Charles University in Prague, Faculty of Science
Department of Cell Biology

Study program: Biology
Branch of study: Cellular and Developmental Biology

Bc. Jana Kubíková

Substrate cleavage by mammalian Dicer isoforms

Štěpení substrátů isoformami savčího Diceru

Master's thesis

Supervisor: Doc. Mgr. Petr Svoboda, Ph.D.
Prague, 2016

Prohlášení:

Prohlašuji, že jsem diplomovou práci zpracovala samostatně a že jsem uvedla všechny použité informační zdroje a literaturu. Tato práce ani její podstatná část nebyla předložena k získání jiného nebo stejného akademického titulu.

V Praze, 1.5.2016

Jana Kubíková

Na tomto místě bych ráda poděkovala mému školiteli, Doc. Mgr. Petru Svobodovi, Ph.D., za svěřený projekt, za cenné rady a připomínky, které mi poskytl během mého projektu a sepisování diplomové práce, a za velmi přátelský a lidský přístup.

Dále děkuji MUDr. Radkovi Malíkovi, Ph.D. za trpělivost, rady i nápady týkající se mého projektu, péči o hmyzí kultury v mé nepřítomnosti a za přípravu některých plasmidů.

Mgr. Janě Urbanové, Ph.D. děkuji za pomoc při přípravě expresních vektorů a za vytváření přátelské atmosféry v laboratoři.

Celému osazenstvu Laboratoře epigenetických regulací děkuji za udržování přátelského prostředí a tolerování některých mých (roztomilých) zlozvyků.

Dále děkuji Mgr. Václavu Urbanovi za pomoc při FPLC, za konzultace týkající se purifikace proteinů a nevratné zapůjčení některých purifikačních resinů a TEV proteázy.

Děkuji Mgr. Pavlíně Řezáčové, Ph.D. za možnost si v její laboratoři vyzkoušet gelovou chromatografii a RNDr. Magdaleně Hořejší, Csc., za její čas, ochotu, rady i pomoc při gelové chromatografii a FPLC.

Děkuji Mgr. Zoře Novákové, Ph.D. a Mgr. Petře Baranové za konzultace ohledně purifikace proteinů pomocí Twin-Strep-tagu.

Děkuji Mgr. Kamile Burdové, Ph.D. za poskytnutí Sf9 buněčné linie a za konzultace týkající bakulovirového expresního systému.

Na závěr bych chtěla poděkovat mé rodině za podporu a mému manželovi za péči, ohleduplnost, toleranci a schopnost šířit kolem sebe dobrou náladu.

Zvláštní poděkování patří mojí sestře Marušce, která mě dokázala v pravou chvíli povzbudit radou nebo příkladem a která mi společně s jezevčicí Mášou vytvořila příjemné pracovní prostředí na dopsání diplomové práce.

Abstrakt

Hostitelské organismy vyvinuly řadu antivirových odpovědí, které rozpoznávají a potlačují virovou infekci. Jedním ze znaků virové infekce je i přítomnost dvouvláknové RNA v buňce. Jedna z drah odpovídajících na dvouvláknovou RNA je RNA interference, která představuje klíčovou složku antivirové obrany u bezobratlých živočichů a rostlin. Savci využívají pro obranu před viry odlišnou dráhu, která rozpoznává dvouvláknovou RNA, nazývající se interferonová odpověd. RNA interference funguje jen v savčích oocytech a raných embryonálních stádiích, ačkoliv potřebné enzymové vybavení se nachází ve všech somatických buňkách, protože je využíváno mikroRNA dráhou. Předchozí studie ukázala, že funkčnost RNA interference v myších oocytech souvisí s přítomností oocytární isoformy Diceru ($Dicer^{O}$), jež je zkrácená na N-konci. V této práci jsem se zaměřila na zhodnocení, zda $Dicer^{O}$ zpracovává substráty typické pro RNA interferenci účinněji než nezkrácená isoforma Diceru ($Dicer^{S}$), která se nachází v somatických buňkách. Za tímto účelem jsem vyvinula protokol na purifikaci Diceru, abych získala obě myší isoformy Diceru o vysoké čistotě. Následně jsem otestovala jejich aktivitu v neradioaktivní enzymové analýze na různých substrátech se strukturními prvky charakteristickými pro substráty RNA interference. Získané výsledky naznačují, že rekombinantní $Dicer^{O}$ a $Dicer^{S}$ se neliší ve zpracování dlouhých komplementárních duplexů ani ve schopnosti štěpení uvnitř substrátu. Pro vysvětlení získaných poznatků je třeba dalších experimentů, protože používaná neradioaktivní enzymová analýza měla řadu technických omezení a izolované rekombinantní Dicery měly nízkou aktivitu, což může být důsledkem přítomnosti purifikačních značek na jejich N-konci.

Klíčová slova: Dicer, RNA interference, mikroRNA, helikáza, dvouvláknová RNA, purifikace proteinů, enzymová analýza

Abstract

Host organisms evolved antiviral responses, which can recognize the viral infection and deal with it. One of the frequent signs of viral infection in a cell is appearance of double-stranded RNA (dsRNA). One of the pathways responding to dsRNA is RNA interference (RNAi), which functions as the key antiviral defence system in invertebrates and plants. Mammals, however, utilize for antiviral defence a different dsRNA-sensing pathway called the interferon response. RNAi functions only in mammalian oocytes and early embryonal stages although its enzymatic machinery is present in all somatic cells, where it is employed in the microRNA pathway. A previous study indicated that the functionality of RNAi in mouse oocytes functions due to an oocyte-specific isoform of protein Dicer ($Dicer^{O}$), which is truncated at the N-terminus. In my thesis, I aimed to assess whether $Dicer^{O}$ processes RNAi substrates more efficiently *in vitro* than the full-length Dicer ($Dicer^{S}$), which is found in somatic cells. Therefore, I developed Dicer purification protocol for obtaining both recombinant mouse Dicer isoforms of high purity. I examined their activity in a non-radioactive cleavage assay using RNA substrates with structural features characteristic of RNAi substrates. My results suggest that recombinant $Dicer^{O}$ and $Dicer^{S}$ do not differ in processing of long perfect duplexes or in their ability of performing an internal cleavage. Nevertheless, further experiments are necessary to explain these results as the non-radioactive cleavage assays suffered from substantial technical limitations and the reduced activity of recombinant Dicers, which might be a consequence of the affinity tags at the N-terminus.

Key words: Dicer, RNAi, miRNA, helicase, dsRNA, protein purification, cleavage assay

Table of contents

List of Abbreviations

ALG	Argonaute-like
Ago/AGO	Argonaute
bp	base pair
DCR-1	Dicer-1 in *Caenorhabditis elegans*
Dcr-1	Dicer-1 in *Drosophila melanogaster*
Dcr-2	Dicer-2 in *Drosophila melanogaster*
DGCR8	DiGeorge syndrome critical region 8
$Dicer^{O}$	oocyte-specific Dicer isoform in *Mus musculus*
$Dicer^{S}$	full-length Dicer in *Mus musculus*
dsRBD	dsRNA binding domain
dsRBP	double-stranded RNA binding protein
dsRNA	double-stranded RNA
DUF283	domain of unknown function
EDTA	ethylenediaminetetraacetic acid
EGFP	Enhanced green fluorescent protein
FPLC	Fast protein liquid chromatography
GFP	green fluorescent protein
GST	Glutathione S-transferase
HRP	horseradish peroxidase
HRV 3C	human *rhinovirus* 3C
IMAC	immobilized metal affinity chromatography
IPTG	isopropyl-β-D-thiogalactopyranoside
LB	lysogeny-broth medium (also known as Luria broth medium)
Loqs	Loquacious
miRNA	microRNA
nt	nucleotide
PAGE	polyacrylamide gel electrophoresis
PAZ domain	Piwi Argonaute Zwille domain
PBS	phosphate-buffered saline
piRNA	piwiRNA
pre-miRNA	precursor miRNA
pri-miRNA	primary miRNA
PMSF	phenylmethylsulfonyl fluoride
PVDF	polyvinylidene difluoride
RdRp	RNA-dependent RNA polymerase

RDE-1	RNAi-Defective 1
RIG-I	retinoic acid-inducible gene 1
RISC	RNA-induced silencing complex
RNAi	RNA interference
RNase	ribonuclease
RI	Ribonuclease Inhibitor
sDCR-1	Small DCR-1 in *C. elegans*
siRNA	short interfering RNA
SDS	sodium dodecyl sulfate
SDS-PAGE	denaturing PAGE
SF	Superfamily
ssRNA	single-stranded RNA
TEMED	tetramethylethylenediamine
TEV protease	Tobacco Etch Virus nuclear-inclusion-a endopeptidase
TTBS	Tween-Tris-buffered saline
TRBP	HIV-1 TAR RNA binding protein
X-gal	5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside

1 Introduction

1.1 Small RNA pathways

Gene expression is regulated both transcriptionally and post-transcriptionally. In eukaryotes, a new level of gene regulation appeared in the form of post-transcriptional regulation mediated by small RNAs. In this regulation, 20-30nucleotide (nt) long RNAs associated with a member of the Argonaute protein family form the core of an effector complex and elicit sequence-specific repression of target mRNAs. There are several small RNA pathways in eukaryotes including RNA interference (RNAi), the microRNA (miRNA) pathway and the piwi RNA (piRNA) pathway. The miRNA pathway and RNAi will be discussed further in more detail because they were in the focus of my work.

Small RNA pathways play a fundamental role in regulation of the transcriptome, in genome integrity maintenance as well as in antiviral defence. As the latter two functions are under considerable selection due to continuous pathogen-host arms race, diversity evolved in these pathways. Consequently, these pathways differ considerably among organisms. Definitions of small RNA pathways are, therefore, somewhat arbitrary and vary according to authors. For the purpose of this thesis, I will use the term miRNA pathway for a pathway controlling expression of endogenous genes through small RNAs encoded in the genome and having a short hairpin intermediate. The term RNAi denotes a sequence-specific RNA degradation initiated by long double-stranded RNA (dsRNA) (Fire et al, 1998).

1.1.1 The miRNA pathway

The miRNA pathway (reviewed for example in Ameres & Zamore, 2013; Ha & Kim, 2014) regulates protein-coding genes at the post-transcriptional level. Majority of miRNAs and their targets are coded in the nuclear genome. There are exceptions to this rule represented by virus-encoded miRNAs and mRNAs. The exceptions can be attributed to the adaptation of the parasite to its host; the virus-encoded miRNAs contribute to viral strategies to increase efficiency of viral replication through regulating the host gene expression via its own pathways.

A canonical miRNA is genome-encoded as a dedicated gene or a miRNA precursor is hosted in a protein coding gene (Rodriguez et al, 2004) (Fig. 1). In both cases, it is transcribed with RNA polymerase II giving rise to a primary miRNA (pri-miRNA) transcript (Cai et al, 2004; Lee et al, 2004a). A pri-miRNA contains one or several internal hairpins carrying miRNAs (Lagos-Quintana et al, 2001; Lee et al, 2002). The internal hairpin is processed in the nucleus by the Microprocessor complex (Denli et al, 2004; Gregory et al, 2004). In case of an intronic pri-miRNA, processing is independent on the splicing machinery (Kim & Kim, 2007). The mammalian Microprocessor complex comprises of Drosha and DiGeorge syndrome critical region 8 (DGCR8) (Gregory et al, 2004). DGCR8 is known as Pasha in other animals like *Drosophila* or *Caenorhabditis elegans* (Denli et al, 2004). DGCR8 is a dsRNA binding protein that recruits and positions Drosha on a pri-miRNA (Gregory et al, 2004; Han et al, 2004;

Han et al, 2006; Landthaler et al, 2004). Drosha is a ribonuclease (RNase) III enzyme, which cleaves a pri-miRNA into a precursor miRNA (pre-miRNA), a short hairpin about 60-70 nt long with a 2nt overhang at the 3'end and a phosphate at 5'end (Lee et al, 2003). These characteristics are important for both export from the nucleus and subsequent Dicer processing. A pre-miRNA is recognized by Exportin 5 (known as Ranbp21 in flies), which in complex with Ran-GTP binds a short dsRNA with a 3'end overhang and facilitates its transport from the nucleus (Bohnsack et al, 2004; Lund et al, 2004; Okada et al, 2009). Upon GTP hydrolysis, the complex dissociates and Exportin 5 releases a pre-miRNA into the cytoplasm. In the cytoplasm, a pre-miRNA is bound and processed by Dicer, a multidomain protein whose two RNase III domains cleave a pre-miRNA into a miRNA duplex, a ~20base pair (bp) dsRNA with 3'end 2nt overhangs and phosphates at 5'ends (Bernstein et al, 2001; Grishok et al, 2001; Hutvagner et al, 2001). A miRNA duplex is then loaded onto a member of Argonaute (Ago) family of proteins forming a core of the RNA-induced silencing complex (RISC) (Hammond et al, 2001). One strand of a miRNA duplex, called a passenger strand, is discarded (Kawamata et al, 2009). The other strand, named either a guide strand or a miRNA, navigates Ago proteins to target mRNAs. Then the target mRNA is translationally repressed and/or destabilized (reviewed in detail in Jonas & Izaurralde, 2015).

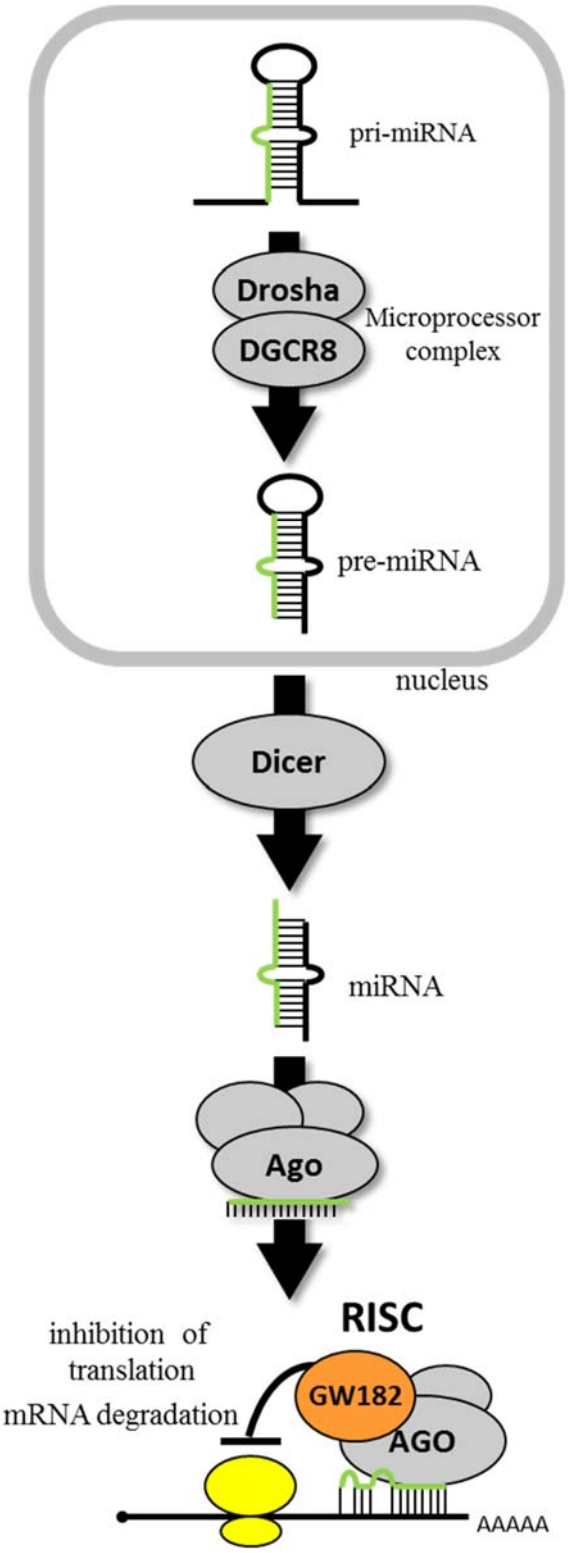

Fig. 1: The miRNA pathway. Adapted from Svoboda, 2014.

A typical miRNA in *Metazoa* is only partially complementary to its target mRNAs (reviewed in Bartel, 2009). The part pairing to the mRNA and, thus, sufficient for target recognition is called "seed" sequence. The seed

sequence is 6-8 nt long and it is localized at the 5'end of a miRNA. miRNAs of the same seed sequence are grouped into miRNA families, whose members might be to a significant extent redundant (Miska et al, 2007).

The origins and evolution of the miRNA pathway are somewhat enigmatic. Although the miRNAs are found in many species across plants and *Metazoa*, miRNAs of these two lineages differ from each other substantially (reviewed in Axtell et al, 2011). Unlike animal miRNAs, biogenesis of plant miRNAs takes place entirely in the nucleus and does not involve any homolog of Drosha (Park et al, 2005). Moreover, plant miRNAs are complementary with their targets to the higher extent (Llave et al, 2002; Tang et al, 2003) and typically have a smaller set of functional targets in comparison with animal miRNAs (Addo-Quaye et al, 2008). Therefore, there are two alternative models for miRNA pathway origins. The miRNA pathway might have emerged from an ancient RNAi twice independently, in the ancestor of *Metazoa* and in the plant lineage (Jones-Rhoades et al, 2006). Alternatively, the common ancestor of those two lineages already had the miRNA pathway and its diversity is a consequence of lineage-specific evolutionary requirements (Axtell et al, 2011).

The miRNA pathway regulates a wide range of essential cellular processes such as proliferation, differentiation and apoptosis. Approximately 60% of mammalian mRNAs contains a conserved miRNA binding site, which illustrates the extent of the miRNA pathway (Friedman et al, 2009). Therefore, it is not surprising that miRNA misregulation is associated with several diseases, most importantly with cancer (reviewed in Lin & Gregory, 2015; Sayed & Abdellatif, 2011).

1.1.2 RNAi

Originally, RNAi has been defined as a mechanism in which long dsRNAs triggers sequence specific degradation of complementary RNAs (Fire et al, 1998). Although some authors use the term more broadly for small RNA pathways, I will use it in its original sense.

RNAi is initiated by processing of a long dsRNA into ~22nt RNA duplexes, which guide degradation of complementary RNAs. A long dsRNA in the cell might come from several sources, e.g. a transcription of repetitive gene loci (Aravin et al, 2001), a bidirectional transcription or transcription of genes with inverted repeats (Sijen & Plasterk, 2003; Watanabe et al, 2008). Apart from endogenous dsRNAs, a viral infection is frequently accompanied by presence of long dsRNAs, either from replication of RNA viruses or from bidirectional transcription of DNA viruses (reviewed in Kumar & Carmichael, 1998). In all these cases, a dsRNA contains a long stretch of perfect complementarity

together with different types of termini, either blunt, with long single-stranded overhangs or with a loop of a variable length (Fig. 2). A long dsRNA is processed by Dicer into 22nt RNA duplexes named short interfering RNAs (siRNA) (Bernstein et al, 2001) (Fig. 3). A siRNA is loaded onto Ago protein of the RISC complex and a passenger strand is removed (Leuschner et al, 2006; Matranga et al, 2005; Rand et al, 2005). A guide strand becomes a part of the RISC complex determining cognate RNAs. If a guide strand is in a complex with a catalytically active Ago protein, a perfect or near-perfect complementarity between the guide strand and a cognate RNA directs cleavage of the cognate RNA (Hutvagner & Zamore, 2002). The Ago protein cleaves a target RNA across the bond between the tenth and eleventh nucleotides of the guide strand when counting from its 5'end (Elbashir et al, 2001).

The RNAi plays different biological roles in different organisms. From an evolutionary perspective, its antiviral role is retained in invertebrates, plants and many other organisms.

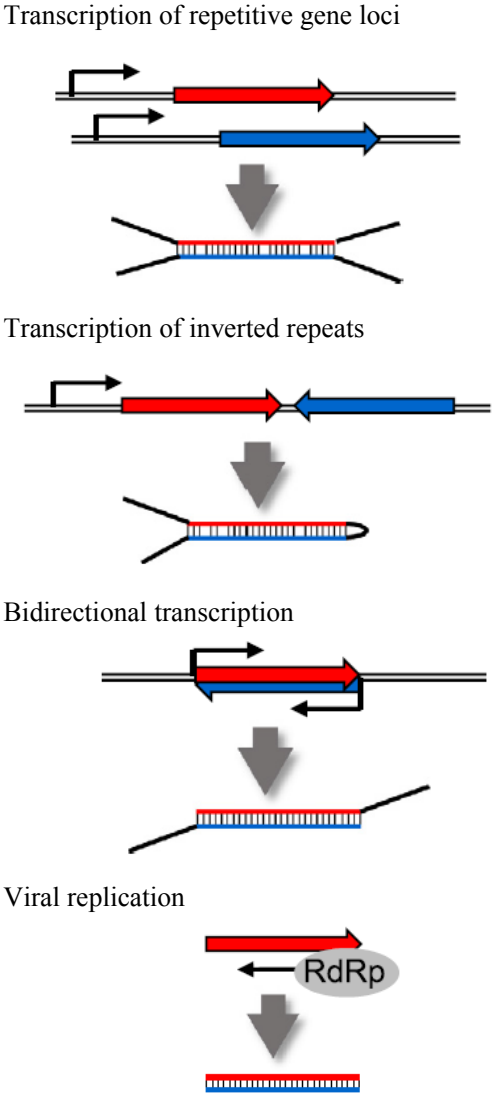

Fig. 2: Substrates of RNAi. A schematic representation of different sources of RNAi substrates in a cell. Note that the origin of a RNAi substrate determines the structure of its ends. Adapted from Svoboda, 2014.

1.1.3 Interplay between the miRNA pathway and RNAi in different animals

As mentioned earlier, the RNAi and miRNA pathways share a common concept; a small RNA generated by Dicer is loaded as a guide on an Argonaute protein, which executes the response. However, organisms differ as to what extent they separate the protein machinery carrying out small RNA biogenesis as well as subsequent target recognition and repression.

In *Drosophila melanogaster*, the RNAi and the miRNA pathway are mechanistically separated; each pathway has its dedicated protein components. Pre-miRNAs are cleaved by Dicer-1 (Dcr-1) and long dsRNAs are processed by Dicer-2 (Dcr-2) (Lee et al, 2004b). Generated miRNA duplexes and siRNAs are then handed over to Ago-1 and Ago-2, respectively (Hammond et al, 2001; Okamura et al, 2004). Interestingly, this partitioning does not depend on Dicer that generated the small RNA duplexes, but on their structure as miRNA duplexes typically have mismatches in central regions, whereas siRNAs tend to be perfectly complementary (Tomari et al, 2007). Accordingly, there are exceptions to the rule,

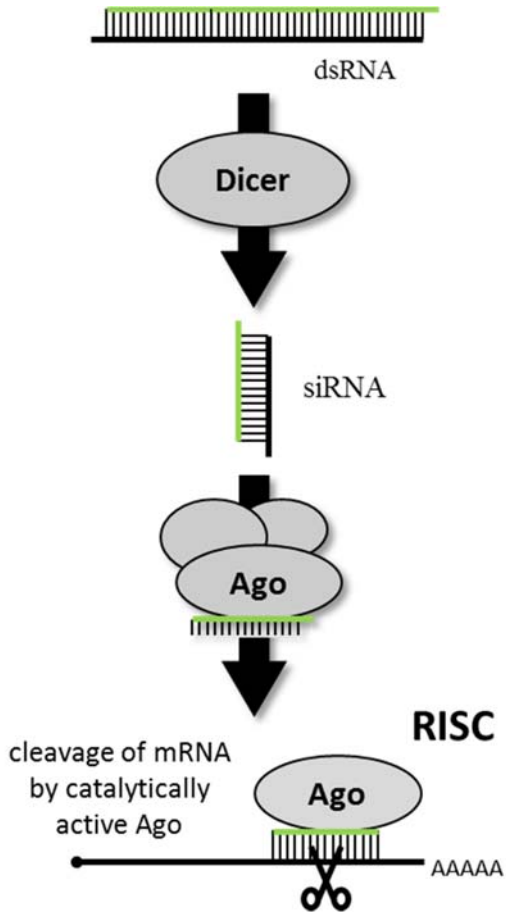

Fig. 3: RNAi. Adapted from Svoboda, 2014.

which can cross the imaginary barrier between the two pathways such as miR-277/miR-277*, which have an atypical secondary structure (Forstemann et al, 2007). In the specific situation of *D. melanogaster* where RNAi forms an antiviral defence, the mechanistic separation seems to be highly advantageous as the endogenous substrates do not occupy the antiviral pathway. This enables to accommodate conflicting evolutionary forces put on the pathways; RNAi should be flexible to cope readily with the pathogen-host arm race, whereas the miRNA pathway should be more conserved as it controls a wide range of developmental processes.

Although both *D. melanogaster* Dicers cleave short hairpins as well as long dsRNA substrates *in vitro* (Bernstein et al, 2001; Cenik et al, 2011; Liu et al, 2003), this versatility is disabled *in vivo* by several mechanisms.

First, it is believed that efficient processing of long dsRNA is connected with the processivity of Dicer as it performs several cleavages on a bound substrate without dissociation. Dcr-2 activity is stimulated in the presence of ATP indicating that there is a domain with ATPase activity, which increases the processing efficiency (Bernstein et al, 2001). In fact, Dcr-2 contains a domain with a putative helicase function, the N-terminal helicase, having all conserved motifs of functional helicases. However, the mechanism by which the helicase domain would increase Dicer activity is still elusive. As Dcr-1 has a poorly conserved N-terminal portion of the helicase, it compromises it for efficient processing of long dsRNA. In any case, the helicase of Dcr-1 binds the terminal loop of a pre-miRNA, which enhances pre-miRNA processing independently on ATP (Tsutsumi et al, 2011).

