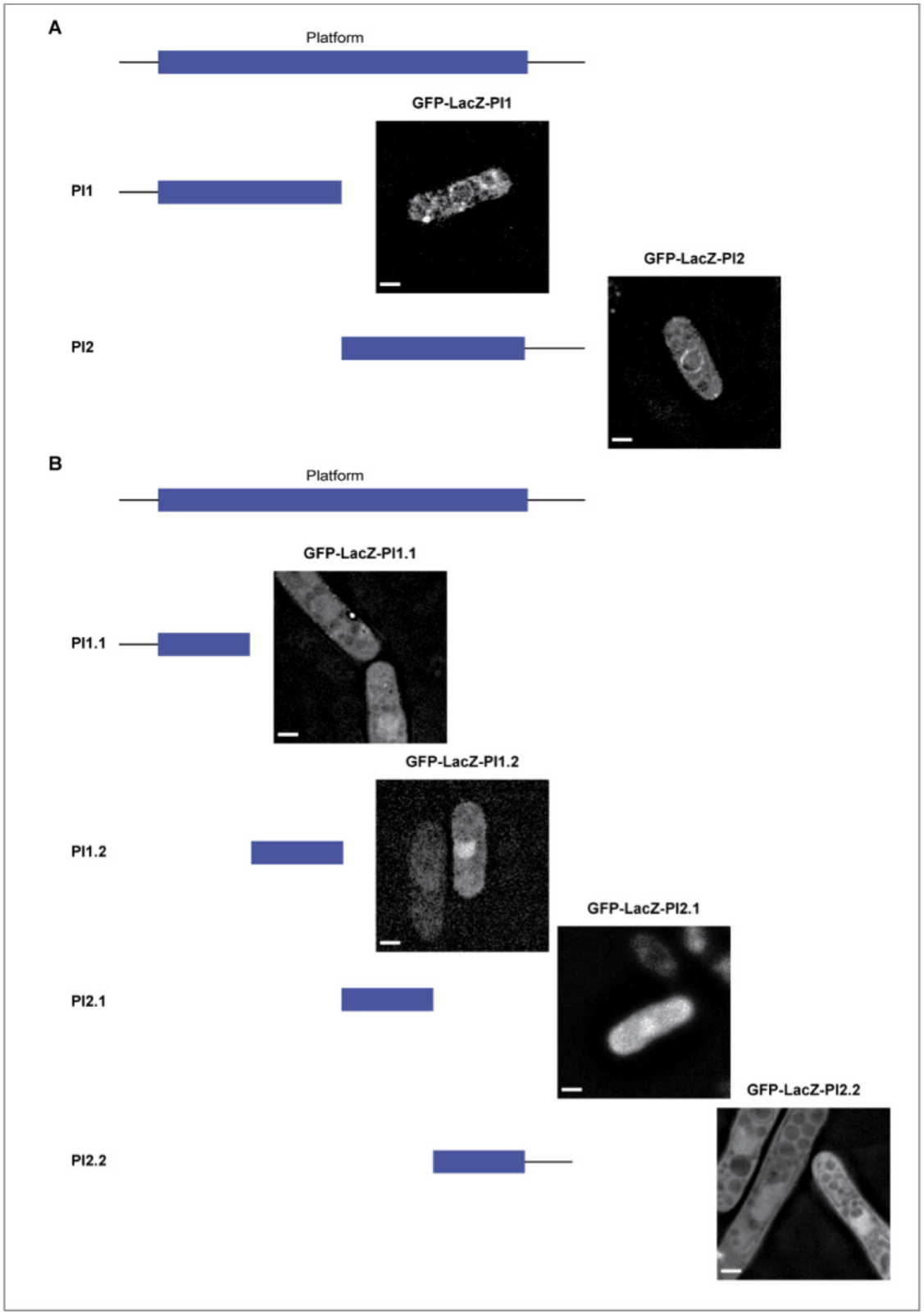


Figure 8: Subcloning the Platform domain C-terminally to a GFP-LacZ reporter
A: Live-cell imaging of wild-type strains transformed with the indicated plasmids. **B:** Live-cell imaging of wild-type strains transformed with the indicated plasmids. Scale bars = 2 μm.

Together, these results demonstrate that the Platform domain is sufficient to recruit a GFP-LacZ reporter to the nuclear periphery. To test whether it would be necessary and sufficient for the association of Dcr1 with the nuclear periphery, I deleted the Platform domain in Dcr1. However, the truncated protein was not imported into the nucleus. By adding a nuclear localization signal (nls), the truncated protein localized to the nucleus and peripheral association was lost. This shows that the Platform domain is both, necessary and sufficient, to recruit Dcr1 to the nuclear periphery (Figure 9A). However, western blot analysis revealed that the truncated proteins were unstable (Figure 9B). Therefore, I could not test whether association of Dcr1 with the nuclear periphery via its Platform domain is required for the formation of centromeric heterochromatin.

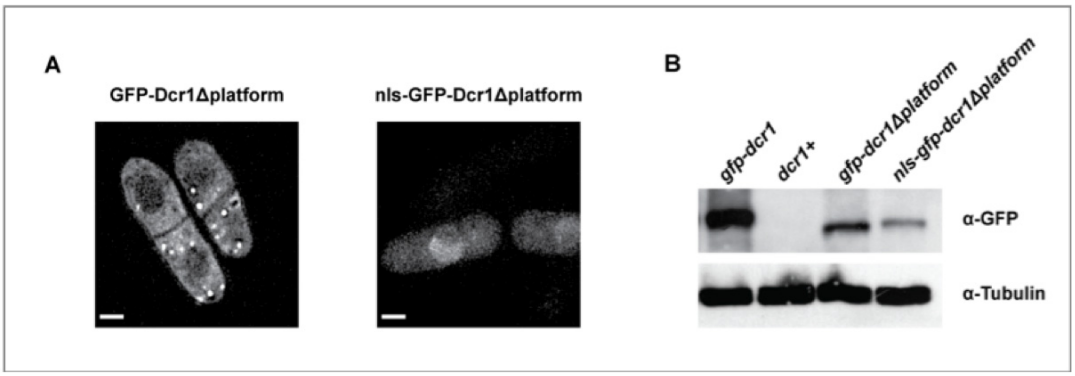


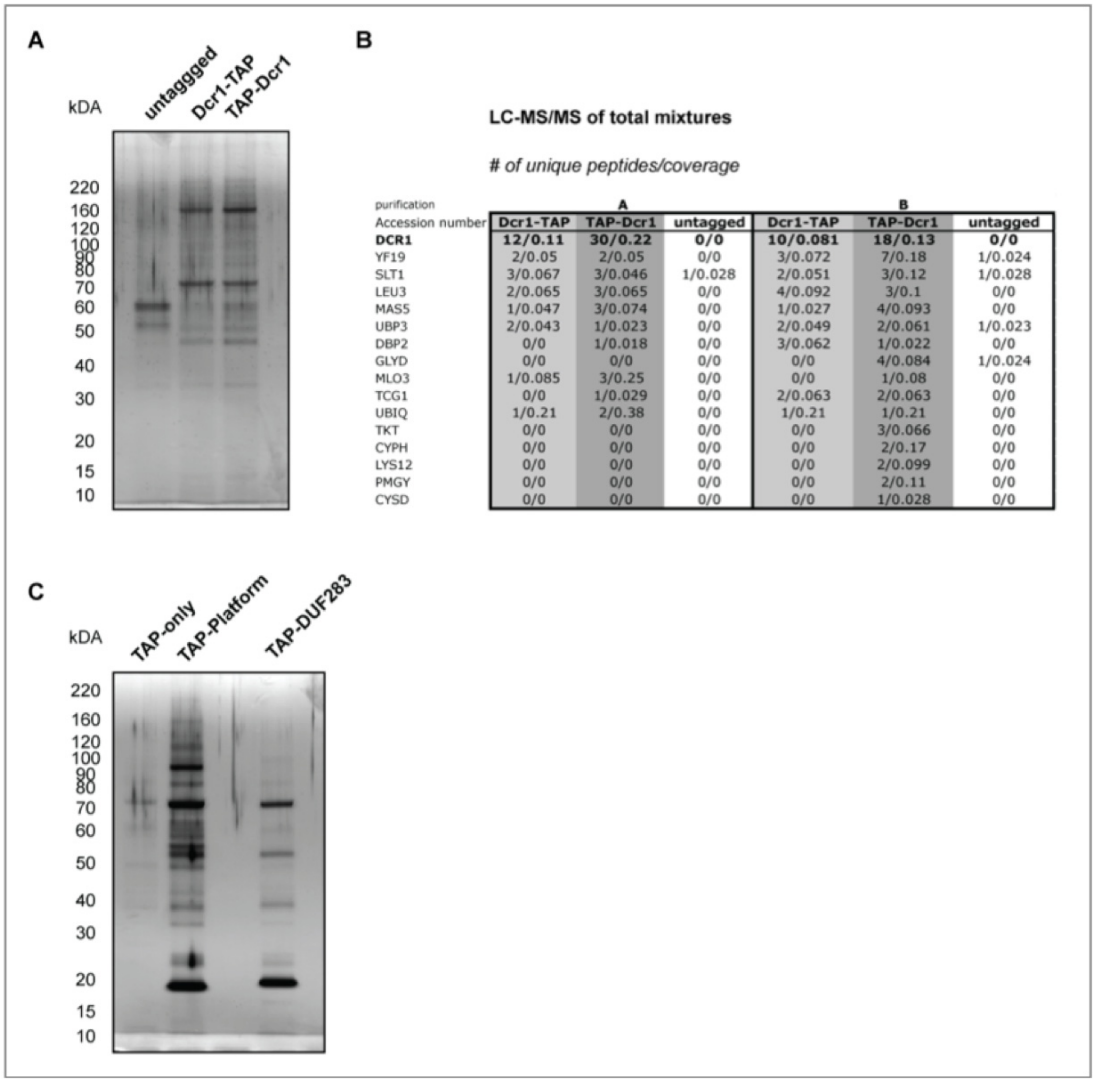
Figure 9: The platform domain recruits Dcr1 to the nuclear periphery
A: Live-cell imaging of cells expressing the indicated GFP-Dcr1 mutants. Scale bars = 2μm. **B:** Western blot analysis of strains expressing the indicated GFP-Dcr1 proteins. Tubulin was used as a loading control.

3.3.3. TAP-tag purification of the Platform domain

Since I was not able to generate a stable version of Dcr1 that loses association with the nuclear periphery, I sought to identify factors recruiting Dcr1 to the nuclear periphery. By deleting the corresponding genes, this would then allow me to address whether peripheral association of Dcr1 is important for centromeric silencing.

I generated strains expressing C- or N-terminally TAP-tagged Dcr1 from the endogenous promoter (Dcr1-TAP or TAP-Dcr1, respectively) and purified Dcr1 as described previously (Keller et al., 2010), with the following modifications: the lysis buffer was supplemented with digitonin to extract

membrane proteins (Prehm et al., 1996), and the cell lysate was spun down at low speed (Figure 10B, purification A). I also purified Dcr1 by additionally sonicating the cell lysate after resuspension in lysis buffer (Figure 10B, purification B) (see also section 3.5.3.). Analysis of both purification experiments by polyacrylamide gel electrophoresis (Figure 10A) and mass spectroscopy (Figure 10B) revealed that Dcr1 could be specifically purified. However, no proteins copurified with Dcr1 that are related to the nuclear periphery.



purification	A			B		
Accession number	Dcr1-TAP	TAP-Dcr1	untagged	Dcr1-TAP	TAP-Dcr1	untagged
DCR1	**12/0.11**	**30/0.22**	**0/0**	**10/0.081**	**18/0.13**	**0/0**
YF19	2/0.05	2/0.05	0/0	3/0.072	7/0.18	1/0.024
SLT1	3/0.067	3/0.046	1/0.028	2/0.051	3/0.12	1/0.028
LEU3	2/0.065	3/0.065	0/0	4/0.092	3/0.1	0/0
MAS5	1/0.047	3/0.074	0/0	1/0.027	4/0.093	0/0
UBP3	2/0.043	1/0.023	0/0	2/0.049	2/0.061	1/0.023
DBP2	0/0	1/0.018	0/0	3/0.062	1/0.022	0/0
GLYD	0/0	0/0	0/0	0/0	4/0.084	1/0.024
MLO3	1/0.085	3/0.25	0/0	0/0	1/0.08	0/0
TCG1	0/0	1/0.029	0/0	2/0.063	2/0.063	0/0
UBIQ	1/0.21	2/0.38	0/0	1/0.21	1/0.21	0/0
TKT	0/0	0/0	0/0	0/0	3/0.066	0/0
CYPH	0/0	0/0	0/0	0/0	2/0.17	0/0
LYS12	0/0	0/0	0/0	0/0	2/0.099	0/0
PMGY	0/0	0/0	0/0	0/0	2/0.11	0/0
CYSD	0/0	0/0	0/0	0/0	1/0.028	0/0

Figure 10: TAP-tag purification to find proteins recruiting Dcr1 to the nuclear periphery
A: Silver-stained gel showing the purification of Dcr1-TAP, TAP-Dcr1, and of an untagged strain. **B:** Results of mass spectrometry sequencing of a mixture of proteins from the two TAP purifications (the protocol used for purification A and B varies slightly, see main text for details). The sequenced proteins are sorted according to the combined number of unique peptides found in the two TAP-tagged strains. Ribosomal proteins and proteins with a high number of peptides in the untagged samples are excluded. Coverage 1 equals 100% coverage **C:** Silver-stained gel showing the purification of TAP-only, TAP-Platform, and TAP-DUF283. See appendix for the results of mass spectrometry sequencing of the different purifications.

Next, I attempted to purify the Platform domain. Therefore, I amino-terminally TAP-tagged this domain and purified it expressed from a plasmid. To control for non-specific protein-protein interactions, I also purified DUF283. I performed the standard TAP-tag purification protocol as well as the modified protocols introduced above, also varying the salt concentrations and analyzed the experiments by polyacrylamide gel electrophoresis (Figure 10C) and mass spectrometry. However, analysis of the purified samples by mass spectrometry did not identify any proteins that specifically associate with the Platform domain, and which would be related to the nuclear periphery (see appendix for the list of proteins identified by mass spectrometry and the details of the six different purifications).

In summary, these purification experiments did not reveal proteins interacting with Dcr1 that would be related to the nuclear periphery.

3.3.4. Localization to the nuclear periphery is not a prerequisite for centromeric silencing

A close relative of *S. pombe* is the fission yeast *Schizosaccharomyces japonicus (S. japonicus)* (Sipiczki, 2000). The genome of this yeast has been recently sequenced (a link to the sequence can be found at: **http://www.broad.mit.edu/annotation).** Like *S. pombe*, *S. japonicus* contains single genes for Dicer, Argonaute, and RNA-directed RNA polymerase (*jdcr1*$^+$, *jago1*$^+$, and *jrdp1*$^+$, respectively). To test if jDcr1 also associates with the nuclear periphery, I N-terminally GFP-tagged it in *S. japonicus* and analyzed the localization by fluorescence microscopy. jDcr1 localized to the nucleus, however no association with the nuclear periphery was detected in *S. japonicus* (Figure 11A). Next, I replaced the genomic sequence of *dcr1*$^+$ in *S. pombe* with GFP-tagged jDcr1 and also fused to it C-terminally the Platform domain. As expected, only the construct with the Platform domain recruited jDcr1 to the nuclear periphery (Figure 11B). It is also worth mentioning that an nls was added, otherwise GFP-jDcr1 was not imported into the nucleus (data not shown). Importantly, GFP-jDcr1 could fully rescue centromeric silencing, suggesting that the localization of Dcr1 to the

nuclear periphery is not necessary for RNAi-mediated heterochromatin formation at the centromeres (Figure 11C).

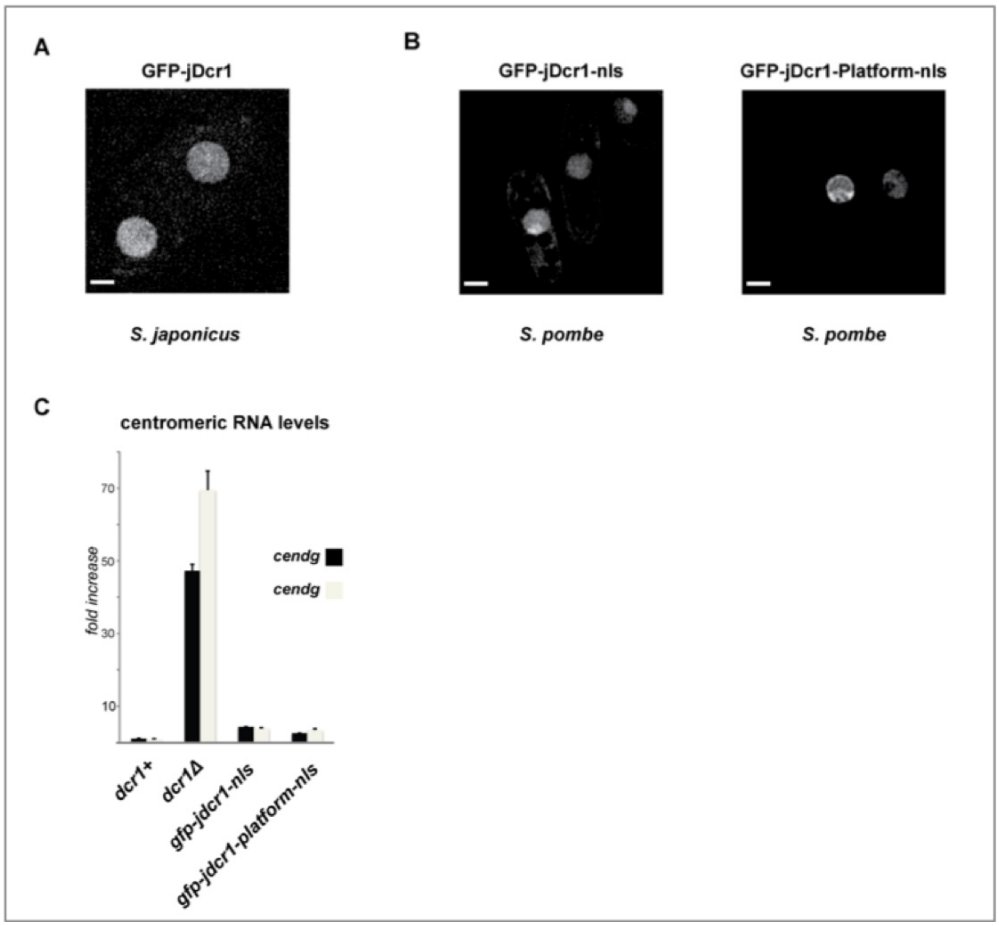


Figure 11: *S. japonicus* Dcr1 can silence centromeres in *S. pombe*
A: Live-cell imaging of *S. japonicus* strains expressing N-terminally GFP-tagged jDcr1. Scale bars = 2μm. **B:** Live-cell imaging of *S. pombe* strains expressing the indicated jDcr1 constructs. Scale bars = 2μm. **D:** Quantitative real-time RT-PCR measuring *cendg* and *cendh* RNA levels in the indicated strains. RNA levels were normalized to *act1+* RNA.

3.4. Manuscript IV: Regulation of Dcr1 localization

Unpublished results.

3.4.1. Introduction

I have shown that Dcr1 associates with the nuclear periphery and this association occurs via the Platform domain (manuscript I, (Emmerth et al., 2010) and manuscript III). However, purification of the Platform domain did not reveal any proteins that would link Dcr1 to the nuclear periphery. Furthermore, experiments where jDcr1 was expressed in *S. pombe* suggest that association of Dcr1 with the nuclear periphery is not required for the formation of centromeric heterochromatin (manuscript III). To address the physiological relevance of peripheral Dcr1 localization, I set out to compare different mutant strains for Dcr1 localization and to expose cells to different environmental stimuli to find conditions that would affect the localization of Dcr1.

3.4.2. Dcr1 localization does not change in RNAi or heterochromatin mutants

Since RNAi is involved in the generation of centromeric heterochromatin, I tested whether the localization pattern of Dcr1 would be affected in different RNAi and heterochromatin mutant backgrounds. Therefore, I examined the localization of GFP-Dcr1 in the RNAi deletion mutants *ago1*, *arb1*, and *rdp1* as well as in deletion mutants for the heterochromatin genes *clr4* and *swi6*. However, I could not detect a change in the localization pattern of GFP-Dcr1 in any of these mutant backgrounds (manuscript I, (Emmerth et al., 2010), and Figure 12).

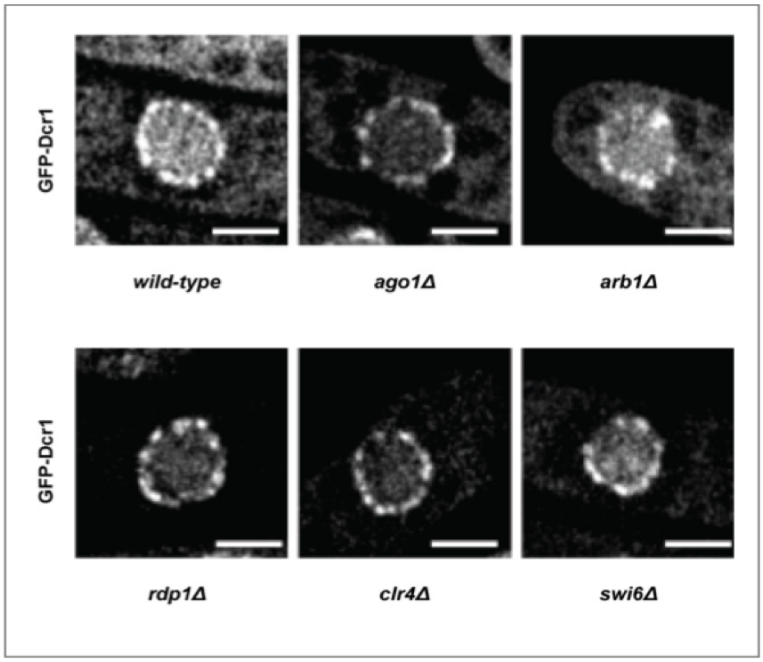


Figure 12: GPF-Dcr1 localization does not change in RNAi and heterochromatin mutants
Live-cell imaging of GFP-Dcr1 expressing cells in the indicated mutant strains. Scale bars = 2μm.

3.4.3. Dcr1 localization changes upon heat shock

Next, I tested whether different environmental conditions would influence the localization pattern of GFP-Dcr1. For this purpose, I grew GFP-Dcr1 expressing cells to high ODs or subjected them to stress, such as heat shock, cold shock, or exposure to H_2O_2. Whereas I could not detect any major changes in the localization pattern of GFP-Dcr1 in cells grown to high ODs, in cells that were exposed to H_2O_2, or in cells grown at 20°C (Figure 13 and R. Stunnenberg, personal communication), I found that growing cells at high temperatures led to the dissociation of Dcr1 from the nuclear periphery and to the export of the protein to the cytoplasm (Figure 13A).

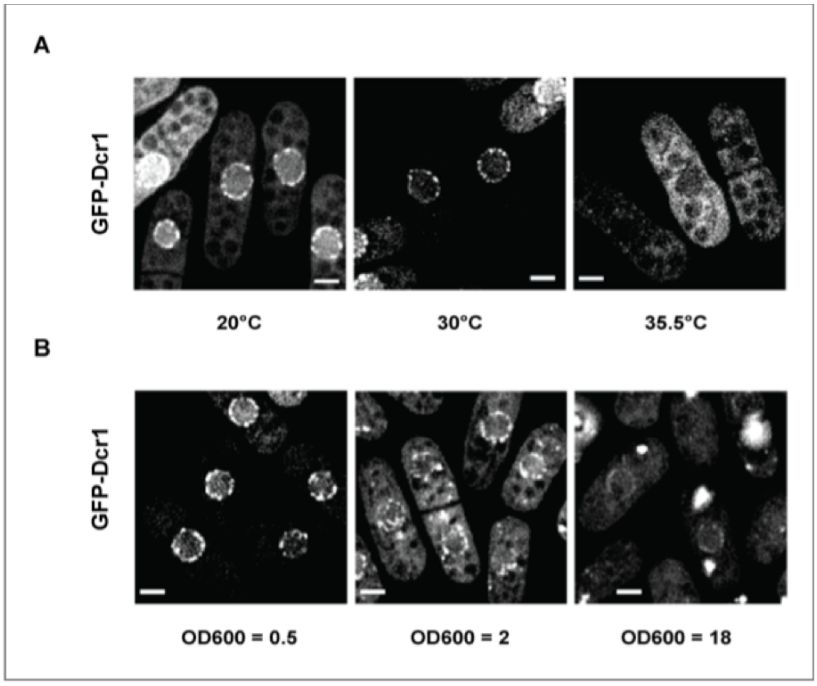


Figure 13: Dcr1 is exported upon heat shock
A: Live-cell imaging of GFP-Dcr1 expressing cells after exposure to different temperatures. Cells were grown overnight at the indicated temperatures. **B:** Live-cell imaging of GFP-Dcr1 expressing cells grown to higher ODs. Scale bars = 2μm.

3.4.4. Heat shock genes are upregulated in *dcr1Δ* cells

Others and I have shown that some RNA levels are different from wild-type in *dcr1Δ* cells (Emmerth et al., 2010; Provost et al., 2002). Strikingly, one of the few RNAs that are upregulated in *dcr1Δ* cells is the transcript corresponding to the heat shock response gene *hsp16* ((Provost et al., 2002) and Figure 14), which is specifically expressed upon heat shock (Taricani et al., 2001). Moreover, it is known that heat shock genes in *Drosophila melanogaster* localize to the nuclear periphery (Kurshakova et al., 2007). I therefore hypothesize that the association of Dcr1 with the nuclear periphery is required to silence heat shock genes. My finding that association of Dcr1 with the nuclear periphery is lost and the protein is exported in cells that are grown at 35.5°C is in agreement with this hypothesis, which however needs to be addressed in more detail by future experiments (see section 5.).

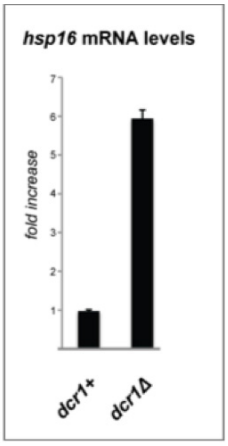


Figure 14: *Hsp16* mRNA levels are upregulated in *dcr1Δ* cells
Quantitative real-time RT-PCR measuring *hsp16* mRNA levels in the indicated strains. RNA levels were normalized to *act1+* RNA levels.

3.5. Material and Methods

Detailed bench protocols have been deposited in the Bühler laboratory protocol database and can be obtained from Marc Bühler upon request.

3.5.1. Manuscript I: The C-terminus of Dcr1 participates in the spatial organization of nuclear RNAi

See experimental procedures in the published manuscript found in the appendix.

3.5.2. Manuscript II: Solution structure of the Dcr1 C-terminus

See experimental procedures in the manuscript found in the appendix.

3.5.3. Manuscript III: Mapping and purifying the perinuclear localization domain of Dcr1

Strains and plasmids

Yeast strains were grown at 30°C in YES. If transformed with plasmids, strains were grown in EMMc-Leu medium for the purifications and EMMc-Leu medium supplemented with 5 µg/ml thiamine for the mapping experiments. All strains were constructed following PCR-based protocols (Bahler et al., 1998; Furuya and Niki, 2009). GFP-LacZ constructs were cloned by fusion PCR-based assembly of the inserts and subsequent PacI/NotI ligation into pMB428 (Emmerth et al., 2010). Plasmids for TAP-tag purifications were generated by AscI/NotI ligation of the PCR-amplified fragments into pREP1NT (pMB115). Constructs on plasmids and in yeast strains were confirmed by sequencing.

Protein purification and mass spectrometry

Protein purifications and mass spectrometry were performed as described previously (Keller et al., 2010), and where indicated the following modifications were included: 1% digitonin (D141, Sigma-Aldrich) was used instead of nonidet P-40 (NP-40), the cell lysate was subjected to sonication (3x30

seconds, 20% amplitude, 1 minute on ice in between, BRANSON Digital Sonifier) after resuspending in lysis buffer, and the cell lysate was spun down at 1800g for 10 minutes at 4°C instead of 4000g for 10 minutes at 4°C.

Yeast live-cell fluorescence microscopy

Yeast live-cell fluorescence microscopy was performed as described previously (Emmerth et al., 2010).

RNA isolation, cDNA synthesis and qPCR

RNA isolation, cDNA synthesis and qPCR were performed as described previously (Emmerth et al., 2010).

Strain table S. pombe

Strain	Genotype	Source
SPB80	*h+ leu1-32 ura4-D18 ori1 ade6-216 imr1R(Nco1)::ura4+*	*2*
SPB81	*h+ leu1-32 ura4-D18 ori1 ade6-216 imr1R(Nco1)::ura4+ dcr1Δ::TAP-Kan*	*2*
SPB147	*h+ leu1-32 ura4-D18 ori1 ade6-216 imr1R(Nco1)::ura4+ dcr1+::TAP-Kan*	*2*
SPB148	*h+ leu1-32 ura4-D18 ori1 ade6-216 imr1R(Nco1)::ura4+ Kan-500bp-native-promoter-TAP::dcr1+*	*2*
SPB278	*h+ leu1-32 ura4-D18 ori1 ade6-216 imr1R(NcoI)::ura4+ Kan-nmt1P3-GFP::dcr1+*	*1*
SPB599	*h+ leu1-32 ura4-D18 ori1 ade6-216 imr1R(Nco1)::ura4+ Kan-nmt1P3-GFP::dcr1Δplatform::NatR*	*1*
SBP621	*h+ leu1-32 ura4-D18 ori1 ade6-216 imr1R(Nco1)::ura4+ Hph-nmt1P3-nls-GFP::dcr1Δplatform::NatR*	*1*
SPB952	*h+ leu1-32 ura4-D18 ori1 ade6-216 imr1R(Nco1)::ura4+ dcr1Δ::Kan-nmt1P3-GFP-jdcr1-nls-NatR*	*1*
SPB954	*h+ leu1-32 ura4-D18 ori1 ade6-216 imr1R(Nco1)::ura4+ dcr1Δ::Kan-nmt1P3-GFP- jdcr1-platform-nls-Hph*	*1*

1 = Bühler lab strain collection, 2 = obtained from Danesh Moazed

Strain table S. japonicus

Strain	Genotype	Source
SJB11	*h+ matsj-2017 Kan-nmt1P3-GFP::dcr1*	*1*

1 = Bühler lab strain collection

Plasmid table

Name	Common Name	Source
pMB115	*pREPnmt1P3-TAP-Leu+*	*2*
pMB446	*pREPnmt1P3-GFP-LacZ-ATPase/helicase-Leu+*	*1*
pMB448	*pREPnmt1P3-GFP-LacZ-DUF283/Platform/PAZ-like-Leu+*	*1*
pMB450	*pREPnmt1P3-GFP-LacZ-RNase-III-domains-Leu+*	*1*
pMB467	*pREPnmt1P3-GFP-LacZ-DUF283-Leu+*	*1*
pMB468	*pREPnmt1P3-GFP-LacZ-Platform-Leu+*	*1*

pMB469	*pREPnmt1P3-GFP-LacZ-PAZ-like-Leu+*	*1*
pMB482	*pREPnmt1P3-GFP-LacZ-Platform-fragment-Pl1-Leu+*	*1*
pMB483	*pREPnmt1P3-GFP-LacZ-Platform-fragment-Pl2-Leu+*	*1*
pMB511	*pREPnmt1P3-GFP-LacZ-Platform-fragment-Pl1.1-Leu+*	*1*
pMB512	*pREPnmt1P3-GFP-LacZ-Platform-fragment-Pl1.2-Leu+*	*1*
pMB513	*pREPnmt1P3-GFP-LacZ-Platform-fragment-Pl2.1-Leu+*	*1*
pMB514	*pREPnmt1P3-GFP-LacZ-Platform-fragment-Pl2.2-Leu+*	*1*
pMB528	*pREPnmt1P3-TAP-Platform-Leu+*	*1*
pMB529	*pREPnmt1P3-TAP-DUF283-Leu+*	*1*

1 = Bühler lab plasmid collection, 2 = obtained from Danesh Moazed

3.5.4. Manuscript IV: Regulation of Dcr1 localization

Strains

Unless otherwise stated, yeast strains were grown at 30°C in YES. Heat and cold shock experiments were performed by growing cells overnight at the indicated temperatures. All strains were constructed following PCR-based protocols (Bahler et al., 1998).

Yeast live-cell fluorescence microscopy

Yeast live-cell fluorescence microscopy was performed as described previously (Emmerth et al., 2010).

RNA isolation, cDNA synthesis and qPCR

RNA isolation, cDNA synthesis and qPCR were performed as described previously (Emmerth et al., 2010).

Strain table S. pombe

Strain	Genotype	Source
SPB200	*h+ leu1-32 ura4-D18 ori ade6? imr1R(NcoI)::ura4+ Nat-nmt1P3-GFP::dcr1+ clr41Δ::TAP-Hph*	*1*
SPB207	*h+ leu1-32 ura4-D18 ori ade6-216 imr1R(NcoI)::ura4+ Nat-nmt1P3-GFP::dcr1+ arb1Δ::TAP-Kan*	*1*
SPB278	*h+ leu1-32 ura4-D18 ori ade6-216 imr1R(NcoI)::ura4+ Kan-nmt1P3-GFP::dcr1+*	*1*
SPB337	*h+ leu1-32 ura4-D18 ori ade6-216 imr1R(NcoI)::ura4+ Nat-nmt1P3-GFP::dcr1+ ago1Δ::TAP-Kan*	*1*
SPB338	*h+ leu1-32 ura4-D18 ori ade6-216 imr1R(NcoI)::ura4+ Nat-nmt1P3-GFP::dcr1+ rdp1Δ::TAP-Kan*	*1*
SPB442	*h+ leu1-32 ura4-D18 ori ade6-216 imr1R(NcoI)::ura4+ Nat-nmt1P3-GFP::dcr1+ swi6Δ::Nat*	*1*

1 = Bühler lab strain collection

4. Discussion

The biochemical properties of the fission yeast RNAi pathway have been characterized in great detail. However, we still lack in depth insights into the spatial organization of RNAi in *S. pombe*. Furthermore, the few studies conducted so far have resulted in data conflicting with biochemical results. In my PhD thesis I employed yeast genetics, biochemical and proteomics approaches to untangle these conflicting data and performed the first careful investigation of Dcr1 localization. I found that Dcr1 localizes to the nuclear periphery and this localization pattern depends on the Platform domain of the enzyme. Moreover, I demonstrated that nuclear localization of Dcr1 is essential for the formation of centromeric heterochromatin, and I provided insights into the physiological function of peripheral Dcr1, which could be involved in the heat shock response.

Furthermore, I found that nuclear accumulation depends on the C-terminus of Dcr1, which can mediate nucleo-cytoplasmic shuttling. Solving the solution structure of this region, which contains a canonical dsRBD, revealed an unusual dsRBD fold containing a novel zinc-binding motif that is conserved among Dicers in yeast. The zinc-binding motif is required for proper folding of the dsRBD, which creates an exposed protein-protein interaction surface that might bind to a nuclear retention factor to prevent nuclear export mediated by the dsRBD. Furthermore, I demonstrated that binding of this dsRBD to dsRNA is dispensable for the formation of centromeric heterochromatin.

Below, I will discuss the implications of my findings for RNAi in *S. pombe* and for RNAi in general.

4.1. Dcr1 localizes to the nucleus

Whereas RNAi pathways in animals are involved in classical cytoplasmic PTGS, it is not clear whether this is also true for siRNAs in *S. pombe* (Buhler et al., 2006; Sigova et al., 2004). Rather, it is believed that RNAi is functionally restricted to the nucleus, where it participates in the establishment and maintenance of heterochromatic domains (Buhler, 2009; Moazed, 2009). Therefore, findings where Dcr1 has been found to localize to the cytoplasm (Carmichael et al., 2006) are difficult to reconcile with the current model of RNAi-mediated heterochromatin formation at the centromeres: it would imply

that siRNAs are generated and loaded onto ARC in the cytoplasm. However, Ago1 localizes to chromatin and only Arb2, but not Arb1, of ARC is partially found in the cytoplasm (Buker et al., 2007; Verdel et al., 2004). Moreover, experiments where cytoplasmic localization of Dcr1 was detected were conducted under non-physiological conditions by massively overexpressing the protein (Carmichael et al., 2006). When expressed at more moderate levels, I observed a clear nuclear fraction, which is associated with the nuclear periphery. Moreover, I demonstrated that nuclear localization of Dcr1 is essential for heterochromatin formation, since mutations that render the protein cytoplasmic resulted in a complete loss of heterochromatin (manuscript I, (Emmerth et al., 2010)). These data establish that Dcr1 has to localize to the nucleus for proper heterochromatin formation and cytoplasmic maturation steps for siRNAs in *S. pombe* seem rather unlikely.

4.2. Compartmentalization – a way to provide substrate specificity for Dcr1 in *S. pombe*

The findings in my PhD thesis show that Dcr1 associates with a specific nuclear compartment, the nuclear periphery (manuscript I, (Emmerth et al., 2010), and manuscript III). Similarly, compartmentalization of RNAi pathways can also be found in animals and plants. In human and *Drosophila melanogaster*, Argonaute proteins have been found to localize to the cytoplasmic surface of multivesicular bodies (MVB), which stimulates the silencing effects of miRNAs and siRNAs (Gibbings et al., 2009; Lee et al., 2009). Additionally, human Argonaute proteins localize to cytoplasmic processing (P-)bodies, although this seems to be rather a consequence than a cause of miRNA-related silencing (Eulalio et al., 2007; Liu et al., 2005). Furthermore, piRNA pathway components localize to structures termed Nuage in *Drosophila melanogaster, C. elegans*, and other organisms, and these structures are required for piRNA activity as well as piRNA biogenesis (Batista et al., 2008; Harris and Macdonald, 2001; Li et al., 2009; Lim and Kai, 2007; Malone et al., 2009). Additionally, in *Arabidopsis thaliana*, nuclear Cajal- and AB-bodies contain proteins involved in RdDM,

whereas miRNA biogenesis in plants is carried out in dicing (D-)bodies (Fang and Spector, 2007; Li et al., 2008; Li et al., 2006; Pontes et al., 2006; Song et al., 2007).

Restricting processes to certain subnuclear compartments is a concept known for many cellular pathways. For example ribosome biogenesis is mainly carried out in the nucleolus where the ribosomes undergo a very sophisticated and tightly controlled maturation process. Briefly, in the nucleolus, the rDNA repeats give rise to the 35S rRNAs, which assemble together with ribosomal proteins, small nucleolar RNAs (snoRNAs), and non-ribosomal factors into the pre-90S ribosome. This is then separated into the pre-60S and the pre-40S ribosome intermediates, which are transported through the nucleoplasm and exported to the cytoplasm where they are processed into the mature 60S and 40S subunits (Fromont-Racine et al., 2003). Restricting the biogenesis of ribosomes mainly to the nucleolus fulfils two purposes. Firstly, factors involved in ribosome biogenesis and ribosome core components are brought together, creating an environment where all the necessary elements to assemble ribosomes are concentrated, thereby enhancing the efficiency of ribosome biogenesis. Secondly, ribosome biogenesis is physically separated from the site of translation, which prevents illegitimate translation events, such as nuclear translation or translation of mRNAs by incompletely assembled ribosomes.

Similarly, compartmentalization of RNAi creates an environment containing all the necessary components, which could enhance the efficiency of RNAi. Moreover, by bringing together RNAi proteins and their corresponding dsRNAs, the specificity of the RNAi response could also be controlled. Of note, in animals the specificity of the RNAi proteins to a subset of dsRNAs can additionally be affected by the dsRBD-containing cofactors of Dicer enzymes (see also below). For *S. pombe* Dcr1 on the other hand, I could not identify any dsRBD-containing cofactors associating with Dcr1 that could provide substrate specificity (manuscript III). Therefore, I speculate that localizing Dcr1 to the nuclear periphery, as well as to some other places (see below) is required to provide the enzyme with substrate specificity by only exposing it to a confined set of targets. This also prevents Dcr1 from processing unwanted substrates thereby avoiding eventual deleterious effects. Accordingly, massive overexpression of Dcr1 is toxic for *S. pombe* (manuscript I, (Emmerth et al.,

2010)) and it would be interesting to assess the fitness of *S. pombe* cells expressing Dcr1 mutants that lose association with the nuclear periphery.

4.3. The physiological relevance of perinuclear Dcr1 localization

Because I could not detect a colocalization between Dcr1 foci and Cnp1 (manuscript I, (Emmerth et al., 2010)), which is the centromere-specific H3 variant, I speculated that centromeric siRNA biogenesis happens off chromatin at the nuclear periphery (manuscript I, (Emmerth et al., 2010)). However, by using a highly sensitive chromatin profiling technique termed DamID, a recent study from our laboratory has shown that Dcr1 is indeed associated with the centromeres and some other loci on chromatin (Woolcock et al., 2011). This suggests that at least two nuclear fractions of Dcr1 exist: a chromatin-associated fraction and a periphery-associated fraction, of which I could only detect the periphery-associated fraction by fluorescence microscopy. However, I propose that at least some parts of these two fractions overlap (see below). The chromatin-associated fraction is most likely involved in centromeric siRNA production and RNAi-mediated heterochromatin formation. Correspondingly, I found that association of Dcr1 with the nuclear periphery does not seem to be required for centromeric heterochromatin formation (manuscript III). As I speculate above, the peripheral association of Dcr1 provides the enzyme with substrate specificity. But what kind of substrates is peripheral Dcr1 processing?

Interestingly, I found that exposing cells to increased temperatures leads to the dissociation of Dcr1 from the nuclear periphery and export to the cytoplasm (manuscript IV). Furthermore, DamID for Dcr1 at high temperatures could not be performed also suggesting that the protein is no longer present in the nucleus (K. Woolcock, personal communication). Together, these results point to a role for Dcr1 in the heat shock response in *S. pombe*. Intriguingly, some of the few transcripts that are upregulated in *dcr1Δ* cells correspond to heat shock response genes (manuscript I, (Emmerth et al., 2010), (Provost et al., 2002), and K. Woolcock, personal communication), and correspondingly, Dcr1 also associates with these regions of the genome (K. Woolcock, personal communication). Notably, it has been shown that heat shock genes in *Drosophila*

melanogaster localize to the nuclear periphery and this localization is important to regulate the expression of the corresponding mRNAs (Kurshakova et al., 2007). Likewise, I propose that the fraction of Dcr1 that is associated with the nuclear periphery is required to downregulate the expression of heat shock genes under normal conditions (this would also imply that some of the periphery- and chromatin-associated fractions of Dcr1 overlap). Upon heat shock, the protein-protein interaction surface on the dsRBD loses the association with a putative nuclear retention factor (manuscript II), the dsRBD exports Dcr1, and heat shock genes are expressed. Alternatively, one could also imagine that zinc ion binding in the C-terminus is lost upon exposure to high temperatures creating a situation similar to mutating the zinc-binding residues (manuscript II), again leading to the dsRBD-mediated export of Dcr1 (see also Figure 15).

