Dissertation zur Erlangung des Doktorgrades

der Fakultät für Chemie und Pharmazie

der Ludwig-Maximilians-Universität München

Contribution of RNA binding proteins to substrate specificity in small RNA biogenesis

Stephanie Marion Fesser

aus

Heidelberg, Deutschland

2013

Erklärung

Diese Dissertation wurde im Sinne von § 7 der Promotionsordnung vom 28. November 2011 von Herrn Prof. Dr. Klaus Förstemann betreut.

Eidesstattliche Versicherung

Diese Dissertation wurde eigenständig und ohne unerlaubte Hilfe erarbeitet.

München, 10.05.2013

Stephanie Marion Fesser

Dissertation eingereicht am 10.05.2013

1. Gutachter: Prof. Dr. Klaus Förstemann
2. Gutachter: Prof. Dr. Michael Sattler

Mündliche Prüfung am 02.07.2013

Summary

Small interfering RNAs (siRNAs) in *Drosophila* can be divided in endo- and exo-siRNAs, depending on their origin: endogenous siRNAs arise from convergent transcription or long hairpin structures formed by highly repetitive transcripts and protect especially somatic cells from the deleterious influence of “jumping” transposons. Exogenous siRNAs defend the cells against external threats and are generated from long dsRNAs which occur for example during infection with RNA viruses. The biogenesis pathways of endo- and exo-siRNAs are highly parallel: Long dsRNA precursors are processed by the RNaseIII enzyme Dicer2 (Dcr-2) with the help of a double stranded RNA binding protein (dsRBP) to 21nt siRNA duplexes. After this processing, Dcr-2, again with the help of a dsRBP, functions as the RISC loading complex (RLC) that loads the duplex in Ago2 to eventually form the RNA induced silencing complex (RISC).

There are two different dsRBPs involved in siRNA biogenesis: The PD isoform of Loquacious (LoqsPD) and R2D2. For which class of siRNAs and at which step of the biogenesis pathway Dcr-2 requires which dsRBP for its proper function is still unclear. No clear classification of LoqsPD and R2D2 can be made based on the type of siRNAs that depend on them. They seem to function redundantly in maturation of some endo-siRNAs, whereas others have differential requirements for LoqsPD or R2D2. Mechanistically, LoqsPD is mainly implicated in the processing- and R2D2 in the loading step. However, there are exceptions to this since in the absence of R2D2 Ago2-loaded siRNAs exist, and in the absence of Loqs a subset of siRNAs are still processed effectively.

In my thesis I wanted to elucidate the differential requirement for LoqsPD in siRNA biogenesis.

I characterized the binding behavior of LoqsPD and found that its two dsRBDs act rather independently in siRNA binding, and that the N- and C-terminal parts of the protein are required to achieve a substrate preference for completely base paired RNA substrates. siRNAs and their precursors are bound with the same affinity, allowing a function of LoqsPD in both the processing and the RISC loading step. LoqsPD binding to RNA substrates is cooperative, which is presumably achieved by protein-protein contacts. The affinity of LoqsPD for siRNAs is in the range of 50nM, and quantification of cellular LoqsPD content implies that local concentration of LoqsPD might be necessary for efficient siRNA binding. In comparison, R2D2 is present at higher concentrations in the cell, but its RNA binding affinities could not be determined since even truncation constructs were unstable without the stabilizing presence of Dcr-2.

Crosslink experiments showed that both LoqsPD and its partner Dicer2 crosslink preferentially to the ends of an siRNA duplex. This coincides well with current models of RNA binding by RNaseIII/dsRBP complexes in small RNA biogenesis, where the siRNA duplex is bound on one side by the RNaseIII enzyme and on the other by the dsRBP, thereby appointing guide and passenger strand. Indeed, LoqsPD can induce asymmetric binding of Dcr-2 on an siRNA duplex, corroborating the idea that the Dcr-2/LoqsPD complex can function as RLC in addition to Dcr-2/R2D2. Interaction of the two proteins is important for this asymmetry sensing.

Contents

1 Fundamentals **1**

1.1 RNA interference . . . 1
1.2 Classes and functions of small RNAs in Drosophila . . . 1
1.3 Roles of double stranded RNA binding proteins in RNAi . . . 2
1.4 Open questions . . . 3

2 Introduction **5**

2.1 Biogenesis of somatic small RNAs and assembly of the RNA induced silencing complex . . . 5
2.1.1 miRNAs . . . 5
2.1.2 siRNAs . . . 6
2.1.3 Roles of LoqsPD and R2D2 . . . 7
2.2 Domain architecture and structure of Dcr-1, Dcr-2 and the dsRBPs . . . 7
2.2.1 dsRBPs . . . 8
2.2.2 Dicer . . . 8
2.3 RNA binding proteins . . . 9
2.3.1 The double-stranded RNA binding domain . . . 10
2.3.1.1 RNA binding interface . . . 10
2.3.1.2 dsRNA specificity . . . 12
2.3.1.3 Sequence dependence of dsRBD binding . . . 12
2.3.1.4 Concerted action of multiple dsRBDs . . . 12
2.4 Measuring RNA-protein interactions . . . 13
2.4.1 Electrophoretic mobility shift assay . . . 13
2.4.2 Anisotropy . . . 14
2.4.3 Thermophoresis . . . 15

3 Aims of this Thesis **17**

4 Material and Methods **19**

4.1 Material . . . 19
4.1.1 Plasmids . . . 19
4.1.2 Primer . . . 20
4.1.3 RNA Oligos . . . 21

4.1.4 Hardware . . . 22
4.1.5 Chemicals . . . 23
4.1.6 Enzymes . . . 24
4.1.7 Buffers and solutions . . . 25
4.1.8 Bacterial strains and media . . . 26
4.1.8.1 Strains . . . 26
4.1.8.2 Media . . . 27
4.1.9 Cell lines and Media . . . 27
4.1.10 Antibodies . . . 27
4.2 Methods . . . 28
4.2.1 Molecular Cloning . . . 28
4.2.1.1 Primer design . . . 28
4.2.1.2 Polymerase chain reaction (PCR) . . . 28
4.2.1.3 Agarose gel electrophoresis . . . 29
4.2.1.4 Restriction digestion and ligation of DNA fragments . . . 29
4.2.1.5 Transformation in E.coli . . . 30
4.2.1.6 Preparation of plasmid DNA . . . 30
4.2.1.7 Sequencing and analysis of results . . . 30
4.2.1.8 Site directed mutagenesis . . . 30
4.2.2 Protein techniques . . . 31
4.2.2.1 Protein lysates . . . 31
4.2.2.2 Denaturing SDS-Gelelectrophoresis . . . 31
4.2.2.3 Western Blotting . . . 31
4.2.2.4 Determination of protein concentration . . . 31
4.2.2.5 Co-Immunoprecipitation . . . 32
4.2.2.6 Mass spectrometry . . . 32
4.2.3 RNA techniques . . . 32
4.2.3.1 Oligo design . . . 32
4.2.3.2 General RNA handling . . . 32
4.2.3.3 Denaturing Urea gel electrophoresis . . . 33
4.2.3.4 RNA extraction . . . 33
4.2.3.5 RNA gel purification . . . 33
4.2.3.6 RNA phosphorylation . . . 33
4.2.3.7 Ligation of pre-miRNA hairpin . . . 34
4.2.3.8 Radioactive labeling of RNA . . . 34
4.2.3.9 Production of dsRNA for gene specific knockdown . . . 34
4.2.4 Cell culture . . . 35
4.2.4.1 Culture of *Drosophila* Schneider 2 cells . . . 35

4.2.4.2 Culture of Sf21 cells . . . 35
4.2.4.3 Culture of High5 cells . . . 35
4.2.4.4 Cell lysates . . . 36
4.2.4.5 Transfection . . . 36
4.2.4.6 RNAi . . . 36
4.2.5 Fluorescence Activated Cell Sorting . . . 36
4.2.6 Immunofluorescence . . . 36
4.2.7 Production of recombinant protein . . . 37
4.2.7.1 Generation of competent BL21 and DH10 cells . . . 37
4.2.7.2 PreScission Protease . . . 38
4.2.7.3 TEV Protease . . . 38
4.2.7.4 Loquacious . . . 39
4.2.7.5 Loquacious, R2D2 and their dsRBDs . . . 40
4.2.7.6 Dicer . . . 41
4.2.8 Binding experiments . . . 42
4.2.8.1 Anisotropy measurements . . . 42
4.2.8.2 Electrophoretic mobility shift assays . . . 42
4.2.8.3 Thermophoresis measurements . . . 43
4.2.9 Crosslink experiments . . . 43
4.2.10 Analysis . . . 43

5 Results **45**
5.1 Characterization of LoqsPD and R2D2 in their cellular environment . . . 45
5.1.1 Subcellular localization of Loquacious and R2D2 . . . 45
5.1.2 R2D2 is present in larger amounts than LoqsPD . . . 47
5.1.3 R2D2 binds Dcr-2 in the linker of the helicase domain . . . 48
5.1.4 Excess of nonspecific dsRNA does not interfere with endo-siRNA function 50
5.2 Optimization of protein expression and purification . . . 51
5.2.1 Loquacious . . . 51
5.2.2 R2D2 . . . 53
5.2.3 dsRBDs of Loquacious and R2D2 . . . 55
5.2.4 Dicer . . . 55
5.3 Binding experiments - Method optimization and validation . . . 55
5.3.1 EMSA . . . 56
5.3.2 Anisotropy . . . 58
5.3.3 Thermophoresis . . . 59
5.3.4 RNA concentration . . . 59
5.3.5 Attached Fluorescein does not interfere with RNA - protein binding . . . 62

5.3.6 Method comparison . . . 62
5.3.7 Thio-Uridine crosslinks replicate natural binding behavior . . . 65
5.4 Binding behavior of the Loquacious double-stranded RNA binding domains . . . 67
5.4.1 Both Loqs RNA binding domains bind RNA with a similar K_D . . . 67
5.4.2 Loqs dsRBDs do not show target specificity . . . 69
5.4.3 Loqs dsRBD binding is rather transient . . . 69
5.4.4 Comparison with R2D2 dsRBDs . . . 70
5.5 Interplay of the two Loquacious double-stranded RNA binding domains . . . 71
5.5.1 Full length LoqsPD and Loqs DeltaNC show enhanced RNA binding affinity . . . 71
5.5.2 Full length LoqsPD has an inherent bias towards siRNA binding . . . 73
5.5.3 Full length LoqsPD has the highest propensity to distinguish between RNA substrates . . . 75
5.5.4 LoqsPD binds two different sequences with similar affinity . . . 77
5.5.5 Comparison with R2D2 DeltaNC . . . 77
5.6 ssRNA binding by double-stranded RNA binding domains . . . 78
5.7 Cooperativity in dsRBD binding . . . 80
5.7.1 Stoichiometry of protein-RNA complexes . . . 80
5.7.2 The Increase in K_D is not as large as expected for cooperative binding . 82
5.7.3 Loqs binding curves require a Hill coefficient for accurate fitting . . . 82
5.7.4 The Hill coefficient does not reflect cooperativity mediated by RNA deformation . . . 84
5.8 Geometry of RNA binding . . . 85
5.8.1 Proteins crosslink preferentially at the extremities of the siRNA duplex . . 85
5.8.1.1 dsRBDs . . . 85
5.8.1.2 Duplex dsRBDs . . . 86
5.8.2 Loqs DeltaNC binding is not completely limited to the duplex ends . . . 86
5.8.3 miRNA binding geometry . . . 88
5.8.4 R2D2 . . . 89
5.9 siRNA duplex asymmetry sensing by the Dcr-2/LoqsPD complex . . . 90
5.9.1 LoqsPD alone does not show preferential binding to one end of the siRNA duplex . . . 90
5.9.2 LoqsPD and Dcr-2 can sense the thermodynamic asymmetry of the RNA duplex . . . 91
5.9.3 The Dcr-2 - LoqsPD interaction is necessary for asymmetry sensing . . . 91
5.9.4 RNA deformation is not the mediator for asymmetric binding of the Dcr-2/LoqsPD complex . . . 93

6 Discussion **95**
6.1 LoqsPD and R2D2 in their cellular environment . . . 95
6.1.1 Concentration . . . 95
6.1.2 Localization . . . 95
6.1.3 Competition for complex formation with Dcr-2 . . . 96
6.2 Protein purification . . . 97
6.3 RNA binding properties of LoqsPD . . . 98
6.3.1 Sequence conservation . . . 98
6.3.2 Binding affinities and interplay of the two Loqs dsRBDs . . . 98
6.3.2.1 Both Loqs dsRBDs contribute equally to RNA binding . . . 99
6.3.2.2 Commitment of LoqsPD to the siRNA pathway . . . 100
6.3.3 Cooperativity of dsRBD binding . . . 100
6.3.3.1 Cooperativity of the two dsRBDs: Single vs tandem dsRBD binding . . . 100
6.3.3.2 Cooperative RNA binding of LoqsPD: The Hill coefficient . . . 102
6.3.4 ssRNA binding . . . 103
6.4 Protein - RNA Binding geometry . . . 104
6.4.1 dsRBD sliding . . . 104
6.4.2 Asymmetry sensing0, . . . 105
6.5 Comparison with R2D2 . . . 106
6.5.1 Binding behavior of R2D2 constructs . . . 106
6.5.1.1 Conservation of R2D2 dsRBDs . . . 106
6.5.1.2 Stability and binding of R2D2 . . . 106
6.6 Roles of LoqsPD and R2D2 in small RNA biogenesis . . . 108

Appendix **127**
Significance of differences in dissociation constants . . . 127
Labtalk script for automated evaluation of Nanotemper Monolith data . . . 127
Mass spectrometry results list . . . 131

1 Fundamentals

1.1 RNA interference

RNA interference (RNAi) as a mechanism to target specific RNA molecules has obtained much attention since its discovery in the late 90ies. Mello's and Fire's initial experiments in *C.elegans* showed that dsRNA was most potent in inhibiting gene expression and not, as previously thought, single stranded antisense RNA [28]. That was the starting signal for a tremendous expansion of the RNA research field. Today, RNAi is known to be involved in a variety of processes: Fine-tuning of gene expression in a post-transcriptional manner [4], defense against mobile genetic elements and viruses [106], heterochromatin assembly [107, 110, 27], and, as shown most recently, it also occurs as a consequence of DNA damage [112, 69]. The RNA silencing machinery is a complex of effector proteins (usually a member of the argonaute family) and a small RNA, which confers target specificity to the complex. It is called the RNA induced silencing complex (RISC).

Due to its programmable specificity, RNAi is also exploited as a mean to turn off single genes. In laboratory practice, RNAi is routinely used to investigate gene function via a gene-specific knock down. In medicine, RNAi based therapeutics were considered to be a hope for treating genetic diseases, and even though problematic delivery has dampened these prospects there is still a lot of potential in this technique.

1.2 Classes and functions of small RNAs in Drosophila

In Drosophila, there are three main classes of small RNAs: piwi-interacting RNAs (piRNAs), microRNAs (miRNAs), and small interfering RNAs (siRNAs), which can be classified as endogenous (endo-) or exogenous (exo-), depending on their origin.

piRNAs protect the genome against transposition of mobile genetic elements, primarily in the germline where disruption of genome integrity would be most detrimental. Their precursors are encoded in piRNA master loci, where sequences of potentially hazardous transposons have accumulated over time. The mature piRNAs are 22-30nt in length and associate with Piwi/Aub and Ago3 effector proteins, which are members of the Piwi subclass of Argonaute proteins. Cleavage of their target RNA is part of their biogenesis, leading to an auto-amplification loop as transposon silencing mechanism [11].

miRNAs act via recruitment of the RISC to their target mRNAs, mostly within the 3'UTR of mRNAs, thereby silencing them post-transcriptionally either by repressing their translation or by facilitating their degradation [14]. miRNAs are also encoded in the genome in individual genes, allowing for regulated expression. Deregulation of miRNA expression is associated with various diseases in humans [50]. In the miRNA database miRbase, 1600 human miRNAs are listed, for *Drosophila* there are 238, and a major part of all mRNAs is thought to be targeted by miRNAs. In contrast to piRNA clusters, miRNA precursors can fold back to form a hairpin structure with a double-stranded stem. This structure is then further processed. In their mature form, miRNAs have a length of ~22nt, and in *Drosophila* they are incorporated mostly into Ago1, a member of the Argonaute subfamily.

siRNAs defend the cells against external threats like invading viruses or endogenous threats like jumping transposons, particularly in the soma where no piRNA system is present to control mobile genetic elements. Endo-siRNAs also play a role in regulating gene expression [56, 55]. In the case of exo-siRNAs, the precursor dsRNA can be either directly introduced, produced as a replication intermediate or by secondary structure formation of viral ssRNA [91]. Endo-siRNA precursors arise by transcription of a head-to-head integrated transposon or other transcripts, that can fold back to generate dsRNA, or by bidirectional, converging transcription, which can either happen stochastically through leaky transcription or via a cryptic promoter present on the partner strand of a transcribed gene [79, 20, 80]. This long dsRNA is then processed into the mature 21nt long siRNA, which is incorporated in a RISC containing Ago2, like Ago1 a member of the Argonaute subfamily. In contrast to Ago1 RISC, Ago2 RISC is capable of efficiently degrading its target RNA.

1.3 Roles of double stranded RNA binding proteins in RNAi

miRNAs and siRNAs follow the same basic biogenesis route: They are excised from their precursors by an RNaseIII enzyme, and subsequently loaded into their respective RISC. dsRBPs play a role in both processes. An overview of their biogenesis pathways is depicted in figure 2.1. The dsRBP LoqsPB assists the RNaseIII enzyme Dcr-1 in pre-miRNA cleavage, by enhancing both the dicing reaction and affinity for its correct target [92]. Another isoform of this protein, LoqsPD, assists Dcr-2 in cleaving endo-siRNA precursors by accelerating the dicing reaction. The dsRBP R2D2 has the same effect on Dcr-2, albeit to a lesser extent [17], and it is mainly involved in exo-siRNA biogenesis. For both, exo- and endo-siRNAs, R2D2 functions in RISC loading, where it assists Dcr-2 in choosing the strand of the siRNA duplex that gets incorporated in the RISC [104].

1.4 Open questions

In most cases, endo-siRNAs are processed by Dcr-2 with the help of LoqsPD and loaded into RISC by the Dcr-2/R2D2 complex. But experiments with cell culture based reporter systems or knock out flies have shown that there are exceptions to this rule, and that some RNA species only depend on one of the dsRBPs in their biogenesis. Exo-siRNAs are both processed and loaded by R2D2. Since the RNaseIII enzyme Dcr-2 is involved in biogenesis of all siRNAs, the dsRBPs seem to be able to discriminate certain siRNA precursors and duplexes from others. The way this distinction is made, and which parameters are involved are still not understood. Insight into this mechanism would not only clarify these aspects of siRNA biogenesis but might also reveal a general way of substrate discrimination.

2 Introduction

2.1 Biogenesis of somatic small RNAs and assembly of the RNA induced silencing complex

2.1.1 miRNAs

Maturation of miRNAs requires the consecutive action of two RNaseIII/dsRBP complexes, that liberate the miRNA/miRNA* duplex from the pri-miRNA. After being transcribed, the pri-miRNA is cleaved by the Drosha/Pasha complex at the stem of the hairpin, releasing the pre-miRNA stemloop. The pre-miRNA is then exported into the cytoplasm via Exportin-5 in a Ran GTP-dependent manner [63], where the Dcr-1/Loqs-PB complex separates the double-stranded stem from the loop, producing a 23nt long RNA duplex containing several mismatches due to the non-complementary nature of the two miRNA strands, and a 2nt 3' overhang on both strands. LoqsPB enhances this dicing reaction, confers specificity for the pre-miRNA substrate and, at least for a subset of miRNAs, ensures production of the correct isomir by helping Dcr-1 to select the correct cleavage position [92, 31].

After processing, the miRNA/miRNA* duplex can be loaded in both Ago1 and Ago2, depending on the amount and the position of mismatches in the duplex [103, 29]. For the Ago1 loading step, neither LoqsPB nor Dcr-1 or 2 seem to be required [60, 42]. It is ATP dependent [42], arguing for a requirement of an additional factor: This has been shown to be the Hsc70/Hsp90 chaperone machinery, which pries the Argonaute protein open in order to accommodate the miRNA/miRNA* duplex [40, 41]. To form a mature miRISC, the passenger strand has to be removed. The incomplete base pairing of the miRNA/miRNA* duplex facilitates this unwinding and compensates the inability of Ago1 to cleave the passenger strand [120, 68].

The decision which strand of the miRNA duplex will confer specificity to the RISC as the guide strand is of importance, since it defines the range of target mRNAs. For Ago1, it is made based on the 5' nucleotide and the thermodynamic stability of the duplex, the strand with the more stable 5' end being chosen as the guide, whereas in Ago2 the duplex pairing status at position 9 and 10 has the greatest influence on strand selection of miRNA duplexes [82].

After loading of Ago1, it associates with GW182 [71], which recruits the miRISC to P-bodies containing enzymes required for mRNA decay. The CCR4-CAF1-NOT deadenylase complex associates with GW182 and is thereby tethered to the miRNA target, destabilizing the tran-

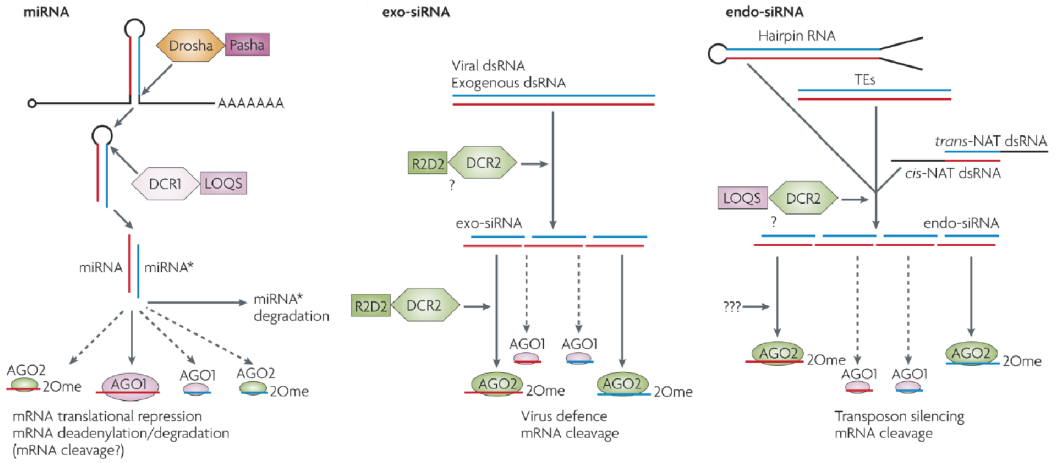

Figure 2.1: Biogenesis and function of somatic small RNAs in Drosophila
Classes of small RNAs in *Drosophila* somatic cells. Their precursors are processed to 21-23nt duplexes by an RNaseIII enzyme (hexagons) with the assistance of a dsRBP (squares), which are subsequently loaded into an effector complex containing Argonaute protein (oval). Ago2 loaded small RNAs become 2-O-methyl modified. Adapted from [81]

script by shortening of the poly-A tail and by mediating translational repression [19, 10].

2.1.2 siRNAs

After being transcribed, endo-siRNA precursors anneal with their partner strands generated either by convergent transcription or secondary structure formation of the RNA precursor itself and are exported to the cytoplasm, presumably also via Exportin-5 [63]. This double stranded RNA is then processed by Dcr-2 in an ATP-dependent manner [58, 17] with the help of the PD isoform of Loquacious, which increases the substrate affinity of Dcr-2 and its dicing efficiency [70, 31, 17].

R2D2, the other dsRBP partner of Dcr-2, has not been implicated in enhancing dsRNA processing by Dcr-2 as strongly as LoqsPD [17], but nevertheless, siRNA production after viral infection and successful RNAi-mediated gene knock-down require Dcr-2 and R2D2.

Both endo- and exo-siRNAs are loaded in Ago2 [43, 109]. Ago2 loading requires a RISC loading complex (RLC), consisting of Dcr-2 and R2D2. The siRNA duplex is bound by the Dcr-2/R2D2 complex in a way that the end with the lower melting temperature interacts with Dcr-2; this asymmetric binding determines which strand of the duplex remains in the RISC as the guide, and which is discarded as the passenger strand [104]. Sensing the thermodynamic stability of the siRNA duplex is conserved in the human Dicer/TRBP complex [34, 78]. Additional factors involved in siRISC formation are Hsp90, analogous to its role in the human

RLC [72, 40] and C3PO, an endoribonuclease which assists in removing the sliced passenger strand [62]. In humans, the autoantigen La has been shown to facilitate the release of cleaved mRNA from RISC, thereby enhancing multiple-turnover RISC catalysis [61]. The siRISC functions in actively degrading its target RNA, a process which is highly efficient due to the capacity of Ago2 to catalyze multiple rounds of cleavage.

2.1.3 Roles of LoqsPD and R2D2

Much evidence points to LoqsPD and R2D2 sharing the work as partners of Dcr-2 in the biogenesis pathway of endo-siRISC: LoqsPD is believed to mainly process the dsRNA precursors to 21nt siRNA duplexes and R2D2 to be responsible for loading it into Ago2. For the human Dicer/TRBP complex it has been shown that the duplex dissociates from the Dicer/dsRBP complex to be rebound in the desired orientation rather than getting reorganized within the complex [78]; this uncoupling of processing and loading could be achieved via the two different dsRBPs in Drosophila. Another example for uncoupling of these processes are miRNAs with extensive double stranded character like miR-277 that are processed by Dcr-1 but loaded in Ago2 via Dcr-2/R2D2 [29].

But there is also data arguing against such a sequential model: Several endo-siRNAs depend solely on LoqsPD, and not R2D2, such as endo-siRNAs derived from the *klarsicht* locus [79], or specific esi-RNAs [22]. In cell culture experiments, the repression of an endo-siRNA reporter consisting of a GFP transgene mimicking a transposon depends solely on LoqsPD, with R2D2 even hampering endo-siRNA mediated repression [37, 38]. Deep sequencing data of small RNAs isolated from LoqsPD or R2D2 mutant flies show Ago2-loading of some endo-siRNAs even in the absence of R2D2, and residual processing of precursor RNA to the 21nt species is retained in LoqsPD mutants. No specific endo-siRNA feature can be made responsible for a differential dependence on R2D2 or LoqsPD. Silencing by added exo-siRNAs depends on R2D2 [37, 38], but LoqsPD has also been associated with viral siRNA biogenesis [36]. Therefore, the roles of LoqsPD and R2D2 in siRNA biogenesis are still not completely defined.

2.2 Domain architecture and structure of Dcr-1, Dcr-2 and the dsRBPs

Proteins involved in RNAi contain several domains, which confer activity and specificity to the pathway. An overview over their domain architecture is given in figure 2.2.

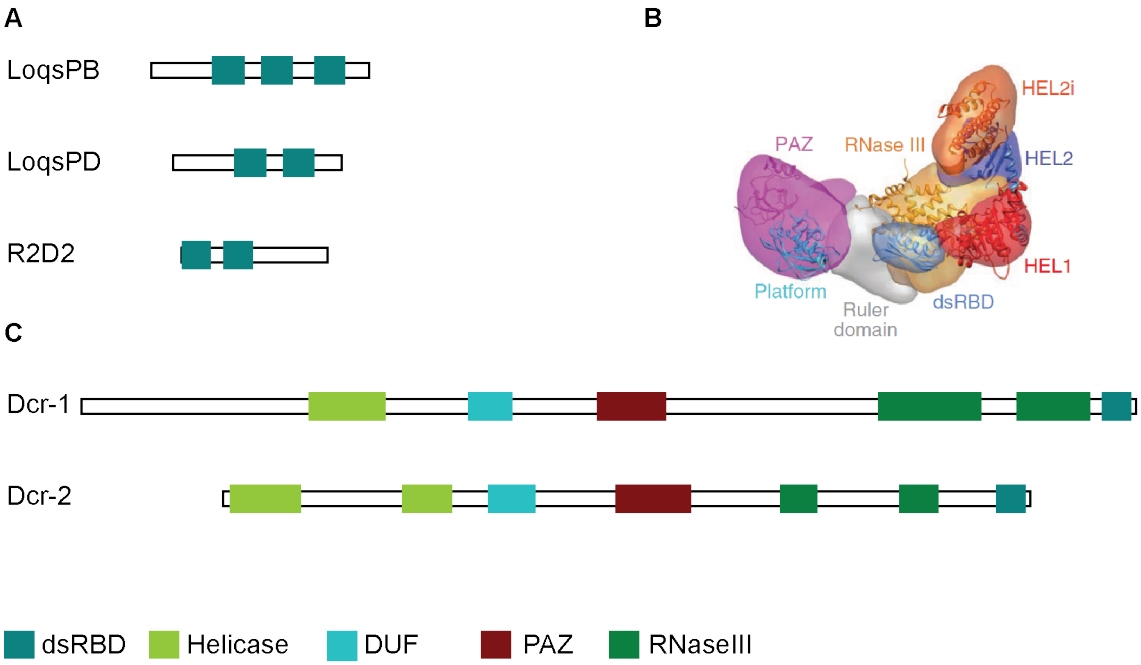

Figure 2.2: Domain architecture of Dicer and dsRBPs
A, C: Schematic representation of domain content and arrangement of dsRBPs and Dicer proteins, respectively. B: Model of human Dicer (from [51]).

2.2.1 dsRBPs

The three dsRBPs involved in the cytoplasmic part of mi- and siRNA biogenesis, LoqsPB, LoqsPD and R2D2, all contain two dsRBDs involved in RNA binding. LoqsPB contains a third dsRBD at the C-terminus, with which it associates with Dcr-1 [119]. LoqsPD shares the first two dsRBDs with LoqsPB, but lacks the third one due to alternative splicing. Instead, it contains a 22AA long isoform specific C-terminus, which mediates association with Dcr-2, assisted by its second dsRBD [38, 70]. R2D2 also interacts with Dcr-2 via its C-terminus [76], which is predicted to fold into a dsRBD-like structure but lacks sequence conservation. Structure and functionality of the dsRBD are discussed in chapter 2.3.1.

2.2.2 Dicer

Dicer are large, multidomain proteins of ~ 200kD. They consist of a Helicase, a PAZ and an RNaseIII domain, a dsRBD and a domain of unknown function (DUF), as depicted in figure 2.2. As seen in EM reconstructions, the structure of Dicer resembles an L (figure 2.2B) [108, 52]: The Helicase domain is positioned in the short branch of the Dicer L-form and has diverse functions and appearances in different Dicers: Dcr-2 contains a complete Helicase domain consisting of the DEXDc and the HELICc subdomain, with the DEXDc subdomain including

the WalkerB motif with the ATP binding site. It is required for repetitive cleavage of long dsRNA under ATP consumption [17] and for efficient processing of dsRNA with blunt or 5' overhanging termini [113]. Dcr-1's Helicase domain consists only of the HELICc subdomain, which by itself cannot form a completely functional helicase. It recognizes the terminal loop of pre-miRNAs, thereby limiting Dicer processing to its authentic substrate[105]. In addition, the Dicer proteins interact with the dsRBPs involved in RNA processing and RISC loading via their Helicase domain, including human Dicer [38, 108].

The RNaseIII domains are located next to the dsRBD domain in the kink of the L, adjacent to the Helicase. They form a conserved dimer, which generates a valley-shaped dsRNA binding surface where the two active sites are separated by ~ 20Å [33, 32]. Each RNaseIII domain cleaves one strand of the RNA duplex: in human Dicer the RNaseIIIA domain cleaves the 3'arm, the RNaseIIIB domain the 5' arm of pre-miRNAs [35], and their separation results in the 2nt overhang of the 3' end. RNA cleavage is Mg^{2+} -dependent [100], and proceeds via a bimolecular nucleophilic substitution-type mechanism [16].

The PAZ domain sits on the end of the long arm of the L [51], separated from the RNaseIII domains by a flat, positively charged surface [65]. This distance serves as a molecular ruler, defining the length of the Dicer products: with the PAZ domain securing the 3' 2nt overhang of one strand, the RNaseIII domains are positioned 21-23nt further away, generating the canonical Dicer products [51, 121]. It folds into a nucleic-acid-binding motif [57]: It favors single stranded RNA and anchors the 2nt 3' overhang of the guide strand of the siRNA duplex. This interaction is supported by additional residues binding to the phosphodiester backbone of the guide and the last nucleotide of the passenger strand [64].

In addition, Dicer also contains a DUF283 domain, which adopts a double-stranded RNA-binding fold and in plants mediates the interaction with the dsRBP [25, 87].

2.3 RNA binding proteins

RNA molecules fulfill diverse functions in a cell: In addition to mRNA, tRNA and rRNA, which are directly involved in protein translation, there are a huge number of regulatory RNAs like siRNAs, miRNAs, snRNAs, and catalytically active RNAs. They can adopt a variety of structures, since compared to DNA, RNA has an increased ability to form hydrogen bonds due to the extra hydroxyl group in the ribose sugar. To control these diverse RNA molecules, a set of RNA binding domains has evolved, conferring specificity for different RNA motifs to the proteins they reside in, including the double stranded RNA binding domain (dsRBD), the RNA recognition motif (RRM), the K-homology (KH) domain, the cold shock domain (CSD) and Zinc fingers.

2.3.1 The double-stranded RNA binding domain

The dsRBD is a well conserved, 65 - 70AA domain that specifically binds dsRNA, irrespective of sequence, and also mediates protein-protein interactions [12, 84]. The first structures of a dsRBD were solved for the *E. coli* RNaseIII dsRBD [44] and the *Drosophila* Staufen dsRBD3 [15] in 1995, and since then a wealth of other dsRBD structures have been solved by NMR or X-ray crystallography, conveying for all the common basic αßßßα-fold where the two α-helices are packed against the three-stranded anti-parallel ß-sheet as depicted in figure 2.3A, with occasional slight variations in the linker regions between α1and ß1 or ß1 and ß2, or with C-terminal extensions [2, 54, 39].

2.3.1.1 RNA binding interface

Structures together with RNA substrates revealed that the dsRBD spans ~16bp, corresponding to 1.5 turns of the helix, and that it binds to only one side, without wrapping around it, thereby covering an area of ~1600Å^2 (see figure 2.3B and [101, 90]). Binding is mediated via the phosphate backbone and the 2'-OH groups in the ribose moiety, with only limited interaction with the bases [90, 39, 116]. Three protein regions mediate contact to the RNA: Region 1 resides in α1 and includes a polar or positively charged AA and a conserved Glutamate which form hydrogen bonds to the 2'-OH of the ribose. In addition, less well conserved residues mediate RNA binding via van der Waals interactions and hydrogen bonds. Region 2 lies in the linker between ß1 and ß2, with a conserved GPxH motif involved in RNA binding, mainly via hydrogen bond formation to the ribose 2'-OH. The N-terminal part of α2 constitutes region 3. Here, a conserved KKxAK motif mediates RNA binding [67, 116]. Along the RNA helix, these regions are arranged in the order 1-3-2, with region 1 binding to the minor groove, region 3 to the major groove, and region 2 to the next minor groove. Since the A-form dsRNA major groove is narrower than the minor groove (10 vs. 15Å), the contacts to the major groove involve mainly the phosphate backbone whereas the minor groove is more accessible and contacts are made with the ribose moieties and even the bases.

In addition to RNA binding, there are residues conserved for the fold: They maintain a hydrophobic core between α helices and ß strands and thereby the overall structure. Among these are two aromatic residues, a Tyrosine in ß1 and a Phenylalanine in ß2 that are also crucial for RNA binding, since they keep the positively charged residues contacting the RNA in their correct orientation [48]. An alignment showing conservation and secondary structure elements is shown in figure 2.3C.

Based on the degree of their conservation, dsRBDs can be divided in type A and type B: type A dsRBDs are characterized by conservation along the whole domain, whereas type B dsRBDs show conservation only in the C-terminal part. Examples for these two types are the first and second dsRBD of PKR, whose dsRBD1 being of type A and dsRBD2 of type B.

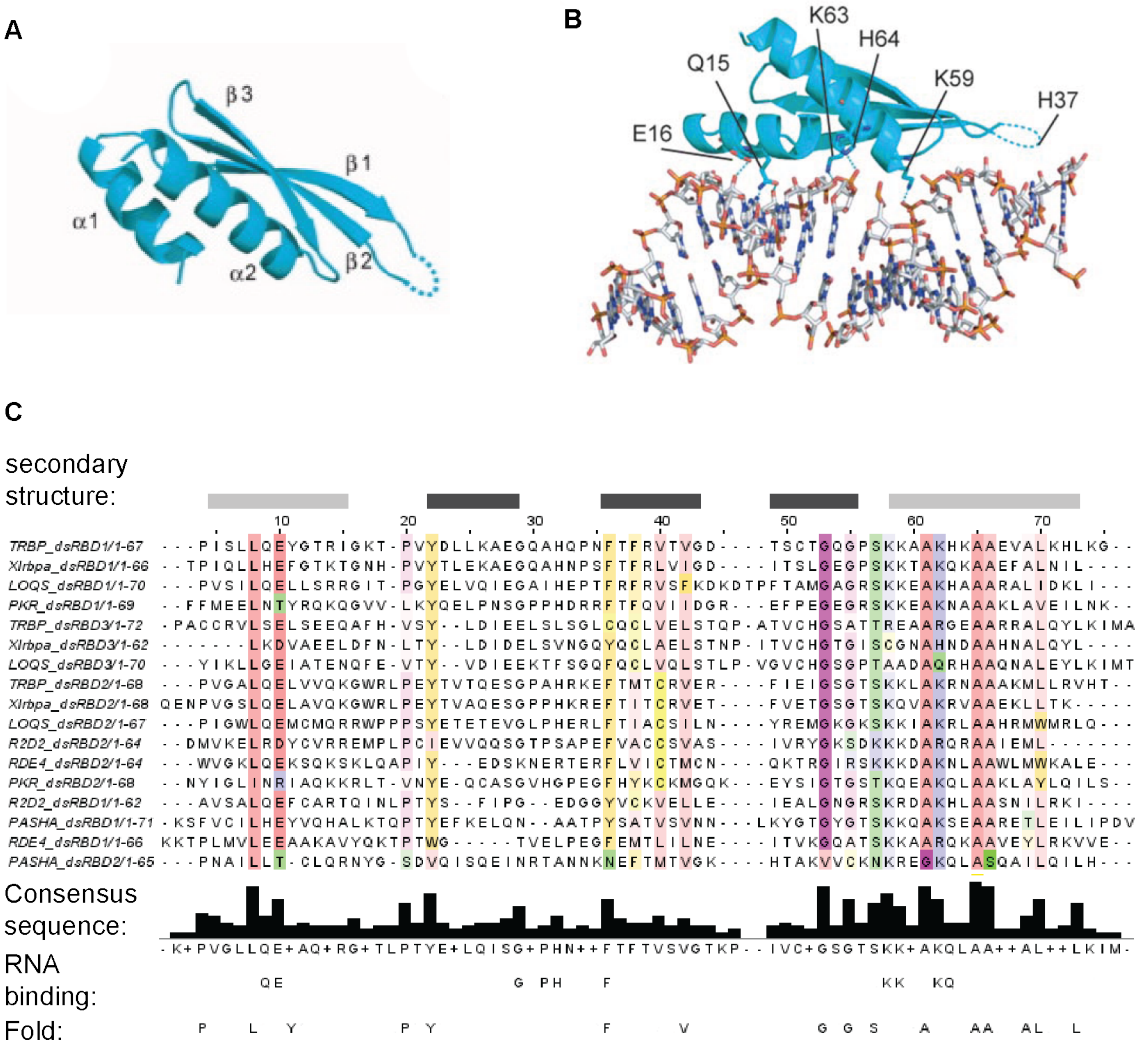

Figure 2.3: The double stranded RNA binding domain

A: Structure of TRBP dsRBD1, two α-Helices are packed against a three stranded ß-sheet (from [116]); B: RNA binding interface of TRBP dsRBD1 (from [116]), C: Conservation of the dsRBD sequence, grey bars indicate the secondary structure (light grey: α-Helices, dark grey: ß-sheets). Conserved residues involved in RNA binding and maintenance of the fold are depicted at the bottom (adapted from [67]).

Usually, type A dsRBDs have a higher affinity for dsRNA than type B, even though exceptions have been found, such as the dsRBDs in DIP1 [24].

2.3.1.2 dsRNA specificity

dsRNA specificity of the dsRBD is achieved via recognition of the A-form helix and specific interactions with the 2'-OH group of the ribose sugar.

Due to the rigid structure of the domain, the distance between region 2 and 3 is fixed and corresponds to the spacing between minor and major groove of A-form dsRNA [18]. In addition, the KKAxK motif in region 3 contacts the phosphodiester backbone of both RNA strands across the major groove, thereby probing its width. Since the major groove of B-form dsDNA is broader than of A-form dsRNA (17 vs 10Å), these contacts also confer specificity for dsRNA.

Contacts with ribose 2'-OH are predominantly made by region 1 and 2 in the minor grooves as discussed above. In addition, almost no electrostatic interactions are involved in dsRNA binding of the dsRBDs, precluding unspecific dsDNA binding [8].

2.3.1.3 Sequence dependence of dsRBD binding

One possible influence of RNA sequence on dsRBD binding includes bulges in the double strand and with them, a distortion of the typical A-form helix. Such bulges in the dsRNA interfere with dsRBD binding [8], but dsRBDs are also able to straighten dsRNA that is kinked due to a bulge [122]. Mismatches that do not influence the overall A-form geometry of the dsRNA helix are also tolerated by the dsRBD [9]. Since dsRBD binding is restricted to one face of the helix, it might be able to avoid structural impediments by binding to the less challenged side [101].

dsRBD binding in general is assumed to be independent of nucleotide sequence, but multiple examples exist that exhibit a high degree of sequence specificity, including the bacterial RNaseIII dsRBD, Drosophila Staufen protein, or the adenosine deaminases that act on RNA (ADARs). In addition to structural imperfection of the dsRNA helix and sequence specific contacts with bases in the major groove, which bulges or internal loops have made accessible [111], high resolution structures have revealed that even though sequence recognition via the base pair edges in the minor groove is uncommon, some specific contacts can be made here to increase the overall sequence specificity [99, 33].

2.3.1.4 Concerted action of multiple dsRBDs

Most dsRBPs encompass several dsRBDs that co-operate in binding to different extents, ranging from predominant binding of only one domain, with the other present to fine-tune or enhance the interaction [102, 74], to simultaneous binding of two domains, resulting in a huge

increase in binding affinity [116]. Whereas the structures of individual dsRBDs have been discussed profoundly, the interplay of two dsRBDs on the structural level has only been considered for a handful of proteins [99, 97, 74]: Plant HEN1 binds dsRNA via two dsRBDs, with dsRBD1 making more extensive contact with the RNA, whereas dsRBD2 shifts ~3Å away from the RNA, probably allowing for binding of bulged duplexes, and they are arranged on opposite faces of the RNA [39]. In DGCR8, the two dsRBDs contact each other via an extra C-terminal helix and adopt a pseudo two-fold symmetry, with the RNA binding sites facing away from each other so that the pri-miRNA has to bend around them [97].

The linker connecting two dsRBDs is usually flexible and ~20 - 70AA long. Linker length can be correlated with the increase in binding affinity of two compared to one domain: Binding affinity of two domains with virtually no separation should be the product of the individual affinities, whereas the binding affinities of two domains separated by an indefinitely long linker would be merely added to yield the affinity of the fusion protein. Shamoo et al. have proposed a model for the correlation of RRM linker lengths with binding affinity, stating that a separation of 60 residues renders RRM affinities fully independent [95].