Second, each Drosophila's Dicer associates with one or more distinct dsRNA binding proteins (dsRBPs), which make the separation of Dicers' functions even stricter. Dcr-1 associates with Loquacious (Loqs) proteins, namely Loqs-PA and Loqs-PB isoforms, which are required for processing of pre-miRNAs (Forstemann et al, 2005), whereas R2D2 restricts Dcr-2 activity to endogenous precursors of siRNAs and viral dsRNAs (Cenik et al, 2011).

In *C. elegans*, small RNA pathways are complex because *C. elegans* genome encodes one Dicer gene, 27 Argonautes and 3 RNA-dependent RNA polymerases (RdRp) (reviewed in Ketting, 2011); some members of the latter group play role in amplification of a substrate for a more efficient repression

of complementary RNAs. Due to emergence of numerous new players, the conventional differentiation on the miRNA pathway and RNAi is blurred and several other pathways emerge instead. Nevertheless, bearing in mind certain simplification, I will still use the terms RNAi and the miRNA pathway in *C. elegans* for the purpose of my thesis.

In *C. elegans*, there is a partial separation of the protein machinery between the miRNA pathway and RNAi. They share Dicer expressed from a single locus, which however gives rise to two Dicer isoforms post-translationally; the full-length Dicer (DCR-1) is cleaved by yet unknown protease behind the Piwi Argonaute Zwille (PAZ) domain and the arisen C-terminal fragment, small DCR-1 (sDCR-1), is catalytically active (Sawh & Duchaine, 2013). The functions of two Dicer isoforms overlap to some degree (Sawh & Duchaine, 2013). DCR-1 processes long dsRNAs, substrates of RNAi, as well as pre-miRNAs (Ketting et al, 2001). DCR-1 shows a stimulation in activity in ATP presence (Ketting et al, 2001), which is in agreement with its domain architecture as it contains an intact helicase. sDCR-1 is believed to be capable of more efficient processing of RNAi substrates, although the mechanistic insight into its processing is still missing (Sawh & Duchaine, 2013). The truncated Dicer isoform appears in the later stages of the development (Sawh & Duchaine, 2013); the timing of its appearance could be correlative of the switch between necessity for the functional miRNA pathway in the early stages when it regulates development and RNAi in later stages when the development is almost finished and reinforcement of antiviral defence is of higher importance (Sawh & Duchaine, 2013). Following Dicer processing, small RNAs are sorted into pathway-specific Ago proteins according to their complementarity similarly to *Drosophila*; miRNA duplexes associate with Argonaute-like (ALG)-1 and ALG-2 and siRNAs associate with RNAi-Defective 1 (RDE-1) (Jannot et al, 2008; Steiner et al, 2007).

In mammals, RNAi and the miRNA pathway differ by their substrate structure and origin, while they share the same protein machinery from Dicer onwards. Once a ~22nt small RNA duplex is generated by Dicer, the pathways essentially merge because a small RNA can be loaded on any AGO protein (Yoda et al, 2010). Subsequently, the response triggered by the RISC complex depends on an AGO protein and on the level of complementarity between a small RNA and a target RNA (Doench et al, 2003; Hutvagner & Zamore, 2002). If the small RNA is perfectly complementary to a target RNA and it is bound to AGO2, the only mammalian AGO protein that is catalytically active (Liu et al, 2004; Meister et al, 2004), then the target RNA is cleaved (Hutvagner & Zamore, 2002; Liu et al, 2004; Meister et al, 2004). If the small RNA is only partially complementary or it is bound to other AGO protein than AGO2, then the target RNA is translationally repressed and destined for degradation (Doench et al, 2003; Meister et al, 2004).

In mammals, the RNAi and the miRNA pathway are essential in different developmental stages and tissues of the adult organism. The miRNA pathway is the dominant RNA silencing pathway in somatic cells, where it plays an essential role in a wide array of cellular processes as is documented on the phenotypes of conditional knockouts of principal components of miRNA protein machinery or miRNAs itself in various tissues (reviewed in Sayed & Abdellatif, 2011). On the other hand, the miRNA

pathway is non-essential for oocyte growth and early development, although several miRNAs are present (Ma et al, 2010; Suh et al, 2010). Mouse oocytes lacking DGCR8 are capable of ovulation, fertilization, zygotic transition and they reach a blastocyst stage without any phenotype or pronounced change in transcriptome (Suh et al, 2010). Nevertheless, the full *Dgcr8* knockout is embryonically lethal; the knockout embryos are resorbed after the 6.5 day stage, which is indicative of the functional requirement for the miRNA pathway in later embryonic stages (Wang et al, 2007).

RNAi seems to play only a minor, if any, role in somatic cells, as shown on the example of cell lines transfected with long RNA hairpins, which give rise to negligible amount of siRNAs, although the necessary machinery is present (Nejepinska et al, 2012). However, RNAi is present in oocytes as shown in several mammalian species by processing of artificially introduced dsRNAs (Anger et al, 2004; Svoboda et al, 2000; Wianny & Zernicka-Goetz, 2000). The capacity of RNAi to trigger sequence specific degradation upon presence of long dsRNAs lasts until blastocyst stage (Wianny & Zernicka-Goetz, 2000). Furthermore, there is a population of endogenous siRNAs found in mouse oocytes, which is derived from the transposable elements (Tam et al, 2008; Watanabe et al, 2008). In addition, RNAi is essential for the normal development of mouse oocytes as oocyte-specific Dicer knockout shows severe meiotic spindle defects in oocytes resulting in female infertility (Murchison et al, 2007; Tang et al, 2007); the precise molecular mechanism behind the phenotype has not been elucidated yet.

The apparent inefficiency of RNAi in somatic cells might be explained by emergence of a new pathway in vertebrates responsible for the antiviral defence in a cell, known as the interferon response (reviewed in Cullen et al, 2013; Svoboda, 2014). Unlike RNAi, the interferon response triggers systemic, complex and sequence independent response upon the presence of dsRNAs in a cell, which could represent products of viral infection (reviewed in Gantier & Williams, 2007). It is possible that the coexistence of two functional antiviral defence pathways might make the overall antiviral response less effective than when only one pathway is in charge leading to RNAi suppression in vertebrate evolution (personal communication with Petr Svoboda).

1.2 Dicer

As described above, Dicer is the central enzyme of two small RNA pathways, RNAi and the miRNA pathway. These pathways differ between themselves in the structure of their substrates, which are cleaved by Dicer. The evolution coped with such difference in substrate structure in several ways, which we can see in different organisms, while the typical mode of Dicer cleavage remained preserved. The Dicer substrates are processed from their termini (Zhang et al, 2002; Zhang et al, 2004). The precise length of a product is determined by the domain arrangement within Dicer as it corresponds to the distance between the PAZ domain binding the terminus of a substrate and the processing centre made of the RNase III domains cleaving the substrate (MacRae et al, 2007; MacRae et al, 2006).

The canonical domain architecture of metazoan Dicer comprises of the N-terminal helicase, the domain of unknown function 283 (DUF283), the PAZ domain, two RNase III domains and the dsRNA

binding domain (dsRBD). The helicase comprises the HEL1, the HEL2 and the HEL2i domains (Section 1.3) (Fig. 3A). Beside these canonical proteins, animals can contain other isoforms originating from separate genes, from the same gene by transcription from an alternative promotor or by proteolytic cleavage of Dicer (Flemr et al, 2013; Lee et al, 2004b; Sawh & Duchaine, 2013).

As revealed by cryo-electron microscopy, Dicer with the canonical domain architecture has an L-shaped structure, which can be divided into three regions: a cap, a platform and a branch (Lau et al, 2012; Lau et al, 2009; Taylor et al, 2013) (Fig. 3B). The cap of Dicer is formed by the PAZ domain and the DUF283 domain (Taylor et al, 2013). The platform comprises of two RNase III domains, dsRBD and the HEL1 domain (Lau et al, 2012; Taylor et al, 2013). Finally, the branch of Dicer is formed by the HEL2 and the HEL2i domains, which together with the HEL1 domain adopt the shape of a clamp (Lau et al, 2012; Taylor et al, 2013).

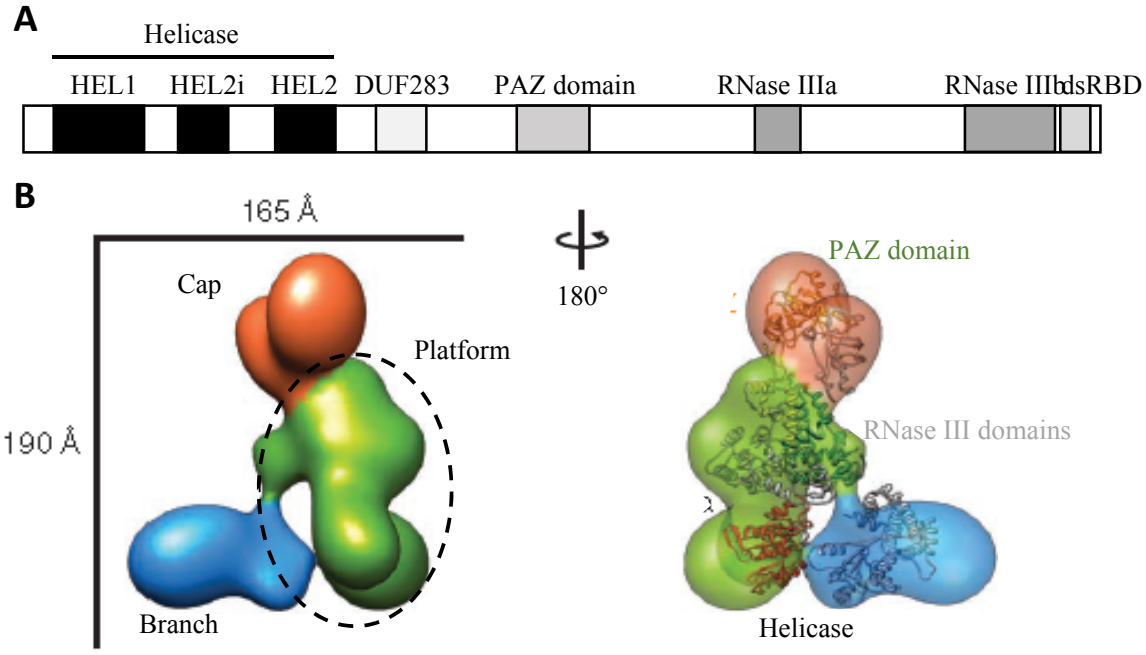

Fig. 3: Dicer. **A** A domain architecture of canonical Dicer. **B** An electron microscopy reconstruction of mammalian Dicer with marked regions (left panel) and several mapped domains (right panel). The pictures of Dicer structure are adapted from Taylor et al, 2013.

Dicer requires divalent cations for substrate cleavage, specifically magnesium cations, which can be replaced by manganese and cobalt cations (Provost et al, 2002; Zhang et al, 2002). The ions are localized in the active centre of RNase III domains where they participate in cleavage of dsRNA (MacRae et al, 2006). Their absence in the active centre, caused e.g. by addition of excess of ethylenediaminetetraacetic acid (EDTA) to the Dicer assay buffer, leads to uncoupling of substrate binding and cleavage by Dicer (Provost et al, 2002; Zhang et al, 2002). This approach is widely used in studies focusing on substrate binding especially when substrates with different structural features are used (e.g. (Taylor et al, 2013)). However, the same result might be achieved when Dicer is incubated with its substrate at 4°C (Provost et al, 2002; Zhang et al, 2002).

In general, substrate processing by an enzyme can be divided into three steps: binding of a substrate, catalysis and a product release. In case of Dicer, this scheme is more complicated because Dicer has a potential to cleave almost any dsRNA (Zhang et al, 2002) and it is important to narrow this ability to only these types of RNA, which are desirable substrates (Ma et al, 2008; Ma et al, 2012). The individual steps of Dicer substrate processing are discussed in the following sections.

As Dicers from various organisms differ considerably in their substrates and, subsequently, in their substrate processing, I will focus on mammalian Dicer in the following sections.

1.2.1 Substrate binding

Dicer employs several distinct domains to bind different structural features of its various substrates. A dsRNA terminus of a pre-miRNA or a perfect duplex is bound by the PAZ domain (MacRae et al, 2007). As substrates with blunt ends as well as substrates with overhangs of different size found on both ends are processed (Bernstein et al, 2001; Park et al, 2011; Vermeulen et al, 2005), it indicates that there is a space for accommodation of different termini into the PAZ domain.

A dsRNA structure is bound by the Dicer dsRBD. The dsRBD is important for substrate binding in the absence of the PAZ domain (Ma et al, 2012). However, whether such case ever occurs for the mammalian Dicer or whether it occurs in case of specific substrates (e.g. with undesirable ends like those discussed in the paragraph about internal cleavage) remains to be elucidated. However, it is tempting to suggest, that it contributes to the binding of so called passive substrates, RNA molecules with stretches of dsRNA, which represent a majority of Dicer binding sites in the cell in spite of the fact that they are not cleaved (Rybak-Wolf et al, 2014). Binding by the dsRBD might be important for substrate processing in case of sDCR-1 in *C. elegans*, which lacks the PAZ domain (Sawh & Duchaine, 2013), although it has not been tested yet.

The N-terminal helicase also participates in substrate binding. It binds both, pre-miRNAs and perfect duplexes, but in a distinct way and with different affinities. The helicase binds a pre-miRNA with a considerable affinity (K_D = 96 nM) (Ma et al, 2012). The interaction is mediated through the terminal loop of a pre-miRNA based on the results of cryo-electron microscopy and biochemical assays (Lau et al, 2009; Liu et al, 2015; Ma et al, 2012; Taylor et al, 2013). Perfect duplexes are bound by the helicase with lower but measurable affinity (K_D = 476 nM) (Ma et al, 2012). It is unknown what structural feature of perfect duplex is the target; the possible candidates are a terminus or a stretch of dsRNA.

1.2.2 Pre-miRNA processing

Although pre-miRNAs do not represent the most abundant Dicer binding sites in a mammalian cell (Rybak-Wolf et al, 2014), they are a preferred substrate of mammalian Dicer as short hairpins are processed with high activity under both single-turnover and multiple-turnover conditions when compared with perfect duplexes (Chakravarthy et al, 2010; Ma et al, 2008; Ma et al, 2012). Pre-miRNA recognition is mediated by the helicase sensing the terminal loop (Ma et al, 2012; Taylor et al, 2013).

In general, a wide array of pre-miRNAs is processed efficiently by mammalian Dicer when tested *in vitro* (Feng et al, 2012). This finding indicates that Dicer alone does not constitute a bottleneck of miRNA biogenesis, which would prevent a production of undesirable miRNAs (Feng et al, 2012). On the contrary, it seems to process pre-miRNAs more or less proportionally to their concentration in a cell (Feng et al, 2012). Nevertheless, other factors in a cell, which might be expressed only in certain contexts, might impart selectivity to Dicer.

The observed variation in the efficiency of pre-miRNA processing can be explained by differences in structural features such as types of ends, the size of the terminal loop, the position of the terminal loop as well as the extent of complementarity in the stem region of a pre-miRNA (Feng et al, 2012; Park et al, 2011). However, the impact of the enumerated structural features on Dicer processing cannot be explained simply by difference in binding affinity (Feng et al, 2012) and an additional layer of regulation seems to be in operation in mammalian Dicer.

A type of end of a pre-miRNA also influences Dicer processing. In general, pre-miRNAs with a 3' 2nt overhang are processed by mammalian Dicer more efficiently than blunt-ended pre-miRNAs or pre-miRNAs with a 5'end overhang, especially when comparing pre-miRNAs derived from one pre-miRNA species, so that the effect of other structural features remains constant (Feng et al, 2012; Park et al, 2011).

The size of the terminal loop and its position relative to the dsRNA terminus play a considerable role in Dicer processing of pre-miRNAs (Feng et al, 2012; Liu et al, 2015)}. There is a positive correlation between the size of the terminal loop and Dicer activity with ~17nt as the upper limit for the effect of the terminal loop size (Feng et al, 2012). This finding is of great importance when designing hairpin expressing vectors for knock-down experiments because a simple increase in the size of the terminal loop could lead to increased efficiency (Feng et al, 2012). The position of the terminal loop relative to the dsRNA terminus of a pre-miRNA does not have a considerable effect on Dicer activity if the stem region is at least 22bp long (Feng et al, 2012). The shorter stem region leads to decrease in Dicer activity as illustrated on the example of pre-miR-151, where editing of the dsRNA stem region by dsRNA-specific adenosine deaminase leads to an increase in the terminal loop size at the expense of the length of the stem region, which reduces the Dicer activity by ~50% (Liu et al, 2015).

Although the data are so far inconclusive, it seems that the extent of complementarity in the stem region has subtle or no effect on Dicer processing efficiency (Feng et al, 2012; Ma et al, 2012). Nevertheless, the presence of bulges and asymmetrical ínternal loops within the stem region leads to generation of longer miRNA duplexes (Starega-Roslan et al, 2011). The change in length has a great impact on the seed sequence of the 3'arm derived miRNAs and, subsequently, on their target sets (Starega-Roslan et al, 2011).

1.2.3 Pre-siRNA processing

As mentioned above, the majority of endogenous RNAi substrates in mammalian cells comprises of long perfectly complementary stretch of dsRNA and heterologous termini with one or two long overhangs. Such substrates require an internal cleavage performed by Dicer to generate two intermediates with preferred dsRNA termini containing a 3' 2nt overhang (Zhang et al, 2002). After that, the intermediates serve as normal Dicer substrates enabling cleavage from the terminus. They bind to the PAZ domain and are cleaved in several Dicer dependent cleavage rounds generating siRNAs of different sequence but complementary to original transcripts.

In any case, the ability of mammalian Dicer to perform an internal cleavage remains still enigmatic. The assay with substrate with 4nt loops instead of free ends shows that Dicer is able to cleave such substrate (Zhang et al, 2002). The resultant cleavage pattern suggests a possibility, that first internal cleavage of such substrate is not absolutely random. It seems to occur in certain positions of the substrate with higher probability, but the actual cleavage might be done within several nucleotides. Therefore, Dicer might orient itself on the substrate in a non-random way or certain features of the substrate enables Dicer to initiate the internal cleavage, therefore producing the free ends from which Dicer can further proceed in usual fashion.

Notwithstanding the character of cellular RNAi substrates, majority of *in vitro* studies have been performed with perfect duplexes of 35 bp, which has a 2nt overhangs on 3'ends, in order to avoid multiple cleavages per substrate in assays (e.g. (Chakravarthy et al, 2010; Ma et al, 2008; Ma et al, 2012; Taylor et al, 2013)). Nevertheless, it is questionable whether such approach could not lead to overlooking a yet unobserved, but considerable aspect of Dicer processing.

Dicer cleaves perfect duplexes of 35 bp with 2nt overhangs on 3'end with low activity when compared to processing of pre-miRNAs under both single-turnover and multiple-turnover conditions (Chakravarthy et al, 2010; Ma et al, 2008; Ma et al, 2012). Under single-turnover conditions, an enzyme is in excess over a substrate, which should theoretically lead to only one round of cleavage per enzyme and the measured cleavage rate should not be, therefore, influenced by the rate of product release. In case of perfect duplexes, ~30% of initial material has been cleaved after 40 min as opposed to ~90% for pre-let-7 (Chakravarthy et al, 2010); the exact proportions of cleaved substrate vary according to used concentration of Dicer and substrates (Ma et al, 2008; Ma et al, 2012). The difference in cleavage rate is even increased under multiple turnover conditions when substrate is in excess over enzyme, which enables several rounds of cleavage by the enzyme. The $t_{1/2}$ value, the time necessary for cleavage of a half of the substrate, is ~15 min for pre-let7 substrate whereas less than 1% of perfect duplexes is processed in that time (Chakravarthy et al, 2010). Taken together, Dicer is able of multiple turnover cleavage in case of pre-miRNAs, but not in case of perfect duplexes (Chakravarthy et al, 2010; Ma et al, 2008).

1.2.4 Product release

Dicer differs from a scheme of a general enzyme model because it associates with its products (Chendrimada et al, 2005; Noland et al, 2011; Zhang et al, 2002). In general, an enzyme quickly releases its product and does not form a stable complex with its product because it would obstruct an initiation of a new round of reaction it catalyses. By contrast, Dicer forms a complex with its products. This rather unusual characteristic has been observed for products of perfect or nearly perfect complementarity, which typically corresponds to siRNA duplexes (Chendrimada et al, 2005; Noland et al, 2011; Zhang et al, 2002). It remains to be determined whether mammalian Dicer is capable of stable association with miRNA duplexes as well. Furthermore, there have been contradictory results regarding the requirement for TRBP for such interaction. The pivotal biochemical study performed with mammalian Dicer shows that recombinant Dicer alone forms a complex with its products coming from cleavage of long perfect duplexes with triphosphate on 5'end as measured by mobility shift (Zhang et al, 2002). More recent biochemical studies show that the formation of heterodimer comprising of Dicer and TRBP or PACT is essential for binding of siRNAs, which are presented (Chendrimada et al, 2005; Noland et al, 2011). It is unclear, whether the absence of substrate processing step might be responsible for the contradictory results.

As mammalian Dicer was believed to be indispensable for loading of RISC with small RNAs, its ability to form a complex with its products was explained as a necessary prerequisite for handing the small RNAs over to an AGO protein (MacRae et al, 2008; Noland et al, 2011). Nevertheless, the recent studies show, that both TRBP and Dicer are dispensable for loading of small RNAs into RISC as well as for the subsequent strand selection (Betancur & Tomari, 2012; Kim et al, 2014; Suzuki et al, 2015). Therefore, it needs to be determined why Dicer, with or without TRBP, forms a complex with its products at the cost of potential detrimental effect on efficiency of its processing.

1.3 Dicer helicase

The Dicer helicase belongs to a large group of helicases which comprises of both RNA and DNA helicases. Despite the name of the group, these proteins have a wide range of functions in the cell, ranging from unwinding dsRNA or dsDNA to facilitating dissociation of proteins from RNA or stabilization the desired intermediate of multiple-step reactions (reviewed in Jankowsky & Fairman, 2007). Unlike DNA helicases, RNA helicases are in general less processive enzymes because majority of RNA forms only short stretches of dsRNA, unwinding of which requires only few reaction steps conducted by RNA helicases. The exception to this rule represents viral RNA helicases. Most RNA helicases require ATP for their activity.

Based on structural, functional and sequence comparison, helicases are divided into five or six superfamilies (SF) depending on author-adopted classification (Gorbalenya & Koonin, 1993; Singleton et al, 2007). SF1 and SF2 represent the largest superfamilies, which are further divided into several families and groups (Fairman-Williams et al, 2010). Dicer helicase belongs to retinoic acid-inducible

gene 1 (RIG-I)-like helicases, a family of the SF2 helicases (Zou et al, 2009). SF1 and SF2 (reviewed in Fairman-Williams et al, 2010) share a structural conservation of the catalytic core together with several conserved motifs. However, the sequence conservation is higher within each superfamily than between.

The typical structure of an SF2 helicase (reviewed in Fairman-Williams et al, 2010) comprises of two highly similar and tandemly arranged domains, which have so called RecA-like folds named after RecA protein. These domains are connected with a flexible linker and form an overall structure resembling a crayfish claw, each domain comprising one half of the claw. A substrate RNA is inserted into the cleft between the two halves of the claw, which is in agreement with the position of conserved motifs responsible for NTP binding, hydrolysis of NTP and substrate binding as they are placed on the surface facing the cleft between the domains. Such arrangement of conserved motifs indicates that both domains are necessary for unwinding. Nevertheless, it is possible, that non-canonical function might not require the presence of both domains.

The HEL1 domain contains several conserved motifs. Motif I and motif II are also known as Walker A NTP-binding motif and Walker B NTP-binding motif, respectively. They are responsible for binding of NTPs. The occurrence of these motifs is not confined to RNA helicases, but to proteins requiring NTP hydrolysis for their action in general. Motif Ia binds sugar-phosphate backbone of RNA substrate. The HEL2 domain contains motifs responsible for substrate binding (motif IV and V) and NTP binding (motif VI). The linker between the Hel1 and the Hel2 domains contains motif III, which is responsible for coupling NTP hydrolysis with the helicase unwinding activity.

In contrast to other SF2 helicases, RIG-I-like helicases have an additional domain called the HEL2i domain, which is inserted inside the HEL2 domain as is apparent from crystal structures (Kowalinski et al, 2011; Luo et al, 2011). In RIG-I, this domain interacts with dsRNA (Kowalinski et al, 2011; Luo et al, 2011) and this interaction seems to be preserved amongst different proteins from this family, as there is rather high sequence conservation of this domain between different members of the family (Luo et al, 2011). In addition, the Hel2i domain seems to provide an interface for evolution of protein-protein interactions without the necessity to meddle with conserved helicase catalytic core; e.g. the HEL2i domain of mammalian Dicer binds the third dsRBD domain of TRBP (Wilson et al, 2015).

The crystal structure of RIG-I indicates that RIG-I is unable of dsRNA unwinding activity. There are two mechanistic models describing unwinding activity in RNA helicases. In the model of local strand separation, RNA helicase binds dsRNA so that it distorts the A-helical structure of dsRNA, which promotes the subsequent strand dissociation (Yang et al, 2007). In the second model, a helicase translocate along a dsRNA and the unwinding is performed by specific β-hairpin motif, which bisects the dsRNA (Gu & Rice, 2010; Lam et al, 2003). Nevertheless, the binding of RIG-I to a dsRNA does not distort A-form conformation of a dsRNA and the β-hairpin motif necessary for unwinding is absent in RIG-I (Jiang et al, 2011; Kowalinski et al, 2011; Luo et al, 2011).