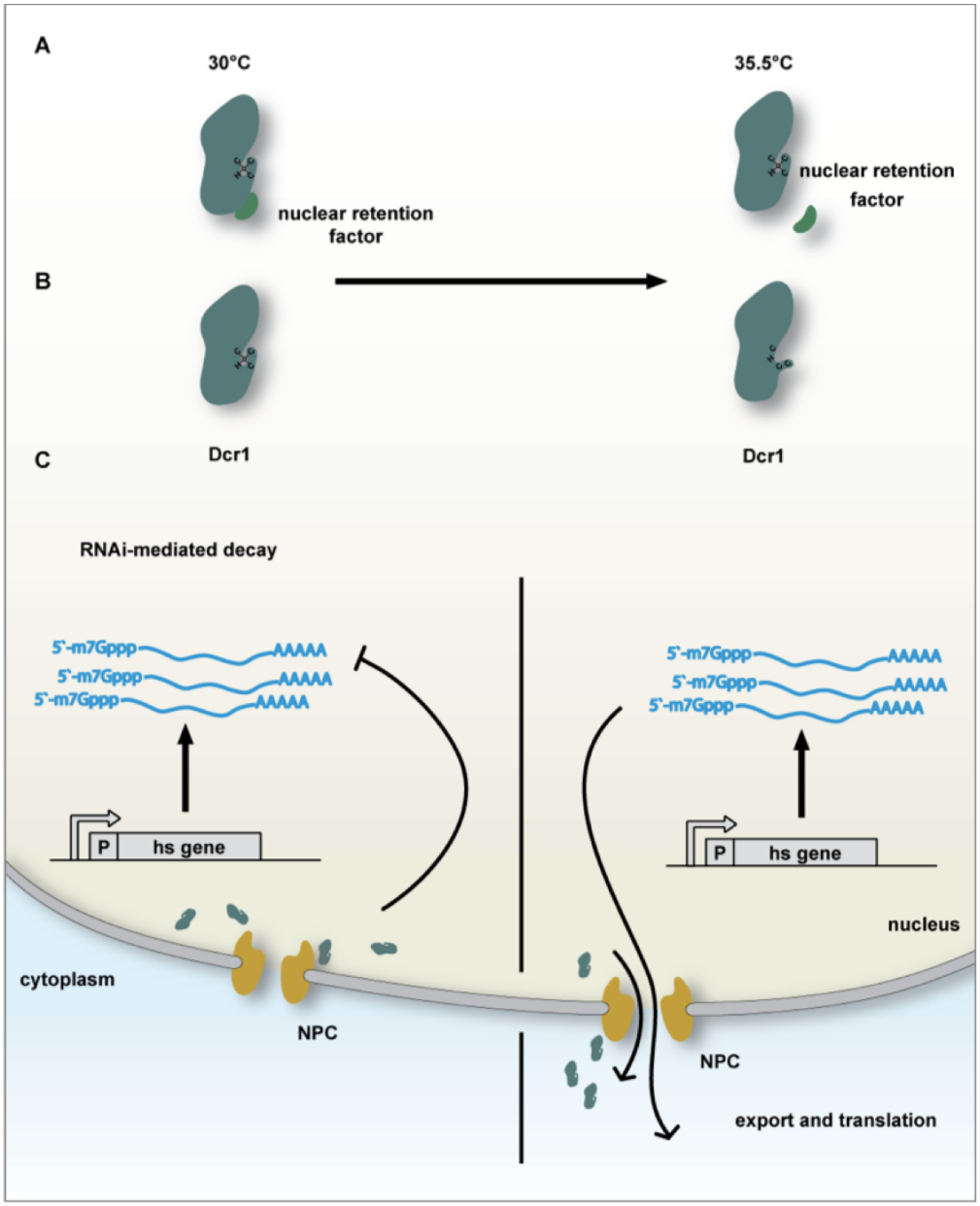


Figure 15: Model for the involvement of Dcr1 in the heat shock response in *S. pombe*
A: The interaction between Dcr1 and a putative nuclear retention factor is lost upon increased temperatures, and the dsRBD can mediate export of Dcr1. **B:** Alternatively, zinc coordination in the C-terminus is lost at high temperatures leading to a structural-induced change unmasking the export-promoting feature of the dsRBD. **C:** Peripheral Dcr1 controls the expression of heat shock genes: upon heat shock, Dcr1 is exported from the nucleus and heat shock genes are expressed. See main text for details. Only the fraction of Dcr1 that is associated with the nuclear periphery is shown.
hs gene: heat shock gene.

What would be the benefit of such a regulation? Genome wide studies assessing H3K9me status in *S. pombe* have not found this mark to be present at loci encoding for heat shock genes (Cam et al., 2005). Therefore, a possible regulation via Dcr1 occurs most likely co-transcriptionally and independently of

heterochromatin, with the heat shock messages constantly generated but rapidly turned over by Dcr1 via CTGS. When Dcr1 leaves the nucleus upon heat shock, the mRNAs are no longer degraded leading to their rapid export and translation in the cytoplasm. On the other hand, in *Drosophila melanogaster* it has been shown that promoters of heat shock genes are bound by paused RNA Pol II, and transcription is resumed upon heat shock (Lee et al., 1992). The situation could be similar in *S. pombe* and Dcr1 could contribute to RNA Pol II pausing. Upon heat shock, Dcr1 would leave the nucleus, RNA Pol II immediately resumes transcription and mRNAs corresponding to heat shock genes are rapidly generated.

In both cases, the messages would be quickly translated. As a result, the cells are able to very rapidly initiate parts of the heat shock response, providing an initial protection until other stress response pathways are activated to cope with the effects of exposing *S. pombe* to high temperatures.

4.4. The dsRBD – a shuttling signal that is conserved in Dicer enzymes?

I found that the dsRBD of Dcr1 can mediate nucleo-cytoplasmic shuttling (manuscript I, (Emmerth et al., 2010)). Moreover, *in vitro* studies have shown that this domain is not required for Dcr1 to process dsRNAs (Colmenares et al., 2007), and I found that binding of the dsRBD to dsRNA is dispensable for the formation of centromeric heterochromatin (manuscript II). This implies that the dsRBD is a regulatory module, which mediates nucleo-cytoplasmic shuttling of Dcr1 but is not necessarily involved in the biogenesis of siRNAs. Notably, human and mouse Dicers associate with ribosomal DNA (Sinkkonen et al., 2010) and nucleo-cytoplasmic shuttling of human Dicer that depends on its dsRBD has also been observed (M. Doyle and W. Filipowicz, personal communication). Furthermore, deleting the dsRBD of human Dicer has only a minor impact on the processivity of the enzyme (Zhang et al., 2004). Together, these data suggest that the dsRBDs of Dicer proteins in general are not necessarily required to bind to RNA. Rather, they could be involved in regulating nucleo-cytoplasmic shuttling of Dicer enzymes, whereas the dsRBD-containing cofactors that usually associate

with Dicer enzymes mediate the association of Dicer proteins with RNA. Accordingly, it has been shown that the interactions of *Drosophila melanogaster* Dicer-1 and Dicer-2 with Loquacious and R2D2, respectively, are important for miRNA- and siRNA-mediated silencing, respectively. Depleting Loquacious or R2D2 interfered with the *in vivo* accumulation of miRNAs and siRNA-mediated silencing, respectively (Forstemann et al., 2005; Liu et al., 2003; Saito et al., 2005). Moreover, *in vitro* assays demonstrate that Loquacious together with Dicer-1 is required to efficiently process pre-miRNAs into miRNAs, and that R2D2 directs the strand-specific incorporation of siRNAs into RISC (Forstemann et al., 2005; Liu et al., 2003). Furthermore, whereas Dicer-1, which is specific for the miRNA pathway, can cleave a long dsRNA *in vitro*, and Dicer-2, which is specific for the siRNA pathway, can process pre-miRNAs *in vitro*, these cleavage events are prevented when Loquacious and R2D2, respectively, are added to the reactions (Cenik et al., 2011; Saito et al., 2005). These data suggest that the dsRBD-containing cofactors are required for specificity and support my hypothesis that the dsRBDs of Dicer enzymes in general are not necessarily involved in RNA binding.

4.5. A nuclear function for mammalian Dicers?

As mentioned above, I propose that the dsRBDs of Dicer enzymes in general are required to control nucleo-cytoplasmic shuttling of these enzymes. But what would be the nuclear function of mammalian Dicer enzymes? Earlier reports suggested that transfected siRNAs induce RNAi-dependent formation of heterochromatin (Janowski et al., 2006; Kim et al., 2006; Ting et al., 2005), and mammalian Dicers were linked to the formation of centromeric heterochromatin and to the regulation of intergenic transcription at the human and chicken β-globin locus (Fukagawa et al., 2004; Giles et al., 2010; Haussecker and Proudfoot, 2005; Kanellopoulou et al., 2005). However, recent data have indicated that some of these effects can be indirectly mediated by miRNAs, demonstrating that results obtained with cells depleted for Dicer or Argonaute proteins have to be treated with caution (Benetti et al., 2008; Sinkkonen et al., 2008). Taken together, so far no satisfying evidence has been presented that would allow the

conclusion that RNAi directly mediates CDGS in mammals similar to what is found in *S. pombe*, plants and *Tetrahymena thermophila*, at least not globally.

However, my findings that Dcr1 is exported at high temperatures point towards a role for the dsRBD-mediated shuttling in stress response pathways. To my knowledge, no experiments have been performed so far where the influence of RNAi pathways on the stress response in the nucleus has been examined in mammals. One could imagine that Dicer enzymes would be imported into the nucleus via their dsRBDs upon stress to perform CTGS, similar to the situation in *S. pombe*, as a method to degrade RNA on chromatin, thereby silencing a subset of genes. However, this is pure speculation and future studies need to show whether this holds true for Dicer enzymes, as well as RNAi pathways in general.

5. Outlook

5.1. Investigate the role of dsRBD-mediated export

Work I have performed in my PhD thesis revealed that Dcr1 is a nucleo-cytoplasmic shuttling protein and that shuttling is mediated by the dsRBD. However, the precise mechanisms of shuttling remain unknown. To identify key factors that mediate nuclear export of Dcr1 via its dsRBD, genetics and proteomics approaches should be employed in parallel.

5.1.1. Reverse genetic screen to identify export deficient Dcr1

A reverse genetic screen should identify key residues in the dsRBD that promote export. A library of randomly mutagenized dsRBD sequences should be transformed into a reporter strain that expresses (cytoplasmic) GFP-Dcr1-SHSS (manuscript II) and that contains an *ade6+* transgene integrated at centromeric heterochromatin that is not silenced. GFP-Dcr1-SHSS would be mutagenized by homologous recombination. Mutations in residues that are involved in nuclear export are expected to silence the *ade6+* reporter transgene, resulting in red colonies. They should be analyzed by fluorescence microscopy and strains with a nuclear localization pattern of GFP-Dcr1 should also be sequenced to identify amino acids important for promoting export (see also Figure 16).

5.1.2. Affinity purification of the Dcr1 dsRBD

To identify receptors that mediate export of the dsRBD, the domain should be fused to a TAP-tag and purified. As a control, the export deficient dsRBD sequence could be used and associated proteins should be identified by mass spectrometry to identify the putative export receptor.

5.1.3. Forward genetic screen to identify export factors

If the export receptor cannot be found by purifying the dsRBD, a forward genetic screen should be conducted. The same reporter strain as described in section 5.1.1. should be crossed to a *S. pombe* haploid deletion library, which is commercially available, and positive hits should be selected based on silencing of

the *ade6+* reporter gene as depicted in section 5.1.1.. The deletion library contains genomic barcodes allowing for rapid PCR-based identification of the deleted genes (see also Figure 16).

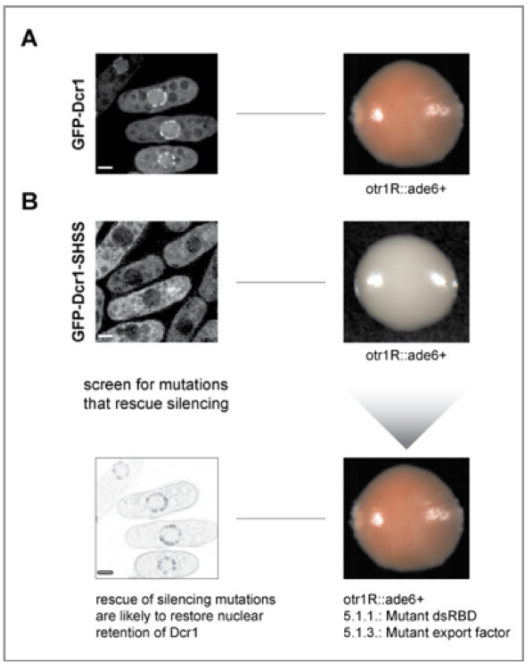


Figure 16: Screen to identify key features mediating the dsRBD-promoted export of Dcr1
A: Nuclear retention of Dcr1 is required for heterochromatic silencing of an *otr1R::ade6+* reporter gene. **B:** Nuclear retention of Dcr1, and therefore heterochromatic silencing, is lost in point mutants of Dcr1 that fail to coordinate zinc (Dcr1-SHSS, see also manuscript II). Mutations that abolish export of Dcr1-SHSS are expected to rescue silencing of the *ade6+* reporter gene. Therefore, red colonies will be selected after mutagenesis and further characterized.

5.1.4. Investigate the physiological consequences of export deficient strains

Once the export receptor and key residues promoting export of the dsRBD would be identified, these mutations should be genomically introduced into wild-type Dcr1 and correspondingly, the export receptor should be deleted. Microarray studies should be performed with cells expressing the export deficient Dcr1 protein and cells that have the export receptor deleted to identify possible new targets, which are regulated by the export of Dcr1. Additionally, the heat shock response of cells expressing export deficient Dcr1 proteins should be assessed (see section 5.2.).

In summary, through genetics and proteomics approaches, the factors involved in dsRBD-mediated export could be identified, and the physiological role of this export pathway could be further examined. These experiments would provide greater insights into the mechanisms and physiological roles of dsRBD-mediated export of Dcr1.

5.2. Investigation of the functional significance of perinuclear Dcr1 localization

Since deleting the Platform domain in Dcr1 rendered the protein unstable (manuscript III) I was not able to analyze the effects of detaching Dcr1 from the nuclear periphery. However, the evidence I have provided suggests that localizing Dcr1 to the nuclear periphery is important to control the expression of heat shock genes. This would predict that the genomic loci of heat shock genes localize to the nuclear periphery, similar to the situation in *Drosophila melanogaster* (Kurshakova et al., 2007). Experiments should be performed that examine the involvement of RNAi in the heat shock response.

5.2.1. Localizing the genomic loci of heat shock genes in the nucleus

To show the localization of the genomic loci of heat shock genes in the nucleus, the LacI/LacO system should be employed: LacO sites should be integrated adjacent to genomic loci of heat shock genes that were found to be associated with Dcr1 (K. Woolcock, personal communication), and the loci should be visualized through a GFP-LacI fusion (Belmont, 2001), which can bind to LacO sites. Furthermore, colocalization studies between mCherry-Dcr1 and GFP-LacI could be performed.

5.2.2. Assessing the heat shock response in strains expressing export deficient Dcr1 protein

The genomic copy of wild-type Dcr1 should be replaced with the sequence of the export deficient Dcr1 protein (see section 5.1.). The resulting strain should be subjected to heat shock and the levels of mRNAs corresponding to heat shock genes as well as the viability of the mutant *S. pombe* strain under heat shock conditions should be assessed.

In summary, the experiments I propose in this section could reveal whether peripheral Dcr1 colocalizes with the genomic loci of heat shock genes, which would further strengthen my hypothesis that Dcr1 associated with the nuclear periphery is required to efficiently silence heat shock genes.

Acknowledgments

First, I would like to thank Marc Bühler for giving me the opportunity to perform my PhD thesis in his laboratory and for the great support he always provided me with during my PhD. I would also like to thank him for his never-ending enthusiasm he always shows - even when the data I presented not seemed to make sense at first -, for his constant support and endless flow of new ideas, hypothesis, and suggestions one could test, and for just being a great boss. And also for always betting (regardless of how hopeless it seemed) on the positive outcome of an experiment (I must have won numerous of these…).

Furthermore, I would like to thank all the current and past members of the group for the great working atmosphere and constant scientific input for my project. Especially, I would like to thank Yukiko Shimada for introducing me to the world of yeast research, for all the little tricks one cannot find in the protocols that make life much easier, for all the work she has performed for my project, and for always taking care that the laboratory does not run out of enzymes, gels, markers, etc. I would also like to thank Tanel Punga for all the great coffee breaks, Fischerstube evenings and jokes, but also scientific suggestions, which made working really enjoyable. Furthermore, I would like to thank Heiko Schober for introducing me to the world of fluorescence microscopy and for the great collaboration we had. I would also like to thank Nathalie Laschet for the effort she made in deciphering my handwriting and for correcting my sometimes-chaotic strain and plasmid annotations. Moreover, I want to thank Katy Woolcock for correcting the Swiss English of this thesis (Swenglish) and eliminating numerous commas, Raghav Kulasegaran for proofreading, and Claudia Keller for always taking care when something was to be organized.

Ultimately, I would like to acknowledge Javier Martinez and Helge Grosshans as members of my thesis committee for investing their time to being in my committee (especially Javier, who always had to fly in from Vienna for the meetings) and for all the suggestions during the thesis committee meetings and for evaluating the thesis.

And I would also like to thank Mio Mülthaler from Sigma-Aldrich for his undetected but invaluable contributions to this project.

References

Allen, E., Xie, Z., Gustafson, A.M., and Carrington, J.C. (2005). microRNA-directed phasing during trans-acting siRNA biogenesis in plants. Cell *121*, 207-221.

Aravin, A., Gaidatzis, D., Pfeffer, S., Lagos-Quintana, M., Landgraf, P., Iovino, N., Morris, P., Brownstein, M.J., Kuramochi-Miyagawa, S., Nakano, T.*, et al.* (2006). A novel class of small RNAs bind to MILI protein in mouse testes. Nature *442*, 203-207.

Aravin, A.A., Sachidanandam, R., Girard, A., Fejes-Toth, K., and Hannon, G.J. (2007). Developmentally regulated piRNA clusters implicate MILI in transposon control. Science *316*, 744-747.

Bahler, J., Wu, J.Q., Longtine, M.S., Shah, N.G., McKenzie, A., 3rd, Steever, A.B., Wach, A., Philippsen, P., and Pringle, J.R. (1998). Heterologous modules for efficient and versatile PCR-based gene targeting in Schizosaccharomyces pombe. Yeast *14*, 943-951.

Baltimore, D. (1970). RNA-dependent DNA polymerase in virions of RNA tumour viruses. Nature *226*, 1209-1211.

Bannister, A.J., Zegerman, P., Partridge, J.F., Miska, E.A., Thomas, J.O., Allshire, R.C., and Kouzarides, T. (2001). Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. Nature *410*, 120-124.

Batista, P.J., Ruby, J.G., Claycomb, J.M., Chiang, R., Fahlgren, N., Kasschau, K.D., Chaves, D.A., Gu, W., Vasale, J.J., Duan, S.*, et al.* (2008). PRG-1 and 21U-RNAs interact to form the piRNA complex required for fertility in C. elegans. Mol Cell *31*, 67-78.

Bayne, E.H., White, S.A., Kagansky, A., Bijos, D.A., Sanchez-Pulido, L., Hoe, K.L., Kim, D.U., Park, H.O., Ponting, C.P., Rappsilber, J.*, et al.* (2010). Stc1: a critical link between RNAi and chromatin modification required for heterochromatin integrity. Cell *140*, 666-677.

Belmont, A.S. (2001). Visualizing chromosome dynamics with GFP. Trends Cell Biol *11*, 250-257.

Benetti, R., Gonzalo, S., Jaco, I., Munoz, P., Gonzalez, S., Schoeftner, S., Murchison, E., Andl, T., Chen, T., Klatt, P.*, et al.* (2008). A mammalian microRNA cluster controls DNA methylation and telomere recombination via Rbl2-dependent regulation of DNA methyltransferases. Nat Struct Mol Biol *15*, 998.

Bernstein, E., Caudy, A.A., Hammond, S.M., and Hannon, G.J. (2001). Role for a bidentate ribonuclease in the initiation step of RNA interference. Nature *409*, 363-366.

Bjerling, P., Silverstein, R.A., Thon, G., Caudy, A., Grewal, S., and Ekwall, K. (2002). Functional divergence between histone deacetylases in fission yeast by distinct cellular localization and in vivo specificity. Mol Cell Biol *22*, 2170-2181.

Brennecke, J., Aravin, A.A., Stark, A., Dus, M., Kellis, M., Sachidanandam, R., and Hannon, G.J. (2007). Discrete small RNA-generating loci as master regulators of transposon activity in Drosophila. Cell *128*, 1089-1103.

Brennecke, J., Stark, A., Russell, R.B., and Cohen, S.M. (2005). Principles of microRNA-target recognition. PLoS Biol *3*, e85.

Brodersen, P., and Voinnet, O. (2006). The diversity of RNA silencing pathways in plants. Trends Genet *22*, 268-280.

Brower-Toland, B., Findley, S.D., Jiang, L., Liu, L., Yin, H., Dus, M., Zhou, P., Elgin, S.C., and Lin, H. (2007). Drosophila PIWI associates with chromatin and interacts directly with HP1a. Genes Dev *21*, 2300-2311.

Buhler, M. (2009). RNA turnover and chromatin-dependent gene silencing. Chromosoma *118*, 141-151.

Buhler, M., Haas, W., Gygi, S.P., and Moazed, D. (2007). RNAi-dependent and -independent RNA turnover mechanisms contribute to heterochromatic gene silencing. Cell *129*, 707-721.

Buhler, M., Spies, N., Bartel, D.P., and Moazed, D. (2008). TRAMP-mediated RNA surveillance prevents spurious entry of RNAs into the Schizosaccharomyces pombe siRNA pathway. Nat Struct Mol Biol *15*, 1015-1023.

Buhler, M., Verdel, A., and Moazed, D. (2006). Tethering RITS to a nascent transcript initiates RNAi- and heterochromatin-dependent gene silencing. Cell *125*, 873-886.

Buker, S.M., Iida, T., Buhler, M., Villen, J., Gygi, S.P., Nakayama, J., and Moazed, D. (2007). Two different Argonaute complexes are required for siRNA generation and heterochromatin assembly in fission yeast. Nat Struct Mol Biol *14*, 200-207.

Cai, X., Hagedorn, C.H., and Cullen, B.R. (2004). Human microRNAs are processed from capped, polyadenylated transcripts that can also function as mRNAs. RNA *10*, 1957-1966.

Cam, H.P., Sugiyama, T., Chen, E.S., Chen, X., FitzGerald, P.C., and Grewal, S.I. (2005). Comprehensive analysis of heterochromatin- and RNAi-mediated epigenetic control of the fission yeast genome. Nat Genet *37*, 809-819.

Canzio, D., Chang, E.Y., Shankar, S., Kuchenbecker, K.M., Simon, M.D., Madhani, H.D., Narlikar, G.J., and Al-Sady, B. (2011). Chromodomain-mediated oligomerization of HP1 suggests a nucleosome-bridging mechanism for heterochromatin assembly. Mol Cell *41*, 67-81.

Carmell, M.A., and Hannon, G.J. (2004). RNase III enzymes and the initiation of gene silencing. Nat Struct Mol Biol *11*, 214-218.

Carmell, M.A., Xuan, Z., Zhang, M.Q., and Hannon, G.J. (2002). The Argonaute family: tentacles that reach into RNAi, developmental control, stem cell maintenance, and tumorigenesis. Genes Dev *16*, 2733-2742.

Carmichael, J.B., Stoica, C., Parker, H., McCaffery, J.M., Simmonds, A.J., and Hobman, T.C. (2006). RNA interference effector proteins localize to mobile cytoplasmic puncta in Schizosaccharomyces pombe. Traffic *7*, 1032-1044.

Carthew, R.W., and Sontheimer, E.J. (2009). Origins and Mechanisms of miRNAs and siRNAs. Cell *136*, 642-655.

Cenik, E.S., Fukunaga, R., Lu, G., Dutcher, R., Wang, Y., Tanaka Hall, T.M., and Zamore, P.D. (2011). Phosphate and R2D2 Restrict the Substrate Specificity of Dicer-2, an ATP-Driven Ribonuclease. Mol Cell *42*, 172-184.

Chen, E.S., Zhang, K., Nicolas, E., Cam, H.P., Zofall, M., and Grewal, S.I. (2008). Cell cycle control of centromeric repeat transcription and heterochromatin assembly. Nature *451*, 734-737.

Chendrimada, T.P., Finn, K.J., Ji, X., Baillat, D., Gregory, R.I., Liebhaber, S.A., Pasquinelli, A.E., and Shiekhattar, R. (2007). MicroRNA silencing through RISC recruitment of eIF6. Nature *447*, 823-828.

Chendrimada, T.P., Gregory, R.I., Kumaraswamy, E., Norman, J., Cooch, N., Nishikura, K., and Shiekhattar, R. (2005). TRBP recruits the Dicer complex to Ago2 for microRNA processing and gene silencing. Nature *436*, 740-744.

Chung, W.J., Okamura, K., Martin, R., and Lai, E.C. (2008). Endogenous RNA interference provides a somatic defense against Drosophila transposons. Curr Biol *18*, 795-802.

Cogoni, C., and Macino, G. (1999). Gene silencing in Neurospora crassa requires a protein homologous to RNA-dependent RNA polymerase. Nature *399*, 166-169.

Colmenares, S.U., Buker, S.M., Buhler, M., Dlakic, M., and Moazed, D. (2007). Coupling of double-stranded RNA synthesis and siRNA generation in fission yeast RNAi. Mol Cell *27*, 449-461.

Cowieson, N.P., Partridge, J.F., Allshire, R.C., and McLaughlin, P.J. (2000). Dimerisation of a chromo shadow domain and distinctions from the chromodomain as revealed by structural analysis. Curr Biol *10*, 517-525.

Crick, F.H. (1958). On protein synthesis. Symp Soc Exp Biol *12*, 138-163.

Czech, B., Malone, C.D., Zhou, R., Stark, A., Schlingeheyde, C., Dus, M., Perrimon, N., Kellis, M., Wohlschlegel, J.A., Sachidanandam, R.*, et al.* (2008). An endogenous small interfering RNA pathway in Drosophila. Nature *453*, 798-802.

Dalmay, T., Hamilton, A., Rudd, S., Angell, S., and Baulcombe, D.C. (2000). An RNA-dependent RNA polymerase gene in Arabidopsis is required for posttranscriptional gene silencing mediated by a transgene but not by a virus. Cell *101*, 543-553.

Debeauchamp, J.L., Moses, A., Noffsinger, V.J., Ulrich, D.L., Job, G., Kosinski, A.M., and Partridge, J.F. (2008). Chp1-Tas3 interaction is required to recruit RITS to fission yeast centromeres and for maintenance of centromeric heterochromatin. Mol Cell Biol *28*, 2154-2166.

Denli, A.M., Tops, B.B., Plasterk, R.H., Ketting, R.F., and Hannon, G.J. (2004). Processing of primary microRNAs by the Microprocessor complex. Nature *432*, 231-235.

Didiano, D., and Hobert, O. (2006). Perfect seed pairing is not a generally reliable predictor for miRNA-target interactions. Nat Struct Mol Biol *13*, 849-851.

Ding, S.W., and Voinnet, O. (2007). Antiviral immunity directed by small RNAs. Cell *130*, 413-426.

Ding, X.C., and Grosshans, H. (2009). Repression of C. elegans microRNA targets at the initiation level of translation requires GW182 proteins. EMBO J *28*, 213-222.

Djupedal, I., Portoso, M., Spahr, H., Bonilla, C., Gustafsson, C.M., Allshire, R.C., and Ekwall, K. (2005). RNA Pol II subunit Rpb7 promotes centromeric transcription and RNAi-directed chromatin silencing. Genes Dev *19*, 2301-2306.

Doench, J.G., and Sharp, P.A. (2004). Specificity of microRNA target selection in translational repression. Genes Dev *18*, 504-511.

Drinnenberg, I.A., Weinberg, D.E., Xie, K.T., Mower, J.P., Wolfe, K.H., Fink, G.R., and Bartel, D.P. (2009). RNAi in budding yeast. Science *326*, 544-550.

Duchaine, T.F., Wohlschlegel, J.A., Kennedy, S., Bei, Y., Conte, D., Jr., Pang, K., Brownell, D.R., Harding, S., Mitani, S., Ruvkun, G.*, et al.* (2006). Functional proteomics reveals the biochemical niche of C. elegans DCR-1 in multiple small-RNA-mediated pathways. Cell *124*, 343-354.

Elbashir, S.M., Harborth, J., Lendeckel, W., Yalcin, A., Weber, K., and Tuschl, T. (2001a). Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. Nature *411*, 494-498.

Elbashir, S.M., Lendeckel, W., and Tuschl, T. (2001b). RNA interference is mediated by 21- and 22-nucleotide RNAs. Genes Dev *15*, 188-200.

Emmerth, S., Schober, H., Gaidatzis, D., Roloff, T., Jacobeit, K., and Buhler, M. (2010). Nuclear retention of fission yeast dicer is a prerequisite for RNAi-mediated heterochromatin assembly. Dev Cell *18*, 102-113.

Eulalio, A., Behm-Ansmant, I., Schweizer, D., and Izaurralde, E. (2007). P-body formation is a consequence, not the cause, of RNA-mediated gene silencing. Mol Cell Biol *27*, 3970-3981.

Fang, Y., and Spector, D.L. (2007). Identification of nuclear dicing bodies containing proteins for microRNA biogenesis in living Arabidopsis plants. Curr Biol *17*, 818-823.

Felsenfeld, G., and Groudine, M. (2003). Controlling the double helix. Nature *421*, 448-453.

Fire, A., Xu, S., Montgomery, M.K., Kostas, S.A., Driver, S.E., and Mello, C.C. (1998). Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. Nature *391*, 806-811.

Forstemann, K., Horwich, M.D., Wee, L., Tomari, Y., and Zamore, P.D. (2007). Drosophila microRNAs are sorted into functionally distinct argonaute complexes after production by dicer-1. Cell *130*, 287-297.

Forstemann, K., Tomari, Y., Du, T., Vagin, V.V., Denli, A.M., Bratu, D.P., Klattenhoff, C., Theurkauf, W.E., and Zamore, P.D. (2005). Normal microRNA maturation and germ-line stem cell maintenance requires Loquacious, a double-stranded RNA-binding domain protein. PLoS Biol *3*, e236.

Fromont-Racine, M., Senger, B., Saveanu, C., and Fasiolo, F. (2003). Ribosome assembly in eukaryotes. Gene *313*, 17-42.

Fukagawa, T., Nogami, M., Yoshikawa, M., Ikeno, M., Okazaki, T., Takami, Y., Nakayama, T., and Oshimura, M. (2004). Dicer is essential for formation of the heterochromatin structure in vertebrate cells. Nat Cell Biol *6*, 784-791.

Furuya, K., and Niki, H. (2009). Isolation of heterothallic haploid and auxotrophic mutants of Schizosaccharomyces japonicus. Yeast *26*, 221-233.

Ghildiyal, M., Seitz, H., Horwich, M.D., Li, C., Du, T., Lee, S., Xu, J., Kittler, E.L., Zapp, M.L., Weng, Z., *et al.* (2008). Endogenous siRNAs derived from transposons and mRNAs in Drosophila somatic cells. Science *320*, 1077-1081.

Gibbings, D.J., Ciaudo, C., Erhardt, M., and Voinnet, O. (2009). Multivesicular bodies associate with components of miRNA effector complexes and modulate miRNA activity. Nat Cell Biol *11*, 1143-1149.

Giles, K.E., Ghirlando, R., and Felsenfeld, G. (2010). Maintenance of a constitutive heterochromatin domain in vertebrates by a Dicer-dependent mechanism. Nat Cell Biol *12*, 94-99; sup pp 91-96.

Giraldez, A.J., Mishima, Y., Rihel, J., Grocock, R.J., Van Dongen, S., Inoue, K., Enright, A.J., and Schier, A.F. (2006). Zebrafish MiR-430 promotes deadenylation and clearance of maternal mRNAs. Science *312*, 75-79.

Girard, A., Sachidanandam, R., Hannon, G.J., and Carmell, M.A. (2006). A germline-specific class of small RNAs binds mammalian Piwi proteins. Nature *442*, 199-202.

Gottschling, D.E. (1992). Telomere-proximal DNA in Saccharomyces cerevisiae is refractory to methyltransferase activity in vivo. Proc Natl Acad Sci U S A *89*, 4062-4065.

Gregory, R.I., Yan, K.P., Amuthan, G., Chendrimada, T., Doratotaj, B., Cooch, N., and Shiekhattar, R. (2004). The Microprocessor complex mediates the genesis of microRNAs. Nature *432*, 235-240.

Grewal, S.I., Bonaduce, M.J., and Klar, A.J. (1998). Histone deacetylase homologs regulate epigenetic inheritance of transcriptional silencing and chromosome segregation in fission yeast. Genetics *150*, 563-576.

Grewal, S.I., and Jia, S. (2007). Heterochromatin revisited. Nat Rev Genet *8*, 35-46.

Grewal, S.I., and Klar, A.J. (1997). A recombinationally repressed region between mat2 and mat3 loci shares homology to centromeric repeats and regulates directionality of mating-type switching in fission yeast. Genetics *146*, 1221-1238.

Grishok, A., Pasquinelli, A.E., Conte, D., Li, N., Parrish, S., Ha, I., Baillie, D.L., Fire, A., Ruvkun, G., and Mello, C.C. (2001). Genes and mechanisms related to RNA interference regulate expression of the small temporal RNAs that control C. elegans developmental timing. Cell *106*, 23-34.

Grivna, S.T., Pyhtila, B., and Lin, H. (2006). MIWI associates with translational machinery and PIWI-interacting RNAs (piRNAs) in regulating spermatogenesis. Proc Natl Acad Sci U S A *103*, 13415-13420.

Gullerova, M., and Proudfoot, N.J. (2008). Cohesin complex promotes transcriptional termination between convergent genes in S. pombe. Cell *132*, 983-995.

Gunawardane, L.S., Saito, K., Nishida, K.M., Miyoshi, K., Kawamura, Y., Nagami, T., Siomi, H., and Siomi, M.C. (2007). A slicer-mediated mechanism for repeat-associated siRNA 5' end formation in Drosophila. Science *315*, 1587-1590.

Guo, H., Ingolia, N.T., Weissman, J.S., and Bartel, D.P. (2010). Mammalian microRNAs predominantly act to decrease target mRNA levels. Nature *466*, 835-840.

Guo, S., and Kemphues, K.J. (1995). par-1, a gene required for establishing polarity in C. elegans embryos, encodes a putative Ser/Thr kinase that is asymmetrically distributed. Cell *81*, 611-620.

Haase, A.D., Jaskiewicz, L., Zhang, H., Laine, S., Sack, R., Gatignol, A., and Filipowicz, W. (2005). TRBP, a regulator of cellular PKR and HIV-1 virus expression, interacts with Dicer and functions in RNA silencing. EMBO Rep *6*, 961-967.

Halic, M., and Moazed, D. (2010). Dicer-independent primal RNAs trigger RNAi and heterochromatin formation. Cell *140*, 504-516.

Hall, I.M., Noma, K., and Grewal, S.I. (2003). RNA interference machinery regulates chromosome dynamics during mitosis and meiosis in fission yeast. Proc Natl Acad Sci U S A *100*, 193-198.

Hall, I.M., Shankaranarayana, G.D., Noma, K., Ayoub, N., Cohen, A., and Grewal, S.I. (2002). Establishment and maintenance of a heterochromatin domain. Science *297*, 2232-2237.

Hamilton, A.J., and Baulcombe, D.C. (1999). A species of small antisense RNA in posttranscriptional gene silencing in plants. Science *286*, 950-952.

Hammond, S.M. (2005). Dicing and slicing: the core machinery of the RNA interference pathway. FEBS Lett *579*, 5822-5829.

Hammond, S.M., Bernstein, E., Beach, D., and Hannon, G.J. (2000). An RNA-directed nuclease mediates post-transcriptional gene silencing in Drosophila cells. Nature *404*, 293-296.

Han, J., Lee, Y., Yeom, K.H., Kim, Y.K., Jin, H., and Kim, V.N. (2004). The Drosha-DGCR8 complex in primary microRNA processing. Genes Dev *18*, 3016-3027.

Hansen, K.R., Ibarra, P.T., and Thon, G. (2006). Evolutionary-conserved telomere-linked helicase genes of fission yeast are repressed by silencing factors, RNAi components and the telomere-binding protein Taz1. Nucleic Acids Res *34*, 78-88.

Harris, A.N., and Macdonald, P.M. (2001). Aubergine encodes a Drosophila polar granule component required for pole cell formation and related to eIF2C. Development *128*, 2823-2832.

Haussecker, D., and Proudfoot, N.J. (2005). Dicer-dependent turnover of intergenic transcripts from the human beta-globin gene cluster. Mol Cell Biol *25*, 9724-9733.

Henderson, I.R., and Jacobsen, S.E. (2007). Epigenetic inheritance in plants. Nature *447*, 418-424.

Henderson, I.R., Zhang, X., Lu, C., Johnson, L., Meyers, B.C., Green, P.J., and Jacobsen, S.E. (2006). Dissecting Arabidopsis thaliana DICER function in small RNA processing, gene silencing and DNA methylation patterning. Nat Genet *38*, 721-725.

Hock, J., Weinmann, L., Ender, C., Rudel, S., Kremmer, E., Raabe, M., Urlaub, H., and Meister, G. (2007). Proteomic and functional analysis of Argonaute-containing mRNA-protein complexes in human cells. EMBO Rep *8*, 1052-1060.

Humphreys, D.T., Westman, B.J., Martin, D.I., and Preiss, T. (2005). MicroRNAs control translation initiation by inhibiting eukaryotic initiation factor 4E/cap and poly(A) tail function. Proc Natl Acad Sci U S A *102*, 16961-16966.

Hunter, C., Sun, H., and Poethig, R.S. (2003). The Arabidopsis heterochronic gene ZIPPY is an ARGONAUTE family member. Curr Biol *13*, 1734-1739.

Hutvagner, G., McLachlan, J., Pasquinelli, A.E., Balint, E., Tuschl, T., and Zamore, P.D. (2001). A cellular function for the RNA-interference enzyme Dicer in the maturation of the let-7 small temporal RNA. Science *293*, 834-838.

Irvine, D.V., Zaratiegui, M., Tolia, N.H., Goto, D.B., Chitwood, D.H., Vaughn, M.W., Joshua-Tor, L., and Martienssen, R.A. (2006). Argonaute slicing is required for heterochromatic silencing and spreading. Science *313*, 1134-1137.

Jannot, G., Boisvert, M.E., Banville, I.H., and Simard, M.J. (2008). Two molecular features contribute to the Argonaute specificity for the microRNA and RNAi pathways in C. elegans. RNA *14*, 829-835.

Janowski, B.A., Huffman, K.E., Schwartz, J.C., Ram, R., Nordsell, R., Shames, D.S., Minna, J.D., and Corey, D.R. (2006). Involvement of AGO1 and AGO2 in mammalian transcriptional silencing. Nat Struct Mol Biol *13*, 787-792.

Jaskiewicz, L., and Filipowicz, W. (2008). Role of Dicer in posttranscriptional RNA silencing. Curr Top Microbiol Immunol *320*, 77-97.

Jenuwein, T., and Allis, C.D. (2001). Translating the histone code. Science *293*, 1074-1080.

Jia, S., Noma, K., and Grewal, S.I. (2004). RNAi-independent heterochromatin nucleation by the stress-activated ATF/CREB family proteins. Science *304*, 1971-1976.

Jinek, M., and Doudna, J.A. (2009). A three-dimensional view of the molecular machinery of RNA interference. Nature *457*, 405-412.

Kagansky, A., Folco, H.D., Almeida, R., Pidoux, A.L., Boukaba, A., Simmer, F., Urano, T., Hamilton, G.L., and Allshire, R.C. (2009). Synthetic heterochromatin bypasses RNAi and centromeric repeats to establish functional centromeres. Science *324*, 1716-1719.

Kanellopoulou, C., Muljo, S.A., Kung, A.L., Ganesan, S., Drapkin, R., Jenuwein, T., Livingston, D.M., and Rajewsky, K. (2005). Dicer-deficient mouse embryonic stem cells are defective in differentiation and centromeric silencing. Genes Dev *19*, 489-501.

Kanoh, J., Sadaie, M., Urano, T., and Ishikawa, F. (2005). Telomere binding protein Taz1 establishes Swi6 heterochromatin independently of RNAi at telomeres. Curr Biol *15*, 1808-1819.

Kato, H., Goto, D.B., Martienssen, R.A., Urano, T., Furukawa, K., and Murakami, Y. (2005). RNA polymerase II is required for RNAi-dependent heterochromatin assembly. Science *309*, 467-469.

Kawamura, Y., Saito, K., Kin, T., Ono, Y., Asai, K., Sunohara, T., Okada, T.N., Siomi, M.C., and Siomi, H. (2008). Drosophila endogenous small RNAs bind to Argonaute 2 in somatic cells. Nature *453*, 793-797.

Keller, C., Woolcock, K., Hess, D., and Buhler, M. (2010). Proteomic and functional analysis of the noncanonical poly(A) polymerase Cid14. RNA *16*, 1124-1129.

Ketting, R.F., Fischer, S.E., Bernstein, E., Sijen, T., Hannon, G.J., and Plasterk, R.H. (2001). Dicer functions in RNA interference and in synthesis of small RNA involved in developmental timing in C. elegans. Genes Dev *15*, 2654-2659.

Kim, D.H., Villeneuve, L.M., Morris, K.V., and Rossi, J.J. (2006). Argonaute-1 directs siRNA-mediated transcriptional gene silencing in human cells. Nat Struct Mol Biol *13*, 793-797.

Kim, H.S., Choi, E.S., Shin, J.A., Jang, Y.K., and Park, S.D. (2004). Regulation of Swi6/HP1-dependent heterochromatin assembly by cooperation of components of the mitogen-activated protein kinase pathway and a histone deacetylase Clr6. J Biol Chem *279*, 42850-42859.

Klattenhoff, C., Xi, H., Li, C., Lee, S., Xu, J., Khurana, J.S., Zhang, F., Schultz, N., Koppetsch, B.S., Nowosielska, A.*, et al.* (2009). The Drosophila HP1 homolog Rhino is required for transposon silencing and piRNA production by dual-strand clusters. Cell *138*, 1137-1149.

Kloc, A., Zaratiegui, M., Nora, E., and Martienssen, R. (2008). RNA interference guides histone modification during the S phase of chromosomal replication. Curr Biol *18*, 490-495.

Kruger, K., Grabowski, P.J., Zaug, A.J., Sands, J., Gottschling, D.E., and Cech, T.R. (1982). Self-splicing RNA: autoexcision and autocyclization of the ribosomal RNA intervening sequence of Tetrahymena. Cell *31*, 147-157.

Kuramochi-Miyagawa, S., Watanabe, T., Gotoh, K., Totoki, Y., Toyoda, A., Ikawa, M., Asada, N., Kojima, K., Yamaguchi, Y., Ijiri, T.W.*, et al.* (2008). DNA methylation of retrotransposon genes is regulated by Piwi family members MILI and MIWI2 in murine fetal testes. Genes Dev *22*, 908-917.

Kurihara, Y., and Watanabe, Y. (2004). Arabidopsis micro-RNA biogenesis through Dicer-like 1 protein functions. Proc Natl Acad Sci U S A *101*, 12753-12758.

Kurshakova, M.M., Krasnov, A.N., Kopytova, D.V., Shidlovskii, Y.V., Nikolenko, J.V., Nabirochkina, E.N., Spehner, D., Schultz, P., Tora, L., and Georgieva, S.G. (2007). SAGA and a novel Drosophila export complex anchor efficient transcription and mRNA export to NPC. EMBO J *26*, 4956-4965.

Lau, N.C., Seto, A.G., Kim, J., Kuramochi-Miyagawa, S., Nakano, T., Bartel, D.P., and Kingston, R.E. (2006). Characterization of the piRNA complex from rat testes. Science *313*, 363-367.

Lee, H., Kraus, K.W., Wolfner, M.F., and Lis, J.T. (1992). DNA sequence requirements for generating paused polymerase at the start of hsp70. Genes Dev *6*, 284-295.

Lee, R.C., Feinbaum, R.L., and Ambros, V. (1993). The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. Cell *75*, 843-854.

Lee, Y., Ahn, C., Han, J., Choi, H., Kim, J., Yim, J., Lee, J., Provost, P., Radmark, O., Kim, S.*, et al.* (2003). The nuclear RNase III Drosha initiates microRNA processing. Nature *425*, 415-419.

Lee, Y., Hur, I., Park, S.Y., Kim, Y.K., Suh, M.R., and Kim, V.N. (2006). The role of PACT in the RNA silencing pathway. EMBO J *25*, 522-532.

Lee, Y., Kim, M., Han, J., Yeom, K.H., Lee, S., Baek, S.H., and Kim, V.N. (2004a). MicroRNA genes are transcribed by RNA polymerase II. EMBO J *23*, 4051-4060.

Lee, Y.S., Nakahara, K., Pham, J.W., Kim, K., He, Z., Sontheimer, E.J., and Carthew, R.W. (2004b). Distinct roles for Drosophila Dicer-1 and Dicer-2 in the siRNA/miRNA silencing pathways. Cell *117*, 69-81.

Lee, Y.S., Pressman, S., Andress, A.P., Kim, K., White, J.L., Cassidy, J.J., Li, X., Lubell, K., Lim do, H., Cho, I.S.*, et al.* (2009). Silencing by small RNAs is linked to endosomal trafficking. Nat Cell Biol *11*, 1150-1156.

Lewis, B.P., Burge, C.B., and Bartel, D.P. (2005). Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. Cell *120*, 15-20.

Li, C., Vagin, V.V., Lee, S., Xu, J., Ma, S., Xi, H., Seitz, H., Horwich, M.D., Syrzycka, M., Honda, B.M.*, et al.* (2009). Collapse of germline piRNAs in the absence of Argonaute3 reveals somatic piRNAs in flies. Cell *137*, 509-521.

Li, C.F., Henderson, I.R., Song, L., Fedoroff, N., Lagrange, T., and Jacobsen, S.E. (2008). Dynamic regulation of ARGONAUTE4 within multiple nuclear bodies in Arabidopsis thaliana. PLoS Genet *4*, e27.

Li, C.F., Pontes, O., El-Shami, M., Henderson, I.R., Bernatavichute, Y.V., Chan, S.W., Lagrange, T., Pikaard, C.S., and Jacobsen, S.E. (2006). An ARGONAUTE4-containing nuclear processing center colocalized with Cajal bodies in Arabidopsis thaliana. Cell *126*, 93-106.

Lim, A.K., and Kai, T. (2007). Unique germ-line organelle, nuage, functions to repress selfish genetic elements in Drosophila melanogaster. Proc Natl Acad Sci U S A *104*, 6714-6719.