In addition to cooperativity of two or multiple dsRBDs in one protein, the complete proteins can also bind in a cooperative manner: RDE-4 for example binds dsRNA cooperatively, TRBP on the other hand does not, reflecting their requirements in siRNA biogenesis, i.e. processing of dsRNA and siRNA transfer to RISC [83].

2.4 Measuring RNA-protein interactions

2.4.1 Electrophoretic mobility shift assay

In an electrophoretic mobility shift assay (EMSA), bound and free RNA species are separated via electrophoresis in a native acrylamide gel. Protein binding changes the RNA movement through the gel, usually retarding it and thereby shifting the RNA band to a higher molecular weight. This is mainly due to an increased size of the complex. Furthermore, both charge and form of the complex can have an impact on the RNA migration behavior. The binding reaction is no more in an equilibrium after separation of bound and unbound RNA, since each species strives to reach equilibrium again in the created subenvironment. Therefore the interaction has to be stable enough to be maintained nevertheless. To improve complex stability, EMSAs are mostly run in a buffer of low ionic strength. In addition, the so-called caging effect of the gel matrix limits diffusion of the protein away from the RNA, which increases the on-rate and therefore prevents complex dissociation to a certain extent. To visualize the RNA, a radioactive or fluorescent label has to be attached.

2.4.2 Anisotropy

In anisotropy measurements, a fluorophore is being excited with linearly polarized light. The polarization change between exciting and emitted light and can be measured and hence, the state of the molecule that is being coupled to the fluorophore can be determined.

Excitation occurs when a photon with sufficient energy interacts with a fluorophore electron and by this energy transfer, the electron is shifted from the ground electronic state to the first excited electronic state. Excess energy from this process is converted into vibrational energy and therefore, the electron occupies a higher vibrational state. Each electronic state is split into several vibrational energy states, as visualized in the Jablonski diagram in figure 2.4. After photon absorption, the electrons relax by non-radiative transitions from the higher to the lowest vibrational state (vibrational relaxation). Transition from the lowest vibrational state of the first excited electronic state to the vibrational states of the ground state results in photon emission, which, due to the relaxation, has a longer wavelength than the exciting photon, a phenomenon called Stoke's shift.

These three important steps in fluorescence, excitation, relaxation and emission, happen at different timescales: excitation occurs in the order of femtoseconds, relaxation happens in picoseconds and fluorescence in the order of nanoseconds. Due to relaxation and a residence

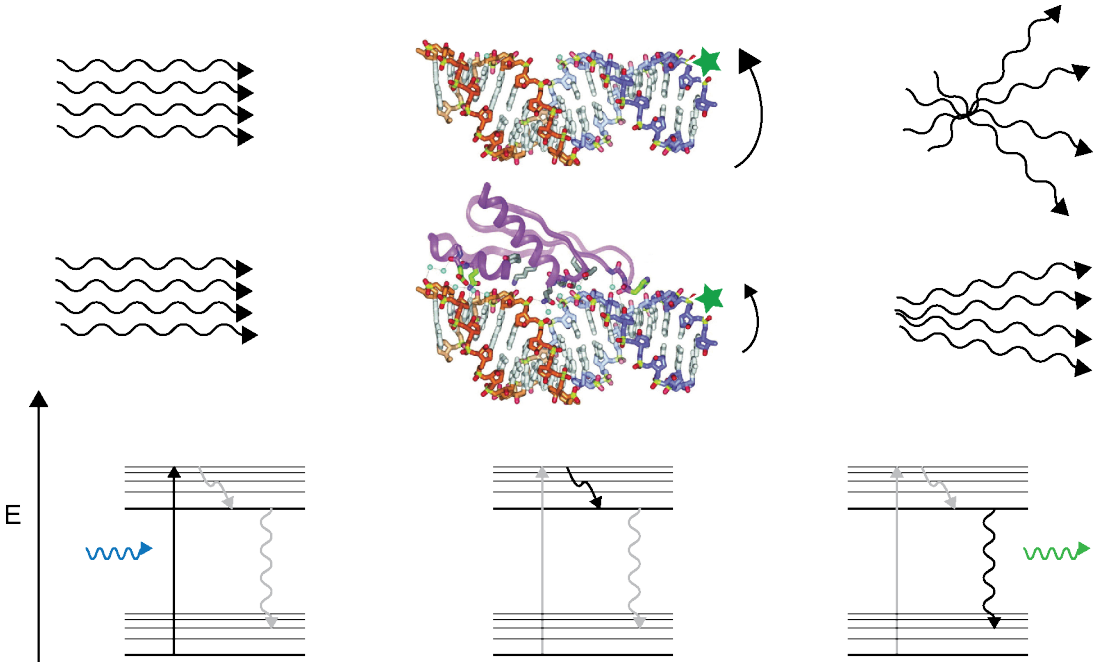

Figure 2.4: Principle of Fluorescence Anisotropy measurements
Between excitation (left) and emission (right) the fluorophore tumbles in solution, and the direction of light polarization is changed (upper row). The Jablonski diagram depicts the states of the fluorophore at the respective time points (lower row).

time of the molecule in the lowest vibrational state of the first excited electronic state, there is a time delay between excitation and emission of the fluorescent photon, which is called fluorescence lifetime.

For excitation to occur, the incident polarized light has to align with the excitation dipole of the fluorophore, so that in a population of randomly oriented fluorophores only the subpopulation with the proper orientation is excited. Emission occurs in the direction of the emission dipole, and in the theoretical case of immobilized fluorophores, the emitted light would be linearly polarized as well. In solution, the fluorophore tumbles in the time delay between excitation and emission, and the polarization of the emitted light is no more completely linear. The tumbling intensity and with it the Anisotropy of the emitted light depends on the size of the molecule the fluorophore is attached to: a larger complex tumbles more slowly than a smaller one, resulting in a smaller decrease of linear polarization of the emitted light and a higher Anisotropy. This way, free and bound RNA labeled with a fluorophore can be distinguished in solution.

2.4.3 Thermophoresis

Thermophoresis, also called Soret effect, describes the movement of molecules along a temperature gradient. In a gaseous environment, this phenomenon can be explained as follows: The gas molecules have acquired different amounts of kinetic energy, depending on the distance to the heat source. The higher energetic molecules close to the heat source push larger particles down the temperature gradient via repeated collisions. Gas molecules pushing from the cold side do not have as much energy, and the resulting net force on the larger particle leads to their movement along the gradient. Thermophoresis of a particle is defined as positive when it moves from the hot to the cold region, and negative in the opposite direction. In gases, larger particles typically show a positive thermophoresis. In aqueous solutions, the mechanism is less well understood. Direction and velocity of the movement depends on the the entropy of ionic shielding, the entropy of hydration and the size of the molecule [26]. Since these parameters are sensitive to even subtle changes of a molecule, binding reactions can be measured via the change in thermophoretic behavior they induce. In the experimental setup, the binding reaction is contained in a capillary, and a temperature gradient is set up via an infrared laser. At this position, the movement of one binding partner upon induction of the temperature gradient is followed via an attached fluorophore. The resulting displacement curve is depicted in figure 2.5. Three different regions give insight into the state of the molecule under observation: In region 1, the so-called temperature jump takes place, which comprises a drop in fluorescence due to changes in its direct vicinity caused by the increase in temperature. Movement of the labeled molecule due to the thermophoretic force happens at a slower timescale and can be seen in region 2. After a while, a steady state is reached

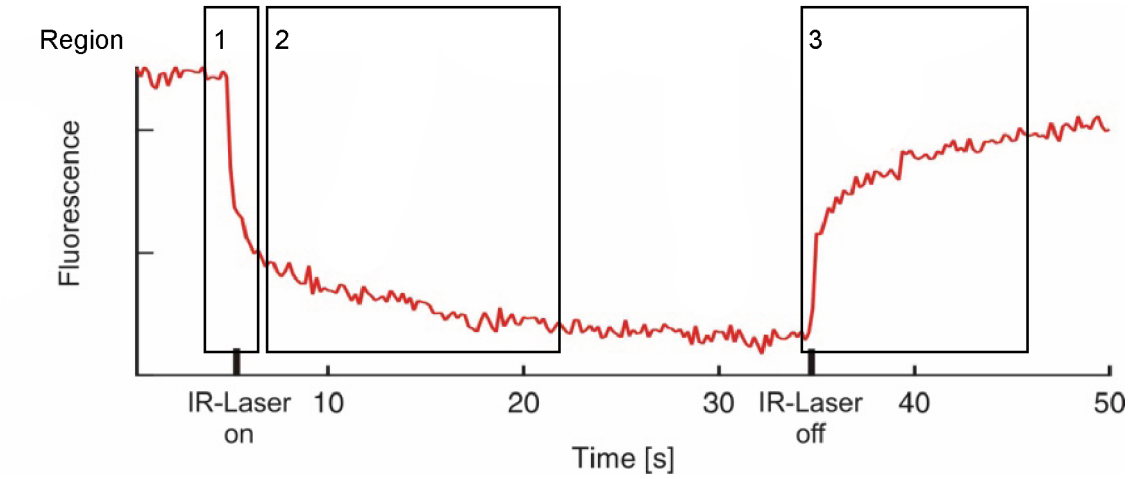

Figure 2.5: Fluorescence trace in thermophoresis experiments
Shape of a typical fluorescence distribution in the IR laser region over time. Region 1 - 3 depict the phenomena of temperature jump, thermophoresis, and backdiffusion, respectively.

where diffusion counteracts the thermophoretic movement. When the laser is switched off, the molecule under observation diffuses back in the previously depleted region, as can be seen in region 3. From each of these three regions, a binding curve can be calculated.

3 Aims of this Thesis

The dsRBPs LoqsPD and R2D2 are both involved in siRNA biogenesis by assisting Dcr-2 in dsRNA processing and, at least R2D2, in the RISC loading step. They resemble each other closely, both contain two dsRBDs and interact with Dcr-2 via their C-terminus, and both are able to increase the dsRNA processing efficiency of Dcr-2. Nevertheless, they are not redundant in their function: Most dsRNAs depend on LoqsPD for the processing and R2D2 for the loading step, but in the absence of either protein, processing and loading of a subset of siRNAs still occurs. A general feature that determines this dependence on the different dsRBPs is missing.

In my thesis I wanted to elucidate the differential requirements for LoqsPD and R2D2 by characterizing their behavior in the cell. One possibility might be a different RNA affinity of the two dsRBDs, which might distinguish low from high abundant siRNA species. Therefore, I measured the RNA binding affinity of LoqsPD and tried to elucidate how specificity it is achieved on the molecular level. R2D2 is more difficult to handle than LoqsPD, but an attempt was made to characterize its binding behavior and compare it to that of LoqsPD. In addition I wanted to see whether LoqsPD is also able to function in the RISC loading step.

4 Material and Methods

4.1 Material

4.1.1 Plasmids

Plasmids for recombinant protein expression

Name	Tag	Selection Marker	Comment
pET LoqsPB	His_6	Kan	
pET LoqsPD	His_6	Kan	
pET R2D2	His_6	Kan	
pET-M11 LoqsPD dsRBD1	His_6, ZZ-tag	Kan	AA129-211
pET-M11 LoqsPD dsRBD2	His_6, ZZ-tag	Kan	AA245-322
pET-M11 LoqsPD DeltaNC	His_6, ZZ-tag	Kan	AA129-322
pET-M11 LoqsPD full length	His_6, ZZ-tag	Kan	
pET-M11 R2D2 dsRBD1	His_6, ZZ-tag	Kan	AA1-74
pET-M11 R2D2 dsRBD2	His_6, ZZ-tag	Kan	AA92-167
pET-M11 R2D2 DeltaNC	His_6, ZZ-tag	Kan	AA1-167
pET SUMO LoqsPB	His_6, SUMO	Kan	
pET SUMO LoqsPD	His_6, SUMO	Kan	
pFastBac Dcr-2	His_6	Amp	
pFastBac Dcr-1	His_6	Amp	
pGex LoqsPB	GST	Amp	
pGex LoqsPD	GST	Amp	
pGex R2D2	GST	Amp	

Name	Tag	Selection Marker	Comment
pMBP R2D2	MBP	Amp	
pPAL7 LoqsPB	Profinity eXact	Amp	
pPAL7 LoqsPD	Profinity eXact	Amp	
pPAL7 R2D2	Profinity eXact	Amp	

Plasmids for expression in S2 cells

Name	Insert	Backbone	tag
pCasper Dcr2	Dcr2	pC5T	Flag
Dcr2 Delta Hel	Dcr2 AA 542 - end	pC5T	Flag
pCasper Dcr2 Hel 1	Dcr2 AA 22 - end	pC5T	Flag
pCasper Dcr2 Hel 2	Dcr2 AA 173 - end	pC5T	Flag
pCasper Dcr2 Hel 3	Dcr2 AA 386 - end	pC5T	Flag
pCasper LoqsPD	LoqsPD	pCasper5	myc
pCasper R2D2	R2D2	pCasper5	myc
pKF63	GFP	pCasper5	myc
RB2	Firefly Luciferase	pC5T	Flag

4.1.2 Primer

Name	Sequence 5’ - 3’
seq-Dcr2-forw1	ATGGAAGATGTGGAAATCAAGCC
seq-Dcr2-forw2	ATGCCTGCGGGATCCGTCG
seq-Dcr2-forw3	CGTGTGAGTACACGGAACACATG
seq-Dcr2-forw4	GTATTTAAACCTCCATCTGCTGCC
seq-Dcr2-forw5	CCCCACAATGTTCTAGCCTTG
seq-Dcr2-forw6	CGAGGAAGCTGATGTTCAGCC
seq-Dcr2-rev1	GGCACTCGTCTATGATGACAACG

Name	Sequence 5’ - 3’
seq-Dcr2-rev2	CAAGGCGTGGCAGCTTGTTC
seq-Dcr2-rev3	TTAGGCGTCGCATTTGCTTAGC
Dcr2-quichange-a	GCATCACGTTTTCCATCTCCTTGGTGTCGGAC
Dcr2-quichange-s	GTCCGACACCAAGGAGATGGAAAACGTGATGC
D10 Bacmid 5’ fw	GTTTTCCCAGTCACGAC
D10 Bacmid 3’ rv	CAGGAAACAGCTATGAC
Dcr2 AA22_Not	CAAGCGGCCGCATGGATTATAAAGATGATGATGA TAAAagcaatggcattgtctacctgccc
Dcr2 AA173_Not	CAAGCGGCCGCATGGATTATAAAGATGATGATGA TAAAggggtgttaatcaagggaaatgaaa
Dcr2 AA386_Not	CAAGCGGCCGCATGGATTATAAAGATGATGATGA TAAAatctgctgtctggtctttgttgaac
Dcr2 mitte_xba rv	TCTAGAACTCGATCCTTTAAATATTCGGC
R2D2_RBD1_fw_nco	tcccatgggaATGGATAACAAGTCAGCCG
R2D2_RBD1_rv_not	taGCGGCCGCttaGTGTATGCCGGGCAG
R2D2_RBD2_fw_nco	tcccatgggaCTCAACCGGGACATGGT
R2D2_RBD2_rv_not	taGCGGCCGCttaATTGTCCGAATTGCTGGA

4.1.3 RNA Oligos

All oligos were 5’ phosphorylated, except bantam base nicked b, bantam si nicked b, and bantam 14nt as. *: fluorescein position, u+ : thiouridine

Name	Sequence	Manufacturer
bantam base	ugagaucauuuugaaagcugau*u	MWG, Dharmacon
bantam si	ucagcuuucaaaaugaucucacu	MWG, Dharmacon
bantam mi	ucgguuuucgauuugguuugacu	MWG, Dharmacon
bantam pre-mi	ucgguuuucgauuugguuugacuguuuuucauacaag	Dharmacon
bantam dsRNA s	gauucauacaagugagaucauuuugaaagcugau*u	Dharmacon
bantam dsRNA as	ucagcuuucaaaaugaucucacuuguaugaaucaa	Dharmacon

Name	Sequence	Manufacturer
bantam 14nt s	ugagaucauuuuga	Dharmacon
bantam 14nt as	u*caaaaugaucucauu	Dharmacon
bantam base nicked a	ugagaucauuuu	MWG
bantam base nicked b	gaaagcugauu	MWG
bantam si nicked a	ucagcuuucaa	MWG
bantam si nicked b	aaugaucucacu	MWG
miR-8 base	aauacugucagguaaagaugu*c	Dharmacon
miR-8 si	caucuuuaccugacaguauuaga	Dharmacon
miR-8 mi	caucuuaccgggcagcauuaga	Dharmacon
Crosslink oligo 1	u+gagaucauuuugaaagcugauu	Dharmacon
Crosslink oligo 2	ucagcuuucaaaaugaucu+cacu	Dharmacon
Crosslink oligo 3	ugagau+cauuuugaaagcugauu	Dharmacon
Crosslink oligo 4	ugagaucau+uuugaaagcugauu	Dharmacon
Crosslink oligo 5	ugagaucauuuu+gaaagcugauu	Dharmacon
Crosslink oligo 6	ucagcu+uucaaaaugaucucacu	Dharmacon
Crosslink oligo 7	ugagaucauuuugaaagcu+gauu	Dharmacon
Crosslink oligo 8	u+cagcuuucaaaaugaucucacu	Dharmacon

4.1.4 Hardware

Agarose gel running chamber	Carl Roth GmbHM; Karlsruhe, Germany
BioLogic™ Low-Pressure Chromatography Systems	BioRad
Bio-photometer	Eppendorf
CL-1000 Ultraviolet Crosslinker	UVP
FACSCalibur flow cytometer	Becton, Dickinson; Franklin Lakes, USA
Infinite® M1000 Plate Reader	Tecan
INTAS UV Imaging System	INTAS; Göttingen, Germany
LAS 3000 mini Western Imager	Fujifilm; Tokyo, Japan

Leica TCS SP2 confocal microscope	Leica Microsystems; Wetzlar, Germany
Monolith, pre-serial model	Nanotemper
Nanodrop1000 spectrometer	Scientific
PAGE electrophoresis gel chamber	BioRad; Hercules, USA
Power supply	BioRad; Hercules, USA
Semi-dry blotter	BioRad; Hercules, USA
SterilGARD cell culture workbench	The Baker Company; Sanford, USA
Table top centrifuge (5417R and 5415R)	Eppendorf AG; Hamburg, Germany
Tank-blotting chamber	BioRad; Hercules, USA
Thermocycler	Sensoquest; Göttingen, Germany
Typhoon 9400 Variable Mode Imager	GE Healthcare; Freiburg; Germany
Vi-Cell XR Cell Viability Analyzer	Beckman Coulter

4.1.5 Chemicals

Acrylamide 40%	Carl Roth GmbH; Karlsruhe, Germany
Agarose	Biozym Scientific GmbH; Oldendorf, Germany
Ampicillin	Carl Roth GmbH; Karlsruhe, Germany
Ammonium peroxodisulfate (APS)	Carl Roth GmbH; Karlsruhe, Germany
Bacto Agar	Becton, Dickinson; Franklin Lakes, USA
Bradford Assay reagent	BioRad; Hercules, USA
Bovine serum albumin (BSA)	New England Biolabs; Ipswich, USA
Chloroform	Merck Biosciences GmbH; Schwalbach, Germany
Complete Protease Inhibitor tablets	Roche Diagnostics; Mannheim, Germany
Coomassie G250	Carl Roth GmbH; Karlsruhe, Germany
Dimethyl sulfoxide (DMSO)	Carl Roth GmbH; Karlsruhe, Germany
Dithiothreitol (DTT)	Carl Roth GmbH; Karlsruhe, Germany
Ethanol (p.a.)	Merck Biosciences GmbH; Schwalbach, Germany
Fetal bovine serum (FBS)	Thermo Fisher Scientific; Waltham, USA

Fugene HD transfection reagent	Roche Diagnostics; Mannheim, Germany
H_2O HPLC quality	VWR; Ismaning, Germany
Hepes	Carl Roth GmbH; Karlsruhe, Germany
Isopropanol (p.a.)	Merck Biosciences GmbH; Schwalbach, Germany
Kanamycin	Carl Roth GmbH; Karlsruhe, Germany
L-Glutathione, reduced	Sigma Aldrich; Taufkirchen, Germany
Methanol (p.a.)	Merck Biosciences GmbH; Schwalbach, Germany
Powdered milk	Rapilait Migros; Zürich, Switzerland
Roti Aqua Phenol/C/I	Carl Roth GmbH; Karlsruhe, Germany
Sodium dodecyl sulphate (SDS)	Merck Biosciences GmbH; Schwalbach, Germany
Syber Safe/Gold	Invitrogen; Karlsruhe, Germany
TEMED	Carl Roth GmbH; Karlsruhe, Germany
Triton X-100	Sigma Aldrich; Taufkirchen, Germany
Tween 20	Carl Roth GmbH; Karlsruhe, Germany
[γ32P] ATP (SRP 501)	Hartmann Analytic; Braunschweig, Germany

All other standard laboratory chemicals were purchased from the Gene Center in house supply (NaCl, KAc, Tris, ...)

4.1.6 Enzymes

DNase I, RNase free	Fermentas; St. Leon-Rot, Germany
Polynucleotidekinase (PNK)	Fermentas; St. Leon-Rot, Germany
Proteinase K	Fermentas; St. Leon-Rot, Germany
Phusion Hot Start DNA Polymerase	Finnzyme
Pfu DNA Polymerase	Fermentas; St. Leon-Rot, Germany
Taq DNA Polymerase	laboratory stock
T7-polymerase	laboratory stock
Restriction enzymes BamHI, BglII, NotI, XbaI, ...	New England Biolabs; Ipswich, USA

4.1.7 Buffers and solutions

Colloidal Coomassie staining solution	50 g/l aluminum sulfate 2% (v/v) H_3PO_4 (conc.) 10% (v/v) 100% ethanol 0.5% (v/v) Coomassie G250 stock solution
Coomassie G250 stock solution	0.5 g/l Coomassie G250 in 100% methanol
Coomassie staining solution	45% (v/v) methanol 10% acetic acid 0.25% (w/v) Coomassie Brilliant Blue
Coomassie destain	45% (v/v) methanol 10% acetic acid
DNA loading buffer (6x)	0.25% (w/v) bromophenol blue 0.25% (w/v) xylene cyanol 30% (w/v) glycerol
Formamide loading dye (2x)	80% (w/v) formamide 10 mM EDTA, pH 8.0 1 mg/ml xylene cyanol 1 mg/ml bromophenol blue
Laemmli SDS loading buffer (2x)	100 mM Tris/HCl, pH 6.8 4% (w/v) SDS 20% (v/v) glycerol 0.2% (w/v) bromophenol blue 200 mM freshly added DTT
Lysis buffer for protein extraction	100 mM KOAc 30 mM Hepes 2 mM $MgCl_2$ 1 mM DTT 1% (v/v) Triton X-100 2x Complete® without EDTA (=protease inhibitor cocktail)
PBS (10x)	137 mM NaCl 2.7 mM KCl 10 mM Na_2HPO_4 2 mM KH_2HPO_4, pH 7.4

RNA gel extraction buffer	0.4 M NaCl 0.5% SDS 50 mM Tris-HCl pH 8
SDS-running buffer (5x)	125 mM Tris/HCl, pH 7.5 1.25 M glycine 5% SDS
TAE (50x)	2 M Tris-base 5.7 1% acetic acid 100 mM EDTA
TBE (10x)	0.9 M Tris base 0.9 M boric acid 0.5 M EDTA (pH 8)
TB (10x)	0.9 M Tris base 0.9 M boric acid
TBS (10x)	50 mM Tris 150 mM NaCl pH 7.4
Western blotting stock (10x)	250 mM Tris/HCl, pH 7.5 1.92 M Glycine
Western blotting buffer (1x)	10% Western blotting stock (10x) 20% Methanol

4.1.8 Bacterial strains and media

4.1.8.1 Strains

XL2-blue	Plasmid amplification
BL21 Gold (DE3; pLys S)	Recombinant protein expression
D10	Baculovirus production
Rosetta	Recombinant protein expression, includes rare tRNAs
ArcticXpress	Recombinant protein expression at low temperatures

4.1.8.2 Media

LB medium	1% (w/v) Tryptone 0.5% (w/v) yeast extract 1% (w/v) NaCl pH 7.2
SOC-medium	0.5% (w/v) yeast extract 2% (w/v) Tryptone 10 mM NaCl 2.5 mM KCl 10 mM $MgCl_2$ 10 mM $MgSO_4$ 20 mM Glucose pH 7.0

4.1.9 Cell lines and Media

Cell line	Organism	Application	Growth medium
S2 B2	*Drosophila melanogaster*	Cell culture experiments	Schneider's medium, Bio&Sell
67-1D	*Drosophila melanogaster*	Cell culture experiments	Schneider's medium, Bio&Sell
63N1	*Drosophila melanogaster*	Cell culture experiments	Schneider's medium, Bio&Sell
Sf21	*Spodoptera frugiperda*	Recombinant protein expression	Sf 900III SFM (1x), Gibco
H5	*Trichoplusia ni*	Recombinant protein expression	ExpressV SFM (1x), Gibco

4.1.10 Antibodies

Epitope	Species	Dilution	Distributor
α R2D2	rabbit, polyclonal	1:5000	abcam
α LoqsPD	rabbit, polyclonal	1:5000	laboratory stock
α Dcr-2	rabbit, polyclonal	1:1000	abcam
α Flag	mouse, monoclonal	1:1000	sigma
α rabbit IgG	goat	1:50000	Pierce (Thermo Scientific)
α mouse IgG	goat	1:50000	Pierce (Thermo Scientific)

4.2 Methods

4.2.1 Molecular Cloning

4.2.1.1 Primer design

For Dcr-2 truncation plasmids, the helicase domain was sub-divided according to NCBI conserved domain search into its DEXDc and HELICc subdomains. The helicase domain was truncated stepwise from the N-terminal end. Primers were designed to amplify the fragments and to introduce the appropriate restriction sites.

To test for successful recombination of pFastBac sequences into D10 bacmid DNA, 5' forward and 3' reverse bacmid primers which anneal adjacent to the recombination sites in the viral DNA were ordered as published [118].

To straighten out a L208M mutation in the Dcr-2 plasmid, primers for site-directed mutagenesis were designed, which are complementary to each other, anneal with 10-15 nucleotides 5' and 3' of the mutated nucleotide and contain the correct nucleotide in the middle.

For R2D2 dsRBD and DeltaNC constructs, the domains were identified via Homology detection & structure prediction by HMM-HMM comparison [96], and primers for amplification of the domains and introduction of restriction sites were designed.

Sequencing primers were designed to cover an estimated sequencing length of 700bp.

Primers were ordered from eurofins MWG.

4.2.1.2 Polymerase chain reaction (PCR)

The standard PCR reaction contained

- 10-50ng DNA template
- 0.2mM of forward and reverse primer
- 0.2mM of each dNTP
- 1x polymerase buffer
- 1.5mM $MgCl_2$
- 0.05U Polymerase

DNA sequences which were subsequently used for cloning were usually amplified in a 50µl mix, using pfu polymerase with proofreading activity to avoid mutations. Colony PCRs served to test bacterial colonies for successful transformation with the correct plasmid, these were carried out in a 10µl reaction. Inoculation with colonies served as DNA template, and Taq polymerase was used for fast amplification. The temperature profile shown in figure 4.1 was followed and adjusted according to melting temperatures of the primers and the length of the amplified sequence.

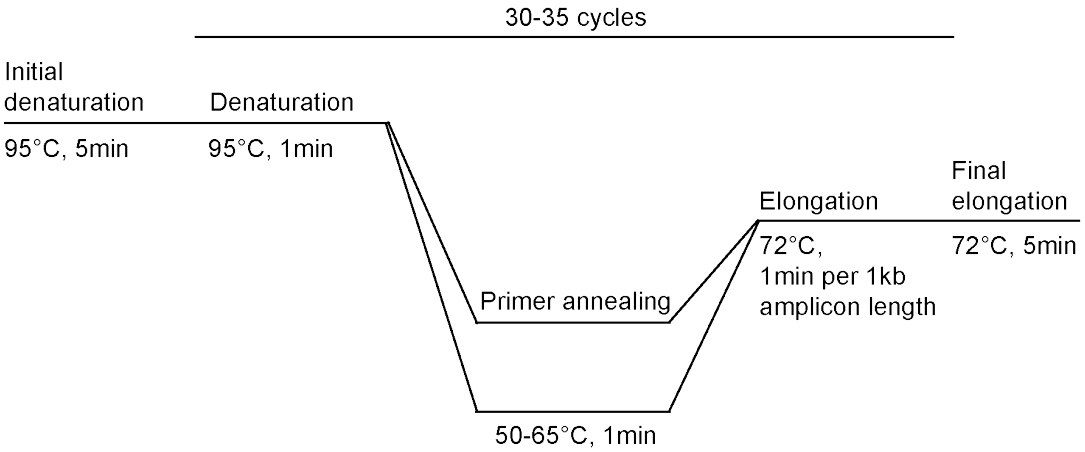

Figure 4.1: PCR temperature profile
Temperature steps in a conventional PCR: After denaturation, the temperature is decreased to allow for primer annealing. For subsequent elongation, the temperature is raised to 72°C. This cycle is repeated about 32 times.

4.2.1.3 Agarose gel electrophoresis

DNA was analyzed on 0.5 - 1.5% agarose gels including 1x SYBR green, depending on fragment size. DNA samples were mixed with DNA loading buffer and loaded, together with the DNA ladder mix for size comparison. Gels were run at 50V for 30min in 1x TAE buffer. The gels were recorded at an Intas UV imaging system; for DNA purification bands were cut out on a blue light transilluminator. Gels with poor contrast were restained in 1x SYBR gold in 1x TAE for 5min.

4.2.1.4 Restriction digestion and ligation of DNA fragments

Digestion of DNA was carried out by restriction endonucleases according to manufacturers' recommendation, usually at 37°C for 1h for analytical, or overnight for preparative digestions. For ligation, the digested vector backbones and inserts were gel purified and combined in a molar ratio of 1 to 6 in the following mix:

$\geq$ 200ng vector backbone

required amount of insert

1x T4 Ligase buffer

20U T4 Ligase

Ligation was carried out in 20µl volumes for 2h at RT or overnight at 18°C.

4.2.1.5 Transformation in E.coli

For DNA preparation, the XL2-blue E.coli strain was used. For transformation of $CaCl_2$ competent bacteria, 50µl aliquots were thawed on ice and 10µl ligation reaction or 1µl plasmid DNA was added. After incubation on ice for 30min, the cells were subjected to a 42°C heat shock for 1min and again cooled down on ice. Antibiotic resistance present on the transformed plasmids was allowed to develop during 1h shaking at 37°C in SOC medium, then the cells were streaked on agar plates containing the appropriate antibiotic for selection of positive transformants.

4.2.1.6 Preparation of plasmid DNA

Mini- or Midi-preps of plasmid DNA were made from 2 or 50ml overnight culture, respectively, using the appropriate Qiagen Kits and following the included protocols. The DNA was stored in H_2O at -20°C.

4.2.1.7 Sequencing and analysis of results

For sequencing, 100–200 ng/µl sample DNA was prepared in a 15µl volume containing 2µM sequencing primer and sent to Eurofins MWG. The resulting sequences were analyzed using the BioEdit or ApE software.

4.2.1.8 Site directed mutagenesis

To introduce the new nucleotide present in the mutagenesis primers as described in 4.2.1.1, a standard PCR reaction with pfu polymerase was carried out with the following deviations: Since the two primers anneal on the DNA directly opposite of each other, not only a fragment but the whole plasmid was amplified, and the amplification time had to be increased significantly compared to the usual DNA fragments. Also, the temperature during amplification was lowered to 68°C and the cycle number was reduced to 16. After amplification, the methylation dependent restriction enzyme Dpn1 was added and the reaction was incubated at 37°C for 90min. The unmutated, methylated plasmid was digested, whereas the remaining mutated, unmethylated plasmid is resistant to digestion and was subsequently transformed in E.coli. Correct plasmids were identified via sequencing.

4.2.2 Protein techniques

4.2.2.1 Protein lysates

For protein lysates, cells were harvested via mild centrifugation (2500 x g, 5min) and washed with 1x PBS. The cell pellet was resuspended in Lysis Buffer (30mM Hepes, pH7.4, 100mM KAc, 2mM $MgCl_2$, 1mM DTT) supplemented with Protease inhibitor and frozen in liquid nitrogen.

4.2.2.2 Denaturing SDS-Gelelectrophoresis

Protein samples were analyzed on 8 - 15% polyacrylamide gels, depending on the size range of interest [49]. Samples were mixed with SDS loading buffer, heated at 95°C for 5min and 10-40µg total protein was loaded. For size estimation, the un- or prestained protein marker was included. Gels were run in 1x SDS Running buffer at 150V for 60 - 90min. Afterwards, they were washed in H_2O for approximately 30min and stained with colloidal coomassie staining solution.

4.2.2.3 Western Blotting

When the protein gels were supposed to be used for western blotting, the staining step described in 4.2.2.2 was omitted. Instead, the gels were assembled in a blotting chamber together with a hydrated polyvinylidenfluoride (PVDF) membrane. The proteins were transferred onto this membrane by blotting at 100V for 1h in 1x Western solution. After transfer and additional immobilization of the proteins on the membrane, the membranes were blocked in 5% milk in PBS, 0.05% Tween for rabbit or TBS, 0.02% Tween for mouse antibodies for 30min at RT. The blocked membranes were incubated with the primary antibody over night at 4°C. Antibody dilutions are summarized in 4.1.10. After three 10min washing steps, the membranes were incubated with the secondary antibody for 4h at RT. Again, the membranes were washed and then immersed in Enhanced chemiluminescence (ECL) substrate. For readout the LAS3000 mini Western Imager System was used.

4.2.2.4 Determination of protein concentration

To determine the concentration of protein mixtures, 1-5µl were mixed with Bradford reagent and incubated for 5min at RT. Afterwards, the concentration was measured via the absorption at 600nm. For recombinant proteins the extinction coefficient was known, so they were quantified directly via their absorption at 280nm. To validate these measurements, recombinant protein concentration was also determined using the Bradford assay and densiometric analysis of PAA gels. In both methods BSA served as the standard.

4.2.2.5 Co-Immunoprecipitation

Cell lysates were prepared as described in 4.2.4.4. The following steps were performed in the cold room at 4°C. All samples in one experiment were diluted to the same concentration in a volume of 300µl. For pre-clearing, the protein samples were incubated with 20µl Protein G agarose beads equilibrated in Hepes Lysis buffer for 1h. Beads and lysate were then separated via centrifugation at 13000rpm for 1min, 10% of the protein were retained as input control and 40µl equilibrated agarose beads carrying the appropriate antibody were added to the rest of the lysate. Binding was allowed to occur for 1h, then the beads were separated from the lysate via spin columns. Corresponding to the input control, a sample was kept as the flowthrough. The beads were washed twice with Hepes Lysis buffer with 1% Triton and once with unsupplemented Hepes Lysis buffer. For elution, the beads were covered in 2x SDS loading buffer without DTT and boiled for 5min at 95°C and centrifuged at 4000rpm for 1min. After adding DTT, the eluate was loaded on a SDS gel and analyzed via western blotting.

4.2.2.6 Mass spectrometry

For Mass spectrometry analysis, bands of interest were cut out of the protein gel and sent to the Core facility of the Biomedical Center at the Ludwig-Maximillians-University of Munich, where LC MS/MS analysis was performed. Results were blasted using the NCBI BLAST tool.

4.2.3 RNA techniques

4.2.3.1 Oligo design

For binding and crosslinking experiments, RNA oligos were ordered from MWG eurofins and Thermo scientific. Their sequences were derived from the miRNAs bantam and miR-8. They are summarized in section 4.1.3 and could be combined to form the canonical miRNA/miRNA* duplex and the pre-miRNA hairpin, but also mimics of the siRNA duplex and its dsRNA precursor. Oligos were ordered unmodified at first, but since phosphorylation did not have an optimal yield (see 4.2.3.6), later oligos were ordered already 5' phosphorylated. For detection in all fluorescence-based assays, a Fluorescein was attached at the second nucleotide from the 3'end of the bantam or miR-8 sequence.

In the two strands of the siRNA mimic, 4-S-Uridines were inserted at 8 different positions along the duplex for site-specific crosslinking (see section 4.1.3).

4.2.3.2 General RNA handling

RNA samples were stored in small aliquots at -80°C and were kept away from light sources to avoid bleaching of the fluorophore or premature crosslinking of the 4-S-Uridines. For pipetting,

RNase free filter tips were used, and all buffers used were prepared with RNase free water and chemicals.

2'-ACE protected RNA oligos were deprotected according to manufacturer's instructions.

4.2.3.3 Denaturing Urea gel electrophoresis

RNA was analyzed on 10 - 20% Sequagel Acrylamide/Urea gels, depending on the size range of interest. Samples were mixed with Formamid loading buffer, heated at 95°C for 5min and 1-5µg RNA was loaded. For size estimation, the miRNA or dsRNA marker was included. Gels were run in 1x TBE at 250V for 60 - 90min. Afterwards, they were stained with 1x SYBR Gold in 1xTBE.

4.2.3.4 RNA extraction

To extract RNA from heterogenous samples, they were treated with 0.5U/µl Proteinase K for 30-60min at 37°C. The RNA was then separated from proteins and other contaminants by adding an equal volume of phenol/chloroform and precipitating the RNA from the aqueous phase with isopropanol. If the RNA was not about to be used in binding experiments, 30µg glycogen was added for better precipitation.

4.2.3.5 RNA gel purification

RNA gel bands were cut out, shredded and incubated in 500µl gel extraction buffer for 2h at 65°C or o/n at RT. The gel fragments were removed via spin columns, the RNA was precipitated with isopropanol and resuspended in H_2O.

4.2.3.6 RNA phosphorylation

RNA oligos were phosphorylated using Polynucleotide kinase. The following reacting mix was incubated at 37°C for 90min:

125µM RNA

5mM ATP

1x PNK Buffer A

10U/50µl T4 Polynucleotide kinase

Following phosphorylation, the RNA was extracted and gel purified as described in 4.2.3.4 and 4.2.3.5. Usually, about 30% of the inserted RNA could be recovered. Successful phosphorylation could be verified by gel electrophoresis (figure 4.2).

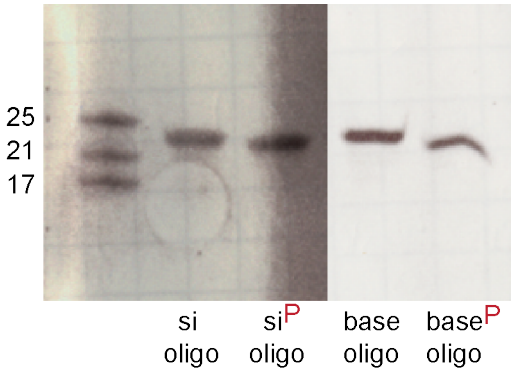

Figure 4.2: Verification of RNA phosphorylation
Phosphorylated and unphosphorylated RNA was run on a 25% Acrylamide Urea gel. The increase in negative charge due to the additional phosphate group leads to faster migration of the phosphorylated oligos.

4.2.3.7 Ligation of pre-miRNA hairpin

To generate the pre-miRNA hairpin for binding experiments, the bantam sequence carrying the fluorophore had to be ligated to the rest of the hairpin. For that, 800pmol of both oligos were mixed, heated for 5min at 95°C and slowly annealed. 10U T4 RNA ligase together with its buffer were added and ligation was carried out for 3h at 37°C. After that the RNA was extracted and gel purified as described in 4.2.3.4 and 4.2.3.5, with a yield of approximately 10%.

4.2.3.8 Radioactive labeling of RNA

For Crosslink- and some EMSA experiments, RNA was labeled radioactively. For that, one RNA strand was phosphorylated with γ-32 ATP as described in 4.2.3.6, and left-over ATP was disposed of via mini quick spin columns (Roche). Assuming a yield of 80%, the RNA was mixed with a 1.2 fold excess of the complementary strand, heated to 95°C for 5min and and the mixture was cooled down for annealing of the two strands. Labeling efficiency and recovery were analyzed on a 15% SDS acrylamide gel, which was dried and used to expose a phosphoimager screen, which was then recorded in a Typhoon scanner.

4.2.3.9 Production of dsRNA for gene specific knockdown

For dsRNA generation, T7-promoter sequences were introduced on both 5' ends of a ~500bp dsDNA fragment via PCR. This DNA was in vitro transcribed o/n at 37°C in 100µl of the following reaction mix:

25µl DNA

5mM DTT

5mM of ATP, CTP, and UTP

10mM GTP

1x T7 buffer

2µl T7 polymerase

After transcription, 2U DnaseI were added and incubated for 1h at 37°C. The formed white precipitate was removed via centrifugation and the RNA was precipitated from the supernatant with isopropanol. The resulting pellet was resuspended in H_2O containing 5mM $MgCl_2$, and after heating on 95°C the dsRNA was slowly annealed. The RNA was loaded on an agarose gel and the concentration was determined by comparison with the standard.

4.2.4 Cell culture

4.2.4.1 Culture of *Drosophila* Schneider 2 cells

Drosophila S2 cells were grown in Schneider's medium supplemented with 10% fetal calf serum in an adhesive culture in 10cm cell culture dishes. For maintenance, the cells were split twice a week 1:10 into fresh medium.

4.2.4.2 Culture of Sf21 cells

Sf21 is a cell line derived from ovaries of the army worm *Spodoptera frugiperda*. For maintenance the cells were grown at 27.5°C in Sf900 medium supplemented with 10µg/ml gentamycine in suspension at a density from $0.6 \cdot 10^6$ to $10 \cdot 10^6$ cells/ml. They were used for baculovirus production and expression of recombinant protein.

4.2.4.3 Culture of High5 cells

High5 (H5) cells are derived from ovarian cells of the cabbage looper, *Trichoplusia ni*. They were grown in ExpressV medium supplemented with 10mM fresh glutamine and 10µg/ml gentamycine. Like Sf21, H5 cells were grown at a density from $0.2 \cdot 10^6$ to $10 \cdot 10^6$ cells/ml. They were used primarily for expression of recombinant protein.

4.2.4.4 Cell lysates

For S2 cell lysates, cells were pelleted at 2000rpm for 3min, washed twice with 1xPBS, resuspended in Hepes Lysis Buffer and frozen in liquid nitrogen. After thawing on ice, the lysate was cleared from cell debris via centrifugation and protein concentration was determined as described in 4.2.2.4.

Small volumes of Sf21 and H5 cells were used to check for protein expression. They were resuspended in Triton Lysis Buffer without EDTA, frozen in liquid nitrogen and subjected to DNase digestion to avoid smearing on the gel.

4.2.4.5 Transfection

S2 cells were transfected with plasmids containing a suitable promoter for expression in *Drosophila* cells. For one transfection, 500ng of DNA and 3µl Fugene reagent were each diluted in 50µl Schneider's medium without serum and then combined. After 30-60min, this mixture was added to the cells growing in 24well plates at a density of $0.25 \cdot 10^6$ cells per well. Cells were split after 3 days and harvested after 5-6 days.

Sf21 and H5 cells were transfected with bacmid DNA containing the baculoviral genome and the protein to be expressed. For transfection, they were seeded in 6-well plates at a density of $0.8 \cdot 10^6$ (Sf21) or $0.5 \cdot 10^6$(H5) cells per well. 5µg DNA and 3µl Fugene reagent were each diluted in 100µl of the suiting medium and then combined. After 30min, the mixture was added to the cells. After 5 days, the virus could be used for subsequent infections.