1.3.1 Roles of the Dicer helicase

Dicer helicase is not required for Dicer substrate cleavage in the strict sense because it is absent or mutated in Dicers of various species which are still capable of dsRNA cleavage, e.g. Dicer of *Giardia intestinalis* lacks the entire helicase, but its overexpression in *Schizosaccharomyces pombe* Dicer delete restores impaired RNAi (MacRae et al, 2006). Moreover, the Dicer helicase lacks a uniform role in the rest of eukaryotic Dicer, which have the helicase, but seems to adopt functions, which enable a rather highly conserved Dicer to adapt to various substrates specific for different pathways. These functions involve induction of processivity and restriction of substrate specificity (Bernstein et al, 2001; Kidwell et al, 2014; Ma et al, 2008).

As mentioned above, the helicase stimulates Dicer processing of long dsRNA at the expense of ATP (Bernstein et al, 2001). This ATP dependent stimulation is found in Dicers, which participates in RNAi by processing long dsRNA, such as *D. melanogaster* Dcr-2 or *C. elegans* DCR-1 (Ketting et al, 2001; Lee et al, 2004b). Although the mechanistic understanding of this ability is still missing, the proposed explanations are that the helicase might drives Dicer translocation along a long dsRNA or that it can help to rearrange a substrate to a conformation more suitable for cleavage (Bernstein et al, 2001; Hutvagner & Zamore, 2002; Nykanen et al, 2001).

The helicase restricts substrate specificity in *D. melanogaster* Dcr-1 and mammalian Dicer; the mechanisms behind its function differ between the two Dicers (Ma et al, 2008; Ma et al, 2012; Tsutsumi et al, 2011). In Drosophila Dcr-1, the regulatory role of the helicase is facilitated through higher affinity of Dicer to a ssRNA, represented by the terminal loop of pre-miRNA (Tsutsumi et al, 2011). Consequently, the higher binding affinity leads to higher activity of pre-miRNA processing by Dcr-1 (Tsutsumi et al, 2011). The efficient binding by the helicase requires the terminal loop and the stem region of certain size; any deviation from this rule leads to decrease in Dcr-1 processing (Tsutsumi et al, 2011). Similarly, *in vitro* deletion of the entire helicase or one of its domains results in impaired Dcr-1 processing of pre-miRNAs (Tsutsumi et al, 2011).

The helicase of mammalian Dicer restricts processing of undesirable substrates, which seems to be everything except of pre-miRNAs (Chakravarthy et al, 2010). In contrast to *Drosophila* Dcr-1, the *in vitro* deletion of mammalian Dicer helicase results in increased Dicer processing of both pre-miRNAs and perfect duplexes (Ma et al, 2008; Ma et al, 2012). Nevertheless, the increase is only mild in case of pre-miRNAs relative to the dramatic change in cleavage rate of perfect duplexes (Ma et al, 2008; Ma et al, 2012). It has been shown that the increase in Dicer processing of perfect duplexes upon the helicase deletion is attributed mainly to the pronounced change in k_{cat}, the catalytic constant for substrate conversion to product, rather than K_m, the Michaelis constant, or K_D (Ma et al, 2008). It indicates, that the effect of the deletion can be explained by increase in speed of catalysis, whereas affinity of the enzyme toward substrate seems to influence the resultant increase only mildly (Ma et al, 2008).

The intriguing behaviour of the helicase of mammalian Dicer has been explained by several mechanistic models. The findings from the biochemical studies support model, in which the helicase

operate as a switch. It forces Dicer into a less active conformation (Ma et al, 2008; Ma et al, 2012). The helicase relieves its coercion only when the proper substrate, pre-miRNA, is bound, which switches Dicer into an active conformation (Ma et al, 2008; Ma et al, 2012). Upon helicase deletion, the Dicer stays in the active conformation more often which leads to the increase in Dicer activity for both pre-miRNAs and perfect duplexes (Ma et al, 2008; Ma et al, 2012).

More recent model compares the helicase of mammalian Dicer to a trap, which is in agreement with data obtained by electron microscopy as well as with biochemical data (Ma et al, 2008; Ma et al, 2012; Taylor et al, 2013). In this model the helicase restricts Dicer processing of undesirable substrates by trapping them into a non-productive conformation (Taylor et al, 2013). The Dicer helicase changes the conformation upon dsRNA binding (Lau et al, 2012; Taylor et al, 2013) (Fig. 4). Interestingly, this conformational change is different depending on the substrate used. In case of pre-miRNA, Dicer forms a more open structure with a branch representing the helicase bending away from the platform when compared to the structure of apo-Dicer (Taylor et al, 2013). This movement is accompanied by the extension of the branch (Taylor et al, 2013). The percentage of opened Dicer particles as well as the processing activity of the protein seems to increase with the size of the terminal loop (Taylor et al, 2013). Furthermore, an additional mass appears near the platform of Dicer, which is interpreted as a bound pre-miRNA (Taylor et al, 2013). By contrast, Dicer undergoes an opposite conformational change upon binding perfect duplexes. The helicase bends toward the Dicer platform and compacts (Taylor et al, 2013). An additional rod-like structure appears bound to the PAZ domain and the helicase in a way that it is not in contact with the platform, a part which is predicted to hold processing centre of RNase III domains (Taylor et al, 2013). Therefore, perfect duplexes are trapped in the conformation, which does not allow Dicer cleavage, thus explaining low efficiency of Dicer processing in biochemical assays (Taylor et al, 2013).

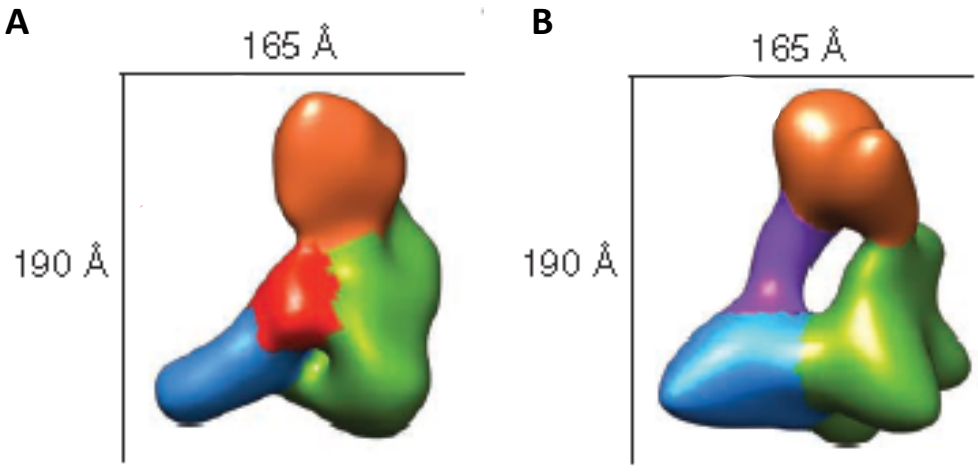

Fig. 4: Differential binding of substrates by mammalian Dicer. A,B Cryo-electron microscopy reconstructions of mammalian Dicer bound to a pre-miRNA (in red) (A) or a perfect duplex (in purple) (B). Adapted from Taylor et al, 2013.

It is important to note that although majority of Dicer particles with bound perfect duplex forms a closed conformation preventing Dicer cleavage, the rest remains either unchanged or undergoes a conformational change similar to the one performed by Dicer bound to a pre-miRNA (Taylor et al, 2013). It is tempting to suggest that this conformational diversity explains the finding that a tiny fraction of perfect duplexes is cleaved by Dicer, whereas majority stays trapped under multiple-turnover conditions (Chakravarthy et al, 2010; Ma et al, 2008).

1.3.2 A restoration of mammalian Dicer ability to process RNAi substrates

The role of the helicase in restricting mammalian Dicer to processing of pre-miRNAs indicates a simple way to extend Dicer substrate specificity when in need by its context dependent omission. Interestingly, this alternative way of overcoming selective substrate processing of mammalian Dicer has been partially employed only in oocytes of species inside the *Muridae* family (Flemr et al, 2013), which are represented by the mouse model.

In mouse oocytes, the transcription starting from an alternative promoter of retrotransposon origin within Dicer gene gives rise to an N-terminally truncated Dicer isoform ($Dicer^{O}$) lacking the HEL1 domain (Flemr et al, 2013). As the retrotransposon-derived promoters are active only in oocytes, the $Dicer^{O}$ transcript level drops rapidly in a two-cell embryo (Flemr et al, 2013). Nevertheless, the change in the transcript level might not correspond with protein level and $Dicer^{O}$ might linger in the embryo longer (Flemr et al, 2013). Interestingly, $Dicer^{S}$ is expressed in oocytes as well, although it represents only a small fraction of total Dicer present in oocytes (Flemr et al, 2013).

$Dicer^{O}$ is essential for oocyte development (Flemr et al, 2013). The specific $Dicer^{O}$ knock-out, generated by deletion of the alternative promoter, leads to female sterility accompanied by serious spindle defects and chromosome misalignments in oocytes (Flemr et al, 2013), phenotypes normally associated with Dicer knockout in oocytes (Murchison et al, 2007; Tang et al, 2007). Moreover, the absence of $Dicer^{O}$ results in substantial increase in levels of transcripts from transposable elements in oocytes (Flemr et al, 2013). Interestingly, the deletion of DicerO does not substantially influence levels of miRNAs (Flemr et al, 2013).

The characterization of $Dicer^{O}$ outside of oocytes showed that $Dicer^{O}$ is more active and has wider substrate specificity *in vivo* when compared to $Dicer^{S}$ (Flemr et al, 2013). The $Dicer^{O}$ ectopic expression in ESC without functional Dicer leads to the production of endo-siRNA from several internal hairpins and a dsRNA generated by a bidirectional transcription, which is not observed in case of $Dicer^{S}$ overexpression (Flemr et al, 2013). The difference in $Dicer^{S}$ and $Dicer^{O}$ siRNA production even increases, when an additional Dicer substrate in form of long hairpin with long flanking ends is introduced into the cell (Flemr et al, 2013). Taken together, the deletion of the HEL1 domain seems to have similar effect on Dicer processing of RNAi substrates *in vivo* as the deletion of the entire helicase (Flemr et al, 2013).

The *in vitro* fluorescent dicing assay showed that Dicer^{O} processes a short fluorescently labelled dsRNA of 28bp with 2nt overhang with higher activity than Dicer^{S} (Flemr et al, 2013). Nevertheless, *in vitro* cleavage assays with wider range of RNAi substrates are necessary to understand the difference between Dicer^{O} and Dicer^{S} and the impact of Dicer^{O} on the functionality of RNAi in oocytes.

2 Aims of the thesis

The main aim of my thesis was to examine whether Dicer^{O}, a new Dicer isoform produced specifically in mouse oocytes, differs from ubiquitously expressed Dicer^{S} in processing of various substrates. Achieving this aim involved two main tasks:

- Purification of Dicer protein isoforms at high purity and integrity

- Testing substrate processing by mouse Dicer isoforms using a non-radioactive cleavage assay

3 Material and Methods

3.1 Preparation of Dicer expression vectors

3.1.1 Ligation

Typically, 250 ng of a linearized dephosphorylated plasmid was ligated with an insert in a molar ratio 1:1 and 1:3 using 2.5 U of T4 DNA ligase (Thermo Scientific). The ligation reactions were incubated at room temperature for 30 min.

3.1.2 Chemical transformation

One vial containing 50 µl of TOP10 chemically competent *Escherichia coli* cells was thawed on ice and 3 µl of a ligation reaction was added. The cells were mixed by few gentle taps on the microtube and incubated on ice for 30 min. The microtube with cells was placed in a mixing block preheated to 42°C, incubated there for 45 s and quickly transferred to a container with ice for 2 min. 250 µl Lysogeny-broth (LB) medium was added to the cells and the cells were incubated vigorously shaking at 37°C for 1 h. Then the cells were spread on an LB agar plate with ampicillin (100 µg/ml). The agar plate was incubated at 37°C overnight.

3.1.3 Preparation of the pFastBacmf-nHisMyc plasmid

To generate a pFastBac plasmid with encoding a 6xHis-tag, a Myc-tag, a Tobacco Etch Virus nuclear-inclusion-a endopeptidase (TEV protease) cleavage site together with start and stop codons, synthesized oligonucleotides were annealed and completed in a PCR cycler (PCR programme: 94°C-4:00, 35x [94°C-0:30, 60°C-0:45, 72°C-1:00], 72°C-5:00) according to the previously published protocol (Yang & Sharan, 2003).

N-tagA 5'-CAAGGATCCACCATGGTACATCACCATCACCATCACAATCTCGAGCACG GTGAACAAAAACTTATTTCTGAAGAAGATCTGCGTCTCGAGGCAGAAAACC-3'

N-tagB 5'-TTGAAGCTTAGCGGCCGCCTGCGTCGACGCCCTGAAAATACAGGTTTTC TGCCTCGAGACGCAGATCTTCTTCAGAAATAAGTTTTTGTTCACCGTGCTC-3'

The PCR product was ligated into pJET1.2. The pJET1.2 plasmid with insert was cut with BamHI and HindIII generating the restriction fragment encoding N-tag, which was inserted into an empty pFastBacmf plasmid with modified multiple cloning site (as described in Flemr, 2013) linearized with BamHI and HindIII.

3.1.4 Preparation of the pFastBacmf-nHisMyc+Dicer$^{S/O}$ plasmids

In order to make the coding region of Dicer isoforms compatible with the backbones of other plasmids in the lab, stop and start codons were deleted and SalI and NotI restriction sites were introduced into the Dicer coding region at a cost of addition of several amino acid residues at the N- and the C-terminus of Dicer^{O} and Dicer^{S}. Expression vectors with Dicer^{S} and Dicer^{O} cDNA were kindly provided

by Matyáš Flemr. The start and stop codons were deleted and replaced with SalI and NotI restriction sites, respectively, by PCR (PCR programme: 95°C-3:30, 15x [95°C-0:30, 58°C-0:45, 68°C-7:00], 68°C-15:00) with Q5 High Fidelity DNA polymerase (Thermo Scientific) using the primers:

mDcrO-Fwd_SalIdelATG	5'-CAGTCGACAGCCGTGATACAGAAGTATACAC-3'
mDcrS-Fwd_SalIdelATG	5'-GTGTCGACGCAGGCCTGCAGCTCATGACCCC-3'
mDcr_Rev_NotI_delSTOP	5'-GTGCGGCCGCTGTTAGGAACCTGAGGCTGG-3'

The PCR products were cloned into a pJET1.2 vector, the corresponding plasmids were isolated from the selected positive colonies and digested with SalI and NotI to obtain inserts with Dicer^{O} and Dicer^{S} cDNAs. Dicer cDNA contains an internal SalI restriction site, thus the insertion of Dicer^{O} and Dicer^{S} was performed in two rounds. In the first round, the fragment encoding the C-terminal part common to both Dicer isoforms was inserted into the pFastBacmf-nHisMyc using SalI and NotI cleavage sites. In the second round, the generated plasmid with C-terminal part of Dicer was linearized with SalI and the remaining fragments encoding N-terminal part of Dicer^{S} or Dicer^{O} were inserted. The colonies were screened so that only positive colonies with N-terminal part of Dicer^{S} or Dicer^{O} in a proper orientation generated PCR product. The Dicer sequence was confirmed by sequencing.

3.1.5 Preparation of the pFastBacmf-nHisMyc+Dicer^{S}+cFlagGST plasmid

(prepared by and Radek Malík)

Preparation of the Dicer^{S} construct with N-terminal 6xHis-tag, Myc-tag and C-terminal 3xFLAG-tag and Glutathione S-transferase (GST) removable with TEV preotease comprised three rounds. In the first round, synthesized oligonucleotides containing 3x FLAG-tag were annealed and generated dsRNA insert was ligated into pFastBacmf with C-terminal placed 6xHis-tag and Myc-tag (not described in the thesis, digested with XhoI), yielding pFastBacmf-cFlag.

Flag_A 5'-TCGACGACTATAAGGACCACGACGGAGACTACAAGGATCATGATATTG
ATTACAAAGACGATGACGATAAGC-3'

Flag_B 5'-TCGAGCTTATCGTCATCGTCTTTGTAATCAATATCATGATCCTTGTAGTC
TCCGTCGTGGTCCTTATAGTCG-3'

In the second round, GST was PCR amplified from the pET4 plasmid using primers with additional XhoI or HindIII restriction sites.

GST_XhoI_Fwd	5'-CAACTCGAGTCCCCTATACTAGGTTATTGGAAAA-3'
GST_HindIII_Rev	5'-CTCAAGCTTATCATTTTGGAGGATGGTCGCCACCAC-3'

The PCR product was ligated into pJET1.2, from which it was restricted with XhoI and HindIII and cloned into XhoI/HindIII-cut pFastBacmf-cFlag yielding pFastBacmf-cFlagGST. In the third round, the fragment encoding 3xFLAG, GST was cut from pFastBacmf-cFlagGST with NotI and HindIII and

inserted into the plasmid pFastBacmf-nHisMyc+Dicer^{S} (pre-digested with NotI and HindIII). The sequence of plasmid insert was verified by sequencing.

3.1.6 Preparation of the pFastBacmf-nHisMyc+DicerS+cEGFP plasmid

(prepared by Jana Urbanová and Radek Malík)

To replace the GST-tag with Enhanced green fluorescent protein (EGFP), sequence of EGFP was obtained by PCR from the pEGFP-C1 plasmid using the primers with additional NotI or HindIII restriction sites:

TEV-EGFP_Fwd_NotI 5‘-CTGCGGCCGCGAAAACCTGTATTTTCAGGGCGTGAGCAAGGGC GAGGAGCTGTTC-3’

EGFP_Rev_HindIII 5‘-CTAAGCTTATCAGAGTCCGGACTTGTACAGCTCGTC-3’

The PCR product was cloned into pJET1.2, from which it was cut off by NotI and HindIII. The fragment was cloned into NotI/HindIII-cut pFastBacmf-nHisMyc+Dicer^{S}+cFlagGST yielding the pFastBacmf-nHisMyc+Dicer^{S}+cEGFP plasmid.

3.1.7 Preparation of the pFastBacmf-nTStrepHA+Dicer$^{S/O}$+cFlagHis plasmids

To generate the inserts encoding construct with a Twin-Strep-tag, an HA-tag, a 3xFLAG-tag and a 8xHis-tag, chemically synthesized oligonucleotides were ordered, which partially overlapped.

N_TwinStrep_HAtag_TEV_Fwd2 5’- ATCTACGGATCCACCATGGTATGGAGCCATCCTC AATTTGAAAAGGGTGGCGGGTCCGGCGGTGGGTCTGGCGGTAGCGCTTGGTCCCACCCCC AGTTCG -3’

N_TwinStrep_HAtag_TEV_Rev 5’-GTAGATGTCGACCAGGCCCTGAAAATACAGGTTT TCGGTACCAGCGTAATCTGGAACATCGTATGGGTAGTCACCCTTCTCGAACTGGGGGTGG GACCAA-3’

C_3C_2xFLAG_8xHis_Fwd 5’- TCTACAGCGGCCGCGGCGGACTCGAAGTGCTCTT CCAGGGACCTGCTAGCGACTATAAGGACCACGACGGAGACTA-3’

C_3C_2xFLAG_8xHis_Rev 5’- GTAGATAAGCTTAGTGATGGTGATGGTGATGGTG GTGGGACCCATCATGATCCTTGTAGTCTCCGTCGTGGTCCTT-3’

After their completion with Q5 High Fidelity DNA polymerase (programme: 94°C-2 min, 5x [94°C-0:25¸ 57°C-0:25, 72°C-1:00], 72°C-2:00), the DNA fragments encoding tags fused to the N-terminus and the C-terminus were digested by SalI, BamHI and EagI, HindIII, respectively, in order to adjust their ends for insertion into existing expression vectors. The fragment encoding the future C-terminus was inserted into the NotI/HindIII-cut plasmid pFastBacmf-nHisMyc+Dicer^{O}+cFlagGST (not described in the thesis). Into newly generated expression vector, the fragment corresponding to the prospective N-terminus was inserted using BamHI, SalI restriction sites. Nevertheless, as SalI cleavage removed the N-terminal part of Dicer, the generated vector contained only the C-terminal part of Dicer. The N-

terminal part of Dicer was inserted using SalI restriction site and the positive colonies were screened for a vector with the Dicer N-terminal fragment inserted in a proper orientation.

3.1.8 Generation of a recombinant bacmid

One vial containing 100 µl of frozen competent DH10Bac *Escherichia coli* cells was thawed on ice, the cells were placed into a 2 ml microtube with a round bottom. From 10 to 100 ng of some pFastBacmf plasmid was added to the bacterial cells. The cells were mixed by few gentle taps on the microtube and incubated on ice for 30 min. The microtube with cells was placed in a mixing block preheated to 42°C, incubated there for 45 s and quickly placed to a container with ice for 2 min. Then 800 µl LB was added to the cells and the whole volume was transferred to a plastic culture tube and incubated shaking vigorously at 37°C for 5 h. During the incubation, the plastic culture tube was only partially closed in order to enable access of oxygen to bacterial culture. Then 150 µl and 300 µl of bacterial culture were spread on LB agar plates (with kanamycin (50 µg/ml), tetracycline (10 µg/ml), gentamicin (7 µg/ml)), which were spread with 40 µl of isopropyl-β-D-thiogalactopyranoside (IPTG, Fermentas, 40 mg/ml) and 40 µl of 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-gal, Thermo Scientific, 20 mg/ml, dissolved in dimethyl sulfoxide) in advance. The agar plates were incubated at 37°C for 48 h. The extended incubation was necessary for negative colonies to turn bright blue.

Several positive, white colonies were picked into 50 µl of LB medium (with kanamycin (50 µg/ml), tetracycline (10 µg/ml), gentamicin (7 µg/ml)) and screened for the presence of a bacmid encoding Dicer (PCR programme: 94°C-2:00, 38x [94°C-0:30, 55°C-0:30, 72°C-2:00], 72°C-7:00). The primers were chosen so that only positive colonies gave PCR product. A pair of primers huDicer_Fwd3 and M13_Rev was used for all constructs apart the one with EGFP, when the huDicer_Fwd3 primer was replaced with EGFP_Fwd_seq primer.

EGFP_Fwd_seq	5'-GACCACTACCAGCAGAACACCC-3'
huDicer_Fwd3	5'-CCAGAAACTGCCAAATTTAGCCC-3'
M13_Rev	5'-CAGGAAACAGCTATGAC-3'

The positive colonies were inoculated into 300 ml of LB medium containing kanamycin (50 µg/ml), tetracycline (10 µg/ml) and gentamicin (7 µg/ml), incubated shaking vigorously at 37°C overnight, and then subjected to DNA isolation by NucleoBond Xtra Midi (Macherey-Nagel) according to the manufacturer's manual.

3.1.9 Transfection of insect cells to produce P1 baculoviral stock

5.4 x 10^6 Sf9 insect cells were centrifuged (900 x g, 3 min), supernatant was removed and they were resuspended in complete TNM-FH (TNM-FH medium with 10% (v/v) Fetal Bovine Serum, both from Sigma-Aldrich) to 450 000 cells/ml. 2 ml of the cell suspension was added to each well of a 6-well plate and the cells were incubated at 27°C for 1 h in order to let the cells attach to the bottom of wells. Medium was removed, the layer of cells was washed with 2ml of TNM-FH and 800 µl of TNM-FH was added

to the cells carefully. In the meantime, 6 µl of Cellfectin (Invitrogen) and 1 µg of bacmid DNA were resuspended separately in 100 µl of TNM-FH. Then they were mixed together and ~210 µl mixture was incubated at room temperature for 15-45 min and transferred to the cells. The cells were incubated with the mixture at 27°C for 5 h. After 5 h, the medium containing Cellfectin:DNA complexes was removed and 2 ml of complete TNM-FH was added. The cells stayed at 27°C for 72 h. During this period, they developed the indications of infection: an increase in cell's size and granularity accompanied with almost no change in a number of cells followed by their death. Medium containing released viruses was collected, centrifuged (500 x g, 3 min) in order to remove dead cells and large debris, supernatant was divided into 400 µl aliquots representing P1 viral stock and stored at -80˚C.

3.1.10 Amplification of a viral stock

900 000 Sf9 cells growing in suspension culture were centrifuged (500 x g, 3 min) and supernatant was removed. The cells were gently resuspended in 9 ml of complete TNM-FH, transferred to a 75cm^2 cell culture flask and incubated at 27°C for 1 h in order to let cells attach to the bottom. After 1 hour, either 400 µl of P1 viral stock or 1 ml of P2 viral stock was added to the attached cells and the infected cells were incubated at 27°C until the cells started dying, usually 4-5 days after infection. Then medium containing viruses was collected into a 15 ml tube, centrifuged (500 x g, 3 min) and the supernatant was carefully transferred to a new 15 ml tube. To avoid access of light, the viral stock was wrapped in aluminium foil and stored at 8°C.

3.1.11 Infection of insect cells

To selected volume of Sf9 cell culture at a density of 1 x 10^{x6} cells/ml, 0.1 volume of P3 viral stock was added. The infected cells were incubated shaking (115 rpm) in complete TNM-FH medium (containing 0.1% (w/v) Pluronic F-127) at 27°C for 48 hours. They were centrifuged (500 x g, 3 min), the supernatant was decanted and pellet, formed by cells expressing a recombinant protein, was used for Dicer purification immediately or stored at -80°C.