Lingel, A., Simon, B., Izaurralde, E., and Sattler, M. (2003). Structure and nucleic-acid binding of the Drosophila Argonaute 2 PAZ domain. Nature *426*, 465-469.

Lingel, A., Simon, B., Izaurralde, E., and Sattler, M. (2004). Nucleic acid 3′-end recognition by the Argonaute2 PAZ domain. Nat Struct Mol Biol *11*, 576-577.

Lippman, Z., May, B., Yordan, C., Singer, T., and Martienssen, R. (2003). Distinct mechanisms determine transposon inheritance and methylation via small interfering RNA and histone modification. PLoS Biol *1*, E67.

Liu, J., Carmell, M.A., Rivas, F.V., Marsden, C.G., Thomson, J.M., Song, J.J., Hammond, S.M., Joshua-Tor, L., and Hannon, G.J. (2004). Argonaute2 is the catalytic engine of mammalian RNAi. Science *305*, 1437-1441.

Liu, J., Rivas, F.V., Wohlschlegel, J., Yates, J.R., 3rd, Parker, R., and Hannon, G.J. (2005). A role for the P-body component GW182 in microRNA function. Nat Cell Biol *7*, 1261-1266.

Liu, Q., Rand, T.A., Kalidas, S., Du, F., Kim, H.E., Smith, D.P., and Wang, X. (2003). R2D2, a bridge between the initiation and effector steps of the Drosophila RNAi pathway. Science *301*, 1921-1925.

Loo, S., and Rine, J. (1994). Silencers and domains of generalized repression. Science *264*, 1768-1771.

Lund, E., Guttinger, S., Calado, A., Dahlberg, J.E., and Kutay, U. (2004). Nuclear export of microRNA precursors. Science *303*, 95-98.

Ma, J.-B., Ye, K., and Patel, D.J. (2004). Structural basis for overhang-specific small interfering RNA recognition by the PAZ domain. Nature *429*, 318-322.

Ma, J.B., Yuan, Y.R., Meister, G., Pei, Y., Tuschl, T., and Patel, D.J. (2005). Structural basis for 5'-end-specific recognition of guide RNA by the A. fulgidus Piwi protein. Nature *434*, 666-670.

MacRae, I.J., Zhou, K., and Doudna, J.A. (2007). Structural determinants of RNA recognition and cleavage by Dicer. Nat Struct Mol Biol *14*, 934-940.

Macrae, I.J., Zhou, K., Li, F., Repic, A., Brooks, A.N., Cande, W.Z., Adams, P.D., and Doudna, J.A. (2006). Structural basis for double-stranded RNA processing by Dicer. Science *311*, 195-198.

Maida, Y., Yasukawa, M., Furuuchi, M., Lassmann, T., Possemato, R., Okamoto, N., Kasim, V., Hayashizaki, Y., Hahn, W.C., and Masutomi, K. (2009). An RNA-dependent RNA polymerase formed by TERT and the RMRP RNA. Nature *461*, 230-235.

Malone, C.D., Brennecke, J., Dus, M., Stark, A., McCombie, W.R., Sachidanandam, R., and Hannon, G.J. (2009). Specialized piRNA pathways act in germline and somatic tissues of the Drosophila ovary. Cell *137*, 522-535.

Margis, R., Fusaro, A.F., Smith, N.A., Curtin, S.J., Watson, J.M., Finnegan, E.J., and Waterhouse, P.M. (2006). The evolution and diversification of Dicers in plants. FEBS Lett *580*, 2442-2450.

Martinez, J., Patkaniowska, A., Urlaub, H., Luhrmann, R., and Tuschl, T. (2002). Single-stranded antisense siRNAs guide target RNA cleavage in RNAi. Cell *110*, 563-574.

Mette, M.F., Aufsatz, W., van der Winden, J., Matzke, M.A., and Matzke, A.J. (2000). Transcriptional silencing and promoter methylation triggered by double-stranded RNA. EMBO J *19*, 5194-5201.

Moazed, D. (2009). Small RNAs in transcriptional gene silencing and genome defence. Nature *457*, 413-420.

Mochizuki, K., Fine, N.A., Fujisawa, T., and Gorovsky, M.A. (2002). Analysis of a piwi-related gene implicates small RNAs in genome rearrangement in tetrahymena. Cell *110*, 689-699.

Motamedi, M.R., Hong, E.J., Li, X., Gerber, S., Denison, C., Gygi, S., and Moazed, D. (2008). HP1 proteins form distinct complexes and mediate heterochromatic gene silencing by nonoverlapping mechanisms. Mol Cell *32*, 778-790.

Motamedi, M.R., Verdel, A., Colmenares, S.U., Gerber, S.A., Gygi, S.P., and Moazed, D. (2004). Two RNAi complexes, RITS and RDRC, physically interact and localize to noncoding centromeric RNAs. Cell *119*, 789-802.

Mourrain, P., Beclin, C., Elmayan, T., Feuerbach, F., Godon, C., Morel, J.B., Jouette, D., Lacombe, A.M., Nikic, S., Picault, N.*, et al.* (2000). Arabidopsis SGS2 and SGS3 genes are required for posttranscriptional gene silencing and natural virus resistance. Cell *101*, 533-542.

Müller (1930). Types of visible variations induced by X-rays in Drosophila J Genet *22*, 299-334.

Nakayama, J., Rice, J.C., Strahl, B.D., Allis, C.D., and Grewal, S.I. (2001). Role of histone H3 lysine 9 methylation in epigenetic control of heterochromatin assembly. Science *292*, 110-113.

Napoli, C., Lemieux, C., and Jorgensen, R. (1990). Introduction of a Chimeric Chalcone Synthase Gene into Petunia Results in Reversible Co-Suppression of Homologous Genes in trans. Plant Cell *2*, 279-289.

Nishida, K.M., Saito, K., Mori, T., Kawamura, Y., Nagami-Okada, T., Inagaki, S., Siomi, H., and Siomi, M.C. (2007). Gene silencing mechanisms mediated by Aubergine piRNA complexes in Drosophila male gonad. RNA *13*, 1911-1922.

Noma, K., Sugiyama, T., Cam, H., Verdel, A., Zofall, M., Jia, S., Moazed, D., and Grewal, S.I. (2004). RITS acts in cis to promote RNA interference-mediated transcriptional and post-transcriptional silencing. Nat Genet *36*, 1174-1180.

Okamura, K., Balla, S., Martin, R., Liu, N., and Lai, E.C. (2008a). Two distinct mechanisms generate endogenous siRNAs from bidirectional transcription in Drosophila melanogaster. Nat Struct Mol Biol *15*, 581-590.

Okamura, K., Chung, W.J., Ruby, J.G., Guo, H., Bartel, D.P., and Lai, E.C. (2008b). The Drosophila hairpin RNA pathway generates endogenous short interfering RNAs. Nature *453*, 803-806.

Olsen, P.H., and Ambros, V. (1999). The lin-4 regulatory RNA controls developmental timing in Caenorhabditis elegans by blocking LIN-14 protein synthesis after the initiation of translation. Dev Biol *216*, 671-680.

Pak, J., and Fire, A. (2007). Distinct populations of primary and secondary effectors during RNAi in C. elegans. Science *315*, 241-244.

Papp, I., Mette, M.F., Aufsatz, W., Daxinger, L., Schauer, S.E., Ray, A., van der Winden, J., Matzke, M., and Matzke, A.J. (2003). Evidence for nuclear processing of plant micro RNA and short interfering RNA precursors. Plant Physiol *132*, 1382-1390.

Park, M.Y., Wu, G., Gonzalez-Sulser, A., Vaucheret, H., and Poethig, R.S. (2005). Nuclear processing and export of microRNAs in Arabidopsis. Proc Natl Acad Sci U S A *102*, 3691-3696.

Park, W., Li, J., Song, R., Messing, J., and Chen, X. (2002). CARPEL FACTORY, a Dicer homolog, and HEN1, a novel protein, act in microRNA metabolism in Arabidopsis thaliana. Curr Biol *12*, 1484-1495.

Parker, J.S., Roe, S.M., and Barford, D. (2005). Structural insights into mRNA recognition from a PIWI domain-siRNA guide complex. Nature *434*, 663-666.

Partridge, J.F., Borgstrom, B., and Allshire, R.C. (2000). Distinct protein interaction domains and protein spreading in a complex centromere. Genes Dev *14*, 783-791.

Partridge, J.F., DeBeauchamp, J.L., Kosinski, A.M., Ulrich, D.L., Hadler, M.J., and Noffsinger, V.J. (2007). Functional separation of the requirements for establishment and maintenance of centromeric heterochromatin. Mol Cell *26*, 593-602.

Partridge, J.F., Scott, K.S., Bannister, A.J., Kouzarides, T., and Allshire, R.C. (2002). cis-acting DNA from fission yeast centromeres mediates histone H3 methylation and recruitment of silencing factors and cohesin to an ectopic site. Curr Biol *12*, 1652-1660.

Petersen, C.P., Bordeleau, M.E., Pelletier, J., and Sharp, P.A. (2006). Short RNAs repress translation after initiation in mammalian cells. Mol Cell *21*, 533-542.

Pillai, R.S., Bhattacharyya, S.N., Artus, C.G., Zoller, T., Cougot, N., Basyuk, E., Bertrand, E., and Filipowicz, W. (2005). Inhibition of translational initiation by Let-7 MicroRNA in human cells. Science *309*, 1573-1576.

Pontes, O., Li, C.F., Nunes, P.C., Haag, J., Ream, T., Vitins, A., Jacobsen, S.E., and Pikaard, C.S. (2006). The Arabidopsis chromatin-modifying nuclear siRNA pathway involves a nucleolar RNA processing center. Cell *126*, 79-92.

Prehm, S., Nickel, V., and Prehm, P. (1996). A mild purification method for polysaccharide binding membrane proteins: phase separation of digitonin extracts to isolate the hyaluronate synthase from Streptococcus sp. in active form. Protein Expr Purif *7*, 343-346.

Provost, P., Silverstein, R.A., Dishart, D., Walfridsson, J., Djupedal, I., Kniola, B., Wright, A., Samuelsson, B., Radmark, O., and Ekwall, K. (2002). Dicer is required for chromosome segregation and gene silencing in fission yeast cells. Proc Natl Acad Sci U S A *99*, 16648-16653.

Rea, S., Eisenhaber, F., O'Carroll, D., Strahl, B.D., Sun, Z.W., Schmid, M., Opravil, S., Mechtler, K., Ponting, C.P., Allis, C.D.*, et al.* (2000). Regulation of chromatin structure by site-specific histone H3 methyltransferases. Nature *406*, 593-599.

Reinhart, B.J., and Bartel, D.P. (2002). Small RNAs correspond to centromere heterochromatic repeats. Science *297*, 1831.

Reinhart, B.J., Weinstein, E.G., Rhoades, M.W., Bartel, B., and Bartel, D.P. (2002). MicroRNAs in plants. Genes Dev *16*, 1616-1626.

Rouget, C., Papin, C., Boureux, A., Meunier, A.C., Franco, B., Robine, N., Lai, E.C., Pelisson, A., and Simonelig, M. (2010). Maternal mRNA deadenylation and decay by the piRNA pathway in the early Drosophila embryo. Nature *467*, 1128-1132.

Ruby, J.G., Jan, C., Player, C., Axtell, M.J., Lee, W., Nusbaum, C., Ge, H., and Bartel, D.P. (2006). Large-scale sequencing reveals 21U-RNAs and additional microRNAs and endogenous siRNAs in C. elegans. Cell *127*, 1193-1207.

Sadaie, M., Kawaguchi, R., Ohtani, Y., Arisaka, F., Tanaka, K., Shirahige, K., and Nakayama, J. (2008). Balance between distinct HP1 family proteins controls heterochromatin assembly in fission yeast. Mol Cell Biol *28*, 6973-6988.

Saetrom, P., Heale, B.S., Snove, O., Jr., Aagaard, L., Alluin, J., and Rossi, J.J. (2007). Distance constraints between microRNA target sites dictate efficacy and cooperativity. Nucleic Acids Res *35*, 2333-2342.

Saito, K., Ishizu, H., Komai, M., Kotani, H., Kawamura, Y., Nishida, K.M., Siomi, H., and Siomi, M.C. (2010). Roles for the Yb body components Armitage and Yb in primary piRNA biogenesis in Drosophila. Genes Dev *24*, 2493-2498.

Saito, K., Ishizuka, A., Siomi, H., and Siomi, M.C. (2005). Processing of pre-microRNAs by the Dicer-1-Loquacious complex in Drosophila cells. PLoS Biol *3*, e235.

Saito, K., Nishida, K.M., Mori, T., Kawamura, Y., Miyoshi, K., Nagami, T., Siomi, H., and Siomi, M.C. (2006). Specific association of Piwi with rasiRNAs derived from retrotransposon and heterochromatic regions in the Drosophila genome. Genes Dev *20*, 2214-2222.

Sasaki, T., Shiohama, A., Minoshima, S., and Shimizu, N. (2003). Identification of eight members of the Argonaute family in the human genome small star, filled. Genomics *82*, 323-330.

Schalch, T., Job, G., Noffsinger, V.J., Shanker, S., Kuscu, C., Joshua-Tor, L., and Partridge, J.F. (2009). High-affinity binding of Chp1 chromodomain to K9 methylated histone H3 is required to establish centromeric heterochromatin. Mol Cell *34*, 36-46.

Schwarz, D.S., Hutvagner, G., Du, T., Xu, Z., Aronin, N., and Zamore, P.D. (2003). Asymmetry in the assembly of the RNAi enzyme complex. Cell *115*, 199-208.

Shankaranarayana, G.D., Motamedi, M.R., Moazed, D., and Grewal, S.I. (2003). Sir2 regulates histone H3 lysine 9 methylation and heterochromatin assembly in fission yeast. Curr Biol *13*, 1240-1246.

Shanker, S., Job, G., George, O.L., Creamer, K.M., Shaban, A., and Partridge, J.F. (2010). Continuous requirement for the Clr4 complex but not RNAi for centromeric heterochromatin assembly in fission yeast harboring a disrupted RITS complex. PLoS Genet *6*, e1001174.

Sigova, A., Rhind, N., and Zamore, P.D. (2004). A single Argonaute protein mediates both transcriptional and posttranscriptional silencing in Schizosaccharomyces pombe. Genes Dev *18*, 2359-2367.

Sijen, T., Fleenor, J., Simmer, F., Thijssen, K.L., Parrish, S., Timmons, L., Plasterk, R.H., and Fire, A. (2001a). On the role of RNA amplification in dsRNA-triggered gene silencing. Cell *107*, 465-476.

Sijen, T., Steiner, F.A., Thijssen, K.L., and Plasterk, R.H. (2007). Secondary siRNAs result from unprimed RNA synthesis and form a distinct class. Science *315*, 244-247.

Sijen, T., Vijn, I., Rebocho, A., van Blokland, R., Roelofs, D., Mol, J.N., and Kooter, J.M. (2001b). Transcriptional and posttranscriptional gene silencing are mechanistically related. Curr Biol *11*, 436-440.

Singh, J., and Klar, A.J. (1992). Active genes in budding yeast display enhanced in vivo accessibility to foreign DNA methylases: a novel in vivo probe for chromatin structure of yeast. Genes Dev *6*, 186-196.

Sinkkonen, L., Hugenschmidt, T., Berninger, P., Gaidatzis, D., Mohn, F., Artus-Revel, C.G., Zavolan, M., Svoboda, P., and Filipowicz, W. (2008). MicroRNAs control de novo DNA methylation through regulation of transcriptional repressors in mouse embryonic stem cells. Nat Struct Mol Biol *15*, 259-267.

Sinkkonen, L., Hugenschmidt, T., Filipowicz, W., and Svoboda, P. (2010). Dicer is associated with ribosomal DNA chromatin in mammalian cells. PLoS One *5*, e12175.

Siomi, M.C., Sato, K., Pezic, D., and Aravin, A.A. (2011). PIWI-interacting small RNAs: the vanguard of genome defence. Nat Rev Mol Cell Biol *12*, 246-258.

Sipiczki, M. (2000). Where does fission yeast sit on the tree of life? Genome Biol *1*, REVIEWS1011.

Smardon, A., Spoerke, J.M., Stacey, S.C., Klein, M.E., Mackin, N., and Maine, E.M. (2000). EGO-1 is related to RNA-directed RNA polymerase and functions in germ-line development and RNA interference in C. elegans. Curr Biol *10*, 169-178.

Song, J.-J., Liu, J., Tolia, N.H., Schneiderman, J., Smith, S.K., Martienssen, R.A., Hannon, G.J., and Joshua-Tor, L. (2003). The crystal structure of the Argonaute2 PAZ domain reveals an RNA binding motif in RNAi effector complexes. Nat Struct Biol *10*, 1026-1032.

Song, J.J., Smith, S.K., Hannon, G.J., and Joshua-Tor, L. (2004). Crystal structure of Argonaute and its implications for RISC slicer activity. Science *305*, 1434-1437.

Song, L., Han, M.H., Lesicka, J., and Fedoroff, N. (2007). Arabidopsis primary microRNA processing proteins HYL1 and DCL1 define a nuclear body distinct from the Cajal body. Proc Natl Acad Sci U S A *104*, 5437-5442.

Stark, B.C., Kole, R., Bowman, E.J., and Altman, S. (1978). Ribonuclease P: an enzyme with an essential RNA component. Proc Natl Acad Sci U S A *75*, 3717-3721.

Steiner, F.A., Hoogstrate, S.W., Okihara, K.L., Thijssen, K.L., Ketting, R.F., Plasterk, R.H., and Sijen, T. (2007). Structural features of small RNA precursors determine Argonaute loading in Caenorhabditis elegans. Nat Struct Mol Biol *14*, 927-933.

Sugiyama, T., Cam, H., Verdel, A., Moazed, D., and Grewal, S.I. (2005). RNA-dependent RNA polymerase is an essential component of a self-enforcing loop coupling heterochromatin assembly to siRNA production. Proc Natl Acad Sci U S A *102*, 152-157.

Sugiyama, T., Cam, H.P., Sugiyama, R., Noma, K., Zofall, M., Kobayashi, R., and Grewal, S.I. (2007). SHREC, an effector complex for heterochromatic transcriptional silencing. Cell *128*, 491-504.

Tahbaz, N., Kolb, F.A., Zhang, H., Jaronczyk, K., Filipowicz, W., and Hobman, T.C. (2004). Characterization of the interactions between mammalian PAZ PIWI domain proteins and Dicer. EMBO Rep *5*, 189-194.

Tam, O.H., Aravin, A.A., Stein, P., Girard, A., Murchison, E.P., Cheloufi, S., Hodges, E., Anger, M., Sachidanandam, R., Schultz, R.M.*, et al.* (2008). Pseudogene-derived small interfering RNAs regulate gene expression in mouse oocytes. Nature *453*, 534-538.

Taricani, L., Feilotter, H.E., Weaver, C., and Young, P.G. (2001). Expression of hsp16 in response to nucleotide depletion is regulated via the spc1 MAPK pathway in Schizosaccharomyces pombe. Nucleic Acids Res *29*, 3030-3040.

Taverna, S.D., Coyne, R.S., and Allis, C.D. (2002). Methylation of histone h3 at lysine 9 targets programmed DNA elimination in tetrahymena. Cell *110*, 701-711.

Temin, H.M., and Mizutani, S. (1970). RNA-dependent DNA polymerase in virions of Rous sarcoma virus. Nature *226*, 1211-1213.

Ting, A.H., Schuebel, K.E., Herman, J.G., and Baylin, S.B. (2005). Short double-stranded RNA induces transcriptional gene silencing in human cancer cells in the absence of DNA methylation. Nat Genet *37*, 906-910.

Tomari, Y., Du, T., and Zamore, P.D. (2007). Sorting of Drosophila small silencing RNAs. Cell *130*, 299-308.

Tomari, Y., and Zamore, P.D. (2005). Perspective: machines for RNAi. Genes Dev *19*, 517-529.

Vagin, V.V., Sigova, A., Li, C., Seitz, H., Gvozdev, V., and Zamore, P.D. (2006). A distinct small RNA pathway silences selfish genetic elements in the germline. Science *313*, 320-324.

van der Krol, A.R., Mur, L.A., Beld, M., Mol, J.N., and Stuitje, A.R. (1990). Flavonoid genes in petunia: addition of a limited number of gene copies may lead to a suppression of gene expression. Plant Cell *2*, 291-299.

van Rij, R.P., Saleh, M.C., Berry, B., Foo, C., Houk, A., Antoniewski, C., and Andino, R. (2006). The RNA silencing endonuclease Argonaute 2 mediates specific antiviral immunity in Drosophila melanogaster. Genes Dev *20*, 2985-2995.

Vazquez, F., Vaucheret, H., Rajagopalan, R., Lepers, C., Gasciolli, V., Mallory, A.C., Hilbert, J.L., Bartel, D.P., and Crete, P. (2004). Endogenous trans-acting siRNAs regulate the accumulation of Arabidopsis mRNAs. Mol Cell *16*, 69-79.

Vella, M.C., Choi, E.Y., Lin, S.Y., Reinert, K., and Slack, F.J. (2004). The C. elegans microRNA let-7 binds to imperfect let-7 complementary sites from the lin-41 3'UTR. Genes Dev *18*, 132-137.

Verdel, A., Jia, S., Gerber, S., Sugiyama, T., Gygi, S., Grewal, S.I., and Moazed, D. (2004). RNAi-mediated targeting of heterochromatin by the RITS complex. Science *303*, 672-676.

Voinnet, O. (2008). Use, tolerance and avoidance of amplified RNA silencing by plants. Trends Plant Sci *13*, 317-328.

Volpe, T.A., Kidner, C., Hall, I.M., Teng, G., Grewal, S.I., and Martienssen, R.A. (2002). Regulation of heterochromatic silencing and histone H3 lysine-9 methylation by RNAi. Science *297*, 1833-1837.

Wang, Y., Juranek, S., Li, H., Sheng, G., Tuschl, T., and Patel, D.J. (2008a). Structure of an argonaute silencing complex with a seed-containing guide DNA and target RNA duplex. Nature *456*, 921-926.

Wang, Y., Juranek, S., Li, H., Sheng, G., Wardle, G.S., Tuschl, T., and Patel, D.J. (2009). Nucleation, propagation and cleavage of target RNAs in Ago silencing complexes. Nature *461*, 754-761.

Wang, Y., Sheng, G., Juranek, S., Tuschl, T., and Patel, D.J. (2008b). Structure of the guide-strand-containing argonaute silencing complex. Nature *456*, 209-213.

Watanabe, T., Takeda, A., Tsukiyama, T., Mise, K., Okuno, T., Sasaki, H., Minami, N., and Imai, H. (2006). Identification and characterization of two novel classes of small RNAs in the mouse germline: retrotransposon-derived siRNAs in oocytes and germline small RNAs in testes. Genes Dev *20*, 1732-1743.

Watanabe, T., Totoki, Y., Toyoda, A., Kaneda, M., Kuramochi-Miyagawa, S., Obata, Y., Chiba, H., Kohara, Y., Kono, T., Nakano, T.*, et al.* (2008). Endogenous siRNAs from naturally formed dsRNAs regulate transcripts in mouse oocytes. Nature *453*, 539-543.

Watson, J.D., and Crick, F.H. (1953). Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. Nature *171*, 737-738.

Wienholds, E., and Plasterk, R.H. (2005). MicroRNA function in animal development. FEBS Lett *579*, 5911-5922.

Williams, R.W., and Rubin, G.M. (2002). ARGONAUTE1 is required for efficient RNA interference in Drosophila embryos. Proc Natl Acad Sci U S A *99*, 6889-6894.

Woolcock, K.J., Gaidatzis, D., Punga, T., and Buhler, M. (2011). Dicer associates with chromatin to repress genome activity in Schizosaccharomyces pombe. Nat Struct Mol Biol *18*, 94-99.

Wu, L., Fan, J., and Belasco, J.G. (2006). MicroRNAs direct rapid deadenylation of mRNA. Proc Natl Acad Sci U S A *103*, 4034-4039.

Xie, Z., Johansen, L.K., Gustafson, A.M., Kasschau, K.D., Lellis, A.D., Zilberman, D., Jacobsen, S.E., and Carrington, J.C. (2004). Genetic and functional diversification of small RNA pathways in plants. PLoS Biol *2*, E104.

Yamada, T., Fischle, W., Sugiyama, T., Allis, C.D., and Grewal, S.I. (2005). The nucleation and maintenance of heterochromatin by a histone deacetylase in fission yeast. Mol Cell *20*, 173-185.

Yan, K.S., Yan, S., Farooq, A., Han, A., Zeng, L., and Zhou, M.M. (2003). Structure and conserved RNA binding of the PAZ domain. Nature *426*, 468-474.

Yi, R., Qin, Y., Macara, I.G., and Cullen, B.R. (2003). Exportin-5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs. Genes Dev *17*, 3011-3016.

Yigit, E., Batista, P.J., Bei, Y., Pang, K.M., Chen, C.C., Tolia, N.H., Joshua-Tor, L., Mitani, S., Simard, M.J., and Mello, C.C. (2006). Analysis of the C. elegans Argonaute family reveals that distinct Argonautes act sequentially during RNAi. Cell *127*, 747-757.

Zamore, P.D., Tuschl, T., Sharp, P.A., and Bartel, D.P. (2000). RNAi: double-stranded RNA directs the ATP-dependent cleavage of mRNA at 21 to 23 nucleotide intervals. Cell *101*, 25-33.

Zhang, H., Kolb, F.A., Jaskiewicz, L., Westhof, E., and Filipowicz, W. (2004). Single processing center models for human Dicer and bacterial RNase III. Cell *118*, 57-68.

Zilberman, D., Cao, X., and Jacobsen, S.E. (2003). ARGONAUTE4 control of locus-specific siRNA accumulation and DNA and histone methylation. Science *299*, 716-719.

Appendix

Manuscript I

Related to section 3.1.

Developmental Cell
Article

Nuclear Retention of Fission Yeast Dicer Is a Prerequisite for RNAi-Mediated Heterochromatin Assembly

Stephan Emmerth,[1,2] Heiko Schober,[1,2] Dimos Gaidatzis,[1] Tim Roloff,[1] Kirsten Jacobeit,[1] and Marc Bühler[1,*]
[1]Friedrich Miescher Institute for Biomedical Research, Maulbeerstrasse 66, CH-4058 Basel, Switzerland
[2]These authors contributed equally to this work
*Correspondence: marc.buehler@fmi.ch
DOI 10.1016/j.devcel.2009.11.011

SUMMARY

RNaseIII ribonucleases act at the heart of RNA silencing pathways by processing precursor RNAs into mature microRNAs and siRNAs. In the fission yeast *Schizosaccharomyces pombe*, siRNAs are generated by the RNaseIII enzyme Dcr1 and are required for heterochromatin formation at centromeres. In this study, we have analyzed the subcellular localization of Dcr1 and found that it accumulates in the nucleus and is enriched at the nuclear periphery. Nuclear accumulation of Dcr1 depends on a short motif that impedes nuclear export promoted by the double-stranded RNA binding domain of Dcr1. Absence of this motif renders Dcr1 mainly cytoplasmic and is accompanied by remarkable changes in gene expression and failure to assemble heterochromatin. Our findings suggest that Dicer proteins are shuttling proteins and that the steady-state subcellular levels can be shifted toward either compartment.

INTRODUCTION

In eukaryotes, small RNAs (sRNAs) function in RNA interference (RNAi) and related RNA silencing mechanisms to regulate gene expression and development, to combat viruses and mobile genetic elements, and to maintain genomic integrity (Baulcombe, 2004; Grishok et al., 2001; Huisinga and Elgin, 2009; Mochizuki et al., 2002). At least three classes of sRNAs have been identified so far: short interfering RNAs (siRNAs), microRNAs (miRNAs), and PIWI-interacting RNAs (piRNAs; Siomi and Siomi, 2009). These sRNAs associate with members of the Argonaute family of proteins, which function as the core components of a diverse set of protein-RNA complexes called RNA-induced silencing complexes (RISCs; Carmell et al., 2002; Hutvagner and Simard, 2008; Pratt and Macrae, 2009). The sRNAs guide RISCs to homologous sequences by base-pairing interactions, which commonly results in inactivation of the target sequence by various mechanisms.

Whereas sRNAs usually function posttranscriptionally, they have also been implicated in chromatin-dependent gene silencing mechanisms (CDGS; Moazed, 2009). The role of sRNAs in CDGS has been well studied in the fission yeast *Schizosaccharomyces pombe* (*S. pombe*), which contains single genes encoding the RNAi proteins Argonaute, Dicer, and RNA-dependent RNA polymerase (*ago1*$^+$, *dcr1*$^+$, and *rdp1*$^+$, respectively). Deletion of any of these genes results in loss of heterochromatic gene silencing, markedly reduced levels of H3K9 methylation (H3K9me) at centromeric repeat regions, and chromosome segregation defects (Provost et al., 2002b; Volpe et al., 2002). *S. pombe* expresses endogenous siRNAs, most of which correspond to heterochromatic regions and are found in an Argonaute-containing complex, called the RNA-induced transcriptional silencing complex (RITS) (Reinhart and Bartel, 2002; Verdel et al., 2004). Importantly, both RITS and the Rdp1-containing complex RDRC are physically linked to heterochromatin, and their interaction as well as siRNA generation depends on the histone H3K9 methyltransferase Clr4 (Cam et al., 2005; Motamedi et al., 2004; Sugiyama et al., 2005; Volpe et al., 2002). These and other observations led to a model in which the association of the RITS complex with chromatin involves base-pairing between siRNAs and chromatin-associated long noncoding RNAs. The associated RITS complex would then recruit the histone methyltransferase Clr4 to methylate H3K9. Subsequently, RITS association with chromatin is further stabilized by binding of its subunit Chp1 to H3K9me and results in high levels of H3K9me and spreading of heterochromatin (Buhler et al., 2006; Djupedal et al., 2005; Kato et al., 2005; Motamedi et al., 2004; Noma et al., 2004; Schalch et al., 2009).

The biogenesis of siRNAs in *S. pombe* involves endonucleolytic processing of double-stranded RNA (dsRNA) precursors by Dcr1. This enzyme belongs to the RNaseIII family of enzymes, which also includes the miRNA-processing enzyme Drosha (Jinek and Doudna, 2009). RNaseIII orthologs have been divided into two classes: class I, which comprises enzymes that contain a single catalytic RNaseIII and a dsRNA-binding domain (dsRBD) and which function as homodimers, and class II, which encompasses enzymes bearing two catalytic RNaseIII domains active as monomers (Jaskiewicz and Filipowicz, 2008; Macrae et al., 2006; Zhang et al., 2004). The simplest RNaseIII enzymes are found in prokaryotes and fungi and belong to class I. Drosha and Dicer RNAi proteins belong to class II and are important for the generation of miRNAs and siRNAs in animals (Bernstein et al., 2001; Hutvagner et al., 2001; Ketting et al., 2001; Lee et al., 2003). Both Dicer and Drosha proteins contain two tandemly arranged RNaseIII domains followed, in most cases, by a single dsRBD in the carboxyl terminus. In addition to these

domains, Dicer enzymes usually have an amino-terminal helicase domain, followed by a small domain of unknown function (DUF283) and a Piwi Argonaute Zwilli (PAZ) domain (Jinek and Doudna, 2009). However, this domain architecture and the requirement of individual domains for sRNA biosynthesis can differ among Dicers from different species (Jaskiewicz and Filipowicz, 2008).

Furthermore, the subcellular localization and function of Dicers can vary from organism to organism. In plants, GFP-fusions of DCL1, DCL3, and DCL4 localize to the nucleus, consistent with the nuclear role of these Dicer proteins (Hiraguri et al., 2005; Xie et al., 2004). On the other hand, human Dicer has been localized in the cytoplasm, where it processes pre-miRNAs into mature miRNAs and exogenous dsRNAs into siRNAs (Billy et al., 2001; Provost et al., 2002a). The subcellular localization of *S. pombe* Dcr1, whose function is clearly nuclear, has been ambiguous. It has been reported to localize almost exclusively to the cytoplasm (Carmichael et al., 2006; Matsuyama et al., 2006), and attempts to crosslink it to heterochromatin have not been successful (Volpe et al., 2002). On the other hand, experimental evidence has indicated that Dcr1 is a component of a self-reinforcing loop driving heterochromatin assembly *in cis* at centromeric repeats, which would favor nuclear localization of Dcr1 (Buhler and Moazed, 2007; Cam et al., 2009; Sugiyama et al., 2005).

In this study, we show that Dcr1 localizes in perinuclear foci which are associated with nuclear pores. We find that Dcr1 has the ability to shuttle between the nucleus and the cytoplasm, and we demonstrate that restriction of *S. pombe* Dcr1 to the nucleus is obligatory for the assembly of heterochromatin at centromeric repeats. Nuclear retention of Dcr1 is mediated by a 33 amino acid motif which also functions to prevent promiscuous activity of Dcr1.

RESULTS

S. pombe Dcr1 Localizes Mainly Perinuclearly

Dcr1 is essential for the biogenesis of endogenous siRNAs and the assembly of heterochromatin at centromeric repeats in *S. pombe* (Volpe et al., 2002). Conflicting data exist about subcellular localization, despite its nuclear function. Therefore, we set out to perform live-cell imaging in order to localize Dcr1 in time and space.

We generated strains expressing Dcr1 fused to GFP under the control of its own or three different versions of the thiamine-controllable nmt1 (3x, 41x, 81x) promoter. The 3x, 41x, and 81x nmt1 promoters are strong, intermediate, and weak, respectively, and they are all repressed when cells are grown in the *S. pombe* medium YES (Forsburg, 1993). Although not detectable when driven by Dcr1's own or the nmt1(81x) promoter, weak GFP-Dcr1 signals were obtained by using the nmt1(41x) promoter under repressive conditions (Figure 1A). These signals were mainly detected at the nuclear periphery (Figure 1A). Expressing GFP-Dcr1 from the stronger nmt1(3x) promoter (see Figure S1 available online) did not result in any changes in this localization pattern when the cells were grown under repressive conditions (Figure 1A). For easier detection, subsequent experiments were therefore performed with strains driving GFP-Dcr1 expression by the nmt1(3x) promoter, repressed by thiamine present in YES.

An inherent risk of live-cell imaging is a possible protein mislocalization caused by the fluorescent tag or overexpression of the fusion protein. For Dcr1, the GFP tag is unlikely to affect localization, as the same localization pattern was observed when using other fluorescent proteins (Figure 3D and data not shown). In addition, all our GFP fusion proteins were fully functional for heterochromatic gene silencing as assessed by 5-Fluoroorotic Acid (5-FOA) silencing assays and quantitative real-time RT-PCR (qRT-PCR) (Figures 1B and 1C). Importantly, we observed that massive overexpression of Dcr1 results in mainly cytoplasmic localization (Figure S1C). In contrast, mainly nuclear localization was observed if the GFP-Dcr1 fusion proteins were expressed at modestly higher levels than endogenous promoter-driven Dcr1-GFP (Figure 1A and Figure S1). From these results, we conclude that the *S. pombe* Dcr1 protein is mainly localized at the nuclear periphery.

Dcr1 Is Associated with Nuclear Pores

Intriguingly, we found GFP-Dcr1 to be enriched in perinuclear foci throughout the cell cycle (Figure 1A). Since Dcr1 is thought to act *in cis* to assemble heterochromatin, we speculated that the perinuclear foci represent Dcr1 localized at heterochromatic loci. However, this seems unlikely because we always observed more Dcr1 than Swi6 foci (Figure 2A). Importantly, Dcr1 foci did not colocalize with Cnp1, which is consistent with previous observations that Dcr1 cannot be crosslinked to centromeres (Volpe et al., 2002; M.B., unpublished data) (Figure 2B). In addition, Dcr1 localization was affected neither by *clr4* nor *swi6* deletion (Figure 2C), further indicating that Dcr1 localization is not dependent on heterochromatin. Finally, deleting genes involved in the *S. pombe* RNAi pathway, such as Ago1 or Rdp1, did not affect Dcr1 localization (Figure 2C). Together, these results show that Dcr1 localizes to the nuclear periphery irrespective of a functional RNAi or heterochromatin assembly pathway.

The punctuate rim-like staining around the nucleus observed for GFP-Dcr1 is reminiscent of the localization pattern of nuclear pore complexes (NPCs) (Bai et al., 2004). To test a possible association of Dcr1 with NPCs, we deleted the gene *nup120+* in strains expressing GFP-Dcr1. Nup120 is part of the scaffold "Y complex," and its absence leads to a NPC clustering phenotype in *S. pombe* (Figure 3A; Bai et al., 2004). We observed a clear Dcr1 clustering phenotype in *nup120Δ* cells, demonstrating that Dcr1 is associated with NPCs (Figure 3B). To assess whether Dcr1 localizes on the nuclear or cytoplasmic side of the nuclear membrane, we compared localization of the nuclear membrane marker Cut11 (West et al., 1998) and Dcr1 tagged with mCherry and GFP, respectively. This analysis revealed that Dcr1 localized more proximal to the nucleoplasm than Cut11 (Figure 3C and data not shown). The same was observed when we compared localization of Dcr1 with the NPC component Nup107 (Figures 3A and 3D; Chen et al., 2004b). This result was not a consequence of the different diffraction patterns of GFP and mCherry, because exchanging the two fluorescent tags did not change the localization pattern (Figure 3D). Thus, Dcr1 is a bona fide nuclear protein that associates with NPCs.

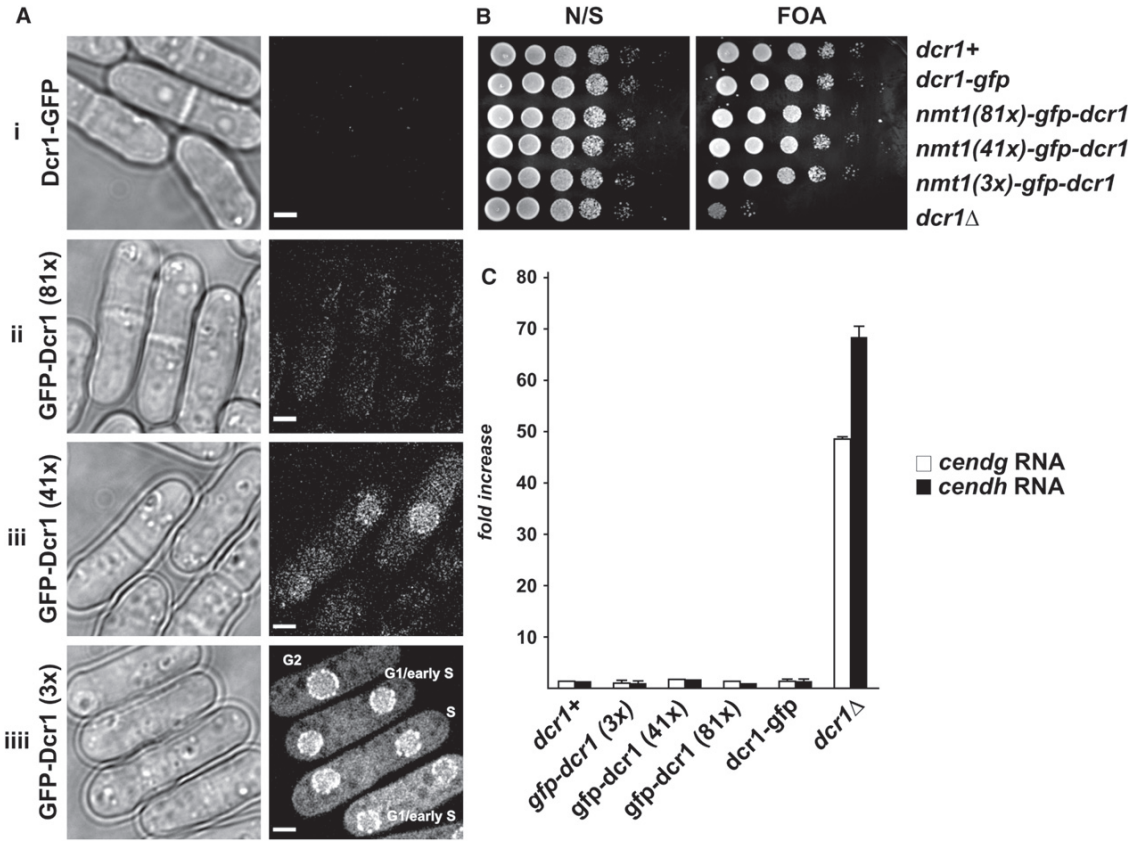


Figure 1. Dcr1 Localizes at the Nuclear Periphery

(A) Microscopy of living *S. pombe* cells expressing C-terminally GFP-tagged Dcr1 driven from the endogenous promoter (i) and N-terminally tagged GFP-Dcr1, driven from (ii) nmt1-81x, (iii) nmt1-41x, or (iiii) nmt1-3x promoters, grown under repressive conditions (i.e., in YES medium) at 30°C. For comparison of different expression levels, see western blot analysis shown in Figure S1A. Scale bars = 2 μm.

(B) Silencing assay showing that GFP-tagged Dcr1 versions shown in (A) are fully functional for heterochromatic silencing of a centromeric *ura4+* reporter (*imr1R::ura4+*).

(C) Quantitative real-time RT-PCR showing that centromeric repeat regions are efficiently silenced in the GFP-Dcr1-expressing cells. Cells were grown under the same conditions as in (A). *cendg* and *cendh* RNA levels are shown relative to *dcr1+* cells and were normalized relative to *act1+* RNA. Error bars represent standard deviations (STDEV). See also Figure S1.

The C Terminus of Dcr1 Mediates Nuclear Localization and Is Required for Heterochromatin Assembly

Fission yeast Dcr1 contains an N-terminal helicase/ATPase domain, followed by a DUF283, two RNaseIII domains, and a long C-terminal domain that bears a divergent dsRBD (Figure 4A; Colmenares et al., 2007). Importantly, the helicase, RNaseIII, and C-terminal domains are all critical for centromeric silencing and RNAi in vivo, whereas the C terminus is not required for dsRNA processing in vitro (Colmenares et al., 2007). This finding, together with the fact that nuclear localization of Dcr1 is RNAi- and heterochromatin independent, led us to speculate that specific regions of Dcr1 mediated subcellular localization and that proper localization would be necessary for centromeric siRNA generation in vivo.

One part of Dcr1 that is required for silencing in vivo but is not essential for dsRNA processing in vitro is its 103 amino acid long C terminus (C103; Colmenares et al., 2007). This C terminus contains the enzyme's dsRBD and a 33 amino acid extension (C33, Figure 4A and Figure S2). To test the possibility that C103 mediates subcellular localization of Dcr1, we generated a strain expressing GFP-tagged Dcr1 lacking C103 (GFP-Dcr1ΔC103). In contrast to full-length GFP-Dcr1, GFP-Dcr1ΔC103 localization was mainly cytoplasmic (Figures 4B and 7C). To determine the contribution of C33, we repeated this analysis on GFP-tagged Dcr1 lacking C33. Importantly, deletion of C33 resulted in the same if not even stronger loss of nuclear localization phenotype as observed for the C103 deletion without affecting protein stability (GFP-Dcr1ΔC33) (Figures 4B, 6B, and 6E; Figure S2). Loss of heterochromatic gene silencing has been previously demonstrated for Dcr1ΔC103 (Colmenares et al., 2007). Similarly, silencing of centromeric heterochromatin and the generation of centromeric siRNAs were abolished in cells expressing Dcr1ΔC33 (Figures 4C–4F and 5D–5H). The observed loss of silencing was due to a failure

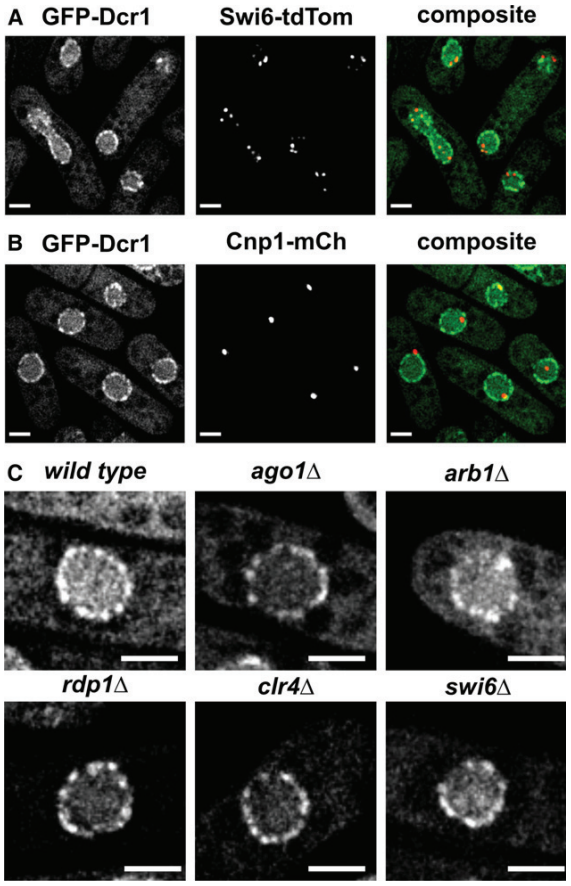


Figure 2. Dcr1 Foci Form Independently of RNAi and Heterochromatin

(A) Two-color live-cell imaging of GFP-Dcr1 with Swi6-tdTomato.