4.2.4.6 RNAi

For RNAi mediated knockdown of target genes, cells were seeded in 24-well plates at a density of $0.25 \cdot 10^6$ cells per well, and 5-10µg dsRNA (4.2.3.9) per well was added. After three days, the cells were split 1:5 and the dsRNA treatment was repeated. After 5-6 days, the cells were harvested.

4.2.5 Fluorescence Activated Cell Sorting

For FACS experiments, 100µl cell suspensions was mixed with 100µl FACS Flow and analyzed in a FACSCalibur flow cytometer.

4.2.6 Immunofluorescence

To visualize the cellular localization of proteins they were marked with a fluorophore conjugated to an appropriate antibody in the usual primary/secondary antibody approach. For

that, 200µl of a cell suspension were pipetted on a glass slide in an area confined with barrier marker, and the cells were allowed to settle. After 30min, the medium was removed and the cells were washed three times with 200µl 1xPBS. Subsequently they were fixed by adding 200µl 4% formaldehyde in 1xPBS for 15min. After that, cells were washed again three times with 1xPBS, 0.2% Triton and blocked with 5% BSA in 1xPBS, 0.2% Triton for 2h. Then the primary antibody diluted in blocking solution was added and incubated o/n at 4°C in a wet chamber. The next day, the primary antibody was removed by washing three times with 1xPBS, 0.2% Triton. The fluorescently labeled secondary antibody diluted in blocking solution was added and incubated for 2h at RT. After three washing steps, the DNA was stained with a 0.5µg/ml DAPI solution for two minutes. Again, the cells were washed three times, embedded in a droplet of DABCO and sealed with a coverslip. Stains were recorded with a Leica confocal microscope.

4.2.7 Production of recombinant protein

Whenever a gradient of wash and elution buffer was employed, a gel-filtration column was used or eluted fractions had to be collected, all steps were carried out with the help of a BioLogic low pressure chromatography system (BioRad).

4.2.7.1 Generation of competent BL21 and DH10 cells

For protein expression, the BL21 *E.coli* strain was used, which is optimized for this purpose and includes a pLys plasmid. BL21 cells were made competent for transformation via the $CaCl_2$ method. Cells were grown to an OD600 of 0.7 in 100ml LB medium supplemented with 34µg/ml chloramphenicol to avoid loosing the pLys plasmid. The culture was cooled down on ice and all subsequent steps were carried out at 4°C. The cells were pelleted at 4500rpm for 10min, resuspended in 50ml sterile, ice-cold 0.1M $CaCl_2$ solution and incubated on ice for 30min. Again, the cells were pelleted and resuspended in 4ml 0.1M $CaCl_2$ solution with 10% glycerol. Aliquots were frozen in liquid nitrogen and stored at -80°C.

DH10 cells were used for recombination and amplification of baculoviral bacmid DNA. For competent cells, cells were grown to an OD600 of 0.4 in 100ml LB medium supplemented with 50 µg/mL Kanamycin and 10 µg/mL Tetracycline. The culture was cooled down on ice and all subsequent steps were carried out at 4°C. After pelleting the cells they were resuspended in 12.5ml 30 mM KOAc, 50 mM $MnCl_2$, 100 mM KCl, 10 mM $CaCl_2$, and 15 % glycerol, pH 5.8. After 10min on ice they were centrifuged again and resuspended in 4ml 10 mM MOPS pH 7, 10 mM RbCl, 75 mM $CaCl_2$, and 15% glycerol. Aliquots were frozen in liquid nitrogen and stored at -80°C.

4.2.7.2 PreScission Protease

PreScission protease is needed to remove the GST-tag from proteins expressed from the pGex plasmid. Cells were grown to an OD600 of 0.5 at 37°C at 120rpm in 500ml LB-medium containing 50 μg/mL Ampicillin, 34μg/ml Chloramphenicol and 0.5% glucose. After induction with 0.5mM IPTG the culture was transferred to 18°C and grown o/n. On the next day the cells were pelleted, resuspended in 12ml Lysis buffer and frozen in liquid nitrogen. For purification, the lysate was cleared by centrifugation and loaded on a GST column equilibrated with Lysis buffer. After washing with 5 column volumes (CV) Wash buffer, the protease was eluted with Elution buffer. Fractions containing the protease were pooled, dialyzed in 150mM Tris-HCl pH 8.0, 150mM NaCl, 2mM DTT, 10mM EDTA, and 20% glycerol, frozen in liquid nitrogen and stored in 30µl aliquots at -80°C.

Lysis buffer	Wash buffer	Elution buffer
50mM Tris-HCl pH 7.5	50mM Tris-HCl pH 7.5	100mM Tris-HCl pH 8.5
500mM NaCl	500mM NaCl	15mM glutathione
1mM EDTA		150mM NaCl

4.2.7.3 TEV Protease

TEV protease is needed to remove the double-Z/His-tag from proteins expressed from the pET-M11 plasmid. Cells were grown to an OD600 of 1 at 37°C at 120rpm in 2l LB medium containing 30µg/ml Kanamycin and 34µg/ml Chloramphenicol. After induction with 1mM IPTG, cells were cooled down to 20°C and grown o/n. The cultures were centrifuged at 6000rpm for 20min and the pellets were resuspended in 40ml Lysis buffer supplemented with lysozyme, DnaseI and proteaseinhibitors. The lysate was sonified 3 times for 1min and cleared by centrifugation at 13000rpm for 10min. Subsequently, it was loaded on a Ni-column equilibrated with Lysis buffer, washed with Lysis buffer and Wash buffer and finally eluted with Elution buffer. Fractions containing the protease were pooled, dialyzed immediately in 50mM Tris-HCl pH 8.0, 150mM NaCl, and 50% glycerol, frozen in liquid nitrogen and stored in 30µl aliquots at -80°C.

Lysis buffer	20mM Imidazole	
50mM Tris-HCl pH 8.5	20% glycerol	Wash buffer
300mM NaCl	0.2% triton	50mM Tris-HCl, pH 8.0

300mM NaCl
50mM imidazole

Elution buffer
Tris-HCl pH 8.0
300mM NaCl

300mM imidazole
20% glycerol

4.2.7.4 Loquacious

Both the PB and PD isoform of Loquacious were expressed and purified from the pGex plasmid, whereas LoqsPD was also expressed from the pET-M11 plasmid (see 4.2.7.5). For expression, cells were grown to an OD600 of 0.4 at 30°C at 90rpm in 1200ml LB-medium containing 50 μg/mL Ampicillin, 34μg/ml Chloramphenicol, and 0.5% glucose. After induction with 0.2mM IPTG, the culture was grown for 4h. After centrifugation, the cell pellet was resuspended in 30ml Lysis Buffer supplemented with proteaseinhibitors and frozen in liquid nitrogen. After thawing, the lysate was sonified 3 times for 1min, rotated at 4°C for 1h and cleared by centrifugation at 13000rpm for 10min. Subsequently, it was loaded on a GST column equilibrated with wash buffer, washed with 20 CV Wash buffer and 10 CV Wash buffer containing 800mM NaCl, and eluted in a gradient of wash buffer and elution buffer.

Lysis buffer
1x PBS, pH 7.0
800mM NaCl
1% triton

GST-Wash buffer
1x PBS, pH 7.0
30mM NaCl

GST-Elution buffer
50mM Tris-HCl pH 8.0
10mM glutathione

Fractions containing the protein were pooled, one aliquot of PreScission protease was added to the protein and incubated o/n at 4°C to remove the GST-tag. During the digestion, the protein solution was dialyzed in SP-Wash buffer. To remove the GST tag and other remaining contaminations, the protein was loaded on an ion exchange column. Dependent on the isoelectric point of LoqsPB and LoqsPD, the pH of the SP-Wash and Elution buffer was 7.0 and 6.0, respectively. After 20 CV SP-Wash buffer, the protein was eluted in a gradient of SP-Wash and -Elution buffer, suitable fractions were combined and after concentration with centrifugal filter units (Millipore) they were dialyzed in Storage buffer and frozen in liquid nitrogen.

SP-Wash buffer
1x PBS, pH 7.0/6.0
30mM NaCl
1% triton

SP-Elution buffer
1x PBS, pH 7.0/6.0
2M NaCl

	10mM Hepes 7.4	5mM DTT
	100mM KAc	50% Glycerol
Storage buffer	2mM MgAc	

4.2.7.5 Loquacious, R2D2 and their dsRBDs

DeltaNC and dsRBD constructs of both Loqs and R2D2 were expressed from the pET-M11 plasmid, LoqsPD also as full length protein. For expression, the cultures were grown to an OD600 of 0.5 at 37°C and 110rpm in 2l of LB medium containing 10µ/ml Kanamycin, 34µg/ml Chloramphenicol, and 0.5% glucose. After induction with 0.5mM IPTG they were cooled down to 20°C, grown over night and harvested. For purification, the cell pellets were resuspended in 30ml Lysis buffer supplemented with DnaseI, Lysozyme and Proteaseinhibitor without EDTA, and the solution was sonified and incubated on ice for 1h. The lysate was cleared via centrifugation and loaded on a Ni-column. The column was washed with 20 CV Wash buffer containing NaCl, LiCl, and again NaCl, and the protein was eluted with 2 CV Elution buffer. TEV protease was added to remove the double-Z/His-tag from the protein o/n at 4°C, and during digestion the protein solution was dialyzed back in Wash buffer. On the next day, the protein was again loaded on the Ni-column to remove the double-Z/His-tag and uncleaved protein.

Lysis buffer	Wash buffer	Elution buffer
50mM Tris, pH 8.0	50mM Tris, pH 7.2	50mM Tris, pH 7.2
500mM NaCl	500mM NaCl / 1M LiCl	500mM NaCl
10mM Imidazole	10mM Imidazole	250mM Imidazole
10mM MercaptoEtOH	10mM MercaptoEtOH	10mM MercaptoEtOH

The flowthrough containing the desired protein was then concentrated and loaded on a gelfiltration column equilibrated with Gefi buffer. Depending on the size of the protein, the flowthrough was collected and suitable fractions were combined. After dialysis in Storage buffer, aliquots were frozen in liquid nitrogen and stored at -80°C.

GeFi buffer

20mM Tris pH 7.2

500mM NaCl

5mM EDTA

5mM DTT

4.2.7.6 Dicer

Dicer proteins were expressed in insect cell culture using the baculovirus system. As cell lines, both Sf21 and H5 were used. Baculoviral bacmid DNA was produced by transforming DH10 cells with pFastBac plasmids containing Dicer. After transformation (see 4.2.1.5), cells were streaked on agar plates containing 10µ/ml Kanamycin, 10µg/ml Tetracyclin, 7µg/ml Gentamycin, 1mM IPTG and 20µg/ml X-gal. Recombination of the plasmid with the bacmid DNA disrupts the LacZ-locus, so positive clones could be identified by their white colour, in addition a PCR with primers located on the integrated part of the plasmid and on the bacmid DNA was performed. Bacmid DNA was extracted following the usual Midi-prep protocol or the Large construct DNA purification protocol of NucleoBond Xtra BAC. To avoid shearing, the reulting DNA was only pipetted with trimmed pipetting tips and stored at 4°C.

To generate the virus, bacmid DNA was transfected in Sf21 or H5 cells (see 4.2.4.5). After five days, signs of infection such as large or dead cells were visible, the virus-containing supernatant was cleared by centrifugation from cells and sterilized by filtration. A 10ml culture of $1 \cdot 10^6$ cells/ml Sf21 or a 50ml culture of $0.5 \cdot 10^6$ cells/ml H5 cells were infected with this virus for amplification. After five days, cell number, fraction of viable cells, and expression of the protein were checked and in case of a successful infection, 500ml of $0.5 \cdot 10^6$ cells/ml Sf21 or H5 were infected with 10ml or 20ml of Sf21 or H5 virus, respectively. The virus could be used up to the fourth generation.

For purification, 1 - 2l of cell culture expressing the protein were harvested by centrifugation at 3000rpm for 10min. The pellets were washed with 1xPBS and subsequently resuspended in four times pellet volume of Lysis buffer supplemented with EDTA free protease inhibitor tablets. To allow the cells to swell in the hypotonic buffer they were incubated on ice for 20min, then the protein was extracted by douncing the cell suspension 40 times on ice. The lysate was then incubated on ice for another 20min with an extra 500mM NaCl added after douncing. After centrifugation, the supernatant was loaded on a Ni-column equilibrated with Wash buffer. Three washing steps followed, with 10 CV wash buffer, 10 CV wash buffer containing 800mM NaCl, and again 10 CV Wash buffer. Then the protein was eluted with a gradient of Wash and Elution buffer. Fractions containing the desired protein were pooled, concentrated and loaded on a size exclusion column equilibrated with GeFi buffer. Again, suitable fractions were pooled, dialyzed in storage buffer and frozen in liquid nitrogen.

Lysis/Wash buffer	Elution buffer	GeFi buffer
10mM Hepes, pH7.4	10mM Hepes, pH7.4	10mM Hepes
10mM KAc	10mM KAc	100mM KAc
2mM MgAc	2mM MgAc	2mM MgAc
20mM Imidazole	250mM Imidazole	2mM DTT

4.2.8 Binding experiments

Binding experiments required optimization, as discussed in 5.3. The following protocols refer to the endpoints of these optimizations, and all K_D values reported in the results were measured with these parameters.

Hepes binding buffer	Tris binding buffer
10mM Hepes, pH7.4	20mM Tris 7.4
100mM KAc	100mM KAc
2mM MgAc	3.5mM $MgCl_2$
5mM DTT	

4.2.8.1 Anisotropy measurements

Anisotropy measurements were conducted both in direct and in competitive forms. In both cases, the binding reactions were prepared in low-binding 96-well plates in Hepes binding buffer. Proteins were transfered in Hepes binding buffer via centrifugal filter units before adding them to the binding reaction. For direct measurements, 10nM fluorescently labeled RNA were mixed with increasing amounts of protein. For competition experiments, increasing amounts of unlabeled RNA were added to 7.5nM fluorescently labeled RNA and protein at a concentration where at least 90% binding had been observed. The binding reactions were incubated at RT for 30min, then both fluorescence polarization and intensity were measured in a Tecan plate reader, with an excitation wavelength of 470nm, an emission wavelength of 520nm, and an excitation and emission bandwidth of 5nm.

4.2.8.2 Electrophoretic mobility shift assays

EMSAs were carried out both with fluorescently and radioactively labeled RNA. Measurements with fluorescently labeled RNA were prepared as described in 4.2.8.1. Of radioactively labeled

RNA, 7.25nM were mixed with increasing amounts of protein in Tris binding buffer and incubated for 30min. Both preparations were mixed 1:4 with loading buffer (binding buffer with 50% glycerol) and loaded on 4% native acrylamide gels (12.2ml H_2O, 2ml 40% Acrylamide, 750ml 10x TB, 75µl APS, and 15µl TEMED). After a 24min run at 200V, gels with fluorescently labeled RNA were recorded directly in the Typhoon scanner (excitation at 488nm, emission detection at 526nm), a phosphoimager screen was exposed with gels containing radioactively labeled RNA, which was then recorded in the Typhoon scanner.

4.2.8.3 Thermophoresis measurements

For thermophoresis measurements, 50nM fluorescently labeled RNA was mixed with increasing amounts of protein in Tris binding buffer. After incubation for 30min, the binding reaction was once again pipetted up and down to ensure a homogenous distribution and approximately 10µl were sucked in a glass capillary via capillary force. These capillaries were arranged on a magnetic slide and the behavior of the fluorophores in temperature jump, thermophoresis and backdiffusion was recorded in a Nanotemper Monolith NT.

4.2.9 Crosslink experiments

To map the binding site of the protein constructs on an RNA duplex, crosslinking experiments were performed. 4-S U containing RNA substrates are summarized in figure 5.29. They were labeled radioactively as described in 4.2.3.8 and annealed to their partner strands to produce perfectly matched or bulged duplexes. 10nM RNA was incubated with proteins at a concentration above their respective dissociation constant in Hepes binding buffer. After incubation at RT for 30min, RNA and proteins were crosslinked with 3x 500mJ/cm^2 of 365nm light, mixed with SDS loading buffer and run for 1h at 170V on a 15% polyacrylamide gel. This was subsequently dried and used to expose a phosphoimager screen o/n, which was then recorded on a Typhoon scanner.

4.2.10 Analysis

Bands in EMSAs and crosslink experiments were quantified using MultiGauge software. Lanes were defined as depicted in figure 4.3 and the intensity histogram of each lane was calculated. The intensity counts of one peak were summed up and used as signal strength for that particular band. Data points were calculated by dividing the crosslinked band or all shifted bands through the total signal. For easier comparison, these data points were normalized between 0 and 1.

Values obtained from Anisotropy or Thermophoresis measurements were likewise normalized,

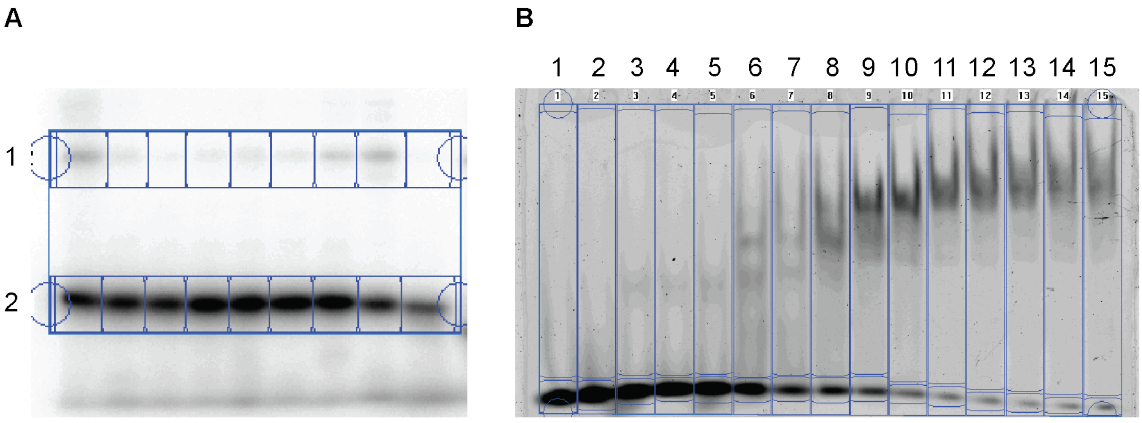

Figure 4.3: MultiGauge Analysis
A: For evaluation of Crosslink experiments, both crosslinked and uncrosslinked bands were quantified and the crosslinked to total ratio was calculated. B: For EMSA experiments, the lanes were divided in bound and unbound, quantified and the bound to total ratio was calculated.

and all binding curves except those obtained from competition experiments were fitted in origin with the Hill formula

$$f(x) = min + \frac{(max - min) \cdot x^n}{K_D^n + x^n} \tag{4.1}$$

Competition curves were fitted as described in [89] with the following function:

$$f(x) = 0.5 \cdot K \cdot \frac{x + B + K_D - \sqrt{(x + B + K_D)^2 - 4 \cdot x \cdot B}}{B} + C \tag{4.2}$$

5 Results

5.1 Characterization of LoqsPD and R2D2 in their cellular environment

LoqsPD and R2D2 are able to discriminate between different siRNA classes and to process them independently of each other [38]. Understanding their behavior in the cell might provide insight into the mechanism involved in this distinction. Localization to specific cellular compartments and the amount of protein present in the cell can be features that distinguish the two dsRBPs. To assess the extent to which LoqsPD and R2D2 compete with each other for Dcr-2 binding, the overlap of their binding site on Dcr-2 is of interest. These questions are addressed in this section.

5.1.1 Subcellular localization of Loquacious and R2D2

Myc-tagged LoqsPB, LoqsPD and R2D2 were expressed in S2-cells by transfecting 1.5 and 7.5µg plasmid DNA per 24-well, and an α-myc immunofluorescence was performed. Titration of the amount of transfected plasmid was critical, since extreme overexpression led to abnormal localization of the dsRBPs. In cells transfected with 1.5µg DNA, all dsRBPs localized to the cytoplasm, whereas the nuclei were devoid of any staining (figure 5.1A). Part of the proteins clustered together in several foci throughout the cytoplasm. Focus formation has been reported for R2D2, but not for Loqs[76]. Here, this clustering behavior was more prominent for the Loqs isoforms than for R2D2. Untransfected controls were performed by other members of the lab, which did not show any punctate pattern in the cytoplasm. α-myc western on transfected cell lysates showed that LoqsPD expression was the strongest, whereas transfection with 1.5µg myc-R2D2 plasmid did not lead to any significant R2D2 overexpression. Transfection with 7.5µg myc-R2D2 plasmid reached expression levels similar to transfection with 1.5µg myc-LoqsPB (figure 5.1B).

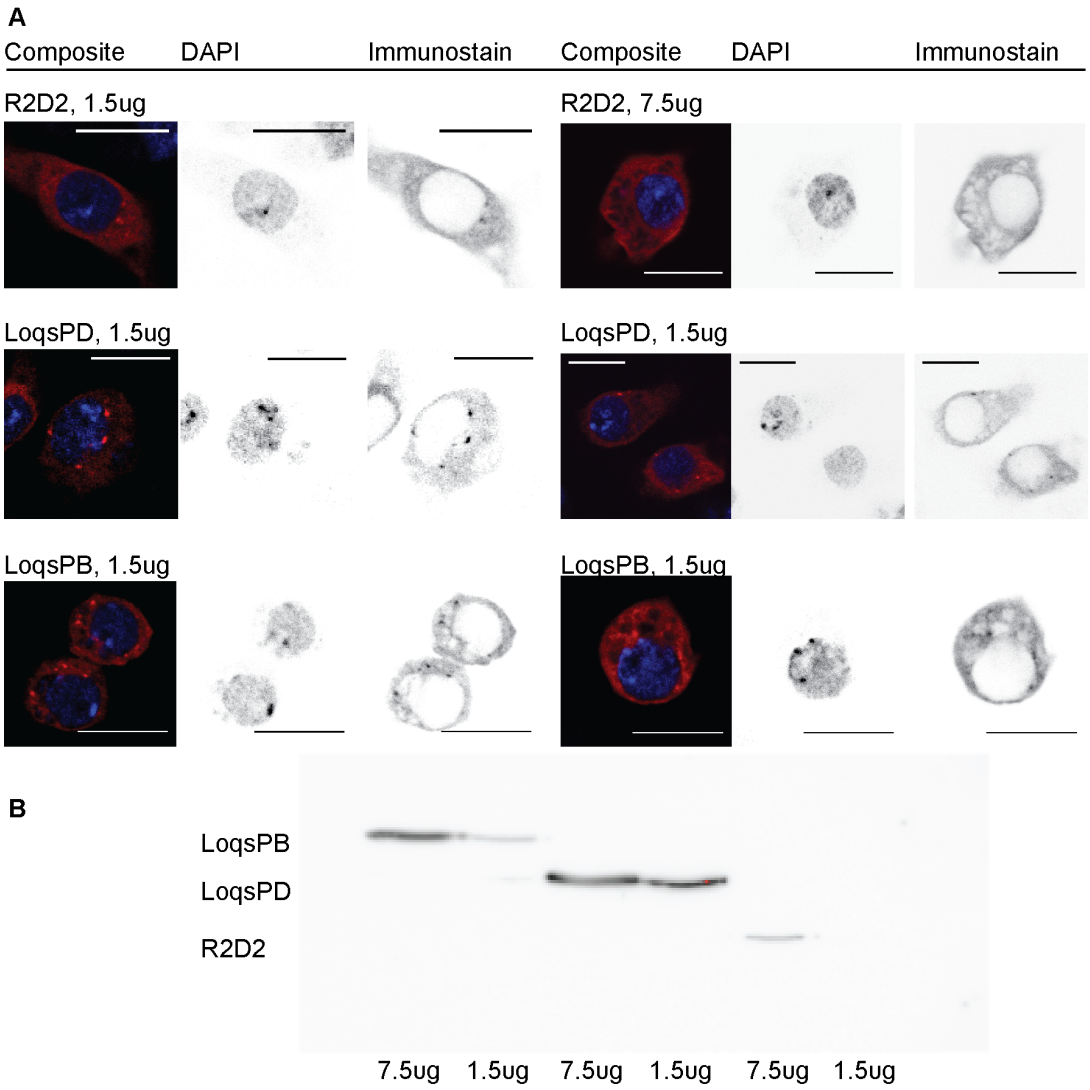

Figure 5.1: Subcellular localization of Loqs and R2D2

A: α-myc immunostain on S2 cells transfected with myc-R2D2, myc-LoqsPD or myc-LoqsPB plasmids. Cells were transfected with 1.5µg or 7.5µg plasmids and after 6 days fixed on a glass slide. They were subsequently incubated with α-myc antibody, diluted 1:100 and an α-mouse antibody conjugated to Alexa488, diluted 1:200. After staining the nuclei with 0.5µg/ml DAPI solution, images were taken with a Leica confocal microscope. Loqs expression was strong enough with 1.5µg plasmid, sufficient R2D2 expression required transfection of 7.5µg plasmid. Scale bars represent 10µm. B: Cell lysates were analyzed via western blot to check for expression levels. LoqsPD showed the strongest overexpression, whereas R2D2 was only expressed when 7.5µg plasmid per 24-well were transfected. α-myc antibody was diluted 1:1000 in TBS-T, exposure time was 100s.

5.1.2 R2D2 is present in larger amounts than LoqsPD

To get an idea about the amounts of R2D2 and LoqsPD present in a cell, I performed western blots on cell lysates using α-Loqs and α-R2D2 antibodies. To calibrate the antibodies, a dilution of recombinant protein was included (figure 5.2). Unfortunately, the antibodies did not result in one defined band in the lysates. In the LoqsPD blot, no distinctive band appeared in the lysate on the appropriate height, and the bands present had approximately the same strength as the one corresponding to 5fmol recombinant protein (figure 5.2A). In the R2D2 blot, a distinct band at 35kD could be observed, and knockdown confirmed this band to be R2D2 (figure 5.2B, right). The R2D2 band had approximately the same strength as that corresponding to 50fmol recombinant protein, therefore R2D2 is at least ten times as abundant as LoqsPD in S2 cells.

For a more precise determination of the R2D2 concentration, I did a densiometric analysis of the R2D2 bands. According to this analysis, 0.00013fmol +/- 0.00004 R2D2 are present in one cell, corresponding to 80105 +/- 21036 molecules. S2 cell size was estimated to be 860µm^3 via the images in figure 5.1. With a nucleus size of ~165µm^3, the volume occupied by R2D2 amounts to 695µm^3. This gives a R2D2 concentration of roughly 100nM.

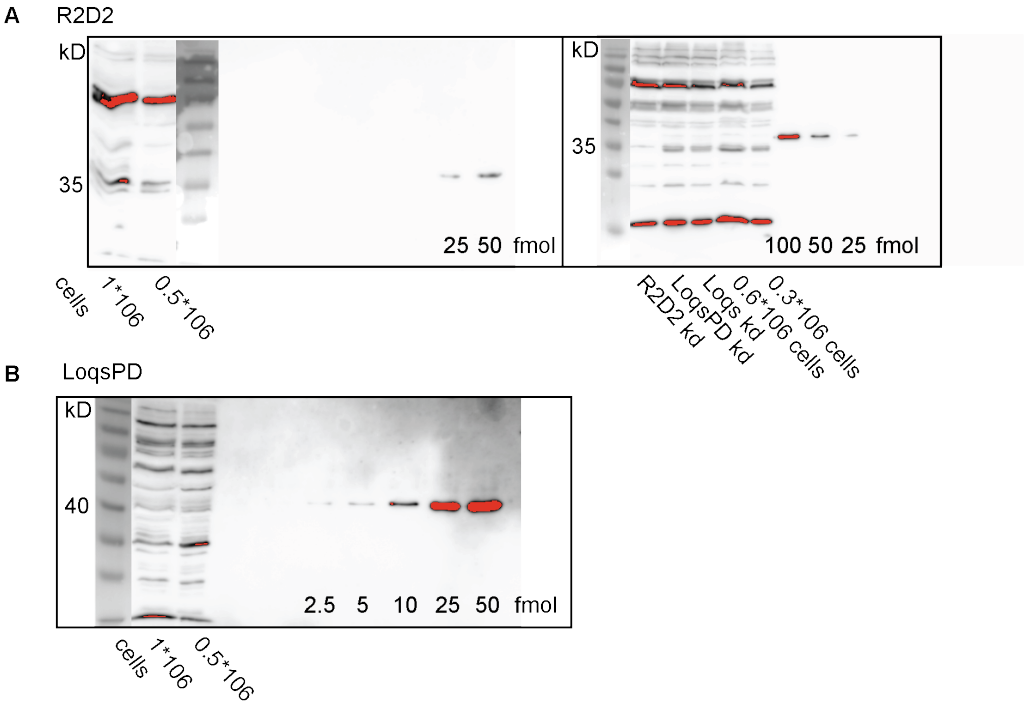

Figure 5.2: Determination of cellular R2D2 concentration

Figure legend on the following page

Figure legend of figure 5.2
A: α-R2D2 western blot on cell lysates and recombinant protein. Abcam α-R2D2 was used in a 1:1000 dilution, exposure times were 10min and 15s for the left and right blot, respectively. A defined number of cells were loaded in each blot, in addition, lysate of R2D2 depleted cells was included in the 1A, right, to identify the correct band in the lysate. To calibrate the antibody, a dilution of recombinant His-tagged R2D2 obtained with the shock-refolding method (see section 5.2.2) was loaded. Densiometric analysis was carried out with the MultiGauge software. B: α-LoqsPD spec. western blot on cell lysates and recombinant protein. LoqsPD specific antibody was used in a 1:1000 dilution, exposure time was 20min. To calibrate the antibody, a dilution of recombinant His-tagged LoqsPD was loaded.

5.1.3 R2D2 binds Dcr-2 in the linker of the helicase domain

Both LoqsPD and R2D2 bind Dcr-2 via its helicase domain [38]. To narrow down the interaction site, Dcr-2 constructs with an increasingly truncated helicase domain (figure 5.3A) were co-expressed with myc-tagged LoqsPD and R2D2 and probed for interaction in immunoprecipitation experiments. Input and bound fraction of myc-R2D2 IPs and the corresponding bound/input ratios are shown in figure 5.3B and C. Removing the linker between the two helicase domains reduced binding of R2D2 to a level comparable to the negative control (construct 3), therefore this region is probably the main interaction site for R2D2.

Alignment of the Dcr-2 sequence on the structure of mouse RigI ATPase [21] shows that the linker is folded in an additional domain on top of the helicase (figure 5.3D, purple) and therefore provides an accessible binding platform for a dsRBP. TRBP and PACT have also bind Dicer's helicase domain in the linker and/or the C-terminal part [53].

Figure 5.3 E and F shows the corresponding IP with myc-tagged LoqsPD instead of R2D2. LoqsPD interacts with Dcr-2 via the Helicase domain [38], therefore the Dcr-2 Delta Helicase construct constitutes the negative control. Unfortunately, this control band intensity was even increased in the bound compared to the input western, therefore LoqsPD experiments were inconsistent.

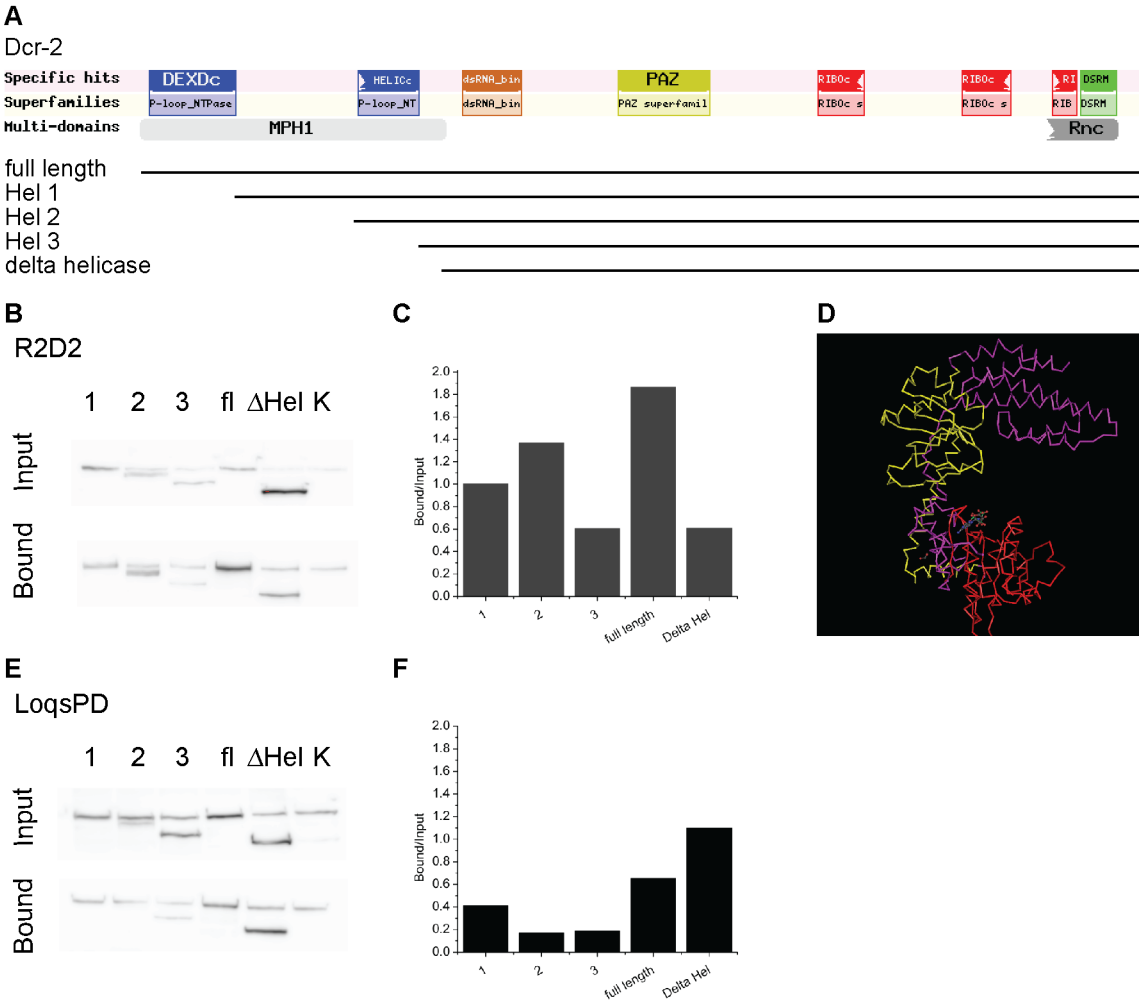

Figure 5.3: R2D2 binds linker of Dcr-2's helicase domain

A: Dcr-2 truncation constructs used for immunoprecipitation experiments in B and E. B: α-myc immunoprecipitation on cells transfected with myc-R2D2 and different flag-Dcr-2 constructs. Flag-tagged luciferase was loaded as a control. 400ng of each plasmid were transfected and the cells were harvested after 6 days. Input and bound fractions were analyzed via α-Dcr-2 western blot (antibody dilution 1:1000, exposure time 20s and 10s for input and bound, respectively).C: Quantitative analysis of B; densiometric analysis was carried out with the MultiGauge software and the bound versus input ratio is shown. Helicase construct 3 shows binding comparable to the negative control. D: Alignment of Dcr-2 helicase domain with mouse RigI ATPase domain, colors according to domains predicted in Dcr-2. red: DEXDc domain, yellow: HELICc domain, purple: linker. E: α-myc immunoprecipitation on cells transfected with myc-LoqsPD and different flag-Dcr-2 constructs. The experiment was carried out analogous to B. F: Quantitative analysis of E.

5.1.4 Excess of nonspecific dsRNA does not interfere with endo-siRNA function

One feature of small RNAs that could determine whether their biogenesis depends more on R2D2 or LoqsPD is their abundance. R2D2 has been shown to be primarily responsible for function of exo-siRNAs and loss of LoqsPD to hamper endo-siRNA function [37, 38], with exo-siRNA being in general more abundant than endo-siRNAs [79]. In this scenario, the R2D2/Dcr-2 complex would not take part in endo-siRNA biogenesis due to a lower affinity for dsRNA than the LoqsPD/Dcr-2 complex, which would have a higher affinity for dsRNA and therefore be able to process rare endo-siRNAs. If that was the case, the endo-siRNA biogenesis machinery would be saturated with dsRNA during a viral infection, and the endo-siRNA mediated control would be lost. To test for this, I used a cell line containing GFP in high copy numbers, which is under endo-siRNA control. Knockdown of LoqsPD led to a derepression of GFP, whereas R2D2 led to a hyperrepression compared to untreated cells, as already reported [37] (figure 5.4A). Adding increasing amounts of unspecific dsRNA did not derepress GFP expression (figure 5.4A), which should be expected if the endo-siRNA machinery would get saturated with exo-siRNA. Compared to untreated cells, addition unspecific dsRNA even hyperrepressed GFP, though not in a dose-dependent manner. As a control, a miR-277 reporter cell line was used, which contains GFP with two perfect match miR-277 binding sites in its 3'UTR. LoqsPD knockdown and unspecific dsRNA addition did not affect the miR-277 reporter, whereas GFP expression was derepressed upon R2D2 knockdown, as expected [29].

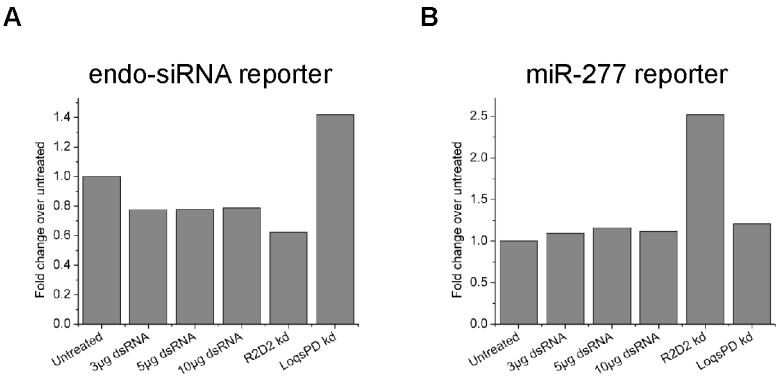

Figure 5.4: Excess of dsRNA does not interfere with endo-siRNA function
A+B: FACS analysis of reporter cell lines treated with different dsRNAs. As unspecific dsRNA, dsRNA targeting the Luciferase sequence was used. A double knockdown was performed, with addition of 3µg R2D2 and LoqsPD dsRNA and the indicated amount of unspecific dsRNA on day 1 and 3. On day 5, FACS data was collected. GFP fluorescence was normalized to that of untreated cells. Values represent the mean of two duplicates. One experiment out of two is shown, the second experiment showed the same trends in fold change over untreated values.

5.2 Optimization of protein expression and purification

To gain more insight in the roles the dsRBPs play in small RNA biogenesis, I wanted to determine their binding strength and geometry to the small RNA precursors present in a cell. To be able to do that, I needed to express and purify the involved proteins in sufficient amount and purity. This section deals with the optimization of expression and purification of Loqs, R2D2, and Dicer proteins.

5.2.1 Loquacious

I started with testing three different vector/tag systems for expression: the pGex vector which adds a GST-tag to the protein that can be cleaved off by digestion with PreScission protease, the pET vector which adds a His_6-tag, and the pPal vector which adds a Profinity eXact tag that is bound and cleaved by subtilisin protease, allowing affinity purification and on-column cleavage of the fusion protein. For both LoqsPB and LoqsPD, the GST fusion protein showed the strongest expression (figure 5.5A), therefore the pGex plasmids were used for further optimization. Since expression was very robust, induction with 0.2mM IPTG was sufficient for optimal expression. Under standard expression conditions (37°C, 110rpm), a maximum expression level was reached between 3 and 5h after induction (figure 5.5B) and 4h were set as standard expression time. To break the cells open, a combination of detergent and freeze/thaw treatment was used: Pellets were resuspended in 1x PBS, 1% Triton and frozen in liquid nitrogen. The lysate was sonified, but the amount of soluble protein in the supernatant was not satisfactory. Therefore, high salt and Urea treatments were tested for improving solubility. Addition of 800mM NaCl to the lysis buffer increased the amount of protein in the soluble fraction substantially, rendering Urea treatment unnecessary (figure 5.5C). The first purification step over the GST column required a high salt wash, otherwise the 260/280 ratio was > 1, suggesting a contamination with nucleic acids. Separate digestion with DNase or RNase showed the contamination to be RNA (figure 5.5D). GST-tag removal was problematic, since a lot of the protein was destabilized and formed insoluble aggregates upon digestion. To improve protein stability, I tried several additives to the digestion reaction (30% glycerol, larger reaction volume/smaller concentration, 1mM DTT, 2mM Arginine, higher salt concentration), which did not lead to any significant improvement. Another approach was to improve proper protein folding by slowing down the translation process, and indeed expression at 21°C led to an increased protein yield, as well as reduction of shaking velocity from 110 to 90rpm. The second purification step over the ion exchange column required no optimization and was conducted as described in 4.2.7.4.

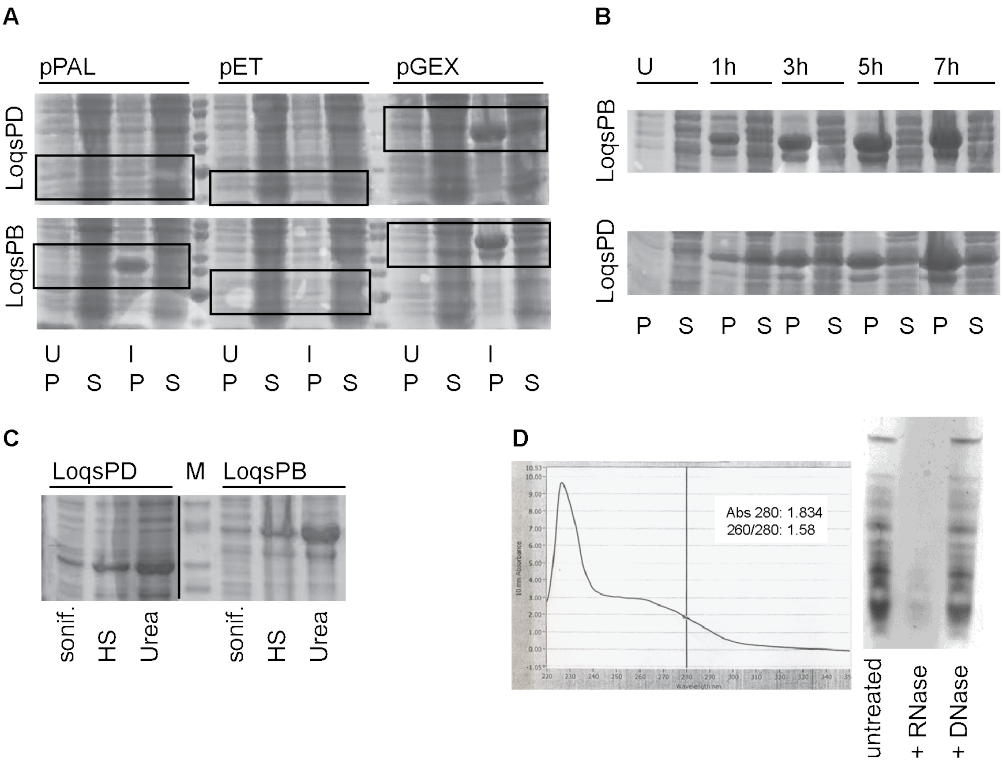

Figure 5.5: Optimization of Loquacious expression and purification

A: Different vectors tested for Loquacious expression. Cultures were grown to OD = 0.6 at 21°C, induced with 0.2mM IPTG and grown for 5h. Boxes indicate the expected size of the tagged proteins. U = uninduced, I = induced, P = pellet, S = soluble lysate. B: Time course of pGex Loqs expression. Cultures were grown at 21°C to OD = 0.6, induced with 0.2mM IPTG and harvested after the designated time. C: Optimization of Loquacious solubility. After thawing the samples, they were - sonified twice for one minute (sample sonif.), - pelleted again and resuspended in 1x PBS, 1% Triton, 800mM NaCl and incubated on ice for 1h (sample HS), - pelleted again and resuspended in 100mM NaH_2PO_4, 10mM Tris, 8M Urea, pH 8.0 and incubated at RT for 1h (sample Urea). Each step improved solubility of the protein. D: RNA content in LoqsPB preparation. left: Absorption scan of a full length LoqsPD preparation. Without a 800mM NaCl wash over the GST-column, there were nucleic acid contaminations in the protein preparation, which can be seen in the absorption scan at 260nm. A 260/280 ratio of 1.58 is rather high, usual values were around 0.7. right: The protein prep was digested with proteinase K and the contaminating nucleic acids were extracted with phenol/chloroform. Subsequently they were divided in three samples, which were either treated with DnaseI, RNaseI or left untreated. An acrylamide urea gel shows that the contaminating nucleic acids were only sensitive to RNase, but not DNase treatment. Therefore, the contamination at 260nm is caused by residual RNA.