3.2 Dicer purification protocols

Dicer purification was performed at 8°C. Dicer purification protocols differed between individual Dicer protein constructs. Therefore, they are described separately.

3.2.1 Single-step Dicer purification using TALON resin

Based on the previously published protocol (Zhang et al, 2002), 15 ml of TALON resin was divided evenly into five Pierce Disposable Plastic Columns and each fraction was equilibrated 3x with 10 ml of W100 Buffer (20 mM Tris-HCl [pH 7.5], 100 mM NaCl, 1 mM $MgCl_2$, 10% (v/v) glycerol, 0.5% (v/v) Triton X-100). 5 x 10^8 Sf9 cells expressing the nHisMyc+Dicer^S construct or the nHisMyc+Dicer^O construct were washed with 20ml of phosphate-buffered saline (PBS) buffer, centrifuged (500 x g, 3 min) and supernatant was discarded. The pellet was resuspended in 25 ml of Lysis buffer (20 mM Tris-

HCl[pH 7.5], 100 mM NaCl, 1 mM $MgCl_2$, 10% (v/v) glycerol, 0.5% (v/v) Triton X-100, 1x Protease Inhibitor Cocktail Set III, Animal-Free (Calbiochem), 2 mM phenylmethylsulfonyl fluoride (PMSF)). The cells were lysed by five passages through a 21G needle and five passages through a 26G needle. The crude lysate was centrifuged (16 000 x g, 30 min, 4˚ C) to pellet the debris. The supernatant was incubated with the TALON resin in the closed Pierce Disposable Plastic Column on a lab rotator mixer for 1 h. Then the suspension sedimented and the unbound fraction of lysate was drained from the column and collected as the flow-through sample. Resin in each column was washed 3x with 10 ml of Lysis buffer, 3x with 10 ml of W800 buffer (20 mM Tris-HCl [pH 7.5], 800 mM NaCl, 1 mM $MgCl_2$, 10% (v/v) glycerol, 0.5% (v/v) Triton X-100) and 3x with 10 ml of W100 buffer. Proteins were eluted from TALON resin in four rounds with 1.5 ml of W100 I200 buffer (20 mM Tris-HCl, [pH 7.5], 100 mM NaCl, 1 mM $MgCl_2$, 200 mM imidazole, 10% (v/v) glycerol, 0.5% (v/v) Triton X-100). In each round, resins were resuspended in W100 I200 buffer prior eluate collection. Eluates of the same round from separate resins were pooled together. The pooled eluates were transferred into two 10k MWCO Slide-A-Lyzer Dialysis Cassettes (Thermo Scientific), which were dialyzed against 2 l of pre-chilled Buffer D (20 mM Tris-HCl [pH 7.5], 100 mM NaCl, 1 mM $MgCl_2$, 50% (v/v) glycerol, 0.1% (v/v) Triton X-100) overnight. The next day, Buffer D was exchanged for a new one, in which the dialysis cassette stayed for 2 h. The recombinant Dicer was collected from each cassette separately and stored at -80°C.

3.2.2 Two-step purification using Ni Sepharose and Glutathione Agarose

2 x 10^8 Sf9 cells expressing the nHisMyc+$Dicer^S$+cFlagGST construct were resuspended in 6 ml of Lysis buffer II (20 mM Tris-HCl [pH 7.5], 500 mM NaCl, 10% (v/v) glycerol, 0.5% (v/v Triton X-100, 1 mM DTT, 20 mM imidazole, 1x Protease Inhibitor Cocktail Set III, Animal-Free (Calbiochem), 1x SigmaFAST Protease Inhibitor Cocktail Tablet, EDTA-free (Sigma-Aldrich)). The cells were lysed by five passages through a 21G needle and five passages through a 26G needle. The crude lysate was centrifuged (16 000 x g, 30 min, 4˚ C) to pellet the debris. As the supernatant was filtered using a 0.2µm filter, the filter clogged and less than three quarters of the lysate spilled. The remaining supernatant was applied to HisTrap High Performance column (filled with Ni Sepharose High performance, 1 ml, GE Healthcare Life Sciences), pre-equilibrated with Buffer A (20 mM Tris-HCl [pH 7.5],, 500 mM NaCl, 10% (v/v) glycerol, 0.5% (v/v) Triton X-100, 1 mM DTT, 20 mM imidazole). The unbound fraction of the lysate was collected as the flow-through sample. The resin with bound $Dicer^S$ was washed with Buffer A and then eluted with gradient of Buffer B (20 mM Tris-HCl [pH 7.5], 500 mM NaCl, 10% (v/v) glycerol, 0.5% (v/v) Triton X-100, 1 mM DTT, 300 mM imidazole). The eluate was collected into 25 0.5 ml fractions. Based on the amount of protein measured by UV absorption at 280 nm, the eluate fractions 5–13 were pooled and used for the second purification step. The pooled eluate fractions were desalted using five serially arranged HiTrap Desalting columns (prepacked with Sephadex G-25 Superfine, 5 ml, GE Healthcare Life Sciences) in order to remove imidazole and then they were applied to Pierce Glutathione Chromatography Cartridge (prepacked with Glutathione Agarose beads, 1 ml,

Thermo Scientific) pre-equilibrated with Buffer C (50 mM Tris-HCl [pH 8], 150 mM NaCl, 1 mM DTT). Bound recombinant Dicer^{S} was washed with Buffer C and eluted with Buffer D (50 mM Tris-HCl [pH 8], 150 mM NaCl, 1 mM DTT, 20 mM reduced glutathione). The eluate was collected in ten 1 ml eluate fractions and stored at -20°C.

3.2.3 Two-step purification using TALON resin and GFP-Trap resin

1 ml TALON resin was transferred into a Pierce Disposable Plastic Column and equilibrated 3x with 10 ml W100 Buffer. 5 x 10^7 Sf9 cells expressing the nHisMyc+Dicer^{S}+cEGFP construct were resuspended in 5 ml of Lysis buffer III (20 mM Tris-HCl [pH 7.5], 100 mM NaCl, 1 mM $MgCl_2$, 10% (v/v) glycerol, 0.5% (v/v) Triton X-100, 1x Protease Inhibitor Cocktail Set III, Animal-Free (Calbiochem), 1x cOmplete, EDTA-free Protease Inhibitor Cocktail (Roche)). The cells were lysed by five passages through a 21G needle and five passages through a 26G needle. The crude lysate was centrifuged (16 000 x g, 30 min, 4˚ C). The supernatant was incubated with the TALON resin in the Pierce Disposable Plastic Column on a lab rotator mixer for 1 h. Then the suspension was sedimented and the unbound lysate fraction was drained from the column and collected as the Flow-through sample. The TALON resin with bound Dicer^{S} was washed 3x with 10 ml of W100 Buffer, 3x with 10 ml of W800 buffer and 3x with 10 ml of W100 buffer. Proteins were eluted from the TALON resin 3x with 700 µl of W100 I200 buffer; the eluate fractions were pooled. 40µl of the GFP-Trap resin (ChromoTek) was transferred into a Micro Bio-Spin Column (Bio-Rad) and equilibrated with 2x 500 µl of EGFP buffer (20 mM Tris-HCl [pH 7.5], 150 mM NaCl). 4x 500 µl of eluate from the TALON resin was incubated with the resin for 15 min on a lab rotator mixer. After each incubation, the supernatant was collected and named Flow-through (1-4). The GFP-Trap resin with bound Dicer was washed 1x with 500µl of EGFP Buffer, 1x with 500µl of EGFP Washing Buffer (20 mM Tris-HCl [pH 7.5], 500 mM NaCl) and 1x with 500µl of EGFP Buffer. Once the resin was resuspended in the third washing step, it was divided into halves, transferred into microtubes, centrifuged (2000 x g, 2 min, 4°C) and the supernatant was removed. 95 µl of TEV protease cleavage buffer (50 mM Tris-HCl [pH 8], 1 mM DTT, 0.5 mM EDTA), 10 µl 100% glycerol and 5 µl either commercially available TEV protease (1 mg/ml, Sigma) or in-house made TEV protease (2.4 mg/ml, kindly provided by Václav Urban) were added to the pelleted resins. The TEV protease cleavage reactions were incubated at 8°C overnight.

3.2.4 Two-step purification using TALON resin and Strep-Tactin resin

500 µl of TALON resin was transferred into a Pierce Disposable Plastic Column and equilibrated 3x with 5 ml of W100 Buffer. 5 x 10^8 Sf9 cells expressing the nTStrepHA+Dicer^{S}+cFlagHis construct or the nTStrepHA+Dicer^{O}+cFlagHis construct were resuspended in 5 ml of Lysis buffer. The cells were lysed by five passages through a 21G needle and five passages through a 26G needle. The crude lysate was centrifuged (16 000 x g, 30 min, 4˚C). The supernatant was incubated with the TALON resin in the closed Pierce Disposable Plastic Column on a lab rotator mixer for 1 h. Then the suspension was sedimented and the rest of lysate was drained from the column. The TALON resin with bound Dicer

was washed 1x with 5 ml of Lysis buffer, 1x with 5 ml of W800 buffer (20 mM Tris-HCl [pH 7.5], 800 mM NaCl, 1 mM $MgCl_2$, 10% (v/v) glycerol, 0.5% (v/v) Triton X-100, 1x Protease Inhibitor Cocktail Set III, Animal-Free (Calbiochem), 2 mM PMSF) and 1x with 5 ml of W100 Buffer. Proteins were eluted from the TALON resin 10x with 500 µl of W100 I200 buffer; the eluate fractions were collected together. The eluate was incubated with 500 µl of Strep-Tactin resin (IBA), placed into the Pierce Disposable Plastic Column and equilibrated 3x with 5ml of Buffer W (100 mM Tris-HCl [pH 7.5], 150 mM NaCl, 1 mM $MgCl_2$). The unbound fraction was drained and collected. The resin with bound Dicer was washed 3x with 5 ml of Buffer W and eluted 10x with 500 µl of Buffer E (100 mM Tris-HCl [pH 7.5], 150 mM NaCl, 1 mM $MgCl_2$, 5 mM desthiobiotin). The eluate fractions were collected separately, but they were pooled during the concentration and buffer exchange step using Amicon® Ultra-0.5 centrifugal filter device (100K, Merck Millipore). The Buffer E was exchanged for Storage buffer (20 mM Tris-HCl [pH 7.5], 100 mM NaCl, 1 mM $MgCl_2$, 50% (v/v) glycerol). Concentrated Dicer preparations were stored at -80°C.

3.3 SDS-Polyacrylamide gel electrophoresis (PAGE)

Samples, diluted with water to 25 µl, were mixed with 5 µl 6x SDS sample buffer (375 mM Tris-HCl, [pH 6.8], 12% (w/v) SDS, 60% (v/v) glycerol, 0.6 M DTT, 0.06% (w/v) bromophenol blue) and denatured at 90°C for 5 min. 3.5 µl of PageRuler Plus Prestained Protein Ladder (Fermentas) and 25 µl of samples were loaded onto gel (unless stated otherwise). Protein samples were resolved on polyacrylamide gels (5-12% separating gel, 5% stacking gel) using 80 -150 V in 1x SDS running buffer (25 mM Tris, 192 mM glycine, 0.1% (w/v) SDS). Once the tracking dye reached the bottom of the gel, the gel sandwich was disassembled and the separating gel was used either for Western blot or for Coomassie staining.

3.4 Western blot

A separating gel was soaked in Semidry transfer buffer (12.5 mM Tris, 96 mM glycine, 10% (v/v) methanol) for 1 min. A polyvinylidene difluoride (PVDF) membrane of desired size was soaked for 1 min in 100% methanol and then transferred equilibrated in Semidry transfer buffer. Then the sandwich was made from bottom to top of two Whatmann 3MM blotting filter papers, two filter papers, a PVDF membrane, a gel, two filter papers and two Whatmann 3MM blotting filter papers. Following addition of every paper, bubbles were pushed out from the sandwich. The transfer was run at 35 V for 55 min. Then the sandwich was disassembled; the PVDF membrane was soaked in 100% methanol and dried in order to increase the affinity between proteins and the PVDF membrane. The membrane was soaked in 100% methanol once again for activation, transferred to a container with Tween-Tris-buffered saline (TTBS) buffer (20mM Tris, [pH 7.5], 150mM NaCl, 0.02% Tween 20) for equilibration and blocked with 20 ml of Blocking buffer (5% fat-free milk dissolved in TTBS buffer) at room temperature

for 1 h. The membrane was incubated with 8 ml of solution of primary antibody rocking at 8°C overnight.

Primary antibodies:

Anti-Dicer antibody (349) (1:5 000, rabbit, polyclonal)

c-myc antibody, HRP conjugate (1:2 000, mouse, monoclonal, 9E10, Exbio)

Anti-Flag M2 antibody (1:5 000, mouse, monoclonal, M2, Sigma)

Anti-His-tag antibody, HRP conjugate (1:2 000, mouse, monoclonal, HIS-1, Sigma)

Anti-HA tag antibody (1:5 000, rat, monoclonal, clone 3F10, Roche)

Secondary antibodies:

Pierce goat anti-rabbit IgG (H+L), HRP conjugate (1:50 000, Thermo Scientific)

Pierce goat anti-rat IgG (H+L), HRP conjugate (1:50 000, Thermo Scientific)

Pierce goat anti-mouse IgG (H+L), HRP conjugate (1:50 000, Thermo Scientific)

The following day, the procedures differed according to the used primary antibody. In case of a primary antibody not conjugated with horseradish peroxidase (HRP), the membrane was washed 4x with 20 ml of Blocking buffer and incubated with solution of secondary antibody conjugated with HRP (1:50 000, 0.25µl of a secondary antibody per 12.5 ml of Blocking buffer) for 1 hour at room temperature. Then the membrane was washed 6 x with 20 ml of TTBS buffer. In case of primary antibody conjugated with HRP, the membrane was directly washed 6x with 20 ml of TTBS buffer. After washing, the membrane was incubated with either 500 µl of SuperSignal™ West Pico Chemiluminescent Substrate (Thermo Fisher Scientific) or SuperSignal™ West Femto Maximum Sensitivity Substrate (Thermo Fisher Scientific) for 1 min and then exposed to a film for a period ranging from 1 s to 10 min. The exposed films were developed using OPTIMAX 2010 (Protec Processor Technology).

3.5 Coomassie staining

Following SDS-PAGE, a separating gel was washed in Destaining solution (10% (v/v) methanol, 25% (v/v) acetic acid) for 5 min. Then it was transferred to a container with Staining solution (50% (v/v) methanol, 10% (v/v) acetic acid, 0.05% (w/v) Coomassie Brilliant Blue R-250), rocking at room temperature overnight. The following day, Staining solution was replaced with Destaining solution, which was exchanged several times until the gel was satisfactory destained so that only blue bands representing separated proteins were visible. For a faster destaining, a piece of soft foam, which adsorbed the dye, was placed into the container with Destaining solution and/or or the gel was heated in a microwave oven for 1 min.

3.6 Substrate preparation for dicing assay

3.6.1 DNA Template preparation

DNA templates for blunt-ended 130nt dsRNA substrate (Zhang et al, 2002) were prepared from plasmid pEGFP-C1 by PCR (PCR programme: 94°C for 2:00 min, 20 x [94°C-0:25, 60°C-0:25, 72°C-0:30], 72°C-7:00) with proofreading Q5 High-Fidelity DNA polymerase (New England Biolabs) according to the published protocol (Zhang et al, 2002). In the first round, the PCR products were resolved on a 2% agarose gel, excised from the gel and isolated using the QIAquick Gel Extraction Kit (Qiagen). In subsequent preparations, the DNA template from the first round was used as the DNA template for another PCR amplification and the PCR products were isolated with the QIAquick PCR Purification Kit (Qiagen) due to higher yield of DNA.

DNA templates necessary for dsRNAtetra substrate generation were produced by PCR (PCR programme: 94°C-2:00, 20x [94°C-0:25, 60°C-0:25, 72°C-0:30], 72°C-7:00) with Q5 High-Fidelity DNA polymerase (New England Biolabs) using primers mounting near the ends of synthesized oligonucleotides (Sigma-Aldrich).

Strand A:

Tetraloop_A 5’-CGCGAAATTAATACGACTCACTATAGGGTGCTCAGGTAGTGGGAAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGA-3’

T7_promotor_primer 5’-TAATACGACTCACTATAGGG-3’

Tetraloop_A_Rev 5’-TCCAGCAGGACCATGTGATCG-3’

Strand B:

Tetraloop_B 5’-CGCGAAATTAATACGACTCACTATAGGCATGGACGAGCTGTACGAAAGTACAGCTCGTCCATGCCTCCAGCAGGACCATGTGATCGCGCTTCTCGTTGGGGTCTTTGCTCAGGGCGGACT-3’

T7prom_Tetraloop_B 5’-TAATACGACTCACTATAGGC-3’

Tetraloop_B_Rev 5’-AGTCCGCCCTGAGCAAAGACC-3’

The DNA templates were resolved on a 2% agarose gel, excised and isolated with the QIAquick Gel Extraction Kit (Qiagen).

The DNA template for pre-let-7a3 substrate was generated using a synthesized oligonucleotide encoding the entire pre-let-7a3 sequence and two specific primers by PCR (PCR programme: 94°C-2:00 min, 20x [94°C-0:25, 60°C-0:25, 72°C-0:30], 72°C-7:00) with Q5 High-Fidelity DNA polymerase (New England Biolabs). The PCR product was resolved on a 2% agarose gel. As the PCR product formed only one band, it was isolated with the QIAquick PCR Purification Kit (Qiagen) in subsequent preparations.

pre-let-7-pcDNA_fwd 5´-GATCTTTAATACGACTCACTATATGAGGTAGTAGGTTGTATAGTTTGGGGCTCTGCCCTGCTATGGGATAACTATACAATCTACTGTCTTTCCTT-3´

T7_promotor_TG	5’-CTTTAATACGACTCACTATATG-3’
pre_let_7a3_rev	5’-GGAAAGACAGTAGATTGTATAG-3’

3.6.2 *In vitro* transcription

RNA substrates for dicing assay were generated by *in vitro* transcription with T7 RNA polymerase (Thermo Scientific) using the reaction set-up:

x µl DNA template (500ng)
1.5 µl T7 RNA polymerase (30 U, Fermentas)
1.25 µl Ribolock Ribonuclease Inhibitor (RI) (50 U, Thermo Scientific)
10 µl NTPs, mix (10 mM each)
10 µl 5x Transcription Buffer
to 50 µl H_2O

The reaction set-up was scaled up as necessary. The transcription reactions were incubated at 37°C overnight. The following day, 1 U DNase I (Fermentas) was added to each reaction to degrade DNA templates; reactions were incubated at 37°C for 15 min. In case of pre-let-7a3 substrate, the treatment with DNase I was coupled with dephosphorylation using 5U Shrimp Alkaline Phosphatase (Fermentas) in order to remove phosphate groups at the 5’end of pre-let-7a3 as only one phosphate group is present at the 5’end of a pre-miRNA as a result of Drosha cleavage.

RNA was isolated using the mirPremier microRNA Isolation Kit (Sigma-Aldrich). Following the procedure for small RNA isolation from mammalian cell cultures, ~50 µl of the transcription reaction was added to 450 µl of Lysis Mix and mixed thoroughly with 500 µl of 96% ethanol. The mixture was applied to a Binding column, centrifuged (16 000 x g, 30 s) and the flow-through liquid decanted. The column with bound RNA was washed once with 700 µl of 96% ethanol and twice with 500 µl of Ethanol-diluted Wash Solution; the column was centrifuged (16 000 x g, 30 s) and the flow-through liquid decanted after each washing step. The column was dried by centrifugation (16 000 x g, 1 min) and the bound RNA was eluted with nuclease-free water. The concentration of isolated RNA was measured on a NanoDrop. The isolated RNA was stored at -20°C until use.

In case of pre-let-7a3, the isolated RNA was phosphorylated at the 5’end. Depending on the pre-let-7a3 concentration, the amount of added T4 Polynucleotide Kinase (Thermo Scientific) was adapted according to the formula - 10 U of T4 Polynucleotide Kinase per 50 pmol pre-let-7a3. The reaction was incubated at 37°C for 30 min. The RNA was isolated using the mirPremier microRNA Isolation Kit as described previously.

3.6.3 Annealing

In case of pre-let-7a3, 0.11 volume of 10x NEBuffer 2, serving as an annealing buffer, was added to the volume of the sample with isolated pre-let-7a3. The annealing reaction was incubated at 95°C for 3 min, quickly placed on ice and stored at -20°C until use.

In case of dsRNA and dsRNAtetra substrates, equimolar amounts of both strands were mixed and 0.11 volume of 10x NEBuffer 2 was added. According to the published protocol (Zhang et al, 2002), both substrates were incubated at 95°C for 3 min and then placed to a water bath preheated to 75°C, which was left to cool down gradually to room temperature lasting 3 to 4 hours. In case of the dsRNAtetra substrate, when the temperature dropped to 50°C, $MgCl_2$ was added to a final concentration of 2 mM in order to help the proper folding of the 4nt loops. The annealed RNA was stored at -20°C until use.

3.7 Dicer cleavage assay

Unless stated otherwise, cleavage reactions (V= 70 µl) contained 50 nM $Dicer^O$ or $Dicer^S$, 70 U Ribolock RI (Thermo Scientific) and Dicer assay buffer (30 mM Tris-HCl [pH 8], 50 mM NaCl, 3 mM $MgCl_2$, 0.25% (v/v) Triton X-100, 15% (v/v) glycerol) (Podolska et al, 2014). The used substrate concentrations are indicated in the legends of corresponding figures. Reactions with inhibited Dicer contained 10 mM EDTA. The cleavage reactions were incubated at 37°C for 96 h (unless stated otherwise). At time points 0, 2, 6, 10, 24 and 96 h, 10 µl were collected from each cleavage reaction. The collected samples were mixed with 1.2 volume of RNA loading buffer (95% (v/v) formamide, 18 mM EDTA, 0.025% (w/v) SDS, 0.1% (w/v) xylene cyanol, 0.1% (w/v) bromophenol blue) to stop Dicer cleavage as published previously (Ma et al, 2008) and stored at -20°C. When it was necessary for lack of substrate, the cleavage reactions were scaled down proportionally.

3.8 Denaturing RNA electrophoresis

Samples from Dicer cleavage assays were denaturated for at 95°C 5 min. Following pre-run (1-2 h, 200-500V) with preheated TBE buffer (100 mM Tris, 100 mM H_3BO_3, 2mM EDTA), 10 µl of denatured samples were loaded onto preheated 15% polyacrylamide gel (with 7M Urea) and resolved using 250-550 V, while the buffer was continually heated by a manual heat exchanger. The temperature of the gel dropped during sample loading and then gradually rose to 40-55°C. Gels were rinsed in water and then incubated with 30 ml 1x SYBR Gold Nucleic Acid Gel Stain (diluted in water, Thermo Scientific) rocking at room temperature for 10-20 min. Stained RNA and DNA were visualized with transiluminator; the exposure ranged from 400 ms to 3 s.

4 Results

4.1 Dicer purification

Dicer preparation of high purity and integrity is fundamental initial step for a biochemical study focusing on Dicer activity and its substrate processing. The purity, the absence of other proteins in a sample, is not only important for the accurate measurement of protein concentration, but also for the validity and reproducibility of obtained data. The presence of other proteins might influence the Dicer activity, stability or the integrity of RNA substrates, which are prone to degradation by ubiquitous RNases. The integrity of a protein denotes presence of a desired protein in a sample in its intact form and without any degradation products. Absence of degradation fragments is of particular importance in the case of mammalian Dicer because its fragments lacking the helicase are potentially more active than the full-length Dicer. As a consequence, they could disguise the activity of intact Dicer in a cleavage assay.

To obtain pure and intact mouse Dicer isoforms, several purification protocols based on affinity chromatography were tested. I started with a previously published single-step purification using a TALON resin (Zhang et al, 2002). Nevertheless, the purity of the Dicer preparation was low, so I tested several two-step purification protocols with various combinations of affinity resins, which are described in this section.

For each purification protocol, a new Dicer construct was prepared with different combination of affinity tags and protease cleavage sites, which enabled to remove the affinity tags during or after the purification. Therefore, descriptions of a purification protocol are accompanied with a schematic representation of the prepared construct. A subset of used affinity tags, e.g. a His-tag, EGFP, GST, were used for purification; the rest of affinity tags, e.g. a Myc-tag, an HA-tag, were introduced into the protein construct in order to enable its detection on Western blot with commercially available validated antibodies when testing purification protocols. Moreover, the arrangement of affinity tags on both termini of Dicer enabled to obtain information about recombinant protein degradation during the purification process.

4.1.1 Single-step Dicer Purification using TALON resin

To establish a benchmark method, to which other purification protocols could be compared, and to produce sufficient amount of $Dicer^{S}$ and $Dicer^{O}$ for ongoing projects in the lab, the previously published single-step Dicer purification protocol (Zhang et al, 2002) was adopted with minor adjustments (described in detail in Material and Methods). It is based on immobilized metal affinity chromatography (IMAC) with a TALON resin, which binds an amino acid motif of six or more histidine residues called a His-tag. Recombinant Dicer isoforms produced throughout the study were fused with His-tags with a variable number of histidine residues as indicated in abbreviations.