(B) Two-color live-cell imaging of GFP-Dcr1 with Cnp1-mCherry (Cnp1 is a centromere-specific histone variant).

(C) Live-cell imaging of GFP-Dcr1 in mutant backgrounds as indicated. Scale bars = 2 μm.

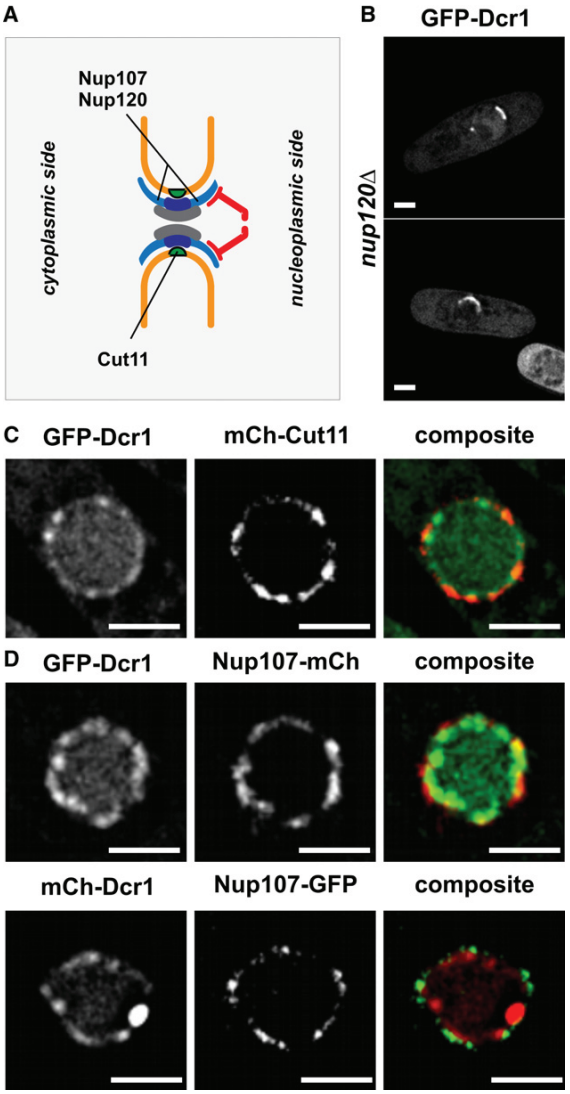


Figure 3. Dcr1 Localizes in NPC-Associated Foci at the Inner Side of the Nuclear Envelope

(A) Cartoon showing the modular assembly of the NPC (adapted from Brohawn et al. [2009]). The positions of nucleoporins shown in (C) and (D) are indicated. Nup107 and Nup120 are part of the scaffold "Y complex" (light blue). Cut11 is an integral nuclear membrane nucleoporin (West et al., 1998). Gray, FG-Nups; green, membrane attachment; red, nuclear basket; orange, nuclear membrane.

(B) GFP-Dcr1 foci cluster in *nup120Δ* cells. Nup120 is required for normal distribution of NPCs, which have been shown to cluster in *nup120Δ* cells (Bai et al., 2004).

(C) Two-color live-cell imaging of GFP-Dcr1 with Cut11-mCherry.

(D) Two-color live-cell imaging of GFP-Dcr1 with Nup107-mCherry (upper panels) or mCherry-Dcr1 with Nup107-GFP (lower panels). Scale bars = 2 μm.

in assembly of heterochromatin at centromeric repeats because H3K9 methylation was also affected in *dcr1ΔC33* cells (Figure 4D). Thus, we have identified a short C-terminal motif (C33), which mediates nuclear localization of Dcr1 and is required for the assembly of heterochromatin at centromeric repeats.

Gene Expression Profiles Change in the Absence of Nuclear Dcr1

Interestingly, we observed that cells expressing Dcr1ΔC33 grew more slowly, are less viable than wild-type (Figure 4C; Figures S2D and S2E), and show morphological abnormalities (data not shown). Importantly, these phenotypes could not be observed in *dcr1Δ* cells, suggesting that deletion of C33 created a gain-of-function allele. To test this, we hybridized total RNA from wild-type, *dcr1*Δ, and *dcr1ΔC33* cells to *S. pombe* tiling arrays. Consistent with previous studies (Hansen et al., 2005; Provost et al., 2002b), we observed that RNA levels, except for centromeric repeat transcripts, were not much different from wild-type in *dcr1Δ* cells. In contrast, many more genes were differentially expressed in *dcr1ΔC33* than in *dcr1Δ* cells when compared to wild-type (Figures 5A and 5B).

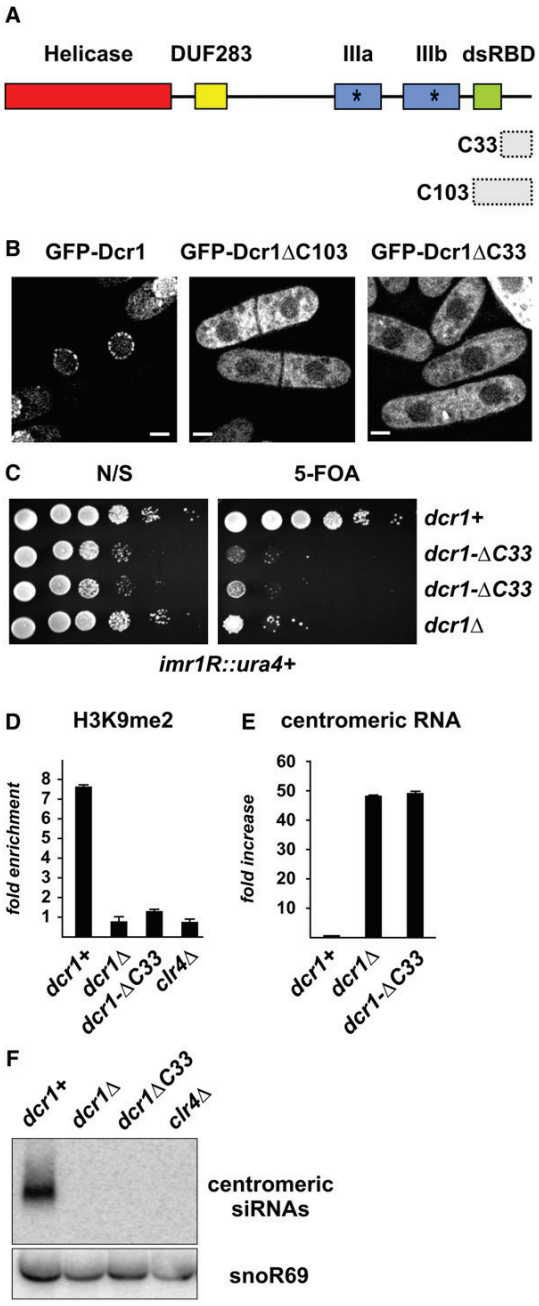


Figure 4. The 33 Most C-Terminal Amino Acids Restrict Dcr1 to the Nucleus and Are Essential for H3K9 Methylation and Silencing

(A) Domain architecture of *S. pombe* Dcr1 (not drawn to scale). C33 and C103 indicate the last 33 and 103 amino acids of Dcr1, respectively.

(B) Live-cell imaging of full-length and C-terminal truncations of GFP-Dcr1. Scale bars = 2 μm.

(C) Silencing assay with Dcr1ΔC33 expressed from its endogenous promoter showing that the truncated protein cannot silence a heterochromatic *ura4+* reporter gene (*imr1R::ura4+*).

(D) ChIP experiment showing that the C33-truncated version of Dcr1 expressed from its endogenous promoter loses H3K9me2 of a heterochromatic *ura4+* transgene (*imr1R::ura4+*). Fold-enrichment values from one representative experiment, normalized to *act1+*, are shown. The value for *clr4Δ* cells was set to 1.

(E) Quantitative real-time RT-PCR with the same strains as in (D) showing that the C33-truncated version of Dcr1 cannot silence *cendh* centromeric repeats. The value for wild-type cells (*dcr1+*) was set to 1. (D and E) Error bars represent standard deviations (STDEV).

(F) Northern blot analysis performed with total RNA preparations enriched for RNAs <200 nt from the same strains as in (D). The membrane was probed with 5′ end-labeled DNA oligos specific for centromeric siRNAs, and the loading control snoR69. See also Figure S2.

The observation that *dcr1ΔC33* and *dcr1Δ* gene expression profiles differ (Figures 5A–5C and 7E) is consistent with the idea that *dcr1ΔC33* is a gain-of-function allele and may be responsible for the slow growth of this mutant. It is possible that this is a consequence of localizing Dcr1 to the cytoplasm, where it might promiscuously process a variety of RNA substrates. If this were the case, one would expect to identify siRNAs specific to *dcr1ΔC33* cells. Therefore, we generated small RNA libraries from total RNA isolations of wild-type, *dcr1Δ*, and *dcr1ΔC33* cells and subjected them to high-throughput sequencing. As expected from the northern blot shown in Figure 4F, only very few reads from either *dcr1Δ* or *dcr1ΔC33* cells mapped to centromeric repeats (Figures 5D and 5E). Interestingly, almost no reads could be found which were specific to *dcr1ΔC33* cells, and the small RNA profiles for *dcr1Δ* and *dcr1ΔC33* appeared very similar (Figures 5D and 5F–5H), suggesting that Dcr1ΔC33 can affect gene expression independently of its dsRNA processing activity. This was confirmed by our finding that the *dcr1ΔC33*-specific gene expression profile did not revert to a *dcr1Δ* profile upon inactivation of the RNaseIII active sites (dcr1ΔC33**) (Figure 5C; Colmenares et al., 2007). Thus, we conclude that Dcr1 can influence gene expression independently of its catalytic activity.

Dcr1ΔC33 Is a Shuttling Protein

Relocation of C-terminally truncated Dcr1 (GFP-Dcr1ΔC33) from the nucleus to the cytoplasm led us to hypothesize that C33 could act as a nuclear localization signal (NLS). To directly test this, we fused C33 to a GFP-LacZ reporter construct, which had previously been shown to localize throughout the cell (Dang and Levin, 2000). In contrast to the SV40 NLS, *S. pombe* C33 did not lead to an enhanced nuclear localization of the GFP-LacZ reporter (Figure 6A). Therefore, C33 is not sufficient for nuclear localization, and, hence, is unlikely to be an NLS.

Surprisingly, although the SV40 NLS was sufficient for nuclear accumulation of a GFP-LacZ reporter (Figure 6A), this sequence, added either C- or N-terminally, was not able to restore nuclear localization of GFP-Dcr1ΔC33 (Figures 6B and 6E and data not shown). Moreover, consistent with our observations above that note the importance of nuclear Dcr1 localization for silencing, the addition of a C-terminal NLS to Dcr1ΔC33 was unable to restore silencing when expressed from Dcr1's endogenous promoter (Figure 6C). However, strong overexpression of Dcr1ΔC33 fully rescued loss of silencing of a centromeric *ura4*$^+$ reporter (Figure S2B). Furthermore, we observed partial rescue of heterochromatic gene silencing upon mild overexpression of Dcr1ΔC33, which was further improved by adding an

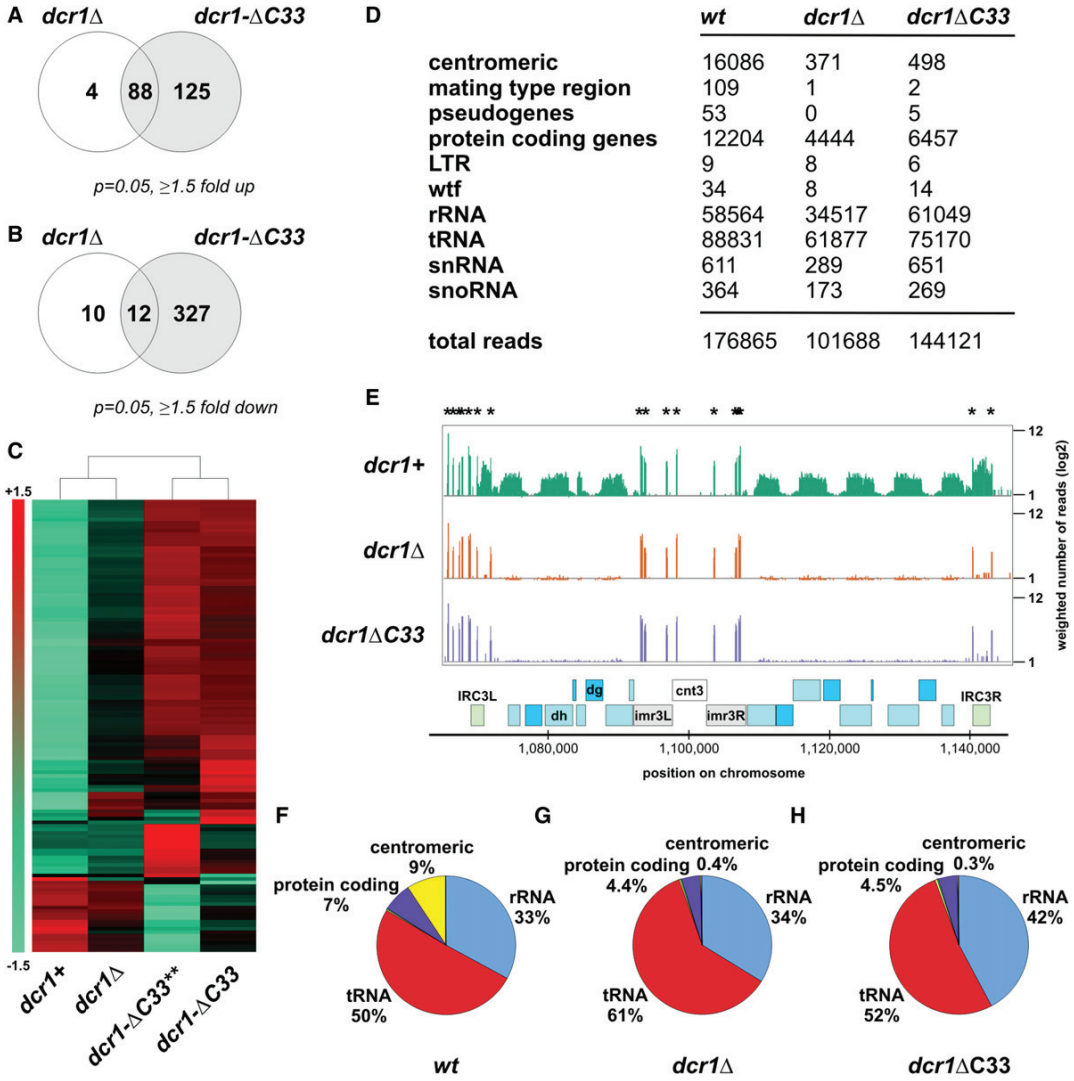


Figure 5. Gene Expression- and siRNA-Profiles in Different Dcr1 Mutants

(A) Venn diagrams showing the number of genes upregulated at least 1.5-fold in *dcr1Δ* or *dcr1ΔC33* cells compared to wild-type (p = 0.05) on *S. pombe* tiling arrays.

(B) Venn diagrams showing the number of genes downregulated at least 1.5-fold in *dcr1Δ* or *dcr1ΔC33* cells compared to wild-type (p = 0.05) on *S. pombe* tiling arrays. (A and B) Six biological replicates were analyzed.

(C) Heatmap displaying the genes which were up- or downregulated at least 2-fold (p = 0.05) in *dcr1ΔC33* cells compared to wild-type on *S. pombe* tiling arrays. Artificially scaled expression values are shown for the strains indicated (−1.5 is set for the gene with the lowest expression, and +1.5 is set for the gene with the highest expression). Two or more biological replicates were performed.

(D) Classification of the sequence reads obtained from deep-sequencing of small RNA libraries derived from wild-type, *dcr1Δ*, and *dcr1ΔC33* cells. The numbers indicate the weighted number of sequenced reads for every class.

(E) Sequence reads mapping to the centromere of chromosome 3 displayed on a genome browser (Affymetrix) for wild-type, *dcr1ΔC33*, and *dcr1Δ* cells. The y axis displays the log(2) of the weighted number of reads. Asterisks indicate tRNA genes.

(F–H) Pie chart illustrating percentages for the individual small RNA classes relative to the total amount of small RNAs sequenced for each condition.

SV40 NLS (Figure 6D). These results strongly suggested that the SV40 NLS did promote nuclear import of Dcr1ΔC33, although it never accumulated to high levels in the nucleus due to faster export kinetics. To directly test this, we performed fluorescence loss in photobleaching (FLIP) experiments, where we bleached the cytoplasm of cells expressing GFP-Dcr1ΔC33 or NLS-GFP-Dcr1ΔC33 while measuring fluorescence in the nucleus. Cytoplasmic bleaching resulted in a rapid decrease of nuclear fluorescence, demonstrating that GFP-Dcr1ΔC33 and NLS-GFP-Dcr1ΔC33 are shuttling proteins (Figures 6E and 6F).

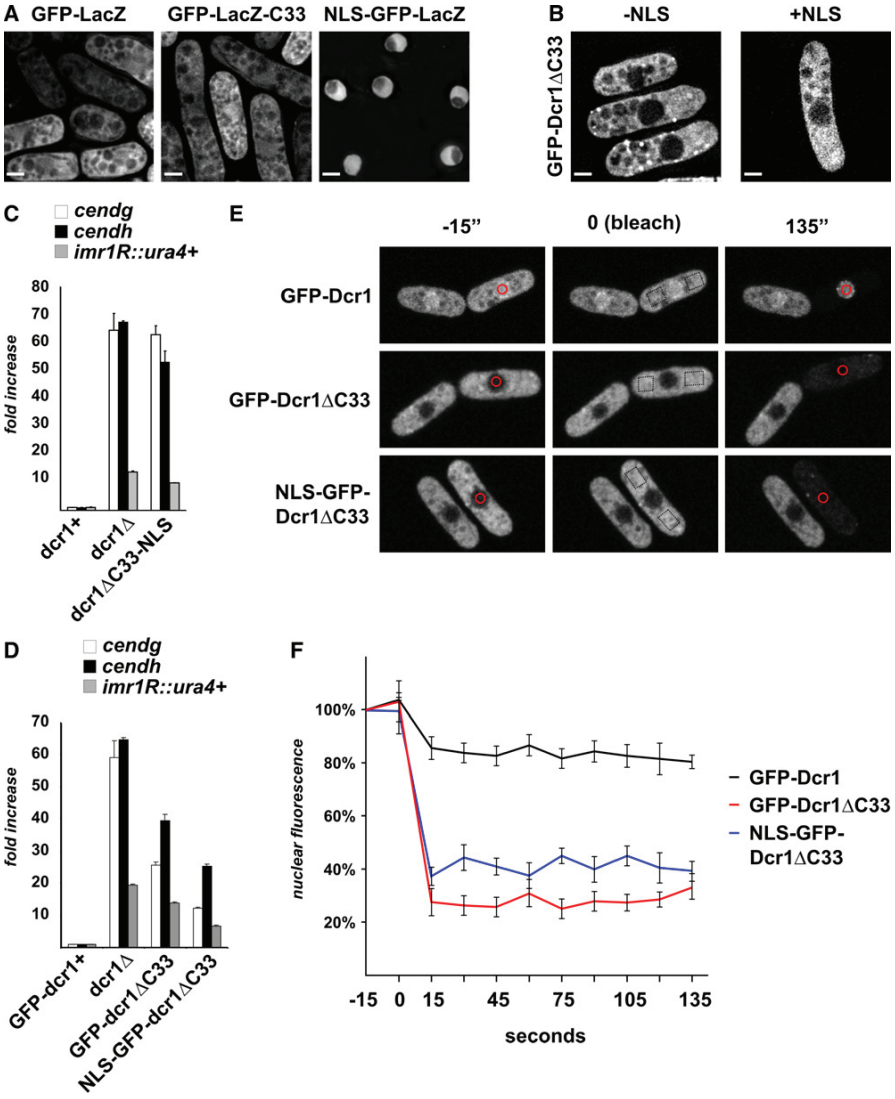


Figure 6. Dcr1ΔC33 Is a Shuttling Protein

(A) Live-cell imaging of wild-type strains transformed with the indicated plasmids.

(B) Live-cell imaging of GFP-Dcr1ΔC33- and NLS-GFP-Dcr1ΔC33-expressing cells. Scale bars = 2 μm.

(C) Quantitative real-time RT-PCR showing that Dcr1ΔC33-NLS fails to silence centromeric repeats when expressed from its endogenous promoter.

(D) Quantitative real-time RT-PCR showing that moderate overexpression of GFP-Dcr1ΔC33 or NLS-GFP-Dcr1ΔC33 results in partial rescue of silencing. *cendg*, *cendh*, and *imr1R::ura4+* RNA levels are shown relative to *dcr1+* cells and were normalized relative to *act1+* RNA. (C and D) Error bars represent standard deviations (STDEV).

(E and F) Nuclear fluorescence loss in (cytoplasmic) photobleaching (FLIP) of different GFP-tagged Dcr1 alleles. (E) Representative images of FLIP. (F) The mean fluorescence of the nucleus was determined (illustrated by red circle in [E], Ø~1 μm) 15 s before, immediately before, and nine times after photobleaching the cytoplasm in 15 s intervals (rectangle = bleach area; bleach iterations = 50). Each value represents the average of ≥3 individual recordings. Error bars represent standard error of the mean (StDev/SQRT(n)). See also Figure S4.

In conclusion, our results demonstrate that fission yeast Dcr1 has the ability to shuttle between the nucleus and the cytoplasm and is very efficiently exported to the cytoplasm if the 33 most C-terminal amino acids are missing. Therefore, rather than acting as an NLS, C33 functions as a nuclear retention signal by inhibiting nucleocytoplasmic export of Dcr1.

The Dcr1 dsRBD Mediates Nucleocytoplasmic Trafficking

Our results suggested that Dcr1 contains nuclear import as well as export signals, although we were not able to predict them. Interestingly, several proteins harboring dsRBDs exhibit nucleocytoplasmic shuttling, and a dsRBD itself can mediate import

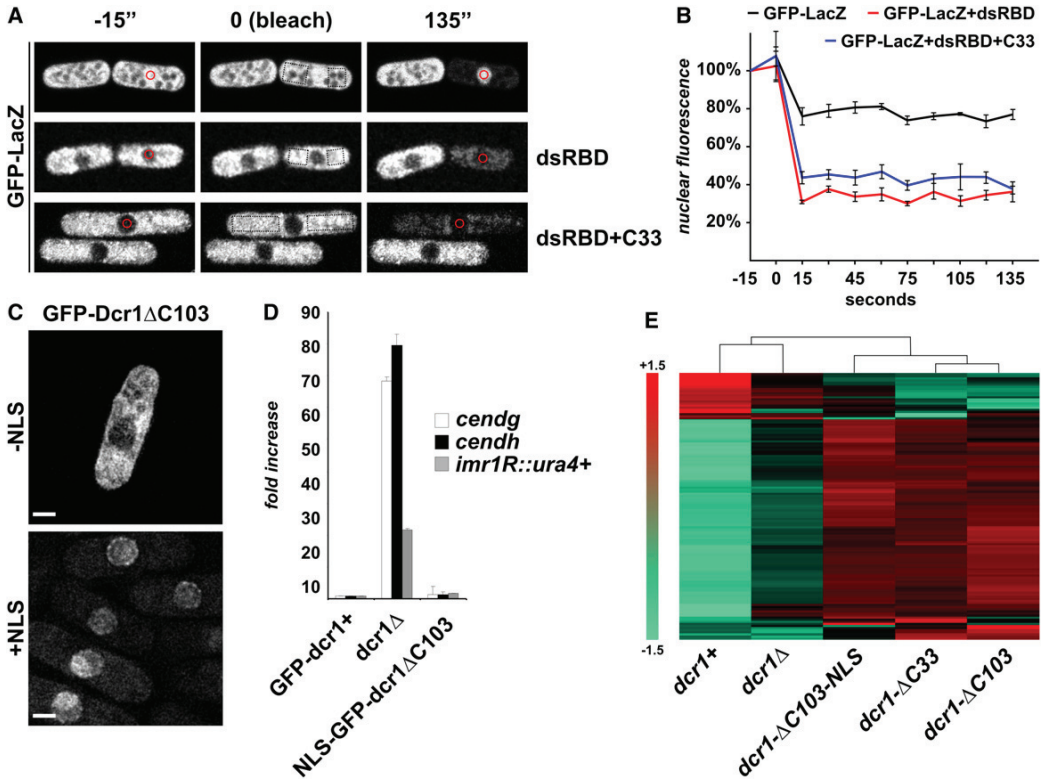


Figure 7. The Dcr1 dsRBD Mediates Nucleocytoplasmic Trafficking

(A and B) FLIP was performed with cells expressing GFP-lacZ, GFP-lacZ-dsRBD, and GFP-lacZ-dsRBD+C33 as in Figure 6.

(C) Live-cell imaging from cells expressing GFP-Dcr1ΔC103 or NLS-GFP-Dcr1ΔC103. Scale bars = 2 μm.

(D) Quantitative real-time RT-PCR to detect heterochromatic RNAs from the strains indicated. *cendg, cendh*, and *imr1R::ura4+* RNA levels are shown relative to *GFP-dcr1+* cells and were normalized to *act1+* RNA. Error bars represent standard deviations (STDEV).

(E) Heatmap displaying the genes which were up- or downregulated at least 2-fold (p = 0.05) in the indicated strains compared to wild-type on *S. pombe* tiling arrays. Artificially scaled expression values are shown for the strains indicated (−1.5 is set for the gene with the lowest expression, and +1.5 is set for the gene with the highest expression). Two to six biological replicates were performed. See also Figure S3.

and/or export (Chen et al., 2004a; Fritz et al., 2009; Gwizdek et al., 2004; Macchi et al., 2004; Strehblow et al., 2002). This led us to test whether the dsRBD of Dcr1 might generally function in nucleocytoplasmic trafficking. Indeed, adding the dsRBD of Dcr1 to a GFP-LacZ reporter resulted in strongly reduced nuclear signal (Figure 7A). This residual nuclear fluorescence was further decreased upon cytoplasmic photobleaching, whereas nuclear fluorescence in GFP-LacZ-expressing cells was not greatly affected (Figures 7A and 7B). This result verified that the dsRBD can function as an export signal. Importantly, C33 did not greatly affect export of GFP-LacZ-dsRBD (dsRBD+C33, Figures 7A and 7B). Thus, C33 can only function efficiently as a nuclear retention signal in the context of full-length Dcr1.

If the dsRBD operates solely as an export signal, GFP-Dcr1ΔC103 would be expected to localize in the nucleus. However, we observed only slightly more nuclear staining for GFP-Dcr1ΔC103 compared to GFP-Dcr1ΔC33 (Figures 4B and 7C and data not shown), and a wild-type Dcr1 localization phenotype could only be restored after adding an SV40 NLS (Figure 7C). Furthermore, nuclear fluorescence was not greatly affected in FLIP experiments performed with cells expressing GFP-Dcr1ΔC103 (Figure S3). These results pointed out that the dsRBD also functions in nuclear import, likely in cooperation with another putative nuclear import-promoting element on Dcr1. Therefore, the dsRBD mediates shuttling of Dcr1 which is under the control of C33.

C33 Prevents Promiscuous Activity of Dcr1

Adding an NLS to GFP-Dcr1ΔC103 restored the formation of perinuclear foci (Figure 7C), showing that neither the dsRBD nor C33 is required for this localization pattern. Expression of NLS-GFP-Dcr1ΔC103 also fully restored heterochromatic gene silencing (Figure 7D), demonstrating that nuclear localization of Dcr1 lacking a dsRBD is sufficient for RNAi-mediated heterochromatin assembly. In contrast, gene expression profiles for Dcr1ΔC103, either with or without an SV40 NLS, were not much different from Dcr1ΔC33 (Figure 7E). Thus, lacking C33 per se seems to be responsible for the observed gain-of-function phenotype, independent of subcellular localization or dsRNA processing activity. Therefore, we conclude that C33, in addition to functioning as a nuclear retention signal, also prevents Dcr1 from functioning promiscuously.

DISCUSSION

In this study, we have analyzed the subcellular localization of the fission yeast Dcr1 enzyme and found that it accumulates in perinuclear foci associated with NPCs throughout the cell cycle. Our findings highlight that proper subcellular localization of Dcr1 is obligatory for RNAi-mediated heterochromatin assembly in *S. pombe* and reveal mechanistic insights into how this is achieved. Below, we discuss the implications of these findings for the mechanism of RNAi-mediated heterochromatin formation and the subcellular organization of RNA silencing.

Nuclear RNAi

Unlike in animals, it is not clear whether siRNAs in *S. pombe* also function in classical, cytoplasmic posttranscriptional gene silencing (PTGS) (Buhler et al., 2006; Sigova et al., 2004). Rather, RNA silencing seems to be functionally restricted to the nucleus in *S. pombe* (Moazed, 2009). Therefore, earlier findings that Dcr1 is predominantly cytoplasmic were difficult to reconcile with current models of RNAi-mediated heterochromatin assembly in fission yeast (Carmichael et al., 2006; Matsuyama et al., 2006). We found that massive overexpression of Dcr1 causes growth defects and strong cytoplasmic staining (Figure S1 and data not shown), which may explain why cytoplasmic localization has been previously reported for Dcr1 (Figure S1; Carmichael et al., 2006; Matsuyama et al., 2006). When Dcr1 was expressed at more moderate levels, we observed a clear nuclear accumulation. Such localization is required for proper centromeric siRNA generation and heterochromatin assembly. Thus, we propose that heterochromatic siRNA biogenesis occurs in the nucleus. We cannot, at present, rule out additional Dcr1 functions in the cytoplasm.

Current models for RNAi-mediated heterochromatin assembly in *S. pombe* propose that siRNAs are generated *in cis*, i.e., at heterochromatic repeats at centromeres. Importantly, we have not been able to demonstrate a persistent colocalization of Dcr1 with the centromere-specific histone variant Cnp1, raising the question of whether siRNA generation might occur elsewhere in the nucleus. Consistent with such a model, neither Dcr1 nor components of the ARC complex, which receives the double-stranded siRNAs from Dcr1, can be crosslinked to chromatin (Buker et al., 2007; Volpe et al., 2002). It is possible that siRNA generation might occur in a specialized nuclear compartment, possibly in the observed peripheral Dcr1 foci. We note that the formation of these foci depends neither on the dsRBD nor C33 (Figure 7C). Thus, neither shuttling nor dsRNA binding is required for targeting Dcr1 to the nuclear periphery.

The Dcr1 foci are reminiscent of subnuclear bodies described in *Arabidopsis thaliana,* although these have been found to be associated with the nucleolus (Fang and Spector, 2007; Li et al., 2006, 2008; Pontes et al., 2006; Song et al., 2007). Plant Cajal- and AB-bodies are rich in RNA-directed DNA methylation (RdDM)-specific proteins and appear to be required to fully silence genes by RdDM (Li et al., 2008). Furthermore, in contrast to animals, miRNA biogenesis in plants is carried out solely in the nucleus in so-called dicing (D-) bodies. It has been proposed that the formation of D-bodies in *Arabidopsis* is crucial for the generation of miRNAs (Fang and Spector, 2007; Papp et al., 2003). The perinuclear Dcr1 foci observed in *S. pombe* resemble such D-bodies, as only Dcr1 and none of the other RNAi proteins exhibited this localization pattern (data not shown). However, Dcr1 foci are distinct from plant D-bodies in that they are associated with NPCs.

In summary, we speculate that siRNA biogenesis might occur in discrete nuclear bodies in *S. pombe*. Such compartmentalization could ensure proper assembly of silencing complexes and prevent illegitimate events from occurring elsewhere, thereby increasing the fidelity of RNA silencing. However, the relevance of Dcr1 foci for a functional RNAi-mediated heterochromatin assembly pathway and the exact mechanism of their formation demand further study.

Nucleocytoplasmic Shuttling of Dicer

Our study revealed that Dicer proteins may have the ability to shuttle between the nucleus and the cytoplasm. This shuttling is mediated by the dsRBD and can be influenced by additional sequences. In *S. pombe*, this additional sequence is contained in the extended C terminus (C33) of Dcr1 and provides a means of regulation of its nucleocytoplasmic distribution. It will be interesting to see whether cellular localization of other RNaseIII enzymes with extended C termini is regulated in a way similar to what we observe for fission yeast Dcr1.

Intriguingly, nucleocytoplasmic shuttling that is dependent on the dsRBD can also be observed for human Dicer (M. Doyle and W. Filipowicz, personal communication). Strikingly, only Dcr1 but not human Dicer has an extended C terminus, which may explain why human Dicer localizes mainly to the cytoplasm (Billy et al., 2001; Provost et al., 2002a). Based on these results, we propose that Dicer proteins in general might have the ability to shuttle between the nucleus and the cytoplasm and that the steady-state subcellular levels can be shifted toward either compartment. This regulation could be achieved by interacting proteins or *cis*-motifs that either promote or inhibit nuclear export/import. For *S. pombe* Dcr1, we favor the second possibility. We suggest that C33 might influence the structural conformation of Dcr1 such that the export-promoting features of the dsRBD are masked, resulting in nuclear retention of Dcr1, an intriguing hypothesis that needs to be directly tested by future structural work. Future studies will also be required to elucidate the precise import/export pathways that mediate Dicer shuttling. We note that this is neither Crm1- nor Exp5 dependent in *S. pombe*, as leptomycin B (LMB) treatment and/or deletion of a putative exportin 5 homolog (SPAC328.01c) did not noticeably affect the subcellular localization of Dcr1ΔC33 (Figure S4).

EXPERIMENTAL PROCEDURES

Strains and Plasmids

Fission yeast strains used in this study are described in Supplemental Experimental Procedures and were grown at 30°C in YES. If transformed with plasmids (see Supplemental Experimental Procedures), strains were grown in EMMc-leu medium containing 5 μg/ml thiamine. All strains were constructed following a standard PCR-based protocol (Bahler et al., 1998). GFP-LacZ constructs were cloned by fusion PCR-based assembly of the inserts and subsequent PacI/NotI ligation into pREPN-3xFLAG (Buker et al., 2007). pREP1-3xFLAG-Dcr1ΔC33 was constructed by exchanging Dcr1 with Dcr1ΔC33 by BglII/NotI cloning into pREP1-3xFLAG-Dcr1 (Colmenares et al., 2007). Primer sequences used for cloning are available upon request. Constructs on plasmids and in yeast strains were confirmed by sequencing.

Yeast Live-Cell Fluorescence Microscopy

If not specified differently, *S. pombe* precultures were grown in YES (sterile filtered components only) for 8 hr at 30°C, diluted to 10^5 cells/ml, and grown for 14–16 hr at 30°C to a concentration of about 5×10^6 cells/ml. Microscopy was performed on cells spread on agarose patches containing YES medium with 3% glucose. Images were captured on a Delta Vision built of an Olympus IX70 widefield microscope equipped with a CoolSNAP HQ2/ICX285 camera. Image stacks of 12–15 μm Z-distance were acquired with a Z-step size of 0.2 μm and deconvolved using the softworks (Delta Vision) software.

Confocal Microscopy and FLIP

Cells expressing the plasmid-encoded fusion proteins GFP-lacZ, GFP-lacZ+dsRBD, and GFP-lacZ+dsRBD+C33 were grown in PMGc-Leu supplemented with 1 μg/ml thiamine for 6–8 hr. From this preculture, cells were diluted into fresh medium to a concentration of 7×10^5 cells/ml and grown for 14 hr. Strains with genomic modifications of the *dcr1+* locus were grown in YES for 6–8 hr. This preculture was diluted to $0.5–1\times10^5$ cells/ml into fresh medium and grown for 14 hr. G2 cells were imaged in liquid media using a ludin chamber and *Bandeiraea simplicifolia* (Sigma L7508, 1 mg/ml solution) coated coverslips. Bleaching was performed on a Zeiss LSM510 microscope. The pinhole was set to 1 airy unit corresponding to 0.7 μm optical thickness. The mean fluorescence of nuclei was determined using the ImageJ software. The values were normalized to control neighboring nuclei and to the first time point.

RNA Isolation and cDNA Synthesis for qRT-PCR

Cells were harvested at OD600 = 0.5, washed once in water, and flash frozen in liquid nitrogen. RNA was isolated using the Absolutely RNA Miniprep Kit from Stratagene (#400805) with the following modification: 600 μl lysis buffer, 4.2 μl Mercaptoethanol, and 500 μl glass beads were added to the pellet, and cells were bead-beaten for 1 min. The cleared lysate was collected, and RNA isolation was continued according to the supplier's manual. cDNA was synthesized using random hexamers with the AffinityScript MultipleTemperature cDNA Synthesis Kit from Stratagene (#200436).

Quantitative RT-PCR/PCR

Quantitative PCR was performed on a 7500 Fast Real-Time PCR System (Applied Biosystems #4351106) using the Fast SYBR Green Master Mix (Applied Biosystems #4385612). Relative RNA levels were calculated from C_T values according to the ΔC_T method and normalized to *act1+* mRNA levels. Primer pairs used for PCR reactions can be found in the Supplemental Experimental Procedures.

Silencing Assays

Serial 10-fold dilutions of the strains indicated were plated on PMGc (nonselective, NS) or on PMGc plates containing 2 mg/ml 5-FOA.

Chromatin Immunoprecipitation and Northern Blot

Chromatin immunoprecipitation was performed as described in Buhler et al. (2006). An antibody against dimethylated H3-K9 (abcam, #ab1220) was used. Centromeric siRNAs were detected by northern blotting as described in Buhler et al. (2006).

S. pombe Tiling Arrays

RNA was isolated from cells collected at OD600 = 0.5 using the hot phenol method (Leeds et al., 1991). The isolated RNA was processed according to the GeneChip Whole Transcript (WT) Double-Stranded Target Assay Manual from Affymetrix using the GeneChip *S. pombe* Tiling 1.0FR. For analysis of the tiling arrays, an R-based script was used, which is available upon request. We used the genome and annotations from the *S. pombe* Genome project (ftp://ftp.sanger.ac.uk/pub/yeast/pombe/GFF/pombe_160708.fasta; ftp://ftp.sanger.ac.uk/pub/yeast/pombe/GFF/pombe_160708.gff). The oligos from the Affymetrix .BPMAP file were remapped using bowtie, and the .GFF file was used to map them to the genes. The resulting .CDF file is available upon request.

Generation of Small RNA Libraries for High-Throughput Sequencing

Total RNA was isolated from cells harvested at OD600 = 0.5 using the hot phenol method (Leeds et al., 1991) and subjected to size fractionation using RNeasy Midi columns (QIAGEN) as previously described (Buhler et al., 2006).

~17–30 nt small RNAs were PAGE purified and were cloned based upon the preactivated, adenylated linkering method described previously (Lau et al., 2001) using a mutant T4 RNA ligase ($Rnl2_{1-249}$) (Ho et al., 2004). All samples were barcoded at the 3′ end of the 5′ adaptor using a hamming distance two code with a 3′ cytosine (AAAC, ACCC, AGGC, ATTC, CACC, CCGC, CGTC, CTAC, GAGC, GCTC, GGAC, GTCC, TATC, TCAC, TGCC, TTGC) and sequenced in one lane of an Illumina GA2 instrument.

In Silico Analysis of Sequencing Data

Individual reads were assigned to their sample based on the first four nucleotides containing the barcode. The 3′ adaptor was removed by aligning it to the read allowing one or two mismatches in prefix alignments of at least seven or ten bases, respectively. Low-complexity reads were filtered out based on their dinucleotide entropy (removing <1% of the reads). All the reads that were shorter than 14 nucleotides were removed. Alignments to the *S. pombe* genome (May 8, 2009, http://www.sanger.ac.uk/Projects/S_pombe/) were performed by the software bowtie (version 0.9.9.1) (Langmead et al., 2009) with parameters -v 2 -a -m 100, tracking up to 100 best alignment positions per query and allowing at most two mismatches. To track genomically untemplated hits (e.g., exon-exon junctions or missing parts in the current assembly), the reads were also mapped to an annotation database containing known sequences (http://www.sanger.ac.uk/Projects/S_pombe/). In that case, all best hits with at most two mismatches were tracked. Each alignment was weighted by the inverse of the number of hits. In the cases where a read had more hits to an individual sequence from the annotation database than to the whole genome, the former number of hits was selected to ensure that the total weight of a read does not exceed one. Genomic read coverage plots were based on weighted alignments.

ACCESSION NUMBERS

Tiling array and deep sequencing data are deposited at GEO (http://www.ncbi.nlm.nih.gov/geo, accession numbers GSE18440 and GSE18582, respectively).

SUPPLEMENTAL INFORMATION

Supplemental Information includes four figures and Supplemental Experimental Procedures and can be found with this article online at doi:10.1016/j.devcel.2009.11.011.

ACKNOWLEDGMENTS

We thank Witold Filipowicz and Michael Doyle for valuable discussions. We are grateful to Christian Beisel, Edward Oakeley, and Michael Stadler for their contribution to setting up deep-sequencing of small-RNA libraries; to Hans-Rudolf Hotz for bioinformatics support; to Jens Rietdorf and Laurent Gelman for imaging support; to Johannes Buchberger, Helge Grosshans, Fred Meins, Olivia Rissland, Florian Steiner, and all members of the Bühler lab for critical reading of the manuscript; to Danesh Moazed, Frank Neumann, and Valérie Doye for strains; and to Yukiko Shimada and Nathalie Laschet for technical support. We deeply appreciate the error-prone oligonucleotide synthesis of Sigma-Aldrich. This work was supported by the Swiss National Science Foundation. The Friedrich Miescher Institute for Biomedical Research is supported by the Novartis Research Foundation.

Received: August 27, 2009
Revised: October 24, 2009
Accepted: November 25, 2009
Published: January 19, 2010

REFERENCES

Bahler, J., Wu, J.Q., Longtine, M.S., Shah, N.G., McKenzie, A., III, Steever, A.B., Wach, A., Philippsen, P., and Pringle, J.R. (1998). Heterologous modules

for efficient and versatile PCR-based gene targeting in Schizosaccharomyces pombe. Yeast *14*, 943–951.

Bai, S.W., Rouquette, J., Umeda, M., Faigle, W., Loew, D., Sazer, S., and Doye, V. (2004). The fission yeast Nup107-120 complex functionally interacts with the small GTPase Ran/Spi1 and is required for mRNA export, nuclear pore distribution, and proper cell division. Mol. Cell. Biol. *24*, 6379–6392.

Baulcombe, D. (2004). RNA silencing in plants. Nature *431*, 356–363.

Bernstein, E., Caudy, A.A., Hammond, S.M., and Hannon, G.J. (2001). Role for a bidentate ribonuclease in the initiation step of RNA interference. Nature *409*, 363–366.

Billy, E., Brondani, V., Zhang, H., Muller, U., and Filipowicz, W. (2001). Specific interference with gene expression induced by long, double-stranded RNA in mouse embryonal teratocarcinoma cell lines. Proc. Natl. Acad. Sci. USA *98*, 14428–14433.

Brohawn, S.G., Partridge, J.R., Whittle, J.R., and Schwartz, T.U. (2009). The nuclear pore complex has entered the atomic age. Structure *17*, 1156–1168.

Buhler, M., and Moazed, D. (2007). Transcription and RNAi in heterochromatic gene silencing. Nat. Struct. Mol. Biol. *14*, 1041–1048.

Buhler, M., Verdel, A., and Moazed, D. (2006). Tethering RITS to a nascent transcript initiates RNAi- and heterochromatin-dependent gene silencing. Cell *125*, 873–886.

Buker, S.M., Iida, T., Buhler, M., Villen, J., Gygi, S.P., Nakayama, J., and Moazed, D. (2007). Two different Argonaute complexes are required for siRNA generation and heterochromatin assembly in fission yeast. Nat. Struct. Mol. Biol. *14*, 200–207.

Cam, H.P., Chen, E.S., and Grewal, S.I. (2009). Transcriptional scaffolds for heterochromatin assembly. Cell *136*, 610–614.

Cam, H.P., Sugiyama, T., Chen, E.S., Chen, X., FitzGerald, P.C., and Grewal, S.I. (2005). Comprehensive analysis of heterochromatin- and RNAi-mediated epigenetic control of the fission yeast genome. Nat. Genet. *37*, 809–819.

Carmell, M.A., Xuan, Z., Zhang, M.Q., and Hannon, G.J. (2002). The Argonaute family: tentacles that reach into RNAi, developmental control, stem cell maintenance, and tumorigenesis. Genes Dev. *16*, 2733–2742.

Carmichael, J.B., Stoica, C., Parker, H., McCaffery, J.M., Simmonds, A.J., and Hobman, T.C. (2006). RNA interference effector proteins localize to mobile cytoplasmic puncta in Schizosaccharomyces pombe. Traffic *7*, 1032–1044.

Chen, T., Brownawell, A.M., and Macara, I.G. (2004a). Nucleocytoplasmic shuttling of JAZ, a new cargo protein for exportin-5. Mol. Cell. Biol. *24*, 6608–6619.