5.2.2 R2D2

Expression of R2D2 was very efficient in all vectors tested (figure 5.6A), but only the MBP-R2D2 fusion protein was soluble in 20mM Tris pH 8.0, 50mM NaCl, and 5mM β-Mercaptoethanol (supplemented with protease inhibitor), after sonification of the lysate [59]. This fusion protein did not bind to RNA in EMSA or Anisotropy experiments (figure 5.6B). Therefore, I tried to improve the solubility of the other R2D2 constructs. Expression in different E.coli strains was not successful. In the Rosetta strain, which overexpresses tRNAs for rare codons, R2D2 was not soluble either and in the Arctic Express strain, in which cold-adapted chaperones should increase the yield of properly folded protein during a 10 ºC expression, the expression level was too low. Other strategies to improve protein folding such as slow expression at 21°C, induction with only 0.05mM IPTG, and expression in insect cells using the baculovirus system proved to be inefficient as well. Purification under denaturing conditions with 8M Urea or 6M GuHCl was successful, and after transfer in non-denaturing buffer for instantaneous refolding by adding a large volume of physiological buffer (10mM Tris pH 7.5, 100mM KAc, 100mM NaH_2PO_4, 20mM Imidazol, 0.5M Arginine +/- 6M GuHCl) or for slow refolding on column (10mM Tris pH 7.0, 100mM NaH_2PO_4, 20mM Imidazol +/- 8M Urea), there was still soluble protein left. Unfortunately, a qualitative run on a size exclusion column repeatedly showed an anomalous profile of the shock-refolded protein (figure 5.6C), and in Anisotropy experiments no proper binding of R2D2 that was refolded on the column could be observed (figure 5.6D). Therefore, efforts to purify full length R2D2 without its stabilizing partner Dcr2 were dismissed as futile and instead truncation constructs of R2D2 were purified, including the individual dsRBDs (AA 1-74 and AA 91–169) and the double domain construct (AA 1-169). Expression and purification of those were straightforward and performed as described in 4.2.7.5.

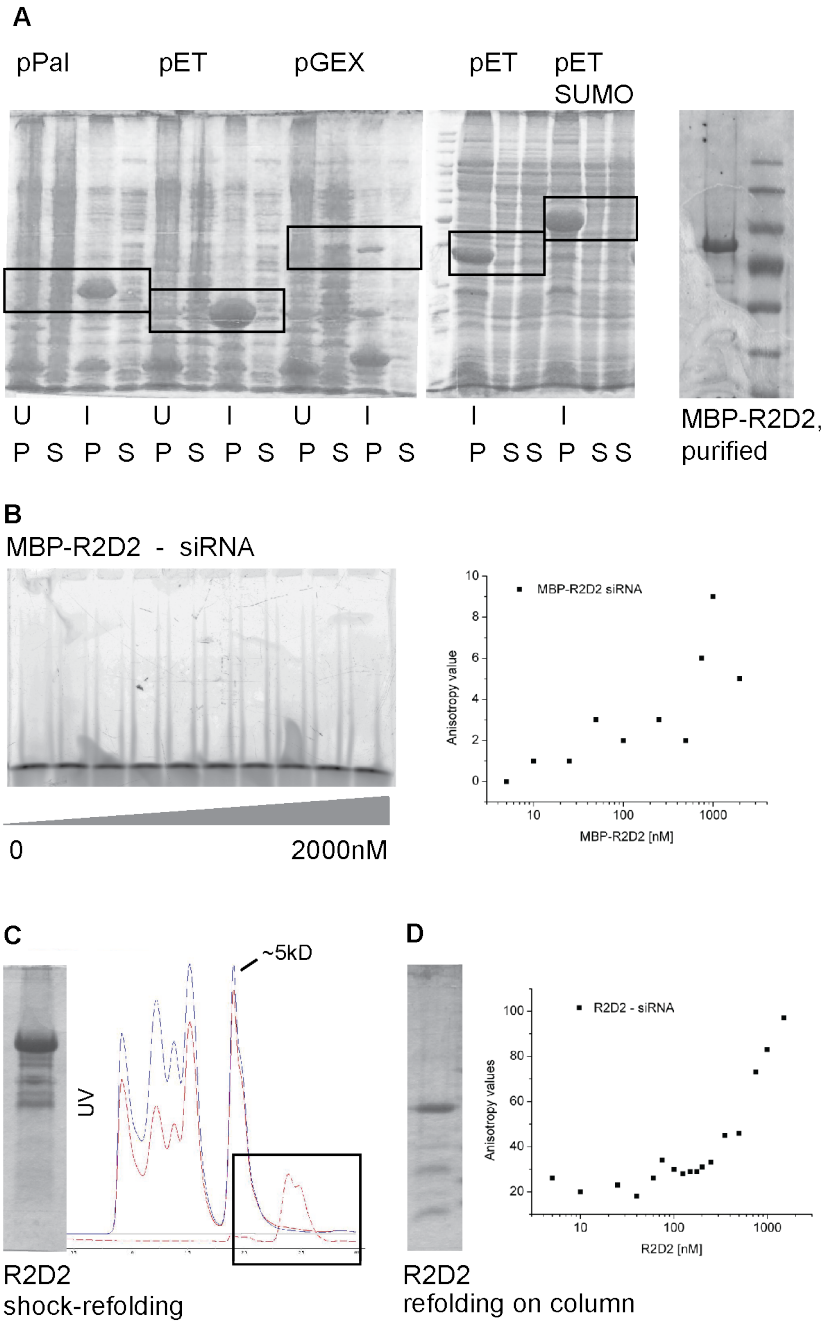

Figure 5.6: Optimization of R2D2 expression and purification

A: Different vectors tested for R2D2 expression. Cultures were grown to OD = 0.6 at 37°C, induced with 5mM IPTG and grown for 4h. Boxes indicate the expected size of the tagged proteins. U = uninduced, I = induced, P = pellet, S = soluble lysate. Instead of the lysate, the already purified MBP-R2D2 protein is shown. B left: EMSA experiment with MBP-R2D2 and a fluorescently labeled siRNA duplex. Increasing amount of protein was titrated to 100nM RNA and incubated for 30min at RT. The binding reaction was then resolved on a 5% native acrylamide gel.

Figure legend is continued on the next page

Continuation of Figure legend 5.2.2
B right: Anisotropy measurements of MBP-R2D2 / siRNA binding. Binding reactions with increasing amounts of protein and 100nM RNA were mixed in a 96-well plate and recorded in a Tecan plate reader after 30min incubation at RT. Note that in the case of binding the Anisotropy values rise up to 200, whereas here they remain as low as 10. C: Recombinant R2D2 after shock-refolding and its gel filtration profile. The R2D2 sample was run on a qualitative size exclusion column, resulting in the depicted profile. The box indicates the signal from the R2D2 protein, the other traces correspond to a gel filtration standard. Refolded R2D2 appears even after the smallest component of the standard, indicating either protein degradation or retention of the protein due to unspecific interaction with the column material. D: Recombinant R2D2 after refolding on column and its binding behavior in Anisotropy measurements. Binding reactions with increasing amounts of R2D2 and 10nM RNA were mixed in a 96-well plate which was recorded in a Tecan plate reader after incubation at RT for 30min.

5.2.3 dsRBDs of Loquacious and R2D2

The only optimization necessary for this protocol was the step from eluting the tagged protein from the Ni agarose by digestion with TEV protease to eluting with 250mM Imidazole followed by TEV digestion in solution, which increased the yield substantially.

5.2.4 Dicer

Dicer proteins were expressed in insect cells using the baculovirus system (figure 5.7A). Optimization was required to obtain enough soluble protein from the lysates. Since normal lysis buffer (1x PBS, 1% Triton) was not sufficient for adequate solubility, several extensions to the protocol were tested, like sonification of varying lengths (2 - 12min), DNase digestion with 10U/ml, and high salt extraction with 800mM NaCl. In addition, cells were resuspended in hypotonic buffer, incubated on ice to allow them to swell and opened by douncing 40 times. As can be seen in figure 5.7B, douncing and high salt extraction rendered the best results. Therefore, a combination of the two was used, where the lysate was incubated on ice for another 20min with an extra 500mM NaCl added after douncing. In all Dcr-2 purifications, a 70kD protein was co-purified. To identify this protein and to exclude contamination with associating dsRBDs endogenous to Sf-21/H5 cells, a mass spectrometry analysis was performed on the isolated band (figure 5.7C). The hits are listed in Appendix 6.6 and identify the protein as a Hsp70 homolog.

5.3 Binding experiments - Method optimization and validation

I tested Electrophoretic mobility shift assays, Anisotropy and Thermophoresis as methods for measuring the strength of protein - RNA interactions and optimized them for the use with dsRBPs and RNA. To determine the geometry of protein - RNA binding I made use

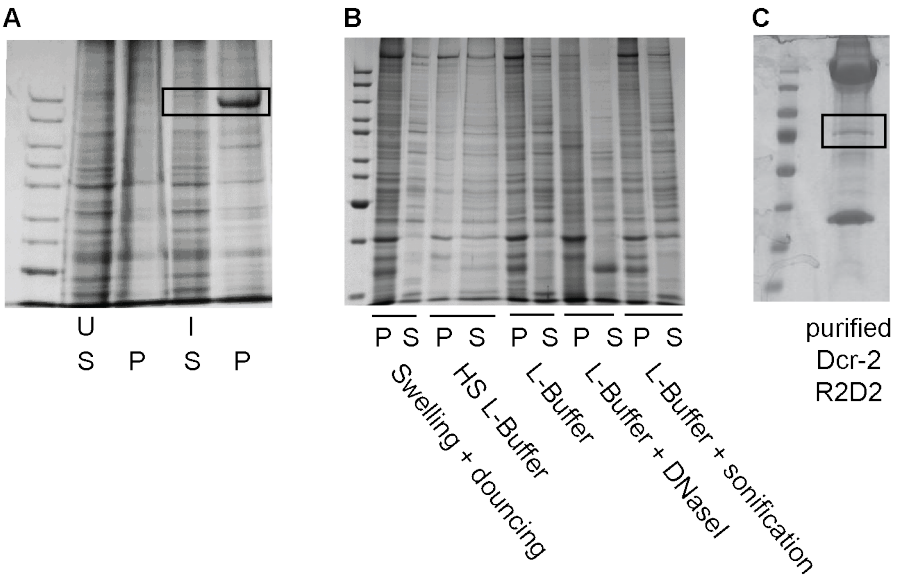

Figure 5.7: Optimization of Dicer expression and purification
A: High5 insect cells were infected with baculoviruses containing the Dcr-2 gene and grown for 5 days at 27.5°C. The cells were then pelleted, resuspended in 1xPBS, 1% Triton and divided in soluble (L) and insoluble (P) fractions (U: uninfected, I: infected). B: Optimization of Dicer solubility. Sf21 cells infected with Dcr-1 virus were harvested and either treated as described in paragraph 4.2.7.6 (swelling and douncing) or resuspended in 1xPBS, 1% Triton (L Buffer) + 800mM NaCl (HS L-Buffer). Samples resuspended in L Buffer were additionally sonified 3x 1min or treated with DnaseI. Soluble (S) and insoluble (P) fractions were separated and compared on the depicted gel. C: A 70kD protein copurifies with Dicer. A Dcr-2/R2D2 preparation is shown on the gel, the band at 70kD was subsequently excised as indicated by the box and analyzed via Mass spectrometry, which identified it as a Hsp70 homolog.

of crosslinkable thiouridines inserted at specific positions in the RNA oligo. Sources for artifacts such as influence of the base pairing environment of the thiouridine or the possibility of capturing non-physiological interactions with the crosslink had to be checked and ruled out.

5.3.1 EMSA

For EMSA experiments, a binding buffer and a gel-system had to be found where the protein-RNA complex remains intact and does not dissociate or aggregate. When aggregated, the protein-RNA complex remained in the gel pockets, as shown in figure 5.8A. This occurred in both Hepes and Tris buffered reactions. (B1: 30mM Hepes, 120mM KAc, 2mM MgAc, 1mM DTT + ribolock, B2: 20mM Tris 7.5, 50mM KAc, 5mM $MgCl_2$, 1mM DTT + ribolock). Additives such as 2mM Arginin or 0.1% Triton did not lead to any improvement. Changing salt concentrations and omission of additives in the Tris Buffer (B3: 20mM Tris 7.4, 100mM KAc, 3.5mM $MgCl_2$) led to reduced aggregation and the shifted band was more clearly defined

(figure 5.8B). Finally, in Hepes buffer including DTT (final Hepes binding buffer: 10mM Hepes, pH7.4, 100mM KAc, 2mM MgAc, 5mM DTT), aggregation was reduced to a minimum and protein-RNA binding was stabilized, so that even secondary and ternary complexes could be resolved (figure 5.8C).

As gel system, a 4% native acrylamide gel was chosen, which was poured and run with 0.5x TB. TBE buffered gel systems were rejected, since EDTA hampered the protein-RNA interaction (figure 5.8D).

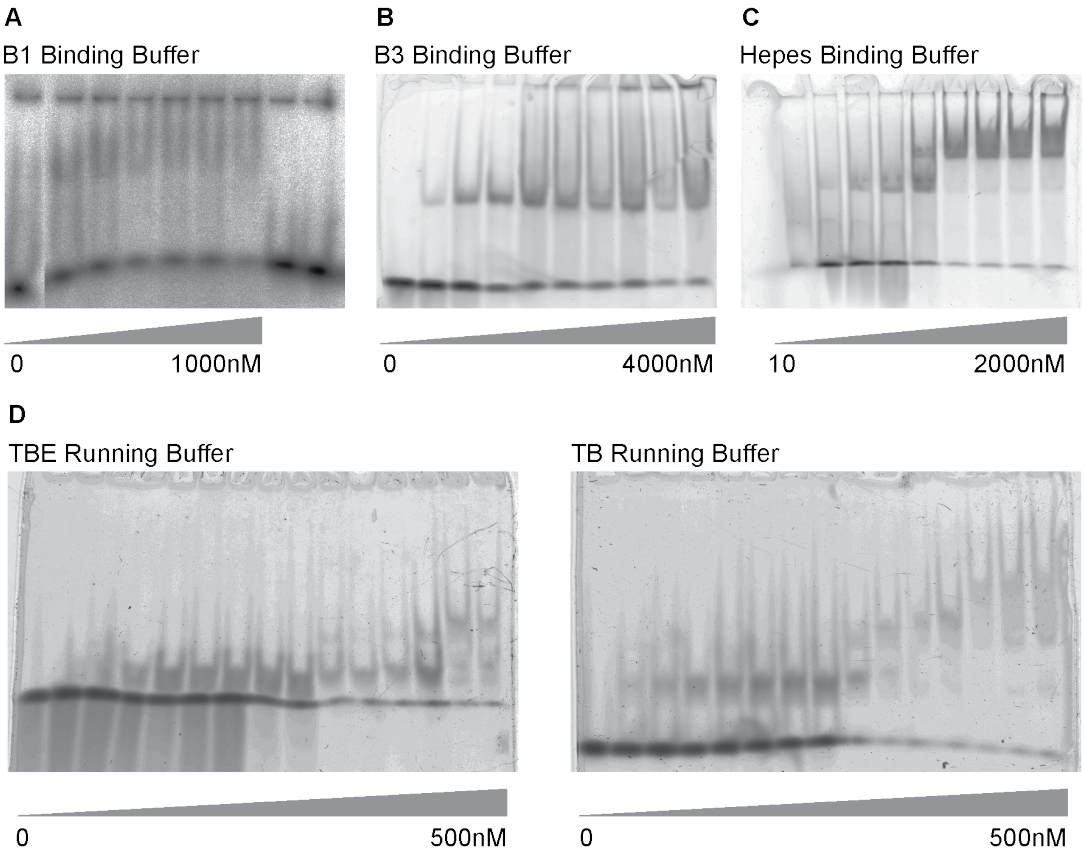

Figure 5.8: Optimization of EMSA experiments
Optimization of Binding and Running buffer for EMSA experiments. A-C: full length LoqsPD was titrated to 5nM radioactively labeled siRNA in B1 Buffer (30mM Hepes, 120mM KAc, 2mM MgAc, 1mM DTT + ribolock), or to 100nM fluorescently labeled RNA in B3 Buffer (20mM Tris 7.4, 100mM KAc, 3.5mM $MgCl_2$) and Hepes binding Buffer (10mM Hepes, pH7.4, 100mM KAc, 2mM MgAc, 5mM DTT) and run on native 4% Acrylamide gel. D: Comparison between TBE and TB running buffer. Loqs DeltaNC was titrated to 10nM fluorescently labeled RNA and the binding reaction was run in 0.5x TBE buffer (left) or 0.5x TB buffer (right). Without EDTA, bands of higher order complexes can be resolved.

5.3.2 Anisotropy

Binding curve quality was equally good with respect to smoothness and stability of the plateaus in all buffers tested (figure 5.9A), so Hepes binding buffer was chosen as standard binding buffer for better comparison with EMSA data. One exception was binding to ssRNA: in 20mM Tris 7.4, 100mM KAc, 3.5mM $MgCl_2$, the anisotropy values decreased upon addition of protein, instead of increase (figure 5.9B).

Other adjustable parameters were the amount of RNA (see 5.3.4) and the plate reader settings: z-position and gain were optimized in each plate. The number of flashes used to determine the fluorescence polarization was set to 100, since with a higher number the standard deviation between repeated measurements at low RNA concentration decreased (24% with 10, 6% with 50 flashes).

For competition experiments, the optimal amount of labeled RNA and added protein had to be identified. For an optimal fit, the binding curve should have a plateau on both ends. If the amount of protein was not sufficient, the anisotropy values dropped immediately after addition of the smallest amount of competitor RNA (figure 5.9C). If too much labeled RNA was provided, or not enough competing RNA was added, the recorded curve stopped before reaching a plateau at the end.

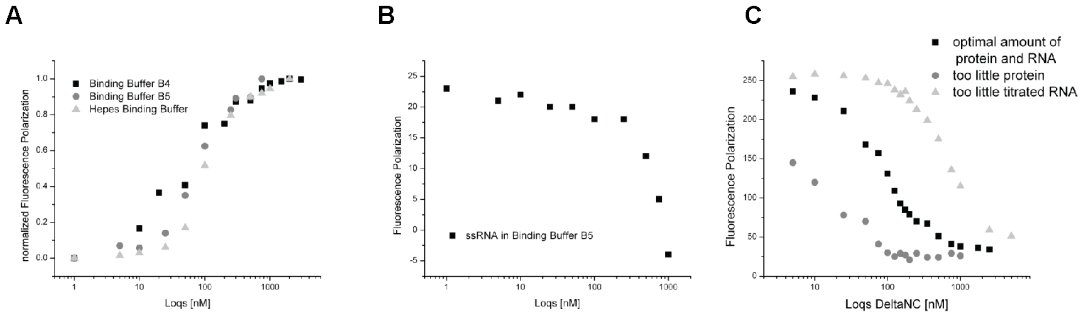

Figure 5.9: Optimization of Anisotropy experiments
A: Binding curves of LoqsPB (black) or LoqsPD (light and dark grey) to siRNAs do not show any significant outliers in any of the Buffers tested (B4: 30mM Hepes pH 7.0, 100mM KAc, 2mM $MgCl_2$, 1mM DTT, 0.1% BSA, 0.01% Tween; B5: 20mM Tris 7.4, 100mM KAc, 3.5mM $MgCl_2$; Hepes binding buffer: 10mM Hepes, pH7.4, 100mM KAc, 2mM MgAc, 5mM DTT). B: Anisotropy measurement of LoqsPD - ssRNA binding in 20mM Tris 7.4, 100mM KAc, 3.5mM $MgCl_2$. Instead of increasing, the Anisotropy values are decreasing. C: Optimization of protein and RNA concentration for competition experiments. Loqs DeltaNC was preincubated with labeled RNA, unlabeled RNA was titrated in and the displacement of the labeled RNA from the protein was followed. The protein - labeled RNA ratio and the amount of titrated unlabeled RNA had to be adjusted to obtain a curve that can be fitted correctly. light grey curve: titration of unlabeled RNA stopped too early; dark grey curve: too little protein (7.5nM RNA, 100nM Loqs DeltaNC); black curve: optimal ratio (7.5nM RNA, 300nM Loqs DeltaNC).

5.3.3 Thermophoresis

Here, parameters for optimization were again the binding buffer, but also the type of capillary and the settings during the experiment, such as strength of the temperature gradient and record length of backdiffusion.

For the temperature gradient, 1.5V were used to achieve the best signal to noise ratio (figure 5.10A), and backdiffusion was observed until a plateau was reached.

The choice between Tris and Hepes buffered measurements was made based on two observations: With the same voltage applied, thermophoresis curves were separated more clearly in Tris than in Hepes buffer, giving rise to a stronger signal difference between bound and unbound RNA (figure 5.10A). In addition, in Hepes buffer unbound ssRNA migrated towards the warm area generated by the laser instead of leaving it. Upon protein binding, the direction of movement normalized. This would not per se be a reason to distrust the binding curves generated from such thermophoresis graphs, but evaluation and comparison with other RNA species would be more difficult, so subsequent measurements were conducted in Tris buffer.

As could be seen with GFP-tagged LoqsPB, adsorption of protein to the capillary wall was an issue under standard conditions (figure 5.10B, black line), which had to be prevented since unspecific adsorption would lower the effective concentration of protein ready for RNA binding in the capillary.

For most measurements, Tris binding buffer complemented with 1% BSA and 0.01% Tween was used. The additives reduce adsorption of protein to the capillary wall, though scanning the capillaries over longer time intervals showed that they could not prevent adsorption completely (figure 5.10B). We therefore recorded the data immediately after filling the capillaries to minimize this effect. In addition, capillaries with a hydrophobic (coated with PEG) and a hydrophilic (coated with Dextran) surface were tested. Hydrophilic capillaries did show reduced protein binding, but instead RNA adsorption occurred, therefore these were dismissed. Between the hydrophobic and the uncoated capillaries no substantial difference was visible after repeated testing, therefore the uncoated ones were used in standard experiments.

For easier and faster evaluation of the data, which usually included 3 consecutive measurements of thermophoresis, temperature jump and backdiffusion behavior, a labtalk routine was written. Data files generated from the Nanotemper software were automatically opened, plotted, and fitted, and the resulting graphs and the fitted parameters were saved. The script is attached in Appendix 6.6.

5.3.4 RNA concentration

RNA concentration was chosen as low as possible, since the amount of protein bound to RNA should be negligible compared to free protein for straightforward calculation of the K_D. Reduc-

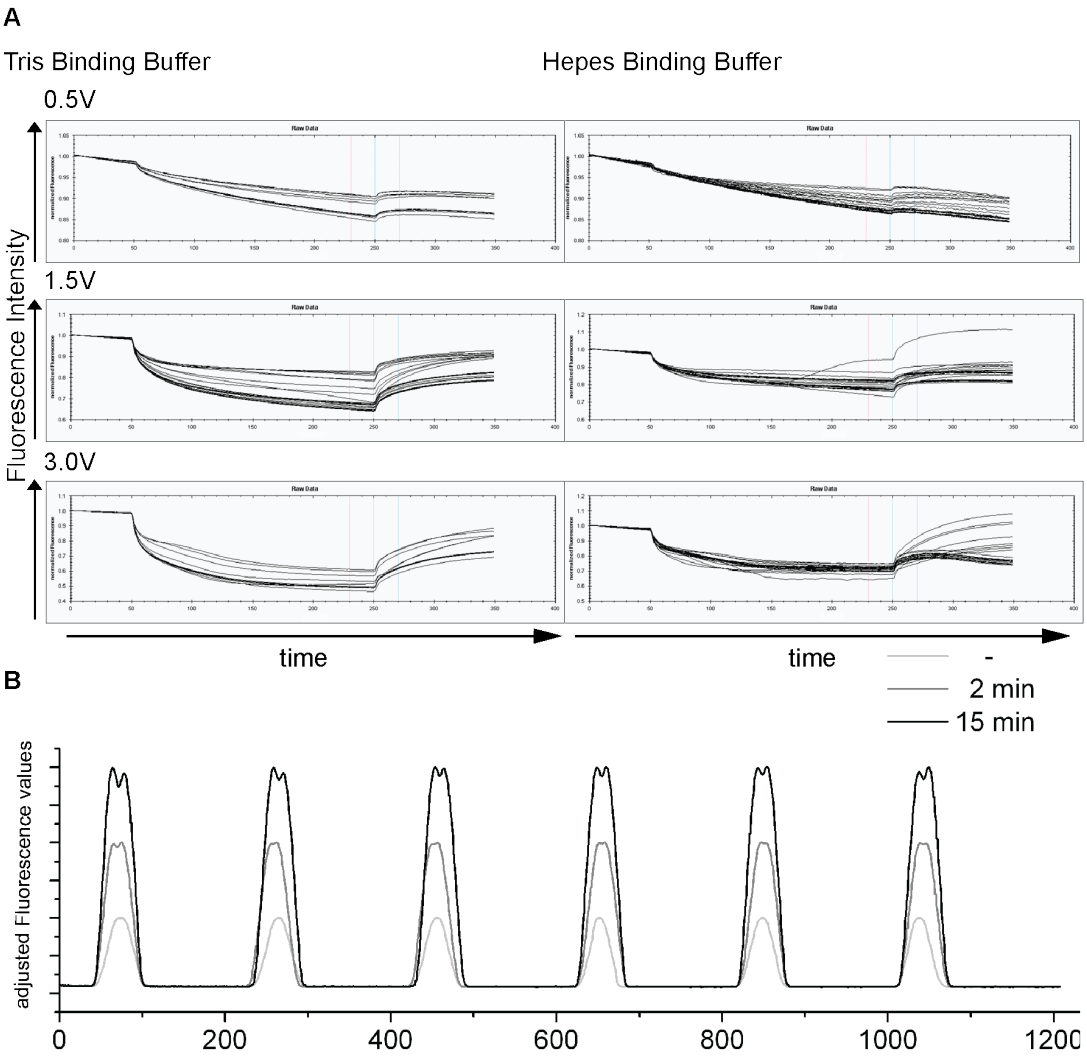

Figure 5.10: Optimization of Thermophoresis measurements
A: 0 - 20µM Loqs dsRBD2 were titrated to 50nM fluorescently labeled RNA in Tris (20mM Tris pH 7.5, 100mM KAc, 3.5mM $MgCl_2$, 0.01% Tween) or Hepes Buffer (30mM Hepes pH 7.0, 100mM KAc, 2mM MgAc, 1mM DTT) and measured in the Nanotemper Monolith with varying Laser strengths (0.5 - 3V). Shown are the displacement curves from the area heated by the laser in fluorescence intensity over time: After recording the steady state before displacement the laser is switched on and the RNA migrates out of the focus. After the fluorophore distribution has again reached a steady state the laser is switched off and the RNA diffuses back in. The curves in Tris buffer with 1.5V showed the best signal to noise ratio and was subsequently used for Thermophoresis measurements. B: Scan of capillaries containing GFP-LoqsPB. Six capillaries were scanned perpendicular to their axis and the fluorescence was recorded. Scans were conducted immediately after filling of the capillaries and 2 and 15 minutes later. Fluorescence profiles were rescaled so that each time point was visible. After 15min a clearly visible dent in the fluorescence profile of the capillaries was appeared, indicating adsorption of protein to the capillary wall.

tion of RNA amount was confined to the detection limit of fluorescein in the Typhoon Scanner, the Tecan plate reader or the Monolith, representing EMSA, anisotropy and thermophoresis measurements, respectively.

In the Monolith, RNA concentrations were tested with an excitation LED power of 200. The capillaries with 5 and 10nM RNA were only detected using the threshold method in the capillary scan, and the thermophoresis curves did not show the expected shape . These problems were eliminated for RNA concentrations above 20nM. 50nM RNA were used in order to be well above the problematic range (figure 5.11A+B).

EMSA and Anisotropy measurements with different RNA concentrations showed that there was indeed an increase in the obtained K_D with higher RNA concentration, whereas the Hill coefficient did not correlate with the amount of RNA (figure 5.11C). At a concentration of 50nM and less, the dissociation constants are relatively constant in the case of fluorescence anisotropy measurements, therefore an RNA concentration under 50nM should be used. Further experiments showed that 10nM RNA were still well above detection level and sufficient

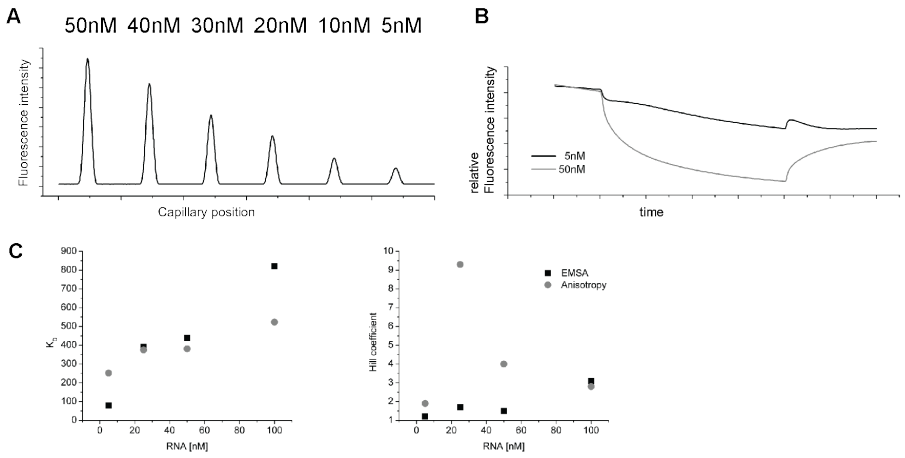

Figure 5.11: Optimization of RNA concentration

A: Thermophoresis. Scans of capillaries containing decreasing concentrations of fluorescently labeled RNA. 5nM and 10nM capillaries were only detected using the threshold method. B: Comparison of displacement curves of 5nM and 50nM RNA. For a sufficiently strong signal an RNA concentration 10nM has to be used. C: K_D dependence on RNA concentration in EMSA and Anisotropy measurements. The K_D values obtained from EMSA and Anisotropy measurements with 5, 25, 50 and 100nM RNA are shown. D: Hill coefficient dependence on RNA concentration in EMSA and Anisotropy measurements. From the same measurements as in C the obtained Hill coefficients are plotted versus the RNA concentration. No correlation could be found.

for noise-free measurements, and at the same time well below the lowest K_D measured in all experiments. Therefore subsequent measurements were performed with a 10nM RNA concentration.

5.3.5 Attached Fluorescein does not interfere with RNA - protein binding

Even though fluorescein is not as bulky as for example GFP, it might still interfere with RNA - protein binding. To exclude this possibility, I compared binding of LoqsPD to a fluorescently and a radioactively labeled siRNA duplex in an EMSA. The resulting binding curve reveals a similar binding behavior of LoqsPD to both types of labeled RNAs, implying that the fluorophore is not disturbing the interaction between protein and RNA (figure 5.12).

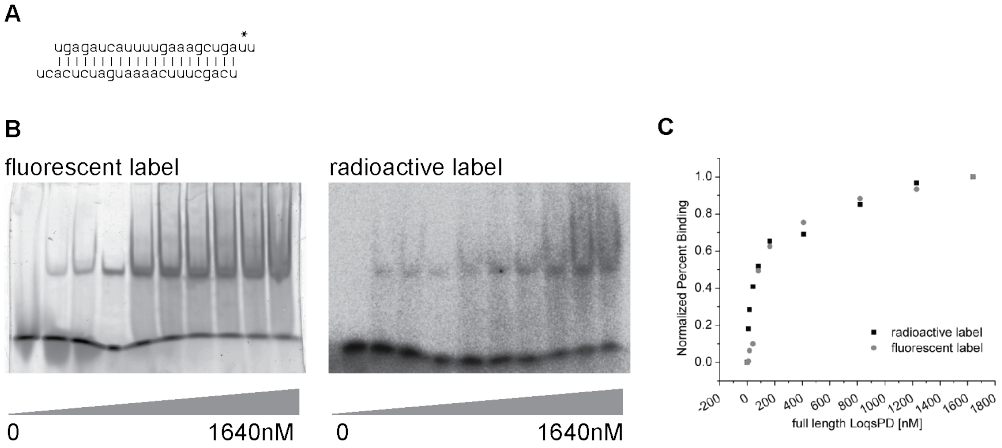

Figure 5.12: Fluorescein does not interfere with RNA-protein binding
A: RNA oligo used for binding experiments. The * depicts the position of the fluorophore in the fluorescently labeled RNA duplex. B: Increasing amounts of full length LoqsPD were titrated to 100nM fluorescently or 6.25nM radioactively labeled siRNA. C: Quantification of the shifts in A. Bands were quantified using the MultiGauge software, the bound/total ratio was calculated and the resulting binding curve normalized. They resulted in a K_D of 96nM for the fluorescent and 163nM for the radioactive EMSA, indicating that the attached fluorescein does not interfere with RNA binding of LoqsPD.

5.3.6 Method comparison

The three methods for measuring binding constants, EMSA, Thermophoresis and Anisotropy, each had their advantages and drawbacks, which are discussed in detail below and summarized in table 5.1. For standard binding experiments, the K_D values obtained with the different methods were comparable. I ended up using Anisotropy and EMSA for my experiments, since these two complemented each other and were the most stable in their performance.

Each of the three techniques consumed comparably little recombinant protein. Since for some

proteins expression and purification were problematic, this was advantageous to other methods like surface plasmon resonance or isothermal titration calorimetry, for which large amounts of protein are necessary.

As already mentioned, the amount of labeled RNA should be as small as possible. In that regard, EMSA was the most favorable technique, since it could be conducted with radioactively labeled RNA for which the detection limit was much lower than for the fluorescently labeled one. Nonetheless, for easier handling and comparability to the other techniques, fluorescently labeled RNA was used for most EMSA experiments. Due to the required RNA concentration of 50nM, thermophoresis experiments required the largest amount of RNA. Perhaps this could have been lowered to 10nM with more optimization, but regarding the K_D values obtained with this method at that time, it still appeared appropriate.

An important difference between the techniques is the state of the binding reaction during detection: In Anisotropy and Thermophoresis, bound and unbound fractions of the binding reaction are in an equilibrium during detection, whereas in EMSAs bound and unbound species have to be separated from each other in order to be distinguishable during detection. This leads to problems with detecting very transient interactions, since once a protein has dissociated from the RNA, it cannot rebind anymore. Indeed, binding of short-lived complexes like single dsRBDs to RNA were difficult to observe in EMSAs. In Anisotropy measurements, but not in EMSAs, even dsRBD binding to ssRNA could be detected, which was not to be expected and has to be interpreted with caution. The sensitive in-equilibrium methods might even report binding events which would be too short-lived to matter in physiological contexts.

Compared with Anisotropy and Thermophoresis, EMSAs have the advantage of visualizing higher order binding stages, since multimers display a different migration behavior than a 1:1 complex. In theory, binding curves obtained from Anisotropy or Thermophoresis should reach a plateau when one binding stage is complete, and rise again when enough protein is added to begin another one. In practice, the different stages overlapped to an extent that made discrimination via the binding curve impossible. In return, the information that can be gathered by Anisotropy and Thermophoresis measurements can be divers in other respects: During an Anisotropy measurement, the fluorescence intensity can be determined along the way, which can also be used to quantify binding if it occurs close enough to the fluorophore to quench it. Thermophoresis measurements result in three different curves, one with the actual thermophoresis curve, one calculated from the temperature jump, and one from backdiffusion. From each, insights into the nature of binding can be gained (see section 2.4.3).

In terms of time and effort, EMSA experiments are the most demanding, since they require several steps from gel loading over separate detection to a more elaborate analysis, which are not required for Anisotropy and Thermophoresis. Also, EMSA analysis can not easily be automated and thus made independent from individual judgment.

Reproducibility is highest in Anisotropy measurements, even different protein or RNA prepa-

	EMSA	Thermophoresis	Anisotropy
Labeling (γ32-ATP, Fluorescein)	radioactive/ fluorescent	fluorescent	fluorescent
Sensitivity (labeled Partner)	2nM / 10nM	50nM	10nM
Required protein (15 data points, K_D = 200nM)	170pmol	125pmol	340pmol
Time expense	2h	20min	20min
Equilibrium?	no	yes	yes

Table 5.1: Method comparison

rations did not result in changes of K_D. EMSAs led to results with a higher standard deviation. For Thermophoresis measurements no statement concerning reproducibility can be made, since before collecting enough data an unknown parameter was changed and no reasonable binding curve could be measured anymore. This unfortunately dismissed Thermophoresis for future experiments, despite promising results during optimization.

Especially with a new method, the obtained results should be tested for reproducibility with established techniques. Under the same conditions, the K_D values resulting from Thermophoresis and Anisotropy measurements are indeed comparable (figure 5.13A), validating Thermophoresis as a method for measuring binding constants. For stable complexes, EMSA and Anisotropy experiments also yielded approximately the same results (figure 5.13B), transient interactions could not be visualized in EMSAs.

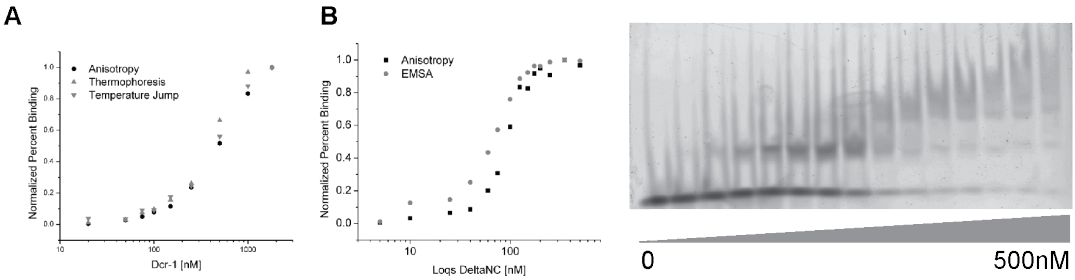

Figure 5.13: Method comparison
A: Dcr-1 was titrated to 100nM fluorescently labeled RNA and Thermophoresis and Anisotropy measurements were performed. The obtained K_D values were 440nM and 406nM from the Anisotropy and the Thermophoresis binding curve, respectively. B: Loqs DeltaNC was titrated to 10nM fluorescently labeled siRNA and EMSA and Anisotropy measurements were performed. The resulting binding curves were fitted and $K_D$values of 70nM and 90nM for EMSA and Anisotropy binding curves, respectively, were determined.

5.3.7 Thio-Uridine crosslinks replicate natural binding behavior

Photochemical crosslinking of proteins and RNA can be induced via irradiation with short-wave UV light (245nm), which converts the amino acids or the nucleotides to reactive species leading to covalent bonds between them. For site specific crosslinking, a Thiouridine can be introduced in the RNA, which is activated at longer wavelengths (365nm), where only these and no unspecific crosslinks occur [98]. The oligos used in this thesis are summarized in section 5.8, figure 5.29.

All thiouridine positions were functional, as could be shown by crosslinking Loqs DeltaNC to single stranded RNA (see figure 5.25). To assess possible impairment of crosslinking efficiency by changes in the spectral properties of thiouridine induced by base pair stacking interactions, absorption of the thiouridines in a single and double-stranded environment were measured, but no change which would explain a preferential crosslinking in single-stranded regions could be observed (figure 5.14A). Also, thiouridine residues in a dsRNA environment were shown to be capable of crosslinking to associated proteins [1], thus the obtained crosslink profiles should be a realistic representation of the protein distribution on the RNA.

To see whether crosslinking also fixates short-lived, abnormal protein binding states, I titrated increasing amounts of Loqs DeltaNC to an siRNA duplex containing a thiouridine at one end and compared the resulting binding curve after crosslinking with one obtained from native EMSA (figure 5.14B). The resulting dissociation constants were 134nM and 120nM, respectively. Therefore crosslinking appears to capture primarily the protein bound in the same configuration as the one detected in our binding studies.

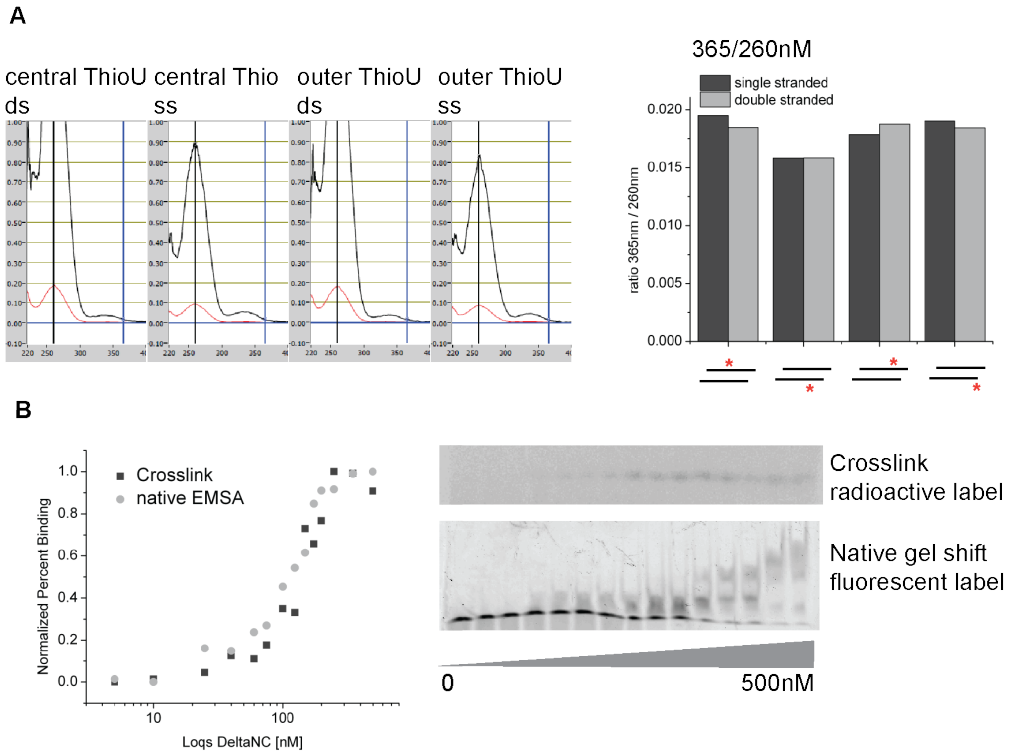

Figure 5.14: Thiouridine crosslinking replicates natural binding behavior
A: left: Absorption measurement of ds and ssRNA oligos containing a Thio-U at different positions. Central Thiouridine: At least 9 base paired nucleotides on both sides of the Thiouridine when the duplex is double stranded; outer Thiouridine: the base pair formed by the Thiouridine and the partner adenosine is the last one in the duplex. 260nm (absorption maximum of RNA) and 365nm (absorption of Thiouridine) are indicated with a black and blue line, respectively. right: A_{365nm}/A_{260nm} ratios are shown for ds and ssRNA oligos with Thiouridines at the positions indicated below the bars. The values constitute a mean of two independent measurements. No difference is seen between the absorption of Thiouridine in a single stranded or a double stranded environment. B: Increasing amounts of Loqs DeltaNC was titrated to 10nM fluorescently labeled siRNA or radioactively labeled siRNA with a Thiouridine at the end. The fluorescent EMSA was run on a 4% native Acrylamide gel, the radioactive EMSA was run on a denaturing Acrylamide gel after crosslinking RNA and protein. The resulting binding curves are shown on the left, the EMSA bands on the right. The obtained K_Ds were 119nM and 134nM for the native and the crosslinked EMSA, respectively.