Using a baculovirus expression system, pFastBac vectors were designed, constructed and used for bacmid generation, leading to preparation of corresponding viral supernatants. 50 ml of the viral supernatants encoding Dicer isoforms fused with a 6xHis-tag and a Myc-tag at their N-terminus (Fig. 5) were used for infection of 5 x 10^8 Sf9 insect cells. After 48 hours, the cells were lysed and the lysate was incubated with 15ml TALON resin, which bound recombinant Dicer isoforms by their 6xHis-tag. Following washing, elution with 200 mM imidazole and dialysis, I obtained approximately 40 mg and 31 mg of purified $Dicer^S$ and $Dicer^O$, respectively. Samples collected during the purification were analysed by SDS-PAGE.

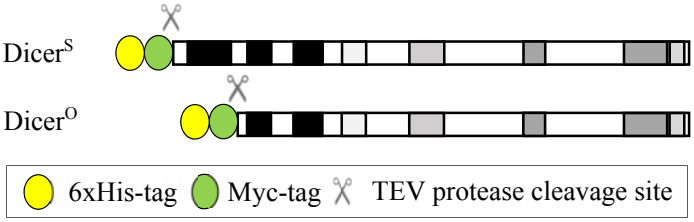

Fig. 5: Dicer isoforms purified with the TALON resin. The tags and TEV cleavage sites are not to scale with the Dicer protein.

As shown in Fig. 6, the purity or integrity of $Dicer^S$ preparation was very low. Estimating from the intensity of a $Dicer^S$ band on a Coomassie-stained gel, intact recombinant $Dicer^S$ comprised less than 20% of the total protein, although the exact proportion could be influenced by Dicer's habit to enter the gel inconsistently resulting in its gradual release from the well forming a smear above the Dicer band (Michael Doyle, personal communication). The remaining fraction of the preparation consisted of proteins of various sizes. They could be contaminants from insect cells or products of Dicer degradation occurring before, during or after purification. The subsequent Western blot analysis confirmed the presence of shorter Dicer fragments in the preparation (data not shown), which did not exclude a

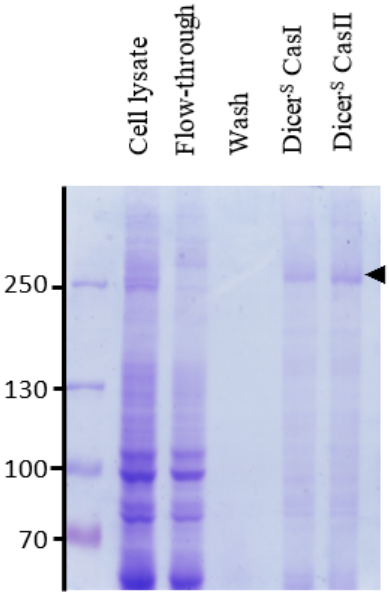

Fig. 6: Analysis of single-step $Dicer^S$ purification using TALON resin. 5 µl of samples collected during the purification were loaded onto gel. The loaded volume corresponded to ~55 µg of cell lysate and 8 µg of purified $Dicer^S$ after dialysis (labeled as $Dicer^S$ CasI and $Dicer^S$ CasII according to two dialysis cassettes comprising different eluate fractions). The gel was stained with Coomassie Brilliant Blue. The full-length $Dicer^S$, indicated by the black arrowhead, migrated slightly slower than 250kDa band of the marker despite its predicted size of 221 kDa. Nevertheless, it was consistent with experiments performed in the lab in the past.

presence of cellular contaminants in the preparation. A similar result was achieved with Dicer^{O} (data not shown).

I did not test the presence of cellular contaminants by any direct methods, e.g. mass spectrometry. Nevertheless, I confirmed it indirectly in the last purification protocol (Section 4.1.4), where I intentionally used a smaller bed volume of the resin for lysate from a comparable number of cells to increase purity of Dicer preparation. IMAC resins in general bind not only proteins fused with a His-tag, but also undesirable proteins naturally containing surface motifs suitable for IMAC resin binding (reviewed in Block et al, 2009). However, such proteins generally bind an IMAC resin with lower affinity and might be, therefore, outcompeted by the desired protein fused with a His-tag. If there is less of the desired protein than is the capacity of the resin, then there are more IMAC functional groups vacant and available for binding of such contaminants, which leads to decreased purity of purified protein. The simple decrease in amount of the TALON resin thus led to the increased purity of the eluate.

The Dicer isoforms purified with the original strategy were used for optimization in a validation screen for selected potential inhibitors and activators of Dicer and the miRNA pathway in general (data not shown). However, the purity and integrity of Dicer preparations was considered suboptimal for the desired aim of my thesis. Therefore, I added another affinity resin into the purification protocol. If the degradation would occur before Dicer purification or while Dicer is bound to the first resin, the successive use of two resins, which bind two different affinity tags fused either with the N- or the C-terminus of Dicer, should ensure that only intact Dicer with both termini should be present in Dicer preparation. In addition, combination of two distinct resins would yield an increased purity of the preparation as a considerably smaller subset of potential contaminants should bind to both resins in a non-specific manner.

4.1.2 Two-step Dicer purification using Ni Sepharose and Glutathione Agarose

I started with the combination of an N-terminal 6xHis-tag and C-terminal GST-tag. The combination was used previously for purifying Rec14, a protein of 32.9 kDa from *Saccharomyces pombe* (Maity et al, 2013), although positions of tags on the recombinant protein and the order of affinity resins differed from the protocol described here.

The pFastBac vector carrying the coding region for Dicer^{S} fused with 6xHis-tag and Myc-tag at the N-terminus and three adjacent FLAG-tags and GST-tag at the C-terminus was produced by my colleague Radek Malík. The affinity

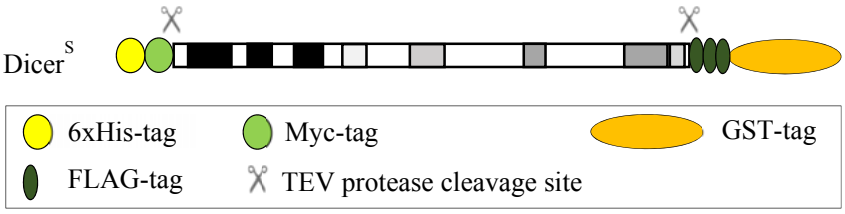

Fig. 7: Dicer^{S} used for purification with combination of the Ni Sepharose and the Glutathione Agarose resins. Depicted purification tags and TEV cleavage sites are not to scale.

tags were separated from Dicer by TEV protease cleavage sites and thus designed as potentially removable (Fig. 7). After generating the corresponding viral supernatant, I tested a small scale purification of Dicer^S from 2.5 x 10^7 insect cells with TALON resin and Pierce Glutathione Agarose and a combination of batch format and gravity-flow column chromatography. Despite the relatively high protein concentration in the eluate from the Glutathione Agarose resin (0.7µg/µl), it was difficult to detect DicerS even with Western blot (data not shown). The outcome did not improve even when a Glutathione Agarose resin from a different manufacturer (Macherey-Nagel, kindly provided by Václav Urban) was used. Therefore, I increased the number of insect cells expressing recombinant Dicer^S and tested fast protein liquid chromatography (FPLC) with Pierce Glutathione Chromatography Cartridge.

Due to accidental loss of approximately three quarters of the lysate, the recombinant protein from ~6 x 10^7 Sf9 cells was purified using FPLC with a HisTrap High Performance (HP) column (prepacked with Ni Sepharose, 1 ml, GE Healthcare Life Sciences). Following elution with increasing concentration of imidazole, the eluate was collected into 24 0.5ml fractions. Based on the protein concentration measured by the UV light absorption at 280 nm in the FPLC machine, the eluate fractions 5 to 13 were pooled, desalted and applied to Pierce Glutathione Chromatography Cartridge (prepacked with Glutathione Agarose beads, 1ml, Thermo Scientific). Recombinant Dicer^S was eluted with steady concentration of reduced glutathione.

I obtained ~700 µg of purified Dicer^S protein using the above mentioned purification protocol. This corresponded to ~15% of the yield obtained from the one-step purification using only the TALON resin, when recalculated per 1 x 10^8 cells. There are several explanations for the reduced protein yield. Apart from unequal Dicer expression in infected cells and increased Dicer loss during two-step purification, the reduced protein yield might also indicate increased purity of the preparation due to decrease in cellular contaminants or Dicer degradation fragments.

The purity of the Dicer^S preparation remained unknown as I did not perform SDS-PAGE followed by Coomassie staining of the elution fractions from the second column due to former poor experiences with detection of Dicer even by the Western blot.

Surprisingly, the two-step purification using His-tag and GST-tag did not yield an intact Dicer^S (Fig. 8). Although the detection with a Dicer-specific antibody indicated absence of a significant amount of degradation fragments, the detection with a FLAG-specific antibody (Fig. 9C) showed the opposite; the shorter fragments of Dicer^S were present in the sample of the purified protein. The observed discrepancy between two antibodies can be attributed to distinct positions of their epitopes on Dicer. Dicer-specific antibody recognizes amino acid sequence localized between the DUF283 and the PAZ domain. Therefore, the shortest possible fragment with intact N-terminus recognizable by the Dicer-specific antibody would have 88 kDa, the fragment with C-terminus would have 163 kDa (the theoretical molecular weights were counted using ExPASy Compute pI/Mw Tool and do not take into account the above mentioned discrepancy in Dicer size when compared with PageRuler Plus Prestained Protein

Ladder). As the fragments detected with FLAG-specific antibody were smaller than 163 kDa, they lacked the epitope necessary for detection with the Dicer-specific antibody.

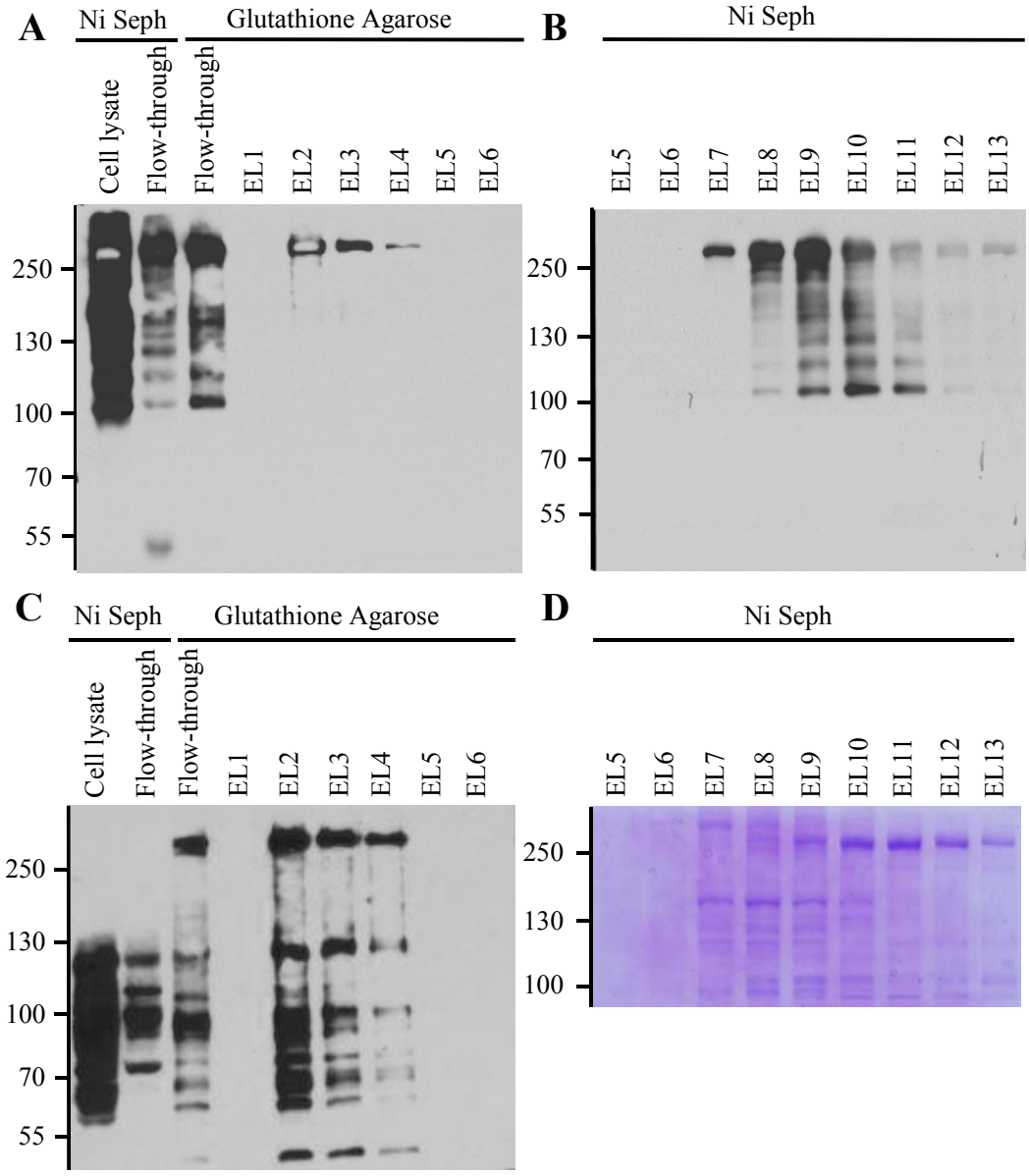

Fig. 8: The analysis of two-step purification with Ni Sepharose and Glutathione Agarose. A-C Western blot analysis and **D** SDS-PAGE followed by Coomassie staining was performed with samples collected from individual purification steps. 50 µg of cell lysate and flow-through and 20 µl of elution fractions from Glutathione Agarose were loaded onto gels. In case of selected eluate fractions from Ni Sepharose, 8 µl and 16 µl were loaded for Western blot and Coomassie staining, respectively. The proteins were resolved on 7% polyacrylamide gels. A Dicer-specific antibody was used for detection of Dicer^{S} in the panels **A** and **B**. The membrane from the panel **A** was subjected to HRP inhibition by 1% sodium azide and Dicer^{S} was detected with a FLAG-specific antibody in panel **C**. Note the distinct pattern of bands in panels **A** and **C**. EL-Eluate fraction, Ni Seph-Ni Sepharose

The presence of shorter Dicer fragments in the Dicer preparation might have two possible explanations, which are not mutually exclusive and are compatible with the Western blot results. First,

the dimerization property of GST-tag fused with Dicer^S might result in pull-down of degraded Dicer fragments when they form a heterodimer with an intact Dicer. GST forms homodimers with very low Kd (<<1nM), so that there are hardly any monomeric forms found in µM concentration (Fabrini et al, 2009). Moreover, the GST-tag dimerization is imposed on the protein with which it is fused (Maru et al, 1996). When the cell lysate contains degradation fragments of the desired protein or some degradation occurs during the purification, the dimers might be formed between the intact protein and degradation fragments lacking the terminus without GST-tag. In that case, the two-step purification is ineffective, because in the first step it enables to pull down intact protein together with its dimerizing partner lacking the N-terminus.

Second, Dicer degradation might have occurred not only before or during cell lysis, but also during the course of purification. The sample could be contaminated with an unknown protease with several cleavage sites within Dicer that could cleave Dicer during the entire purification process. Nevertheless, the set-up would have enabled that only degradation fragments with the intact C-terminus generated during or after the elution from the first column would have passed the second purification step. Degradation fragments with the N-terminus would be present in the Dicer preparation only in case that the protease would cleave during or after elution from the second resin.

In any case, I decided to avoid the problem with the strong dimerization tendency of GST by its replacement by EGFP in the construct. EGFP differs from wild type Green Fluorescent Protein (GFP) in several mutations enhancing its fluorescence intensity (Cormack et al, 1996; Thastrup et al, 2001). Notably, the wildtype GFP forms dimers in concentration above 5 mg/ml (Chalfie, 1995), which indicates that EGFP might dimerize as well.

4.1.3 Two-step Dicer purification using TALON resin and GFP-Trap

The pFastBac vector containing Dicer^S fused with 6xHis-tag and Myc-tag at its N-terminus and EGFP at its C-terminus was generated by Jana Urbanová and Radek Malík. Both affinity tags were removable by TEV protease cleavage (Fig. 9), which was particularly important, as I intended to avoid the elution from the second resin, the GFP-Trap. The GFP-Trap consists of a fragment of alpaca's antibody against GFP coupled with agarose. Once the EGFP-tagged protein is bound to the GFP-Trap, it can only be eluted under denaturing conditions of low pH (2.5), quickly followed by neutralization of the eluate. Such treatment might, however, result in loss of activity or the degradation of the protein of interest. To release the Dicer from the GFP-Trap, I intended to replace the elution step with an on-column TEV protease cleavage of both EGFP and N-

Fig. 9: Dicer^S purified with TALON resin and GFP-Trap. The tags and TEV cleavage sites are not to scale with Dicer.

terminal affinity tags. The short N-terminal tag might be removed from the sample of purified Dicer by adding into the purification protocol a second TALON resin, a size exclusion chromatography or a concentration through the Amicon Ultra Centrifugal Filter (Merck Millipore)

I performed the two-step purification from 5 x 10^7 cells expressing Dicer^S using 1 ml TALON resin (Clontech). The eluate was divided into 4 parts, which were successively applied to 40 µl GFP-trap (ChromoTek). The GFP-trap with bound EGFP-tagged Dicer was divided into halves and incubated with either 10 µg of commercially available TEV protease (Sigma) or 12 µg of in-house made TEV protease (kindly provided by Václav Urban) at 8°C overnight. The following day, the cleavage reactions were shortly spinned and supernatant, which should have contained Dicer, was separated from beads forming pellet.

As shown in Fig. 10, the on-column TEV protease cleavage seemed to be inefficient for both TEV proteases as no visible amount of full-length Dicer was released from the GFP-Trap to the supernatant samples. Unlike the commercial TEV protease, the in-house made TEV protease cleaved inside Dicer, which resulted in generation of several smaller Dicer fragments. This could be caused by contamination of the in-house made TEV protease with another protease. The subsequent experiments showed that the on-column cleavage with the TEV protease (Sigma) yielded full-length Dicer in supernatant, where it could be detected by Western blot stained with Dicer-specific antibody (data not shown). However, the efficiency of the on-column TEV protease cleavage was unacceptably low. In addition, a c-myc specific antibody also detected the released Dicer, which could be possible only due to inefficient removal of N-terminal affinity tags.

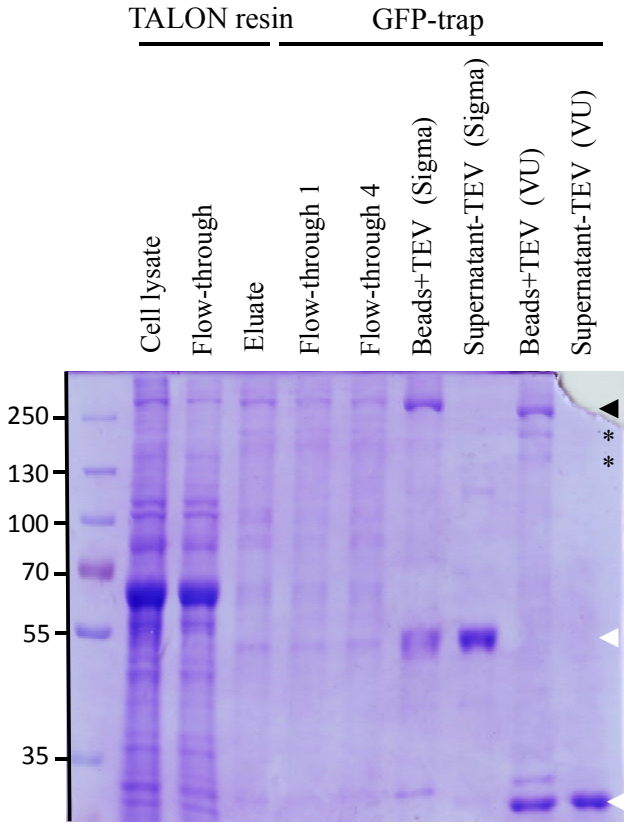

Fig. 10: Analysis of Dicer^S purification using TALON resin and GFP-Trap. SDS-PAGE followed with Coomassie staining was performed with samples collected from individual purification steps. 25 µl of reactions with TEV proteases and 13.6 µl of rest of the samples were loaded onto gel. It corresponded to Dicer^S from ~5 µl of beads and to ~55 µg of cell lysate. The proteins were resolved on 8% gel. The intact Dicer^S is marked by the black arrowhead and its degradation fragments by the black asterisks. TEV proteases are marked by the white arrowheads.

To avoid the inefficient on-column cleavage, I tested the elution with low pH mentioned above, but no detectable amount of Dicer was

recovered. Because of high purity and integrity of Dicer, I tried whether it would be possible to use the complex of Dicer with GFP-Trap for the cleavage assay. The impossibility to measure Dicer protein concentration when bound to GFP-Trap resin by standard methods would be circumvented, the amount of Dicer in the reaction should be proportional to the intensity of fluorescence emitted by EGFP, which could be used for Dicer quantification. I used an *in vitro* dicing fluorescent assay established in our lab. Briefly, Dicer is incubated with a 27bp duplex with 3' 2nt overhang on one end and a pair of fluorescent dye Cy5 and quencher Iowa Black-RQ on the other end (Podolska et al, 2014). The Dicer-cleaved substrate releases the fluorescent dye from the quencher, thus resulting in an increased fluorescence of the reaction mixture. As the emission maximum of EGFP (509 nm) differs from the emission maximum of Cy5 (665 nm), the assay can be performed with EGFP-tagged Dicer while EGFP can serve for calibration. Nevertheless, I encountered two problems. First, there was high variability among samples because it was difficult to add the same amount of beads accurately as they sedimented quickly. Second, there was high variability in the fluorescence even between multiple measurements from one sample, which could be caused by uneven resuspension of the particles with bound Dicer between individual measurements leading to uneven exposure of EGFP or varying degree of shadowing (data not shown). Therefore, the idea to use the complex of Dicer bound to the GFP-Trap for the cleavage assay was abandoned.

The purity of TEV protease (Sigma) treated Dicer^{S} bound to GFP-trap seemed to increase considerably when compared to the purified Dicer^{S} obtained by the single-step purification (Fig. 6), although the negative control in form of resin with bound Dicer^{S} was not included. To circumvent the problematic recovery of Dicer from the GFP-Trap while maintaining its high purity, the GFP-trap was replaced with a Strep-Tactin resin in the purification protocol.

4.1.4 Two-step purification using TALON resin and Strep-Tactin resin

The Strep-Tactin-Twin-Strep-tag system uses immobilized tetrameric Strep-Tactin, a specially engineered streptavidin, which binds to its biotin binding pocket a short peptide called Strep-tag II with high affinity (Schmidt et al, 1996; Voss & Skerra, 1997). The use of two tandemly arranged Strep-Tags II, called Twin-Strep-tag, increases affinity of the recombinant protein to Strep-Tactin (Schmidt et al, 2013). The Twin-Strep-tag is displaced from the resin by an analogue of biotin, desthiobiotin. Unlike biotin, desthiobiotin binds to Strep-Tactin reversibly, thus enabling resin regeneration.

Following generation of expression vectors encoding Dicer^{S} and Dicer^{O} fused with Twin-Strep-tag and 8xHis-tag to the N-terminus and the C-terminus, respectively (Fig. 11), I performed two-step purification. The lysate from 5 x 10^{8} insect cells expressing the corresponding Dicer isoform was mixed with 500 µl of the TALON resin. The eluate was incubated with 500 µl of the Strep-Tactin resin. After addition of desthiobiotin, the eluate fractions were collected separately to take samples for downstream analysis. To obtain the highest possible protein yield, the eluate fractions were pooled during the concentration and buffer exchange step.

I obtained ~300 µg and ~100 µg of Dicer^S and Dicer^O, respectively. The smaller yield of Dicer^O could be caused by its lower expression in the infected cells when compared to Dicer^S. Moreover, a considerable amount of Dicer^O did not bind to the TALON resin as judged from the amount of Dicer^O in the flow-through sample from the TALON resin (Fig. 12). Although Dicer isoforms differed only in the N-terminus and the 8xHis-tag was fused with the C-terminus, it cannot be ruled out at the moment that the absence of the N-terminal HEL1 domain might have imposed less frequent conformation on Dicer, which obstructed the C-terminus, hence caused a lower yield of Dicer^O.

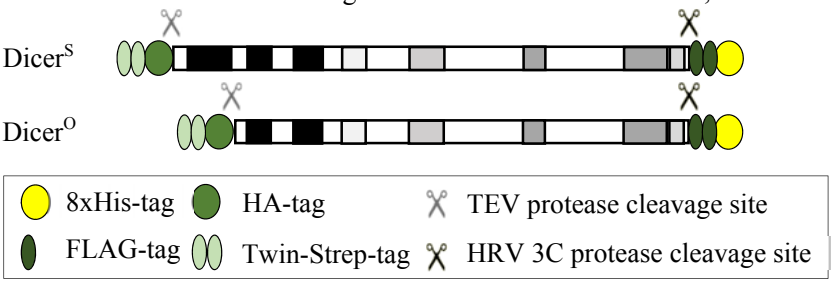

Fig. 11: Dicer isoforms purified with the TALON resin and the Strep-Tactin resin. The affinity tags and protease cleavage sites are not to scale with Dicer.

I obtained a protein of high purity as the majority of impurities in eluate from the TALON resin apparently did not bind to Strep-Tactin and were, therefore, removed in the second purification step (Fig. 13). In any case, the sample of purified Dicer contained a tiny amount of a protein around ~70 kDa and small amounts of proteins of size ranging from 180 kDa to 240 kDa. The identity of the smaller protein remained unknown, but the longer fragments contained FLAG-tag as evidenced by Western blot analysis (data not shown). Therefore, they originated from the Dicer construct. Their smaller size indicated that they lacked the N-terminus. Such Dicer fragments should not bind to the Strep-Tactin resin and should have been lost during purification. Therefore, it seemed that a small but consistent degradation occurred during or after elution from the second resin. I excluded the influence of inappropriate storage as similar results were obtained for Dicer preparation stored at 8 °C or undergoing one freeze-thaw cycle.