Chen, X.Q., Du, X., Liu, J., Balasubramanian, M.K., and Balasundaram, D. (2004b). Identification of genes encoding putative nucleoporins and transport factors in the fission yeast Schizosaccharomyces pombe: a deletion analysis. Yeast *21*, 495–509.

Colmenares, S.U., Buker, S.M., Buhler, M., Dlakic, M., and Moazed, D. (2007). Coupling of double-stranded RNA synthesis and siRNA generation in fission yeast RNAi. Mol. Cell *27*, 449–461.

Dang, V.D., and Levin, H.L. (2000). Nuclear import of the retrotransposon Tf1 is governed by a nuclear localization signal that possesses a unique requirement for the FXFG nuclear pore factor Nup124p. Mol. Cell. Biol. *20*, 7798–7812.

Djupedal, I., Portoso, M., Spahr, H., Bonilla, C., Gustafsson, C.M., Allshire, R.C., and Ekwall, K. (2005). RNA Pol II subunit Rpb7 promotes centromeric transcription and RNAi-directed chromatin silencing. Genes Dev. *19*, 2301–2306.

Fang, Y., and Spector, D.L. (2007). Identification of nuclear dicing bodies containing proteins for microRNA biogenesis in living Arabidopsis plants. Curr. Biol. *17*, 818–823.

Forsburg, S.L. (1993). Comparison of Schizosaccharomyces pombe expression systems. Nucleic Acids Res. *21*, 2955–2956.

Fritz, J., Strehblow, A., Taschner, A., Schopoff, S., Pasierbek, P., and Jantsch, M.F. (2009). RNA-regulated interaction of transportin-1 and exportin-5 with the double-stranded RNA-binding domain regulates nucleocytoplasmic shuttling of ADAR1. Mol. Cell. Biol. *29*, 1487–1497.

Grishok, A., Pasquinelli, A.E., Conte, D., Li, N., Parrish, S., Ha, I., Baillie, D.L., Fire, A., Ruvkun, G., and Mello, C.C. (2001). Genes and mechanisms related to RNA interference regulate expression of the small temporal RNAs that control *C. elegans* developmental timing. Cell *106*, 23–34.

Gwizdek, C., Ossareh-Nazari, B., Brownawell, A.M., Evers, S., Macara, I.G., and Dargemont, C. (2004). Minihelix-containing RNAs mediate exportin-5-dependent nuclear export of the double-stranded RNA-binding protein ILF3. J. Biol. Chem. *279*, 884–891.

Hansen, K.R., Burns, G., Mata, J., Volpe, T.A., Martienssen, R.A., Bahler, J., and Thon, G. (2005). Global effects on gene expression in fission yeast by silencing and RNA interference machineries. Mol. Cell. Biol. *25*, 590–601.

Hiraguri, A., Itoh, R., Kondo, N., Nomura, Y., Aizawa, D., Murai, Y., Koiwa, H., Seki, M., Shinozaki, K., and Fukuhara, T. (2005). Specific interactions between Dicer-like proteins and HYL1/DRB-family dsRNA-binding proteins in Arabidopsis thaliana. Plant Mol. Biol. *57*, 173–188.

Ho, C.K., Wang, L.K., Lima, C.D., and Shuman, S. (2004). Structure and mechanism of RNA ligase. Structure *12*, 327–339.

Huisinga, K.L., and Elgin, S.C. (2009). Small RNA-directed heterochromatin formation in the context of development: what flies might learn from fission yeast. Biochim. Biophys. Acta *1789*, 3–16.

Hutvagner, G., McLachlan, J., Pasquinelli, A.E., Balint, E., Tuschl, T., and Zamore, P.D. (2001). A cellular function for the RNA-interference enzyme Dicer in the maturation of the let-7 small temporal RNA. Science *293*, 834–838.

Hutvagner, G., and Simard, M.J. (2008). Argonaute proteins: key players in RNA silencing. Nat. Rev. Mol. Cell Biol. *9*, 22–32.

Jaskiewicz, L., and Filipowicz, W. (2008). Role of Dicer in posttranscriptional RNA silencing. Curr. Top. Microbiol. Immunol. *320*, 77–97.

Jinek, M., and Doudna, J.A. (2009). A three-dimensional view of the molecular machinery of RNA interference. Nature *457*, 405–412.

Kato, H., Goto, D.B., Martienssen, R.A., Urano, T., Furukawa, K., and Murakami, Y. (2005). RNA polymerase II is required for RNAi-dependent heterochromatin assembly. Science *309*, 467–469.

Ketting, R.F., Fischer, S.E., Bernstein, E., Sijen, T., Hannon, G.J., and Plasterk, R.H. (2001). Dicer functions in RNA interference and in synthesis of small RNA involved in developmental timing in C. elegans. Genes Dev. *15*, 2654–2659.

Langmead, B., Trapnell, C., Pop, M., and Salzberg, S.L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol. *10*, R25.

Lau, N.C., Lim, L.P., Weinstein, E.G., and Bartel, D.P. (2001). An abundant class of tiny RNAs with probable regulatory roles in Caenorhabditis elegans. Science *294*, 858–862.

Lee, Y., Ahn, C., Han, J., Choi, H., Kim, J., Yim, J., Lee, J., Provost, P., Radmark, O., Kim, S., et al. (2003). The nuclear RNase III Drosha initiates microRNA processing. Nature *425*, 415–419.

Leeds, P., Peltz, S.W., Jacobson, A., and Culbertson, M.R. (1991). The product of the yeast UPF1 gene is required for rapid turnover of mRNAs containing a premature translational termination codon. Genes Dev. *5*, 2303–2314.

Li, C.F., Henderson, I.R., Song, L., Fedoroff, N., Lagrange, T., and Jacobsen, S.E. (2008). Dynamic regulation of ARGONAUTE4 within multiple nuclear bodies in Arabidopsis thaliana. PLoS Genet. *4*, e27.

Li, C.F., Pontes, O., El-Shami, M., Henderson, I.R., Bernatavichute, Y.V., Chan, S.W., Lagrange, T., Pikaard, C.S., and Jacobsen, S.E. (2006). An ARGONAUTE4-containing nuclear processing center colocalized with Cajal bodies in *Arabidopsis thaliana*. Cell *126*, 93–106.

Macchi, P., Brownawell, A.M., Grunewald, B., DesGroseillers, L., Macara, I.G., and Kiebler, M.A. (2004). The brain-specific double-stranded RNA-binding protein Staufen2: nucleolar accumulation and isoform-specific exportin-5-dependent export. J. Biol. Chem. *279*, 31440–31444.

Macrae, I.J., Zhou, K., Li, F., Repic, A., Brooks, A.N., Cande, W.Z., Adams, P.D., and Doudna, J.A. (2006). Structural basis for double-stranded RNA processing by Dicer. Science *311*, 195–198.

Matsuyama, A., Arai, R., Yashiroda, Y., Shirai, A., Kamata, A., Sekido, S., Kobayashi, Y., Hashimoto, A., Hamamoto, M., Hiraoka, Y., et al. (2006).

ORFeome cloning and global analysis of protein localization in the fission yeast Schizosaccharomyces pombe. Nat. Biotechnol. *24*, 841–847.

Moazed, D. (2009). Small RNAs in transcriptional gene silencing and genome defence. Nature *457*, 413–420.

Mochizuki, K., Fine, N.A., Fujisawa, T., and Gorovsky, M.A. (2002). Analysis of a piwi-related gene implicates small RNAs in genome rearrangement in tetrahymena. Cell *110*, 689–699.

Motamedi, M.R., Verdel, A., Colmenares, S.U., Gerber, S.A., Gygi, S.P., and Moazed, D. (2004). Two RNAi complexes, RITS and RDRC, physically interact and localize to noncoding centromeric RNAs. Cell *119*, 789–802.

Noma, K., Sugiyama, T., Cam, H., Verdel, A., Zofall, M., Jia, S., Moazed, D., and Grewal, S.I. (2004). RITS acts in cis to promote RNA interference-mediated transcriptional and post-transcriptional silencing. Nat. Genet. *36*, 1174–1180.

Papp, I., Mette, M.F., Aufsatz, W., Daxinger, L., Schauer, S.E., Ray, A., van der Winden, J., Matzke, M., and Matzke, A.J. (2003). Evidence for nuclear processing of plant micro RNA and short interfering RNA precursors. Plant Physiol. *132*, 1382–1390.

Pontes, O., Li, C.F., Nunes, P.C., Haag, J., Ream, T., Vitins, A., Jacobsen, S.E., and Pikaard, C.S. (2006). The Arabidopsis chromatin-modifying nuclear siRNA pathway involves a nucleolar RNA processing center. Cell *126*, 79–92.

Pratt, A.J., and Macrae, I.J. (2009). The RNA-induced silencing complex: a versatile gene-silencing machine. J. Biol. Chem. *284*, 17897–17901.

Provost, P., Dishart, D., Doucet, J., Frendewey, D., Samuelsson, B., and Radmark, O. (2002a). Ribonuclease activity and RNA binding of recombinant human Dicer. EMBO J. *21*, 5864–5874.

Provost, P., Silverstein, R.A., Dishart, D., Walfridsson, J., Djupedal, I., Kniola, B., Wright, A., Samuelsson, B., Radmark, O., and Ekwall, K. (2002b). Dicer is required for chromosome segregation and gene silencing in fission yeast cells. Proc. Natl. Acad. Sci. USA *99*, 16648–16653.

Reinhart, B.J., and Bartel, D.P. (2002). Small RNAs correspond to centromere heterochromatic repeats. Science *297*, 1831.

Schalch, T., Job, G., Noffsinger, V.J., Shanker, S., Kuscu, C., Joshua-Tor, L., and Partridge, J.F. (2009). High-affinity binding of Chp1 chromodomain to K9 methylated histone H3 is required to establish centromeric heterochromatin. Mol. Cell *34*, 36–46.

Sigova, A., Rhind, N., and Zamore, P.D. (2004). A single Argonaute protein mediates both transcriptional and posttranscriptional silencing in Schizosaccharomyces pombe. Genes Dev. *18*, 2359–2367.

Siomi, H., and Siomi, M.C. (2009). On the road to reading the RNA-interference code. Nature *457*, 396–404.

Song, L., Han, M.H., Lesicka, J., and Fedoroff, N. (2007). Arabidopsis primary microRNA processing proteins HYL1 and DCL1 define a nuclear body distinct from the Cajal body. Proc. Natl. Acad. Sci. USA *104*, 5437–5442.

Strehblow, A., Hallegger, M., and Jantsch, M.F. (2002). Nucleocytoplasmic distribution of human RNA-editing enzyme ADAR1 is modulated by double-stranded RNA-binding domains, a leucine-rich export signal, and a putative dimerization domain. Mol. Biol. Cell *13*, 3822–3835.

Sugiyama, T., Cam, H., Verdel, A., Moazed, D., and Grewal, S.I. (2005). RNA-dependent RNA polymerase is an essential component of a self-enforcing loop coupling heterochromatin assembly to siRNA production. Proc. Natl. Acad. Sci. USA *102*, 152–157.

Verdel, A., Jia, S., Gerber, S., Sugiyama, T., Gygi, S., Grewal, S.I., and Moazed, D. (2004). RNAi-mediated targeting of heterochromatin by the RITS complex. Science *303*, 672–676.

Volpe, T.A., Kidner, C., Hall, I.M., Teng, G., Grewal, S.I., and Martienssen, R.A. (2002). Regulation of heterochromatic silencing and histone H3 lysine-9 methylation by RNAi. Science *297*, 1833–1837.

West, R.R., Vaisberg, E.V., Ding, R., Nurse, P., and McIntosh, J.R. (1998). cut11(+): A gene required for cell cycle-dependent spindle pole body anchoring in the nuclear envelope and bipolar spindle formation in Schizosaccharomyces pombe. Mol. Biol. Cell *9*, 2839–2855.

Xie, Z., Johansen, L.K., Gustafson, A.M., Kasschau, K.D., Lellis, A.D., Zilberman, D., Jacobsen, S.E., and Carrington, J.C. (2004). Genetic and functional diversification of small RNA pathways in plants. PLoS Biol. *2*, E104.

Zhang, H., Kolb, F.A., Jaskiewicz, L., Westhof, E., and Filipowicz, W. (2004). Single processing center models for human Dicer and bacterial RNase III. Cell *118*, 57–68.

Developmental Cell *18*

Supplemental Information

Nuclear Retention of Fission Yeast Dicer

Is a Prerequisite for RNAi-Mediated

Heterochromatin Assembly

Stephan Emmerth, Heiko Schober, Dimos Gaidatzis, Tim Roloff, Kirsten Jacobeit, and Marc Bühler

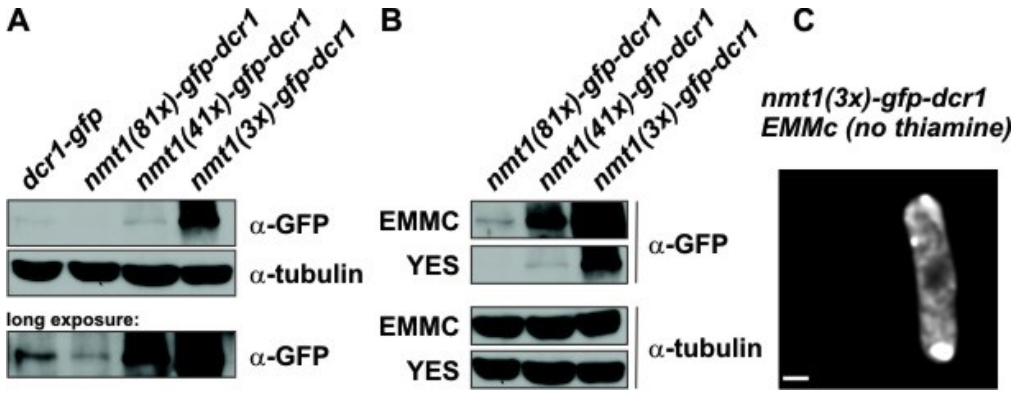


Figure S1, related to Figure 1. Assessment of GFP-Dcr1 expression levels under different conditions and cytoplasmic Dcr1 as a result of massive overexpression of GFP-Dcr1

(A) Western blot analysis comparing the expression levels of GFP-Dcr1 expressed from different versions of the nmt1 promoter and Dcr1-GFP driven by its endogenous promoter. Tubulin was used as a loading control.

(B) Western blot analysis comparing the expression levels of GFP-Dcr1 from cells grown under repressed conditions in YES (thiamine present) or grown under non-repressed conditions in EMMc. Tubulin was used as a loading control.

(C) Live cell microscopy of a *S. pombe* cell expressing N-terminally GFP-tagged Dcr1 driven from the nmt1(3x) promoter. Cells were grown under non-repressive conditions (i.e. in EMMc medium) at 30°C, resulting in massive overexpression of GFP-Dcr1. Scale bar = 2μm. Compare this Figure with Figures 1A iii and iiii. This shows that GFP-Dcr1 localizes in perinuclear foci when expressed at more moderate levels. Note that previous studies, which demonstrated mainly cytoplasmic localization for Dcr1 (as shown in (C)), were performed with plasmid encoded, *adh1* promoter driven Dcr1-GFP, or from a single copy genomic integration of Dcr1-YFP expressed from the nmt1(3x) promoter under non-repressed conditions (Matsuyama et al., 2006). Both approaches lead to a massive overproduction of Dcr1. *adh1* promoter-driven expression of Dcr1-GFP can be expected to be at least 800-fold or 160-fold stronger than nmt1(41x)-or nmt1(3x)-driven expression under repressive conditions, respectively (Forsburg, 1993). Expression from the nmt1(3x) promoter is ~300-fold stronger in non-repressed compared to repressed conditions (S1B; Forsburg, 1993).

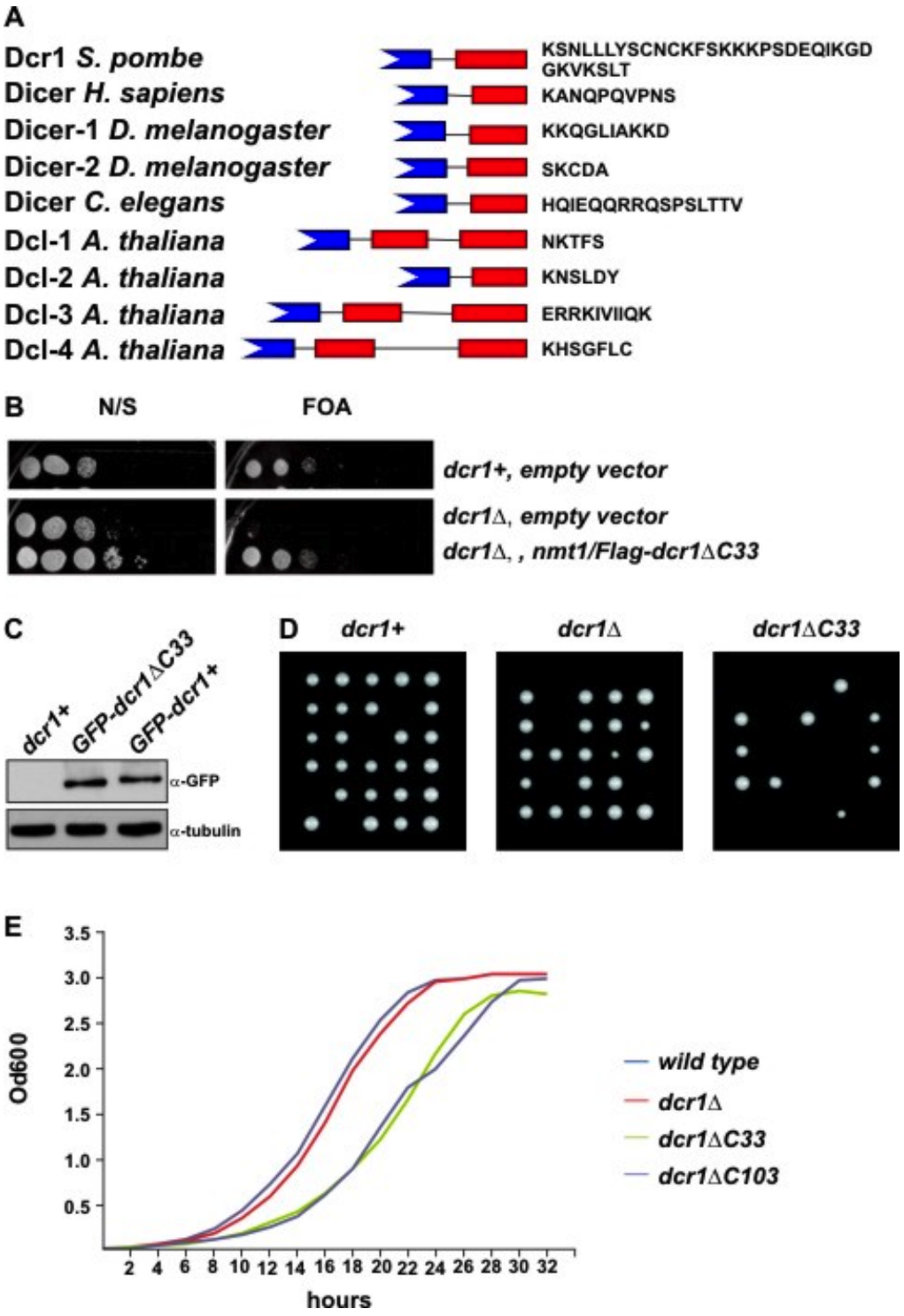


Figure S2, related to Figure 4. Silencing, protein stability, and cell growth assessed for cells expressing Dcr1ΔC33

(A) C-terminal structure of dicer homologues. Protein sequences were collected from UniProtKB (release15.6; The UniProt Consortium, 2009). Isoforms were ignored, as they had no differences in the C-terminal regions. The C-terminal region is defined as the sequence after the second RNaseIII domain (represented here as blue fragments), which was determined by a scan against the Pfam database (release 23.0; Finn et al., 2008). The number and location of double-stranded RNA binding motifs (represented as red bars) within the C-Terminal fragments were determined by using the Pfam dsrm (PF00035) Hidden Markov Model (allowing E-values of up to 0.54 in the case of the *S. pombe* sequence). Amino acid sequences following the (potential) double-stranded RNA binding motifs are displayed in capital letters.

(B) Overexpression of Dcr1ΔC33 rescues silencing of a centromeric *ura4+* reporter (*imr1R::ura4+*). Serial 10-fold dilutions of the strains indicated were

plated on PMGc-leu (non selective, NS) or on PMGc-leu plates containing 2 mg/mL 5-FOA, a substance toxic for cells which express *ura4*+.
(C) Western blot showing that protein stability of GFP-Dcr1 is not affected by deleting C33. Tubulin served as a loading control.
(D) 6x5 single haploid cells for each indicated strain were dissected onto YES plates and grown at 30°C for ~2-3 days.
(E) Growth of wild type, *dcr1Δ*, *dcr1ΔC33* and *dcr1ΔC103* cells was monitored for 32h by OD600 measurements every 2 hours. Over-night cultures grown in YES were diluted to OD600=0.03 in YES.

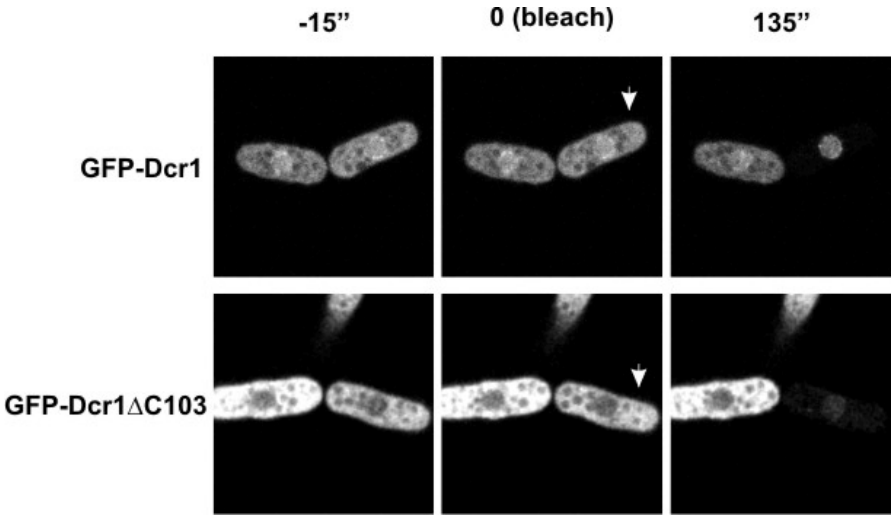


Figure S3, related to Figure 7. Nuclear fluorescence is not greatly affected after cytoplasmic photobleaching in Dcr1 and Dcr1ΔC103 expressing cells
Representative images of nuclear fluorescence loss in (cytoplasmic) photobleaching (FLIP) of cells expressing GFP-Dcr1 and GFPDcr1ΔC103 (GFP-Dcr1 images are also shown in Figure 6). White arrow indicates the bleached cell. Bleach iterations = 50.

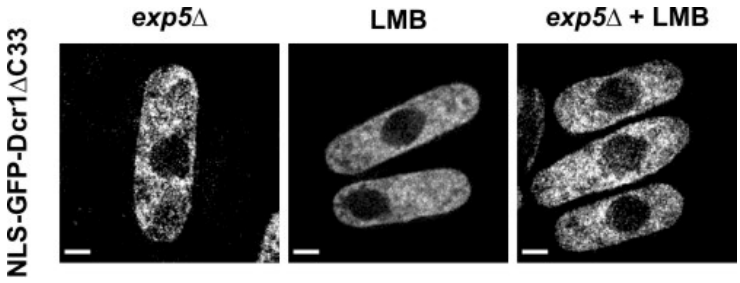


Figure S4, related to Figure 6. NLS-GFP-Dcr1DC33 does not accumulate in the nucleus after *exp5* deletion and/or LMB treatment
Live cell microscopy of *S. pombe* cells expressing NLS-GFP-Dcr1ΔC33. Where indicated, cells were grown for 3 hours in 100 ng/ml Leptomycin B (LMB) prior to live cell imaging. Scale bar = 2 μm.

SUPPLEMENTAL EXPERIMENTAL PROCEDURES

Strains used in this study

Strain	Genotype	Source
SPB80	*h+ leu1-32 ura4-D18 ori1 ade6-216 imr1R(Nco1)::ura4+*	2
SPB81	*h+ leu1-32 ura4-D18 ori1 ade6-216 imr1R(Nco1)::ura4+ dcr1Δ::TAP-Kan*	2
SPB125	*h- clr4Δ::Nat imr1R::ura4+*	2
SPB287	*h+ leu1-32 ura4-D18 ori1 ade6-216 imr1R(Nco1)::ura4+ dcr1ΔC33::Nat*	1
SPB311	*h+ leu1-32 ura4-D18 ori1 ade6-216 imr1R(Nco1)::ura4+ dcr1-D937A D1127AΔC33::Nat*	1
SPB322	*h+ leu1-32 ura4-D18 ori1 ade6-216 imr1R(Nco1)::ura4+ dcr1ΔC33::NLS-Hph*	1
SPB185	*h+ leu1-32 ura4-D18 ade6-216 imr1R(NcoI)::ura4+ Nat-nmt1P3-GFP::dcr1+*	2
SPB267	*h+ imr1R(NcoI)::ura4+ ura4D18 ade6-216 leu1-32 Kan-nmt1P41-GFP::dcr1+*	1
SPB269	*h+ imr1R(NcoI)::ura4+ ura4D18 ade6-216 leu1-32 Kan-nmt1P81-GFP::dcr1+*	1
SPB427	*h+ imr1R(NcoI)::ura4+ dcr1+::GFP-Kan ura4D18 ade6-M216 leu1-32*	1
SPB337	*h+ Nat-nmt1P3-GFP::dcr1+ ago1Δ::TAP-Kan imr1R(NcoI)::ura4+ ura4D18 ade6-(?) leu1-32*	1
SPB338	*h+ Nat-nmt1P3-GFP::dcr1+ rdp1Δ::TAP-Kan imr1R(NcoI)::ura4+ ura4D18 ade6-(?) leu1-32*	1
SPB207	*h+ Nat-nmt1P3-GFP::dcr1+ arb1Δ::TAP-Kan imr1R(NcoI)::ura4+ ura4D18 ade6-(?) leu1-32*	1
SPB200	*h+ Nat-nmt1P3-GFP::dcr1+ clr4Δ::Hph imr1R(NcoI)::ura4+ ura4D18 ade6-(?) leu1-32*	1
SPB408	*h+ swi6+::tdTomato-Hph Nat-nmt1P3-GFP::dcr1+ lys1::GFP-lacI-his7+(?) ade6-(?) ura4-(?) leu1-(?)*	1
SPB198	*h- cnp1+::mCherry-Nat Nat-nmt1P3-GFP::dcr1+ura4-D18 leu1-32 ade6-(?)*	1
SPB328	*h+ nup107+::mCherry-Nat Nat-nmt1P3-GFP::dcr1+ imr1R(NcoI)::ura4+ ura4-D18 leu1-32 ade6-(?)*	1
SBP290	*h+ imr1R(Nco1)::ura4+ Kan-nmt1P3-GFP::dcr1ΔC33::Nat leu1-32 ura4-D18 ori1 ade6-216*	1
SPB353	*h+ imr1R(Nco1)::ura4+ Kan-nmt1P3-NLS-GFP::dcr1ΔC33::Nat leu1-32 ura4-D18 ori1 ade6-216*	1
SPB451	*h? nup107+::GFP-Kan Kan-nmt1P3-mCh::dcr1+ imr1R::GFP-Nat ura4-(?) leu1-(?) ade6-(?) GFP-lacI(?)*	1
SPB454	*h+ imr1R(Nco1)::ura4+ Kan-nmt1P3-GFP::dcr1ΔC103::Nat leu1-32 ura4-D18 ori1 ade6-216*	1
SPB455	*h+ imr1R(Nco1)::ura4+ Kan-nmt1P3-NLS-GFP::dcr1ΔC103::Nat leu1-32 ura4-D18 ori1 ade6-216*	1
SPB531	*h? ade6- ura4-D18 leu1-32 nup120Δ::Hph Kan-nmt1P3-GFP::dcr1*	1
SPB196	*h- cut11+::mCherry-Nat, Nat-nmt1P3-GFP::dcr1+ ura4-D18 leu1-32 ade6-216*	1
SPB486	*h- ade6- ura4-D18 leu1-32 his? SPAC328.01cΔ::Hph Kan-nmt1P3-NLS-GFP::dcr1ΔC33::Nat*	1

1 = this study, 2 = Danesh Moazed

Plasmids used in this study

Name	common name
pMB279	pREP1-3xFLAG-Dcr1ΔC33-Leu+
pMB428	p3-nmt1-gfp-LacZ-Leu+
pMB429	p3-nmt1-nls-gfp-LacZ-Leu+
pMB430	p3-nmt1-gfp-LacZ-Dcr1-C103-Leu+
pMB432	p3-nmt1-gfp-LacZ-Dcr1-dsRBD-Leu+
pMB434	p3-nmt1-gfp-LacZ-Dcr1-C33-Leu+

Primers used for qRT-PCR

Name	Sequence
Cendg fwd	AAGGAATGTGCCTCGTCAAATT
Cendg rev	TGCTTCACGGTATTTTTTGAAATC
Cendh fwd	GTATTTGGATTCCATCGGTACTATGG
Cendh rev	ACTACATCGACACAGAAAAGAAAACAA
Ura4 fwd	TACAAAATTGCTTCTTGGGCTCAT
Ura4 rev	AGACCACGTCCCAAAGGTAAAC
Act fwd	TCCTCATGCTATCATGCGTCTT
Act rev	CCACGCTCCATGAGAATCTTC

SUPPLEMENTAL REFERENCES

The UniProt Consortium, T.U. (2009). The Universal Protein Resource (UniProt) 2009. Nucleic Acids Res *37*, D169-174.

Finn, R.D., Tate, J., Mistry, J., Coggill, P.C., Sammut, S.J., Hotz, H.R., Ceric, G., Forslund, K., Eddy, S.R., Sonnhammer, E.L.*, et al.* (2008). The Pfam protein families database. Nucleic Acids Res *36*, D281-288.

Manuscript II

Related to section 3.2.

Manuscript II

An Extended dsRBD with a Novel Zinc-Binding Motif Mediates Nuclear Retention of Fission Yeast Dicer

Pierre Barraud[1,*], Stephan Emmerth[2,3,*], Yukiko Shimada[2,3], Hans-Rudolf Hotz[2,3], Frédéric H.-T. Allain[1,†] and Marc Bühler[2,3,†]

[1]Institute of Molecular Biology and Biophysics, ETH Zürich, CH-8093 Zürich, Switzerland
[2]Friedrich Miescher Institute for Biomedical Research, Maulbeerstrasse 66, CH-4058 Basel, Switzerland
[3]Universität Basel, Petersplatz 10, CH-4003 Basel, Switzerland

* These authors contributed equally to this work

† Correspondence: marc.buehler@fmi.ch; allain@mol.biol.ethz.ch

Running title: Dicer harbors a novel zinc-binding motif

ABSTRACT

Dicer proteins function in RNA interference (RNAi) pathways by generating small RNAs (sRNAs). Here we report the solution structure of the C-terminal domain of *Schizosaccharomyces pombe* Dicer (Dcr1). The structure reveals an unusual double-stranded RNA binding domain (dsRBD) fold embedding a novel zinc-binding motif that is conserved among dicers in yeast. Although the C-terminal domain of Dcr1 still binds nucleic acids, this property is dispensable for proper functioning of Dcr1. In contrast, disruption of zinc coordination renders Dcr1 mainly cytoplasmic and leads to remarkable changes in gene expression and loss of heterochromatin assembly. In summary, our results reveal novel insights into the mechanism of nuclear retention of Dcr1 and raise the possibility that this new class of dsRBDs might generally function in nucleo-cytoplasmic trafficking and not substrate binding. The C-terminal domain of Dcr1 constitutes a novel regulatory module that might represent a potential target for therapeutic intervention with fungal diseases.

Keywords:
Dicer / dsRBD / heterochromatin / RNA interference / zinc-binding domain

INTRODUCTION

RNA interference (RNAi) designates a set of conserved pathways found in many eukaryotes. These pathways are involved in various processes ranging from the control of gene expression to silencing of mobile genetic elements, combating viruses and maintaining genomic integrity (Baulcombe, 2004; Grishok et al, 2001; Huisinga & Elgin, 2009; van Wolfswinkel & Ketting). Additionally, RNAi is also being pursued as a promising tool for the treatment of human diseases. Common to all RNAi pathways is the association of small RNAs (sRNA) with members of the Argonaute family of proteins, forming the core component of a diverse set of protein-RNA complexes called RNA-induced silencing complexes (RISCs; Carmell et al, 2002; Hutvagner & Simard, 2008). To date, at least three classes of RISC-associated sRNAs have been identified (Siomi & Siomi, 2009): short interfering RNAs (siRNAs), microRNAs (miRNAs) and PIWI-interacting RNAs (piRNAs). These sRNAs guide the RISCs through base-pairing interactions to homologous sequences, which usually results in reduced activity of the target sequence (Ghildiyal & Zamore, 2009).

In *Schizosaccharomyces pombe (S. pombe)* the RNAi pathway functions in association with chromatin to trigger co-transcriptional gene silencing (CTGS) and the formation of heterochromatin. Deletions of any of the three genes encoding the RNAi proteins Dicer, Argonaute and RNA-dependent RNA polymerase (*dcr1*$^+$, *ago1*$^+$ and *rdp1*$^+$, respectively) result in loss of CTGS, chromosome segregation defects, and greatly reduced levels of H3K9 methylation (H3K9me) at centromeric repeats (Volpe et al, 2003; Volpe et al, 2002; Woolcock et al, 2011). siRNAs corresponding to heterochromatic regions are loaded into an Argonaute-containing complex referred to as RNA-induced transcriptional silencing (RITS) complex in *S. pombe* (Reinhart & Bartel, 2002; Verdel et al, 2004). Importantly, the RITS complex is physically linked to heterochromatin and this interaction, as well as siRNA generation, depends on the histone H3K9 methyltransferase Clr4 (Cam et al, 2005; Motamedi et al, 2004; Sugiyama et al, 2005). These and other observations have led to a model in which the association of the RITS complex with chromatin is proposed to involve base-pairing between siRNAs and the nascent RNA polymerase II (RNApII) transcripts. Subsequently, RITS would recruit

histone-modifying enzymes such as Clr4, leading to the generation and spreading of heterochromatin (Buhler et al, 2006; Moazed, 2009).

The biogenesis of siRNAs is mediated by Dcr1, which processes double-stranded RNA (dsRNA) precursors endonucleolytically. Resembling canonical Dicers in animals, *S. pombe* Dcr1 contains an N-terminal helicase/ATPase domain, followed by a DUF283 domain, a Platform domain (Macrae et al, 2006) and a PAZ-like domain, two RNase III domains, and a long C-terminal domain that bears a divergent double-stranded RNA binding domain (dsRBD) and a short motif referred to as C33 (Figure 1A). Notably, the C-terminus of Dcr1, comprising the dsRBD and C33, is dispensable for processing dsRNA *in vitro* but not *in vivo,* where it functions to control the subcellular localization of Dcr1 (Colmenares et al, 2007; Emmerth et al, 2010). We have shown previously that C33 functions to counteract dsRBD-mediated export of Dcr1 to the cytoplasm, resulting in nuclear accumulation of Dcr1. This is crucial for a functional RNAi pathway and establishment of heterochromatin in *S. pombe*, as mutations rendering Dcr1 cytoplasmic abolish siRNA generation and the methylation of H3K9 (Emmerth et al, 2010).

That dsRBDs can mediate nucleo-cytoplasmic trafficking has been demonstrated for other proteins (Chen et al, 2004; Gwizdek et al, 2004; Macchi et al, 2004; Strehblow et al, 2002). However, mechanistically this process is poorly understood. For instance, it is not clear to what extent binding to dsRNA contributes to this rather unexpected function of dsRBDs. Similarly, how C33 of Dcr1 contributes to the inhibition of the export-promoting activity of the dsRBD remains unknown. To address these questions, we determined the NMR solution structure of the C-terminal domain of Dcr1. Although the structure reveals a typical dsRBD fold, additional structural elements are present. This includes a novel zinc-binding motif, formed by four residues that are encoded in both the dsRBD and C33. We demonstrate that zinc coordination by this motif is required for the formation of a protein-protein interaction surface that is required for proper Dcr1 localization. Surprisingly, although the dsRBD binds to dsRNA strongly, this property is dispensable for proper functioning of Dcr1 in the *S. pombe* RNAi pathway. This raises the possibility that dicer dsRBDs might generally function in nucleo-cytoplasmic transport and not necessarily substrate binding.

RESULTS

The dsRBD of Dcr1 embeds an unexpected zinc-binding motif

To obtain novel insights into the role of C33 in promoting nuclear accumulation of Dcr1, we studied the Dcr1 C-terminal domain with NMR spectroscopy (1259-1374, long dsRBD construct; Figure 1B). We assigned backbone and side-chain NMR chemical shifts using standard triple-resonance experiments as described in Material and Methods and could achieve 80.1 % of proton assignment. NOE distance restraints were obtained from four NOESY spectra and used to run initial structure calculations. As expected, the preliminary folds revealed a typical dsRBD fold for residues 1262 to 1341. More surprisingly, the first half of C33 appeared to contact the dsRBD. Strikingly, we observed that four side-chains with two residues from the dsRBD and two from C33 (in total 3 cysteines and 1 histidine; highlighted in red on Figure 1B) appear in close proximity in this preliminary structure, strongly suggesting that these four side-chains might be coordinating a zinc ion. Subsequent atomic absorption spectroscopy (AAS) and mass spectrometry measurements confirmed that the Dcr1 C-terminal domain containing the dsRBD and the C33 indeed include one zinc ion per molecule (Supplementary Figure 1 and Table I). $^{15}N\{^{1}H\}$-NOE measurement (Supplementary Figure 2) further confirmed that the first half of C33 fragment is structured and participates in the overall fold of the domain, while the last 16 residues (1359-1374) are flexible. To reduce signal overlap and therefore improve the precision of the structure, we decided to investigate a construct lacking the last 16 residues of Dcr1 (1259-1358, short dsRBD construct; Figure 1B). Also, $ZnCl_2$ was added during the protein preparation, which helped to stabilize the recombinant short dsRBD construct and improved the overall quality of the NMR spectra.

Final protein structure determination

The NMR solution structure of the Dcr1 C-terminal domain (1259-1358, short dsRBD construct) was determined using 2308 NOE-derived constraints. The structure is very precise with a backbone rmsd 0.46 ± 0.12 Å for the NMR ensemble of conformers.

Hydrogen-bonded amides were established as slowly exchanging protons in the presence of D_2O. Their hydrogen-bond acceptors were identified from preliminary structures as well as from analysis of the characteristic NOE pattern found in α-helices and β-sheets. Characteristic NOE contacts between the ligands of the zinc ion further confirmed the composition of the coordination sphere that involves C1275 and H1312 from the dsRBD, and C1350 and C1352 from C33 (Supplementary Figure 3). In addition, long-range ^{1}H-^{15}N HSQC shows a characteristic pattern of cross peaks that allowed us to unambiguously identify H1312 as coordinating the zinc ion via its Nε (Supplementary Figure 4; Legge et al, 2004; Pelton et al, 1993). Therefore, we added distance constraints to the four zinc-binding residues to impose a tetrahedral coordination around the zinc ion. NMR experimental constraints and refinement statistics are presented in Table II.

The Dcr1 C-terminal domain adopts an extended dsRBD fold due to a novel zinc-binding motif

The residues ranging from I1262 to D1341 form the dsRBD core domain with the canonical $\alpha\beta\beta\beta\alpha$ topology (Bycroft et al, 1995; Kharrat et al, 1995), with the two α-helices packed against the 3-stranded anti-parallel β-sheet (Figure 1C and D). The residues of C33 that immediately follow helix α2 are well structured (Figure 1C and D and Supplementary Figure 2) and fold back on the dsRBD. N1344 to L1347 fold in a short α helical turn (helix 3) that is followed by C1350 and C1352 coordinating zinc with C1275 and H1312 from loop 1 and loop 3 of the dsRBD, respectively (Figure 1C and D). This zinc-binding site is unusual because of the long spacing between the first and the last residue coordinating the zinc atoms (76 residues) as well as the spacing between the individual coordinating side-chains (C-X_{36}-H-X_{37}-C-X-C). Typically, zinc-binding motifs are found within 20 to 40 residues (Auld, 2001) with the most abundant form being the C2H2 zinc fingers that adopt a $\beta\beta\alpha$ fold (Laity et al, 2001). To our knowledge, the present ligand arrangement in the Dcr1 C-terminal domain is unprecedented among zinc-binding domains. Rather than being an independent fold of the dsRBD, this novel zinc-binding domain is formed jointly by dsRBD and C33 and therefore constitutes an

extended dsRBD fold (Figure 1E).

Binding of zinc to a protein is unlikely to sustain the large entropy changes that would result from the ordering of a very mobile peptide chain (Auld, 2001). Thus, in addition to the CHCC zinc-binding motif, other structural elements are required to maintain proper folding of the C-terminus. Indeed, the short segment joining helix 2 to the C-X-C part of the zinc-binding motif contains three hydrophobic residues (L1346, L1347 and Y1348) that form hydrophobic interactions with residues from the dsRBD (helix 1, helix 2 and the β-sheet, Figure 2). The two leucines interact with other leucine side-chains, constituting most of the hydrophobic core of the protein. Importantly, five leucine residues distributed along the three helices together with other hydrophobic residues of helix 1 (I1262 and V1266) bring the three α-helices, as well as the hydrophobic surface of the β-sheet, which is mainly composed of aromatic residues, in close contact (Figure 2). Most surprisingly, the side-chain of R1311 is also buried in the hydrophobic core of the protein, in close proximity to the zinc coordination site (Figure 2). The vicinity of this positively charged arginine to the CHCC motif would counterbalance the overall -1 net charge of a zinc ion coordinated by three deprotonated cysteines.

Zinc coordination is required for nuclear localization of Dcr1

We have previously demonstrated that C33 functions to counteract dsRBD-mediated export of Dcr1 to the cytoplasm, resulting in nuclear accumulation of Dcr1 (Emmerth et al, 2010). This could be achieved by binding of C33 to a putative nuclear retention factor. However, our structural analysis strongly suggests that C33 rather contributes to the overall folding of the C-terminus of Dcr1 by providing two cysteines to form the newly identified zinc-binding domain. To test whether disruption of zinc coordination would result in a similar phenotype as deleting C33, we mutated the three cysteines in the CHCC motif of endogenous Dcr1 into serines (Dcr1-CHCC to Dcr1-SHSS). This resulted in mainly cytoplasmic localization of the protein without changing the amount of protein, similar to what we observed previously when C33 was deleted entirely (Figure 3A and Supplementary Figure 5A). Therefore, it appears that it is the extended dsRBD structure formed upon zinc coordination, rather than the amino acid sequence of C33 *per se,* that

is responsible for preventing nuclear export of Dcr1. This is further supported by our observation that deletion of the last 16 amino acids of Dcr1 did not lead to a change in localization of the protein (Supplementary Figure 5B).

To test whether disrupting the coordination of zinc has the same functional consequences as deleting C33, we investigated heterochromatic gene silencing and the generation of centromeric siRNAs by quantitative RT-PCR and sRNA deep-sequencing, respectively. Consistent with a failure to retain Dcr1 in the nucleus, silencing of centromeric heterochromatin and the generation of centromeric siRNAs were lost in cells expressing Dcr1-SHSS (Figure 3B and C, Supplementary Figure 6 and Supplementary Table I). Concomitant with the observed loss of silencing, H3K9 methylation was lost in Dcr1-SHSS expressing cells (Figure 3D). In conclusion, coordination of zinc by the C-terminus of Dcr1 is required for its nuclear retention and thus RNAi-mediated heterochromatin assembly.

The Dcr1 dsRBD binds dsRNA and dsDNA

It has been demonstrated for other dsRBDs that they can mediate protein import and/or export (Chen et al, 2004; Gwizdek et al, 2004; Macchi et al, 2004; Strehblow et al, 2002), and it has been discussed that RNA may regulate this process (Fritz et al, 2009). In canonical dsRBDs, helix 1 and the loop between β1 and β2 both contact the minor groove of a dsRNA helix (Supplementary Figure 7A) (Ramos et al, 2000; Ryter & Schultz, 1998; Stefl et al; Wu et al, 2004). Interestingly, the dsRBD of Dcr1 has an unusually long loop between β1 and β2 (14 residues compared to 4 to 6 residues in most dsRBDs), raising the question as to whether the C-terminal domain of Dcr1 would bind dsRNA at all. We thus investigated the nucleic acid binding properties of the extended dsRBD by isothermal titration calorimetry (ITC). Similar to canonical dsRBDs (Bass et al, 1994; St Johnston et al, 1992), the isolated domain (Figure 4A) binds rather strongly to a regular A-form RNA helix (Kd = 0.8 μM, with two dsRBDs bound per 24 bp dsRNA helix) (Figure 4B), but binds only weakly to ssRNA or ssDNA (Supplementary Figure 7B). The absence of binding to a regular B-form DNA helix has been reported for various dsRBDs (Burd & Dreyfuss, 1994; Fierro-Monti & Mathews, 2000; Saunders &

Barber, 2003), although competition binding studies revealed that some dsRBDs can also bind to dsDNA, albeit with weaker affinity (Bass et al, 1994). Surprisingly, the extended dsRBD of Dcr1 showed a fairly strong affinity to a regular B-form DNA helix (Kd = 1.5 μM), although only one molecule is bound per dsDNA (Supplementary Figure 7B). This raises the question of how this dsRBD can adapt to bind both a B-form DNA helix as well as an A-form RNA helix. We speculate that the unusually long β1-β2 loop and the unusual position of helix 1 relative to the rest of the domain (see discussion) might be responsible for this dual binding capacity.