5.4 Binding behavior of the Loquacious double-stranded RNA binding domains

Since direct comparison of full length LoqsPD and R2D2 were not possible due to the instability of R2D2, I wanted to get a general idea of how the dsRBPs recognize and bind their substrates. This section deals with the binding behavior of the building block of the dsRBPs, an isolated double stranded RNA binding domain, with the example of the two dsRBDs found in LoqsPD. A comparison with the R2D2 dsRBDs is included as well.

5.4.1 Both Loqs RNA binding domains bind RNA with a similar K_D

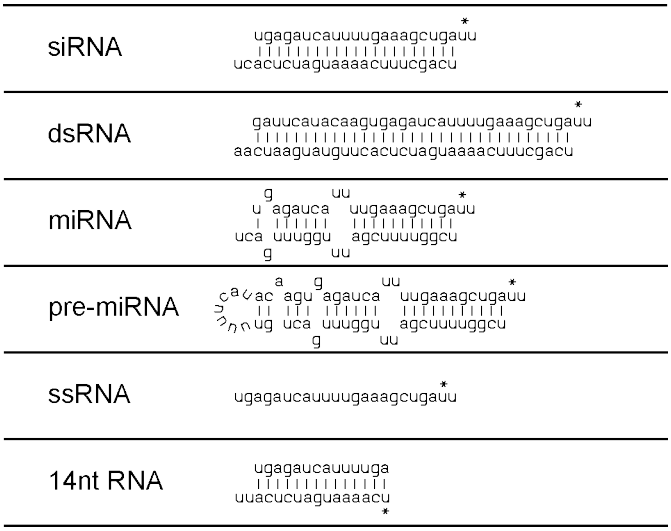

Figure 5.15: RNA substrates used for binding experiments
List of the RNA substrates used in binding experiments. Their sequence is derived from miRNA bantam, for the siRNA substrates the mismatches were removed. The * indicates the position of a Fluorescein.

The two RNA binding domains are very similar, both on the sequence and the structural level, which suggests a similar binding behavior. I performed fluorescence anisotropy measurements to determine the binding affinity of both domains to an siRNA-like substrate, a completely base-paired RNA duplex derived from the bantam sequence with the characteristic 3' 2nt overhangs and 5' phosphates (see figure 5.15). The obtained binding curves were fitted with the Hill equation and results from at least 3 experiments were averaged.

Both domains bind siRNA with a K_D of 228nM and 278nM (see table 5.2), which is in accordance with the K_D values published for the TRBP dsRBDs [116]. dsRBD1 has a trend for tighter binding compared with dsRBD2, but this difference is too small to be completely resolved with this type of measurement and judged not significant by an unpaired, heteroscedas-

	Loqs dsRBD1	Loqs dsRBD2	Loqs DeltaNC	LoqsPD full length
siRNA	228nM +/- 27nM	278nM +/- 24nM	82nM +/- 21nM	50nM +/- 6nM
dsRNA	191nM +/- 29nM	229nM +/- 48nM	86nM +/- 18nM	62nM +/- 5nM
miRNA	233nM +/- 17nM	241nM +/- 75nM	97nM +/- 21nM	118nM +/- 17nM
pre-miRNA	221nM +/- 43nM	218nM +/- 61nM	121nM +/- 38nM	100nM +/- 12nM
14nt RNA	365nM +/- 62nM	334nM +/- 8nM	128nM +/- 41nM	118nM +/- 15nM
ssRNA	265nM +/- 164nM	259nM +/- 40nM	138nM +/- 47nM	328nM +/- 204nM

Table 5.2: K_D values

K_D values of one protein - RNA pair were obtained by fitting each binding curve separately with the Hill formula (equation 4.1) and averaging the obtained values. The indicated errors represent the standard deviation of these values.

tic t-test. Even though the dissociation constants of the two dsRBDs do not differ significantly, some differences in their binding behavior can be observed. The binding curves depicted in figure 5.16A show a steeper progression for dsRBD1, and the curves intersect before the half-maximum is reached, also potentially indicating a slightly stronger binding by dsRBD1. To determine possible interference of unproperly folded protein, I also conducted competition experiments to determine the K_D (figure 5.16B). In competition titrations, the amount of active protein does not affect determination of the K_D. The datapoints were fitted with equation 4.2, and the resulting parameters were averaged. Even though the K_D values obtained with this method were generally lower than with direct measurements (dsRBD1: 24nM +/- 18, dsRBD2: 39nM +/- 9), they were still comparable for both dsRBDs. The amount of properly folded protein in the reaction was also a parameter of the fit function, and it conveyed that 48% of dsRBD1 but only 18% of dsRBD2 contributed to the binding reaction. If this was indeed the case, the K_D of dsRBD2 should be significantly lower than of dsRBD1, since similar affinities were determined for both in direct measurements, where the active concentration of both proteins is assumed to be the same. Since equation 4.2 models a system where one binding partner binds to one available binding site, it is not completely applicable in this scenario: up to 4 dsRBDs have been shown to bind to one miRNA duplex [5]. The influence of this stoichiometry on the fitting parameters is not readily predictable, therefore the results obtained from direct measurements are more reliable.

Even though both dsRBDs bind to siRNA with similar affinity, in EMSA experiments only dsRBD2 could induce a shift of the siRNA substrate (figure 5.17). This suggests a qualitatively different binding behavior of both dsRBDs that is not necessarily reflected in the K_D .

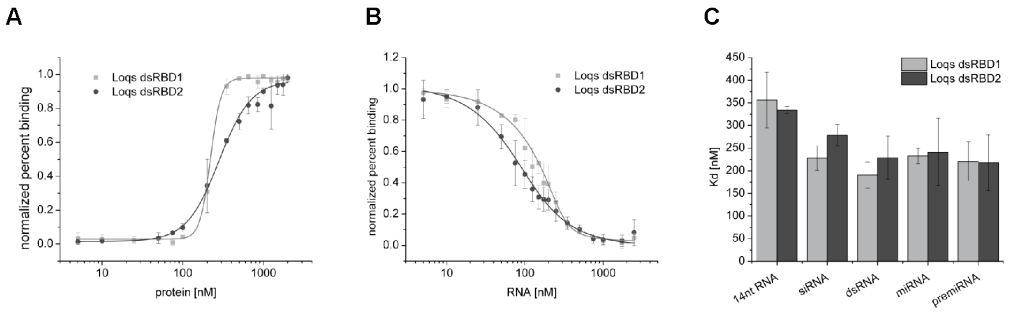

Figure 5.16: siRNA binding of Loqs dsRBDs

A: Binding curves of Loqs dsRBD1 and dsRBD2 to siRNA. Protein was added in increasing amounts to 10nM fluorescently labeled RNA and Anisotropy values were measured. At least three independent measurements are averaged. The error bars represent the standard deviation of each data point. Both dsRBDs bind siRNA in the same nM range. B: Competition experiments of Loqs dsRBD1 and dsRBD2. 500nM Protein was mixed with 7.5nM fluorescently labeled siRNA and increasing amounts of unlabeled siRNA were added. Anisotropy values were measured in three independent measurements and the averaged displacement curve is shown. The error bars represent the standard deviation of each data point. C: K_D values for dsRBD binding to five different RNA substrates. K_D values of one protein - RNA pair were obtained by fitting each Anisotropy binding curve separately with the Hill formula (equation 4.1) and averaging the obtained values. The indicated errors represent the standard deviation of these values. The dsRBDs do not exhibit substrate specificity.

5.4.2 Loqs dsRBDs do not show target specificity

siRNA duplexes are the products of the dicing reaction by Dcr-2 that is aided by LoqsPD [70]. To see whether also the educt is bound, and whether binding is influenced by mismatches in the RNA duplex, I also measured the dissociation constants of the dsRBDs to 35nt long dsRNA, miRNA and pre-miRNA substrates, all derived from the bantam sequence. Additionally, we engineered a short, 14nt duplex with the Fluorophore attached directly at a blunt end (see figure 5.15).

The resulting dissociation constants are summarized in table 5.2 and figure 5.16C. All substrates were bound with approximately the same affinity by both dsRBDs (for significance of differences in the K_D values, see Appendix 6.6). Therefore, distinction of structurally different substrates is not occurring on the level of individual dsRBDs.

5.4.3 Loqs dsRBD binding is rather transient

When probed in an EMSA, the ability of both dsRBDs to shift various substrates differ: with dsRNA, a strong shift could be observed for both dsRBDs, whereas miRNA was not shifted at all (see figure 5.17). As already mentioned, only dsRBD2 was able to shift siRNA. Since in EMSAs the binding equilibrium is disrupted, lack of binding observed with this method

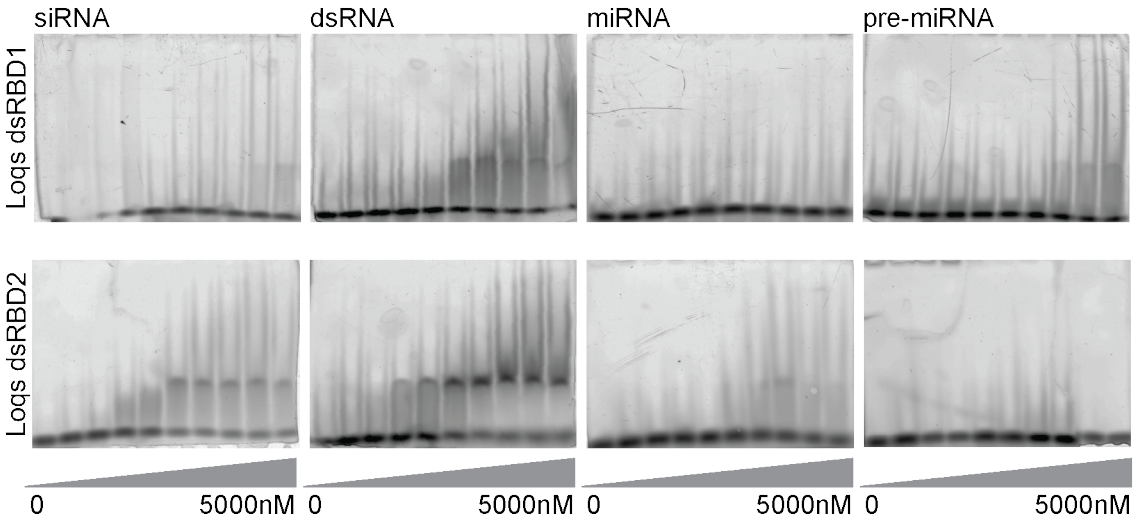

Figure 5.17: Binding of individual Loqs dsRBDs is rather transient
EMSA experiments with Loqs dsRBDs and various RNA substrates. Protein was added in increasing amounts to 10nM fluorescently labeled RNA and the binding reaction was loaded on a 4% native Acrylamide gel. Gels were run at 200V for 20 min in 0.5x TB buffer and read out with a Typhoon scanner. Both dsRBDs are able to shift the 35nt dsRNA, but only dsRBD2 induced a shift of the 23nt siRNA.

suggests a transient complex that falls apart too quickly to be maintained during the separation of bound and unbound species. dsRBD EMSAs were not very reproducible and most of the time no shift could be observed for any substrate, therefore only qualitative statements can be made: In general, dsRBD2 shifts were more prone to succeed, and dsRNA is bound best.

5.4.4 Comparison with R2D2 dsRBDs

Since we were interested in the differences between LoqsPD and R2D2 that could explain differential processing of siRNA groups, I tried to gain insight into R2D2 binding as well.

Binding experiments with R2D2 dsRBDs were noisier than those with the Loqs dsRBD, as can be seen in the error bars of the binding curves (figure 5.18A). It is conceivable that even the isolated dsRBDs of R2D2 might be unstable and that in every protein preparation a different amount of active protein is left. Given the fact that R2D2 is only stable together with Dcr-2, a general frailty of the protein fold might be possible. Therefore, the dissociation constants should not be interpreted directly. However, a comparison of the same dsRBD binding to different RNA substrates can be made, since the protein should be similarly active in all binding reactions performed in parallel. It seems that the R2D2 dsRBDs, like the Loqs dsRBDs, can not distinguish between different substrates (figure 5.18). R2D2 dsRBD2 might have a tendency to prefer dsRNA, though considering the large error bars this is not significant.

It can be stated that R2D2 dsRBDs are either rather unstable or bind to RNA substrates much

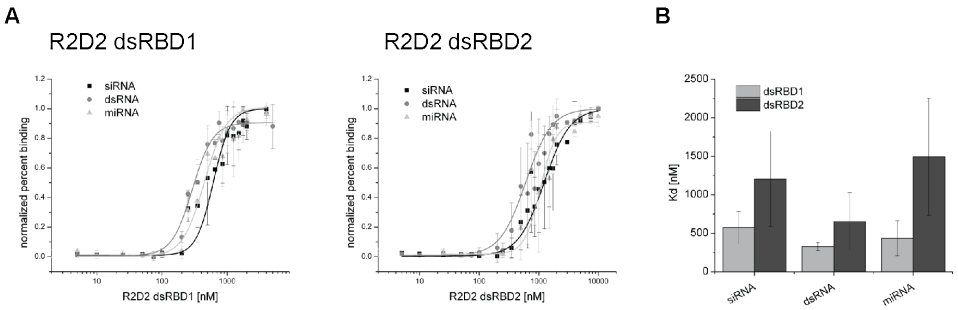

Figure 5.18: R2D2 dsRBD binding
A: Binding curves of R2D2 dsRBD1 and dsRBD2 to siRNA. Binding curves were generated as described. B: K_D values for dsRBD binding to five different RNA substrates. K_D values of one protein - RNA pair were obtained by fitting each Anisotropy binding curve separately with the Hill formula (equation 4.1) and averaging the obtained values. The indicated errors represent the standard deviation of these values. The dsRBDs do not seem to exhibit substrate specificity.

more weakly than Loqs dsRBDs. The missing conservation of AA residues involved in RNA binding supports the latter. Like Loqs dsRBDs, R2D2 dsRBDs do not seem to preferentially bind a particular RNA substrate.

5.5 Interplay of the two Loquacious double-stranded RNA binding domains

This section describes the way the two dsRBDs of LoqsPD interact to achieve the affinity and substrate specificity of full length LoqsPD.

5.5.1 Full length LoqsPD and Loqs DeltaNC show enhanced RNA binding affinity

To see how the two dsRBDs act together in RNA binding, I performed Anisotropy measurements to determine the affinity of full length LoqsPD and Loqs DeltaNC, which entailed both domains connected by the linker (figure 5.20), for the siRNA substrate. The binding curves of single and double domain constructs were clearly distinguishable (figure 5.19A), and LoqsPD full length and Loqs DeltaNC exceeded the isolated dsRBDs in binding strength by a factor of 2 - 3 (figure 5.19B). EMSAs yielded values in the same nM range, corroborating the validity of the results obtained from anisotropy measurements. EMSAs also confirmed the tighter siRNA binding of the double domain proteins compared to the single domains (figure 5.19C), since LoqsPD full length and Loqs DeltaNC - siRNA complexes were far more stable and always induced an RNA shift in the EMSAs.

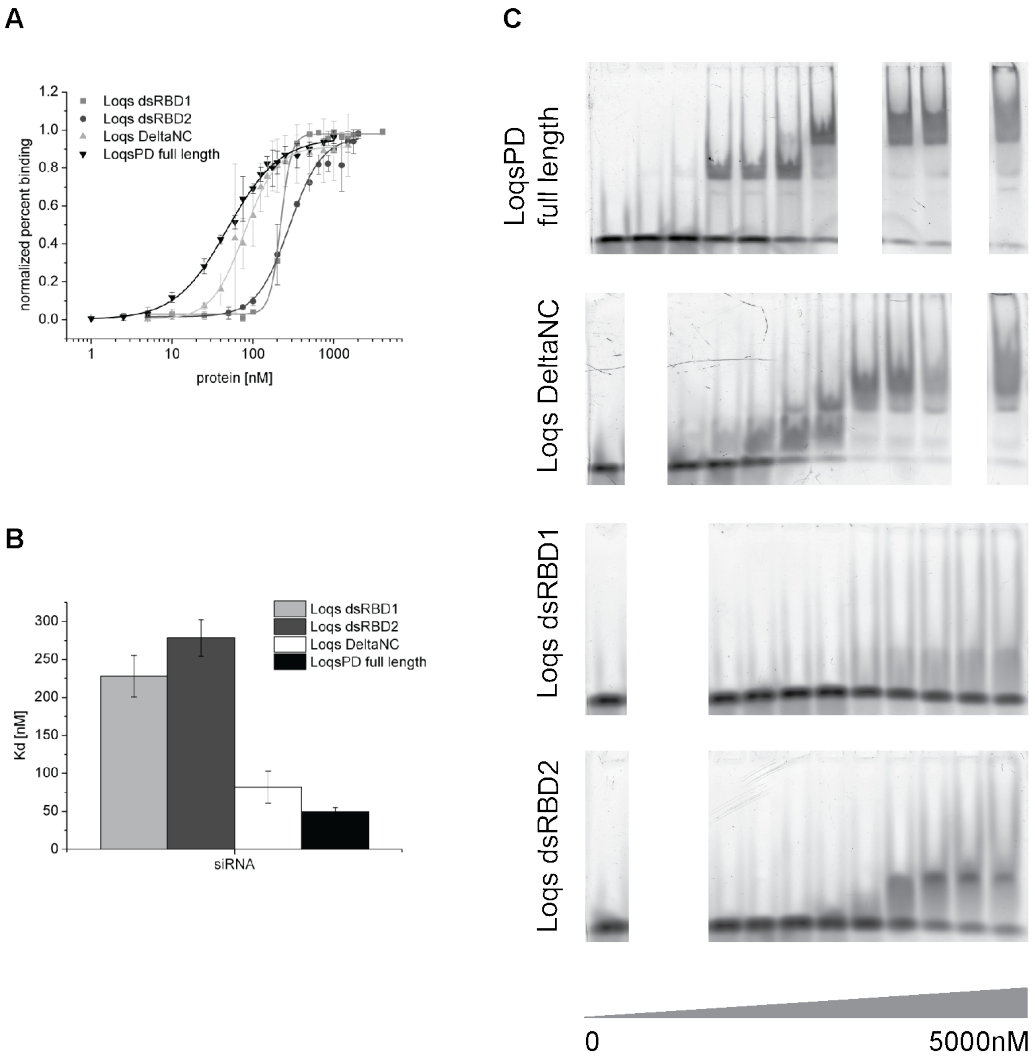

Figure 5.19: Full length LoqsPD and Loqs DeltaNC show enhanced RNA binding affinity
A: Binding curves of all Loqs constructs to siRNA. Binding curves were generated as described. LoqsPD full length and Loqs DeltaNC - siRNA binding curves are shifted to lower protein concentrations compared to those of the dsRBDs, indicating a lower K_D. B: K_D values for siRNA binding by all the Loqs constructs. K_D values of protein - siRNA pairs were obtained as described. The double domain constructs bind up to five times stronger than the individual dsRBDs. C: EMSA experiments with all Loqs constructs and siRNA. EMSAs were conducted as described. The lanes are arranged in a way that the added protein amount is the same for lanes at the same position. EMSAs confirmed the tighter siRNA binding of the double domain proteins compared to the single domains.

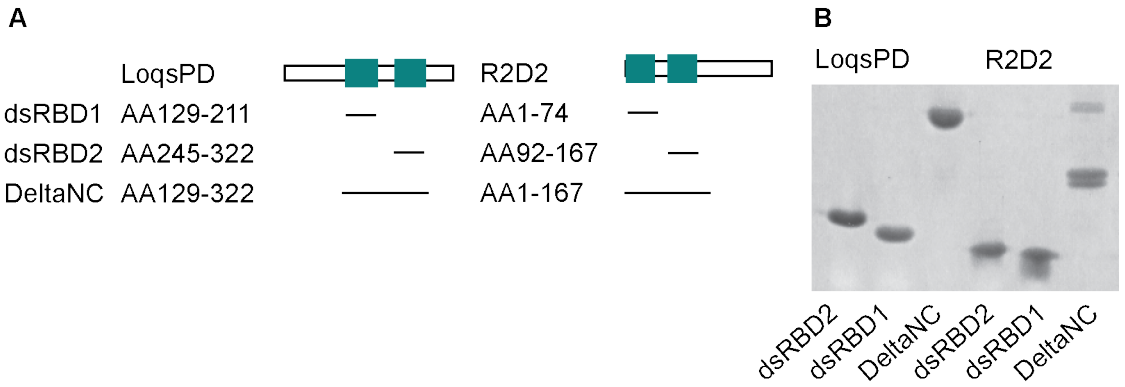

Figure 5.20: Protein constructs used in binding experiments
A: Schematic representation of LoqsPD and R2D2. dsRBDs are depicted as blue squares. The amino acids constituting the respective protein constructs are depicted below. B: Purified protein constructs. The respective molecular masses are: LoqsPD dsRBD1: 9kD, dsRBD2: 9.1kD, DeltaNC: 21kD; R2D2 dsRBD1: 7.1kD, dsRBD2: 7.8kD, DeltaNC: 17.4kD. Purification of R2D2 DeltaNC always resulted in the double band seen here.

5.5.2 Full length LoqsPD has an inherent bias towards siRNA binding

The commitment of LoqsPD to the siRNA biogenesis pathway might arise from substrate specificity of LoqsPD, from interaction of LoqsPD with Dcr-2 or a combination of both. In anisotropy measurements, the individual dsRBDs did not show any substrate specificity, therefore the interplay of both domains might be necessary for substrate distinction. Since Loqs DeltaNC constitutes the core of LoqsPB as well as LoqsPD, the PD specific parts might also be necessary to increase the affinity for siRNA compared to miRNA. To test this, I determined the binding affinities of Loqs DeltaNC and LoqsPD full length to the complete set of RNA substrates (figure 5.15). Loqs DeltaNC bound all RNA substrates with similar affinity, with a slight trend for stronger si and dsRNA binding (figure 5.21A), which was not significant (see Appendix 6.6). Full length LoqsPD showed a significantly increased affinity for siRNA and dsRNA compared to Loqs DeltaNC, whereas their affinity for the mismatched substrates was approximately the same. The preference for perfectly matched substrates can also be seen in EMSAs of full length LoqsPD with the various RNAs (figure 5.21B).

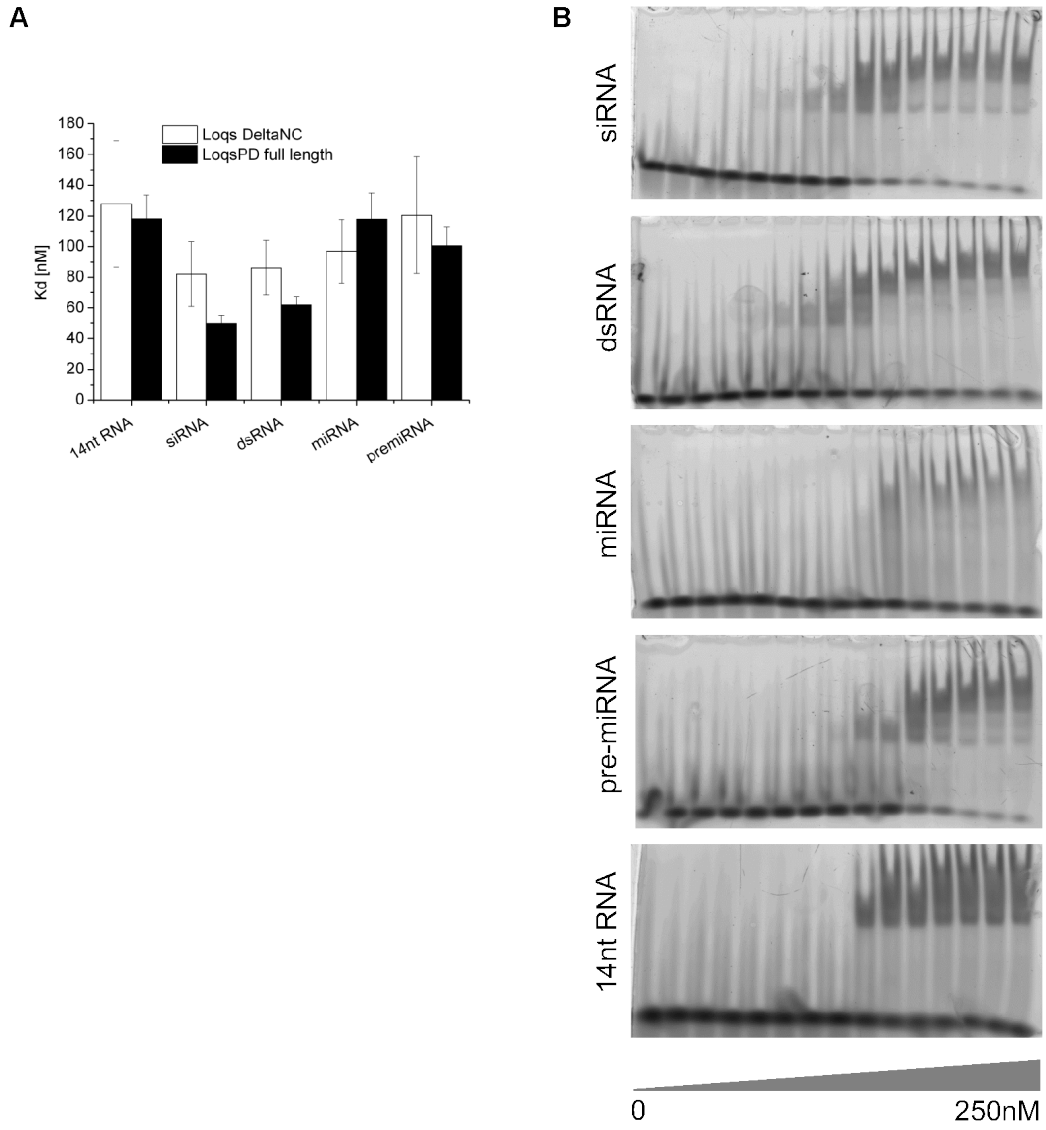

Figure 5.21: Full length LoqsPD has an inherent bias towards siRNA binding
A: K_D values for Loqs DeltaNC and LoqsPD full length binding to five different RNA substrates. K_D values of one protein - RNA pair were obtained as described. LoqsPD full length exhibits significantly stronger binding of completely base paired RNA substrates of a sufficient length compared to mismatched RNA substrates and the 14nt dsRNA. B: EMSA experiments with LoqsPD full length and various RNA substrates. EMSAs were conducted as described. EMSAs confirmed the tighter siRNA and dsRNA binding by LoqsPD full length compared with the other RNA substrates.

5.5.3 Full length LoqsPD has the highest propensity to distinguish between RNA substrates

As can already be inferred from its increased siRNA binding strength, the ability to discriminate between RNA substrates is most pronounced in full length LoqsPD, and the differences in binding strengths for si- and miRNA targets were significant in an unpaired heteroscedastic t-test (see Appendix 6.6). This can already be seen in the averaged binding curves of the four protein constructs, which run very close to each other for all the substrates, except those of full length LoqsPD, which separate to some extent (see figure 5.22A).

To visualize the different binding behavior to the various substrates, the K_D values of each protein construct were normalized to their affinity for siRNA, and the change in affinity for substrates with different structure and length were visualized (figure 5.22B, left and right, resp.). Only two dsRBDs together showed a preferential binding to the siRNA mimic compared to miRNA-like structures, and only the distinction made by full length LoqsPD is significant. This argues for involvement of residues beyond the dsRBDs in limiting the substrate range of LoqsPD. For no protein construct the affinity for the longer dsRNA is increased, whereas the affinity for the 14nt RNA is marginally decreased for all proteins. The affinity of LoqsPD full length for the 14nt RNA is decreased significantly compared with siRNA and dsRNA substrates, endorsing the idea of additional residues involved in binding, which the shorter oligo might not be able to accommodate. Nevertheless, the 14nt oligo is bound remarkably well by all Loqs constructs.

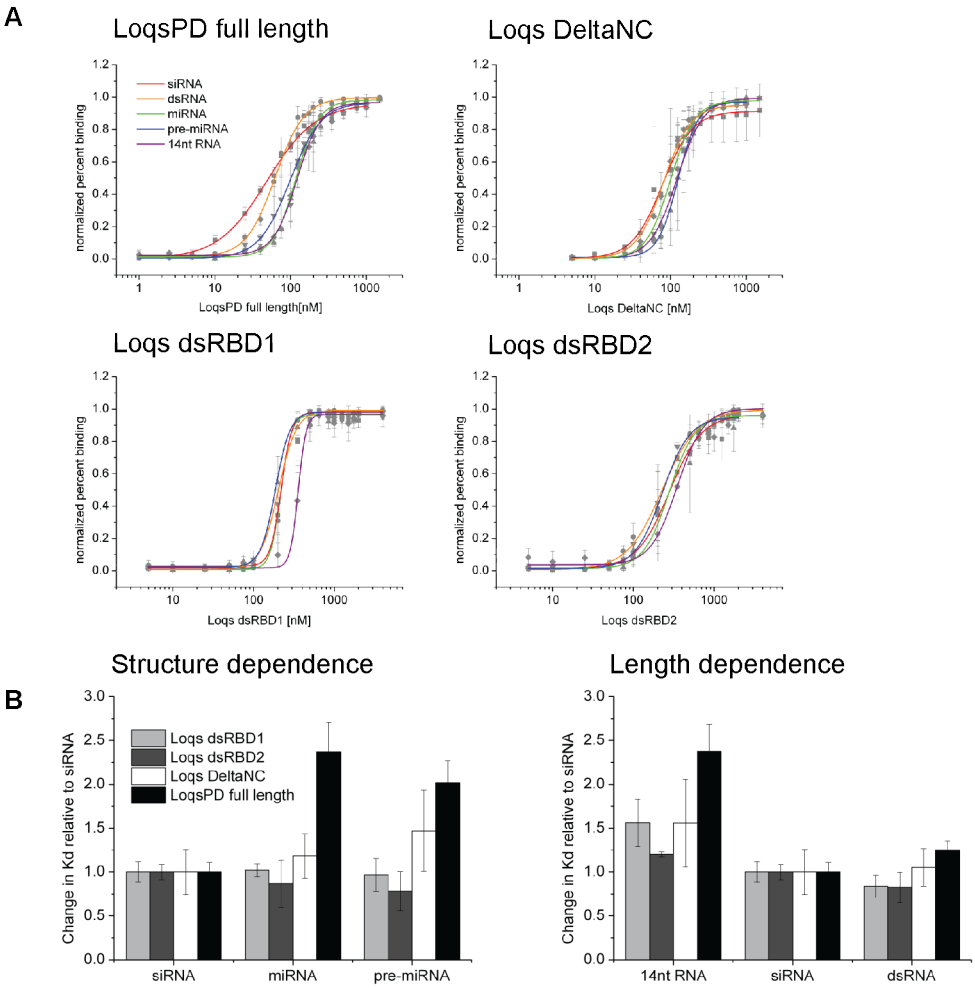

Figure 5.22: Full length LoqsPD has the highest propensity to distinguish between RNA substrates

A: Overview over binding curves of all Loqs constructs to all RNA substrates. Binding curves were generated as described. Compared to the dsRBD2 binding curves, the LoqsPD full length binding curves are more expanded, indicating more variability in the K_D values. B: For better visualization of substrate specificity of the Loqs constructs, the K_Ds shown in the bar graphs of figure 5.16 and 5.21 were normalized to the K_D of the respective siRNA binding of each protein construct. Left: Comparison of substrate preference based on the dsRNA structure. Only the double domain constructs show decreased binding affinity for mismatched dsRNA substrates, with LoqsPD full length making the largest difference. Right: Comparison of substrate preference based on dsRNA length. Length increase does not lead to an increased binding affinity of any substrate, Length decrease results in decreased binding affinity, particularly of LoqsPD full length.

5.5.4 LoqsPD binds two different sequences with similar affinity

To exclude any sequence specific effects, the binding affinity of LoqsPD full length to a siRNA and miRNA mimic based on the sequence of miR8 (figure 5.23A) was measured, which yielded $K_D$values comparable to the ones determined for the bantam derived oligos (figure 5.23B). In addition, the difference between the siRNA and miRNA substrates was more pronounced than the difference between the two sequences, as can be seen in the corresponding binding curves (figure 5.23C).

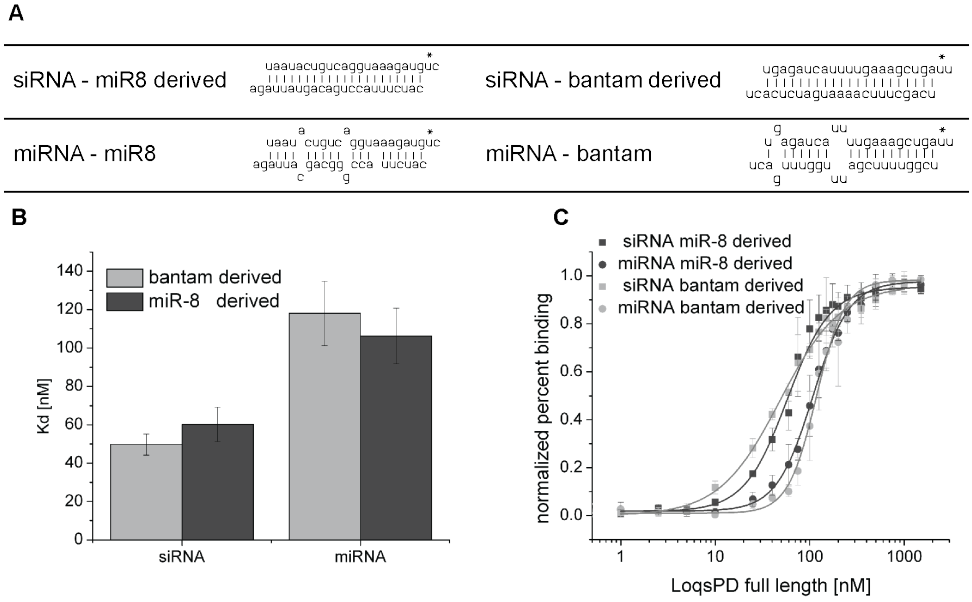

Figure 5.23: LoqsPD full length binding to two different sequences
A: The two different RNA sequences tested for binding to LoqsPD full length. The bantam-derived sequences were used in the preceding binding experiments, the miR8-derived sequences are introduced here. B: K_D values for LoqsPD full length binding to the two different siRNA and miRNA substrates. K_D values of one protein - RNA pair were obtained as described. C: Binding curves of LoqsPD full length to the two different siRNA and miRNA substrates. Binding curves were generated as described. K_D values are approximately the same for both sequences, the differences between the substrates is bigger than the differences between the two sequences.

5.5.5 Comparison with R2D2 DeltaNC

I also measured binding of R2D2 DeltaNC to the RNA substrates. The purified protein always presented a double-band, which might indicate confined degradation. Even with 10µM protein added, no real plateau in the binding curve could be reached (figure 5.24). Either R2D2 alone is not able to bind siRNA at physiological concentrations (as postulated, [59]), or it is simply

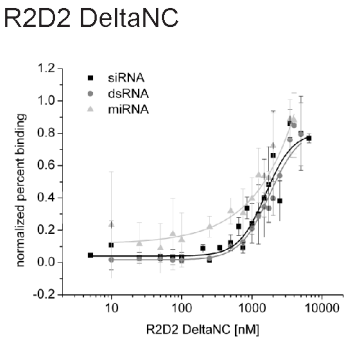

Figure 5.24: R2D2 DeltaNC binding
Binding curve of R2D2 full length to the siRNA, dsRNA and miRNA substrate. Binding curves were generated as described. No plateau was reached for any of the substrates, indicating either a K_D above the range covered in the protein titration or unspecific protein-RNA absorption due to misfolded protein.

not stable enough to retain its proper fold without Dcr-2 and the increase in Anisotropy values is due to trace amounts of correctly folded protein or unspecific protein-RNA adsorption.

5.6 ssRNA binding by double-stranded RNA binding domains

As the name implies, dsRBDs are specialized for dsRNA binding, and their structure is adapted to the geometries of an A-form dsRNA helix (see paragraph 2.3.1.2). Nevertheless, binding experiments in solution showed ssRNA binding by all Loqs protein constructs. Here, I want to characterize this binding further and see whether it has any physiological relevance.

The binding curves and K_D values derived from anisotropy measurements are shown in figure 5.25A and B and are summarized in table 5.2. ssRNA binding is surprisingly strong, about as strong as siRNA binding for dsRBD1+2 and Loqs DeltaNC, whereas the distinction made by LoqsPD full length is more pronounced (figure 5.25A). The K_D values are afflicted with large errors, as are the data points in the binding curves (figure 5.25B), which might reflect very transient ssRNA binding. ssRNA binding is not a feature inherent to the bantam sequence, since also the miR8 ssRNA is bound by Loqs DeltaNC (figure 5.25C). In addition, only high concentrations of full length LoqsPD could induce a ssRNA shift in an EMSA, which was not even reproducible (figure 5.25D). To make another point in favor of transient ssRNA binding, I want to anticipate the crosslinking experiments discussed in section 5.8: To pinpoint the protein binding positions on the RNA, crosslinkable Thiouridines were inserted at different positions in the RNA (see figure 5.29A). The binding reaction was subjected to 365nm light, which induced crosslinks when the protein was in direct contact with the respective Thiouridine. In the case of ssRNA, the crosslinking efficiency for the individual dsRBDs and the deltaNC construct was the same for every position (figure 5.25E). LoqsPD full length seems

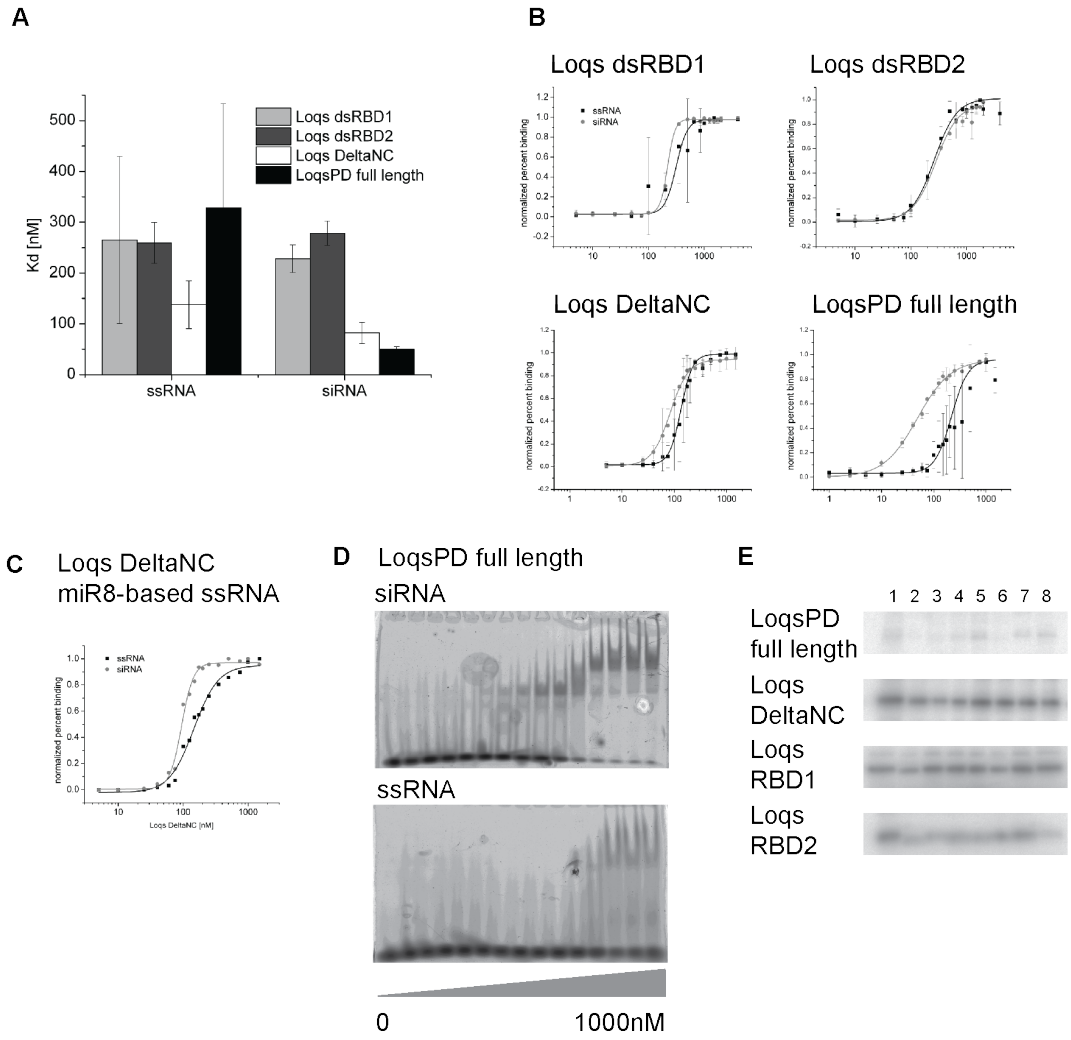

Figure 5.25: ssRNA binding by dsRBDs

A: K_D values for all Loqs constructs binding to ssRNA. For comparison, their K_D values for siRNA are also shown. K_D values of each protein - RNA pair were obtained as described. B: Binding curves of all Loqs constructs to ssRNA (black) and siRNA (grey). Binding curves were generated as described. Except for LoqsPD full length, which shows a decreased ssRNA binding affinity compared to siRNA, all Loqs constructs bind siRNA and ssRNA with a comparable K_D in Anisotropy measurements. C: Binding curve of Loqs DeltaNC to miR8 based ssRNA. miR8 ssRNA is also bound by Loqs DeltaNC in Anisotropy experiments. D: EMSA experiments with LoqsPD full length and ssRNA and siRNA. EMSAs were conducted as described in figure 5.16. LoqsPD full length alone was able to induce a shift of ssRNA at high protein concentrations. E: Crosslink pattern of Loqs protein constructs on ssRNA. Crosslink oligos are depicted in figure 5.29.

to constitute an exception. It has different affinities for ssRNA and siRNA, was the only protein that was able to shift ssRNA in an EMSA, and does not crosslink equally to all thiouridine positions.