The purity of the Dicer preparation seemed to be similar regardless of the order of TALON and Strep-Tactin resin in the purification protocol (data not shown). Nevertheless, the resin order had a profound effect on the purity of the eluate from the first resin. In case of Strep-Tactin as the first resin, the eluate contained very little or no impurities at all and the estimated purity of the samples was around 90%, whereas the estimated purity of Dicer in the eluate from TALON as a first resin ranged from 50 to 60 %. However, due to smaller amount of protein loaded onto gel for eluate from the first Strep-Tactin resin when compared to its counterpart in Fig. 13, it would be difficult to recognize the impurities against the background, even if they were present.

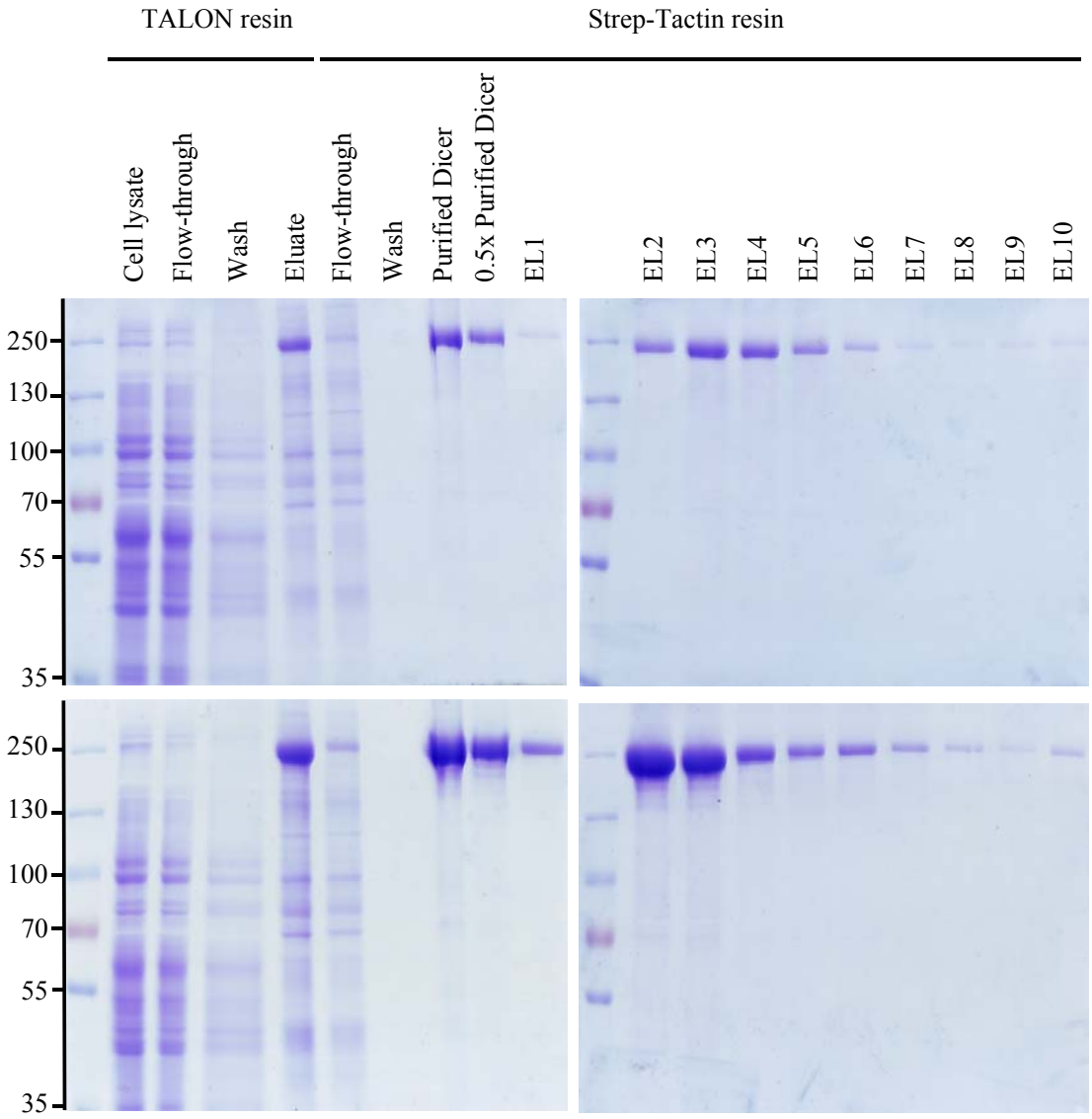

Fig. 12: Analysis of Dicer^{O} and Dicer^{S} purification using TALON and Strep-Tactin resins. The upper panels and lower panels show samples from purification of Dicer^{O} and Dicer^{S}, respectively. SDS-PAGE analysis was performed with samples collected from individual purification steps. 50 µg of the cell lysate and the flow-through from the TALON resin and 20.8 µl of the rest of the samples were loaded onto gel. The sample labelled 0.5x purified Dicer represented 2x diluted purified Dicer sample. Proteins were resolved on 7.5% gels and stained with Coomassie Brilliant Blue. Note the distinct appearance of the bands corresponding to intact Dicer of purified Dicer when compared to EL2, which is a consequence of higher percentage of glycerol in the sample. EL-Eluate fraction

As it seemed that the purity of Dicer was independent on the order of the resins, I decided to use the purification protocol in which the TALON resin preceded the Strep-Tactin resin for several reasons. First, the TALON resin is prone to nonspecific protein binding when its binding capacity is not reached (reviewed in Block et al, 2009). As the protein concentration decreases after the first resin, the TALON resin might be even less selective when used as the second resin. Second, the lysate might contain traces of biotin from the complete TNM-FH medium, where biotin is present in a nanomolar

concentration. As the Strep-Tactin resin binds to biotin irreversibly, it might be advisable to use Strep-Tactin as the second resin to prevent the permanent decrease in Strep-Tactin binding capacity.

The purified Dicer isoforms from Fig.13 were used for the cleavage assays (sections 4.2.1-4.2.3). I tested the affinity tags removal by TEV and human *rhinovirus* 3C (HRV 3C) protease cleavage several times on several different Dicer constructs and the obtained results were inconsistent ranging from complete tag removal to apparently no removal at all (data not shown). The observed difference in the efficiency of proteases cleavage could be caused by different sequence contexts adjacent to their cleavage sites, because I used different Dicer constructs. Moreover, the Dicer constructs differed in their purity and some impurities could prevent the TEV protease adhesion to the surface of a microtube similarly to Bovine Serum Albumin in the restriction reaction, thus resulting in the different cleavage efficiency. In any case, I omitted the affinity tags removal at all, so the Dicer isoforms used for the cleavage assay contained all affinity tags as shown in Fig. 11.

4.1.5 Dicer concentration comparison

Estimating intact Dicer concentration in the samples with purified Dicer isoforms was a necessary prerequisite for obtaining a reproducible data from the following cleavage assays. However, measurement of total protein concentration might be inaccurate in determining concentration of intact Dicer isoforms for at least two reasons. First, purified Dicer isoforms might differ from each other in the proportion of the degradation fragments or impurities, so the total concentration would not be proportional to exact concentration. Second, different methods of total protein concentration measurement depend on or are influenced by the sequence of the protein and thus might not be optimal for comparison of concentration of two proteins with different sequences. Therefore, I determined the ratio between Dicer isoforms with unequal concentrations, when the amounts of Dicer isoforms equal, based on a gel stained with Coomassie Brilliant Blue as well as by Western blot (Fig. 13). Finally, I chose a ratio of 3x diluted $Dicer^S$: $Dicer^O$ 5 : 6, which amounts to a ratio $Dicer^S$: $Dicer^O$ 5 : 18. Notably, the accuracy of this method is also limited, the actual ratio of $Dicer^S$ to $Dicer^O$ might deviate from the estimated one by 20%. Repeating of the experiment with the range of $Dicer^S$ to

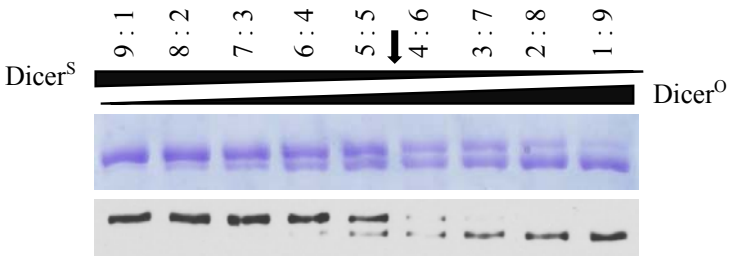

Fig 13: A comparison of concentrations between $Dicer^S$ and $Dicer^O$. The gel stained with Coomassie Brilliant Blue (upper panel) and the developed film from Western blot (lower panel) show different volume ratios of purified 3x diluted $Dicer^S$ and $Dicer^O$ loaded onto gel, which are indicated as $Dicer^S$:$Dicer^O$ above images. Five times smaller amount of protein was loaded onto the gel used for Western blot. Dicer was detected with a FLAG specific antibody (Sigma, 1:5000). The arrow indicates approximate ratio I selected for following cleavage assay. The $Dicer^S$ dilution by factor three was chosen based on the comparison of the amount of purified $Dicer^S$ and $Dicer^O$ on the Coomassie-stained gel (Fig. 12).

Dicer^{O} ratios closer to the observed one could make the estimation more precise. Nevertheless, even 20 % deviation was small in comparison with measurement of total protein concentration in the sample of Dicer^{O} by absorbance at 280 nm with NanoDrop, where there was a 4-fold difference between minimal and maximal measured value when measuring one sample several times due to low concentration of the protein.

4.2 Dicer cleavage assay

To investigate the difference between Dicer^{S} and Dicer^{O}, I opted for a nonradioactive cleavage assay set-up. It enabled to observe Dicer products together with the processing intermediates similarly to a radioactive cleavage assay without the requirement to follow the extra security measures necessary for handling radioactive material. Furthermore, as it was not used before, I wanted to test its potential.

I performed several cleavage assays with various substrates, which could reveal differences in Dicer activity or the ability of internal cleavage. The cleavage reactions were resolved on 15% polyacrylamide-7M urea gels and gels were stained with SYBR Gold Nucleic Acid Gel Stain (Invitrogen). As shown in Fig. 14, the amount of a 21nt RNA visible on gel stained with SYBR Gold Nucleic Acid Gel Stain ranged from 2 to 4ng, which corresponds to ~300 and 600 fmol, respectively; the detection limit also depended on the length of exposure.

To compensate for the high detection limit, I extended the time of the cleavage assay to up to 96 hours. Nevertheless, the extension of the time might have resulted in accentuation of drawbacks of *in vitro* assays, which might have a negative impact on obtained data. The drawbacks involve protein destabilization, loss of enzymatic activity, substrate degradation as well as excessive evaporation. Although they might be present in biochemical assays as well, the short observation time might prevent them or minimize their influence so, that they can be ignored. Nevertheless, in case that difference in processing between two Dicer isoforms exists, even a cleavage assay performed under suboptimal conditions could provide some evidence.

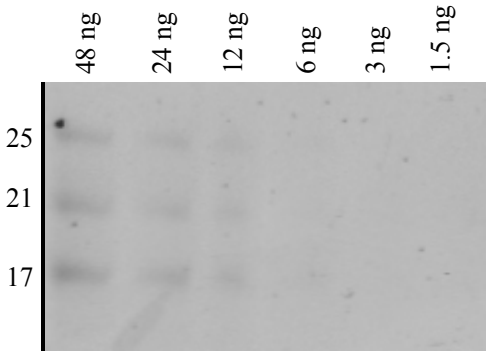

Fig. 14: Detection limit of staining with SYBR Gold Nucleic Acid Gel Stain. Indicated amount of miRNA Marker (New England Biolabs) was loaded onto 15% polyacrylamide-7M urea gel. The amount of 21nt ssRNA was counted as one third of the total amount of miRNA Marker/well.

4.2.1 Dicer processing of pre-let-7a3

To confirm that both purified Dicer isoforms had not lost the activity during purification and to optimize reaction conditions, I performed a cleavage assay with a miRNA precursor, a typical substrate

of mammalian Dicer. It was represented by human pre-let-7a3 (70nt), which had been used previously (Zhang et al, 2004).

4.2.1.1 Test of contamination by ubiquitous RNases

Due to high sensitivity of RNA to degradation by contaminating RNases, I tested the presence of nonspecific RNases in my cleavage reactions by performing the cleavage assay with pre-let-7a3 and $Dicer^S$ with and without Ribolock RI (Thermo Scientific), which is derived from the mammalian ribonuclease inhibitor protein.

The cleavage reactions contained 500 nM pre-let-7a3, which was incubated with 50 nM $Dicer^S$, so the pre-let-7a3 substrate was in 10-fold excess over $Dicer^S$. Notwithstanding long cleavage time, $Dicer^S$ cleaved only ~ 60 % of pre-let-7a3 in 24 hours. Therefore, I added a 96-hour time interval into the cleavage assay to see how Dicer isoform would be able to cleave the substrate within this time range, which is admittedly extremely long for a biochemical reaction.

The comparison of pre-let-7a3 cleavage pattern generated with and without Ribolock RI showed that the samples were contaminated with a nonspecific RNase (Fig. 15). The reaction with Ribolock RI generated products characteristic for Dicer cleavage: miRNA duplex comprising small 21-23nt RNA and the ~25nt fragment corresponding to the cleaved-off terminal loop. In contrast, cleavage products from the reaction without the Ribolock RI contained two additional cleavage fragments of around ~30 and ~35nt. The additional fragments were likely to arise from single or multiple cleavages inside the terminal loop of pre-let-7a3, suggesting that the nonspecific nuclease is a single-stranded RNA (ssRNA)-cleaving endoribonuclease. The specificity of contaminating RNase for ssRNA was further supported by the results of the same experiment with dsRNA substrate. The dsRNA substrate does not contain a stretch of ssRNA enabling a cleavage by ssRNA-cleaving nonspecific RNase, so it could be degraded only by dsRNA-cleaving nonspecific RNase. However, the cleavage pattern of dsRNA substrate with and without Ribolock RI did not differ (data not shown).

As shown in Fig. 16, I observed a higher Dicer cleavage rate in the presence of Ribolock RI than without it. There are two possible explanations. First, the ribonuclease inhibitor binds RNase A so that it prevented the access of substrate to its active site (Kobe & Deisenhofer, 1995). Therefore, it is plausible, that nonspecific RNase could not inhibit Dicer processing by competitive binding to pre-let-7a3. Second, the nonspecific cleavage of pre-let-7a3 in the terminal loop region resulted in generation of pseudo dsRNA substrate, which might have been more prone to form a non-productive conformation with Dicer than intact pre-let-7a3. It was improbable that Ribolock RI would directly stimulate Dicer cleavage rate in general as I did not see the stimulation for other substrates (data not shown). Nevertheless, the possibility of specific stimulation Dicer cleavage of pre-let-7a3 by Ribolock RI could not be ruled out.

Despite many precautions against RNase contamination, some appeared in the cleavage reactions, but their source remained elusive. Interestingly, the performed cleavage assays with pre-let-

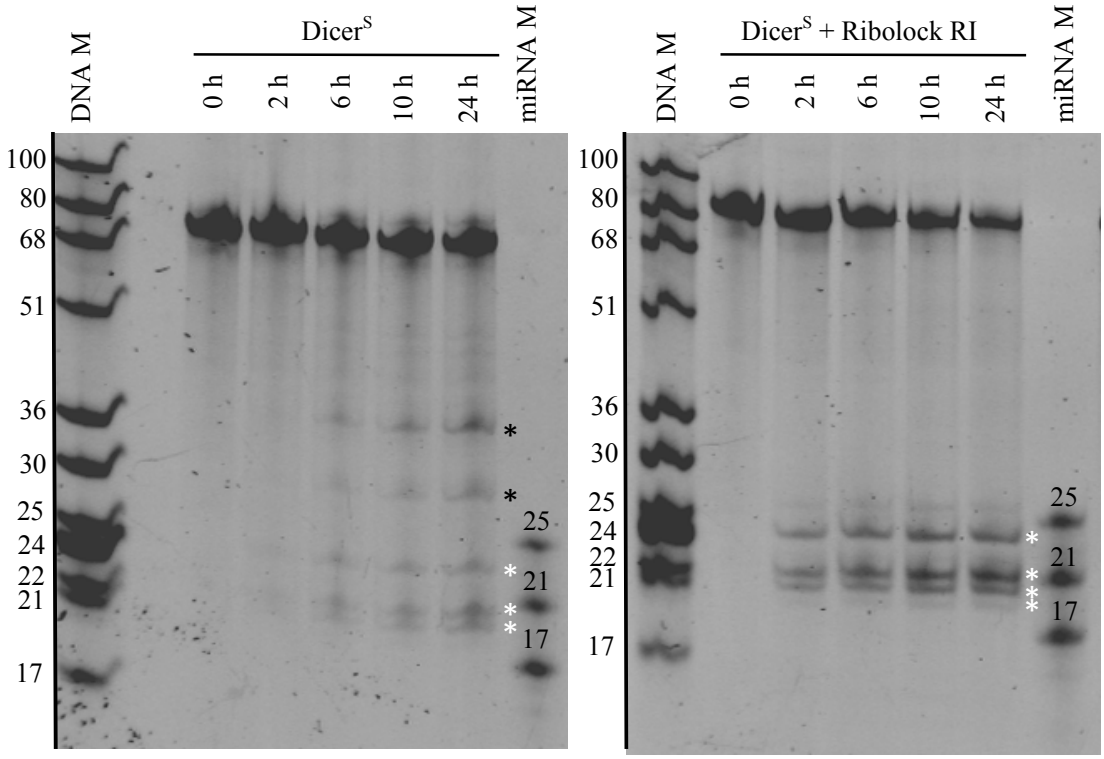

Fig. 15: Test of the presence of a nonspecific RNase. The cleavage assay with 50 nM Dicer^S and 500 nM pre-let-7a3 was performed under two different conditions: with and without Ribolock RI, inhibitor of RNase A. Only in this assay, I used the pre-let-7a3 substrate with 5'end triphosphate; the subsequent assays were performed with the pre-let-7a3 with monophosphate at the 5'end. The Dicer cleavage products are indicated with white asterisks, the cleavage products of nonspecific Ribonucleases are indicated with black asterisks. DNA and miRNA Marker served only as an approximate indicator of small RNA length; DNA marker tended to migrate slightly faster than RNA and miRNA Marker tended to migrate slightly slower than the rest of RNA due to different buffer composition. Furthermore, the migration of markers differed from gel to gel presumably because of varying temperature of gels. DNA M-DNA marker, miRNA M-miRNA Marker

7a3 substrate differed from each other in the rate of cleavage by nonspecific RNase in the reactions with Dicer isoforms and negative controls without Dicers and Ribolock RI. Thus, the substrate might be the source of nonspecific RNase, as each preparation might have differed in the amount of nonspecific RNase. However, the cleavage assay with the lowest possible cleavage rate by nonspecific protease as shown by negative controls had contaminated cleavage reaction with Dicer without Ribolock RI, which would indicate, that the Dicer preparation was the source. Nevertheless, it is contradictory with the presence of nonspecific RNase in the negative controls without Dicer. As Ribolock RI had a positive effect on nonspecific RNase inhibition, I used it routinely in the cleavage assays.

4.2.1.2 Test of Dicer-specific cleavage

I tested whether pre-miRNA cleavage in the presence of Ribolock RI is specific to Dicer by addition of excess of divalent cations chelating agent, EDTA, which inhibits Dicer cleavage (Provost et al, 2002; Zhang et al, 2002). Although the divalent cations are required for cleavage of many

ribonucleases, the combined characteristic of EDTA-inhibition together with generation of characteristic ~22nt products constitute a solid evidence for the Dicer specific cleavage. To verify that the ~22nt products are indeed produced by Dicer cleavage, it would be best to generate catalytically dead mutants. However, I did not prepare it and, to my knowledge, there is no commercially available small compound inhibitor specifically inhibiting Dicer cleavage activity, which could be used instead.

As shown in Fig. 16, the addition of EDTA decreased production of characteristic 21 - 22nt small RNA fragments bud did not completely prevent it. It was probably due to nonspecific RNase

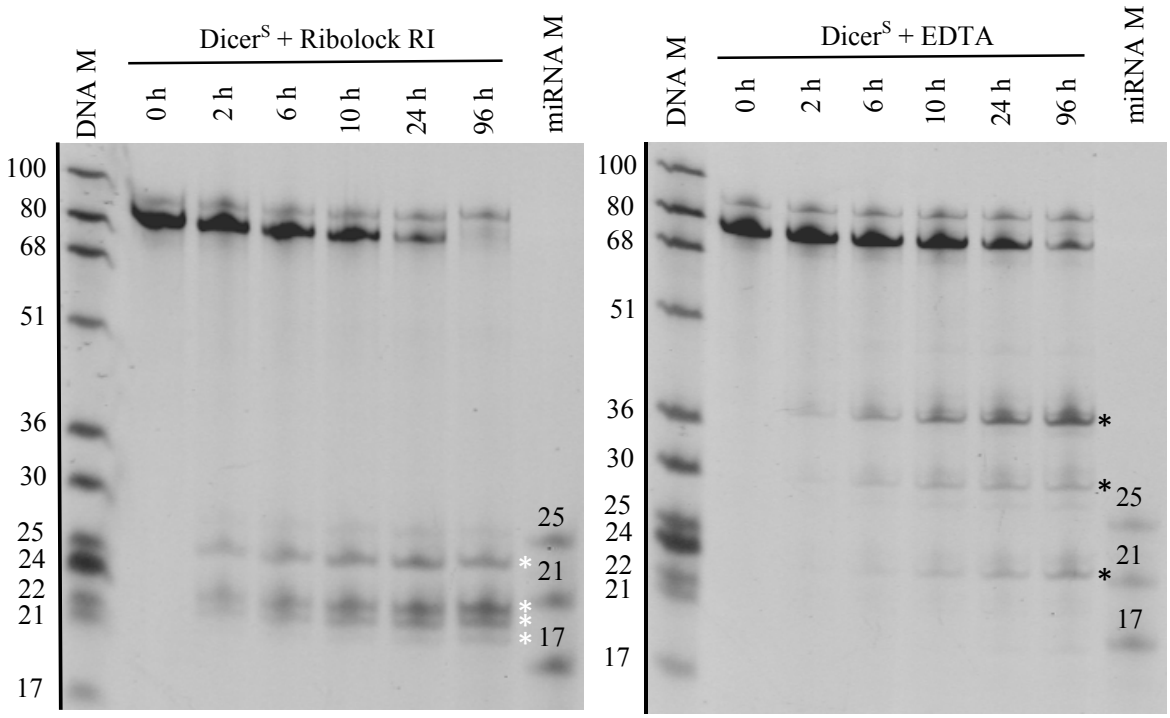

Fig. 16 : Test of specificity of Dicer cleavage. The cleavage assay with 50 nM Dicer^S and 500 nM pre-let-7a3 was performed with Ribolock RI or EDTA. The Dicer cleavage products are indicated with white asterisks, the cleavage products of nonspecific RNase are indicated with black asterisks. DNA M-DNA marker, miRNA M-miRNA Marker

cleavage in the sample without Ribolock RI. As is shown in Fig.17, there is a possible combination which might generate 21nt cleavage fragments as well. The same experiment with dsRNA did not yield ~22nt cleavage product when treated with EDTA (data not shown).

The addition of the 96h time point showed that Dicer retained at least a part of its activity during several days at 37°C (Fig. 16). Nevertheless, such prolonged incubation led to Ribolock RI inactivation and excessive evaporation even from closed microtubes. As the evaporation was uneven across samples, I decided to take into consideration only the results from shorter time points. However, I included the 96h time point in the cleavage assay and also used it for RNA PAGE, because it was helpful for determining the position of Dicer products of ~22nt, which were weakly visible in case of other substrates.

4.2.1.3 Comparison of Dicer^{O} and Dicer^{S} pre-let-7a3 processing

To compare processing of the Dicer isoforms, I performed cleavage assay with the pre-let-7a substrate, which represents a typical substrate of mammalian Dicer in a cell. As in previous assays, the pre-let-7a3 substrate was in 10-fold excess over Dicer isoforms. As shown in Fig. 18, the Dicer isoforms did not differ significantly in their processing of pre-let-7a.

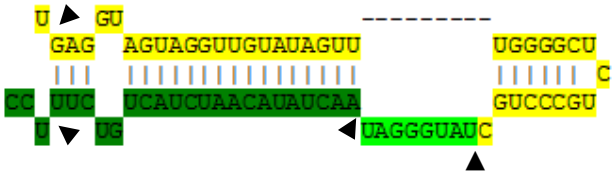

Fig. 17: Possible cleavage sites of nonspecific RNAse. At least two cleavage events by nonspecific RNase are necessary to generate fragments corresponding to the observed RNA species in the Fig, 16. Possible degradation fragments are highlighted in different colours: 38, 32 and 24nt fragments in yellow, green and dark green, respectively. If nonspecific RNAse cleaves off also the ssRNA overhangs, the generated products could be 37, 29 and 21nt long. Arrowheads indicate possible cleavage sites of nonspecific RNase. Secondary structure of pre-let-7a3 was adapted from the miRBase database.