RNA binding by the extended dsRBD is dispensable for proper functioning of Dcr1

Having shown that the extended dsRBD of Dcr1 binds dsRNA with high affinity, we set out to test if RNA binding is important for import/export-promoting activity of the dsRBD of Dcr1. We created mutations in three regions of the dsRBD (K1265A in helix 1, Δloop 2 and R1322A in loop 4, Figure 4A and Figure 1B, Supplementary Figure 5A for stability) that can be predicted to be involved in RNA binding based on previous structures of dsRBDs in complex with dsRNA (Ramos et al, 2000; Ryter & Schultz, 1998; Stefl et al; Wu et al, 2004). As expected, none of these dsRBD mutants could bind dsRNA *in vitro* (Figure 4C), despite being properly folded (as shown by 1D ^{1}H NMR spectra). This further demonstrates that the extended dsRBD of Dcr1 binds RNA in a canonical mode despite its rather unusual structural features. Moreover, binding to dsDNA is also lost in these three mutants (Supplementary Figure 7C). This allowed us to generate mutant strains expressing Dcr1 that has lost its ability to bind dsRNA via the extended dsRBD.

Interestingly, nuclear localization of Dcr1 was not affected in any of the three mutant strains GFP-Dcr1-R1322A (Figure 4D), -K1265A or -Δloop 2 (Figure 4E), demonstrating that nuclear import of Dcr1 is unlikely to be dependent on binding of dsRNA via the extended dsRBD. Instead, one might speculate that dsRNA binding could be important to mediate the rapid export observed in the absence of C33 or of a functional CHCC motif. However, this is also highly unlikely because we assume that a dsRBD lacking C33 or the zinc-binding motif is unstructured and therefore cannot bind RNA. Most

importantly, nuclear localization of the mainly cytoplasmic Dcr1ΔC33 could not be restored by mutating R1322 to alanine (Figure 4D, Supplementary Figure 5A for stability). Thus, the dsRBD of Dcr1 functions in nucleo-cytoplasmic trafficking independently of its affinity for dsRNA.

Surprisingly, we found that dsRNA binding through the dsRBD is also dispensable for heterochromatin silencing and siRNA generation (Figure 4F and 4G, Supplementary Figure 6). Consistent with this observation, *in vitro* processing of dsRNA into siRNAs by Dcr1 has previously been shown not to require the extended dsRBD (Colmenares et al, 2007). Therefore, rather than substrate binding and processing, the main function of the unusual and elongated dsRBD of Dcr1 appears to be to regulate the subcellular localization of Dcr1 independently of its affinity for dsRNA.

The extended dsRBD of Dcr1 exposes a protein interaction surface that functions in nuclear retention

Our finding that mutating the zinc coordination motif has the same functional consequences as deleting C33, together with the fact that mutations in the zinc-binding motif render the protein insoluble *in vitro*, strongly suggested that coordination of zinc is required for the proper folding of the protein. Because nuclear retention of Dcr1 is lost in both C33 and CHCC mutants, and loss of RNA binding does not affect nuclear localization, we speculated that an exposed protein surface unique to the extended dsRBD might be crucial for this nuclear retention.

In order to identify such a putative protein interaction surface, we aligned the sequence of the C-terminal domain of *S. pombe* Dcr1 with the two closest annotated proteins (*S. octosporus* Dcr1 and *S. cryophilus* Dcr1; 31 % and 33 % sequence identity respectively; see Supplementary Figure 8 for the sequence alignment). This revealed a patch of conserved residues (shown in red in Figure 5A) that cluster on one side of the domain around the third most N-terminal residue of C33 (N1344). In order to test whether this conserved surface is important for Dcr1 localization, we mutated residues in the center of the surface (N1344A, Y1348A, S1349A; belonging to C33) or at its periphery (R1334A; belonging to α-helix 2). Whereas mutating R1334 to A did not affect Dcr1

localization (Figure 5B), the three residues in the middle of the putative protein-protein interaction surface turned out be crucial for proper Dcr1 localization resulting in reduced nuclear localization when mutated singly and in a complete loss of nuclear localization when mutated together (Figure 5C and D).

These results are consistent with the idea that nuclear accumulation of Dcr1 is mediated by binding to a putative nuclear retention factor, which is likely to be mediated by this conserved surface. Importantly, C33 is required to establish the Dcr1 protein-protein interaction surface by contributing residues that are either important for the folding of the extended dsRBD or that belong to the exposed protein-protein interaction surface itself.

The novel zinc-binding motif is conserved in yeast dicers

In contrast to *S. pombe* Dcr1, human Dicer (DICER1) localizes mainly to the cytoplasm (Billy et al, 2001; Provost et al, 2002). Strikingly, only Dcr1 but not DICER1 has an extended C-terminus and we have therefore previously attempted to restrict DICER1 to the nucleus by the addition of C33 to its C-terminus. However, this had no effect on DICER1 localization, neither in human nor in *S. pombe* cells (S.E. and M.B. unpublished observation), further corroborating our conclusion that C33 is unlikely to bind a nuclear retention factor on its own. With the structure of the Dcr1 dsRBD solved, our failure to restrict DICER1 to the nucleus can be explained. Because C1275 and H1312 in the dsRBD of Dcr1 are not conserved in the dsRBD of DICER1, addition of C33 is not sufficient to coordinate zinc. Thus, DICER1+C33 cannot fold into an extended dsRBD as observed for the C-terminus of Dcr1.

Intrigued by this difference between the two proteins, we performed a refined search for proteins with at least one RNase III domain, followed by a dsRBD and a C-terminal extension of at least 15 amino acids, and looked for the presence of a potential zinc coordination motif similar to the one found in *S. pombe* Dcr1. This search revealed 39 proteins of the dicer family in which the CHCC motif is found. Surprisingly, except for the CHCC motif, most amino acids of the dsRBD or C33 differ between *S. pombe* and other dicer proteins (Supplementary Figure 9). Strikingly, many of the dicers in which the four zinc-binding ligands are present belong to human or plant pathogens (Figure 6A and

Supplementary Table II). This is particularly interesting because a *S. pombe* mutant strain that fails to coordinate zinc grows much more slowly and is less viable than wild-type or *dcr1Δ* cells (Figure 6B). Concomitantly, dramatic changes in gene expression can be observed upon mutating the zinc-coordinating residues but not upon deletion of the *dcr1+* gene (Figure 6C), in accordance with our previous observation that a functional C33 prevents Dcr1 from acting promiscuously (Emmerth et al, 2010). Thus, if growth would be similarly reduced in pathogenic yeasts, disruption of zinc coordination by Dcr1 in these organisms might represent an appealing strategy to fight pathogenic fungi.

DISCUSSION

In this study, we determined the solution structure of the C-terminal domain of *S. pombe* Dcr1 which revealed an extended dsRBD embedding a novel zinc-binding motif that is conserved among dicers in yeast. This zinc-binding motif is formed jointly by two side-chains of the dsRBD and two from C33, and could represent a potential drug target for therapeutic intervention with fungal diseases. The insights gained from this study have important implications for our mechanistic understanding of RNAi and will be discussed below.

Dcr1's dsRBD is an elongated, non-canonical dsRBD

In this study we found that the Dcr1 C-terminal domain adopts an extended dsRBD structure. Similarly, the dsRBD of the budding yeast RNase III protein Rnt1p has an extended fold with the region immediately C-terminal to the dsRBD forming a long C-terminal helix (α3, Figure 7A). This additional α-helix 3 is indispensable for the dsRBD fold and was proposed to contribute indirectly to RNA binding by helping position helix α1, which is the primary determinant of RNA recognition by Rnt1p (Leulliot et al, 2004; Wu et al, 2004). Despite this additional element, the dsRBD of Rnt1p adopts a very similar dsRBD fold compared to other dsRBDs (Figure 7B) like dsRBD2 of Xlrbpa (Ryter & Schultz, 1998) or the two dsRBDs of ADAR2 (Stefl et al, 2006). This is not what we observe in the structure of the dsRBD of *S. pombe* Dcr1. By comparing the Dcr1

extended dsRBD with others, one can see the different position of α-helix 1 relative to the rest of the domain (Figure 7B and 7C). Furthermore, the β-sheet of the Dcr1 dsRBD adopts a different curvature compared to other dsRBDs due to the involvement in the β2–β3 loop of H1312 in zinc coordination (Figure 7C). In summary, the Dcr1 dsRBD is not only extended by C33 and zinc coordination but adopts a very distinct fold compared to other dsRBDs. Remarkably, the Dcr1 dsRBD still binds dsRNA strongly, despite an extended fold, an unusual position of α-helix 1, and an unusually long β1-β2 loop. These unique features might explain why the Dcr1 dsRBD also binds dsDNA. Finally, we note that it is possible that a regular dsRNA might not be the optimal RNA target of the Dcr1 dsRBD. Similar to the dsRBD of Rnt1p that recognizes AGNN tetraloops (Leulliot et al, 2004; Wu et al, 2004), or dsRBDs of ADAR2 that have been shown to be partly sequence-specific (Stefl et al, 2010), the dsRBD of Dcr1 might preferentially recognize a particular RNA motif. It will be interesting to investigate such RNA sequence and/or structural preferences of the Dcr1 dsRBD in future studies.

Mechanism of nuclear retention

In our previous study we demonstrated that the C-terminus of Dcr1 functions to control the subcellular localization of Dcr1. We have shown that the dsRBD can function as a strong nuclear export signal, which is under negative control by C33 (Emmerth et al, 2010). The C-terminal structure of Dcr1 determined in this study provides an elegant explanation for why our previous attempts to use C33 as a nuclear retention signal on heterologous proteins have failed. Rather than interacting with a nuclear retention factor directly, C33 contributes to the overall fold of the C-terminus by providing two of the four zinc-binding residues. This results in the formation of a putative protein-protein interaction surface that is characterized by a patch of conserved residues that cluster around the side-chain of N1344 at the beginning of helix 3. This protein interaction surface would allow the attachment of Dcr1 to an as yet unidentified nuclear protein, resulting in nuclear accumulation of Dcr1. Even though we favour such a model, we cannot rule out the possibility that export-promoting features of the dsRBD are buried in the overall fold of the C-terminus and that this also contributes to nuclear retention of

Dcr1. Only if zinc coordination was disrupted or C33 was completely missing, would the export signal become exposed and Dcr1 rapidly exported. The former model is particularly appealing to us because it would allow regulation of nucleo-cytoplasmic localization of Dcr1 via modulating the interaction between Dcr1 and the nuclear retention factor. This is an intriguing possibility, which we will investigate more closely in our future studies.

The role of dicer dsRBDs

The solution structure of the Dcr1 dsRBD allowed us to design specific mutations that interfere with binding to dsRNA without disturbing the overall structure of the domain. Intriguingly, we found that key residues in the dsRBD that are required for dsRNA binding are dispensable for proper functioning of Dcr1 in the RNAi pathway. This is surprising because one would assume that dicer proteins contain dsRBDs to get a hold on their dsRNA substrates. Thus, rather than substrate binding and processing, the main function of the Dcr1 dsRBD appears to be regulation of the subcellular localization of Dcr1 independently of its affinity for dsRNA. Interestingly, dicer proteins in animals are known to interact with other dsRBD-containing proteins, such as DICER1 with TRBP (Chendrimada et al, 2005; Haase et al, 2005), or Drosophila Dcr-1 and Dcr-2 with Loquacious and R2D2, respectively (Forstemann et al, 2005; Liu et al, 2003; Saito et al, 2005). We speculate that dicer dsRBDs might generally function in nucleo-cytoplasmic transport and not necessarily substrate binding. Dicers in some organisms might therefore have acquired auxiliary dsRBD proteins that would increase their specificity or efficiency in dsRNA processing.

Conservation of the zinc-binding motif

The unusual spacing between C and H residues in the zinc-binding motif makes it difficult to search for homologies. Therefore, we restricted our analysis to proteins with at least one RNase III domain, followed by a dsRBD and a C-terminal extension of at least 15 amino acids. This analysis revealed that the CHCC zinc-binding motif is highly conserved in the C-terminal domains of yeast dicers (Figure 6 and Supplementary

Figure 9).

Whereas the CHCC zinc-binding motif is highly conserved, the remaining sequences can be rather divergent even within the yeast group. The region between helix 2 and the C-X-C motif, as well as the sequence that follows the C-X-C motif, is neither well conserved in amino acid composition nor in length (Supplementary Figure 9). For example, the region between helix 2 and the C-X-C motif in *S. pombe* Dcr1 is three amino acids shorter than in many other dicer sequences (Figure 6A and Supplementary Figure 9), suggesting the presence of an additional turn to helix 3 in other species. We also noted the presence of a buried arginine (R1311) side-chain in the core of the *S. pombe* domain, whose positive charge would not be favourable in such a hydrophobic environment. However, we note that the vicinity of the CHCC zinc-binding site, with an overall -1 net charge, is likely to counterbalance this positive charge. Interestingly, in any other CHCC zinc-binding motif-containing dsRBD, this arginine is replaced by a hydrophobic residue (Figure 6A and Supplementary Figure 9), thus forming a more classical hydrophobic core.

Finally, the high degree of conservation and its mutational intolerance make the zinc-binding motif a prime antifungal target. Ejection of zinc from HIV type 1 nucleocapsid protein by disulfide benzamides or other molecules has been shown to have an anti-viral effect (Pannecouque et al; Rice et al, 1995). Similarly, compounds reacting with the Dcr1 zinc-binding motif may offer new anti-fungal strategies to cope with the increasing incidence of invasive mycoses.

MATERIALS and METHODS

Strains and Plasmids

Fission yeast strains used in this study are listed in Supplementary Data and were grown at 30 °C in YES. All strains were constructed following a standard PCR-based protocol (Bahler et al, 1998) and confirmed by sequencing. For protein expression, the different constructs, consisting of residues 1258-1374 (long-dsRBD) or residues 1258-1358 (short-dsRBD) of *S. pombe* Dcr1 (Figure 1B), were sub-cloned in *E. coli* expression vector pET28a+ between *NdeI* and *XhoI* cloning sites. The constructs hold a

N-terminal tag whose sequence MGSSHHHHHHSSGLVPRGSHM includes a 6 histidine stretch used for protein purification. Mutagenesis was performed using the Quickchange Kit (Stratagene) following the manufacturer's recommendations.

Protein expression and purification

Proteins were overexpressed in BL21(DE3) Codon-plus (RIL) cells in either LB media or M9 minimal media supplemented with $^{15}NH_4Cl$ and ^{13}C-labeled glucose. All media were supplemented with 0.1 mM $ZnCl_2$. The cells were grown at 37 °C to OD600 ~0.4, cooled down at 18 °C and induced at OD600 ~0.6 by adding isopropyl-β-D-thiogalactopyranoside to a final concentration of 0.25 mM. Cells were harvested 20 h after induction by centrifugation. The protein of interest was purified from both the soluble and insoluble fraction of *E. coli.* Cell pellets were resuspended in lysis buffer (Tris-HCl pH 7.5 50 mM, NaCl 500 mM, imidazole 20 mM, $ZnCl_2$ 0.1 mM, DTT 1 mM, Tween-20 0.2 %), and lysed by sonication. Cell lysates were centrifuged 30 min at 45,000 g. When the protein is expressed in LB media, about 40 % remains in the soluble fraction, whereas it expresses almost exclusively in the inclusion bodies when expressed in M9 minimal media. On the one hand, the supernatant was loaded on a Ni-NTA column on a ÄKTA Prime purification system (Amersham Biosciences), and the protein of interest was eluted with an imidazole gradient. The fractions containing the protein were pooled, dialyzed against the NMR Buffer (NaPi pH 7.0 25 mM, KCl 75 mM, DTT 2 mM and $ZnCl_2$ 10 μM), and concentrated to ~0.5 mM using Vivaspin 5000 MWCO (Sartorius Stedim Biotech). On the other hand, the pellet was dissolved in 6 M guanidinium chloride and DTT was added to a final concentration of 1 mM. The resuspended proteins from the insoluble fraction were then loaded on a Ni-NTA column equilibrated with Tris-HCl pH 7.5 50 mM, guanidinium chloride 4 M. The bound protein was then eluted with an imidazole gradient. The fractions containing the protein of interest were pooled, concentrated and subjected to a refolding protocol. Shortly, small amounts of purified protein were added step by step in a large amount of refolding buffer (Tris-HCl pH 7.5 50 mM, KCl 300 mM, 2-mercaptoethanol 1 mM and $ZnCl_2$ 0.1 mM) at 4 °C under stirring. The protein solution was then dialyzed against the NMR buffer and

concentrated as for the protein purified from the soluble fraction. The protein purified from the insoluble fraction gives exactly the same ^{1}H NMR spectrum as the one purified from the soluble fraction. This shows that both proteins adopt the same structure and thus validate our refolding protocol.

NMR Spectroscopy

All NMR spectra were recorded at 298 K in a buffer containing 25 mM NaPi pH 7.0, 75 mM KCl, 2 mM DTT, 10 μM $ZnCl_2$ on Bruker AVIII-500 MHz, AVIII-600 MHz, AVIII-700 MHz and Avance-900 MHz spectrometers (all equipped with a cryoprobe). The data were processed using TOPSPIN 2.1 (Bruker) and analyzed with Sparky (http://www.cgl.ucsf.edu/home/sparky/). Protein resonances were assigned using 2D (^{1}H-^{15}N)-HSQC, 2D (^{1}H-^{13}C)-HSQC, 3D HNCA, 3D HNCACB, 3D CBCA(CO)NH, 3D TOCSY-(^{1}H-^{15}N)-HSQC, 3D (H)CCH-TOCSY, 3D NOESY-(^{1}H-^{15}N)-HSQC and two 3D NOESY-(^{1}H-^{13}C)-HSQC optimized for the observation of protons attached to aliphatic carbons and to aromatic carbons, respectively. In addition, the assignment of aromatic protons was conducted using 2D (^{1}H-^{1}H)-TOCSY and 2D (^{1}H-^{1}H)-NOESY measured in D_2O. We recorded all 3D NOESY spectra with a mixing time of 150 ms and the 2D NOESY spectra with a mixing time of 120 ms. Long-range ^{1}H-^{15}N HSQC spectra was measured at 600 MHz and the delay during which ^{15}N and ^{1}H signals become antiphase was set to 22 ms to refocus magnetization arising from $^{1}J_{NH}$ couplings (Pelton et al, 1993) .

Protein structure determination

Automated NOE cross-peak assignments (Herrmann et al, 2002a) and structure calculations with torsion-angle dynamics (Guntert et al, 1997) were performed using the macro noeassign of the software package CYANA 2.1 (Guntert, 2004). Peak lists of the four NOESY spectra were generated as input with the program ATNOS (Herrmann et al, 2002b) and manually cleaned to remove artefact peaks. The input also contained 41 hydrogen-bond restraints and restraints that defined the coordination geometry around the zinc. Hydrogen bonded amides were identified as slowly exchanging protons in

presence of D_2O. Their bonding partner was identified from preliminary structures. We calculated 100 independent structures that we refined in a water shell with the program CNS 1.21 (Brunger, 2007; Brunger et al, 1998) using a water-refinement protocol successfully employed to refine a large set of NMR structures calculated with CYANA (Nederveen et al, 2005). The coordination geometry around the zinc has been defined with 4 distance restraints between the metal ion and the four ligand atoms that directly contact the metal center (Cysteins $S\gamma$ and Histidine $N\varepsilon$), 10 bond angle and one dihedral angle restraints that defined a tetrahedral coordination sphere (Neuhaus et al, 1992). The 20 best energy structures were analyzed using PROCHECK-NMR (Laskowski et al, 1996). Structures were visualized and figures were prepared using program MOLMOL (Koradi et al, 1996) and PYMOL (www.pymol.org).

Atomic Absorption Spectroscopy

Samples of unlabeled long-dsRBD purified from the soluble protein fraction, were extensively dialyzed against Ammonium acetate pH 7.0 100 mM, and analyzed for Zn^{2+} content using an atomic absorption spectrometer from Varian (model spectrAA 220 FS). Zn^{2+} standards of 0—1.5 ppm were prepared from a 1000 ppm calibrated stock solution in the same buffer as the protein sample. Absorption measurement was repeated 5 times for standards and sample. The relative standard deviation for each sample was below 1 %.

Mass Spectrometry

Samples for mass spectrometry were prepared either under denaturing conditions in aqueous 50 % acetonitrile and 0.1 % formic acid to a final pH of 2.0 to assess the mass of the apo form of the protein, or under native conditions in 10 mM ammonium acetate pH 7.0 to measure the mass of the holo form. Samples were analyzed at the Functional Genomics Center Zurich on an electrospray time-of-flight mass spectrometer. Deconvoluted mass spectra were obtained using the MaxEnt1 software.

Nucleic acid samples used in ITC binding assays

To assess the nucleic acid binding properties of the dsRBD, we produced two complementary RNA fragments of arbitrary sequence by *in vitro* transcription with T7 polymerase (RNAfwd of sequence 5' GGGAUCAAUAUGCUAAGCGAUCCC 3' and RNArev being the reverse complement). RNA was purified by anion-exchange high-pressure liquid chromatography under denaturing conditions. The two complementary RNA strands were mixed to a 1:1 ratio, heated-up at 95 °C for 5 min and slowly cooled down to allow formation of dsRNA. DNA oligonucleotides of the same sequence (DNAfwd of sequence 5' GGGATCAATATGCTAAGCGATCCC 3' and DNArev being the reverse complement) were purchased from Microsynth AG (Switzerland) and purified by anion-exchange high-pressure liquid chromatography under denaturing conditions. The dsDNA duplex was formed as described for the dsRNA. ITC experiments with ssRNA and ssDNA were performed with RNAfwd and DNAfwd respectively.

Isothermal titration calorimetry

ITC experiments were performed on a VP-ITC instrument (MicroCal) calibrated according to the manufacturer's instructions. The samples of protein and nucleic acids were prepared in and dialyzed against the ITC buffer (NaPi pH 7.0 25 mM, KCl 75 mM, 2-mercaptoethanol 2 mM). The concentration of protein and nucleic acid was determined using OD absorbance at 280 and 260 nm, respectively. The sample cell (1.4 mL) was loaded with 2 μM of nucleic acid (dsRNA, dsDNA, ssRNA, ssDNA); dsRBD and mutant concentration in the syringe were between 90 and 130 μM. Titration experiments were done at 25 °C and typically consisted of 35-45 injections, each of 6-8 μL volume with a 5-min interval between additions. Stirring rate was 307 r.p.m. Raw data were integrated, corrected for nonspecific heats, normalized for the molar concentration. Three parameters were fitted (the association constant Ka, the binding enthalpy ΔH and the number of site N) using the equation for 1:1 binding model.

Yeast live-cell fluorescence microscopy

S. pombe pre-cultures were grown in YES (sterile filtered components only) for 8 h at 30

°C, diluted to 10^5 cells/mL, and grown for 14-16 h at 30 °C to a concentration of about 5×10^6 cells/mL. Microscopy was performed on cells spread on agarose patches containing YES medium with 3 % glucose. Images were captured on a Delta Vision built of an Olympus IX70 widefield microscope equipped with a CoolSNAP HQ2/ICX285 camera. Image stacks of 12–15 µm Z-distance were acquired with a Z-step size of 0.2 µm and deconvolved using the softworks (Delta Vision) software.

RNA analysis

RNA isolation, quantitative real-time RT-PCR, tiling arrays, as well as small RNA deep sequencing and analysis of the data was performed as described (Emmerth et al. 2010).

Multiple sequence alignment

All entries in UniProtKB (release 2010_10) were searched for Pfam domain PF00636 (RNase3). Sequences containing one or more RNase3 domains were cut after the (last) predicted RNase3 domain, and the Pfam domain PF00035 (double-stranded RNA binding motif) was aligned to the remaining C-terminal part. The fragments were further selected for the presence of at least 15 amino acids following the putative RNA binding motif and the presence of a Cys, followed by a His followed by Cys-X-Cys distributed over the whole length of the fragment. The distances between the first Cys and the His, and the His and the Cys-X-Cys in the remaining 58 fragments were manually compared with the numbers from *S. pombe* and *S. japonicus*. Thereby, another 15 proteins were eliminated. Only one protein was selected from proteins originating from the same species, but different strains. Proteins originating from the same species and the same strain, which were only different in 5 amino acids were reduced to one fragment as well. At this stage, C5WYA8 (a putative uncharacterized protein from *Sorghum bicolor*, the last non-fungi sequence) was removed, as it distorted the multiple sequence alignment created with Clustalw2 (Larkin et al, 2007). The multiple sequence alignment was manually adjusted and colored ('Clustalx') with Jalview (Waterhouse et al, 2009) and sequences to the left and the right of the dsRNA binding motif were removed as indicated in the figure. The UniProt accession for the *S. pombe* dicer protein is

highlighted in red. Accessions in bold are proteins expressed in human or plant pathogenic yeast as listed in Supplementary Table II.

SUPPLEMENTARY DATA

Supplementary data are available online.

ACCESSION NUMBERS

The chemical shifts of Dcr1 dsRBD have been deposited in the BioMagResBank under accession number 17315. The coordinates of the structure have been deposited in the Protein Data Bank under accession code 2L6M.

Tiling array and deep sequencing data are deposited at GEO (http://www.ncbi.nlm.nih.gov/geo, accession numbers GSE30586, and GSE30549, respectively).

ACKNOWLEDGMENT

We are grateful to Serge Chesnov (Functional Genomics Center Zürich) for the measurements of ESI-MS spectrum, Daniel Fodor (ETH Zürich) for technical assistance with AAS, Heinz Gut and Jeremy Keusch (FMI Basel) for help with construct design and protein expression, and Irene Beusch (ETH Zürich) for her involvement in chemical shift assignment. We thank Helge Grosshans, Susan Gasser and Yvon Jaillais for critical reading of the manuscript and all the members of the Allain and Bühler laboratories for fruitful discussions. Research in the Bühler laboratory is supported by the Novartis Research Foundation, the Swiss National Science Foundation, the Gebert Rüf Foundation and the EMBO YIP program. Research in the Allain laboratory is supported by the Swiss National Science Foundation (Nr 310030-131031) and the SNF-NCCR structural biology. P.B. is supported by the Postdoctoral ETH Fellowship Program.

AUTHOR CONTRIBUTIONS

M.B. and F.H.-T.A. designed the project; P.B. prepared protein and RNA samples for structural studies and ITC, measured and analyzed NMR data, and performed structure

calculations and ITC measurements; P.B. and F.H.-T.A. analyzed the structure; S.E. and Y.S. generated mutant yeast strains; Y.S. cloned Dcr1 fragments for recombinant protein production in E. coli and insect cells, and performed protein expression tests. S.E. performed all the functional analysis; H-R. H. performed homology searches. P.B., S.E., F.H.-T.A. and M.B. wrote the manuscript; all authors discussed the results and approved the manuscript.

CONFLICT OF INTEREST

Patent application (EP10191173.3) has been filed.

REFERENCES

Auld DS (2001) Zinc coordination sphere in biochemical zinc sites. *Biometals* **14:** 271-313

Bahler J, Wu JQ, Longtine MS, Shah NG, McKenzie A, III, Steever AB, Wach A, Philippsen P, Pringle JR (1998) Heterologous modules for efficient and versatile PCR-based gene targeting in Schizosaccharomyces pombe. *Yeast* **14:** 943-951

Bass BL, Hurst SR, Singer JD (1994) Binding properties of newly identified Xenopus proteins containing dsRNA-binding motifs. *Curr Biol* **4:** 301-314

Baulcombe D (2004) RNA silencing in plants. *Nature* **431:** 356-363

Billy E, Brondani V, Zhang H, Muller U, Filipowicz W (2001) Specific interference with gene expression induced by long, double-stranded RNA in mouse embryonal teratocarcinoma cell lines. *Proc Natl Acad Sci U S A* **98:** 14428-14433

Brunger AT (2007) Version 1.2 of the Crystallography and NMR system. *Nat Protoc* **2:**

2728-2733

Brunger AT, Adams PD, Clore GM, DeLano WL, Gros P, Grosse-Kunstleve RW, Jiang JS, Kuszewski J, Nilges M, Pannu NS, Read RJ, Rice LM, Simonson T, Warren GL (1998) Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr D Biol Crystallogr* **54:** 905-921

Buhler M, Verdel A, Moazed D (2006) Tethering RITS to a Nascent Transcript Initiates RNAi- and Heterochromatin-Dependent Gene Silencing. *Cell* **125:** 873-886

Burd CG, Dreyfuss G (1994) Conserved structures and diversity of functions of RNA-binding proteins. *Science* **265:** 615-621

Bycroft M, Grunert S, Murzin AG, Proctor M, St Johnston D (1995) NMR solution structure of a dsRNA binding domain from Drosophila staufen protein reveals homology to the N-terminal domain of ribosomal protein S5. *EMBO J* **14:** 3563-3571

Cam HP, Sugiyama T, Chen ES, Chen X, FitzGerald PC, Grewal SI (2005) Comprehensive analysis of heterochromatin- and RNAi-mediated epigenetic control of the fission yeast genome. *NatGenet* **37:** 809-819

Carmell MA, Xuan Z, Zhang MQ, Hannon GJ (2002) The Argonaute family: tentacles that reach into RNAi, developmental control, stem cell maintenance, and tumorigenesis. *Genes Dev* **16:** 2733-2742

Chen T, Brownawell AM, Macara IG (2004) Nucleocytoplasmic shuttling of JAZ, a new cargo protein for exportin-5. *Mol Cell Biol* **24:** 6608-6619

Chendrimada TP, Gregory RI, Kumaraswamy E, Norman J, Cooch N, Nishikura K, Shiekhattar R (2005) TRBP recruits the Dicer complex to Ago2 for microRNA processing

and gene silencing. *Nature* **436:** 740-744

Colmenares SU, Buker SM, Buhler M, Dlakic M, Moazed D (2007) Coupling of double-stranded RNA synthesis and siRNA generation in fission yeast RNAi. *Mol Cell* **27:** 449-461

Emmerth S, Schober H, Gaidatzis D, Roloff T, Jacobeit K, Buhler M (2010) Nuclear retention of fission yeast dicer is a prerequisite for RNAi-mediated heterochromatin assembly. *Dev Cell* **18:** 102-113

Fierro-Monti I, Mathews MB (2000) Proteins binding to duplexed RNA: one motif, multiple functions. *Trends Biochem Sci* **25:** 241-246

Forstemann K, Tomari Y, Du T, Vagin VV, Denli AM, Bratu DP, Klattenhoff C, Theurkauf WE, Zamore PD (2005) Normal microRNA maturation and germ-line stem cell maintenance requires Loquacious, a double-stranded RNA-binding domain protein. *PLoS Biol* **3:** e236

Fritz J, Strehblow A, Taschner A, Schopoff S, Pasierbek P, Jantsch MF (2009) RNA-regulated interaction of transportin-1 and exportin-5 with the double-stranded RNA-binding domain regulates nucleocytoplasmic shuttling of ADAR1. *Mol Cell Biol* **29:** 1487-1497

Ghildiyal M, Zamore PD (2009) Small silencing RNAs: an expanding universe. *Nat Rev Genet* **10:** 94-108

Grishok A, Pasquinelli AE, Conte D, Li N, Parrish S, Ha I, Baillie DL, Fire A, Ruvkun G, Mello CC (2001) Genes and mechanisms related to RNA interference regulate expression of the small temporal RNAs that control C. elegans developmental timing. *Cell* **106:** 23-34

Guntert P (2004) Automated NMR structure calculation with CYANA. *Methods Mol Biol* **278:** 353-378

Guntert P, Mumenthaler C, Wuthrich K (1997) Torsion angle dynamics for NMR structure calculation with the new program DYANA. *J Mol Biol* **273:** 283-298

Gwizdek C, Ossareh-Nazari B, Brownawell AM, Evers S, Macara IG, Dargemont C (2004) Minihelix-containing RNAs mediate exportin-5-dependent nuclear export of the double-stranded RNA-binding protein ILF3. *J Biol Chem* **279:** 884-891

Haase AD, Jaskiewicz L, Zhang H, Laine S, Sack R, Gatignol A, Filipowicz W (2005) TRBP, a regulator of cellular PKR and HIV-1 virus expression, interacts with Dicer and functions in RNA silencing. *EMBO Rep* **6:** 961-967

Herrmann T, Guntert P, Wuthrich K (2002a) Protein NMR structure determination with automated NOE assignment using the new software CANDID and the torsion angle dynamics algorithm DYANA. *J Mol Biol* **319:** 209-227

Herrmann T, Guntert P, Wuthrich K (2002b) Protein NMR structure determination with automated NOE-identification in the NOESY spectra using the new software ATNOS. *J Biomol NMR* **24:** 171-189

Huisinga KL, Elgin SC (2009) Small RNA-directed heterochromatin formation in the context of development: what flies might learn from fission yeast. *Biochim Biophys Acta* **1789:** 3-16

Hutvagner G, Simard MJ (2008) Argonaute proteins: key players in RNA silencing. *Nat Rev Mol Cell Biol* **9:** 22-32

Kharrat A, Macias MJ, Gibson TJ, Nilges M, Pastore A (1995) Structure of the dsRNA binding domain of E. coli RNase III. *EMBO J* **14:** 3572-3584

Koradi R, Billeter M, Wuthrich K (1996) MOLMOL: a program for display and analysis of macromolecular structures. *J Mol Graph* **14:** 51-55, 29-32

Laity JH, Lee BM, Wright PE (2001) Zinc finger proteins: new insights into structural and functional diversity. *Curr Opin Struct Biol* **11:** 39-46

Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG (2007) Clustal W and Clustal X version 2.0. *Bioinformatics* **23:** 2947-2948

Laskowski RA, Rullmannn JA, MacArthur MW, Kaptein R, Thornton JM (1996) AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. *J Biomol NMR* **8:** 477-486

Legge GB, Martinez-Yamout MA, Hambly DM, Trinh T, Lee BM, Dyson HJ, Wright PE (2004) ZZ domain of CBP: an unusual zinc finger fold in a protein interaction module. *J Mol Biol* **343:** 1081-1093

Lei QP, Cui X, Kurtz DM, Jr., Amster IJ, Chernushevich IV, Standing KG (1998) Electrospray mass spectrometry studies of non-heme iron-containing proteins. *Anal Chem* **70:** 1838-1846

Leulliot N, Quevillon-Cheruel S, Graille M, van Tilbeurgh H, Leeper TC, Godin KS, Edwards TE, Sigurdsson ST, Rozenkrants N, Nagel RJ, Ares M, Varani G (2004) A new alpha-helical extension promotes RNA binding by the dsRBD of Rnt1p RNAse III. *EMBO J* **23:** 2468-2477

Liu Q, Rand TA, Kalidas S, Du F, Kim HE, Smith DP, Wang X (2003) R2D2, a bridge between the initiation and effector steps of the Drosophila RNAi pathway. *Science* **301:** 1921-1925

Macchi P, Brownawell AM, Grunewald B, DesGroseillers L, Macara IG, Kiebler MA (2004) The brain-specific double-stranded RNA-binding protein Staufen2: nucleolar accumulation and isoform-specific exportin-5-dependent export. *J Biol Chem* **279:** 31440-31444

Macrae IJ, Zhou K, Li F, Repic A, Brooks AN, Cande WZ, Adams PD, Doudna JA (2006) Structural basis for double-stranded RNA processing by Dicer. *Science* **311:** 195-198

Moazed D (2009) Small RNAs in transcriptional gene silencing and genome defence. *Nature* **457:** 413-420

Motamedi MR, Verdel A, Colmenares SU, Gerber SA, Gygi SP, Moazed D (2004) Two RNAi complexes, RITS and RDRC, physically interact and localize to noncoding centromeric RNAs. *Cell* **119:** 789-802

Nederveen AJ, Doreleijers JF, Vranken W, Miller Z, Spronk CA, Nabuurs SB, Guntert P, Livny M, Markley JL, Nilges M, Ulrich EL, Kaptein R, Bonvin AM (2005) RECOORD: a recalculated coordinate database of 500+ proteins from the PDB using restraints from the BioMagResBank. *Proteins* **59:** 662-672

Neuhaus D, Nakaseko Y, Schwabe JW, Klug A (1992) Solution structures of two zinc-finger domains from SWI5 obtained using two-dimensional 1H nuclear magnetic resonance spectroscopy. A zinc-finger structure with a third strand of beta-sheet. *J Mol Biol* **228:** 637-651

Pannecouque C, Szafarowicz B, Volkova N, Bakulev V, Dehaen W, Mely Y, Daelemans

D (2010) Inhibition of HIV-1 replication by a bis-thiadiazolbenzene-1,2-diamine that chelates zinc ions from retroviral nucleocapsid zinc fingers. *Antimicrob Agents Chemother* **54:** 1461-1468

Pelton JG, Torchia DA, Meadow ND, Roseman S (1993) Tautomeric states of the active-site histidines of phosphorylated and unphosphorylated IIIGlc, a signal-transducing protein from Escherichia coli, using two-dimensional heteronuclear NMR techniques. *Protein Sci* **2:** 543-558

Provost P, Dishart D, Doucet J, Frendewey D, Samuelsson B, Radmark O (2002) Ribonuclease activity and RNA binding of recombinant human Dicer. *EMBO J* **21:** 5864-5874

Ramos A, Grunert S, Adams J, Micklem DR, Proctor MR, Freund S, Bycroft M, St Johnston D, Varani G (2000) RNA recognition by a Staufen double-stranded RNA-binding domain. *EMBO J* **19:** 997-1009

Reinhart BJ, Bartel DP (2002) Small RNAs correspond to centromere heterochromatic repeats. *Science* **297:** 1831

Rice WG, Supko JG, Malspeis L, Buckheit RW, Jr., Clanton D, Bu M, Graham L, Schaeffer CA, Turpin JA, Domagala J, Gogliotti R, Bader JP, Halliday SM, Coren L, Sowder RC, 2nd, Arthur LO, Henderson LE (1995) Inhibitors of HIV nucleocapsid protein zinc fingers as candidates for the treatment of AIDS. *Science* **270:** 1194-1197

Ryter JM, Schultz SC (1998) Molecular basis of double-stranded RNA-protein interactions: structure of a dsRNA-binding domain complexed with dsRNA. *EMBO J* **17:** 7505-7513

Saito K, Ishizuka A, Siomi H, Siomi MC (2005) Processing of pre-microRNAs by the

Dicer-1-Loquacious complex in Drosophila cells. *PLoS Biol* **3:** e235

Saunders LR, Barber GN (2003) The dsRNA binding protein family: critical roles, diverse cellular functions. *FASEB J* **17:** 961-983

Siomi H, Siomi MC (2009) On the road to reading the RNA-interference code. *Nature* **457:** 396-404

St Johnston D, Brown NH, Gall JG, Jantsch M (1992) A conserved double-stranded RNA-binding domain. *Proc Natl Acad Sci U S A* **89:** 10979-10983

Stefl R, Oberstrass FC, Hood JL, Jourdan M, Zimmermann M, Skrisovska L, Maris C, Peng L, Hofr C, Emeson RB, Allain FH (2010) The solution structure of the ADAR2 dsRBM-RNA complex reveals a sequence-specific readout of the minor groove. *Cell* **143:** 225-237

Stefl R, Xu M, Skrisovska L, Emeson RB, Allain FH (2006) Structure and specific RNA binding of ADAR2 double-stranded RNA binding motifs. *Structure* **14:** 345-355

Strehblow A, Hallegger M, Jantsch MF (2002) Nucleocytoplasmic distribution of human RNA-editing enzyme ADAR1 is modulated by double-stranded RNA-binding domains, a leucine-rich export signal, and a putative dimerization domain. *Mol Biol Cell* **13:** 3822-3835

Sugiyama T, Cam H, Verdel A, Moazed D, Grewal SI (2005) RNA-dependent RNA polymerase is an essential component of a self-enforcing loop coupling heterochromatin assembly to siRNA production. *ProcNatlAcadSciUSA* **102:** 152-157

van Wolfswinkel JC, Ketting RF (2011) The role of small non-coding RNAs in genome stability and chromatin organization. *J Cell Sci* **123:** 1825-1839

Verdel A, Jia S, Gerber S, Sugiyama T, Gygi S, Grewal SI, Moazed D (2004) RNAi-mediated targeting of heterochromatin by the RITS complex. *Science* **303:** 672-676

Volpe T, Schramke V, Hamilton GL, White SA, Teng G, Martienssen RA, Allshire RC (2003) RNA interference is required for normal centromere function in fission yeast. *Chromosome Res* **11:** 137-146

Volpe TA, Kidner C, Hall IM, Teng G, Grewal SI, Martienssen RA (2002) Regulation of heterochromatic silencing and histone H3 lysine-9 methylation by RNAi. *Science* **297:** 1833-1837

Waterhouse AM, Procter JB, Martin DM, Clamp M, Barton GJ (2009) Jalview Version 2--a multiple sequence alignment editor and analysis workbench. *Bioinformatics* **25:** 1189-1191

Woolcock KJ, Gaidatzis D, Punga T, Buhler M (2011) Dicer associates with chromatin to repress genome activity in Schizosaccharomyces pombe. *Nat Struct Mol Biol* **18:** 94-99

Wu H, Henras A, Chanfreau G, Feigon J (2004) Structural basis for recognition of the AGNN tetraloop RNA fold by the double-stranded RNA-binding domain of Rnt1p RNase III. *Proc Natl Acad Sci U S A* **101:** 8307-8312

FIGURE LEGENDS

Figure 1

NMR-solution structure of the Dcr1 C-terminus

(A) Domain architecture of *S. pombe* Dcr1 (not drawn to scale).

(B) Sequence of the C-terminus (dsRBD + C33) with the corresponding secondary structures indicated (color-code is the same as in Figure 3). The CHCC zinc coordination motif is highlighted in red. Arrows indicate the C-terminal ends of the different constructs used for structure determination. The residues highlighted in yellow are involved in dsRNA binding. Amino acid numbers refer to the *S. pombe* Dcr1 protein (UniprotKB Q09884).

(C) NMR ensemble. Overlay of the 20 final structures with color-coded secondary-structure elements: α1 in blue, β-strands 1-3 in red-orange-yellow, α2 in green and α3 in purple. The four zinc ligands (C1275, H1312, C1350 and C1352) are represented as sticks in light-blue. The zinc ion is represented as a black dot.

(D) Cartoon representation of the lowest-energy structure. The same color code is used for secondary structure elements. The zinc ion is represented as a grey sphere.

(E) Visualization of the extended dsRBD fold on the protein surface. The canonical elements of the dsRBD are represented in gray and the C-terminal extension in red. (Left panel). Same orientation of the domain as in (A) and (B). (Right panel) Side view.

Figure 2

Close-up view of the hydrophobic core of the domain.

(A) Front view. (B) Side view.

Side-chains belonging to the three α helices and to the β sheet are shown as green and yellow sticks, respectively. The four zinc ligands are colored in light-blue. The zinc ion is represented as a black dot.

Figure 3

Structuring of the C-terminus of Dcr1 via coordination of zinc is essential for RNAi-mediated heterochromatin formation.

(A) Live-cell imaging of wild-type GFP-Dcr1 and the zinc motif mutant GFP-Dcr1-SHSS. Scale bars = 2 µm.

(B) Quantitative real-time RT-PCR showing that *dcr1-SHSS* cannot silence centromeric repeats. *cendg*, *cendh* and *imr1R::ura4+* RNA levels were normalized to *act1+* RNA and are shown in relation to *dcr1+* cells. Error bars represent standard deviations (STDEV).

(C) Total small RNA profiles of wild-type, *dcr1Δ* and *dcr1-SHSS* cells as determined by Illumina Sequencing. Major small RNA classes relative to the total amount of small RNAs sequenced are indicated. See Supplementary Table I for read numbers.

(D) ChIP experiment showing that mutations in the zinc coordinating motif of Dcr1 (Dcr1-SHSS) abolish H3K9 methylation at a centromeric *ura4+* transgene (*imr1R::ura4+*). Fold-enrichment values from one representative experiment, normalized to *act1+*, are shown. The value for *dcr1Δ* cells was set to 1. Error bars represent standard deviations (STDEV).

Figure 4

Key residues in the dsRBD that are required for dsRNA binding are dispensable for proper functioning of Dcr1 in the RNAi pathway.

(A) Location of residues important for the binding of dsRNA.

(B) Affinity of the Dcr1 dsRBD for dsRNA as determined by ITC.