5.7 Cooperativity in dsRBD binding

dsRBPs with more than one dsRBD display cooperative binding of their dsRBDs, as introduced in paragraph 2.3.1.4. Since LoqsPD contains two dsRBDs, the extent to which they affect each other's binding was surveyed and the various aspects that provide information about this question are discussed in this section.

5.7.1 Stoichiometry of protein-RNA complexes

How many proteins can a 23nt or 35nt RNA duplex accommodate? Two full length TRBPs or four individual dsRBDs can bind to a miRNA precursor [5]. A 35nt dsRNA is approximately as long as the stem of a pre-miRNA, therefore, it should be able to accommodate the same number of dsRBDs.

EMSA experiments can resolve complexes with different numbers of proteins on the RNA. Dividing the gel in bands with a different molecular weight, up to three different Loqs DeltaNC - siRNA complexes can be distinguished, and up to four different complexes of Loqs DeltaNC and dsRNA (figure 5.26). This has to be interpreted with caution, since different species than increasing numbers of proteins on one RNA molecule can form, like one full length or DeltaNC construct binding two different RNA duplexes. The EMSA pattern indicates repeated binding of a molecule with constant molecular weight (as can be deduced from the uniform decrease in distance between higher order complexes).But with Loqs DeltaNC having a molecular weight of 21kD and siRNA and dsRNA one of 13kD and 20kD, respectively, binding of multiple proteins on one RNA would be hard to discriminate from other possible multimerizations. Nevertheless, assuming a 2:1 complex of Loqs DeltaNC on siRNA and a 3:1 complex on dsRNA is consistent with the available space on the dsRNA, when not every dsRBD is bound directly but is only tethered to the RNA via its partner dsRBD. LoqsPD full length binding is restricted to 2 different complexes with siRNA and 3 with dsRNA (see figure 5.19C and 5.21B), consistent with the stoichiometry mentioned for Loqs DeltaNC. Higher order Loqs DeltaNC complexes are unlikely to involve additional dsRBD-RNA contacts.

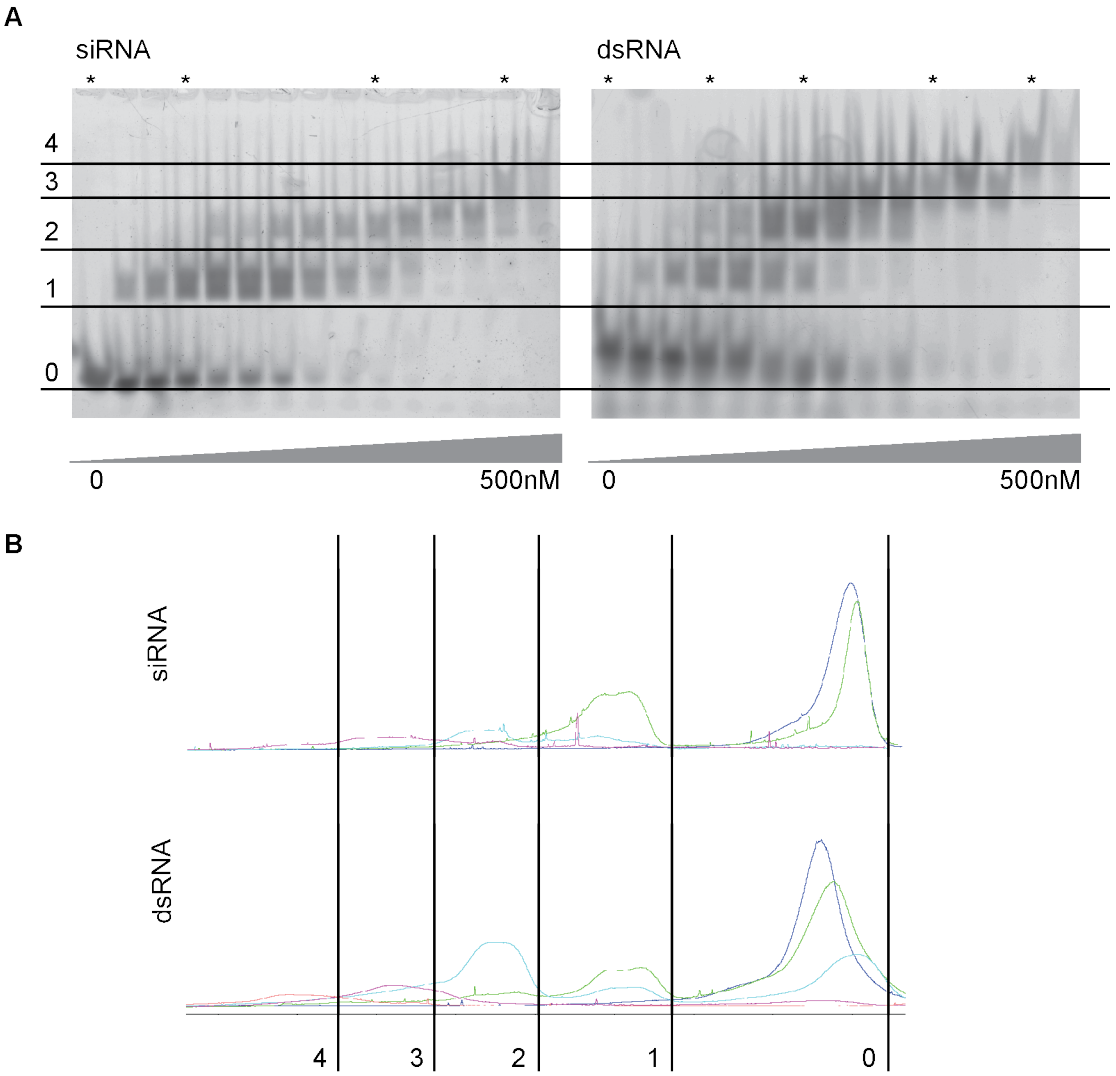

Figure 5.26: Stoichiometry of DeltaNC - dsRNA complexes
A: EMSA experiments with Loqs DeltaNC and dsRNA and siRNA substrates. EMSAs were conducted as described. The lines separate the different protein - RNA species resolved in the EMSA. The indices refer to the number of proteins bound to the RNA, assuming that additional bound protein is the only reason for an increased shift. Free RNA runs at 0. B: Intensity profiles of EMSA lanes indicated with a "*", subdivided in the same regions as in A.

	Loqs dsRBD1	Loqs dsRBD2	Loqs DeltaNC	LoqsPD full length
siRNA	7.9 +/- 3.0	3.4 +/- 2.1	3.3 +/- 0.9	1.3 +/- 0.2
dsRNA	3.9 +/- 1.0	3.0 +/- 1.2	3.2 +/- 0.8	2.5 +/- 0.2
miRNA	5.1 +/- 1.6	3.8 +/- 1.6	5.1 +/-2.9	3.3 +/- 0.8
pre-miRNA	4.5 +/- 0.8	2.6 +/- 0.7	4.6 +/- 0.9	2.4 +/- 0.7
14nt RNA	7.7 +/- 2.4	2.0 +/- 1.7	3.9 +/- 1.0	2.9 +/- 0.7
ssRNA	3.9 +/- 2.6	3.3 +/- 0.5	6.3 +/- 1.9	23.5 +/- 24

Table 5.3: Overview over Hill coefficients

5.7.2 The Increase in K_D is not as large as expected for cooperative binding

The two dsRBDs in Loquacious are connected via a 46AA linker. This is below the 60AA limit calculated for RRMs, at which the two domains interact with RNA independent from each other [95]. Nevertheless, the K_D of Loqs DeltaNC for siRNA is only three times lower than that of the individual dsRBDs, which does not infer a strong binding cooperativity between the domains. In TRBP for example, the individual dsRBDs bind siRNA with a K_D of 220nM and 113nM, whereas full length TRBP binds with a K_D of 0.77nM [83, 116]. This is not the case for the Loqs constructs, where the linker between the two dsRBDs seems flexible enough to uncouple binding of the two dsRBDs.

5.7.3 Loqs binding curves require a Hill coefficient for accurate fitting

When fitting the Loqs-RNA binding curves, a Hill coefficient had to be included to obtain accurate fitting results. When kept fixed at n=1, no satisfactory approximation of the binding curve could be achieved (figure 5.27A, grey vs black curve). The Hill coefficients obtained from the fits are summarized in table 5.3 and figure 5.27B: The largest Hill coefficients are observed for dsRBD1, the lowest for full length LoqsPD. In general, binding to mismatched dsRNAs required larger Hill coefficients. This is not an inherent feature of the method used for measuring the binding curves, since the Hill coefficient for full length LoqsPD binding to siRNA lies between 1 and 2, as does the one obtained for Dcr-1 binding to miRNA and pre-miRNA (figure 5.27C). In addition, Hill coefficients between 2 and 5 were also obtained from EMSA data (figure 5.27D). Wostenberg et al. obtain similar Hill coefficients for siRNA binding by human Dicer's dsRBD, but reject it as a biologically meaningful fitting parameter [114]. Nevertheless, binding of one protein might favor binding of another, via a mechanism that is independent of the linker connecting the two dsRBDs in the case of the double domain constructs.

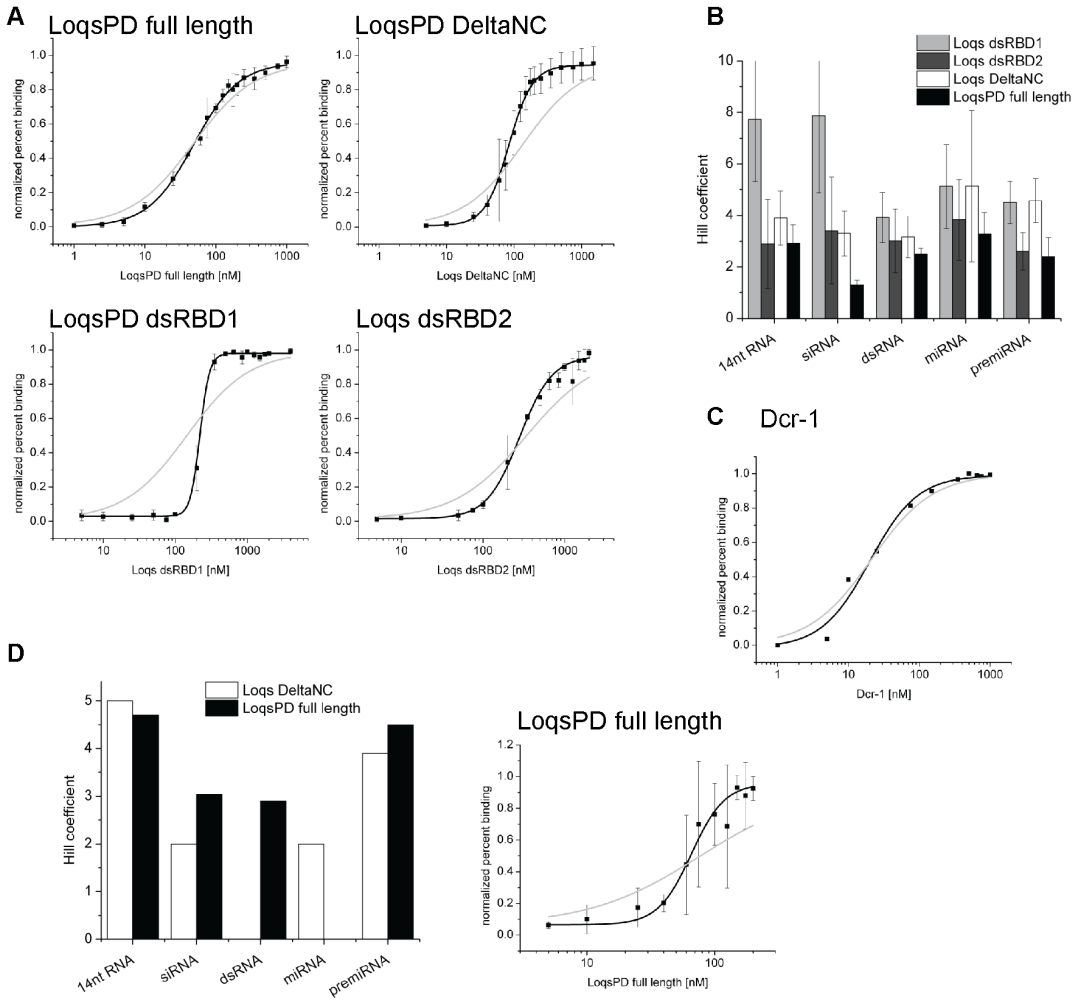

Figure 5.27: Loqs binding curves require a Hill coefficient for accurate fitting
A: Binding curves of all Loqs constructs to siRNA. Binding curves are the same as in figure 5.16 and 5.19. In addition, they are fitted with the Hill equation with the Hill coefficient fixed at 1 (grey curve). For appropriate fitting, a Hill coefficient > 1 is required. B: Hill coefficients describing the binding curves of Loqs constructs to the RNA substrates. Each binding curve was fitted independently and the resulting Hill coefficients were averaged. Only LoqsPD full length binding to siRNA can be described with a Hill coefficient of ~1, all other binding reactions require a far greater Hill coefficient for an optimal fit, especially dsRBD1 binding to 14nt RNA and siRNA. C: Dcr-1- miRNA binding curve fitted with the Hill equation, without any parameter constraints (black curve) and with the Hill coefficient fixed at 1 (grey curve). The grey curve also fits the data reasonably well. D: left: Hill coefficients obtained from fitting EMSA binding curves. EMSA experiments were evaluated using MultiGauge software, and the resulting binding curve were fitted with the Hill equation. The resulting Hill coefficients were averaged. Right: exemplary EMSA binding curve (LoqsPD full length to siRNA), fitted with both unconstrained parameters (black line) and the Hill coefficient fixed at 1 (grey line).

5.7.4 The Hill coefficient does not reflect cooperativity mediated by RNA deformation

One possible way to mediate cooperative binding is deformation of the RNA duplex by the first protein facilitating binding of the second. This has been shown for allosteric protein binding on DNA [46].

To test for this possibility, a nick was introduced in the siRNA duplex, which should interrupt propagation of RNA deformation (figure 5.28A). EMSA experiments showed a shift of the nicked siRNA duplex similar to the one of normal siRNA, indicating that the two halves of the partner strand annealed strong enough to survive this treatment (figure 5.28B). The binding curves of Loqs DeltaNC and dsRBD2 to the nicked siRNA duplex were congruent with the ones to the normal siRNA (figure 5.28C), and fitting resulted in similar Hill coefficients. Therefore, cooperativity mediated by RNA deformation is not the cause for the Hill coefficient.

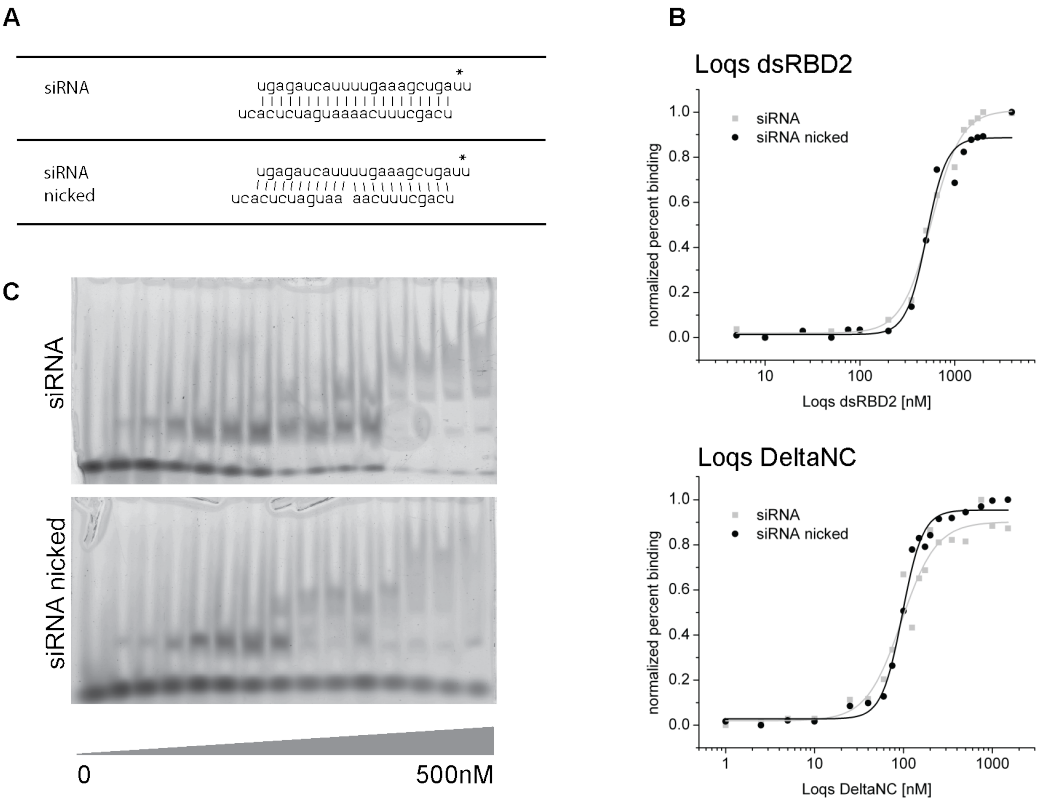

Figure 5.28: Cooperativity is not mediated via RNA deformation
A: intact and nicked RNA oligos used for the measurement. For the nicked oligo, the three RNA parts were mixed in equimolar amounts and annealed by slowly lowering the temperature from 95°C to RT. B: Binding curves of Loqs DeltaNC and dsRBD2 to intact and nicked siRNA. No significant difference between the two RNA substrates can be seen. C: EMSA experiments with Loqs DeltaNC and the two siRNA substrates. EMSAs were conducted as described.

5.8 Geometry of RNA binding

To gain insight into the positioning of the proteins on the RNA, I performed crosslinking experiments with RNA substrates that contained thiouridines at 8 different positions (figure 5.29A). Protein binding in close proximity to the respective Thiouridine resulted in a covalent link upon irradiation with 365nm light, and the complex could be resolved on a denaturing gel.

5.8.1 Proteins crosslink preferentially at the extremities of the siRNA duplex

5.8.1.1 dsRBDs

1µM of the individual dsRBDs were incubated with 10nM radioactively labeled crosslink oligos for 30min at RT in hepes binding buffer. Subsequently, they were irradiated with 365nm light. Crosslinked and uncrosslinked RNA were separated via a 15% SDS Acrylamide gel and read

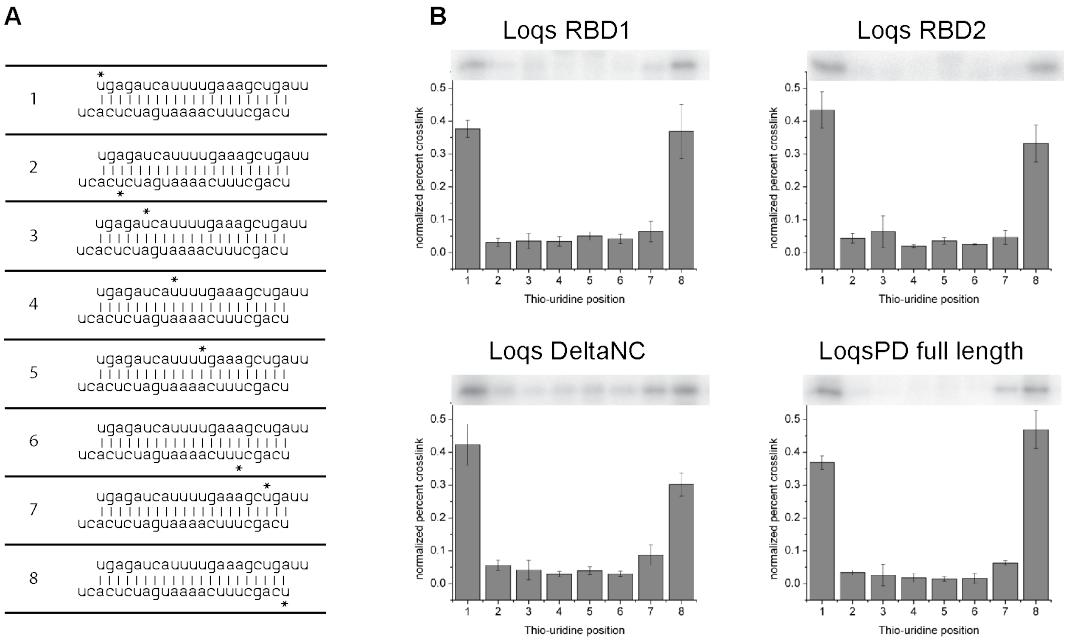

Figure 5.29: Proteins bind preferentially at the extremities of the siRNA duplex
A: siRNA oligos used for crosslink experiments. The "*" indicate Thiouridine positions. B: Crosslink profiles of all Loqs constructs. Proteins were incubated with the respective crosslink oligo for 30min at RT and were subsequently crosslinked with 3x 500mJ/cm^2 of 365nm light. The crosslinked binding reaction was run on a 15% denaturing acrylamide gel for 1h at 170V and used to expose a phosphoimager screen. The banding pattern was analyzed using MultiGauge software, the bound/total ratio was calculated and the values normalized, so that the value for each Thiouridine position constitutes the percentage of protein that is crosslinked there. An example for every crosslink pattern is shown below. All Loqs protein constructs crosslinked preferentially at the ends of the siRNA duplex.

out with a phosphoimager screen. The bands were quantified using MultiGauge software.

Both dsRBDs crosslinked to the siRNA preferentially at the extremities of the duplex (figure 5.29B): 70 - 80% of all crosslinking events occur at nt 1 of the 5'end of either strand, without any preference for a particular end of the duplex.

5.8.1.2 Duplex dsRBDs

500nM Loqs DeltaNC and full length LoqsPD were incubated with 10nM radioactively labeled crosslink oligos. These binding reactions were treated as described above.

The same binding geometry was found for Loqs DeltaNC and full length LoqsPD (figure 5.29C), which were also crosslinking at the ends of the siRNA duplex. It has recently been shown that proteins containing tandem dsRBDs diffuse along their RNA substrates [47]. Movement of the dsRBPs has been visualized via changes in FRET signal of fluorophores attached to one end of the RNA and to the protein. Even though FRET signals are very sensitive for small changes in the distance between the two fluorophores, the transition between minimal and maximal signal was very sharp. Together with our data, this shows that, even though the proteins might move along the RNA helix, they reside most of the time at the ends of the duplex.

5.8.2 Loqs DeltaNC binding is not completely limited to the duplex ends

Even though Loqs DeltaNC crosslinks preferentially at the periphery, residual crosslinking is found towards the middle of the duplex. To see whether this central binding represents a second binding event or whether the first binding event can occur also in the middle of the duplex, Loqs DeltaNC was titrated to different crosslink oligos and the dissociation constants were determined. The resulting crosslinking patterns with the corresponding binding curves are depicted in figure 5.30A and B. Even though the intensity of the crosslink is stronger at the end of the duplex, the affinities are roughly equivalent for all positions (254nM, 206nM, and 357nM for oligo 5, 6, and 8, resp.). This implies that Loqs DeltaNC does not exclusively bind to the duplex ends, but also in the middle of the duplex, albeit at a lower time-averaged occupancy.

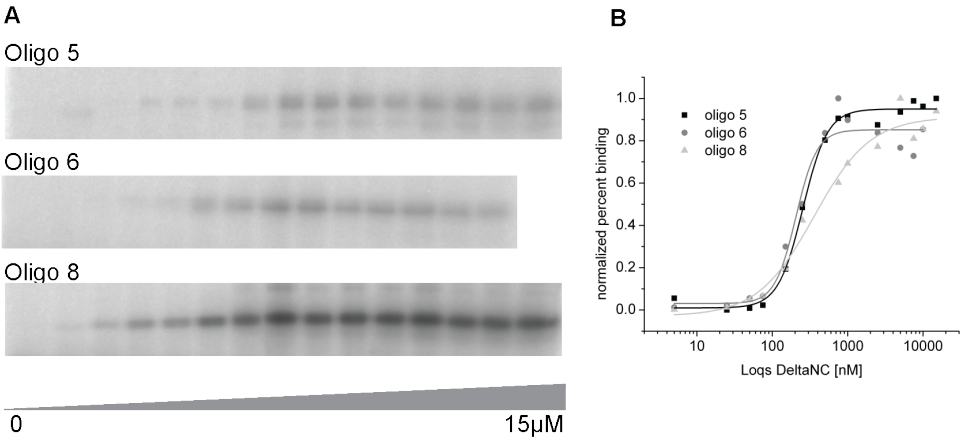

Figure 5.30: Loqs DeltaNC binds also in the center of the siRNA duplex
A: Increasing amounts of Loqs DeltaNC was titrated to crosslink oligo 5, 6, and 8, and crosslinks were performed as described. The resulting bands were analyzed with MultiGauge software. B: Binding curves obtained from the bands in A. All oligos are crosslinked to the same amount over the concentration range.

5.8.3 miRNA binding geometry

For testing miRNA binding geometry, 5 different Thiouridine positions were available (figure 5.31A). Loqs dsRBD1, dsRBD2 and DeltaNC did not show preferential binding to the miRNA duplex ends, whereas LoqsPD full length and Dcr-2 retained approximately the same binding pattern as with siRNA duplexes (figure 5.31B). Remarkably, crosslink oligo 4, which contains a mismatched Thiouridine, is also not crosslinked to LoqsPD or Dcr-2. This further demonstrates that binding to the extremities of the siRNA duplex is not only caused by Thiouridines being inaccessible for crosslinking in a double stranded environment.

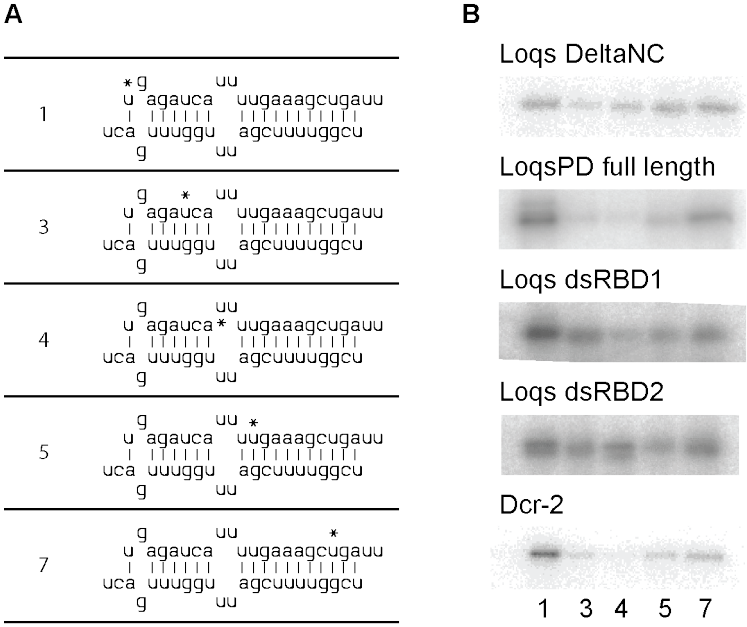

Figure 5.31: miRNA binding geometry
A: miRNA oligos used for crosslink experiments. The "*" indicate Thiouridine positions. B: Crosslink pattern of all Loqs constructs and Dcr-2. Crosslinks were performed as described. Loqs DeltaNC, dsRBD1 and 2 crosslink equally well to all crosslink positions, whereas Dcr-2 and LoqsPD full length rather bind at the extremities of the miRNA duplex.

5.8.4 R2D2

Even though R2D2 DeltaNC did not show reliable binding behavior, I tested whether it could be crosslinked to the RNA. As shown in figure 5.32A, R2D2 DeltaNC was able to crosslink to the siRNA duplex. Like the other dsRBPs, R2D2 DeltaNC also crosslinked preferentially to the duplex ends. To see whether fixing the binding reaction with crosslinks could stabilize R2D2 DeltaNC binding, I titrated increasing amounts to crosslink oligo 1 and analyzed the binding reaction after crosslinking on a denaturing gel. However, the binding curve still did not reach a plateau.

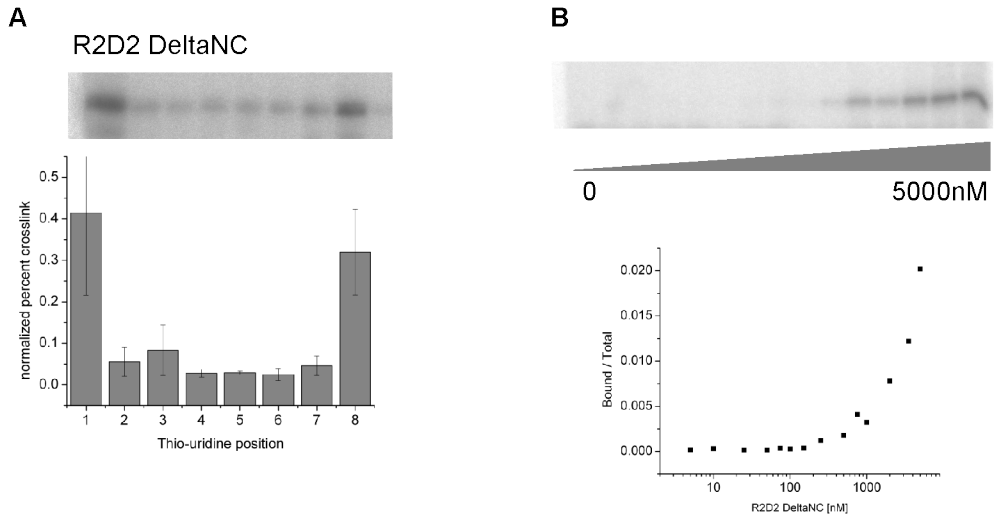

Figure 5.32: R2D2 crosslink pattern
A: Crosslink profile of R2D2 DeltaNC with an exemplary crosslink pattern below. The crosslink profile was obtained as described. B: R2D2 DeltaNC was titrated to crosslink oligo 1, incubated for 30min at RT and crosslinked. The resulting crosslink bands were analyzed with MultiGauge software, the quantification is shown below. No plateau is reached using this method, indicating either a K_D above the range covered in the protein titration or unspecific protein-RNA absorption due to misfolded protein.

5.9 siRNA duplex asymmetry sensing by the Dcr-2/LoqsPD complex

The Dcr-2/R2D2 complex has been shown to be the predominant RISC loading complex in Drosophila, which is able to sense the thermodynamic asymmetry of the siRNA duplex and thereby to discriminate guide and passenger strand[85, 104]. Since the Dcr-2/LoqsPD complex can also function as RLC at least for a subset of endo-siRNAs (Mirkovic-Hösle, personal communication), we wanted to see whether LoqsPD can also induce asymmetric binding of Dcr-2. The siRNA duplex used in the crosslink experiments is asymmetric, with a $\Delta\Delta G$ = -1.07 kcal/mol (see figure 5.34A), and can therefore be used to test the Dcr-2/LoqsPD complex for this ability.

5.9.1 LoqsPD alone does not show preferential binding to one end of the siRNA duplex

TRBP recognizes asymmetry on its own and binds preferentially to the more stably hybridized end of the siRNA duplex [34]. To test if LoqsPD also has this ability, I titrated full length LoqsPD to oligos 1 and 8 and analyzed the crosslink products on a denaturing gel (figure 5.33A), which yielded the same K_D for both duplex ends (figure 5.33B). Therefore, LoqsPD alone is not able to identify the duplex end with the smaller ΔG.

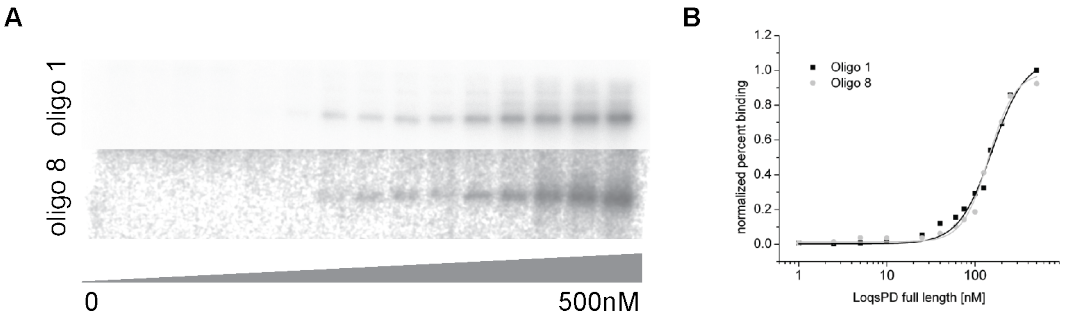

Figure 5.33: LoqsPD alone does not sense thermodynamic asymmetry
A: LoqsPD full length was titrated to crosslink oligos 1 and 8, incubated for 30min at RT, and crosslinked. The crosslinked bands were analyzed with MultiGauge software. B: Quantification of both crosslink profiles. Values constitute the mean of two independent experiments. LoqsPD full length crosslinked to the same amount to crosslink oligo 1 and 8 over the concentration range. The fitted parameters were a K_D of 155nM and a Hill coefficient of 2.3 for crosslink oligo 1, and a K_D of 143nM and a Hill coefficient of 2.9 for crosslink oligo 8.

5.9.2 LoqsPD and Dcr-2 can sense the thermodynamic asymmetry of the RNA duplex

Like the dsRBPs, Dcr-2 crosslinks preferentially to the duplex ends, covering also the neighbor of nt 3 on the opposing 3' strand, probably due to its greater size (figure 5.34B). Alone, Dcr-2 does also not show any preference for one duplex end. However, when crosslinked together with an excess of full length LoqsPD, Dcr-2 showed preferential binding to the thermodynamic less stable end of the siRNA duplex (figure 5.34C). Due to the excess of LoqsPD, its binding behavior in complex with Dcr-2 is superposed by binding of LoqsPD alone. Therefore, preferential binding of LoqsPD is not visible in this experimental setup. Nevertheless, these results show that the Dcr-2/LoqsPD complex has asymmetry-sensing properties, which corroborates the idea that it can function as RLC in addition to Dcr-2/R2D2.

5.9.3 The Dcr-2 - LoqsPD interaction is necessary for asymmetry sensing

LoqsPD interacts with Dcr-2 via its specific C-terminus. To see whether a protein-protein interaction between Dcr-2 and LoqsPD is necessary for their asymmetry-sensing ability, I performed crosslink experiments with Dcr-2 together with an excess of Loqs DeltaNC, which should have reduced ability to bind Dcr-2 due to the missing C-terminus. Loqs DeltaNC could not induce asymmetric Dcr-2 binding (figure 5.34D), which implies that the Dcr-2 - LoqsPD interaction is necessary for that.

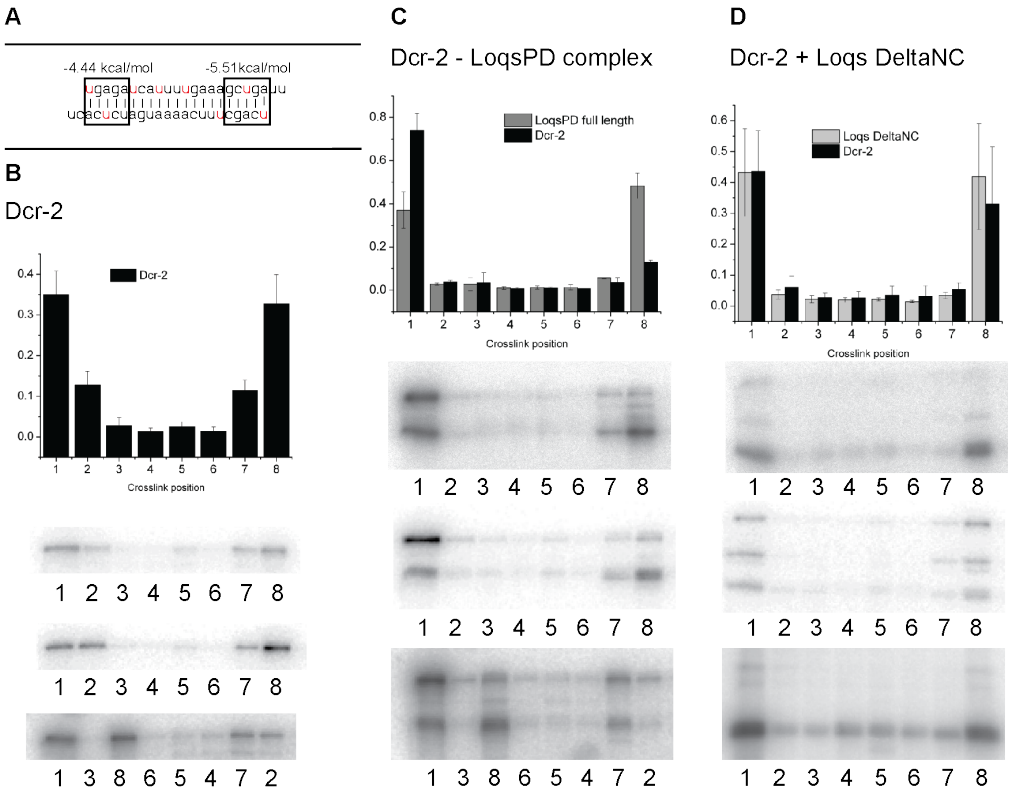

Figure 5.34: LoqsPD and Dcr-2 can sense the thermodynamic asymmetry of the RNA duplex A: Thermodynamic properties of the siRNA oligo used for crosslink experiments. Red uridines are substituted with thiouridines in the respective crosslink oligos. The thermodynamic stabilities of both ends add up to a $\Delta\Delta G$ of 1.07 kcal/mol. B: Crosslink profile of Dcr-2 with three example crosslink patterns below. The crosslink profile was obtained as described. Dcr-2 crosslinks preferentially to the extremities of the siRNA duplex. C: Dcr-2 together with an excess of LoqsPD full length was incubated with the crosslink oligos, and the crosslink profile was obtained as described. All crosslink patterns are shown below. Together with LoqsPD, Dcr-2 binds preferentially to the less stable end of the duplex. D: Dcr-2 together with an excess of Loqs DeltaNC was incubated with the crosslink oligos. When the interaction between Dcr-2 and LoqsPD is weakened, Dcr-2 looses preferential binding to the less stable duplex end.

5.9.4 RNA deformation is not the mediator for asymmetric binding of the Dcr-2/LoqsPD complex

How this asymmetric binding of Dcr-2 and the dsRBP occurs on the mechanistic level is unknown. Since not the absolute but the relative thermodynamic stability of each end is read out and neither Dicer nor the dsRBP knows about the stability of the opposing duplex end, this is no trivial problem. dsRNA binding by the RLC should be envisioned as a dynamic process, since immediate binding in the correct orientation should only happen by chance in 50% of all cases. One possible way might be communication between the duplex ends via the RNA: Dcr-2 might twist the RNA duplex upon binding, which would be easier from the thermodynamically less stable end, and thereby facilitate binding of the dsRBP on the other end. To test for this possibility, I checked for impairment of asymmetry sensing when communication between the duplex ends is weakened via introduction of a nick in the RNA duplex. However, no change in efficiency of asymmetry sensing could be observed in crosslink experiments with nicked RNA duplexes (figure 5.35), therefore some other mechanism has to underlay this phenomenon.

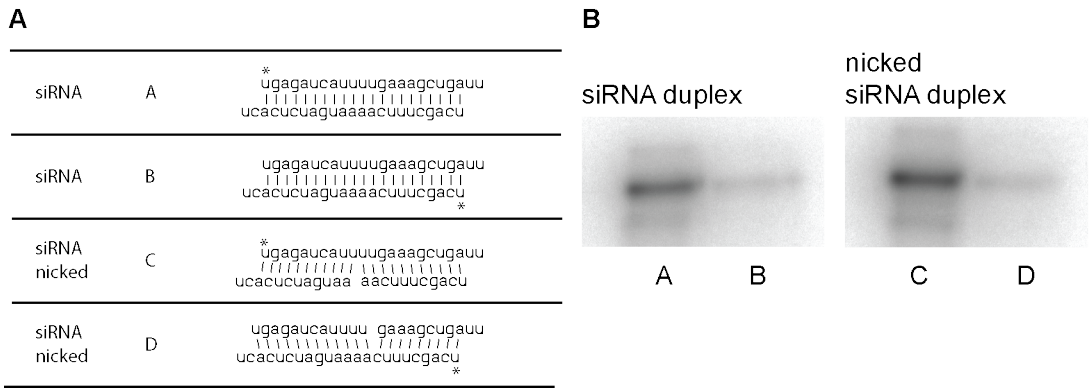

Figure 5.35: Role of RNA deformation in asymmetry sensing
A: Oligos used for crosslink experiments. The siRNA oligos A and B correspond to crosslink oligos 1 and 8, resp. (see figure 5.29). For the nicked siRNA oligo, the thiouridine containing strand was annealed with a partner strand that was split in two. B: Crosslink Pattern of Dcr-2 after incubation with LoqsPD and the respective thiouridine oligos. Asymmetric binding is maintained also with the nicked siRNA duplex.

6 Discussion

6.1 LoqsPD and R2D2 in their cellular environment

The function of LoqsPD and R2D2 is influenced by several factors. Their concentration, their affinity for their target RNA, and target RNA concentration are related by the law of mass action. This relationship allows conclusions about the specifications of dsRBP activity. Their localization might favor encounter with specific RNA species and effector proteins, and has an influence on their local concentration.

6.1.1 Concentration

We probed the concentration of endogenous LoqsPD and R2D2 in S2 cell lysates using antibodies calibrated with the help of recombinant protein. R2D2 occurs in the cell at a concentration of ~ 100nM, which is comparable to the concentration range determined in previous experiments [104]. Since R2D2 is unstable without Dcr-2, this also indicates a lower limit for the Dcr-2 concentration. The determined concentration agree well with kinetic data available on RNA processing by Dcr-2/R2D2 [17], since it is clearly above the K_M of the complex. The concentrations obtained here likely underestimate the actual protein concentration in the cell, since probably the extraction efficiency was not 100%. LoqsPD concentration seems to be lower than that of R2D2, approximately by a factor of 10, giving 10nM as a rough estimate of LoqsPD concentration.

Protein expression levels depend on the cell type, and the R2D2 - LoqsPD ratio determined here may not be valid in every case. S2 cells originate from *Drosophila* embryos and presumably derive from a macrophage-like lineage [94]. Defense against invading viruses by RNAi primarily depends on R2D2. Macrophages as vanguards of the immune system might be equipped with a particularly strong defense against viruses that escape the usual phagocytic pathway and therefore contain more R2D2 than other cell types.