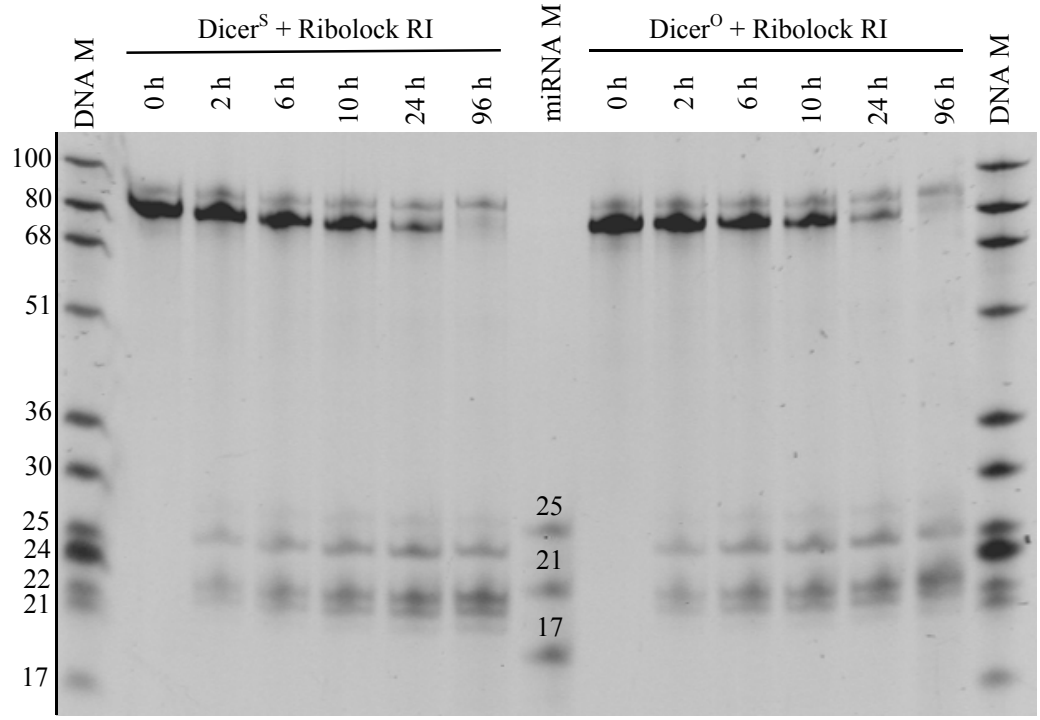

Fig. 18: Comparison of pre-let-7a processing between Dicer^{S} and Dicer^{O}. Reactions containing 50 nM Dicer^{S} or Dicer^{O} and 500 nM pre-let-7a were incubated at 37°C. DNA M-DNA Marker, miRNA M-miRNA Marker

4.2.2 Dicer processing of perfectly complementary dsRNA

Next, I tested Dicer processing of 130bp long, blunt ended, perfectly complementary dsRNA, which had been used previously (Zhang et al, 2002). It represents a type of RNAi substrate resembling a blunt-ended replicated RNA virus.

As shown in Fig. 19, both Dicer isoforms were inefficiently cleaving the 130bp perfect duplex. Although reactions contained equal molar amount of Dicer to substrate at the beginning, the length of the substrate would enable several cleavages and thus the multiple rounds of cleavage. As expected, Dicer^{S} apparently cleaved less than 20 % of the perfect duplex in 96 hours, which corresponded to less than one cleavage per one Dicer molecule. The cleavage intermediates of ~85nt were probably products of two rounds of Dicer cleavage.

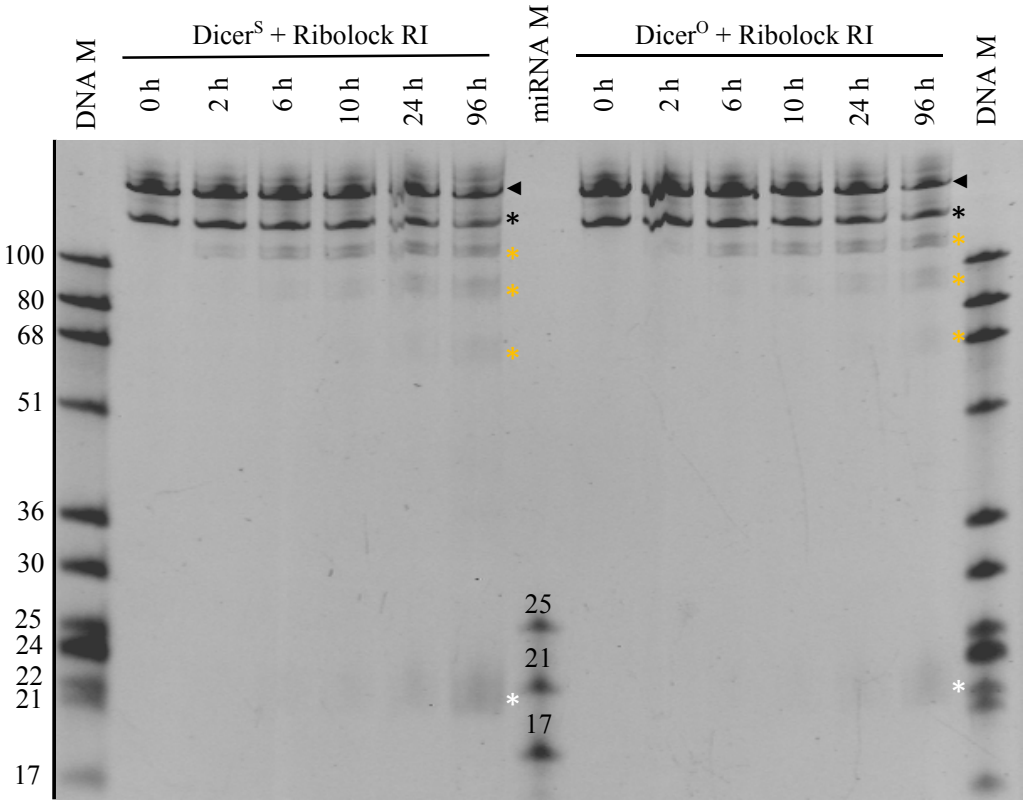

Fig. 19: Comparison of Dicer^{S} and Dicer^{O} processing of perfectly complementary dsRNA. Reactions contained 50 nM dsRNA and 50 nM Dicer^{S} or Dicer^{O}. Substrate in the form of dsRNA and ssRNA is marked by black arrowheads and black asterisks, respectively. Cleavage intermediates are denoted by orange asterisks and cleavage product by white asterisks. DNA M-DNA Marker, miRNA M-miRNA Marker

Interestingly, there was no significant difference in Dicer^{S} and Dicer^{O} processing. The subtle difference in Dicer processing in favour of Dicer^{S}, which seemed to be more active, has been reversed once I switched the order of the loading of cleavage reactions on the gels. This reversal indicated that, apart from other artefacts, the uneven staining or uneven diffusion of small RNA from the gel according to their position on the gel might play a role as well.

The substrate did not form only one band on the gel, which would correspond to 2 strands of the same size, but formed two bands instead. It indicated, despite the high urea content, pre-run and preheating of the TBE buffer to 50-55°C that the electrophoresis might not have been denaturing enough, especially for RNA samples above 100nt. Therefore, the upper band likely corresponded to dsRNA and the lower band to denatured strands.

4.2.3 Dicer processing of dsRNAtetra substrate

As the RNAi substrates usually lack open helical ends and initiation of their processing would require internal cleavage, I wanted to investigate how Dicer isoforms differ in the ability of making an internal cleavage. For this purpose, I used a derivative of a substrate lacking free ends (Zhang et al, 2002). It comprises two RNA strands. A 5'end of each strand basepairs with the internal region of the strand, thus forming a stem region with 4-nt terminal loop and long 60nt single stranded extension which pairs with the second RNA oligo. After annealing, the substrate has two tetranucleotide terminal loops instead of open helical ends (Fig. 20), which prevents substrate binding by the PAZ domain of Dicer.

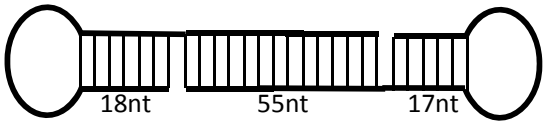

Fig. 20: A schematic representation of the dsRNAtetra substrate.

The efficiency of product generation was very low (Fig. 21), because the 22nt products were readily visible only after 96 hours. However, it is important to note that the formation of one 22nt products from dsRNAtetra required 2 cleavages. Therefore, the generation of products should have been less frequent than in case of previous dsRNA substrate where only one cleavage event was necessary for generation of ~22bp small RNA as it was cleaved from the terminus. Moreover, in the case of dsRNAtetra, it was plausible to expect that the first internal cleavage was rather random and did not formed intermediates of the same size. As a consequence, the detection method would have to be extra sensitive in order to detect so infrequent and random event. For future experiments, it might be advisable to increase the concentration of Dicers in the reaction in order to increase the frequency of product generation or to use more sensitive method (e.g. radiolabelled substrates) enabling visualization of small amount of products.

There seemed to be no difference in rates of ~22bp products generation by the Dicer isoforms, which implies that there was no difference in the ability of internal cleavage between $Dicer^{S}$ and $Dicer^{O}$. In the samples representing 24-hour time point, there was only a very faint band of the 22nt detectable in both samples and only in case of a very long exposure. The 96h time interval with detectable 22nt product was inadvisable to use for comparison as the samples differed in their concentration of the substrate and products due to uneven evaporation after such long time, which would strongly disturb reaction conditions.

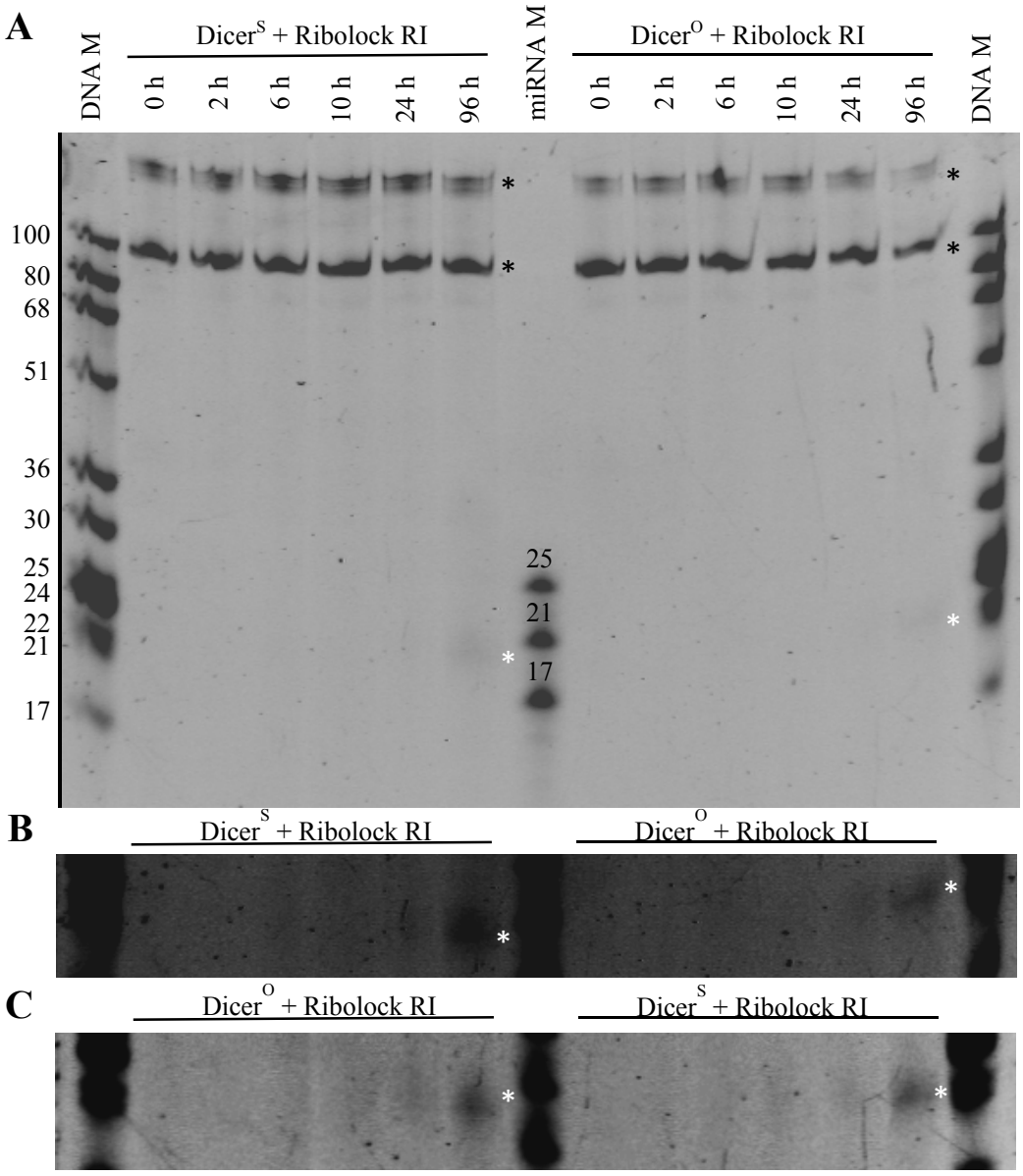

Fig. 21: Comparison of Dicer^{S} and Dicer^{O} processing of the dsRNAtetra substrate. Reactions contained 50 nM dsRNAtetra and 50 nM Dicer^{S} or Dicer^{O}. Substrate is marked by black asterisks and the cleavage product by white asterisks. **A** A picture of the entire gel with resolved cleavage reactions. **B** A section of the gel from **A** showing Dicer products. **C** A section of the gel from the second independent experiment showing Dicer products. Sections are from photograps with longer exposure than in **A** and adjusted brightness and contrast. DNA M-DNA Marker, miRNA M-miRNA Marker

5 Discussion

The functionality of RNAi in mouse oocytes coincides with the expression of Dicer^{O}, a truncated Dicer isoform. To better understand contribution of Dicer^{O} to functionality of RNAi in mouse oocytes, I intended to assess the difference in processing of RNAi substrate between Dicer^{O} and Dicer^{S} *in vitro*. Therefore, I established a protocol for isolation of pure recombinant Dicer and performed non-radioactive cleavage assays with substrates containing features characteristic of RNAi substrates, such as long stretch of perfectly complementary dsRNA or absence of blunt dsRNA ends.

Below, I will discuss the issues I encountered during testing different Dicer purification protocols and continue with the discussion of the last protocol, comprising TALON and Strep-Tactin resins, which yielded Dicer of the highest purity and integrity. The second part of discussion is devoted to the analysis of the non-radioactive cleavage assay and to the interpretation of the results.

5.1 Testing different Dicer purification protocols

In order to obtain Dicer of high purity and integrity, several different purification protocols were tested. The purity and integrity evaluation of some Dicer preparations was complicated by the low concentration of Dicer in eluates. Therefore, I was routinely forced to use Western blot, which has higher detection sensitivity. As a consequence, I was evaluating only Dicer integrity because the used antibody visualized only one specific protein and the rest of the proteins present in the sample remained hidden. Because of the expected higher activity of some Dicer degradation fragments and the potential risk they posed to the validity and reproducibility of obtained data, the estimating precisely Dicer integrity was more critical than purity. However, it should be noted that Western blot has certain limitations, which could distort the data and subsequent interpretation if used alone.

First, it is very sensitive, so it detects a minuscule amount of present proteins. The potential drawback of such sensitivity lays in the question, whether such amounts would be relevant. For example, I used Western blotting to assess whether I reached the binding capacity of the TALON resin. As I always detected Dicer in the flow-through sample, I scaled down the number of Dicer expressing cells because I assumed that the binding capacity of the resin had not been reached. Apart from an increase in the cost of the purification, such arrangement led to the above mentioned drawback of increased binding of nonspecific proteins, which however could not be detected by Western blot.

Second, the high sensitivity of the Western blot is coupled with a narrowed linear dynamic range of the detection signal. Therefore, the integrity of certain Dicer preparation might have looked lower because, unlike the signal from the degradation fragments, the signal from the intact Dicer might have already been saturated.

Third, the Western blot comprises of several steps, which are prone to generation of artefacts. Specifically, I encountered problems with Dicer transfer especially when I overloaded the gel with proteins. As smaller fragments in general tend to be transferred more efficiently, it might have added to

the underestimation of Dicer integrity, e.g. in case of Dicer preparation obtained by the two-step purification with TALON and Glutathione Agarose resin. Additional distortion of results can be caused by an uneven substrate distribution across membrane, which manifests as white patches with weak or no signal on developed films.

Taken together, it is advisable to combine the Western blot with the Coomassie staining or the silver staining if conditions allow it. If the Western blot alone is used for analysis, then it is necessary to be cautious when interpreting the obtained results.

Despite numerous obstacles emerging during work with various purification protocols, I established a functioning two-step purification protocol based on the TALON and the Strep-Tactin resins, which allowed for preparing Dicer at high purity and integrity. In fact, concerning the purity of Dicer, the described two-step purification is comparable to, or even outperforms, the widely used Dicer purification protocol comprising of IMAC combined with the size exclusion chromatography (MacRae et al, 2008) when comparing images of Coomassie-stained gels showing Dicer preparations from various studies (Liu et al, 2015; Ma et al, 2008; Ma et al, 2012; MacRae et al, 2008).

The described two-step purification protocol represents an accessible and affordable alternative to the previous purification protocol (MacRae et al, 2008). The size exclusion chromatography in general requires equipment of high acquisition costs that is not available in all laboratories. In contrast, the described two-step purification does not need any special equipment as it is performed in a batch or a gravity-flow format. Moreover, both TALON and Strep-Tactin resins can be reused multiple times when handled properly, which further reduces the cost of the purification.

To obtain ultrapure Dicer without affinity tags, however, it is necessary to optimize the two-step purification protocol further. Moreover, the outcome of the optimization might provide insights into the lower purity and integrity of Dicer preparations obtained by other protocols tested in the thesis.

First, it needs to be determined whether the degradation fragments are result of residual protease activity. Therefore, purification with a wide range of protease inhibitors present during the whole procedure needs to be tested as some could be found, which would inhibit the residual protease activity. Nevertheless, it is also possible that the putative protease responsible for Dicer degradation would not be inhibited by any of the available inhibitors. If the putative protease originates from insect cells, it would be useful to test different methods of cell lysis in order to prevent the protease release into the lysate.

Second, I would like to test whether Dicer with affinity tags forms monomers or dimers. Although Dicer with His-tag does not have a dimerization tendency (Zhang et al, 2004), there is a possibility that the other affinity tags or the composition of used buffers might promote its dimerization. Consequently, it would lead to unintentional copurification of intact Dicer together with its degraded dimerization partner as suggested for GST-tagged Dicer.

Third, an additional experiment should be performed to test the effect of high temperature on Dicer stability as Dicer might be degraded during the denaturation step preceding loading on the gel.

Although I tested the effect of denaturation temperature several times and saw no effect, I used either Dicer preparations of low integrity and purity or Dicer, which was bound to beads (data not shown). In the first case, the appearance of new bands or intensification of the existing ones due to additional degradation might not have been detectable due to low purity of Dicer. In the second case, the temperature seemed to correlate with the release of Dicer from the beads rather than with its degradation, however, this notion should be examined more systematically.

Fourth, the protocol for protease cleavage needs to be optimized so that the affinity tags could be removed from the recombinant Dicer. However, it might be advisable to construct a new expression vector containing cleavage sites of only one protease in order to simplify the purification protocol.

5.2 Non-radioactive cleavage assay

I established a non-radioactive cleavage assay which enables to directly visualize the cleavage products of Dicer processing on the gel using SYBR Gold Nucleic Acid Gel Stain. There are several advantages of the non-radioactive cleavage assay over the traditional radiolabelling. The samples from the non-radioactive cleavage assay do not need to be processed immediately but can be stored at -80 for a long time; the duration of the storage is limited only by the stability of the RNA. Moreover, the nucleic acid visualization and subsequent photographing is simple as it consumes only several minutes and does not require any gel drying or prolonged exposition. In contrast to radioactive nucleotides used in radioactive assays, SYBR Gold Nucleic Acid Gel Stain is not hazardous to health, its handling requires only standard safety measures and waste disposal regulations are minimal.

There are, however, two drawbacks of the developed non-radioactive assay: (1) the high detection limit, which is in the nanogram range, together with (2) the occurrence of some artefacts, which might distort the obtained data. The potential artefacts involve i) an effect of extensive and uneven evaporation caused by extended incubation of the samples at 37°C, which was used to accumulate sufficient amounts of cleavage products, and ii) an effect of the sample position on the gel, probably caused by uneven gel staining or uneven diffusion of small RNAs from the gel during its staining.

Using the non-radioactive cleavage assay, I confirmed that both recombinant Dicer isoforms isolated by two-step purification using TALON and Strep-Tactin resins were active. They generated the characteristic ~22nt small RNA species from various substrates. Moreover, both Dicer preparations retained their ability to process pre-miRNA in a multiple-turnover fashion. Under condition of ten-fold excess of substrate over the enzyme, both Dicer isoforms processed more than 50% of initial substrate in 24 hours. Nevertheless, the cleavage rate of both mouse Dicer isoforms was very low when compared with published results, where the full-length human Dicer, equivalent of Dicer^{S}, processed half of the initial amount of pre-miRNA in 15 minutes under the same Dicer to substrate ratio.

There could be several reasons behind such a striking difference in efficiency of Dicer processing. First, the two assays differed in the used substrate; I used human pre-let-7a3, whose dsRNA end has 1nt 5'end overhang and a 3nt 3'end overhang, whereas the published study utilized pre-let-7a1

modified so as to form a 3‘ 2nt overhang. Therefore, the substrates differ from each other in the terminal loop and the structure of their ends bound by the PAZ domain; difference in both structures could influence the Dicer activity (Feng et al, 2012; Park et al, 2011). Second, due to persisting issue with RNase contamination even during RNA purification, individual pre-let-7a3 preparations differed from each other in the presence of shorter RNA fragments. The fragments could imitate short dsRNAs with their inhibitory effects on the Dicer cleavage efficiency (Chakravarthy et al, 2010; Ma et al, 2008; Ma et al, 2012; Taylor et al, 2013). Although the reduced cleavage rate was most pronounced in the reactions without Ribolock RI, it is possible that even a mild contamination with degradation fragments in reactions with Ribolock RI could cause the discrepancy between the published data and my assay. Third, the assays differed in pH and composition of used buffer. I used a buffer optimized for the *in vitro* fluorescent Dicer assay described above (Podolska et al, 2014), which was with minor adjustments used in a study reporting the discovery of Dicer^{O} isoform (Flemr et al, 2013). The pH of the buffer was 8, which was by 1.5 units higher than in case of the buffer used in Ma et al, 2008. Although it represents a considerable difference, the test of buffer of a similar pH (6.8) decreased Dicer activity in the *in vitro* fluorescent assay, when compared to the used buffer (Podolska et al, 2014). Therefore, the higher pH of the used buffer should have either neutral or positive effect on Dicer processing. Nevertheless, the cleavage assay with both buffers needs to be performed to exclude the detrimental effect of different pH and buffer composition on Dicer activity in my assay. Fourth, recombinant Dicers used in the cleavage assay contained affinity tags, so that I cannot exclude the potential interference with the Dicer activity which might have manifested in the reduced Dicer processing. This issue needs to be resolved in further experiments.

Both recombinant Dicer isoforms showed similar, but reduced activity when processing pre-let-7a3, which indicated that they were equally affected by the above suggested factors. Therefore, they could be used in other cleavage assays as the potential difference in their processing of these substrates would be caused by differences in their inherent characteristic rather than the above suggested external factors.

When I performed non-radioactive cleavage assay with a long, perfectly complementary dsRNA, both Dicer isoforms generated ~22 nt small RNAs and longer dsRNA corresponding to cleavage intermediates, while there was no apparent difference in their mode of processing or in their activity. Nevertheless, no conclusion can be drawn from the result at the moment due to assay limitations. Given the high detection limit of the assay and artefacts arising from an uneven gel staining or diffusion, I probably would not be able to see a 50% difference in Dicer activity, if it would have existed. Therefore, additional assay optimization needs to be done to improve the sensitivity of the assay. The artefacts might be reduced or prevented by gel fixation, shortened duration of the gel staining or mirrored arrangement of samples in the gel. The signal to background ratio of Dicer cleavage products might be improved by increasing Dicer concentration, which should result in a higher amount of products and cleavage intermediates. If the implementation of suggested improvements does not lead to

increased sensitivity of the assay, it will be necessary to repeat the experiments with a radioactively labelled substrate.

The cleavage assay with dsRNAtetra substrate testing the ability of internal cleavage showed no difference between Dicer^{S} and Dicer^{O}. However, it suffered from the same drawbacks as the previous assay and they were even more pronounced due to structural features of dsRNAtetra preventing efficient Dicer processing.

There is a possibility that even improved non-radioactive cleavage assay would not show any difference in Dicer^{S} and Dicer^{O} processing of RNAi substrates. In that case, there are two possible explanations. First, the recombinant Dicer isoforms I prepared could be impaired, e.g. due to presence of affinity tags at the N-terminus, which might interfere with the helicase function, or due to prolonged purification, so they could not process the dsRNA substrate differently. Second, the published difference in activity of Dicer^{O} and Dicer^{S} was an artefact generated by the *in vitro* fluorescent cleavage assay, which measures the Dicer activity indirectly by the increase in fluorescence. However, nonspecific cleavage by contaminating RNase or spontaneous dissociation of the strands comprising the fluorescent substrate would also lead to an increase in fluorescence. Moreover, the bulky structure of the fluorescent dye and the quencher could have differential effect on the activity of Dicer^{O} and Dicer^{S}.