(C) ITC measurements with dsRNA and Dcr1 dsRBD mutants (R1322A, ∆loop 2 and K1265A). Data for ∆loop 2 and K1265A are shifted on the y-axis for clarity.

(D) Live-cell imaging of GFP-Dcr1-R1322A and GFP-Dcr1-R1322A-∆C33. Scale bars = 2 µm.

(E) Live-cell imaging of GFP-Dcr1-∆loop2. Scale bars = 2 µm.

(F) Quantitative real-time RT-PCR showing that heterochromatic gene silencing is not affected in dsRBD mutants that cannot bind dsRNA. *cendg*, *cendh* and *imr1R::ura4+* RNA levels are shown relative to *dcr1+* cells and were normalized to *act1+* RNA. Error

bars represent standard deviations (STDEV).

(G) Size distribution and the 5'-most nucleotide of centromeric sRNAs from *dcr1+* and *dcr1-Δloop2* cells.

Figure 5

A cluster of conserved residues is required for proper localization of Dcr1.

(A) Solvent exposed residues with conserved biochemical properties among *S. pombe*, *S. octosporus* and *S. cryophilus* are displayed in color on the surface of the domain. Residues that might be conserved for dsRNA binding are shown in yellow, H1312 which is conserved for zinc coordination is colored in blue. All other residues are shown in red. (Left) Front view. (Right) Back view.

(B) Live-cell imaging of GFP-Dcr1-R1334A. Scale bars = 2 µm.

(C) Live-cell imaging of GFP-Dcr1-Y1348A, GFP-Dcr1-N1344A, and GFP-Dcr1-S1349A. Scale bars = 2 µm.

(D) Live-cell imaging of a strain expressing GFP-Dcr1 containing the mutations Y1348A, N1344A, S1349A (GPF-Dcr1-triple). Scale bars = 2 µm.

Figure 6

Mutations in the conserved zinc coordination motif of Dcr1 are toxic in *S. pombe.*

(A) Conservation of the Dcr1 CHCC motif in dicer proteins of human pathogenic yeasts. The highly conserved residues are indicated by asterisks. Amino acid numbers refer to the *S. pombe* Dcr1 protein (UniprotKB Q09884).

(B) Single wild-type (CHCC) and zinc coordination motif mutant (SHSS) cells were dissected onto a 5x6 matrix on EMMc-agar plates and grown at 30 °C for ~5-6 days.

(C) Heatmap displaying the genes that were up- or downregulated at least 1.5-fold (p=0.05) in *dcr1SHSS* cells compared to wild-type on *S. pombe* tiling arrays. Artificially scaled expression values are shown for the strains indicated. Three biological replicates were performed.

Figure 7

Structural comparison with other dsRBDs.

(A) Structures of Xlrbpa dsRBD2 in gray (PDB code 1DI2), human ADAR2 dsRBD1 in green (PDB code 2L3C) and yeast RNase III Rnt1p in red (PDB code 1T4O) were superimposed over Cα atoms over the entire domain. *S. pombe* Dcr1 dsRBD in blue was superimposed with the other structures over Cα atoms of the C-terminal part of β-strand 3, β3-α2 loop and helix 2 in order to emphasize the differences in the position and the relative orientation of helix 1 and the β-sheet surface.

(B) Relative position of helix 1 and helix 2. Only helices are displayed as ribbon. (Left panel) Xlrbpa dsRBD2 and *A. aeolicus* RNase III (PDB code 1RC7) are displayed in gray, human ADAR2 dsRBD1 in green, and *S. pombe* Dcr1 dsRBD in blue. (Right panel) Yeast RNase III Rnt1p is displayed in red, and *S. pombe* Dcr1 dsRBD in blue.

(C) Superposition of Xlrbpa dsRBD2 in yellow and *S. pombe* Dcr1 dsRBD in blue over Cα atoms of the C-terminal part of β-strand 3, β3-α2 loop and helix 2, showing the unusual position of helix 1 and the β-sheet relative to helix 2.

TABLES

Table I

Zinc binding determination by ESI-MS and Atomic Absorption Spectroscopy[a]

observed apoprotein mass (Da)	observed holoprotein mass (Da)	expected apoprotein mass (Da)	expected holoprotein mass with 1 Zn^{2+} (Da)	atomic absorption spectroscopy (zinc/protein ratio)
15499.0 ± 0.5	15562.0 ± 0.5	15498.8	15562.2	1.2 ± 0.1

[a] the long-dsRBD protein construct was analyzed for bound metal by ESI-MS in 10 mM ammonium acetate at pH 7.0 to observe the holo form of the protein and under denaturing conditions (0.1 % formic acid, 50 % acetonitrile) to observe the apo form of the protein. The expected neutral mass of the holoprotein is calculated according to holo = apo + Zn^{2+} - $2H^{+}$, where M_{Zn}=65.4 Da (Lei et al, 1998). For AAS, the protein concentration was determined by UV absorbance at 280 nm taking a theoretical extinction coefficient ε = 8940 M^{-1} cm^{-1}.

Table II

NMR experimental restraints and structural statistics

Distance restraints	
Total NOE	2308
Intraresidue	520
Sequential	600
Medium range (|i-j| < 5 residues)	521
Long range (|i-j| ≥ 5 residues)	667
Hydrogen bond	41
Zinc coordination restraints	
Distance restraints	4
Bond angles	10
Dihedral angles	1
Structure statistics	
NOE violations (mean ± s.d.)	
Number of NOE violations > 0.2 Å	4.7 ± 1.8
Maximum NOE violation (Å)	0.30 ± 0.05
R.m.s.d. from average structure[a] (Å)	
Backbone	0.46 ± 0.12 Å
Heavy atoms	0.95 ± 0.14 Å
Mean deviation from ideal covalent geometry	
Bond length (Å)	0.007
Bond angles (°)	0.9
Ramachandran analysis	
Most favored region	79.6 %
Allowed region	20.2 %
Disallowed region	0.2 %

[a] Protein r.m.s.d. was calculated using residues 1262-1286,1303-1353 for the ensemble of 20 refined structures.

Figure 1

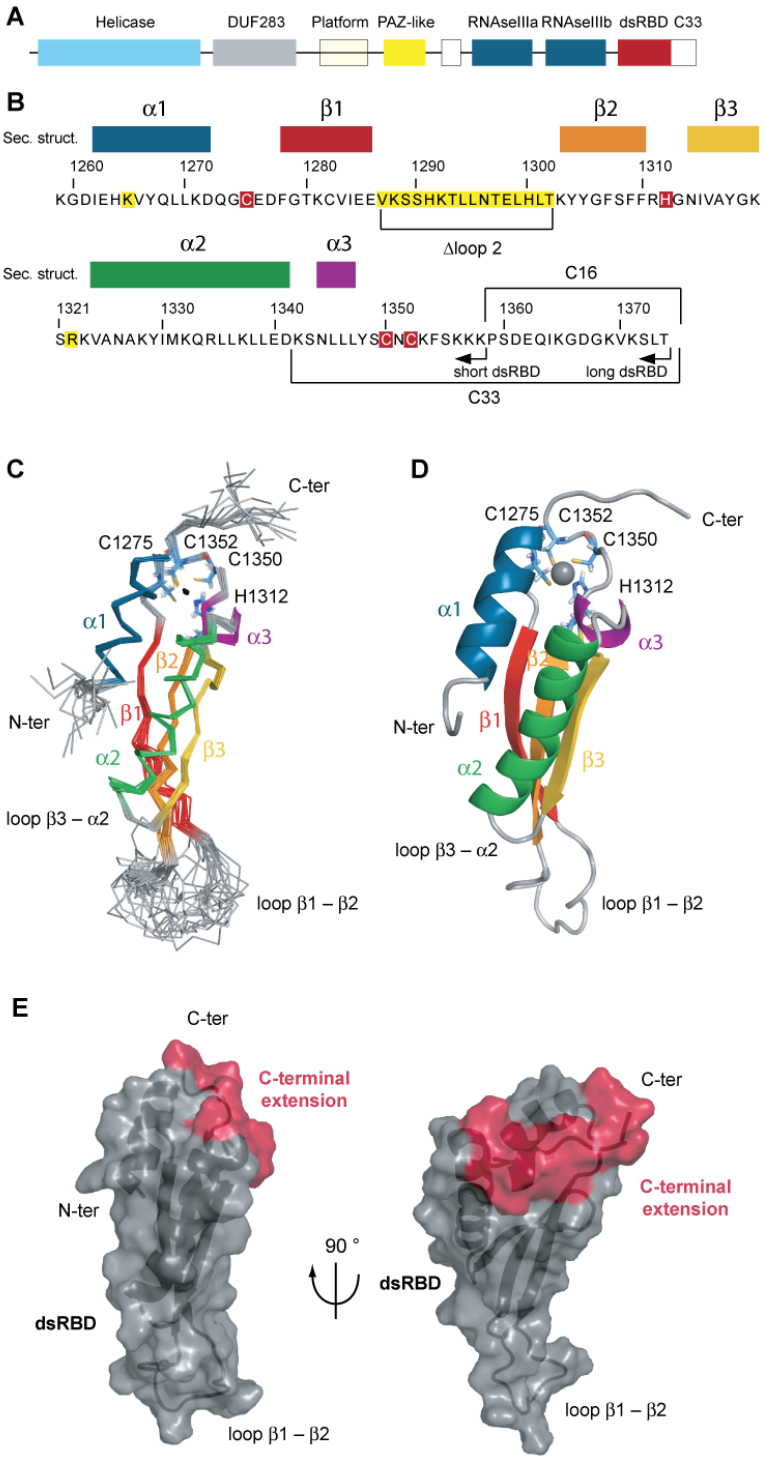

A
Helicase
DUF283
Platform
PAZ-like
RNaseIIIa
RNaseIIIb
dsRBD
C33
B
α1
β1
β2
β3
Sec. struct.
1260
1270
1280
1290
1300
1310
KGDIEHKVYQLLKDQGCEDFGTKCVIEEVKSSHKTLLNTELHLTKYYGFSFFRHGNIVAYGK
Δloop 2
α2
α3
Sec. struct.
C16
1321
1330
1340
1350
1360
1370
SRKVANAKYIMKQRLLKLLEDKSNLLLYSCNCKFSKKKPSDEQIKGDGKVKSLT
short dsRBD
long dsRBD
C33
C
C-ter
C1275
C1352
C1350
H1312
α1
α3
β2
β1
N-ter
α2
β3
loop β3 – α2
loop β1 – β2
D
C1275
C1352
C-ter
C1350
H1312
α1
α3
β2
N-ter
β1
α2
β3
loop β3 – α2
loop β1 – β2
E
C-ter
C-terminal extension
N-ter
90 °
dsRBD
dsRBD
C-ter
C-terminal extension
loop β1 – β2

Figure 2

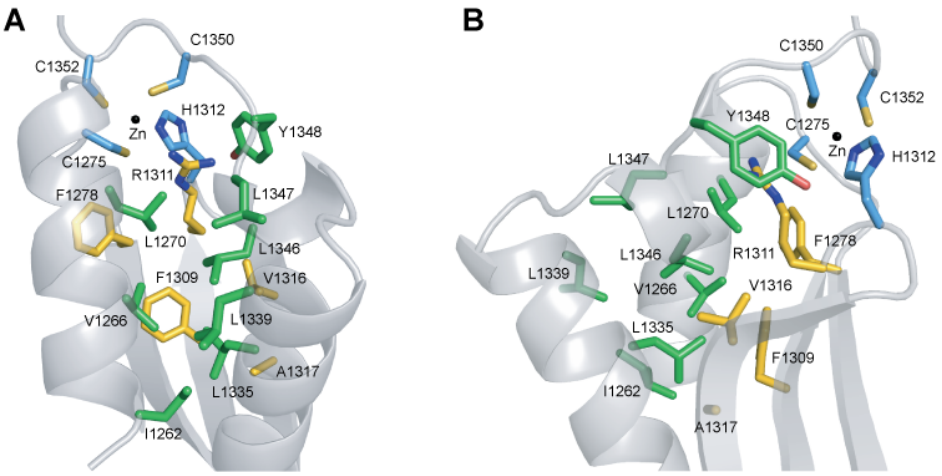


Figure 3

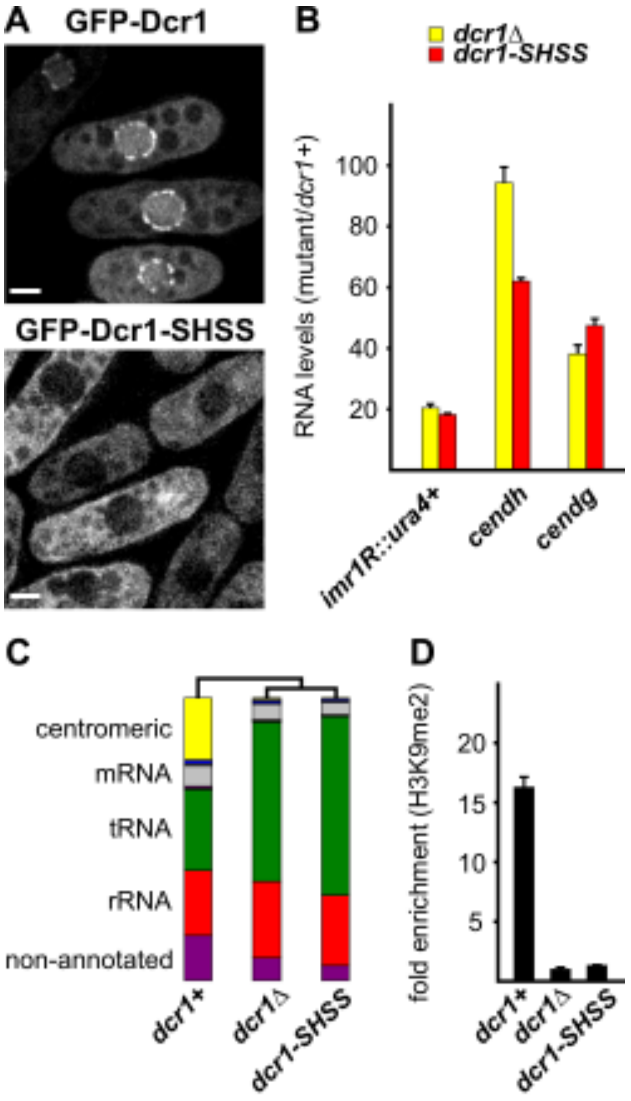


Figure 4

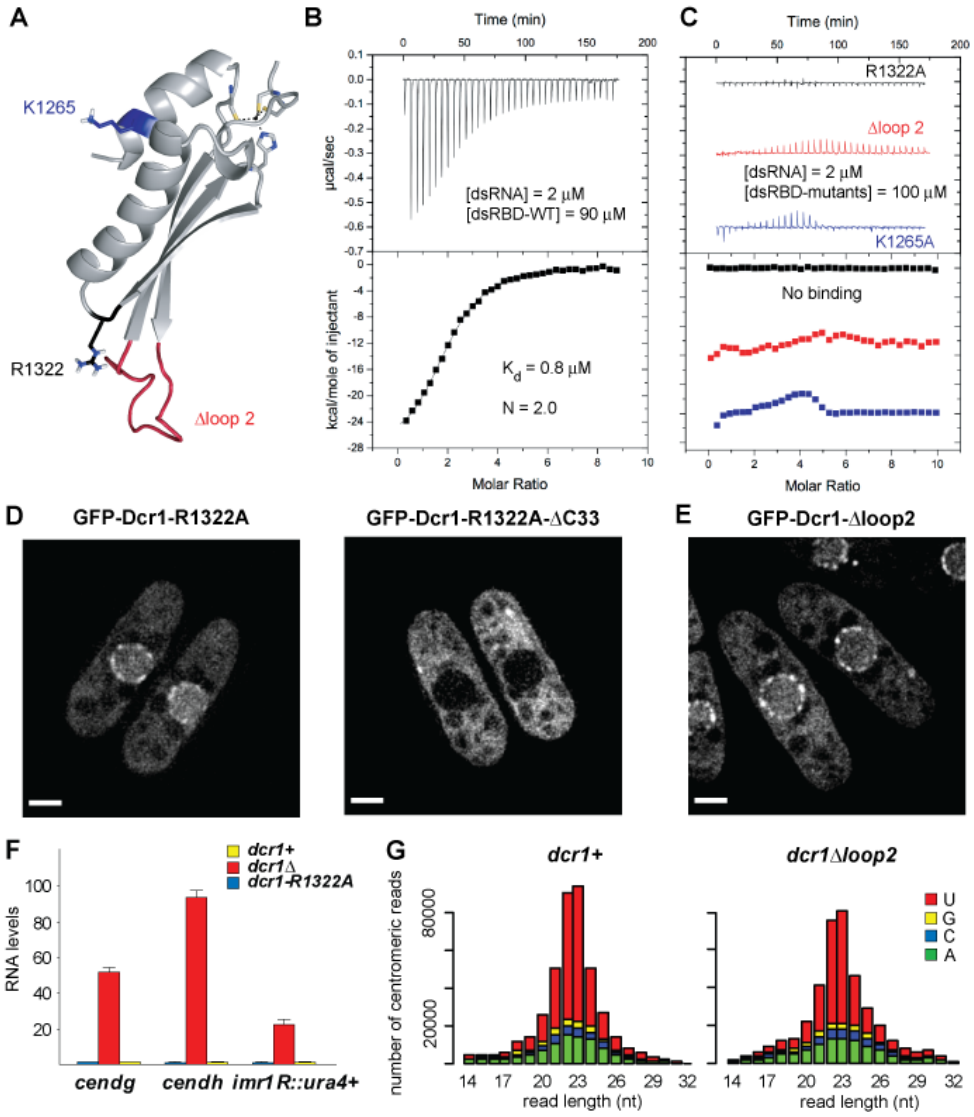

A
K1265
R1322
Δloop 2
B
Time (min)
µcal/sec
[dsRNA] = 2 µM
[dsRBD-WT] = 90 µM
kcal/mole of injectant
K_d = 0.8 µM
N = 2.0
Molar Ratio
C
Time (min)
R1322A
Δloop 2
[dsRNA] = 2 µM
[dsRBD-mutants] = 100 µM
K1265A
No binding
Molar Ratio
D
GFP-Dcr1-R1322A
GFP-Dcr1-R1322A-ΔC33
E
GFP-Dcr1-Δloop2
F
dcr1+
dcr1Δ
dcr1-R1322A
RNA levels
cendg
cendh
imr1R::ura4+
G
dcr1+
dcr1Δloop2
number of centromeric reads
U
G
C
A
read length (nt)
read length (nt)

Figure 5

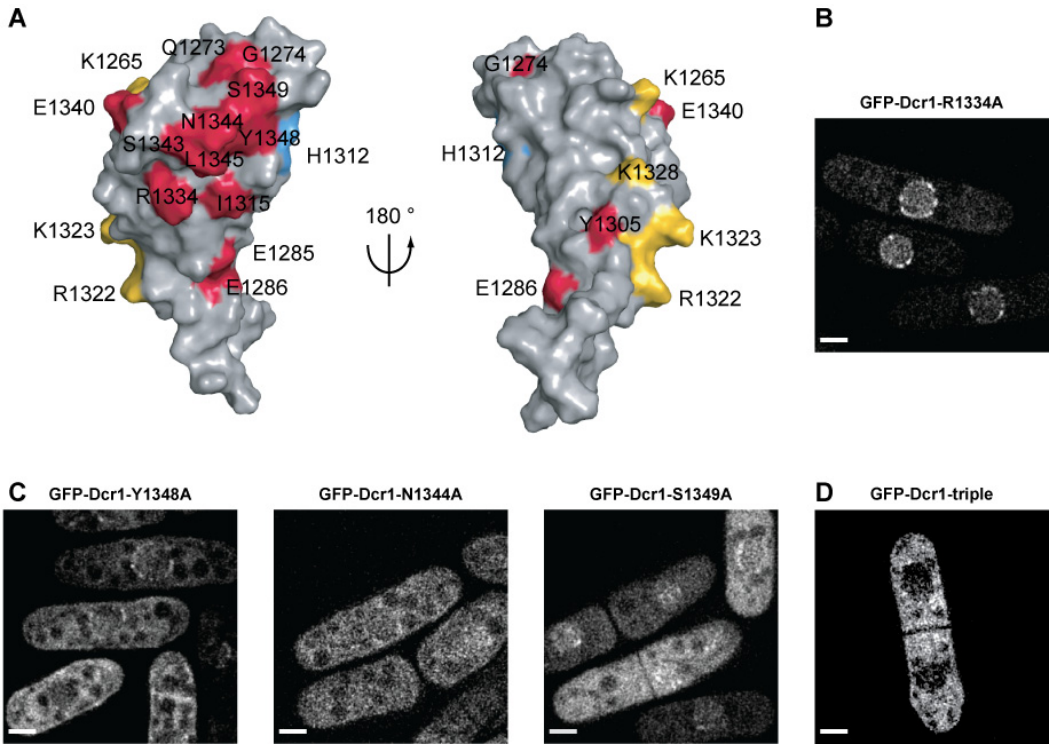


Figure 6

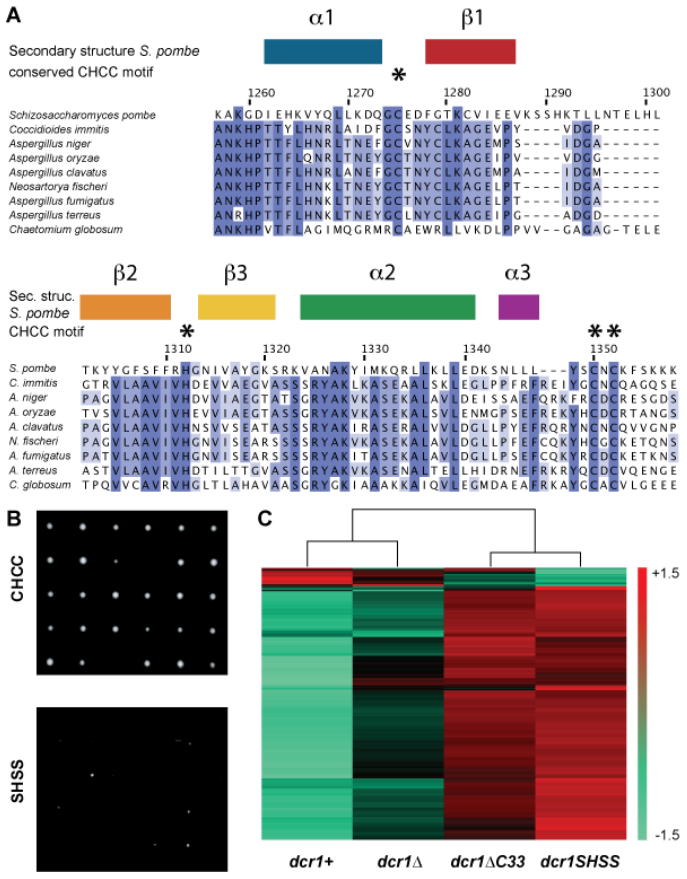
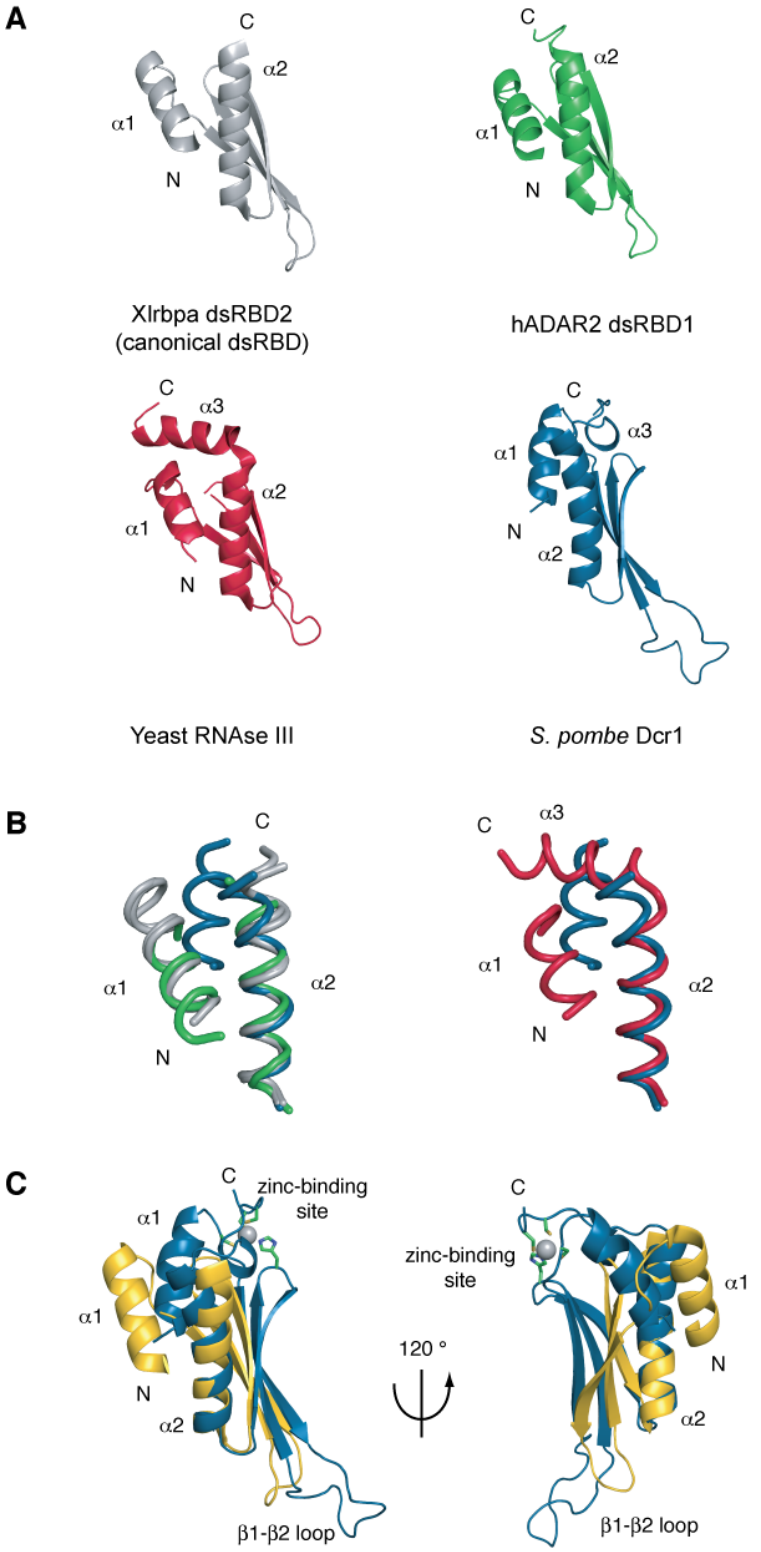


Figure 7

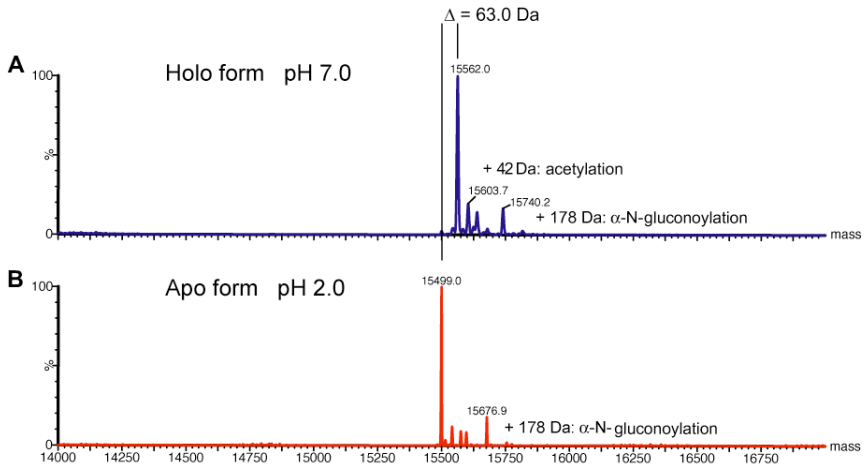

A
C
α2
α1
N
Xlrbpa dsRBD2
(canonical dsRBD)
C
α2
α1
N
hADAR2 dsRBD1
C
α3
α2
α1
N
Yeast RNAse III
C
α3
α1
N
α2
S. pombe Dcr1
B
C
α1
α2
N
C
α3
α1
α2
N
C
zinc-binding
site
α1
α1
N
α2
β1-β2 loop
120 °
C
zinc-binding
site
α1
N
α2
β1-β2 loop

Supplementary Data

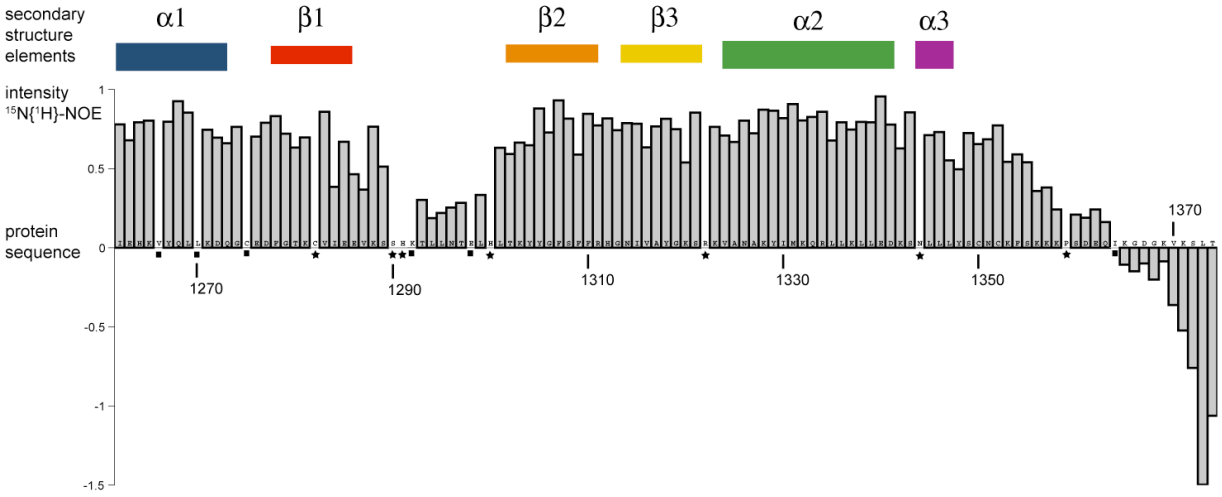
An Extended dsRBD with a Novel Zinc-Binding Motif Mediates Nuclear Retention of Fission Yeast Dicer

Pierre Barraud, Stephan Emmerth, Yukiko Shimada, Hans-Rudolf Hotz,
Frédéric H.-T. Allain, and Marc Bühler

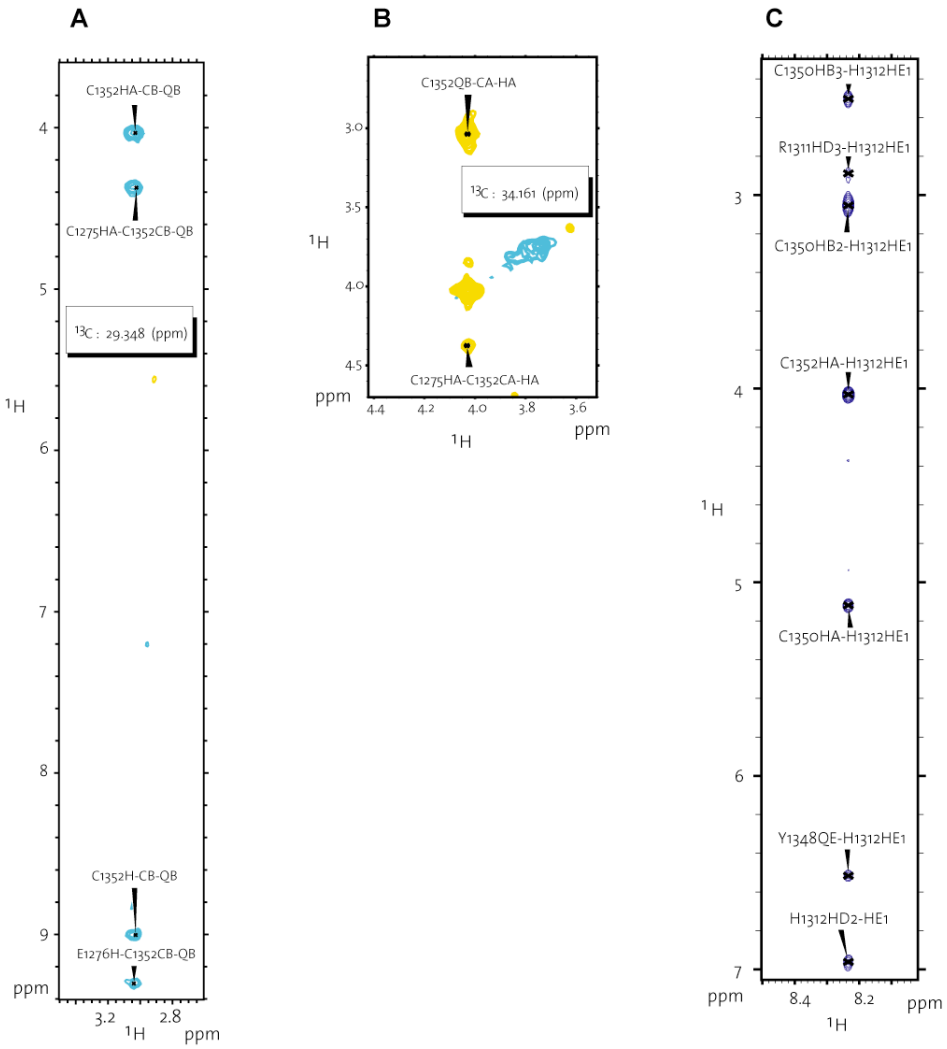
Supplementary Figures



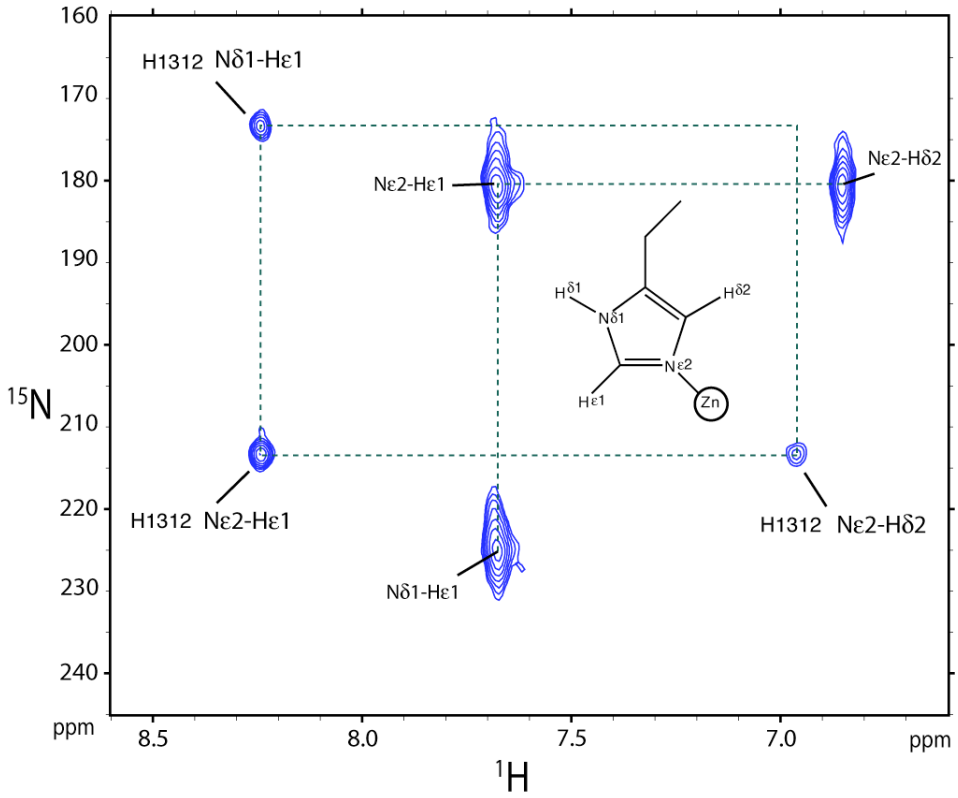
Supplementary Figure 1.
The C-terminal domain of Dcr1 binds one zinc ion per protein domain.
(A) ESI-TOF deconvoluted spectrum obtained under non-denaturing conditions (10 mM ammonium acetate pH 7.0). (B) ESI-TOF deconvoluted spectrum obtained under denaturing conditions (50 % acetonitrile and 0.1 % formic acid pH 2.0). Minor peaks come from acetylation, α-N-gluconoylation of the histidine tag and phosphorylation.



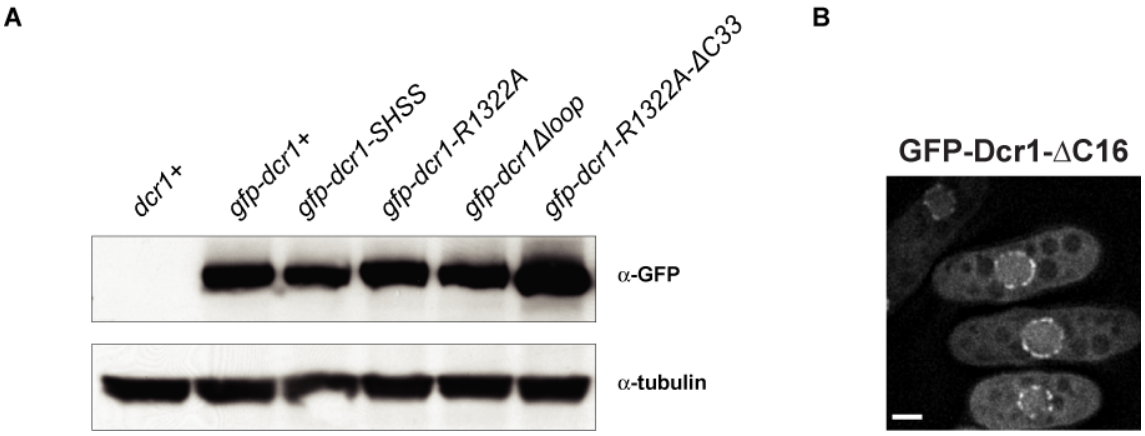
Supplementary Figure 2.
Residues of C33 following the canonical dsRBD fold are structured.
$^{15}N\{^{1}H\}$-NOE values plotted against the Dcr1 dsRBD sequence showing that the residues following the canonical dsRBD fold are structured. Values for residues marked with black squares could not be determined due to resonance overlap. Residues marked with an asterisk correspond to the one for which the amide assignment is missing. The secondary structure elements are represented as rectangles above the diagram. The color code refers to the one used in Figure 1.



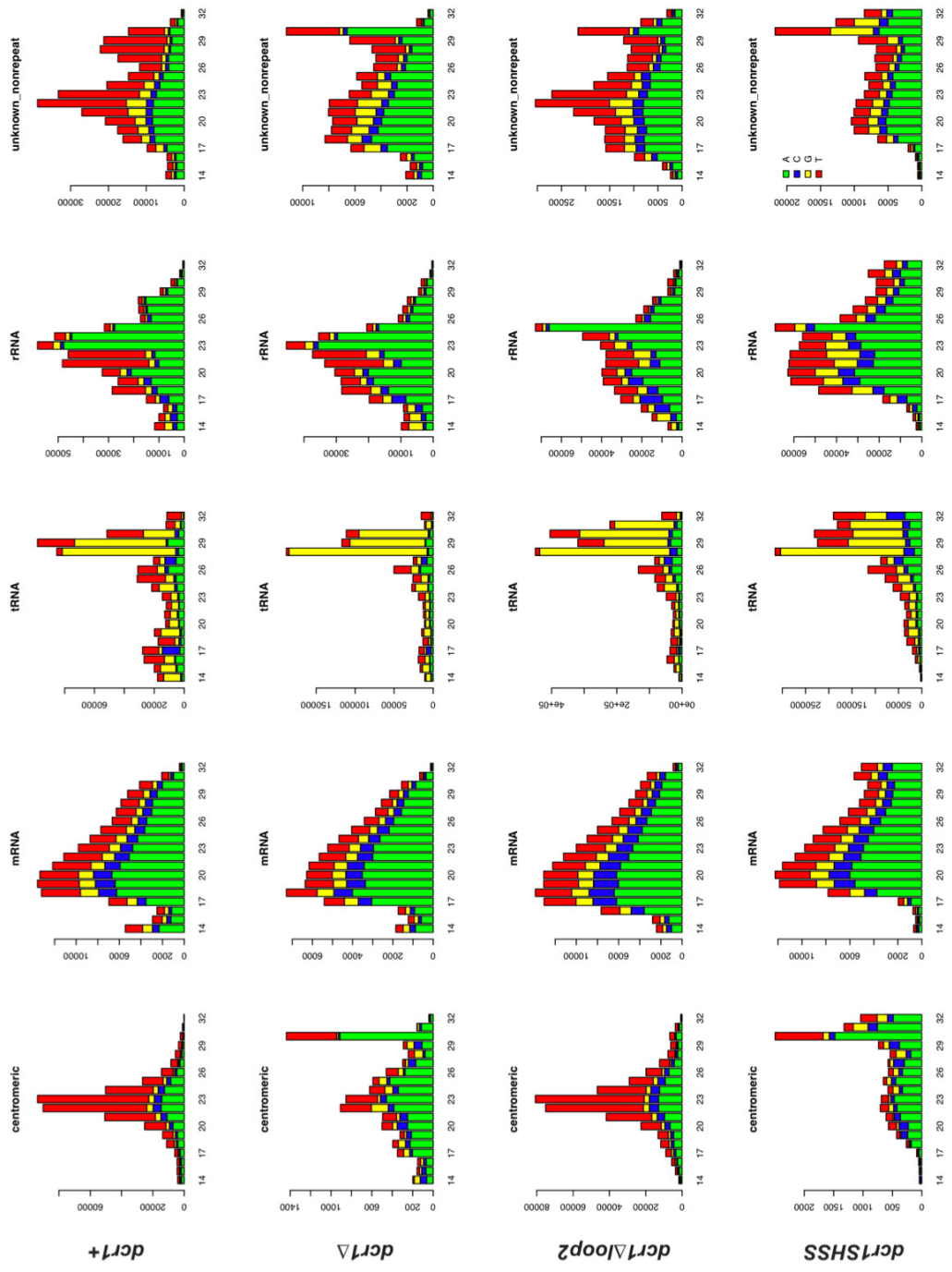
Supplementary Figure 3.
Coordination sphere around the zinc. (A-C) NOE contacts supporting the coordination sphere around the zinc ion. (A) Slice of the aliphatic 3D NOESY-(^{1}H-^{13}C)-HSQC showing contacts between C1352 and C1275. (B) Slice of the aliphatic 3D NOESY-(^{1}H-^{13}C)-HSQC showing contacts between C1352 and C1275. (C) Region of the 2D (^{1}H-^{1}H)-NOESY showing contacts between H1312 Hε1 and C1350 and C1352.