6.1.2 Localization

R2D2 resides in D2-bodies together with Dcr-2 and Ago2, which prevents misdirection of endo-siRNAs into Ago1 [76]. So far, LoqsPD has not been described as localized to cytoplasmic bodies [70]. We could show that both myc-tagged LoqsPB and LoqsPD localize to speckles

in the cytoplasm (see figure 5.1), even more so than R2D2, whose localization to distinct foci we could not confirm to the extent seen by Nishida et al [76]. This punctate localization pattern of LoqsPD and LoqsPB has also been seen in independent experiments by other members in our lab. Since no co-staining of Dcr-2 has been made, we can only speculate that these foci are D2-bodies. D2-bodies disassemble when R2D2 is knocked down in S2 cells, whereas LoqsPD knockdown had no effect, leading to the assumption that R2D2 is the core component of this bodies [76]. This might be simply due to the lower LoqsPD concentration in S2 cells, since upon removal of LoqsPD there would be enough R2D2 left to maintain localization of the majority of Dcr-2 in the D2 bodies. Concentration of LoqsPD in such foci would be reasonable, since its K_D for siRNA determined in this study requires a higher LoqsPD concentration for RNA binding than what can be deduced from quantification experiments in S2 cells. This way, LoqsPD could bind dsRNA without being in a complex with Dcr-2, with the complex presumably having a higher affinity for dsRNA than either protein alone [17], and recruit dsRNA to Dcr-2. Depletion of R2D2 leads to misdirection of endo-siRNAs into Ago1, since enrichment of RLC and Ago2 in the D2-bodies does not occur any more. However, there is still endo-siRNA loading in Ago2 upon R2D2 depletion, which indicates LoqsPD / Dcr-2 RLC activity in residual D2-bodies. Taken together, LoqsPD might play a similar role in D2 body formation as R2D2, which might have been missed due to lower expression of LoqsPD than R2D2 in S2 cells.

Immunostaining of TRBP did not show localization to cytoplasmic foci in a study by Daniels et al. [23]. In the human system, local enrichment of RLC components in subcellular foci might not be necessary since no sorting of small RNAs in specific Argonautes has to occur.

6.1.3 Competition for complex formation with Dcr-2

R2D2 is unstable without Dcr-2, therefore it will not act as single protein but only in complex with Dcr-2. LoqsPD also interacts with Dcr-2 to perform its function, even though it is stable on its own and may bind dsRNA alone to recruit it to Dcr-2. In S2 cells, R2D2 and LoqsPD compete for Dcr-2 binding: Reporters for their activity show increased function of one of the dsRBPs when the other is eliminated and can not titrate away Dcr-2 anymore [38]. Over-expressed LoqsPD not only co-immunoprecipitated Dcr-2 but also R2D2, which led to the assumption of a ternary complex comprising Dcr-2 and both dsRBPs [70]. Both LoqsPD and R2D2 interact with Dcr-2's helicase domain, a 62kD subdomain of the Dcr-2 enzyme [38]. I showed that R2D2 presumably interacts with the linker between the two helicase subdomains. Binding of the delta helicase construct in figure 5.3 may represent unspecific binding or background signal. Unfortunately a corresponding control for this was not included, therefore the results should be interpreted with caution. Nevertheless, they are supported by the work of Nishida et al., who interpret Dcr-2 and R2D2 interaction via localization of Dcr-2 to D2-bodies

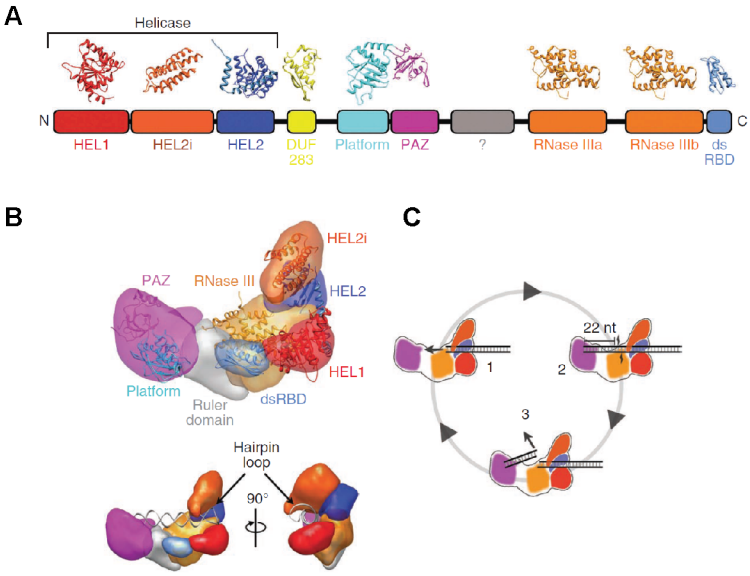

Figure 6.1: Structure and orientation of the Helicase domain in the context of full length human Dicer
A: Domain structure of human Dicer. B: Upper: 3D architecture of human Dicer, including the x-ray structure of the individual subdomains. Lower: Model with bound pre-miRNA hairpin. C: Model for processive Dicer cleavage of dsRNA. Adapted from [51].

and see that only the complete Helicase domain of Dcr-2, but not the N-terminal part alone, is localized there [76]. The linker adopts a domain-like fold and also constitutes the TRBP interaction site in human Dicer [23]. If the linker between the helicase and the ATPase domain of Dicer as interaction site for dsRBPs is a common theme and LoqsPD also interacts there, would there be enough space to accommodate both dsRBPs at the same time? With its 24.7kD, the linker is smaller compared to LoqsPD (38kD) or R2D2 (35kD). The helicase domain linker is located at the tip of the protein, therefore it is accessible from several sides (figure 6.1). The 3D model with a bound RNA substrate does at least not exclude the possibility of two dsRBPs contacting the helicase domain [51].

6.2 Protein purification

dsRBP purification in general was rather variable in yield and stability of the proteins. Removal of a solubility-enhancing tag like GST or a ZZ-tag often led to aggregation of full length proteins. It was important that no RNA remained in the protein preparation, since that would have influenced the RNA binding experiments. High salt treatment with 800mM NaCl or even 1M LiCl was necessary to get rid of the RNA, indicating a strong interaction (with a dissociation constant of 50nM in the case of full length LoqsPD). Bound RNA might contribute to dsRBP stability, and its removal might therefore enhance their aggregation.

Instability of the dsRBPs culminated in R2D2, which could not be purified as a functional RNA binding protein at all. A MBP-fusion of the protein could be purified but did not show RNA binding in the physiological concentration range. This has also been observed by Liu et al [59], who deduce that R2D2 is not able to bind dsRNA on its own. Given that a protein containing two instable yet functional dsRBDs (see section 5.4.4) should in theory be able to bind RNA, it can be assumed that also in this context R2D2 is instable and only kept soluble via the MBP-tag.

In the Dcr-2 preparations from *Spodoptera frugiperda*, a 70kD protein copurified which was identified as a heat shock protein via mass spectrometry (see Appendix). Hsc70 and Hsp90 are involved in RISC loading in Drosophila [40, 72]. They interact with Ago1 and Ago2 and are specifically required for loading, but not for unwinding of the small RNA duplex. In addition to Ago, the chaperones might also interact with Dicer, which could be mediated via a conserved motif that is strong enough to tether a related Hsp70 to *Drosophila* Dcr-2 during the purification process.

6.3 RNA binding properties of LoqsPD

In addition to binding and crosslinking experiments, NMR analyses of the LoqsPD dsRBDs and the DeltaNC constructs were performed by Dr. Thomas Kern and Jan-Niklas Tants from the group of Prof. Dr. Michael Sattler. They focused on RNA binding of the two domains and their orientation towards each other in the RNA bound state. Whenever NMR data is mentioned in the course of this discussion, it refers to their work.

6.3.1 Sequence conservation

The amino acid sequence of both Loqs dsRBDs aligns well with those of other dsRBDs (figure 6.2). Residues implicated in RNA binding are conserved throughout the entire domain, classifying them as type A dsRBDs. An unusual 4AA insertion between ß2 and ß3 is present in dsRBD1. The second dsRBD of DCL1, the Dicer homolog in plants, and of DHX9, an RNA helicase, contain a similar extension. It is located on the opposite side of the expected RNA-binding interface and is not involved in RNA binding, but might provide additional stabilization for the dsRBD fold [13, 30].

6.3.2 Binding affinities and interplay of the two Loqs dsRBDs

The two dsRBDs of LoqsPD bind RNA with similar affinity (228nM and 278nM, respectively). The affinity of the two domains is increased ~2-fold when the two domains are fused together via their endogenous linker and 4-fold in the context of the full length protein. The influence of

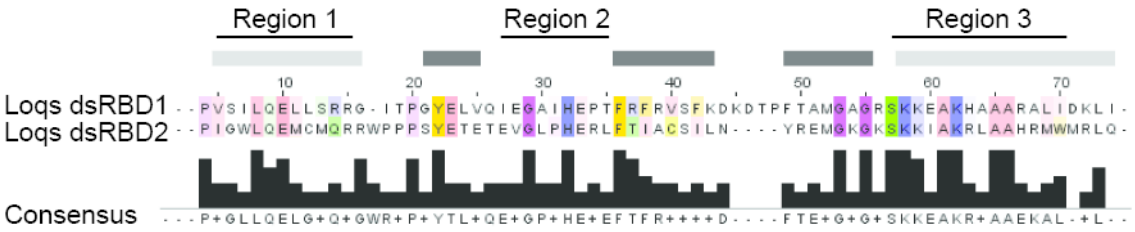

Figure 6.2: Loqs dsRBD conservation -
Alignments of the Loqs dsRBDs with the dsRBD consensus sequence from figure 2.3. RNA binding regions are labeled and the secondary structure is indicated by the grey bars (light grey: α-Helices, dark grey: ß-Sheets).

combining two dsRBDs on their affinity has been studied for several dsRBPs, including TRBP (see table 6.1).

6.3.2.1 Both Loqs dsRBDs contribute equally to RNA binding

The two dsRBDs of both PKR and HYL1 differ significantly in their RNA binding affinities, which has functional implications. The first dsRBD of PKR has a significantly higher affinity for dsRNA than the second [102]. It fluctuates with motions on the milli to microsecond timescale, whereas the second dsRBD interacts with the kinase domain and thereby keeps the PKR inactive. When the first dsRBD binds dsRNA, it induces cooperative RNA binding of the second dsRBD, which now exposes the kinase domain and activates the PKR [75]. Due to the cooperative binding of the second dsRBD, the affinity of the double domain construct exceeds that of the first dsRBD by a factor of 10-15.

The second dsRBD of HYL1 is a non-canonical dsRBD with structural deviations from dsRBD1 at the RNA binding interface, and consequently does not bind to RNA substrates under physiological conditions [117]. The affinity of the double domain construct has been reported to be stronger or equal to the affinity of the isolated first dsRBD [117, 88], but RNA binding of HYL1 seems to only be mediated via its first dsRBD. dsRBD2 harbors a putative dimerization interface, and HYL1 is implicated to bind dsRNA as a functional dimer, therefore the second dsRBD seems to function in mediating protein-protein contacts.

For the Loqs dsRBDs, no such division of labor seems to exist, since the difference in binding affinity of the two dsRBDs is only marginal. Their affinities approximately add up to the affinity of the double domain construct, indicating that both dsRBDs are involved in RNA binding of Loqs to similar extents. The NMR structures show a slightly tighter binding of dsRBD1 to the RNA, which is reflected in a marginally decreased K_D of dsRBD1 for siRNA obtained from both direct and competition experiments. LoqsPD misses the third dsRBD-like fold with which LoqsPB and R2D2 bind to their Dicer partners [119, 76]. The C-terminus of LoqsPD

confers its specificity for Dcr-2 over Dcr-1 and can promote Dcr-2 binding on its own, even though other residues of the proteins are involved to reach full Dcr-2 binding [38]. An N-terminal truncation starting directly after the first dsRBD binds Dcr-2 as strong as the full length protein, suggesting the linker and the second dsRBD as additional anchors in Dcr-2 binding [70]. Therefore, in addition to RNA binding the second dsRBD might also be involved in protein-protein interactions with Dcr-2.

6.3.2.2 Commitment of LoqsPD to the siRNA pathway

In TRBP, the second dsRBD has in general a higher affinity for RNA than the first, ranging from 2 to 4-fold, depending on the RNA substrate and the study. TRBP has a significantly higher affinity for completely base-paired dsRNA compared to mismatched dsRNA substrates even on the level of the individual dsRBDs [116, 5]. It has to be kept in mind that the dissociation constants for siRNA and miRNA binding are the results from two independent studies, even though both were obtained with ITC measurements. In the case of LoqsPD, neither the individual dsRBDs nor their fusion protein show preferential binding of siRNA over miRNA. Only full length LoqsPD has a higher propensity for binding siRNA substrates. This implies involvement of residues outside the dsRBDs which modulate the orientation of the dsRBDs in a way that enhances preference for completely base-paired RNA. These residues are likely to be located in the C-terminus specific to the PD isoform of Loquacious, since the PB isoform, which contains the same dsRBDs, is involved in miRNA processing and therefore should not prefer siRNAs over mismatched RNA substrates. Anyway, the affinity of LoqsPD for siRNA over miRNA substrates differs only by a factor of 2, which is far from excluding mismatched RNA binding by LoqsPD. Since endo-siRNAs arising from transcripts that are folded back on themselves also contain mismatches, LoqsPD binding of imperfectly base paired RNA substrates is also necessary. Apart from the slight preference of LoqsPD for siRNAs, LoqsPD is committed to the siRNA pathway via its specific interaction with Dcr-2.

6.3.3 Cooperativity of dsRBD binding

6.3.3.1 Cooperativity of the two dsRBDs: Single vs tandem dsRBD binding

When two dsRBDs are tethered together by a linker of limited length, they should influence each others binding. Upon binding of the first dsRBD the probability that the second dsRBD also binds is enhanced: Its movement is not arbitrary anymore but confined to a limited region around the RNA defined by the linker length, which increases the effective dsRBD concentration around the RNA. DHX9, or RNA helicase A, promotes RISC assembly in humans and interacts with the RISC components via its two dsRBDs in an RNA dependent manner [30]. Its two dsRBDs bind RNA in the µM range, and even though they are connected by a rather long

Protein	dsRBD1	dsRBD2	dsRBD1+2	Full length	Substrate	Linker [AA]	Method	cit.
Loqs	227nM	212nM	89nM	50nM	siRNA	48	Anisotropy	this study
TRBP	15.4µM	3.7µM	3.4µM 0.5µM + 6.2µM°		pre-miRNA	60	ITC	[5]
TRBP			1µM + 5.7µM°		miRNA duplex	60	ITC	[5]
TRBP	220nM	113nM	0.25nM + 121nM°	0.24nM + 13.3nM°	siRNA	62*	ITC	[116]
TRBP			0.51nM		siRNA	62*	EMSA	[83]
TRBP			0.51nM		dsRNA 650bp	62*	EMSA	[83]
HYL1	low binding affinity	very low binding	40x tighter than dsRBD1		miRNA duplex, pre-miRNA	17*	EMSA	[117]
HYL1	740nM	>10µM	similar to first domain		pre-miRNA	17*	EMSA	[88]
PKR	50nM	200nM	5nM		poly(rI.rC)	33*	filter binding assay	[102]
PKR	380nM	no binding			dsRNA	33*	EMSA	[102]
PKR	1050nM	~80µM	70 nM		HIV TAR RNA		ITC	[45]
DHX9	5.9µM	13.6µM	106nM + 4.4µM °		siRNA	109*	ITC	[30]

Table 6.1: Summary of K_D values for dsRBD - dsRNA interaction

Overview over the dissociation constants of dsRBDs obtained in other studies.
*: derived from http://www.uniprot.org/uniprot/P19525; °: fitted with two-site binding model

linker (~100AA), the affinity of the double domain construct is ~50 fold higher. The first two dsRBDs of TRBP bind siRNA substrates cooperatively, which results in the K_D of the double domain construct being one order of magnitude smaller than that of the individual dsRBDs [116]. For miRNA substrates there is also a ~10-fold increase in affinity from the single to the double domain construct [5]. The 60AA linker connecting the two TRBP dsRBDs is highly flexible, and no interdomain contacts are formed between the two dsRBDs [5]. Therefore, even though linker length is one factor that should impact the degree to which two dsRBDs affect each others binding [95], there are additional determinants that influence cooperation of two dsRBDs beyond the influence of the linker length or interdomain contacts.

In LoqsPD, the affinities of the two dsRBDs approximately add up to the affinity of the double domain construct, indicating that one dsRBD does not influence binding of the other. This is corroborated by results obtained from NMR experiments which show no interaction of the two domains both in the RNA bound and unbound state. Apparently, in this case an unstructured linker of 46AA is sufficient to render the binding of both domains completely independent. Since the binding curves are fitted with a binding model that assumes the same affinity for all binding events, the obtained K_D values represent the macroscopic binding constant, which includes a possibly weaker binding of a second protein to the RNA. This might conceal a higher affinity of the first binding event.

6.3.3.2 Cooperative RNA binding of LoqsPD: The Hill coefficient

A Hill coefficient n > 1 implies that once one protein is bound to the RNA, its affinity for another increases. Hill coefficients have been observed in other studies of dsRBD containing proteins, like DGCR8, Dicer, and RDE-4 [83, 115, 114, 97]. Whereas Wostenberg et al. do not interpret the Hill coefficient of 2 - 4 in their Dicer dsRBD binding data, they accept cooperative binding of more than one DGCR8 dsRBD1 on one pri-miRNA substrate (n = 2.3), which they attribute to an intrinsic capability of the domain to assemble [115, 114]. Sohn et al. also argue for cooperative binding of the DGCR8 core to pri-miRNA (n = 2.1): They hypothesize that binding of one dsRBD of DGCR8 to a pri-miRNA induces a conformational change in the stem region that allows the other dsRBD to interact with the second binding site on the pri-miRNA more efficiently [97]. RDE-4 binding to dsRNA also resulted in a Hill coefficient > 2, depending on the length of the RNA. This cooperativity derives either from protein-protein interactions or from production of highly favorable binding sites adjacent to bound proteins due to slight perturbations of the dsRNA structure [83]. For TRBP, no cooperative binding could be observed in this study. The two ways to obtain cooperative binding of proteins on dsRNA are therefore protein-protein interactions or deformation of the RNA. DNA deformation has recently been shown to mediate allostery [46], making RNA deformation more than a hypothetical source for cooperativity.

Almost all binding curves of dsRBPs obtained in this work required a Hill coefficient for accurate fitting, independent of the technique used for the measurements. Hill coefficients derived from Anisotropy experiments are summarized in table 5.3. The only exception is LoqsPD full length binding to siRNA, with n = 1.3. In general, LoqsPD full length binding resulted in the lowest Hill coefficients for all substrates, which might indicate that the RNA can accommodate less of the full length protein due to its bigger size. Comparing the EMSA shifts of Loqs DeltaNC and LoqsPD full length supports this idea, as does the amplitude of the Anisotropy signal: Even though LoqsPD full length is the biggest protein which should slow RNA tumbling down the most, it reaches the lowest Anisotropy signal.

If the Hill coefficient stems from cooperative binding mediated by RNA deformation, a nick in the RNA duplex should decrease it. Nicked dsRNA cannot propagate deformation as well as an intact duplex, even though base stacking interactions could keep the bases on both sides of the nick together to some extent. Since the Hill coefficient did not decrease in binding experiments with a nicked duplex (see section 5.7.4), cooperativity probably arises mostly through protein-protein interactions.

6.3.4 ssRNA binding

ssRNA binding by dsRBDs is rather unusual, especially with affinities comparable to dsRNA substrates. Both ssRNAs, bantam and miR8 derived, were bound, excluding features inherent to one specific sequence as reasons for binding. Fold predictions of the RNAs show double-stranded portions of 4 - 5nt in length which are unlikely to provide a regular binding site for the dsRBDs (figure 6.3, upper panel). A large error in ssRNA binding curves and the fact that no binding could be observed in EMSA experiments suggests very transient binding, which might be used to probe for substrate authenticity. It would be interesting to see whether ssRNA is able to compete siRNA away from the dsRBDs in a competition experiment. Self-dimerization of the ssRNA oligo is also possible to some extent, resulting in duplexes with a high amount of mismatched base-pairs and unusual lengths of the 3' overhang (figure 6.3, lower panel). Since only LoqsPD full length distinguishes between completely base-paired and mismatched substrates this might be a reason why only LoqsPD has a reduced affinity for ssRNA compared to siRNA. Low affinity ssRNA binding has been reported for TRBP, with the suggestion that it may provide stabilization of ssRNA-containing RLCs prior to Ago2 loading [34]. In Drosophila siRNA unwinding can also be initiated by the RLC, which might bring LoqsPD in contact with ssRNA species in vivo.

bantam derived ssRNA miR8 derived ssRNA

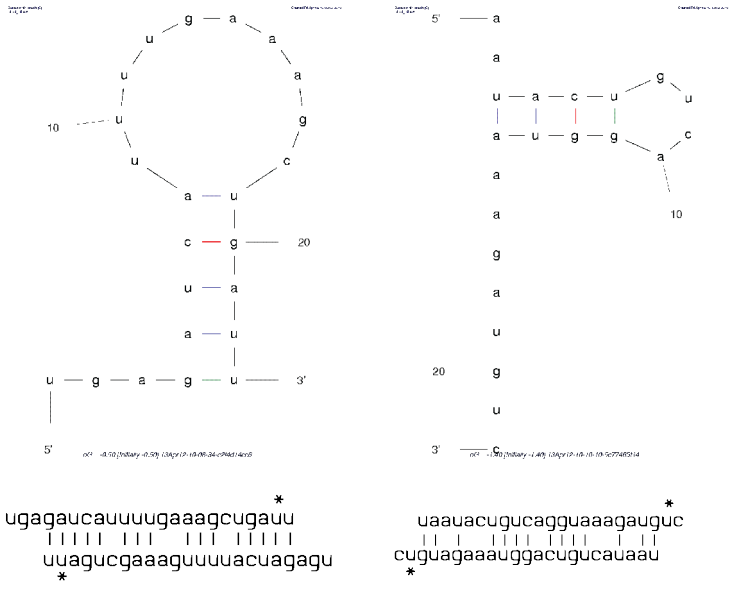

Figure 6.3: Predicted fold of ssRNAs
Upper panel: Predictions of bantam and miR8-derived ssRNA folds calculated with the M-fold software. Lower panel: Possible conformation upon self-annealing of the ssRNA oligo. GU wobble base pairs are depicted as normal base pairs.

6.4 Protein - RNA Binding geometry

All protein constructs were found to crosslink at the extremities of the RNA duplex, and artifacts introduced by crosslinking were ruled out as far as possible (see section 5.3.7). Nevertheless, a reduced crosslink efficiency in the double-stranded region of the RNA duplex can not be excluded completely.

6.4.1 dsRBD sliding

Broadened resonances at the protein-RNA interface in NMR measurements suggest movement of the dsRBDs along the RNA. Sliding of dsRBPs on RNA has been observed before: TRBP can move along dsRNA independent of ATP, and for this diffusion both of its dsRBDs are necessary [47]. *Drosophila* ADAR dsRBD, which does not show a sequence specificity as high as human ADAR2 dsRBDs, might also slide back and forth on the RNA [3]. The aspect of dsRBD sliding on the RNA is compatible with the results from crosslink experiments, since they provide an average of the residence time at each point of the duplex. FRET traces in the studies of Koh et al show that the protein resides longer at the ends of the RNA than it takes it to travel along the helix to the other end. In addition, sliding proteins might be captured in the

residual crosslinking in the middle of the RNA duplex. Koh et al only see sliding on duplexes with canonical Watson-crick base pairs, and not on dsRNAs containing mismatches or loops [47]. The crosslink pattern on miRNA duplexes shows protein binding also in the middle of the RNA, most prominently for the individual dsRBDs and Loqs DeltaNC. This might be due to a stalling of dsRBD movement along the helix due to the mismatches, which would increase residence time in the middle and thereby the probability of being crosslinked there.

6.4.2 Asymmetry sensing0,

Guide strand selection is of major importance for target recognition of the RISC. This decision is made by the RISC loading complex (RLC), consisting of Dicer and a dsRBP, by binding asymmetrically to the siRNA duplex, thereby licensing one strand as the guide. During siRNA production, a Dicer protein processes the long double stranded precursor with the help of a dsRBP by repeatedly cleaving off 21nt duplexes from one end. This leads to a directionality of protein binding on the RNA, but this is not the source for guide strand selection [86]. Rather, the cleaved siRNA duplex is released into the bulk solvent and rebound by the Dicer/dsRBP complex in an orientation based on the thermodynamic characteristics of its two ends: the strand whose 5' end has the lower thermodynamic stability is bound by Dicer and retained in the RISC as the guide strand [78]. In *Drosophila*, the Dcr-2/R2D2 complex has been shown to be the sensor for siRNA asymmetry [104], and this study shows that also the Dcr-2/LoqsPD complex can function in guide strand determination (see section 5.9.2). In humans both TRBP and PACT together with Dicer are capable of sensing the thermodynamic asymmetry of siRNA ends [78]. TRBP has been shown to recognize siRNA asymmetry on its own [34], whereas LoqsPD is not able to do this (see section 5.9.1). Recent studies have even shown that human Ago2 alone is able to select the guide strand without the help of Dicer [77, 7].

In *Drosophila*, only the complex of Dcr-2 and the dsRBP is able to sense the asymmetry, and for the Dcr-2/LoqsPD complex interaction of the two proteins is of importance (see section 5.9.3). Cryo-EM reconstructions of the human RISC loading complex show that the RNA duplex can be bound on the less stable end by the Dicer PAZ domain and on the more stable end by the Ago PAZ domain during siRNA transfer [93]. After anchoring the 3' end of the guide strand in the Ago PAZ domain, Dicer can hand over its 5' end to the Ago mid domain, and passenger strand cleavage can occur. When looking at the EM reconstructions, the Dicer helicase domain points approximately towards the Ago PAZ domain. Assuming the RNA to position along the helicase domain and being bound by the Dicer PAZ domain on one and the dsRBP on the other end in the Dicer/dsRBP complex, this would put the dsRBP bound to the 3' end of the guide strand in close proximity to the Ago PAZ domain, enabling RNA hand-over between the two proteins.

How is asymmetric binding of Dcr-2 and the dsRBP achieved? We could show that RNA

deformation is presumably not the way asymmetric binding is mediated. However, a nicked RNA duplex might still be able to partially transmit deformation across the nick due to base stacking interactions. Another approach to restrict RNA deformation might be via backbone modifications that render the RNA more rigid. In Locked nucleic acids, the ribose is locked in the 3'-endo conformation by an extra bridge between the 2'O and the 4'C, thereby increasing base-stacking interactions and the melting temperature. The influence of such a modification might be more pronounced than that of a nicked siRNA duplex.

In addition to passenger strand cleavage and subsequent unwinding of the siRNA duplex in the Ago protein, the RLC is also able to initiate duplex unwinding [104]. Dcr-2's PAZ domain could start this unwinding process more effectively when bound to the thermodynamically less stable end. One could imagine that fixation of the other end of the duplex by LoqsPD is a prerequisite for unwinding, since otherwise the RNA might be propelled out of the PAZ domain.

6.5 Comparison with R2D2

6.5.1 Binding behavior of R2D2 constructs

6.5.1.1 Conservation of R2D2 dsRBDs

The R2D2 dsRBDs deviate more from the common consensus sequence than the Loqs dsRBDs (figure 6.4). In the first dsRBD, four amino acids are missing in binding region 2, including a conserved histidine implicated in RNA binding [116, 90]. This shortens the separation between binding region 1 and 2, which has to span the major groove of an A-form helix to allow binding of regions 1 and 2 to the minor grooves. In R2D2 dsRBD2, several highly conserved residues are replaced, including the glutamine in region 1 and a lysine in region 3, which both make contacts with the RNA. Therefore, RNA binding by R2D2 could be limited compared with other dsRBPs. Also, conserved residues involved in maintaining the fold of the dsRBD are missing in the R2D2 dsRBDs. Both lack the stabilizing proline right at the beginning of the dsRBD, and in dsRBD2 three additional highly conserved residues are exchanged against a non-substituting amino acid. Therefore, in particular dsRBD2 might not adopt exactly the canonical dsRBD fold.

6.5.1.2 Stability and binding of R2D2

R2D2 is highly unstable when expressed without Dcr-2, as knockdown of Dcr-2 also results in R2D2 depletion in vivo. This is probably due to exposed hydrophobic residues in the Dcr-2 interacting region which aggregate without its binding partner. Deletion of the Dcr-2 interaction

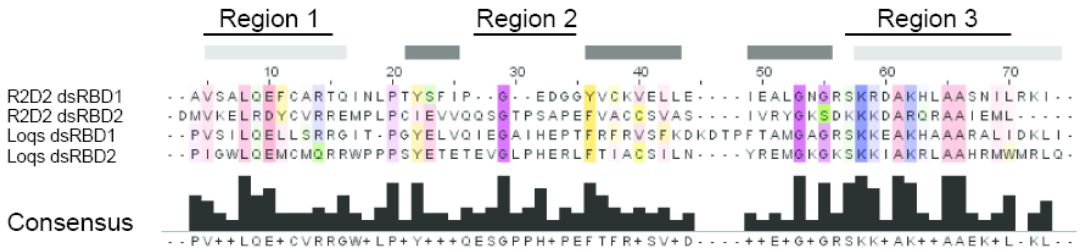

Figure 6.4: Comparison Loqs and R2D2 dsRBDs

site indeed increased the yield of soluble protein. This dependence on Dcr-2 ensures Dcr-2/R2D2 complex formation and prevents the presence of unbound R2D2.

Binding experiments with R2D2 constructs were afflicted with large errors, which might be due to varying amounts of properly folded protein in each preparation. Binding curves of the isolated dsRBDs reached a plateau, arguing for specific RNA binding rather than unspecific aggregation on the RNA (see section 5.4.4). No substrate specificity is obvious, which might only arise in the interplay of the two dsRBDs comparable to LoqsPD or in complex with Dcr-2. The higher K_D values of R2D2 compared to Loqs dsRBDs might either be due to a higher amount of misfolded protein in the preparation or might actually reflect weaker binding of the R2D2 dsRBDs, or constitute a combination of both. The omission of four AAs in dsRBD1 and the incomplete conservation of residues involved in RNA binding and dsRBD fold in dsRBD2 suggest a decrease in RNA binding affinity of the R2D2- compared with the Loqs dsRBDs. The tendency to a decreased RNA binding strength of dsRBD2 compared with dsRBD1 is supported by its greater divergence from the canonical dsRBD sequence, but both R2D2 dsRBDs are equally necessary for binding siRNA together with Dcr-2 [59].

Connecting the two dsRBDs by the endogenous 28AA R2D2 linker results in even weaker or unspecific binding (see section 5.5.5). Could the linker be too short to allow proper arrangement of the dsRBDs on the RNA? DsRBPs with comparable linker lengths between their dsRBDs are PKR (23AA) and Xlrbpa (20AA). In PKR, primarily the first dsRBD mediates RNA binding, whereas the second does not show strong RNA binding properties but it also contacts the RNA to fine-tune the interaction of the first. Xlrbpa is the *Xenopus laevis* homologue of Loqs and TRBP and might therefore make comparable use of its dsRBDs. These examples imply that a short linker between the dsRBDs should not impede their binding. Another explanation for the reduced stability of R2D2 DeltaNC might be that the linker is involved in contacting Dcr-2 in addition to the R2D2 C-terminus. Analysis of the linker sequence however predicts a net charge of -3.0 at pH 7 and good water solubility, comparable to that of LoqsPD (see hydropathy plots in figure 6.5A). In the linker alignment no common motif of TRBP, PACT and Loqs can be found that is not present in R2D2 (figure 6.5B). The amino acid composition of the linker differs between human and fly: TRBP and PACT linkers contain a high amount of

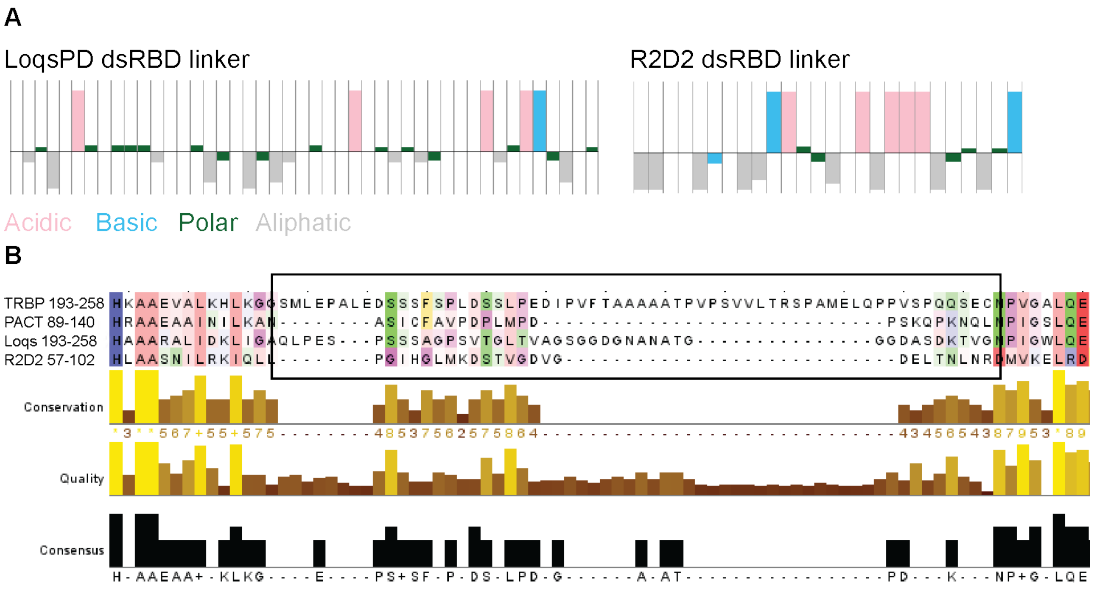

Figure 6.5: dsRBD linker analysis
A: Hydropathy plots of Loqs and R2D2 dsRBD linkers. Hydropathy plots were calculated on http://www.innovagen.se/custom-peptide-synthesis/peptide-property-calculator/peptide-property-calculator.asp B: dsRBD linker alignment of TRBP, PACT, Loqs and R2D2. The linker sequence is marked with the black rectangle.

prolines which cause kinks in the secondary structure and are unable to form hydrogen bonds. On the other hand, the Loqs and R2D2 linkers contain a lot of glycines that promote flexibility of the local protein structure. But again, no reason for an increased propensity of the R2D2 linker to induce aggregation could be found.

6.6 Roles of LoqsPD and R2D2 in small RNA biogenesis

The roles of LoqsPD and R2D2 have been categorized in two different ways. One is the division based on their function, namely assisting Dcr-2 in dsRNA processing or RISC loading. The other is based on the RNA class they are processing, namely endo- or exo-siRNAs.

At the time of their discovery, involvement of the miRNA biogenesis factor Loqs in endo-siRNA biogenesis came as a surprise, and their dependence on Loqs rather than R2D2 was demonstrated repeatedly [37, 80, 22, 123]. In addition, R2D2 instead of LoqsPD was shown to be involved in exo-siRNA biogenesis [38, 109]. From an evolutionary viewpoint, R2D2's primary role might indeed be in viral defense. While Loqs, TRBP and PACT derive from a common ancestor, R2D2 is more distantly related (figure 6.6A) [73]. This is reflected in the intron positions in their genes: the paralogs TRBP and PACT share the same intron structure, whereas Loqs and R2D2 do not (figure 6.6B). This might imply a specialization of Loqs for processing

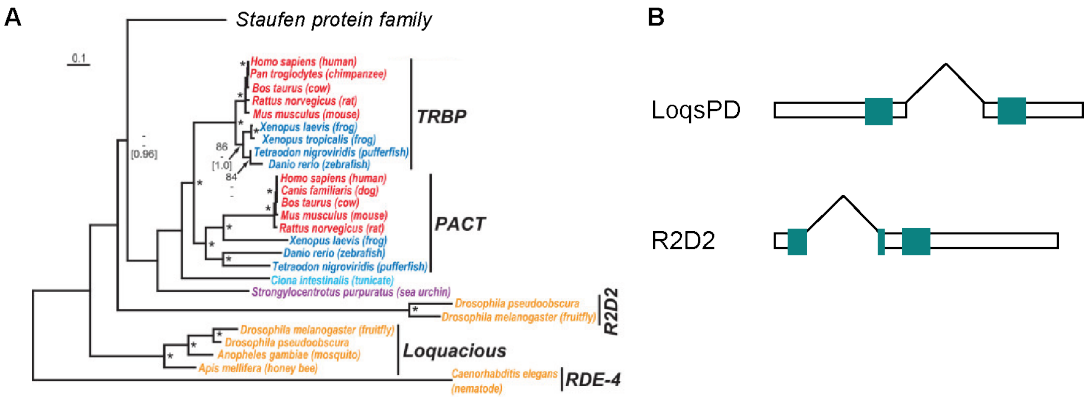

Figure 6.6: Evolution of dsRBPs in small RNA biogenesis
A: Intron structure of LoqsPD and R2D2. Whereas LoqsPD contains a splice site between its dsRBDs, in R2D2 it is located in the middle of dsRBD1. B: Neighbor-joining phylogenetic tree of double-stranded RNA-binding proteins. Scale bar represents 0.1 expected amino acid residue substitutions per site. Adapted from [73].

of endogenous RNAs, including miRNAs, that was maintained during evolution, and of R2D2 for viral defense, that was lost upon acquisition of the interferon system. In addition, fly Ago2, Dcr-2 and R2D2 belong to the top 3% of rapidly evolving genes in Drosophila, potentially enabling them to react on viral evasion strategies [6].

Other studies have postulated that LoqsPD and R2D2 function sequentially in siRNA biogenesis, with LoqsPD assisting Dcr-2 in dsRNA processing and R2D2 in RISC loading [66].

The parameters determining the function of LoqsPD and R2D2 are still not completely understood. Both increase Dcr-2's substrate affinity [17], and both are able to function in the RLC. In general, LoqsPD seems indeed to be primarily involved in dsRNA processing and R2D2 in RISC loading. However, recent findings confirm that the two dsRBPs function at least in part redundantly (Mirkovic-Hösle, personal communication).

This work supports LoqsPD's role in RISC loading. It shows that LoqsPD binds Dcr-2 substrates and products with equal affinity, indicating that it is not only involved in recognizing and processing long dsRNAs like RDE-4 [83]. In contrast to R2D2, it can bind RNA without Dcr-2, which might enable it to recruit dsRNA precursors to Dcr-2. LoqsPD is also able to direct asymmetric Dcr-2 binding to an siRNA duplex, and this asymmetric binding requires protein-protein interaction between the two proteins. Therefore, LoqsPD can substitute for R2D2 in the RLC.

Outlook

I recapitulated in vitro what has been implicated by deep sequencing analyses before, that LoqsPD is able to perform in each step of siRNA biogenesis, including RISC loading. Nevertheless, a parameter for classifying which dsRBP is involved in maturing which RNA is still missing. For TRBP, an impact of RNA sequence on binding has been suggested [34]. There is no correlation between the first nucleotide bias and the dsRBP involved in processing (Mirkovic-Hösle, personal communication), but analysis of small RNA deep sequencing libraries from LoqsPD or R2D2 mutant flies might reveal sequence motifs within the duplex. In addition, individual analysis of different tissue types should be included in the investigation, since transcript levels of endo-siRNA precursors and the protein levels of LoqsPD and R2D2 are presumably variable between different cell types. This way, previously superposed effects like dependence on RNA abundance might be revealed.

Regarding the binding behavior of LoqsPD, it would be interesting to see how a LoqsPD DeltaN construct would behave in the binding and crosslink experiments conducted in this thesis. Is the PD specific C-terminus enough to promote specificity for siRNAs? Is the N-terminus of LoqsPD in any way required for asymmetry sensing together with Dcr-2?

To get a better understanding of the mechanism of asymmetry sensing, the effect of modified RNA oligos like locked nucleic acids and thereby the influence of RNA rigidity should be tested. One might also check how many nucleotides one can extend the siRNA duplex until asymmetry sensing is lost. Since our oligos were already 2nt longer than canonical siRNA duplexes, there does not seem to be a stringent requirement for a defined length. This experiment would have to be conducted under Mg^{2+} deprivation to avoid dicing of the longer dsRNA. The ability of Dcr-2 and LoqsPD to bind asymmetrically might either decrease continuously with increasing RNA length, indicating an involvement of the RNA structure, or it might be lost as soon as the two proteins cannot contact both ends anymore because they are too far apart. It might be interesting to see if Dcr-2 or LoqsPD would be the one to retract from the RNA end to keep the interaction with its binding partner.

To identify protein-protein interactions between Dcr-2 and LoqsPD in addition to the C-terminus of LoqsPD, mass spectrometric analysis of crosslinking sites between the two proteins might be employed. Expanding this method to the Dcr-2/LoqsPD complex bound on RNA might visualize where on the duplex the complex is bound exactly. This knowledge might help to develop new strategies to probe for the mechanism underlying the asymmetry sensing phenomenon.

Bibliography

[1] S. L. Ameres, J. Martinez, and R. Schroeder. Molecular basis for target RNA recognition and cleavage by human RISC. *Cell*, 130(1):101–112, Jul 2007.

[2] P. Barraud, S. Emmerth, Y. Shimada, H. R. Hotz, F. H. Allain, and M. Buhler. An extended dsRBD with a novel zinc-binding motif mediates nuclear retention of fission yeast Dicer. *EMBO J.*, 30(20):4223–4235, Oct 2011.

[3] P. Barraud, B. S. Heale, M. A. O'Connell, and F. H. Allain. Solution structure of the N-terminal dsRBD of Drosophila ADAR and interaction studies with RNA. *Biochimie*, 94(7):1499–1509, Jul 2012.

[4] D. P. Bartel. MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell*, 116(2):281–297, Jan 2004.

[5] M. P. Benoit, L. Imbert, A. Palencia, J. Perard, C. Ebel, J. Boisbouvier, and M. J. Plevin. The RNA-binding region of human TRBP interacts with microRNA precursors through two independent domains. *Nucleic Acids Res.*, Feb 2013.

[6] S. A. Bernhardt, M. P. Simmons, K. E. Olson, B. J. Beaty, C. D. Blair, and W. C. Black. Rapid intraspecific evolution of miRNA and siRNA genes in the mosquito Aedes aegypti. *PLoS ONE*, 7(9):e44198, 2012.

[7] J. G. Betancur and Y. Tomari. Dicer is dispensable for asymmetric RISC loading in mammals. *RNA*, 18(1):24–30, Jan 2012.

[8] P. C. Bevilacqua and T. R. Cech. Minor-groove recognition of double-stranded RNA by the double-stranded RNA-binding domain from the RNA-activated protein kinase PKR. *Biochemistry*, 35(31):9983–9994, Aug 1996.

[9] P. C. Bevilacqua, C. X. George, C. E. Samuel, and T. R. Cech. Binding of the protein kinase PKR to RNAs with secondary structure defects: role of the tandem A-G mismatch and noncontiguous helixes. *Biochemistry*, 37(18):6303–6316, May 1998.

[10] J. E. Braun, E. Huntzinger, M. Fauser, and E. Izaurralde. GW182 proteins directly recruit cytoplasmic deadenylase complexes to miRNA targets. *Mol. Cell*, 44(1):120–133, Oct 2011.

[11] J. Brennecke, A. A. Aravin, A. Stark, M. Dus, M. Kellis, R. Sachidanandam, and G. J. Hannon. Discrete small RNA-generating loci as master regulators of transposon activity in Drosophila. *Cell*, 128(6):1089–1103, Mar 2007.

[12] A. M. Brownawell and I. G. Macara. Exportin-5, a novel karyopherin, mediates nuclear export of double-stranded RNA binding proteins. *J. Cell Biol.*, 156(1):53–64, Jan 2002.