To test the functionality of my Dicer preparations it would be advisable to include a positive control in the non-radioactive cleavage assay to be able to distinguish, whether the assay conditions enables detection of the difference in Dicer activity. The positive control could be mouse Dicer without the entire helicase, which is known to be more active in processing perfectly complementary dsRNA than Dicer^{S} (Ma et al, 2008; Ma et al, 2012), mouse Dicer^{S} partially digested with Proteinase K (Zhang et al, 2002) or commercially available recombinant human Dicer (Genlantis). Moreover, I should test the original Dicer preparations for comparison, despite the fact that they are more than three years old so that they could already lose their activity.

To test the reproducibility of the *in vitro* fluorescent cleavage assay with previously published substrates, it should be performed with the Dicer preparations described in the thesis in two parallel set-ups. In the first set-up, the rate of Dicer cleavage would be measured as published previously by the increase in fluorescence (Podolska et al, 2014). In the second set-up, a non-radioactive assay described in this thesis would be performed with the fluorescently labelled substrate to visualize the cleavage products and to determine whether the increase in fluorescence correlates with the increase in an amount of the Dicer cleavage products.

6 Conclusions

The main aim of this thesis was to examine whether two mouse Dicer isoforms, $Dicer^{O}$ and $Dicer^{S}$, differ in *in vitro* processing of various substrates. In order to achieve the aim,

- I tested several Dicer purification protocols with varying number and combination of affinity resins. Finally, I established a Dicer purification protocol comprising TALON and Strep-Tactin resins, which yielded mouse Dicer isoforms of high purity and integrity.

- I developed a non-radioactive cleavage assay enabling visualization of Dicer cleavage products and processing intermediates similarly to radioactive cleavage assays, although with lower sensitivity. It involves separation of RNA species with urea-PAGE and subsequent staining of gels with SYBR Gold Nucleic Acid Gel Stain.

- I performed non-radioactive cleavage assays with various substrates. I observed no difference between two mouse Dicer isoforms, $Dicer^{O}$ and $Dicer^{S}$, in *in vitro* processing of a pre-let-7a3, a perfectly complementary dsRNA and a dsRNA lacking free ends. The observed results indicate that deletion of the HEL1 domain might not be sufficient for the increase of *in vitro* Dicer activity observed for the human Dicer without the entire helicase. Nevertheless, the validity of the observed results is limited due to technical limitations of the assay together with low activity of recombinant Dicers, possibly caused by the presence of affinity tags at the N-terminus. Therefore, further experiments are necessary to confirm these preliminary results.

7 References

Addo-Quaye C, Eshoo TW, Bartel DP, Axtell MJ (2008) Endogenous siRNA and miRNA targets identified by sequencing of the Arabidopsis degradome. *Current biology : CB* **18:** 758-762

Ameres SL, Zamore PD (2013) Diversifying microRNA sequence and function. *Nature reviews Molecular cell biology* **14:** 475-488

Anger M, Klima J, Kubelka M, Prochazka R, Motlik J, Schultz RM (2004) Timing of Plk1 and MPF activation during porcine oocyte maturation. *Molecular reproduction and development* **69:** 11-16

Aravin AA, Naumova NM, Tulin AV, Vagin VV, Rozovsky YM, Gvozdev VA (2001) Double-stranded RNA-mediated silencing of genomic tandem repeats and transposable elements in the D. melanogaster germline. *Current biology : CB* **11:** 1017-1027

Axtell MJ, Westholm JO, Lai EC (2011) Vive la difference: biogenesis and evolution of microRNAs in plants and animals. *Genome biology* **12:** 221

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

Bernstein E, Caudy AA, Hammond SM, Hannon GJ (2001) Role for a bidentate ribonuclease in the initiation step of RNA interference. *Nature* **409:** 363-366

Betancur JG, Tomari Y (2012) Dicer is dispensable for asymmetric RISC loading in mammals. *RNA (New York, NY)* **18:** 24-30

Block H, Maertens B, Spriestersbach A, Brinker N, Kubicek J, Fabis R, Labahn J, Schafer F (2009) Immobilized-metal affinity chromatography (IMAC): a review. *Methods in enzymology* **463:** 439-473

Bohnsack MT, Czaplinski K, Gorlich D (2004) Exportin 5 is a RanGTP-dependent dsRNA-binding protein that mediates nuclear export of pre-miRNAs. *RNA (New York, NY)* **10:** 185-191

Cai X, Hagedorn CH, Cullen BR (2004) Human microRNAs are processed from capped, polyadenylated transcripts that can also function as mRNAs. *RNA (New York, NY)* **10:** 1957-1966

Cenik ES, Fukunaga R, Lu G, Dutcher R, Wang Y, Tanaka Hall TM, Zamore PD (2011) Phosphate and R2D2 restrict the substrate specificity of Dicer-2, an ATP-driven ribonuclease. *Mol Cell* **42:** 172-184

Cormack BP, Valdivia RH, Falkow S (1996) FACS-optimized mutants of the green fluorescent protein (GFP). *Gene* **173:** 33-38

Cullen Bryan R, Cherry S, tenOever Benjamin R (2013) Is RNA Interference a Physiologically Relevant Innate Antiviral Immune Response in Mammals? *Cell Host & Microbe* **14:** 374-378

Denli AM, Tops BB, Plasterk RH, Ketting RF, Hannon GJ (2004) Processing of primary microRNAs by the Microprocessor complex. *Nature* **432:** 231-235

Doench JG, Petersen CP, Sharp PA (2003) siRNAs can function as miRNAs. *Genes & development* **17:** 438-442

Elbashir SM, Lendeckel W, Tuschl T (2001) RNA interference is mediated by 21- and 22-nucleotide RNAs. *Genes & development* **15:** 188-200

Fabrini R, De Luca A, Stella L, Mei G, Orioni B, Ciccone S, Federici G, Lo Bello M, Ricci G (2009) Monomer-dimer equilibrium in glutathione transferases: a critical re-examination. *Biochemistry* **48:** 10473-10482

Fairman-Williams ME, Guenther UP, Jankowsky E (2010) SF1 and SF2 helicases: family matters. *Curr Opin Struct Biol* **20:** 313-324

Feng Y, Zhang X, Graves P, Zeng Y (2012) A comprehensive analysis of precursor microRNA cleavage by human Dicer. *Rna-a Publication of the Rna Society* **18:** 2083-2092

Fire A, Xu S, Montgomery MK, Kostas SA, Driver SE, Mello CC (1998) Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. *Nature* **391:** 806-811

Flemr M, Malik R, Franke V, Nejepinska J, Sedlacek R, Vlahovicek K, Svoboda P (2013) A retrotransposon-driven dicer isoform directs endogenous small interfering RNA production in mouse oocytes. *Cell* **155:** 807-816

Forstemann K, Horwich MD, Wee L, Tomari Y, Zamore PD (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 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 biology* **3:** e236

Friedman RC, Farh KK, Burge CB, Bartel DP (2009) Most mammalian mRNAs are conserved targets of microRNAs. *Genome research* **19:** 92-105

Gantier MP, Williams BR (2007) The response of mammalian cells to double-stranded RNA. *Cytokine & growth factor reviews* **18:** 363-371

Gorbalenya AE, Koonin EV (1993) HELICASES - AMINO-ACID-SEQUENCE COMPARISONS AND STRUCTURE-FUNCTION-RELATIONSHIPS. *Current Opinion in Structural Biology* **3:** 419-429

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

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

Gu M, Rice CM (2010) Three conformational snapshots of the hepatitis C virus NS3 helicase reveal a ratchet translocation mechanism. *Proceedings of the National Academy of Sciences of the United States of America* **107:** 521-528

Ha M, Kim VN (2014) Regulation of microRNA biogenesis. *Nature reviews Molecular cell biology* **15:** 509-524

Hammond SM, Boettcher S, Caudy AA, Kobayashi R, Hannon GJ (2001) Argonaute2, a link between genetic and biochemical analyses of RNAi. *Science (New York, NY)* **293:** 1146-1150

Han J, Lee Y, Yeom KH, Kim YK, Jin H, Kim VN (2004) The Drosha-DGCR8 complex in primary microRNA processing. *Genes & development* **18:** 3016-3027

Han J, Lee Y, Yeom KH, Nam JW, Heo I, Rhee JK, Sohn SY, Cho Y, Zhang BT, Kim VN (2006) Molecular basis for the recognition of primary microRNAs by the Drosha-DGCR8 complex. *Cell* **125:** 887-901

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

Hutvagner G, Zamore PD (2002) A microRNA in a multiple-turnover RNAi enzyme complex. *Science (New York, NY)* **297:** 2056-2060

Chakravarthy S, Sternberg SH, Kellenberger CA, Doudna JA (2010) Substrate-Specific Kinetics of Dicer-Catalyzed RNA Processing. *Journal of Molecular Biology* **404:** 392-402

Chalfie M (1995) Green fluorescent protein. *Photochemistry and photobiology* **62:** 651-656

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

Jankowsky E, Fairman ME (2007) RNA helicases--one fold for many functions. *Curr Opin Struct Biol* **17:** 316-324

Jannot G, Boisvert ME, Banville IH, Simard MJ (2008) Two molecular features contribute to the Argonaute specificity for the microRNA and RNAi pathways in C. elegans. *RNA (New York, NY)* **14:** 829-835

Jiang F, Ramanathan A, Miller MT, Tang GQ, Gale M, Jr., Patel SS, Marcotrigiano J (2011) Structural basis of RNA recognition and activation by innate immune receptor RIG-I. *Nature* **479:** 423-427

Jonas S, Izaurralde E (2015) Towards a molecular understanding of microRNA-mediated gene silencing. *Nature reviews Genetics* **16:** 421-433

Jones-Rhoades MW, Bartel DP, Bartel B (2006) MicroRNAS and their regulatory roles in plants. *Annual review of plant biology* **57:** 19-53

Kawamata T, Seitz H, Tomari Y (2009) Structural determinants of miRNAs for RISC loading and slicer-independent unwinding. *Nat Struct Mol Biol* **16:** 953-960

Ketting RF (2011) The many faces of RNAi. *Developmental cell* **20:** 148-161

Ketting RF, Fischer SE, Bernstein E, Sijen T, Hannon GJ, Plasterk RH (2001) Dicer functions in RNA interference and in synthesis of small RNA involved in developmental timing in C. elegans. *Genes & development* **15:** 2654-2659

Kidwell MA, Chan JM, Doudna JA (2014) Evolutionarily conserved roles of the dicer helicase domain in regulating RNA interference processing. *The Journal of biological chemistry* **289:** 28352-28362

Kim Y, Yeo J, Lee Jung H, Cho J, Seo D, Kim J-S, Kim VN (2014) Deletion of Human tarbp2 Reveals Cellular MicroRNA Targets and Cell-Cycle Function of TRBP. *Cell reports* **9:** 1061-1074

Kim YK, Kim VN (2007) Processing of intronic microRNAs. *The EMBO journal* **26:** 775-783

Kobe B, Deisenhofer J (1995) A structural basis of the interactions between leucine-rich repeats and protein ligands. *Nature* **374:** 183-186

Kowalinski E, Lunardi T, McCarthy AA, Louber J, Brunel J, Grigorov B, Gerlier D, Cusack S (2011) Structural basis for the activation of innate immune pattern-recognition receptor RIG-I by viral RNA. *Cell* **147:** 423-435

Kumar M, Carmichael GG (1998) Antisense RNA: function and fate of duplex RNA in cells of higher eukaryotes. *Microbiology and molecular biology reviews : MMBR* **62:** 1415-1434

Lagos-Quintana M, Rauhut R, Lendeckel W, Tuschl T (2001) Identification of novel genes coding for small expressed RNAs. *Science (New York, NY)* **294:** 853-858

Lam AM, Keeney D, Frick DN (2003) Two novel conserved motifs in the hepatitis C virus NS3 protein critical for helicase action. *The Journal of biological chemistry* **278:** 44514-44524

Landthaler M, Yalcin A, Tuschl T (2004) The human DiGeorge syndrome critical region gene 8 and Its D. melanogaster homolog are required for miRNA biogenesis. *Current biology : CB* **14:** 2162-2167

Lau PW, Guiley KZ, De N, Potter CS, Carragher B, MacRae IJ (2012) The molecular architecture of human Dicer. *Nature Structural & Molecular Biology* **19:** 436-440

Lau PW, Potter CS, Carragher B, MacRae IJ (2009) Structure of the Human Dicer-TRBP Complex by Electron Microscopy. *Structure* **17:** 1326-1332

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

Lee Y, Jeon K, Lee JT, Kim S, Kim VN (2002) MicroRNA maturation: stepwise processing and subcellular localization. *The EMBO journal* **21:** 4663-4670

Lee Y, Kim M, Han JJ, Yeom KH, Lee S, Baek SH, Kim VN (2004a) MicroRNA genes are transcribed by RNA polymerase II. *Embo Journal* **23:** 4051-4060

Lee YS, Nakahara K, Pham JW, Kim K, He Z, Sontheimer EJ, Carthew RW (2004b) Distinct roles for Drosophila Dicer-1 and Dicer-2 in the siRNA/miRNA silencing pathways. *Cell* **117:** 69-81

Leuschner PJ, Ameres SL, Kueng S, Martinez J (2006) Cleavage of the siRNA passenger strand during RISC assembly in human cells. *EMBO reports* **7:** 314-320

Lin SB, Gregory RI (2015) MicroRNA biogenesis pathways in cancer. *Nat Rev Cancer* **15:** 321-333

Liu J, Carmell MA, Rivas FV, Marsden CG, Thomson JM, Song JJ, Hammond SM, Joshua-Tor L, Hannon GJ (2004) Argonaute2 is the catalytic engine of mammalian RNAi. *Science (New York, NY)* **305:** 1437-1441

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 (New York, NY)* **301:** 1921-1925

Liu Z, Wang J, Li G, Wang HW (2015) Structure of precursor microRNA's terminal loop regulates human Dicer's dicing activity by switching DExH/D domain. *Protein & cell* **6:** 185-193

Llave C, Xie Z, Kasschau KD, Carrington JC (2002) Cleavage of Scarecrow-like mRNA targets directed by a class of Arabidopsis miRNA. *Science (New York, NY)* **297:** 2053-2056

Lund E, Guttinger S, Calado A, Dahlberg JE, Kutay U (2004) Nuclear export of microRNA precursors. *Science (New York, NY)* **303:** 95-98

Luo D, Ding SC, Vela A, Kohlway A, Lindenbach BD, Pyle AM (2011) Structural insights into RNA recognition by RIG-I. *Cell* **147:** 409-422

Ma E, MacRae IJ, Kirsch JF, Doudna JA (2008) Autoinhibition of human dicer by its internal helicase domain. *Journal of Molecular Biology* **380:** 237-243

Ma EB, Zhou KH, Kidwell MA, Doudna JA (2012) Coordinated Activities of Human Dicer Domains in Regulatory RNA Processing. *Journal of Molecular Biology* **422:** 466-476

Ma J, Flemr M, Stein P, Berninger P, Malik R, Zavolan M, Svoboda P, Schultz RM (2010) MicroRNA activity is suppressed in mouse oocytes. *Current biology : CB* **20:** 265-270

MacRae IJ, Ma E, Zhou M, Robinson CV, Doudna JA (2008) In vitro reconstitution of the human RISC-loading complex. *Proceedings of the National Academy of Sciences of the United States of America* **105:** 512-517

MacRae IJ, Zhou K, Doudna JA (2007) Structural determinants of RNA recognition and cleavage by Dicer. *Nature Structural & Molecular Biology* **14:** 934-940

MacRae IJ, Zhou KH, Li F, Repic A, Brooks AN, Cande WZ, Adams PD, Doudna JA (2006) Structural basis for double-stranded RNA processing by dicer. *Science (New York, NY)* **311:** 195-198

Maity R, Pauty J, Krietsch J, Buisson R, Genois MM, Masson JY (2013) GST-His purification: a two-step affinity purification protocol yielding full-length purified proteins. *Journal of visualized experiments : JoVE***:** e50320

Maru Y, Afar DE, Witte ON, Shibuya M (1996) The dimerization property of glutathione S-transferase partially reactivates Bcr-Abl lacking the oligomerization domain. *The Journal of biological chemistry* **271:** 15353-15357

Matranga C, Tomari Y, Shin C, Bartel DP, Zamore PD (2005) Passenger-strand cleavage facilitates assembly of siRNA into Ago2-containing RNAi enzyme complexes. *Cell* **123:** 607-620

Meister G, Landthaler M, Patkaniowska A, Dorsett Y, Teng G, Tuschl T (2004) Human Argonaute2 mediates RNA cleavage targeted by miRNAs and siRNAs. *Mol Cell* **15:** 185-197

Miska EA, Alvarez-Saavedra E, Abbott AL, Lau NC, Hellman AB, McGonagle SM, Bartel DP, Ambros VR, Horvitz HR (2007) Most Caenorhabditis elegans microRNAs are individually not essential for development or viability. *PLoS genetics* **3:** e215

Murchison EP, Stein P, Xuan Z, Pan H, Zhang MQ, Schultz RM, Hannon GJ (2007) Critical roles for Dicer in the female germline. *Genes & development* **21:** 682-693

Nejepinska J, Malik R, Filkowski J, Flemr M, Filipowicz W, Svoboda P (2012) dsRNA expression in the mouse elicits RNAi in oocytes and low adenosine deamination in somatic cells. *Nucleic Acids Res* **40:** 399-413

Noland CL, Ma E, Doudna JA (2011) siRNA repositioning for guide strand selection by human Dicer complexes. *Mol Cell* **43:** 110-121

Nykanen A, Haley B, Zamore PD (2001) ATP requirements and small interfering RNA structure in the RNA interference pathway. *Cell* **107:** 309-321

Okada C, Yamashita E, Lee SJ, Shibata S, Katahira J, Nakagawa A, Yoneda Y, Tsukihara T (2009) A high-resolution structure of the pre-microRNA nuclear export machinery. *Science (New York, NY)* **326:** 1275-1279

Okamura K, Ishizuka A, Siomi H, Siomi MC (2004) Distinct roles for Argonaute proteins in small RNA-directed RNA cleavage pathways. *Genes & development* **18:** 1655-1666

Park JE, Heo I, Tian Y, Simanshu DK, Chang H, Jee D, Patel DJ, Kim VN (2011) Dicer recognizes the 5' end of RNA for efficient and accurate processing. *Nature* **475:** 201-205

Park MY, Wu G, Gonzalez-Sulser A, Vaucheret H, Poethig RS (2005) Nuclear processing and export of microRNAs in Arabidopsis. *Proceedings of the National Academy of Sciences of the United States of America* **102:** 3691-3696

Podolska K, Sedlak D, Bartunek P, Svoboda P (2014) Fluorescence-based high-throughput screening of dicer cleavage activity. *Journal of biomolecular screening* **19:** 417-426

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

Rand TA, Petersen S, Du F, Wang X (2005) Argonaute2 cleaves the anti-guide strand of siRNA during RISC activation. *Cell* **123:** 621-629

Rodriguez A, Griffiths-Jones S, Ashurst JL, Bradley A (2004) Identification of mammalian microRNA host genes and transcription units. *Genome research* **14:** 1902-1910

Rybak-Wolf A, Jens M, Murakawa Y, Herzog M, Landthaler M, Rajewsky N (2014) A Variety of Dicer Substrates in Human and C-elegans. *Cell* **159:** 1153-1167

Sawh AN, Duchaine TF (2013) A truncated form of dicer tilts the balance of RNA interference pathways. *Cell reports* **4:** 454-463

Sayed D, Abdellatif M (2011) MicroRNAs in development and disease. *Physiological reviews* **91:** 827-887

Schmidt TG, Batz L, Bonet L, Carl U, Holzapfel G, Kiem K, Matulewicz K, Niermeier D, Schuchardt I, Stanar K (2013) Development of the Twin-Strep-tag(R) and its application for purification of recombinant proteins from cell culture supernatants. *Protein expression and purification* **92:** 54-61

Schmidt TG, Koepke J, Frank R, Skerra A (1996) Molecular interaction between the Strep-tag affinity peptide and its cognate target, streptavidin. *J Mol Biol* **255:** 753-766

Sijen T, Plasterk RH (2003) Transposon silencing in the Caenorhabditis elegans germ line by natural RNAi. *Nature* **426:** 310-314

Singleton MR, Dillingham MS, Wigley DB (2007) Structure and mechanism of helicases and nucleic acid translocases. *Annual review of biochemistry* **76:** 23-50

Starega-Roslan J, Krol J, Koscianska E, Kozlowski P, Szlachcic WJ, Sobczak K, Krzyzosiak WJ (2011) Structural basis of microRNA length variety. *Nucleic Acids Res* **39:** 257-268

Steiner FA, Hoogstrate SW, Okihara KL, Thijssen KL, Ketting RF, Plasterk RH, Sijen T (2007) Structural features of small RNA precursors determine Argonaute loading in Caenorhabditis elegans. *Nat Struct Mol Biol* **14:** 927-933

Suh N, Baehner L, Moltzahn F, Melton C, Shenoy A, Chen J, Blelloch R (2010) MicroRNA Function Is Globally Suppressed in Mouse Oocytes and Early Embryos. *Current Biology* **20:** 271-277

Suzuki HI, Katsura A, Yasuda T, Ueno T, Mano H, Sugimoto K, Miyazono K (2015) Small-RNA asymmetry is directly driven by mammalian Argonautes. *Nat Struct Mol Biol* **22:** 512-521

Svoboda P (2014) Renaissance of mammalian endogenous RNAi. *FEBS letters* **588:** 2550-2556

Svoboda P, Stein P, Hayashi H, Schultz RM (2000) Selective reduction of dormant maternal mRNAs in mouse oocytes by RNA interference. *Development (Cambridge, England)* **127:** 4147-4156

Tam OH, Aravin AA, Stein P, Girard A, Murchison EP, Cheloufi S, Hodges E, Anger M, Sachidanandam R, Schultz RM, Hannon GJ (2008) Pseudogene-derived small interfering RNAs regulate gene expression in mouse oocytes. *Nature* **453:** 534-538

Tang F, Kaneda M, O'Carroll D, Hajkova P, Barton SC, Sun YA, Lee C, Tarakhovsky A, Lao K, Surani MA (2007) Maternal microRNAs are essential for mouse zygotic development. *Genes & development* **21:** 644-648

Tang G, Reinhart BJ, Bartel DP, Zamore PD (2003) A biochemical framework for RNA silencing in plants. *Genes & development* **17:** 49-63

Taylor DW, Ma E, Shigematsu H, Cianfrocco MA, Noland CL, Nagayama K, Nogales E, Doudna JA, Wang HW (2013) Substrate-specific structural rearrangements of human Dicer. *Nat Struct Mol Biol* **20:** 662-670

Thastrup O, Tullin S, Poulsen LK, Bjørn SP. (2001) Fluorescent proteins. Google Patents.

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

Tsutsumi A, Kawamata T, Izumi N, Seitz H, Tomari Y (2011) Recognition of the pre-miRNA structure by Drosophila Dicer-1. *Nat Struct Mol Biol* **18:** 1153-1158

Vermeulen A, Behlen L, Reynolds A, Wolfson A, Marshall WS, Karpilow J, Khvorova A (2005) The contributions of dsRNA structure to Dicer specificity and efficiency. *RNA (New York, NY)* **11:** 674-682

Voss S, Skerra A (1997) Mutagenesis of a flexible loop in streptavidin leads to higher affinity for the Strep-tag II peptide and improved performance in recombinant protein purification. *Protein engineering* **10:** 975-982

Wang Y, Medvid R, Melton C, Jaenisch R, Blelloch R (2007) DGCR8 is essential for microRNA biogenesis and silencing of embryonic stem cell self-renewal. *Nature Genetics* **39:** 380-385

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

Wianny F, Zernicka-Goetz M (2000) Specific interference with gene function by double-stranded RNA in early mouse development. *Nat Cell Biol* **2:** 70-75

Wilson RC, Tambe A, Kidwell MA, Noland CL, Schneider CP, Doudna JA (2015) Dicer-TRBP Complex Formation Ensures Accurate Mammalian MicroRNA Biogenesis. *Molecular Cell* **57:** 397-407

Yang Q, Del Campo M, Lambowitz AM, Jankowsky E (2007) DEAD-box proteins unwind duplexes by local strand separation. *Mol Cell* **28:** 253-263

Yang Y, Sharan SK (2003) A simple two-step, 'hit and fix' method to generate subtle mutations in BACs using short denatured PCR fragments. *Nucleic Acids Res* **31:** e80

Yoda M, Kawamata T, Paroo Z, Ye X, Iwasaki S, Liu Q, Tomari Y (2010) ATP-dependent human RISC assembly pathways. *Nat Struct Mol Biol* **17:** 17-23

Zhang HD, Kolb FA, Brondani V, Billy E, Filipowicz W (2002) Human Dicer preferentially cleaves dsRNAs at their termini without a requirement for ATP. *Embo Journal* **21:** 5875-5885

Zhang HD, Kolb FA, Jaskiewicz L, Westhof E, Filipowicz W (2004) Single processing center models for human dicer and bacterial RNase III. *Cell* **118:** 57-68

Zou J, Chang M, Nie P, Secombes CJ (2009) Origin and evolution of the RIG-I like RNA helicase gene family. *BMC evolutionary biology* **9:** 85