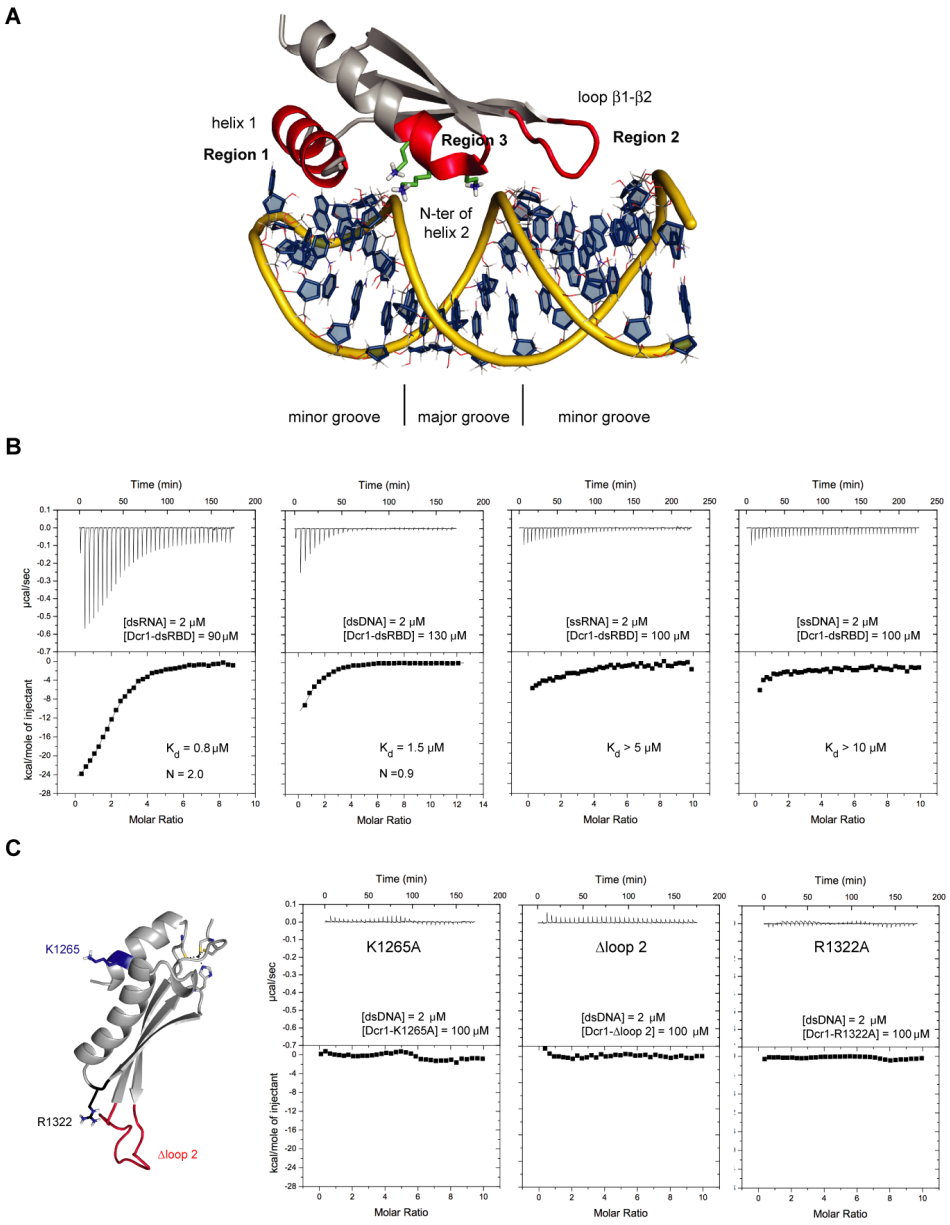
Supplementary Figure 4.
Long-range ^{1}H-^{15}N HSQC spectrum for the determination of the coordination mode of histidine 1312.
The pattern of cross peaks for H1312 is uniquely compatible with protonation on Nδ1 and therefore coordination through the Nε2 atom (c.f. Legge et al. (2004) J. Mol. Biol. 343, 1081-1093). The other peaks correspond to histidines from the N-terminal 6-histidine tag and to histidines present in the unstructured loop 2 (See Figure 1B). The pattern is characteristic of the histidine tautomer protonated on the Nε2.



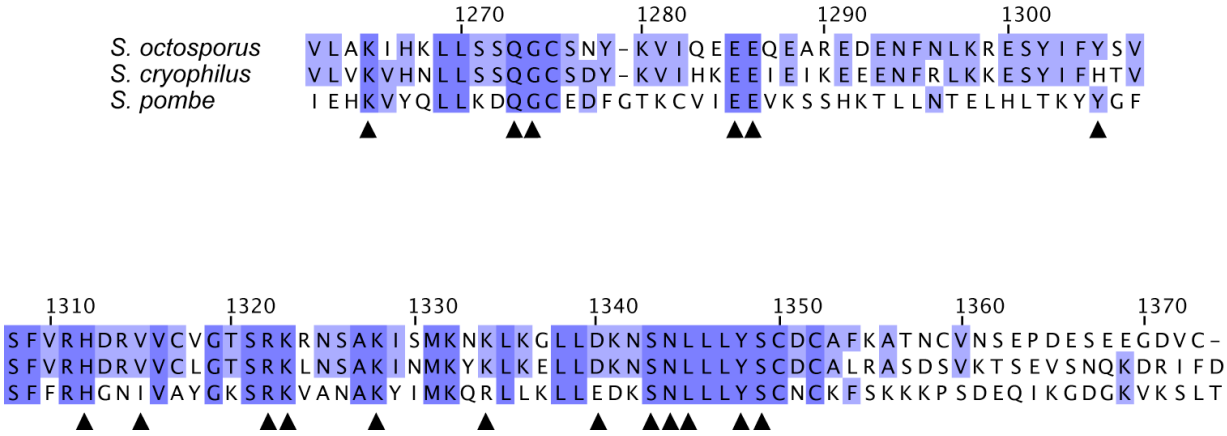
Supplementary Figure 5.
Western blot analysis and imaging of GFP-Dcr1-ΔC16. (A) Western blot showing that stability is not affected for the zinc coordination motif mutant or the RNA-binding mutants. (B) Live-cell imaging of GFP-Dcr1-ΔC16. Scale bars = 2 μm.



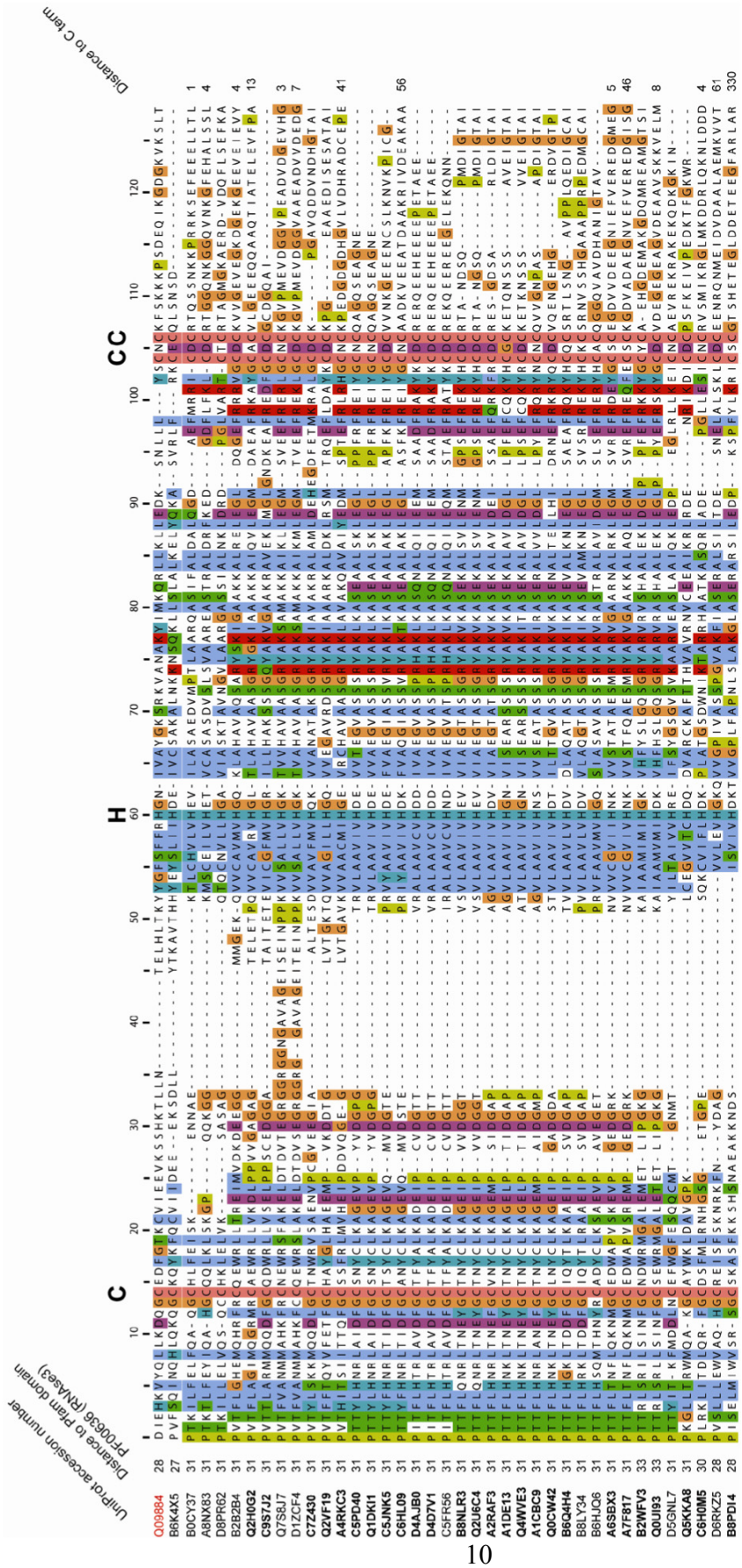
Supplementary Figure 6.
Size distribution and the 5'-most nucleotide of major classes of sRNAs sequenced from *dcr1+*, *dcr1Δ*, *dcr1Δloop2*, and *dcr1-SHSS* cells.



Supplementary Figure 7.
The Dcr1 dsRBD binds preferentially to dsRNA. (A) Regions of interaction between a classical dsRBD and a dsRNA target shown on the structure of ADAR2 dsRBD1 in complex with RNA (PDB code 2L3C). The three regions of interactions are represented in red on the protein structure (helix 1 and loop β1-β2 interact with the minor groove at one turn of interval and the N-terminal part of helix 2 interacts across the major groove). (B) ITC measurements of *S. pombe* Dcr1 dsRBD with different nucleic acid targets. From left to right: dsRNA (24 bp), dsDNA (24 bp), ssRNA (24 nt) and ssDNA (24 nt). (C) ITC measurements with dsDNA and Dcr1 dsRBD mutants (R1322A, Δloop 2 and K1265A).



Supplementary Figure 8.
Identification of the solvent exposed and conserved residues of Dcr1 dsRBD.
Black arrows show residues with a solvent exposed side chain and conserved biochemical properties among *S. pombe*, *S. octosporus* and *S. cryophilus*. Those residues are displayed in color on Figure 7A. Residue numbers refer to the *S. pombe* Dcr1 protein (UniprotKB Q09884).



Supplementary Figure 9.
Conservation of CHCC motif in C-termini of RNase III enzymes. Alignment of all annotated RNase III proteins with C termini that contain the CHCC motif. UniProt accession for the *S. pombe* Dcr1 protein is highlighted in red. Accessions in bold are proteins expressed in human or plant pathogenic yeast as listed in Supplementary Table II.

Supplementary Table I

Small RNAs sequenced from wild-type and Dcr1 mutant cells

	dcr1+		**dcr1Δ**		***dcr1-SHSS***		***dcr1Δloop2***	
	reads	%	reads	%	reads	%	reads	%
rRNA+	437259	23	345175	27	693100	25	499969	14
rRNA-	1122	0.1	158	0	739	0	451	0
tRNA+	544340	29	719518	56	1752112	63	2155557	61
tRNA-	460	0	244	0	365	0	226	0
snRNA+	5728	0.3	7382	0.6	9529	0.3	5786	0.2
snRNA-	26	0	4	0	12	0	22	0
snoRNA+	2074	0.1	1165	0.1	2938	0.1	2136	0.1
snoRNA-	201	0	112	0	338	0	326	0
3UTR+	14172	0.7	6866	0.5	10732	0.4	11557	0.3
3UTR-	1312	0.1	533	0	439	0	1042	0
5UTR+	2835	0.1	1301	0.1	2269	0.1	2627	0.1
5UTR-	163	0	56	0	138	0	194	0
mRNA+	126176	6.6	64120	5	117254	4.2	135565	3.8
mRNA-	12114	0.6	5243	0.4	6013	0.2	10885	0.3
intron+	10031	0.5	7536	0.6	10126	0.4	15736	0.4
intron-	465	0	324	0	1006	0	432	0
centromeric+	401625	21	7123	0.6	9712	0.3	382979	11
centromeric-	15473	0.8	1357	0.1	2286	0.1	15895	0.4
telomeric+	1	0	0	0	0	0	2	0
telomeric-	3	0	1	0	1	0	0	0
mating_type_region+	798	0	8	0	2	0	756	0
mating_type_region-	1315	0.1	17	0	7	0	1292	0
rep_origin+	56	0	21	0	20	0	59	0
rep_origin-	5	0	1	0	2	0	12	0
pseudogene+	316	0	35	0	89	0	315	0
pseudogene-	261	0	14	0	21	0	204	0
LTR+	138	0	75	0	81	0	104	0
LTR-	55	0	45	0	30	0	72	0
wtf+	450	0	298	0	435	0	539	0
wtf-	6	0	4	0	5	0	5	0
misc_RNA+	18258	1	11769	0.9	19463	0.7	23950	0.7
misc_RNA-	5558	0.3	861	0.1	1255	0	3151	0.1
unknown_non-repeat	303920	16	103173	8	148691	5.3	274583	7.7
sum	1906716	100	1284539	100	2789210	100	3546429	100

Weighted number of small RNA reads for different genomic features from deep sequencing of total sRNAs in wild-type, *dcr1Δ, dcr1-SHSS, and dcr1Δloop2* mutant cells.

Supplementary Table II

Pathogenic yeasts expressing at least one dicer protein containing the CHCC motif

Accession	Organism	Pathology
Q5KKA8	**Cryptococcus neoformans**	Cryptococcosis (opportunitstic infection in HIV+ patients, fungal meningitis)
Q2H0G2	Chaetomium globosum	Peritonitis, Onychomycosis
C5PD40	Coccidioides posadasii	Coccidioidomycosis (California Disease, Valley Fever)
Q1DKI1	Coccidioides immitis	Coccidioidomycosis (California Disease, Valley Fever)
C5JNK5	**Ajellomyces dermatitidis**	Blastomycosis (Chicago disease, Gilchrist's disease)
C6HL09	Ajellomyces capsulata	Darling's Disease
D4AJB0	Arthroderma benhamiae	Dermatophytosis (Ringworm)
B8NLR3	Aspergillus flavus	Aspergillosis (lung, occasionally identified as the cause of corneal, otomycotic and nasoorbital infections)
Q2U6C4	Aspergillus oryzae	Aspergillosis (only occasionlly pathogenic though)
A2RAF3	Aspergillus niger	Aspergillosis
Q4WVE3	**Aspergillus fumigatus**	Major cause of Aspergillosis (both invasive and allergic Aspergilloses), Mycotoxicosis, Farmer's Lung
A1DE13	Neosartorya fischeri	Aspergillosis
A1CBC9	Aspergillus clavatus	Aspergillosis, Mycotoxicosis
Q0CW42	**Aspergillus terreus**	Aspergillosis of the lungs and or disseminated aspergillosis, Onychomycosis, Mycotoxicosis (thermotolerant yeast)
B6Q4H4	Penicillium marneffei	Penicilliosis (third most common opportunistic infection in HIV+ patients)
D4D7V1	Trichophyton verrucosum	Dermatophytosis (Ringworm)
C7Z430	Nectria haematococca	Plant diseases, opportunistic infections in animals
C9S7J2	Verticillium albo-atrum	Verticillium Wilt Disease
Q2VF19	Cryphonectria parasitica	Chestnut Blight
A4RKC3	**Magnaporthe grisea**	Blast Disease (rice)
A6SBX3	Botryotinia fuckeliana	Gray mold disease
A7F817	Sclerotinia sclerotiorum	White Mold
B2WFV3	Pyrenophora tritici-repentis	Tan spot
Q0UI93	Phaeosphaeria nodorum	Glume blotch (major pathogen of wheat)
C6H0M5	Mucor c. f. lusitanicus	plant diseases
B8PDI4	Postia placenta	destructive decay of wood in buildings and other structures

UniProt release 2010_10 - Oct 5, 2010. Gold, human or animal pathogens. Yellow, plant pathogens.

Strains used in this study

Strain	Genotype	Source
SPB80	*h+ leu1-32 ura4-D18 ori1 ade6-216 imr1R(Nco1)::ura4+*	*2*
SPB81	*h+ leu1-32 ura4-D18 ori1 ade6-216 imr1R(Nco1)::ura4+ dcr1Δ::TAP-Kan*	*2*
SPB287	*h+ leu1-32 ura4-D18 ori1 ade6-216 imr1R(Nco1)::ura4+ dcr1-ΔC33::Nat*	*1*
SPB278	*h+ leu1-32 ura4-D18 ade6-216 imr1R(NcoI)::ura4+ Kan-nmt1P3-GFP::dcr1+*	*1*
SBP777	*h+ leu1-32 ura4-D18 ade6-216 imr1R(NcoI)::ura4+ Kan-nmt1P3-GFP::dcr1-SHSS::Nat*	*1*
SPB776	*h+ leu1-32 ura4-D18 ori1 ade6-216 imr1R(Nco1)::ura4+ dcr1-SHSS::Nat*	*1*
SPB773	*h+ leu1-32 ura4-D18 ori1 ade6-216 imr1R(Nco1)::ura4+ Kan-nmt1P3-GFP::dcr1-ΔC16::Nat*	*1*
SBP771	*h+ leu1-32 ura4-D18 ade6-216 imr1R(NcoI)::ura4+ Kan-nmt1P3-GFP::dcr1-Δloop2::Nat*	*1*
SPB770	*h+ leu1-32 ura4-D18 ori1 ade6-216 imr1R(Nco1)::ura4+ dcr1-Δloop2::Nat*	*1*
SBP773	*h+ leu1-32 ura4-D18 ade6-216 imr1R(NcoI)::ura4+ Kan-nmt1P3-GFP::dcr1-ΔC16::Nat*	*1*
SPB772	*h+ leu1-32 ura4-D18 ori1 ade6-216 imr1R(Nco1)::ura4+ dcr1-ΔC16::Nat*	*1*
SPB919	*h+ leu1-32 ura4-D18 ade6-216 imr1R(NcoI)::ura4+ Kan-nmt1P3-GFP::dcr1-R1322A::Hph*	*1*
SPB918	*h+ leu1-32 ura4-D18 ori1 ade6-216 imr1R(Nco1)::ura4+ dcr-R1322A::Hph*	*1*
SPB919	*h+ leu1-32 ura4-D18 ade6-216 imr1R(NcoI)::ura4+ Kan-nmt1P3-GFP::dcr1-R1322A-ΔC33::Hph*	*1*
SPB949	*h+ leu1-32 ura4-D18 ori1 ade6-216 imr1R(Nco1)::ura4+ dcr-K1265A::Hph*	*1*
SPB775	*h+ leu1-32 ura4-D18 ori1 ade6-216 imr1R(Nco1)::ura4+ Kan-nmt1P3-GFP::dcr1-Y1348A::Nat*	*1*
SPB1079	*h+ leu1-32 ura4-D18 ori1 ade6-216 imr1R(Nco1)::ura4+ Kan-nmt1P3-GFP::dcr1-R1334A::Nat*	*1*
SPB1080	*h+ leu1-32 ura4-D18 ori1 ade6-216 imr1R(Nco1)::ura4+ Kan-nmt1P3-GFP::dcr1-S1349A::Nat*	*1*
SPB1082	*h+ leu1-32 ura4-D18 ori1 ade6-216 imr1R(Nco1)::ura4+ Kan-nmt1P3-GFP::dcr1-N1344A::Nat*	*1*
SPB1116	*h+ leu1-32 ura4-D18 ori1 ade6-216 imr1R(Nco1)::ura4+ Kan-nmt1P3-GFP::dcr1-Y1348A-N1344A-S1349A::Nat*	*1*

1 = Bühler lab strain collection, 2 = obtained from Danesh Moazed

Mass spectrometry data

Related to section 3.3.

Explanation to the mass spectrometry data

The purifications were performed as described in section 3.5.3..

Additional information (see also section 3.5.3.):

Purification A: sonication, low speed spin.

Purification B: sonication, low speed spin, 500mM NaCl.

Purification C: 1% digitonin, sonication, low speed spin.

Purification D: 1% digitonin, sonication, low speed spin, 500mM NaCl.

Purification E: 1% digitonin.

Purification F: no changes as compared to the standard TAP-protocol.

The sequenced proteins are sorted according to the combined number of unique peptides found in the six TAP-Platform purifications. Ribosomal proteins and proteins with a high number of peptides in the TAP-only samples (named TAP in the tables) are excluded. The names of the identified proteins represent the accession number as displayed by Scaffold. The numbers for the sequenced proteins are read as follows: # of unique peptides/coverage, where coverage 1 equals 100% coverage.

The raw data can be found on the G-drive in DATA sets, Proteomics data, Emmerth, TAP-tag, Dicer platform project.

Mass spectrometry data

Purification:	A		B			C			D			E			F		
Name	**Platform**	**TAP**	**Platform**	**DUF**	**TAP**	**Platform**	**DUF**	**TAP**	**Platform**	**DUF**	**TAP**	**Platform**	**DUF**	**TAP**	**Platform**	**DUF**	**TAP**
EF2	20/0.31	0/0	24/0.29	0/0	0/0	39/0.5	0/0	0/0	15/0.18	0/0	0/0	37/0.59	8/0.16	0/0	32/0.5	9/0.19	0/0
K6PF	22/0.34	2/0.037	5/0.089	1/0.019	0/0	26/0.35	9/0.15	0/0	6/0.099	4/0.07	1/0.019	33/0.47	30/0.43	0/0	38/0.54	48/0.62	3/0.04
HSP90	14/0.27	0/0	17/0.24	0/0	0/0	12/0.21	1/0.027	0/0	4/0.054	0/0	0/0	21/0.37	7/0.12	0/0	25/0.4	28/0.46	0/0
ENO11	21/0.62	4/0.16	6/0.22	0/0	2/0.073	15/0.48	9/0.31	4/0.14	3/0.087	3/0.087	2/0.052	20/0.59	15/0.51	2/0.068	22/0.66	20/0.57	6/0.23
HSP75	11/0.27	1/0.02	3/0.067	0/0	0/0	20/0.36	7/0.18	8/0.18	9/0.19	1/0.02	0/0	21/0.43	13/0.24	4/0.083	20/0.42	17/0.39	7/0.2
KPYK	17/0.43	1/0.033	1/0.026	0/0	0/0	11/0.35	4/0.15	1/0.033	4/0.1	1/0.018	0/0	30/0.78	20/0.64	1/0.026	21/0.63	30/0.78	5/0.14
METE	12/0.24	0/0	0/0	0/0	0/0	9/0.18	2/0.043	0/0	3/0.046	2/0.03	1/0.016	24/0.4	16/0.26	0/0	22/0.41	23/0.43	4/0.088
THI2	12/0.46	0/0	3/0.18	0/0	0/0	14/0.62	3/0.2	0/0	5/0.22	2/0.091	0/0	15/0.6	11/0.46	0/0	19/0.73	14/0.58	1/0.04
HSP60	7/0.2	0/0	12/0.27	0/0	0/0	15/0.39	5/0.15	0/0	0/0	0/0	0/0	11/0.26	9/0.24	0/0	18/0.44	20/0.46	1/0.031
PGK	13/0.42	2/0.068	1/0.046	0/0	1/0.022	11/0.35	11/0.38	1/0.027	3/0.075	0/0	1/0.029	17/0.48	16/0.44	2/0.072	18/0.52	28/0.74	4/0.12
MAS5	10/0.29	0/0	0/0	0/0	0/0	19/0.52	8/0.35	0/0	0/0	0/0	0/0	17/0.48	13/0.41	0/0	15/0.42	18/0.48	0/0
DCR1	**8/0.057**	**0/0**	**14/0.06**	**6/0.032**	**0/0**	**16/0.086**	**16/0.055**	**0/0**	**3/0.023**	**7/0.031**	**0/0**	**11/0.081**	**10/0.058**	**0/0**	**8/0.055**	**10/0.058**	**0/0**
LEU3	8/0.33	0/0	1/0.035	0/0	0/0	14/0.52	2/0.081	0/0	0/0	0/0	0/0	15/0.56	7/0.3	0/0	18/0.7	9/0.34	0/0
PMA1	18/0.25	0/0	1/0.012	0/0	0/0	2/0.025	1/0.011	1/0.012	2/0.023	1/0.012	0/0	7/0.11	7/0.12	0/0	23/0.33	25/0.35	2/0.023
YF19	8/0.26	0/0	1/0.026	0/0	0/0	8/0.24	0/0	0/0	2/0.05	0/0	0/0	16/0.49	8/0.22	0/0	17/0.51	7/0.2	0/0
ACT	11/0.42	1/0.043	8/0.26	2/0.048	0/0	4/0.18	0/0	0/0	3/0.085	2/0.056	2/0.048	5/0.26	9/0.38	0/0	19/0.66	21/0.7	1/0.043
METK	9/0.32	0/0	0/0	0/0	0/0	10/0.32	0/0	0/0	3/0.097	1/0.039	1/0.039	14/0.56	7/0.29	0/0	14/0.54	11/0.45	1/0.039
PDC2	13/0.29	0/0	0/0	0/0	0/0	9/0.18	1/0.021	0/0	4/0.065	1/0.016	0/0	7/0.15	10/0.24	0/0	10/0.26	17/0.42	3/0.06
ADH	10/0.41	0/0	1/0.043	0/0	0/0	6/0.32	4/0.21	0/0	1/0.051	0/0	0/0	7/0.29	10/0.42	0/0	15/0.61	14/0.55	2/0.066
RUVB2	8/0.23	0/0	0/0	0/0	0/0	4/0.11	0/0	0/0	1/0.024	0/0	0/0	11/0.3	9/0.24	0/0	16/0.37	12/0.3	0/0
C1TC	2/0.045	0/0	0/0	0/0	0/0	3/0.039	1/0.017	0/0	0/0	0/0	0/0	15/0.19	1/0.015	0/0	12/0.19	15/0.26	0/0
EF3	3/0.037	0/0	0/0	0/0	0/0	4/0.094	0/0	0/0	0/0	0/0	0/0	12/0.18	8/0.14	0/0	13/0.19	16/0.23	0/0
UBIQ	3/0.41	0/0	11/0.95	4/0.47	0/0	6/0.79	1/0.3	0/0	3/0.47	2/0.33	0/0	4/0.53	6/0.64	0/0	5/0.62	4/0.64	0/0
RAD25	4/0.2	0/0	5/0.23	0/0	0/0	5/0.28	2/0.11	0/0	0/0	0/0	0/0	8/0.46	8/0.35	0/0	9/0.37	17/0.66	0/0
GLYD	2/0.056	0/0	0/0	0/0	0/0	1/0.024	1/0.058	1/0.024	1/0.024	0/0	0/0	17/0.35	8/0.19	1/0.039	9/0.19	15/0.38	2/0.062
ILVB	3/0.082	0/0	2/0.043	0/0	0/0	3/0.064	0/0	0/0	0/0	0/0	0/0	9/0.23	5/0.12	0/0	13/0.34	21/0.47	0/0
DED1	0/0	0/0	0/0	0/0	0/0	4/0.09	1/0.038	0/0	0/0	0/0	0/0	15/0.33	6/0.19	3/0.069	9/0.26	6/0.15	5/0.086
PRS6B	2/0.072	0/0	0/0	0/0	0/0	3/0.12	0/0	0/0	0/0	0/0	0/0	10/0.38	8/0.26	0/0	13/0.46	10/0.38	0/0
FAS2	4/0.03	8/0.08	0/0	0/0	0/0	2/0.019	0/0	0/0	0/0	0/0	0/0	5/0.033	3/0.024	0/0	15/0.12	10/0.083	14/0.11
ECM17	2/0.016	0/0	1/0.0081	0/0	0/0	3/0.029	0/0	0/0	1/0.0081	0/0	0/0	8/0.088	5/0.052	0/0	10/0.11	14/0.15	1/0.0081
SYTC	4/0.074	0/0	0/0	0/0	0/0	1/0.016	0/0	0/0	0/0	0/0	0/0	10/0.16	2/0.043	0/0	8/0.15	8/0.14	0/0
RUVB1	3/0.07	0/0	0/0	0/0	0/0	2/0.072	0/0	0/0	0/0	0/0	0/0	8/0.31	6/0.18	0/0	9/0.28	8/0.26	0/0
THI4	1/0.031	0/0	0/0	0/0	0/0	3/0.093	0/0	0/0	0/0	0/0	0/0	9/0.29	1/0.025	0/0	9/0.23	5/0.15	0/0
ALF	5/0.2	1/0.05	2/0.12	0/0	0/0	3/0.14	3/0.14	1/0.05	0/0	0/0	0/0	6/0.3	5/0.25	0/0	5/0.25	6/0.27	2/0.1
DYR	1/0.048	0/0	0/0	0/0	0/0	3/0.13	1/0.043	0/0	1/0.026	1/0.026	0/0	11/0.31	5/0.13	0/0	5/0.15	4/0.16	0/0
LYS9	2/0.056	0/0	0/0	0/0	0/0	5/0.16	1/0.031	0/0	0/0	0/0	0/0	8/0.3	5/0.2	0/0	6/0.24	5/0.18	0/0
CDC48	2/0.047	0/0	0/0	0/0	0/0	7/0.15	0/0	0/0	0/0	0/0	0/0	5/0.091	4/0.094	0/0	6/0.13	13/0.22	0/0
SAHH	2/0.053	0/0	0/0	0/0	0/0	2/0.088	0/0	0/0	0/0	0/0	0/0	9/0.2	6/0.16	0/0	7/0.16	6/0.17	1/0.032

Mass spectrometry data

	Platform	TAP	Platform	DUF	TAP	Platform	DUF	TAP	Platform	DUF	TAP	Platform	DUF	TAP	Platform	DUF	TAP
PDX1	5/0.3	0/0	1/0.054	4/0.17	1/0.03	1/0.051	3/0.11	0/0	0/0	4/0.16	0/0	4/0.26	10/0.44	0/0	8/0.36	14/0.59	0/0
THIL	3/0.12	0/0	0/0	0/0	0/0	1/0.038	1/0.038	0/0	0/0	0/0	0/0	5/0.2	2/0.073	0/0	10/0.44	10/0.47	0/0
THI3	1/0.064	0/0	0/0	0/0	0/0	2/0.12	0/0	0/0	0/0	0/0	0/0	11/0.43	7/0.24	0/0	4/0.19	13/0.44	0/0
HOSM	3/0.11	0/0	0/0	0/0	0/0	1/0.029	0/0	0/0	0/0	1/0.033	0/0	3/0.081	2/0.05	0/0	10/0.31	14/0.5	0/0
SERA	2/0.054	0/0	0/0	0/0	0/0	2/0.047	0/0	0/0	1/0.024	0/0	0/0	6/0.21	3/0.13	0/0	6/0.21	6/0.21	0/0
TCPG	0/0	0/0	0/0	0/0	0/0	1/0.021	0/0	0/0	0/0	0/0	0/0	9/0.21	3/0.061	0/0	7/0.16	12/0.26	0/0
DHAS	2/0.067	1/0.036	0/0	0/0	0/0	2/0.07	0/0	1/0.036	0/0	0/0	0/0	5/0.2	2/0.064	1/0.036	7/0.24	3/0.14	0/0
PRS8	0/0	0/0	0/0	0/0	0/0	1/0.022	0/0	0/0	0/0	0/0	0/0	5/0.17	8/0.28	0/0	10/0.34	12/0.38	0/0
SLT1	4/0.13	0/0	1/0.023	0/0	0/0	1/0.023	0/0	0/0	0/0	0/0	0/0	5/0.14	5/0.18	0/0	4/0.15	1/0.023	0/0
ATPA	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	4/0.086	4/0.086	0/0	10/0.21	6/0.17	0/0
MPG1	1/0.05	0/0	0/0	0/0	0/0	1/0.05	0/0	0/0	0/0	0/0	0/0	8/0.44	2/0.083	0/0	3/0.12	8/0.44	0/0
PYRD	3/0.12	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	1/0.032	0/0	0/0	9/0.3	3/0.099	0/0
SYMC	1/0.041	0/0	0/0	0/0	0/0	1/0.041	0/0	0/0	1/0.01	0/0	0/0	7/0.13	2/0.037	0/0	3/0.064	3/0.078	0/0
TOM70	3/0.085	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	1/0.03	3/0.094	0/0	9/0.27	11/0.32	0/0
RAD24	1/0.052	0/0	1/0.052	1/0.052	0/0	3/0.2	1/0.052	0/0	0/0	0/0	0/0	5/0.35	5/0.21	0/0	2/0.13	6/0.37	0/0
GPD1	1/0.036	0/0	0/0	0/0	0/0	2/0.12	1/0.042	0/0	0/0	0/0	0/0	6/0.27	4/0.17	0/0	2/0.091	12/0.48	0/0
PRS6A	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	3/0.11	6/0.21	0/0	8/0.22	10/0.27	0/0
GLT1	0/0	0/0	0/0	0/0	0/0	1/0.0033	0/0	0/0	0/0	0/0	0/0	2/0.019	1/0.0076	0/0	7/0.059	1/0.0076	0/0
PYR1	4/0.029	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	1/0.004	1/0.0053	0/0	5/0.036	7/0.051	0/0
CYSD	1/0.061	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	3/0.15	1/0.058	0/0	5/0.24	5/0.27	0/0
PRS4	1/0.027	0/0	0/0	0/0	0/0	1/0.027	1/0.027	0/0	0/0	0/0	0/0	1/0.027	3/0.11	0/0	6/0.18	7/0.25	0/0
YGI4	1/0.026	0/0	1/0.034	0/0	0/0	1/0.034	1/0.034	0/0	0/0	0/0	0/0	3/0.07	2/0.059	0/0	3/0.07	3/0.07	0/0
6PGD	2/0.063	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	4/0.12	3/0.083	0/0	2/0.063	8/0.19	1/0.026
ATPB	2/0.059	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	1/0.025	0/0	6/0.21	5/0.15	0/0
PMGY	1/0.066	0/0	0/0	0/0	0/0	1/0.066	0/0	0/0	0/0	0/0	0/0	4/0.2	0/0	0/0	2/0.066	5/0.33	0/0
PUR9	1/0.024	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	5/0.1	1/0.019	0/0	2/0.043	3/0.056	0/0
PYC	1/0.0093	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	2/0.018	1/0.0084	0/0	5/0.057	1/0.018	0/0
TCPZ	1/0.041	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	4/0.13	2/0.064	0/0	3/0.095	4/0.13	0/0
TPIS	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	6/0.45	2/0.1	0/0	2/0.12	4/0.22	0/0
ACL2	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	5/0.14	0/0	0/0	2/0.057	3/0.11	0/0
ARO1	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	4/0.038	1/0.007	0/0	3/0.025	3/0.029	0/0
H4	1/0.078	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	4/0.41	4/0.39	0/0	2/0.17	1/0.097	0/0
PRS7	1/0.034	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	2/0.055	5/0.16	0/0	4/0.12	8/0.26	0/0
RIR1	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	5/0.094	0/0	0/0	2/0.037	0/0	0/0
TCPQ	1/0.029	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	3/0.075	1/0.029	0/0	3/0.1	5/0.15	0/0
ADH4	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	3/0.097	0/0	0/0	3/0.097	6/0.19	0/0
ARG56	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	1/0.014	2/0.028	0/0	5/0.081	2/0.029	0/0
CARA	1/0.029	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	1/0.029	0/0	0/0	4/0.12	4/0.13	0/0
COPG	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	4/0.074	0/0	0/0	2/0.024	1/0.018	0/0

Mass spectrometry data

	Platform	TAP	Platform	DUF	TAP	Platform	DUF	TAP	Platform	DUF	TAP	Platform	DUF	TAP	Platform	DUF	TAP
G3P2	0/0	0/0	1/0.036	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	3/0.14	1/0.036	0/0	2/0.099	4/0.21	0/0
GCST	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	3/0.14	2/0.11	0/0	3/0.14	1/0.041	0/0
IF2A	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	4/0.18	0/0	0/0	2/0.1	1/0.049	0/0
MAOX	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	5/0.1	5/0.12	0/0	1/0.025	7/0.14	0/0
TCPB	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	3/0.1	0/0	0/0	3/0.1	1/0.025	0/0
TCPD	0/0	0/0	0/0	0/0	0/0	1/0.03	0/0	0/0	0/0	0/0	0/0	1/0.03	0/0	0/0	4/0.18	3/0.11	0/0
VDAC	0/0	0/0	0/0	0/0	0/0	3/0.24	2/0.16	0/0	0/0	0/0	0/0	3/0.11	2/0.11	0/0	0/0	0/0	0/0
YFM8	1/0.043	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	2/0.077	3/0.18	0/0	3/0.18	2/0.15	0/0
YLF4	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	3/0.1	1/0.022	0/0	3/0.12	1/0.022	0/0
ADT	1/0.04	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	1/0.034	0/0	4/0.13	6/0.22	0/0
DHE4	1/0.035	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	1/0.035	0/0	0/0	3/0.098	1/0.035	0/0
DNJC7	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	4/0.12	0/0	0/0	1/0.036	0/0	0/0
EF1A2	1/0.074	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	2/0.074	1/0.074	0/0	2/0.074	2/0.074	0/0
EIF3I	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	2/0.076	1/0.034	0/0	3/0.16	1/0.034	1/0.034
FBRL	1/0.036	0/0	0/0	0/0	0/0	1/0.036	5/0.21	0/0	0/0	0/0	0/0	1/0.036	0/0	0/0	2/0.098	1/0.036	1/0.036
ILV5	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	3/0.11	1/0.062	1/0.027	2/0.092	4/0.17	0/0
NACA	1/0.075	0/0	0/0	0/0	0/0	1/0.075	0/0	0/0	0/0	0/0	0/0	2/0.14	1/0.064	0/0	1/0.075	2/0.14	0/0
OAT	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	3/0.059	0/0	0/0	2/0.055	1/0.025	0/0
PUT2	1/0.042	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	1/0.029	0/0	1/0.042	3/0.095	0/0	3/0.095	8/0.2	0/0
SYDC	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	5/0.17	8/0.17	0/0
TBA1	1/0.033	0/0	0/0	0/0	0/0	1/0.033	0/0	0/0	0/0	0/0	0/0	2/0.066	2/0.055	0/0	1/0.031	4/0.15	0/0
YAY1	2/0.082	2/0.069	0/0	0/0	0/0	0/0	1/0.02	0/0	0/0	0/0	0/0	1/0.048	2/0.084	0/0	2/0.071	8/0.23	3/0.092
ACO1	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	4/0.16	5/0.22	0/0
COPD	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	2/0.11	0/0	0/0	2/0.11	0/0	0/0
CUE4	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	2/0.088	0/0	0/0	2/0.088	1/0.079	0/0
EIF3G	1/0.039	0/0	0/0	0/0	0/0	1/0.05	0/0	0/0	0/0	0/0	0/0	1/0.039	1/0.067	0/0	1/0.039	2/0.11	2/0.11
H2B1	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	3/0.23	1/0.095	0/0	1/0.095	1/0.095	0/0
HSP7F	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	1/0.017	0/0	0/0	2/0.046	0/0	0/0	1/0.018	2/0.046	0/0
IPYR	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	3/0.14	0/0	0/0	1/0.038	1/0.038	0/0
MET3	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	2/0.033	0/0	0/0	2/0.057	0/0	0/0
MMF1	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	2/0.19	1/0.068	0/0	2/0.2	4/0.31	0/0
PUR2	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	1/0.023	1/0.02	0/0	3/0.049	0/0	0/0
TBB	0/0	0/0	0/0	0/0	0/0	1/0.04	0/0	0/0	1/0.033	0/0	0/0	1/0.04	3/0.1	0/0	1/0.033	2/0.074	0/0
VATA	1/0.024	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	1/0.032	0/0	3/0.11	2/0.057	0/0
YHY6	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	1/0.034	0/0	0/0	3/0.088	3/0.11	0/0
AK	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	1/0.035	0/0	0/0	2/0.056	1/0.035	0/0
CYPH	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	2/0.08	0/0	0/0	1/0.08	0/0	0/0
DLDH	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	1/0.027	2/0.047	0/0	2/0.049	3/0.09	0/0
G6PD	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	1/0.024	0/0	0/0	2/0.056	0/0	0/0

Mass spectrometry data

	Platform	TAP	Platform	DUF	TAP	Platform	DUF	TAP	Platform	DUF	TAP	Platform	DUF	TAP	Platform	DUF	TAP
H2A1	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	3/0.3	1/0.068	0/0	0/0	0/0	0/0
IDHP	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	1/0.031	2/0.055	0/0	2/0.055	3/0.081	0/0
LYS12	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	1/0.044	0/0	0/0	2/0.088	3/0.13	0/0
PROA	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	3/0.12	2/0.055	0/0
RPB1	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	1/0.008	0/0	0/0	2/0.018	0/0	0/0
RS6A	1/0.071	0/0	0/0	0/0	0/0	0/0	2/0.084	0/0	0/0	0/0	0/0	1/0.071	1/0.071	0/0	1/0.071	1/0.071	0/0
SYFA	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	1/0.022	0/0	0/0	2/0.044	0/0	0/0
TAL1	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	1/0.056	1/0.056	0/0	2/0.16	1/0.056	0/0
TCPE	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	3/0.13	0/0	0/0	0/0	1/0.038	0/0
UREG	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	2/0.1	0/0	0/0	1/0.07	0/0	0/0
VTC4	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	1/0.026	0/0	0/0	2/0.037	0/0	0/0
YEAA	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	3/0.13	2/0.081	0/0
ASSY	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	1/0.032	0/0	2/0.12	1/0.032	0/0
BIOB	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	2/0.055	1/0.022	0/0
CYSK1	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	1/0.06	0/0	0/0	1/0.06	2/0.1	0/0
DAK1	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	1/0.021	0/0	0/0	1/0.021	2/0.067	0/0
EFTU	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	2/0.08	2/0.073	0/0
HIS1	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	2/0.14	0/0	0/0
HXK1	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	1/0.048	1/0.048	0/0	1/0.048	2/0.079	0/0
IDH1	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	1/0.042	0/0	0/0	1/0.042	3/0.16	0/0
IF4A	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	2/0.061	0/0	0/0	0/0	2/0.071	0/0
MDJ1	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	2/0.055	0/0	0/0	0/0	4/0.12	0/0
MLO3	1/0.085	0/0	0/0	0/0	0/0	1/0.085	4/0.36	0/0	0/0	0/0	0/0	0/0	2/0.17	0/0	0/0	3/0.17	0/0
NOG2	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	2/0.047	1/0.026	0/0
P25	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	2/0.14	0/0	0/0	0/0	1/0.054	0/0
PLB1	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	1/0.021	0/0	1/0.021	1/0.021	0/0	1/0.021	4/0.072	0/0
PRS10	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	1/0.028	3/0.1	0/0	1/0.036	3/0.11	0/0
PSB1	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	2/0.1	0/0	0/0	0/0	2/0.1	0/0
RNC1	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	1/0.04	0/0	0/0	1/0.04	2/0.075	0/0
RPIA	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	1/0.036	0/0	2/0.11	0/0	0/0
RPN1	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	1/0.013	0/0	2/0.038	1/0.025	0/0
SERC	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	2/0.054	0/0	0/0	0/0	0/0	0/0
SOU1	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	2/0.1	1/0.043	0/0
SYG	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	2/0.054	1/0.026	0/0
TCPA	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	2/0.058	0/0	0/0
THI22	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	1/0.031	1/0.031	0/0	1/0.053	4/0.14	0/0
TRP	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	1/0.029	0/0	2/0.053	6/0.12	0/0
YBB2	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	2/0.077	0/0	0/0	0/0	0/0	0/0
YC0C	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	2/0.04	0/0	0/0

	Platform	TAP	Platform	DUF	TAP	Platform	DUF	TAP	Platform	DUF	TAP	Platform	DUF	TAP	Platform	DUF	TAP
ACAC	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	1/0.0048	0/0	0/0	1/0.012	2/0.014	0/0
ACL1	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	1/0.02	0/0	0/0	0/0	0/0	0/0	0/0	2/0.05	0/0
GHT5	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	1/0.033	2/0.068	0/0
LEUC	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	1/0.022	0/0	0/0	0/0	0/0	0/0	0/0	2/0.045	0/0
RPN2	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	1/0.016	0/0	0/0	0/0	0/0	1/0.016	4/0.044	0/0
RPN3	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	1/0.036	2/0.054	0/0
SLBR2	0/0	0/0	0/0	0/0	0/0	0/0	0/0	1/0.028	0/0	0/0	0/0	0/0	0/0	0/0	0/0	3/0.15	0/0
TIM44	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	2/0.075	0/0	0/0	0/0	0/0	1/0.035	3/0.11	0/0
TOM20	0/0	0/0	1/0.092	3/0.5	0/0	0/0	3/0.23	0/0	0/0	5/0.48	0/0	0/0	0/0	0/0	0/0	0/0	0/0
YE0C	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	1/0.079	2/0.12	0/0
2AD1	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	2/0.04	0/0	0/0	0/0	0/0	0/0	1/0.024	0/0
DPOD	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	2/0.024	0/0
EF1G	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	1/0.022	0/0	0/0	0/0	0/0	0/0	2/0.12	0/0
EIF3A	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	2/0.021	0/0
ERG5	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	5/0.11	0/0
IF5	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	3/0.15	0/0
IMA3	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	2/0.027	0/0	0/0	0/0	0/0	0/0	2/0.027	0/0
LCF1	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	4/0.084	0/0
MU157	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	2/0.067	0/0
P5CR	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	2/0.082	0/0
PTR1	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	4/0.017	0/0
TOM22	0/0	0/0	0/0	5/0.41	0/0	0/0	2/0.14	0/0	0/0	4/0.36	0/0	0/0	0/0	0/0	0/0	0/0	0/0
TOM40	0/0	0/0	0/0	3/0.13	0/0	0/0	1/0.047	0/0	0/0	7/0.2	0/0	0/0	0/0	0/0	0/0	0/0	0/0
YAC4	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	3/0.074	0/0
YBY6	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	4/0.48	0/0
YPT1	0/0	0/0	0/0	0/0	0/0	0/0	1/0.059	0/0	0/0	2/0.12	0/0	0/0	0/0	0/0	0/0	0/0	0/0
YQ52	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	1/0.021	0/0	0/0	0/0	0/0	0/0	2/0.044	0/0