[13] P. Burdisso, I. P. Suarez, N. G. Bologna, J. F. Palatnik, B. Bersch, and R. M. Rasia. Second Double-Stranded RNA Binding Domain of Dicer-like Ribonuclease 1: Structural and Biochemical Characterization. *Biochemistry*, 51(51):10159–10166, Dec 2012.

[14] N. Bushati and S. M. Cohen. microRNA functions. *Annu. Rev. Cell Dev. Biol.*, 23:175–205, 2007.

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

[16] F. E. Campbell, A. G. Cassano, V. E. Anderson, and M. E. Harris. Pre-steady-state and stopped-flow fluorescence analysis of Escherichia coli ribonuclease III: insights into mechanism and conformational changes associated with binding and catalysis. *J. Mol. Biol.*, 317(1):21–40, Mar 2002.

[17] E. S. Cenik, R. Fukunaga, G. Lu, R. Dutcher, Y. Wang, T. M. Tanaka Hall, and P. D. Zamore. Phosphate and R2D2 restrict the substrate specificity of Dicer-2, an ATP-driven ribonuclease. *Mol. Cell*, 42(2):172–184, Apr 2011.

[18] K. Y. Chang and A. Ramos. The double-stranded RNA-binding motif, a versatile macromolecular docking platform. *FEBS J.*, 272(9):2109–2117, May 2005.

[19] M. Chekulaeva, H. Mathys, J. T. Zipprich, J. Attig, M. Colic, R. Parker, and W. Filipowicz. miRNA repression involves GW182-mediated recruitment of CCR4-NOT through conserved W-containing motifs. *Nat. Struct. Mol. Biol.*, 18(11):1218–1226, Nov 2011.

[20] W. J. Chung, K. Okamura, R. Martin, and E. C. Lai. Endogenous RNA interference provides a somatic defense against Drosophila transposons. *Curr. Biol.*, 18(11):795–802, Jun 2008.

[21] F. Civril, M. Bennett, M. Moldt, T. Deimling, G. Witte, S. Schiesser, T. Carell, and K. P. Hopfner. The RIG-I ATPase domain structure reveals insights into ATP-dependent antiviral signalling. *EMBO Rep.*, 12(11):1127–1134, Nov 2011.

[22] B. Czech, C. D. Malone, R. Zhou, A. Stark, C. Schlingeheyde, M. Dus, N. Perrimon, M. Kellis, J. A. Wohlschlegel, R. Sachidanandam, G. J. Hannon, and J. Brennecke. An endogenous small interfering RNA pathway in Drosophila. *Nature*, 453(7196):798–802, Jun 2008.

[23] S. M. Daniels, C. E. Melendez-Pena, R. J. Scarborough, A. Daher, H. S. Christensen, M. El Far, D. F. Purcell, S. Laine, and A. Gatignol. Characterization of the TRBP domain required for dicer interaction and function in RNA interference. *BMC Mol. Biol.*, 10:38, 2009.

[24] D. DeSousa, M. Mukhopadhyay, P. Pelka, X. Zhao, B. K. Dey, V. Robert, A. Pelisson, A. Bucheton, and A. R. Campos. A novel double-stranded RNA-binding protein, disco interacting protein 1 (DIP1), contributes to cell fate decisions during Drosophila development. *J. Biol. Chem.*, 278(39):38040–38050, Sep 2003.

[25] M. Dlaki? DUF283 domain of Dicer proteins has a double-stranded RNA-binding fold. *Bioinformatics*, 22(22):2711–2714, Nov 2006.

[26] S. Duhr and D. Braun. Why molecules move along a temperature gradient. *Proc. Natl. Acad. Sci. U.S.A.*, 103(52):19678–19682, Dec 2006.

[27] D. Fagegaltier, A. L. Bouge, B. Berry, E. Poisot, O. Sismeiro, J. Y. Coppee, L. Theodore, O. Voinnet, and C. Antoniewski. The endogenous siRNA pathway is involved in heterochromatin formation in Drosophila. *Proc. Natl. Acad. Sci. U.S.A.*, 106(50):21258–21263, Dec 2009.

[28] A. Fire, S. Xu, M. K. Montgomery, S. A. Kostas, S. E. Driver, and C. C. Mello. Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. *Nature*, 391(6669):806–811, Feb 1998.

[29] K. Forstemann, M. D. Horwich, L. Wee, Y. Tomari, and P. D. Zamore. Drosophila microRNAs are sorted into functionally distinct argonaute complexes after production by dicer-1. *Cell*, 130(2):287–297, Jul 2007.

[30] Q. Fu and Y. A. Yuan. Structural insights into RISC assembly facilitated by dsRNA-binding domains of human RNA helicase A (DHX9). *Nucleic Acids Res.*, 41(5):3457–3470, Mar 2013.

[31] R. Fukunaga, B. W. Han, J. H. Hung, J. Xu, Z. Weng, and P. D. Zamore. Dicer partner proteins tune the length of mature miRNAs in flies and mammals. *Cell*, 151(3):533–546, Oct 2012.

[32] J. Gan, G. Shaw, J. E. Tropea, D. S. Waugh, D. L. Court, and X. Ji. A stepwise model for double-stranded RNA processing by ribonuclease III. *Mol. Microbiol.*, 67(1):143–154, Jan 2008.

[33] J. Gan, J. E. Tropea, B. P. Austin, D. L. Court, D. S. Waugh, and X. Ji. Structural insight into the mechanism of double-stranded RNA processing by ribonuclease III. *Cell*, 124(2):355–366, Jan 2006.

[34] J. A. Gredell, M. J. Dittmer, M. Wu, C. Chan, and S. P. Walton. Recognition of siRNA asymmetry by TAR RNA binding protein. *Biochemistry*, 49(14):3148–3155, Apr 2010.

[35] A. M. Gurtan, V. Lu, A. Bhutkar, and P. A. Sharp. In vivo structure-function analysis of human Dicer reveals directional processing of precursor miRNAs. *RNA*, 18(6):1116–1122, Jun 2012.

[36] Y. H. Han, Y. J. Luo, Q. Wu, J. Jovel, X. H. Wang, R. Aliyari, C. Han, W. X. Li, and S. W. Ding. RNA-based immunity terminates viral infection in adult Drosophila in the absence

of viral suppression of RNA interference: characterization of viral small interfering RNA populations in wild-type and mutant flies. *J. Virol.*, 85(24):13153–13163, Dec 2011.

[37] J. V. Hartig, S. Esslinger, R. Bottcher, K. Saito, and K. Forstemann. Endo-siRNAs depend on a new isoform of loquacious and target artificially introduced, high-copy sequences. *EMBO J.*, 28(19):2932–2944, Oct 2009.

[38] J. V. Hartig and K. Forstemann. Loqs-PD and R2D2 define independent pathways for RISC generation in Drosophila. *Nucleic Acids Res.*, 39(9):3836–3851, May 2011.

[39] Y. Huang, L. Ji, Q. Huang, D. G. Vassylyev, X. Chen, and J. B. Ma. Structural insights into mechanisms of the small RNA methyltransferase HEN1. *Nature*, 461(7265):823–827, Oct 2009.

[40] S. Iwasaki, M. Kobayashi, M. Yoda, Y. Sakaguchi, S. Katsuma, T. Suzuki, and Y. Tomari. Hsc70/Hsp90 chaperone machinery mediates ATP-dependent RISC loading of small RNA duplexes. *Mol. Cell*, 39(2):292–299, Jul 2010.

[41] M. Johnston, M. C. Geoffroy, A. Sobala, R. Hay, and G. Hutvagner. HSP90 protein stabilizes unloaded argonaute complexes and microscopic P-bodies in human cells. *Mol. Biol. Cell*, 21(9):1462–1469, May 2010.

[42] T. Kawamata, H. Seitz, and Y. Tomari. Structural determinants of miRNAs for RISC loading and slicer-independent unwinding. *Nat. Struct. Mol. Biol.*, 16(9):953–960, Sep 2009.

[43] Y. Kawamura, K. Saito, T. Kin, Y. Ono, K. Asai, T. Sunohara, T. N. Okada, M. C. Siomi, and H. Siomi. Drosophila endogenous small RNAs bind to Argonaute 2 in somatic cells. *Nature*, 453(7196):793–797, Jun 2008.

[44] A. Kharrat, M. J. Macias, T. J. Gibson, M. Nilges, and A. Pastore. Structure of the dsRNA binding domain of E. coli RNase III. *EMBO J.*, 14(14):3572–3584, Jul 1995.

[45] I. Kim, C. W. Liu, and J. D. Puglisi. Specific recognition of HIV TAR RNA by the dsRNA binding domains (dsRBD1-dsRBD2) of PKR. *J. Mol. Biol.*, 358(2):430–442, Apr 2006.

[46] S. Kim, E. Brostromer, D. Xing, J. Jin, S. Chong, H. Ge, S. Wang, C. Gu, L. Yang, Y. Q. Gao, X. D. Su, Y. Sun, and X. S. Xie. Probing allostery through DNA. *Science*, 339(6121):816–819, Feb 2013.

[47] H. R. Koh, M. A. Kidwell, K. Ragunathan, J. A. Doudna, and S. Myong. ATP-independent diffusion of double-stranded RNA binding proteins. *Proc. Natl. Acad. Sci. U.S.A.*, 110(1):151–156, Jan 2013.

[48] B. C. Krovat and M. F. Jantsch. Comparative mutational analysis of the double-stranded RNA binding domains of Xenopus laevis RNA-binding protein A. *J. Biol. Chem.*, 271(45):28112–28119, Nov 1996.

[49] U. K. Laemmli. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*, 227(5259):680–685, Aug 1970.

[50] E. Lages, H. Ipas, A. Guttin, H. Nesr, F. Berger, and J. P. Issartel. MicroRNAs: molecular features and role in cancer. *Front. Biosci.*, 17:2508–2540, 2012.

[51] P. W. Lau, K. Z. Guiley, N. De, C. S. Potter, B. Carragher, and I. J. MacRae. The molecular architecture of human Dicer. *Nat. Struct. Mol. Biol.*, 19(4):436–440, Apr 2012.

[52] P. W. Lau, C. S. Potter, B. Carragher, and I. J. MacRae. Structure of the human Dicer-TRBP complex by electron microscopy. *Structure*, 17(10):1326–1332, Oct 2009.

[53] Y. Lee, I. Hur, S. Y. Park, Y. K. Kim, M. R. Suh, and V. N. Kim. The role of PACT in the RNA silencing pathway. *EMBO J.*, 25(3):522–532, Feb 2006.

[54] N. Leulliot, S. Quevillon-Cheruel, M. Graille, H. van Tilbeurgh, T. C. Leeper, K. S. Godin, T. E. Edwards, S. T. Sigurdsson, N. Rozenkrants, R. J. Nagel, M. Ares, and G. Varani. A new alpha-helical extension promotes RNA binding by the dsRBD of Rnt1p RNAse III. *EMBO J.*, 23(13):2468–2477, Jul 2004.

[55] D. H. Lim, L. Lee, C. T. Oh, N. H. Kim, S. Hwang, S. J. Han, and Y. S. Lee. Microarray analysis of Drosophila dicer-2 mutants reveals potential regulation of mitochondrial metabolism by endogenous siRNAs. *J. Cell. Biochem.*, 114(2):418–427, Feb 2013.

[56] D. H. Lim, C. T. Oh, L. Lee, J. S. Hong, S. H. Noh, S. Hwang, S. Kim, S. J. Han, and Y. S. Lee. The endogenous siRNA pathway in Drosophila impacts stress resistance and lifespan by regulating metabolic homeostasis. *FEBS Lett.*, 585(19):3079–3085, Oct 2011.

[57] A. Lingel, B. Simon, E. Izaurralde, and M. Sattler. Structure and nucleic-acid binding of the Drosophila Argonaute 2 PAZ domain. *Nature*, 426(6965):465–469, Nov 2003.

[58] Q. Liu, T. A. Rand, S. Kalidas, F. Du, H. E. Kim, D. P. Smith, and X. Wang. R2D2, a bridge between the initiation and effector steps of the Drosophila RNAi pathway. *Science*, 301(5641):1921–1925, Sep 2003.

[59] X. Liu, F. Jiang, S. Kalidas, D. Smith, and Q. Liu. Dicer-2 and R2D2 coordinately bind siRNA to promote assembly of the siRISC complexes. *RNA*, 12(8):1514–1520, Aug 2006.

[60] X. Liu, J. K. Park, F. Jiang, Y. Liu, D. McKearin, and Q. Liu. Dicer-1, but not Loquacious, is critical for assembly of miRNA-induced silencing complexes. *RNA*, 13(12):2324–2329, Dec 2007.

[61] Y. Liu, H. Tan, H. Tian, C. Liang, S. Chen, and Q. Liu. Autoantigen La promotes efficient RNAi, antiviral response, and transposon silencing by facilitating multiple-turnover RISC catalysis. *Mol. Cell*, 44(3):502–508, Nov 2011.

[62] Y. Liu, X. Ye, F. Jiang, C. Liang, D. Chen, J. Peng, L. N. Kinch, N. V. Grishin, and Q. Liu. C3PO, an endoribonuclease that promotes RNAi by facilitating RISC activation. *Science*, 325(5941):750–753, Aug 2009.

[63] E. Lund, S. Guttinger, A. Calado, J. E. Dahlberg, and U. Kutay. Nuclear export of microRNA precursors. *Science*, 303(5654):95–98, Jan 2004.

[64] J. B. Ma, K. Ye, and D. J. Patel. Structural basis for overhang-specific small interfering RNA recognition by the PAZ domain. *Nature*, 429(6989):318–322, May 2004.

[65] I. J. Macrae, K. Zhou, F. Li, A. Repic, A. N. Brooks, W. Z. Cande, P. D. Adams, and J. A. Doudna. Structural basis for double-stranded RNA processing by Dicer. *Science*, 311(5758):195–198, Jan 2006.

[66] J. T. Marques, K. Kim, P. H. Wu, T. M. Alleyne, N. Jafari, and R. W. Carthew. Loqs and R2D2 act sequentially in the siRNA pathway in Drosophila. *Nat. Struct. Mol. Biol.*, 17(1):24–30, Jan 2010.

[67] G. Masliah, P. Barraud, and F. H. Allain. RNA recognition by double-stranded RNA binding domains: a matter of shape and sequence. *Cell. Mol. Life Sci.*, Aug 2012.

[68] C. Matranga, Y. Tomari, C. Shin, D. P. Bartel, and P. D. Zamore. Passenger-strand cleavage facilitates assembly of siRNA into Ago2-containing RNAi enzyme complexes. *Cell*, 123(4):607–620, Nov 2005.

[69] K. M. Michalik, R. Bottcher, and K. Forstemann. A small RNA response at DNA ends in Drosophila. *Nucleic Acids Res.*, 40(19):9596–9603, Oct 2012.

[70] K. Miyoshi, T. Miyoshi, J. V. Hartig, H. Siomi, and M. C. Siomi. Molecular mechanisms that funnel RNA precursors into endogenous small-interfering RNA and microRNA biogenesis pathways in Drosophila. *RNA*, 16(3):506–515, Mar 2010.

[71] K. Miyoshi, T. N. Okada, H. Siomi, and M. C. Siomi. Characterization of the miRNA-RISC loading complex and miRNA-RISC formed in the Drosophila miRNA pathway. *RNA*, 15(7):1282–1291, Jul 2009.

[72] T. Miyoshi, A. Takeuchi, H. Siomi, and M. C. Siomi. A direct role for Hsp90 in pre-RISC formation in Drosophila. *Nat. Struct. Mol. Biol.*, 17(8):1024–1026, Aug 2010.

[73] D. Murphy, B. Dancis, and J. R. Brown. The evolution of core proteins involved in microRNA biogenesis. *BMC Evol. Biol.*, 8:92, 2008.

[74] S. Nanduri, B. W. Carpick, Y. Yang, B. R. Williams, and J. Qin. Structure of the double-stranded RNA-binding domain of the protein kinase PKR reveals the molecular basis of its dsRNA-mediated activation. *EMBO J.*, 17(18):5458–5465, Sep 1998.

[75] S. Nanduri, F. Rahman, B. R. Williams, and J. Qin. A dynamically tuned double-stranded RNA binding mechanism for the activation of antiviral kinase PKR. *EMBO J.*, 19(20):5567–5574, Oct 2000.

[76] K. M. Nishida, K. Miyoshi, A. Ogino, T. Miyoshi, H. Siomi, and M. C. Siomi. Roles of R2D2, a Cytoplasmic D2 Body Component, in the Endogenous siRNA Pathway in Drosophila. *Mol. Cell*, Jan 2013.

[77] C. L. Noland and J. A. Doudna. Multiple sensors ensure guide strand selection in human RNAi pathways. *RNA*, Mar 2013.

[78] C. L. Noland, E. Ma, and J. A. Doudna. siRNA repositioning for guide strand selection by human Dicer complexes. *Mol. Cell*, 43(1):110–121, Jul 2011.

[79] K. Okamura, S. Balla, R. Martin, N. Liu, and E. C. Lai. Two distinct mechanisms generate endogenous siRNAs from bidirectional transcription in Drosophila melanogaster. *Nat. Struct. Mol. Biol.*, 15(6):581–590, Jun 2008.

[80] K. Okamura, W. J. Chung, J. G. Ruby, H. Guo, D. P. Bartel, and E. C. Lai. The Drosophila hairpin RNA pathway generates endogenous short interfering RNAs. *Nature*, 453(7196):803–806, Jun 2008.

[81] K. Okamura and E. C. Lai. Endogenous small interfering RNAs in animals. *Nat. Rev. Mol. Cell Biol.*, 9(9):673–678, Sep 2008.

[82] K. Okamura, N. Liu, and E. C. Lai. Distinct mechanisms for microRNA strand selection by Drosophila Argonautes. *Mol. Cell*, 36(3):431–444, Nov 2009.

[83] G. S. Parker, T. S. Maity, and B. L. Bass. dsRNA binding properties of RDE-4 and TRBP reflect their distinct roles in RNAi. *J. Mol. Biol.*, 384(4):967–979, Dec 2008.

[84] R. C. Patel and G. C. Sen. Requirement of PKR dimerization mediated by specific hydrophobic residues for its activation by double-stranded RNA and its antigrowth effects in yeast. *Mol. Cell. Biol.*, 18(12):7009–7019, Dec 1998.

[85] J. W. Pham and E. J. Sontheimer. Molecular requirements for RNA-induced silencing complex assembly in the Drosophila RNA interference pathway. *J. Biol. Chem.*, 280(47):39278–39283, Nov 2005.

[86] J. B. Preall, Z. He, J. M. Gorra, and E. J. Sontheimer. Short interfering RNA strand selection is independent of dsRNA processing polarity during RNAi in Drosophila. *Curr. Biol.*, 16(5):530–535, Mar 2006.

[87] H. Qin, F. Chen, X. Huan, S. Machida, J. Song, and Y. A. Yuan. Structure of the Arabidopsis thaliana DCL4 DUF283 domain reveals a noncanonical double-stranded RNA-binding fold for protein-protein interaction. *RNA*, 16(3):474–481, Mar 2010.

[88] R. M. Rasia, J. Mateos, N. G. Bologna, P. Burdisso, L. Imbert, J. F. Palatnik, and J. Boisbouvier. Structure and RNA interactions of the plant MicroRNA processing-associated protein HYL1. *Biochemistry*, 49(38):8237–8239, Sep 2010.

[89] M. H. Roehrl, J. Y. Wang, and G. Wagner. A general framework for development and data analysis of competitive high-throughput screens for small-molecule inhibitors of protein-protein interactions by fluorescence polarization. *Biochemistry*, 43(51):16056–16066, Dec 2004.

[90] J. M. Ryter and S. C. Schultz. Molecular basis of double-stranded RNA-protein in-

teractions: structure of a dsRNA-binding domain complexed with dsRNA. *EMBO J.*, 17(24):7505–7513, Dec 1998.

[91] L. R. Sabin, Q. Zheng, P. Thekkat, J. Yang, G. J. Hannon, B. D. Gregory, M. Tudor, and S. Cherry. Dicer-2 processes diverse viral RNA species. *PLoS ONE*, 8(2):e55458, 2013.

[92] K. Saito, A. Ishizuka, H. Siomi, and M. C. Siomi. Processing of pre-microRNAs by the Dicer-1-Loquacious complex in Drosophila cells. *PLoS Biol.*, 3(7):e235, Jul 2005.

[93] D. G. Sashital and J. A. Doudna. Structural insights into RNA interference. *Curr. Opin. Struct. Biol.*, 20(1):90–97, Feb 2010.

[94] I. Schneider. Cell lines derived from late embryonic stages of Drosophila melanogaster. *J Embryol Exp Morphol*, 27(2):353–365, Apr 1972.

[95] Y. Shamoo, N. Abdul-Manan, and K. R. Williams. Multiple RNA binding domains (RBDs) just don't add up. *Nucleic Acids Res.*, 23(5):725–728, Mar 1995.

[96] J. Soding, A. Biegert, and A. N. Lupas. The HHpred interactive server for protein homology detection and structure prediction. *Nucleic Acids Res.*, 33(Web Server issue):W244–248, Jul 2005.

[97] S. Y. Sohn, W. J. Bae, J. J. Kim, K. H. Yeom, V. N. Kim, and Y. Cho. Crystal structure of human DGCR8 core. *Nat. Struct. Mol. Biol.*, 14(9):847–853, Sep 2007.

[98] E. J. Sontheimer. Site-specific RNA crosslinking with 4-thiouridine. *Mol. Biol. Rep.*, 20(1):35–44, Jul 1994.

[99] R. Stefl, F. C. Oberstrass, J. L. Hood, M. Jourdan, M. Zimmermann, L. Skrisovska, C. Maris, L. Peng, C. Hofr, R. B. Emeson, and F. H. Allain. The solution structure of the ADAR2 dsRBM-RNA complex reveals a sequence-specific readout of the minor groove. *Cell*, 143(2):225–237, Oct 2010.

[100] W. Sun, A. Pertzev, and A. W. Nicholson. Catalytic mechanism of Escherichia coli ribonuclease III: kinetic and inhibitor evidence for the involvement of two magnesium ions in RNA phosphodiester hydrolysis. *Nucleic Acids Res.*, 33(3):807–815, 2005.

[101] B. Tian, P. C. Bevilacqua, A. Diegelman-Parente, and M. B. Mathews. The double-stranded-RNA-binding motif: interference and much more. *Nat. Rev. Mol. Cell Biol.*, 5(12):1013–1023, Dec 2004.

[102] B. Tian and M. B. Mathews. Functional characterization of and cooperation between the double-stranded RNA-binding motifs of the protein kinase PKR. *J. Biol. Chem.*, 276(13):9936–9944, Mar 2001.

[103] Y. Tomari, T. Du, and P. D. Zamore. Sorting of Drosophila small silencing RNAs. *Cell*, 130(2):299–308, Jul 2007.

[104] Y. Tomari, C. Matranga, B. Haley, N. Martinez, and P. D. Zamore. A protein sensor for siRNA asymmetry. *Science*, 306(5700):1377–1380, Nov 2004.

[105] A. Tsutsumi, T. Kawamata, N. Izumi, H. Seitz, and Y. Tomari. Recognition of the pre-miRNA structure by Drosophila Dicer-1. *Nat. Struct. Mol. Biol.*, 18(10):1153–1158, Oct 2011.

[106] R. P. van Rij and E. Berezikov. Small RNAs and the control of transposons and viruses in Drosophila. *Trends Microbiol.*, 17(4):163–171, Apr 2009.

[107] T. A. Volpe, C. Kidner, I. M. Hall, G. Teng, S. I. Grewal, and R. A. Martienssen. Regulation of heterochromatic silencing and histone H3 lysine-9 methylation by RNAi. *Science*, 297(5588):1833–1837, Sep 2002.

[108] H. W. Wang, C. Noland, B. Siridechadilok, D. W. Taylor, E. Ma, K. Felderer, J. A. Doudna, and E. Nogales. Structural insights into RNA processing by the human RISC-loading complex. *Nat. Struct. Mol. Biol.*, 16(11):1148–1153, Nov 2009.

[109] X. H. Wang, R. Aliyari, W. X. Li, H. W. Li, K. Kim, R. Carthew, P. Atkinson, and S. W. Ding. RNA interference directs innate immunity against viruses in adult Drosophila. *Science*, 312(5772):452–454, Apr 2006.

[110] M. Wassenegger. The role of the RNAi machinery in heterochromatin formation. *Cell*, 122(1):13–16, Jul 2005.

[111] K. M. Weeks and D. M. Crothers. Major groove accessibility of RNA. *Science*, 261(5128):1574–1577, Sep 1993.

[112] W. Wei, Z. Ba, M. Gao, Y. Wu, Y. Ma, S. Amiard, C. I. White, J. M. Rendtlew Danielsen, Y. G. Yang, and Y. Qi. A role for small RNAs in DNA double-strand break repair. *Cell*, 149(1):101–112, Mar 2012.

[113] N. C. Welker, T. S. Maity, X. Ye, P. J. Aruscavage, A. A. Krauchuk, Q. Liu, and B. L. Bass. Dicer's helicase domain discriminates dsRNA termini to promote an altered reaction mode. *Mol. Cell*, 41(5):589–599, Mar 2011.

[114] C. Wostenberg, J. W. Lary, D. Sahu, R. Acevedo, K. A. Quarles, J. L. Cole, and S. A. Showalter. The role of human Dicer-dsRBD in processing small regulatory RNAs. *PLoS ONE*, 7(12):e51829, 2012.

[115] C. Wostenberg, K. A. Quarles, and S. A. Showalter. Dynamic origins of differential RNA binding function in two dsRBDs from the miRNA "microprocessor" complex. *Biochemistry*, 49(50):10728–10736, Dec 2010.

[116] S. Yamashita, T. Nagata, M. Kawazoe, C. Takemoto, T. Kigawa, P. Guntert, N. Kobayashi, T. Terada, M. Shirouzu, M. Wakiyama, Y. Muto, and S. Yokoyama. Structures of the first and second double-stranded RNA-binding domains of human TAR RNA-binding protein. *Protein Sci.*, 20(1):118–130, Jan 2011.

[117] S. W. Yang, H. Y. Chen, J. Yang, S. Machida, N. H. Chua, and Y. A. Yuan. Structure of Arabidopsis HYPONASTIC LEAVES1 and its molecular implications for miRNA processing. *Structure*, 18(5):594–605, May 2010.

[118] X. Ye and Q. Liu. Expression, purification, and analysis of recombinant Drosophila Dicer-1 and Dicer-2 enzymes. *Methods Mol. Biol.*, 442:11–27, 2008.

[119] X. Ye, Z. Paroo, and Q. Liu. Functional anatomy of the Drosophila microRNA-generating enzyme. *J. Biol. Chem.*, 282(39):28373–28378, Sep 2007.

[120] M. Yoda, T. Kawamata, Z. Paroo, X. Ye, S. Iwasaki, Q. Liu, and Y. Tomari. ATP-dependent human RISC assembly pathways. *Nat. Struct. Mol. Biol.*, 17(1):17–23, Jan 2010.

[121] H. Zhang, F. A. Kolb, L. Jaskiewicz, E. Westhof, and W. Filipowicz. Single processing center models for human Dicer and bacterial RNase III. *Cell*, 118(1):57–68, Jul 2004.

[122] X. Zheng and P. C. Bevilacqua. Straightening of bulged RNA by the double-stranded RNA-binding domain from the protein kinase PKR. *Proc. Natl. Acad. Sci. U.S.A.*, 97(26):14162–14167, Dec 2000.

[123] R. Zhou, B. Czech, J. Brennecke, R. Sachidanandam, J. A. Wohlschlegel, N. Perrimon, and G. J. Hannon. Processing of Drosophila endo-siRNAs depends on a specific Loquacious isoform. *RNA*, 15(10):1886–1895, Oct 2009.

Abbreviations

°C	degrees Celsius
AA	amino acid
ADAR	adenosine deaminase acting on RNA
Ago	Argonaute
Amp	Ampicillin
APS	ammonium peroxodisulfate
ATP	adenosine triphosphate
BLAST	Basic Local Alignment Search Tool
bp	base pair(s)
BSA	bovine serum albumine
co-IP	co-immunoprecipitation
CSD	Cold shock domain
CV	column volume
Da	Dalton
Dcr	Dicer
DMSO	dimethyl sulfoxide
DNA	desoxy-ribonucleic acid
dNTP	desoxy-nucleotide-tri-phosphate
ds	double-stranded
dsRBP	double-stranded RNA binding domain protein
DsRed	Discosoma striata red fluorescence protein
dsRNA	double-stranded RNA
DTT	dithiothreitol
DUF	Domain of unknown function
E.coli	Escherichia coli
EMSA	Electrophoretic mobility shift assay
endo-	endogenous
exo-	exogenous
FACS	Fluorescence Activated Cell Sorting
FBS	Fetal Bovine Serum
fw	forward
GFP	Green Fluorescent Protein
h	hour(s)
HEPES	(4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
IP	immunoprecipitation
IPTG	Isopropyl-β-D-thiogalactopyranosid
k	kilo
k.d.	knock down
Kan	kanamycin
KH domain	K-homology domain
Loqs	loquacious
Luc	luciferase
mg	milligram
min	minute
miRNA	micro RNA
ml	milliliter
mRNA	messenger RNA
ng	nanogram
nt	nucleotide(s)
NTP	nucleotide-tri-phosphate
o/n	over night
PA/PB/PC/PD	protein isoform A/B/C/D

PAGE	Polyacrylamide Gel Electrophoresis	RRM	RNA recognition motif
		RT	room temperature
PAZ	Piwi-Argonaute-Zwille domain of Dicer and Argonaute proteins	S2 cell	Schneider-2 cell
		SDS	sodium docecyl sulfate
PCR	Polymerase Chain Reaction	siRNA	small interfering RNA
piRNA	Piwi-interacting RNA	ss	single stranded
PNK	polynucleotide kinase	TAR	transactivation response RNA
Pol II	DNA polymerase II	TRBP	TAR RNA binding protein
R2D2	2 dsRBD-containing protein interacting with Dcr-2	TRIS	Tris(hydroxymethyl)-aminomethane
rb	rabbit	tub.	tubulin
rev	reverse	UTR	untranslated region
RISC	RNA induced silencing complex	V	Volt
RLC	RISC loading complex	α	anti
RNA	ribonucleic acid	Δ	deletion
RNAi	RNA interference	μ	micro
RNaseIII	endoribonuclease class III	μg	microgram
rpm	rounds per minute		

Danksagung

Zuallererst gebührt mein Dank natürlich meinem Doktorvater Klaus Förstemann, der mir mit meinem Thema interessante Fragestellungen überlassen hat und mich mit seinem Vertrauen, seinem Optimismus und seinen tollen Ideen immer motivieren konnte.

Thomas Kern, Jan-Niklas Tants und Michael Sattler danke ich für die Zusammenarbeit am LoqsPD Projekt. Die Treffen mit euch haben zu lehrreichen Eindrücken aus ungewohnten Blickwinkeln geführt. Michael Sattler danke ich außerdem sehr herzlich für die Übernahme des Zweitgutachtens.

Stefan Kunzelmann, Erik Tilch und Anselm Gruber, danke euch für eure Mitarbeit während eurer Praktikumszeit.

Danke an alle Genzentrums-kollegen, für eure bereitwillige Hilfe, aufheiternde Gespräche auf dem Gang und einfach eine gute Zeit.

Ein riesen Dankeschön natürlich allen Förstegirls: Ihr habt diese Jahre unvergesslich gemacht. Ich weiß es sehr zu schätzen, daß ich einen so großen Teil meiner Zeit mit Leuten verbringen durfte mit denen ich soviel lachen und so unbeschwert sein konnte wie mit euch.

Romy, Danke für Deine Hilfe während deinem Exkurs in die Proteinaufreinigung und für ungezähltes Animpfen.

Danke Katha, für die vielen Autofahrten, die fruchtbaren Diskussionen und die Inspiration zu allen möglichen Dingen.

Meiner Familie, vorneweg natürlich meinen Eltern, Danke daß ich immer euren Rückhalt habe, und für eure großzügige Unterstützung.

Florian, ich danke Dir für Deine Ausdauer im Pendeln nach München, für Deine Hilfe gerade in den letzten Wochen und Deine Art mich wieder auf den Boden zu bringen wenns nötig ist. Und für den ganzen Rest.

Appendix

Significance of differences in dissociation constants

	Loqs dsRBD1				
	siRNA	miRNA	dsRNA	premiRNA	14nt RNA
siRNA	1.00				
miRNA	0.82	1.00			
dsRNA	0.18	0.11	1.00		
premiRNA	0.82	0.69	0.38	1.00	
14nt RNA	0.05	0.06	0.03	0.04	1.00

	Loqs dsRBD2				
	siRNA	miRNA	dsRNA	premiRNA	14nt RNA
siRNA	1.00				
miRNA	0.48	1.00			
dsRNA	0.21	0.83	1.00		
premiRNA	0.22	0.70	0.81	1.00	
14nt RNA	0.05	0.16	0.06	0.08	1.00

	Loqs DeltaNC				
	siRNA	miRNA	dsRNA	premiRNA	14nt RNA
siRNA	1.00				
miRNA	0.35	1.00			
dsRNA	0.77	0.46	1.00		
premiRNA	0.14	0.33	0.17	1.00	
14nt RNA	0.11	0.25	0.14	0.80	1.00

	LoqsPD full length				
	siRNA	miRNA	dsRNA	premiRNA	14nt RNA
siRNA	1.00				
miRNA	0.00	1.00			
dsRNA	0.02	0.00	1.00		
premiRNA	0.00	0.15	0.00	1.00	
14nt RNA	0.00	0.98	0.00	0.12	1.00

Significance of differences in dissociation constants between the four Loqs protein constructs and the various RNA substrates. Significances were calculated using an unpaired, heteroscedastic t-test. Dark green: difference is significant, $p < 5\%$. Light green: $5\% < p < 10\%$.

Labtalk script for automated evaluation of Nanotemper Monolith data

```
dlgFile multi:=4 group:=ascii init:="C:\Documents and Settings\fesser\My Documents\OriginLab\81\User Files\Stephanie\";
int anzahldateien = fname.getnumtokens(CRLF); //selected files are saved in fname, separated by CRLF
StringArray dateinamen;

getn -s (How often were individual experiments repeated?) AnzahlMessungen; //Number of repetitions of one measurement is saved under the variable AnzahlMessungen, which is important for sorting the data points

//filenames are saved in the stringarray “dateinamen”, with dateinamen.getAt(position)$ the entries can be accessed.
loop (qq, 1, anzahldateien){
   dateinamen.add(fname.gettoken(qq,crlf)$);
};
```

```
//importing the data, loops over all selected files, each file is imported in its own sheet and sorted there.
loop (ii,1, anzahldateien){
   string index$ = $(ii);//generates the name of the graph
   string graphname$ = "fitgraph" + index$;
   string AktuelleDatei$ = dateinamen.getAt(ii)$; //gets the path of the current file from the stringarray “dateinamen”
   string filepath$ = AktuelleDatei.GetFilePath()$;

   impasc AktuelleDatei$ //file is imported and opened in a workbook
   options.FileStruct.NumericSeparator:= 0 //defining , as a 1000-separator
   options.sparklines:=0;

   win -r %H Evaluation; //renames the workbook in Evaluation

   //sorts entries in column A
   int AnzahlEintraege = count(col(A),1); //counts number of entries
   int AnzahlMesspunkte = AnzahlEintraege/AnzahlMessungen;
   for (ww=1, jj=1; ww<=AnzahlEintraege; ww+AnzahlMessungen, jj++){//loops through all entries in column A and sorts them (in the nanotemper output file of repeated experiments the data points are not sorted by repetition but by x-value)
   col(A)[jj]=col(A)[ww];
   }
   range fordeletion = 1[AnzahlMesspunkte+1 : AnzahlEintraege];
   del fordeletion; //deletes remaining numbers, only one set of x-values is needed

   //separating and sorting column B
   string CommentName$ = AktuelleDatei.GetFileName(1)$; //Backdiffusion etc is written in comment panel, including the number of the measurement
   int NumToken = AktuelleDatei.getNumTokens('\');
   NumToken -=1;
   string LongName$ = AktuelleDatei.getToken(NumToken,'\')$; //Name of the folder is written in LongName panel of the column, includes date, protein and RNA type
   for (rr=1; rr<=AnzahlMessungen; rr++){
   string Number$ = $(rr);
   string Comment$ = CommentName$ + " Nr" + Number$;
   for (ee=rr, yy=1; ee<=AnzahlEintraege; ee+AnzahlMessungen, yy++){
      col(Comment$)[yy]=col(B)[ee]};
      Col(Comment$)[C]$= Comment$;
      Col(Comment$)[L]$= Longname$;
   }
   del -s col(B);

   //plotten
   int min = col(1)[2];
   int anzahlcurves = AnzahlMessungen+1;
   plotxy iy:=(1,2:end) plot:= 201;//walk parameter should better be variable
   axis -ps x s 1; //log scale
   layer.x.from = min;
```

```
expGraph type:=tif path:=filepath$ filename:=%H
tr1.unit:=2
tr1.width:=640;

win -a Evaluation;

for (mm=1; mm<=AnzahlMessungen+5; mm++){//add enough columns for fitcurves and fit results
    wks.addcol();
}

//fitting and plotting
range xx = wcol(AnzahlMessungen+2);//generates x-values for the fitcurve
int cnt= count(col(A),1);
int max = col(1)[cnt];
xx=data(0,max,1); //x-values for the fitcurve are saved in xx
xx.type = 4;//designating column with these x-values as X-value-column
xx[L]$="Fitted X";//longname of the x-column

int counter = 0;//counter that is increased after readout of the fit results, so that they do not overwrite the ones of the previous measurement
int counterfitimage = 0;

for(dd=2; dd<=AnzahlMessungen+1; dd++){//go through all measurements in one sheet and fit each one individually
    win -a Evaluation; //activate the correct window, since later in the loop a graph window becomes activated
    range yy = wcol(dd+AnzahlMessungen+1);//assigning a column to the y-values that will be generated by the fit
    yy[L]$="Fitted Y";//longname of this column
    nlbegin (1,wcol(dd)) logistic tt init:=enable; //actual fit, fit function: logistic, tt is the data tree in which the fitparameters are saved.
    nlfit; // fit
    yy=fit(xx); //function “fit” takes parameter from the last nlfit and applies them to the values saved in xx.

    //writing the fitparameter in the corresponding columns in the Evaluation window
    oo = counter;

    wcol(AnzahlMessungen+AnzahlMessungen+2+1)[oo+1]$="A1";
    wcol(AnzahlMessungen+AnzahlMessungen+2+2)[oo+1]=int(tt.A1);
    wcol(AnzahlMessungen+AnzahlMessungen+2+3)[oo+1]=int(tt.e_A1);
    wcol(AnzahlMessungen+AnzahlMessungen+2+1)[oo+2]$="A2";
    wcol(AnzahlMessungen+AnzahlMessungen+2+2)[oo+2]=int(tt.A2);
    wcol(AnzahlMessungen+AnzahlMessungen+2+3)[oo+2]=int(tt.e_A2);
    wcol(AnzahlMessungen+AnzahlMessungen+2+1)[oo+3]$="x0";
    wcol(AnzahlMessungen+AnzahlMessungen+2+2)[oo+3]=int(tt.x0);
    wcol(AnzahlMessungen+AnzahlMessungen+2+3)[oo+3]=int(tt.e_x0);
    wcol(AnzahlMessungen+AnzahlMessungen+2+1)[oo+4]$="p";
    wcol(AnzahlMessungen+AnzahlMessungen+2+2)[oo+4]=int(tt.p);
```

```
      wcol(AnzahlMessungen+AnzahlMessungen+2+3)[oo+4]=int(tt.e_p);
      wcol(AnzahlMessungen+AnzahlMessungen+2+1)[oo+5]$="EC20";
      wcol(AnzahlMessungen+AnzahlMessungen+2+2)[oo+5]=int(tt.EC20);
      wcol(AnzahlMessungen+AnzahlMessungen+2+3)[oo+5]$="-";
      wcol(AnzahlMessungen+AnzahlMessungen+2+1)[oo+6]$="EC50";
      wcol(AnzahlMessungen+AnzahlMessungen+2+2)[oo+6]=int(tt.EC50);
      wcol(AnzahlMessungen+AnzahlMessungen+2+3)[oo+6]$="-";
      wcol(AnzahlMessungen+AnzahlMessungen+2+1)[oo+7]$="EC80";
      wcol(AnzahlMessungen+AnzahlMessungen+2+2)[oo+7]=int(tt.EC80);
      wcol(AnzahlMessungen+AnzahlMessungen+2+3)[oo+7]$="-";
      wcol(AnzahlMessungen+AnzahlMessungen+2+1)[oo+8]$="Chisqr";
      wcol(AnzahlMessungen+AnzahlMessungen+2+2)[oo+8]=int(tt.chisqr);
      wcol(AnzahlMessungen+AnzahlMessungen+2+3)[oo+8]$="-";
      nlend;

      counter = counter+9; //raising the counter, results from the next measurement are written
9 windows below the ones from the current measurement.

      range aa = (1,wcol(dd)), bb=(wcol(AnzahlMessungen+2),wcol(dd+AnzahlMessungen+1));
//plots datapoints and fitcurve to one graph, all graphs belonging to one measurement are plot-
tet in one window.
         plotxy aa 201 color:=dd ogl:=[graphname$];
         plotxy bb 200 color:=dd ogl:=[graphname$];

      counterfitimage = counterfitimage+1;

      if(counterfitimage == AnzahlMessungen){
         win -a %(graphname$);
         expGraph type:=tif path:=filepath$ filename:=%H
         tr1.unit:=2
         tr1.width:=640;
      }
   }

   win -a Evaluation;
   string Dateiname$ = Longname$ + "_" + commentName$ + "_results.txt"; //-> each data file
has an own result file.
   string resultsavename$ = filepath$ + Dateiname$;
   range results =     wcol(AnzahlMessungen+AnzahlMessungen+2+1):wcol(AnzahlMessungen+AnzahlMessu
//first column is not exportet

   wcellsel results c:=ge v:=0.5;

   expAsc
   type:=1 /* txt format */
   select:= 1 /* export the selected data */
   path:=resultsavename$
   shortname:=1 /* use short name as the headlines */
   longname:=1
   separator:=3 /* set (,TAB) as the separator */
```

```
    comment:=1; /* no comment */

    win -a Evaluation;
    newsheet; //neues sheet
};
```

Mass spectrometry results list

	Nominal mass	pI value	Sequence Coverage	
1	199463	6.41	43%	Dicer-2 [Drosophila melanogaster]
2	71372	5.36	27%	heat shock protein cognate 1, isoform A [Drosophila melanogaster]
				heat shock protein 70 cognate [Spodoptera litura]
3	72330	5.22	22%	heat shock protein cognate 1, isoform A [Drosophila melanogaster]
4	70871	5.34	21%	heat shock protein cognate 1, isoform A [Drosophila melanogaster]
				heat shock protein 70 cognate [Spodoptera litura]
5	70043	5.63	19%	heat shock protein 68 [Drosophila melanogaster]
6	70516	5.54	11%	Heat-shock-protein-70Aa [Drosophila melanogaster]
7	70078	6.76	15%	heat shock protein cognate 2 [Drosophila melanogaster]
				heat shock protein 70 [Spodoptera exigua]
8	35503	5.36	39%	r2d2, isoform A [Drosophila melanogaster]
9	48836	9.58	1%	CG6404, isoform B [Drosophila melanogaster]
10	11374	11.36	21%	Histone H4 [Cricetulus griseus]